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 essential cognitive ability. Categorization, which involves compression of information, is one solution to handle an almost infinite number of entities efficiently. Humans have great linguistic abilities and have excellently categorized vocabulary. How do we infer meanings of novel nouns? Quine (1960) explained the difficulty of word learning without prior knowledge. If we hear a novel word of unknown language in one situation, how do we infer it? For example, suppose we heard a novel word 'gavagai' when we looked at a rabbit in a field. 'Gavagai' might mean *rabbit* ,but it might mean *rabbit's color*. How can we infer the meaning of 'gavagai'? This problem occurs when children acquire the word meaning in the early stage. Parent's daily words to their children will be spoken with many possible interpretations. How do children learn word meanings in that situation? Children have to logically reject many useless possibilities, so they can not acquire the meaning once. However, in effect, children do not consider useless possibilities. Therefore, they can acquire temporary word meaning presented only once.

Constraints to acquire word meanings

Landau, Smith and Jones (1988) explained that phenomenon with the idea 'constraints'. They claimed that children can learn words so fast, because they use the prior knowledge about vocabulary and entities as constraints. Landau et al. (1988) showed that shape is a important property to categorize objects and they called this 'shape bias'. Colunga & Smith (2000) and Samuelson (2002) suggested that children attended to perceptual features depending on the solidity of objects. In other words, children know the nature of entities and use them to generalize the novel words. We focus on how children acquire the knowledge about the nature of entities and ontological categories.

Much research suggested that children know ontological category before speaking words. For example, babies seem to understand individuation of entities (Spelke, 1990). **Individuation** is entities' concreteness or continuity. Entities are divided into two global ontological categories by individuation. These two categories are 'objects' and 'substance'. Objects are concrete entities which we recognize as 'one isolated entities', and parts of objects do not have the same nature as whole objects. For examples, cats or dishes are kinds of objects; legs of cats are not cats and fragments of dishes are not dishes. On the other hand, substance can not be individuated. Part of substance have basically same nature. Water is one kind of substance; water poured into the different glasses is still water. That is why objects and substance have very different nature. Colunga & Smith (2000) and Samueslon (2002) suggested the relation between solidity and shape bias in their experiment.

Animacy is another important ontological category which babies can understand in early stages of development. In this paper, we define two categories of animacy as 'animates' and 'inanimates'. Spelke (1990) suggested that new-born babies know different laws of causality between animates and inanimates. Most of animates have multiple complex features, and move by themselves. On the other hand, most inanimates have simple features and move passively.

Linguistic categories and ontological categories

Some researchers suggested a deep relation between ontological categories and linguistic categories. In particular, the relation between count/ mass noun syntax in English and objects/ substance ontology are typical.

According to Quine (1960), words reference have uncertainty, therefore we will infer the word meaning by using syntactic cues. As a result, count/ mass syntax helps us categorize ontology. However, Soja, Carey and Spelke (1991) criticized him, because their experiments suggested that 2-year-old children who can not judge count/ mass syntax knew ontological categories. Therefore they claimed that ontological categories were learned before syntactic categories and , opposite to Quine, ontological knowledge help children understand syntactic categories.

Imai & Gentner (1997) expanded upon the experi-

Table 1: The boundary shift hypothesis (revised from Yoshida & Smith, 2003). Individuation continuum (Lucy, 1992) has animates at one end, substance at the other end, and objects in the middle.

Japanese English
▽ ▽

Linguistic Individuation	individuals		masses
Conceptual Distinction	animates	objects	substances
Perceptual Cues	multiple similarities	shape	materials
examples	dog	cup	milk

ments of Soja et al. to verify the difference between English and Japanese speakers. English has syntax compatible with solidity ontological category, but Japanese do not have such syntax, so comparing them will reveal the influence of count/ mass syntax to ontological category. The results suggested the different categorization of simple objects between English and Japanese speakers. Imai & Gentner considered these simple objects to be near the boundary between objects and substance, because they were objects, but their fragments also had a similar nature. Their experiments showed the linguistic influence on ontological categories of ambiguous entities.

