Title
Ad-hoc Meaning Substitution Causes Shape and Material Biases: Computational Explanation for Emergence of Word Learning Biases

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 28(28)

ISSN
1069-7977

Authors
Kurosaki, Kosuke
Omori, Takashi

Publication Date
2006

Peer reviewed

Ad-hoc Meaning Substitution Causes Shape and Material Biases: Computational Explanation for Emergence of Word Learning Biases

Kosuke Kurosaki (kurosaki@complex.eng.hokudai.ac.jp)
Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido, JAPAN

Takashi Omori (omori@lab.tamagawa.ac.jp)
Intelligent Robot Research Center, Tamagawa University Research Institute, Tokyo, JAPAN

Abstract

We illustrated a computational mechanism of *shape bias* and *material bias* with "ad-hoc meaning substitution (AMS)" hypothesis and verified it by computer simulations. AMS represents that when given a novel word and a instance object/substance, children substitute a known noun meaning nearest to the instance and the instance itself as an ad-hoc template of the novel noun meaning. The substitution enables fast mapping and subsequent vocabulary spurt. To describe internal process of AMS, we introduced "word distributional prototype (WDP)" as an explicit representation of word meaning with an inductive learning function. Simulation 1 revealed that when neural networks with WDP and AMS were given biased vocabularies reflecting those of young children, it demonstrated shape, material, and overgeneralized shape biases which means wrong shape bias over material bias. This result suggests that the triad of word meaning induction, AMS, and early biased vocabulary is essential for emergence of the biases. Simulation 2 introduced a notion of *maturity* that denote a degree of learning convergence for each word meaning, and then networks showed neither shape nor material bias during early small vocabulary. This result indicates that the age of bias emerges is decided by the maturity. These results suggest that phenomena concerning shape and material biases are explicable with the simple ad-hoc learning mechanism instead of meta learning or built-in language-specific ones.

Introduction

When we encounter a novel word such as *Gavagai* and guess its meaning, too many logically possible meanings exist (Quine, 1960). Nevertheless, children can estimate words meanings very fast (Carey & Bartlett, 1978). Such *fast mapping* can't be explained by existing machine learning algorithms based on trial and error. The ability appears after child's productive noun vocabulary exceeds about 50 words, and the vocabulary starts to grow quite rapidly(*vocabulary spurt*).

To explain these phenomena, developmental psychologists have suggested *word learning biases*. When applying a novel name to an object, the biases make children focus on particular features instead of other possible features and estimate words' meaning accurately (e.g. Landau, Smith, & Jones, 1988; Soja, Carey, & Spelke, 1991). Problem is that these biases are just phenomenological explanations that can't explain why they exist or how they are processed in human brain. In this paper, we try to illustrate a computational mechanism of *shape bias* (Landau, Smith, & Jones, 1988) and *material bias* (Dickinson, 1988; Soja et al., 1991). Those biases are investigated by novel noun generalization task (Samuelson, 2002; Samuelson & Smith, 1999). In the task, experimenter prepares solid and nonsolid sets. Each set consists of three stimuli: a novel target stimulus, shape-match stimuli that have the same *shape* as the target stimulus, and material-match stimuli that have the same *material*. First, an experimenter assigns a novel noun to the target stimulus in front of a subject. Next, the experimenter presents a corresponding shape-match stimulus and a material-match stimulus together and asks the subject to select one that can be called by the same noun as the target stimulus. When either is selected significantly more often by some subjects, we conclude that subjects have a bias to generalize novel nouns based on similarity in shape or material.

Shape bias is a behavior that when a novel *solid* target stumulus is named with a novel noun, people tend to extend the noun to shape-match stimulus (Landau, Smith, & Jones, 1988). Material bias is a behavior that when a novel *nonsolid* target stimulus is named with a novel noun, they tend to generalize the noun to the material-match stimulus (Dickinson, 1988; Soja et al., 1991). We can summarize some experimental findings as below: (1.1.1a) shape bias doesn't appear to solid stimuli when subjects have small vocabulary (Samuelson & Smith, 1999; Smith, 1995); and (1.1.1b) it does after middle vocabulary (Landau, Smith, & Jones, 1988; Soja et al., 1991; Dickinson, 1988; Imai & Gentner, 1997); (1.1.2a) material bias doesn't appear during small vocabulary (Samuelson & Smith, 1999); or (1.1.2b) "overgeneralized shape bias," which means novel noun named to nonsolid target stimulus is extended also to shape-match stimuli, happens to appear during small vocabulary (Samuelson, 2002); and (1.1.2c) material bias appears after sufficiently large vocabulary (Soja et al., 1991; Dickinson, 1988).

From these findings, we argue that the biases result from simple learning ("learned bias account [LBA]") because: (1.2a) The fact that shape and material biases appear after children have acquired certain number of words indicates that they emerge as a consequence of vocabulary learning; (1.2b) emergence of the overgeneralized shape bias suggests that both biases consist of common mechanism instead of separate modules and it causes the overgeneralization; (1.2c) both biases arise in almost identical situations except for the *solidity* of target stimuli and it supports possibility of their common mechanism.

Linda Smith and colleagues led LBA studies (Smith, 1995; Samuelson & Smith, 1999; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002; Samuelson, 2002; Colunga & Smith, 2005). They explained that when experiencing word learning, a "nonlinear attentional system (NAS)" organized attention to particular dimensions (e.g. shape dimensions), and the selective attention resulted in word learning bias (Smith, 1995). Recently, they advanced their hypothesis as "higher-order abstraction (HOA)" (Smith et al., 2002). It explains that from learned noun meanings, children abstract higher-order knowledge which represents meaning nouns generally have; children use it as a template to estimate meanings of novel nouns, which accelerates vocabulary acquisition; when children with this mechanism experience early vocabulary that is dominated by nouns organized by similarity in shape (Samuelson & Smith, 1999), learned higher-order knowledge becomes shape-based, which subsequently encourages acquisition of shape-based meanings.

Their hypothesis is meaningful because it intended to illustrate mechanism that had been compressed into NAS. But their hypothesis contains two points of debate: the age at which cognitive functions become available and the concreteness of their hypothesis. First, Smith (1989) demonstrated that children begin to pay selective attention to particular sensory dimensions in nonnaming categorization tasks after five years old. So we consider it difficult to explain shape and material bias at 24 months by NAS, although experiencing novel nouns may encourage the organization of selective attention (Smith, 1995). And they also didn't provide evidence that HOA is possible at 24 months[1]. The second issue about concreteness means that their hypothesis still retains some unclear points. Therefore though they verified NAS or HOA by simulations with neural networks (Smith, 1995; Samuelson, 2002; Colunga & Smith, 2005), it remains unclear whether those studies could have verified their hypothesis because they didn't describe its essence, that is, how to acquire the attention or abstraction.

So, in this study we manifest requirements for an alternative hypothesis of shape and material biases as follows: (1.3a) concrete description of computational process of the biases; (1.3b) explanation for why their emergence depends on children's vocabulary size; (1.3c) explanation for the existence of interference between them; (1.3d) clarification of essential factors that enable shape bias to appear at 24 months. Based on them, we hypothesize about these biases as below. Shape and material biases arise from simple learning and consist of two common primitive abilities available from infancy: (1) ability to learn the meanings of words by induction, and (2) ability to instantly estimate novel noun meaning based on already learned words meanings and given input stimulus with the noun. We call the second ability "ad-hoc meaning substitution (AMS)." Both have no specific mechanism to generate biased behavior. But when exposed to the statistically biased vocabulary of toddlers, learner with the two abilities shows shape and material biases. This triad of the word meaning induction, AMS, and the early biased vocabulary is essential for the emergence of biases. But this triad can cause stable biases from very early stages of development, while young children actually show inefficient word learning and no bias (1.1.1a; 1.1.2a). The delay of bias emergence is caused not by the triad but by a secondary factor: *maturity* in learning word meanings.

[1] Children begin to use words of superordinate-level categories at about four years. If the knowledge of a superordinate-level category word is abstracted from the knowledge of basic-level category words, the abstraction process should resemble HOA and both available ages can hardly be expected to be so different.