Japanese have animacy syntax which shows animacy by verb form. For example in sentences,(1) '*Animates*-ga **iru**.' and (2) '*Inanimates*-ga **aru**.', 'iru' and 'aru' have almost the same meaning, but an animate subject needs 'iru' and an inanimate one needs 'aru'. In this paper, we call this syntax 'iru'/ 'aru' syntax. Yoshida & Smith (2001,2003) verified the influence of Japanese syntax by using objects simulating animates. The results suggested that English and Japanese speakers had different categorical criterion.

They proposed 'the boundary shift hypothesis' (BSH) (Table 1). This theory states that the linguistic cues influence the ontological boundaries on 'individuation continuum' , which explains ontological categories by individuation (Lucy, 1992). However, the mechanism of shifting boundary is still unclear. Therefore we quantify the ontological organization and simulate the influence of linguistic cues on it.

Experiment 1

In Experiment 1, we surveyed adults to analyze the statistical structure of vocabulary. Especially, we quantified perceptual properties of objects (e.g. shape, color, texture, etc.) and analyzed their statistical structure. The main goal in this experiment is to examine ontological structure and to propose the computational model to explain BSH. We choose the Semantic Differential (the SD; Osgood, 1957) technique to quantify complex perceptual knowledge.

We surveyed 48 nouns listed in MacArthur Communicative Development Inventory (Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994) [1] by 5 degrees scales of adjectives pairs.

Method

Participants In the pilot survey, we recruited 12 volunteers (from 23 to 25 years old) from Kyoto university. In the main 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

- perceptual properties of objects
  shape, material, color, texture, sound, temperature, flavor, movement, smell, and function.
- Adjective pairs (linguistic scales)
  16 pairs of adjectives expressing perceptual properties of objects as follows: dynamic-static, noisy-silent, light-heavy, large-small, complex-simple, crafted-rough, solid-soft, stable-unstable, strong-weak, natural-artificial, round-square, warm-cool, quick-slow,straight-curved, smooth-irregular and wet-dry.
- Children's canonical acquired vocabulary
  48 nouns were selected evenly from 9 categories [2] of MCDI (see also Table 2).

Procedure We did two surveys, one pilot survey (perceptual properties (10) × adjectives (16)) and a survey using the SD technique (adjectives (16) × vocabularies (48)).

Pilot survey

We asked how 16 pairs of adjectives express the properties of objects. The main purpose is to obtain the basis for transforming linguistic scales from the SD data into perceptual scales. That is why we asked participants, 'How do you use these words to express familiar objects' perceptual features'. We gave participants electronic files to rate the questionnaire. Most participants finished answering the questionnaire in about 30 minutes.

Vocabulary survey

We used the SD technique with 16 linguistic scales to evaluate 48 nouns. The questionnaires had one noun and 16 scales in one page. 16 verbal scales have 5 levels (e.g. very small, small, ambiguous, big and very big). We made questionnaires of 5 different order type to cancel out the order effect. Participants finished answering by one hour.

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.

---

[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 2: 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

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 2) 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.

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

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.

Perceptual weights of the English and Japanese spaces We estimated perceptual weight in the English and Japanese ontological spaces. Tables 3 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.

Experiment 2

In Experiment 1, we quantified the English and Japanese ontological spaces by analysis of adults’ vocabulary knowledge and estimated perceptual weights of the En-

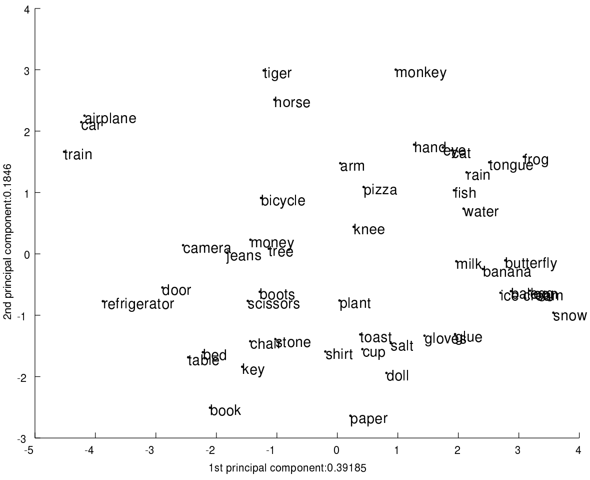

Figure 1: The first two principal components for the neutral condition. The first principal component was interpreted as 'solidity' or 'size' of objects. The second principal component was interpreted as 'animacy' or 'movement' of objects.