Proposed Hypothesis

Input and Word Representation

Following Samuelson (2002), input consisted of three attributes: SOLIDITY, SYNTAX, and FEATURE. Their attributes are represented as 3, 3, and 30 dimensions, respectively. SOLIDITY, which denotes solidness of objects/substances presented to learners, has three discrete attributes, SOLID, NONSOLID, and AMBIGUOUS[2]. The attributes were represented as vectors of (0.95, 0, 0), (0, 0, 0.95), and (0, 0.95, 0), respectively. SYNTAX, which represents a contextual attribute in a sentence given in parallel with the named object[3], has three discrete attributes: COUNT, MASS, and AMBIGUOUS[4]. They are represented as vectors of (0.95, 0, 0), (0, 0, 0.95), and (0, 0.95, 0), respectively. FEATURE is an attribute that denotes other information and consists of three attributes: SHAPE, MATERIAL, and OTHER[5]. They are represented as separate 10-dimentional vectors, respectively.

We assumed children constructed category knowledge for nouns. Based on the definition of FEATURE, we defined three categories: SHAPE-BASED, MATERIAL-BASED, and OTHER-BASED, organized on the basis of similarity in SHAPE, MATERIAL, and OTHER, respectively. These categories can overlap for each noun. Inputs belonging to a SHAPE-BASED category had arbitrary but constant values of SHAPE dimensions and random values of MATERIAL and OTHER dimensions. To the fixed SHAPE dimensions, random noises between [-0.05, 0.05] was added. Inputs belonging to MATERIAL- and OTHER-BASED categories were made in the same way as SHAPE-BASED category, respectively.

Word Distributional Prototype (WDP)

We introduced explicit representation of word meaning, which enable us to describe process of AMS concretely. We assume that a word meaning is defined as an extended *prototype* (Rosch & Mervis, 1975) that consists of distribution formed by inputs co-occurring with the word

[2] AMBIGUOUS denotes borderline cases in which SOLIDITY is neutral or both SOLID and NONSOLID are possible.

[3] Broad sense of syntax such as articles and determiners.

[4] SYNTAX AMBIGUOUS also denotes borderline cases.

[5] OTHER includes miscellaneous information other than SHAPE and MATERIAL: color, taste, temperature, fun, etc.

(Kurosaki & Omori, 2005a,b). For example, *banana* is generally used for crescent-shaped yellow fruits, but not for red or globular things. It's so sensitive to shape and usage information that even slight discrepancy in them is unacceptable. Meanwhile, such information as emotion isn't crucial for the recognition of *banana*, though it's also presented simultaneously with the name. Such characteristics of word meaning can be expressed by two factors: mean value in each input dimension and allowable deviation from the mean value. Children must learn them based on experience.

So we used a multidimensional normal distribution as one of the simplest possible representations for a word's meaning to fulfill the above requirements. We called it word distributional prototype (WDP). Its mean vector denoted the word's standard appearance in input space. Its variance matrix denoted the allowable range of fluctuation from the mean vector. We used diagonal variance matrices assuming that input dimensions had no correlation each other. Though the simplification might have some problems, we thought WDP was sufficient for word meaning representation of young children.

WDP described each word meaning as below. $\vec{x} \in \Re^{\mathrm{M}}$, $\vec{x} = (x_1 \ x_2 \ \cdots \ x_{\mathrm{M}})^{\mathrm{T}}$ is an input vector in M-dimensional input space, j is an ID number for $\mathrm{WDP_j}$, $\vec{\mu}_{\mathrm{j}} \in \Re^{\mathrm{M}}$, $\vec{\mu}_{\mathrm{j}} = (\mu_{\mathrm{j}\,1} \ \mu_{\mathrm{j}\,2} \ \cdots \ \mu_{\mathrm{j}\,\mathrm{M}})^{\mathrm{T}}$ is a mean vector of $\mathrm{WDP_j}$, $\mu_{\mathrm{j}\,\mathrm{i}} \in \Re$ is a mean value of input unit i of $\mathrm{WDP_j}$, $\Sigma_{\mathrm{j}} \in \Re^{\mathrm{M}} \times \Re^{\mathrm{M}}$ is a diagonal variance matrix of $\mathrm{WDP_j}$, and $\sigma_{\mathrm{j}\,\mathrm{i}} \in \Re$ is a standard deviation of input unit i of $\mathrm{WDP_j}$. Then, likelihood of $\mathrm{WDP_j}$ to an input $\vec{x}$ is calculated as:

$$p_{\mathrm{j}}(\vec{x}) = \frac{1}{(2\,\pi)^{\frac{\mathrm{M}}{2}}|\Sigma_{\mathrm{j}}|^{\frac{1}{2}}} \exp\Big(-\frac{1}{2}(\vec{x}-\vec{\mu}_{\mathrm{j}})^{\mathrm{T}}\Sigma_{\mathrm{j}}^{-1}(\vec{x}-\vec{\mu}_{\mathrm{j}})\Big). \quad (1)$$

WDPs were trained by following algorithm: (2a) Initially, each word is paired with an individual WDP; (2b) Given an input vector $\vec{x}$ and a corresponding word, all WDPs calculate likelihood $p_{\mathrm{j}}(\vec{x})$ for the input, and the winning $\mathrm{WDP_c}$ that outputs the highest likelihood is chosen; (2c) If the $\mathrm{WDP_c}$ is the correct paired one with the given word, then it learns to increase its likelihood for the input (Eq. (3) and (4)), and the others don't; (2c') If $\mathrm{WDP_c}$ is incorrect WDP for the given word, then it learns to decrease its likelihood $p_{\mathrm{j}}(\vec{x})$ for the input. Its update rules correspond to those of the opposite direction of (2c). The correct WDP, which is paired with a given word and should give the highest likelihood, simultaneously learns by the update rule (2c); (2d) Repeat (2b), (2c), and (2c') depending on word inputs. The learning corresponds to extension of "learning vector quantization (LVQ)" (Kohonen, 1995). To maintain stable learning, we set lower limit of $\sigma_{\mathrm{j}\,\mathrm{i}}$ to 0.1 and range of $\mu_{\mathrm{j}\,\mathrm{i}}$ to $[-1,\,1]$. The initial value of every $\sigma_{\mathrm{j}\,\mathrm{i}}$ is set to a sufficiently large value so that all WDPs don't output higher likelihood to particular inputs in the initial state. Loss function $\varepsilon_{\mathrm{c}}(\vec{x})$ and update rules for each parameter are defined as:

$$\varepsilon_{\mathrm{c}}(\vec{x}) = -\log\left(p_{\mathrm{c}}(\vec{x})\right) \quad (2)$$

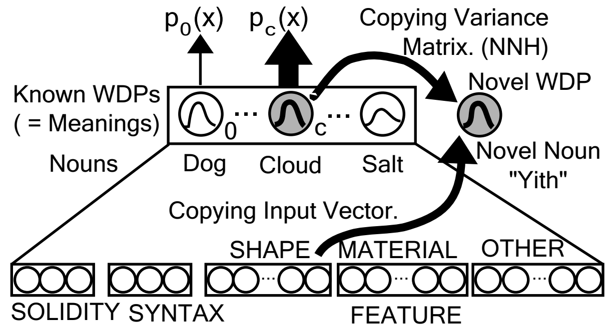

Figure 1: Ad-hoc meaning substitution (AMS).

$$\Delta\mu_{\mathrm{c}\,\mathrm{i}} = -\alpha\frac{\partial\varepsilon_{\mathrm{c}}(\vec{x})}{\partial\mu_{\mathrm{c}\,\mathrm{i}}} = -\alpha\frac{\mu_{\mathrm{c}\,\mathrm{i}} - x_{\mathrm{i}}}{{\sigma_{\mathrm{c}\,\mathrm{i}}}^2} \quad (3)$$

$$\Delta\sigma_{\mathrm{c}\,\mathrm{i}} = -\beta\frac{\partial\varepsilon_{\mathrm{c}}(\vec{x})}{\partial\sigma_{\mathrm{c}\,\mathrm{i}}} = -\beta\left(\frac{1}{\sigma_{\mathrm{c}\,\mathrm{i}}} - \frac{(\mu_{\mathrm{c}\,\mathrm{i}} - x_i)^2}{{\sigma_{\mathrm{c}\,\mathrm{i}}}^3}\right). \quad (4)$$

Ad-hoc Meaning Substitution (AMS)

When a child is given a novel objects/substances and a corresponding noun, the single experience may indicate the mean vector of the noun, but not the standard deviation. Nevertheless, the existence of word learning bias demonstrates that children do complement the missing information of standard deviations in some way. Children must do it by use of mechanisms nonspecific to vocabulary learning because as discussed in introduction, the biases seems to be learned.