Table 3: 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

glish and Japanese ontological space. In this experiment, we used these perceptual weights to simulate the experiment suggesting BSH and provided the quantitative evidence for our model proposed in Experiment 1.

Experiment to be simulated Yoshida & Smith (2003) conducted three experiments showing ontology difference between Japanese and English monolingual children. Here we introduce their second experiment to be simulated. Participants of Yoshida & Smith's experiment were 3-year-old English and Japanese monolingual children. Experimenters presented them exemplars with pipes like legs of animates and named it novel label (e.g. in Japanese 'Kore-wa __ dayo', in English 'This is __.'). Experimenters did not tell any syntactic cue like 'iru/ aru' which tells children animacy of the label. Then experimenters presented them test objects and asked them whether the test object had a novel label (e.g. in Japanese 'Kore-wa __-kana?', in English 'Is this __ ?'). Exemplars and test objects were under control to be matched or not matched in three perceptual features (Table 4).

The results showed different responses between English speakers and Japanese speakers. English speakers tended to generalize novel labels to test objects matched in shape, but Japanese speaker did not. Yoshida & Smith claim that this difference reflects the difference in ontological category criterion between English and Japanese. Furthermore they proposed 'Shifting ontological categories' hypothesis. Their hypothesis claims that the boundaries of ontological categories shift on the individuation continuum proposed by Lucy(1992) because of syntactic cues. We simulated this experiment with our results of Experiment 1 to provide quantitative evidence.

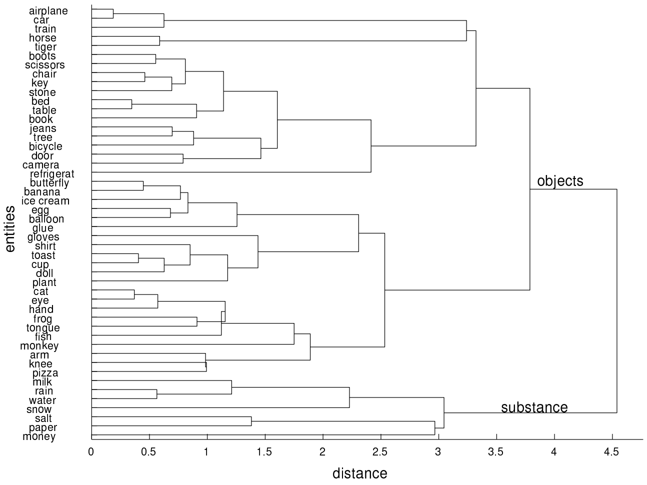

Figure 2: The result of cluster analysis for the English condition. We estimated 'objects' cluster and 'substance' cluster in superior hierarchy.

Method

The task of Yoshida & Smith's experiment is to respond 'yes' or 'no' when two feature-controlled objects is presented. We assumed that the objects have three features shape, color and texture, and the other features will have no difference between two objects. We also assumed children make 'yes' response based on the psychological distance between two objects. 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).

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

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

$b > 0$ 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 4). w_{il} is the value of ith perceptual weight of l ($l \subset \{English, Japanese\}$) participant(see also Table

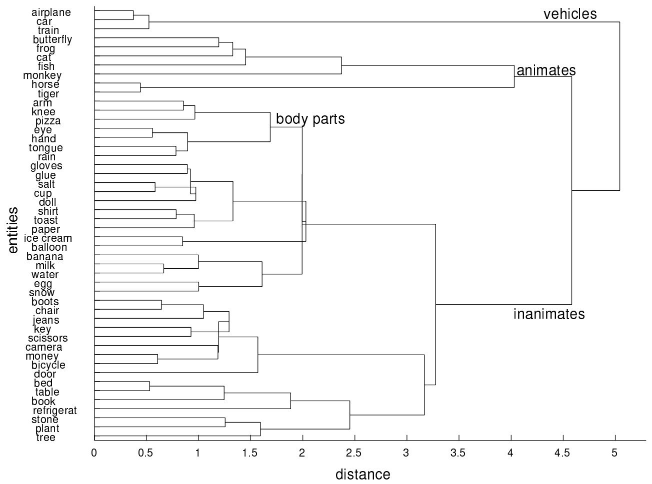

Figure 3: The result of cluster analysis for the Japanese condition. We estimated 'animates' cluster and 'inanimates' cluster in superior hierarchy.