We explain a process of novel noun generalization as below. Children determine the noun's meaning based on meaning of a known noun: The variance matrix of a known noun WDP, which output the highest likelihood to the input information (Eq. (5)), is copied as the variance matrix of the novel noun WDP (Eq. (6)). We call the copying process "nearest neighbor hypothesis (NNH)" (Fig. 1), which denotes that WDP which has the nearest Mahalanobis distance to the input vector was chosen. Meanwhile, the input vector presented with the novel noun was copied as the mean vector of the novel noun WDP (Eq. (7)). Though they may not work always correctly, algorithm which always complements the missing information correctly doesn't exist. We suggest that one of probable strategies for the complements is to substitute the nearest known word and input vector for the information. We call the ad-hoc construction of novel noun meaning ad-hoc meaning substitution (AMS).

$$c = \arg\max_{\mathrm{j}} \ p_{\mathrm{j}}(\vec{x}) \quad (5)$$

$$\Sigma_{\mathrm{new}} = \Sigma_{\mathrm{c}} \quad (6)$$

$$\vec{\mu}_{\mathrm{new}} = \vec{x} \quad (7)$$

Table 1: Structural ratio in early vocabulary evaluated in this study. Upper table shows ratio of each attribute value in SOLIDITY, SYNTAX, and FEATURE. Lower tables show conditional ratio of each attribute value.

	SOLIDITY			SYNTAX			FEATURE		
	SOLID	NON-SOLID	AMBIGUOUS	COUNT	MASS	AMBIGUOUS	SHAPE	MATERIAL	OTHER
	.63	.04	.32	.74	.10	.16	.48	.16	.39
SOLID				.88	.03	.09	.71	.09	.19
NONSOLID				.07	.43	.50	.07	.86	.14
AMBIGUOUS				.56	.19	.25	.24	.36	.45

Simulation 1

Inductive Learning Phase

We divided the word learning process into two phases. In "inductive learning phase," learners received sets of words and corresponding input and constructed WDPs with Eq. (3) and (4). Then in "generalizing phase," they were conducted novel noun generalization tasks on with AMS and the learned WDPs. We prepared six groups whose learners were given 18, 50, 102, 213, 281, and 312 words, respectively (see Samuelson & Smith, 1999). We called them groups 1, 2, 3, 4, 5, and 6, respectively. Each learner in a group learned different words each other.

Vocabulary used in this phase should contain words generally heard and produced by young children. Hence, we used MCDI (Fenson et al., 1994), which include a typical vocabulary for 16- to 30-month-olds. Based on previous studies (Samuelson & Smith, 1999; Samuelson, 2002), we used 312 words from nine categories[6] of MCDI. We evaluated the words ourselves in terms of SOLIDITY, SYNTAX, and FEATURE (Table 1) so that the evaluated structural ratio in the vocabulary was consistent with previous study (Samuelson & Smith, 1999) and made the 36-dimensional input data.

We set learning parameters α and β to 0.001, the initial values of $\mu_{\mathrm{j\,i}}$ to [0.4999, 0.501] and $\sigma_{\mathrm{j\,i}}$ to 1.0. We prepared 50 instances for a word, and a corresponding WDP learned the instances. In an *epoch* , a learner experienced all instances of all words prepared for the learner. The learning iterated the epoch 30 times. We confirmed that the learning of all groups almost converged at 30th epoch, and that the learned parameters in each WDP were correctly estimated to form distribution of corresponding noun.

Generalizing Phase

We prepared solid and nonsolid sets and conducted novel noun generalization tasks separately with them. Each set had 21 stimulus sets, and a stimulus set consisted of a target stimulus and 20 pairs of shape-match and material-match stimuli. Shape-match stimuli were made as input vectors that had same SOLIDITY and SHAPE values as the target stimulus, random MATERIAL and OTHER values, and no SYNTAX[7] value. Random noises between $[-0.05, 0.05]$ were added to the SHAPE values. The material-match stimuli were made in the same way. Both sets were initially prepared and shared by all learners in all groups.

In our model, shape choice probabilities for each group were calculated as below. Given a novel target stimulus and a novel noun, $\mathrm{Learner}_{\mathrm{g\,i}}$ of $\mathrm{Group}_{\mathrm{g}}$ made a novel noun WDP applying AMS. Then we gave $\mathrm{Learner}_{\mathrm{g\,i}}$ 20 pairs of shape-match and material-match stimuli and compared which evoked higher likelihood for each pair. Shape choice probability to Target $\mathrm{stimulus}_{\mathrm{j}}$ by $\mathrm{Learner}_{\mathrm{g\,i}}$ was calculated by $p(g, i, j) =$ (winning number of shape-match stimuli) / (total number of pairs). Shape choice probability to all target stimuli by $\mathrm{Learner}_{\mathrm{g\,i}}$ was calculated by $p_{\mathrm{L}}(g, i) = \left(\sum_j p(g, i, j)\right)$ / (total number of target stimuli). Mean probability of shape choice in $\mathrm{Group}_{\mathrm{g}}$ was calculated by $p(g) = \left(\sum_i p_{\mathrm{L}}(g, i)\right)$ / (total number of learners).

For the solid set, t-test confirmed that the shape choice probability for each group was significantly larger than chance: They showed shape bias; $t(20)$=12.95, $p \prec .001$; $t(20)$ =23.02, $p \prec .001$; $t(20)$ =26.77, $p \prec .001$; $t(20)$= 27.98, $p \prec .001$; $t(20)$=27.89, $p \prec .001$; and $t(20)$=32.45, $p \prec .001$, respectively. For the nonsolid set, the probabilities of shape choice for groups 1, 2, and 3 were significantly larger than chance even for the nonsolid set: They showed overgeneralized shape bias; $t(20)$=2.71, $p \prec .05$; $t(20)$=3.27, $p \prec .01$; and $t(20)$=4.68, $p \prec .001$, respectively. But those for groups 4, 5, and 6 were significantly smaller than chance: They showed material bias; $t(20)$ $=-3.54$, $p \prec .01$; $t(20)=-3.57$, $p \prec .01$; and $t(20)=-4.82$, $p \prec .001$, respectively (Fig. 2).

Results of the stable shape bias during somewhat larger vocabulary (1.1.1b), the overgeneralized shape bias during small vocabulary (1.1.2b), and the material bias during large vocabulary (1.1.2c) is explained as follows. WDPs having same SOLIDITY as a target stimulus, which were SOLID WDPs for solid set and NONSOLID WDPs for nonsolid set, were likely to be chosen as the nearest WDP by NNH because they output higher likelihood. Since SOLID WDPs tended to be SHAPE-

[6] animals, vehicles, toys, food and drink, clothing, body parts, small household items, furniture and rooms, outside things

[7] SYNTAX information was withheld because we intended to compare our results to previous experimental results without syntax (see Samuelson & Smith, 1999; Samuelson, 2002).