Table 4: 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

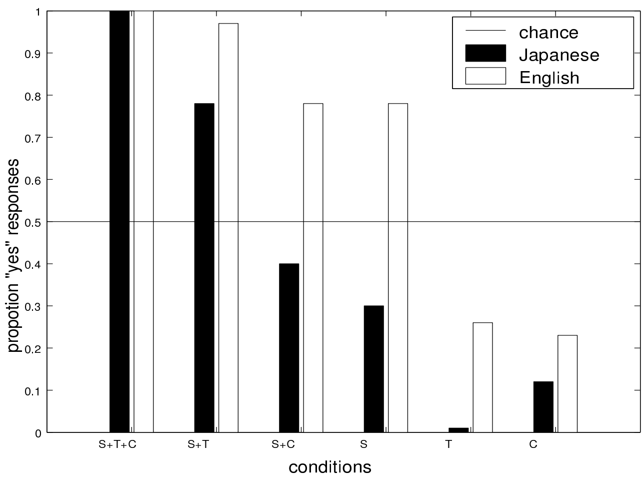

Figure 4: The behavioral data of Yoshida & Smith (2003). The English speakers categorize stimuli based on shape, while the Japanese speakers categorize them based on multiple features.

3). 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 & 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 (3). 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 inanimate objects and the Japanese speakers categorize them as animates. These results provide the evidence of BSH because they suggested the difference of criteria between English and Japanese. In this point of view, our model fitted the behavioral results well, so we could provide the simple model which accounted for the behavioral experiment based on the results of Experiment 1.

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, shape bias is stronger with solid objects than nonsolid ones, and Children speaking English show stronger shape bias than those speaking Japanese. These findings are explained by postulating children's knowledge on ontological category, and linguistic and cultural influence on the boundary of ontological categories. For example, Yoshida & Smith (2003) proposed 'boundary shift hypothesis' to account for the effect of language on children's object categorization, based on an idea of 'individuation continuum' proposed by Lucy (1992).

This study proposed a theory on the underlying mechanism of this boundary shift on the individuation continuum. Our theory proposes that (1) individuation continuum is not a special and abstract dimension, but an emergent property derived from multidimensional perceptual and linguistic features, and (2) boundary shift by different languages 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 multidimensional feature space, such as PCA. To evaluate whether our theory can account for the behavioral findings, we conducted a survey to obtain the multidimensional feature space of objects, and a series of quantitative analysis to obtain the language

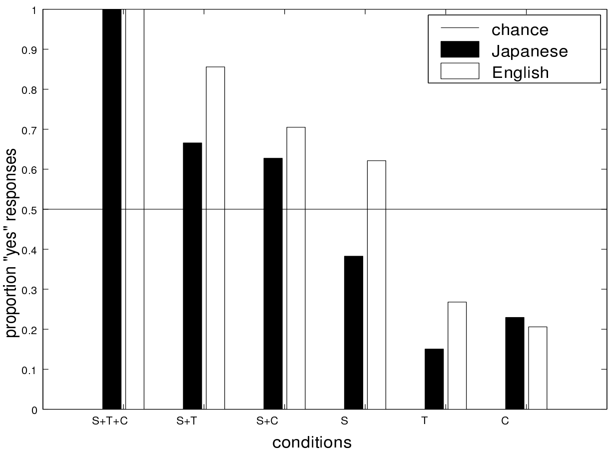

Figure 5: The result of simulation.
The coefficient of determination of 12 responses pattern (R^2) is 0.916

specific ontological spaces. Without linguistic features, the compressed perceptual space spanned by two principal components was organized by objects' solidity or size. Thus, solidity-dominant space can be derived from perceptual feature space, but there was no principal component representing 'individual continuum'. This result may be consistent with Soja, et al. (1991) showing that 2-year-old children seemed to have solidity ontological knowledge before they could distinguish count/mass syntax. However, given our data were obtained from adults, we need more direct evidence to support Soja et al.'s finding. 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-domininant and color-and-texture-weighted, which is consisitent with Yoshida & Smith (2001, 2003) showing that 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 could 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 between English and Japanese speaking children with various different stimulus conditions could be simulated quite well.

Expanded 'boundary shift hypothesis' Our theory is beyond a simple quantitative implementation of boundary shift on the individuation continuum. It 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

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.

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.

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.

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.