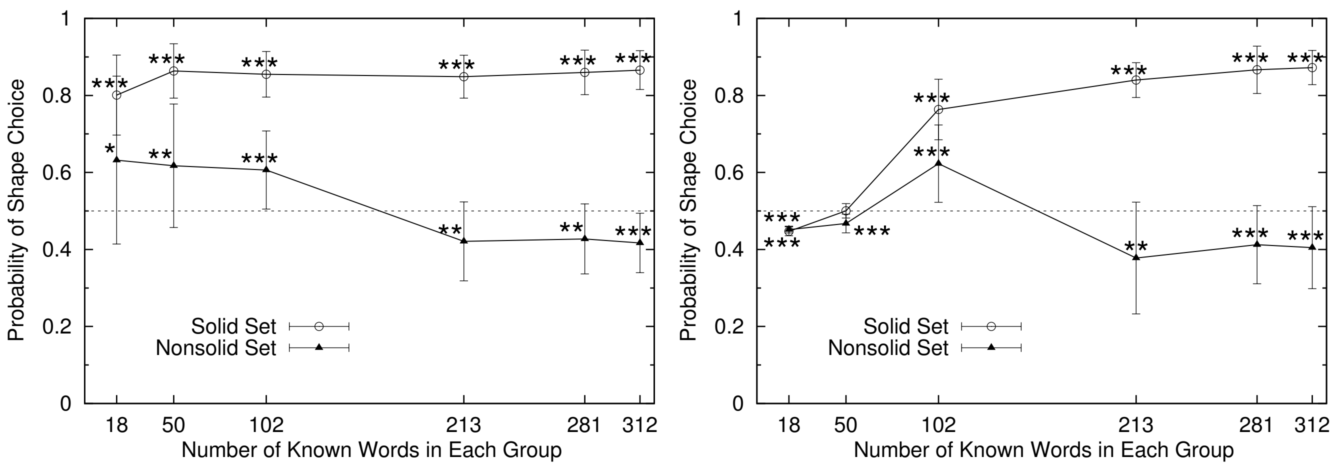

Figure 2: Probabilities of shape choices in generalizing phase of simulation 1 (left) and 2 (right). Vertical lines depict standard errors. Horizontal broken line depicts chance level (=.5). $^*p \prec .05$. $^{**}p \prec .01$. $^{***}p \prec .001$.

BASED (see Table 1), novel WDP was inclined to copy the SHAPE-BASED variance matrix by NNH and it led to shape bias. Meanwhile, NONSOLID WDPs couldn't be the nearest to nonsolid set during small vocabulary. NONSOLID WDPs were far less than SOLID WDPs and the NONSOLID field was too *sparse* (see Table 1) for them to be chosen as the nearest. Therefore SOLID WDPs could be chosen alternatively, though they had a disadvantage of having different SOLIDITY from the target stimulus. Then as with the case of solid set, it led to overgeneralized shape bias. When the sparseness disappeared, NONSOLID WDPs could be the nearest to nonsolid stimuli and it caused material bias.

Simulation 2

Inductive Learning Phase

We introduced *maturity* to produce a situation in which convergence of word meaning learning was more insufficient for learners who had small vocabularies than large vocabularies. We realized the situation by reducing the number of instances for each word depending on the vocabulary size of each group: 5, 10, 20, 40, 50, and 55, respectively, in ascending order of group number. Other conditions were identical to simulation 1.

Learning hadn't converged even at 30th epochs in groups 1 and 2. Their parameters hadn't been estimated correctly, compared with simulation 1. Though the learning in group 3 apparently converged better than groups 1 and 2, their parameters hadn't been estimated as well as simulation 1. After group 4, their learning had almost converged and their parameters had been estimated well enough. From these results, the expected effects by introducing the maturity were realized.

Generalizing Phase

The generalizing phase procedure was identical to simulation 1. For the solid set, the probability of shape choice for group 1 was significantly lower than chance even for the solid set, demonstrating material bias: $t(20)= -20.39$, $p \prec .001$. But in group 2, there was no significant differences to chance, that is, it showed no bias: $t(20)= 0.081$, $p \succ .05$. After that, the probabilities for groups 3, 4, 5, and 6 were significantly higher than chance, showing shape bias: $t(20)=15.01$, $p \prec .001$; $t(20)=33.59$, $p \prec .001$; $t(20)=26.63$, $p \prec .001$; and $t(20)=37.49$, $p \prec .001$, respectively. For the nonsolid set, the shape choice probabilities for groups 1, 2, 4, 5, and 6 were significantly smaller than chance, showing material bias; $t(20)=-24.54$, $p \prec .001$; $t(20)=-5.98$, $p \prec .001$; $t(20)=-3.77$, $p \prec .01$; $t(20)= -3.86$, $p \prec .001$; and $t(20)=-4.01$, $p \prec .001$, respectively. But in group 3, we observed a significantly larger shape choice than chance, demonstrating shape bias even for the nonsolid set; $t(20)=5.48$, $p \prec .001$ (Fig. 2).

Only the results during small vocabulary were different from simulation 1. So, we discuss the disappearance of robust shape bias during small vocabulary (1.1.1a, 1.1.2a) and the alternative appearance of material bias. WDPs in the very early groups were closer to hyperspheres because learning of their variance matrices hadn't progressed at all due to the effect of maturity. Novel WDPs that copied the variance matrix output low likelihood to every input equally. It led to no shape and material bias. Actually the robust shape bias during small vocabulary in simulation 1 disappeared and $p(g = 1)$ and $p(g = 2)$ became almost close to chance. But statistically, early groups showed significant material choice. The problem seemed to be due to our simplified setting in each group. Exactly the same setting of the maturity and the number of words for learners in a group caused almost same $p_{\mathrm{L}}(g, i)$ and their small standard deviation. In that case, just slight difference than chance lead to shape/material bias. But the problem could be resolved by modifying the current simplification. In sum, we could replicate some previous findings in simulation 1 (1.1.1b; 1.1.2b; 1.1.2c) and the others in simulation 2 (1.1.1a; 1.1.2a) and also explain the process computationally.

Position of AMS

HOA and AMS were proposed as LBA instead of accounts based on innateness of the biases. It's well known that children reason about animals and plants by analogy with humans familiar to them (Inagaki & Hatano, 1987). Since AMS means that children construct knowledge of unfamiliar nouns based on nouns familiar to them, AMS resembles in it. Therefore AMS is natural for them as a method to estimate novel noun meaning. And AMS has simpler process, but it can illustrate wider phenomena of shape bias, material bias, and overgeneralized shape bias than other accounts. Besides, since applicability of HOA to them all is unclear and it has the age problem, we argue that AMS is more appropriate explanation for the biases. But we shouldn't exclude HOA because it can apparently explain acquisition of words belonging to superordinate categories or more abstract concepts well. And it is quite plausible that HOA and AMS cooperatively engage in bias emergence with other functions (e.g. *theory of mind*). We consider AMS to be the simplest function within them, and thus it takes an important role as the foundation of bias emergence from the early stage of development.

Conclusion

In this paper, we presented an integrated explanation of ad-hoc meaning substitution (AMS) for behaviors that had been described separately as shape and material biases and verified it by computer simulations. Besides, to describe AMS's processing, we introduced word distributional prototype (WDP) with the inductive learning function. We consider the explicit meaning representation is valid methodology for illustrating the computational mechanism of word learning bias, which is deeply committed to word meaning.

Simulation 1 revealed that when learners that possess (1) WDP and (2) AMS were exposed to (3) early biased vocabulary, they showed shape, material, and overgeneralized shape bias. This result suggested that the triad is essential for emergence of the biases. Simulation 2 revealed that when maturity was introduced, learners showed neither shape nor material bias during the early small vocabulary. This result indicated that the period of bias emergence is decided not by the triad but by maturity. So we can reply to the requirements in introduction: (1.3a) AMS (especially, NNH); (1.3b) bias emergence is influenced by maturity and sparseness of SOLID and NONSOLID WDPs; (1.3c) shape and material biases derive from common mechanisms; (1.3d) the triad and maturity. Our results suggest that phenomena concerning shape and material biases, which have been explained by meta learning (HOA) or built-in language-specific mechanism, are explicable with a simple ad-hoc learning mechanism.

References

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17–29.

Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*, *112*(2), 347–382.

Dickinson, D. K. (1988). Learning names for materials: Factors constraining and limiting hypotheses about word meaning. *Cognitive Development*, *3*(1), 15–35.

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*, *5*, 242.

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: universal ontology and linguistic influence. *Cognition*, *62*(2), 169–200.

Inagaki, K., & Hatano, G. (1987). Young children's spontaneous personification as analogy. *Child Development*, *58*, 1013–1020.

Kohonen, T. (1995). *Self-organizing maps*. New York: Springer-Verlag.

Kurosaki, K., & Omori, T. (2005a). Computational modeling of word learning biases by using known word meanings. *Proceedings of the 9th IASTED International Conference on Artificial Intelligence & Soft Computing* (pp. 201–206), Calgary: ACTA Press.

Kurosaki, K., & Omori, T. (2005b). *Children construct novel word meaning ad-hoc based on known words: Computational model of shape and material biases*. Manuscript submitted for publication.

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

Quine, W. V. (1960). *Word and object*. Cambridge: MIT Press.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605.

Samuelson, L. K. (2002). Statistical regularities in vocabulary guide language acquisition in connectionist models and 15-20-month-olds. *Developmental Psychology*, *38*(6), 1016–1037.

Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: do ontology, category structure and syntax correspond? *Cognition*, *35*(3), 1–33.

Smith, L. B. (1989). A model of perceptual classification in children and adults. *Psychological Review*, *98*, 125–144.

Smith, L. B. (1995). Self-organizing processes in learning to learn words: Development is not induction. In C. A. Nelson (Ed.), *The Minnesota Symposia on Child Psychology, Vol. 28*. Mahwah, NJ: Erlbaum.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. K. (2002). Object name learning provides on-the-Job training for attention. *Psychological Science*, *13*(1), 13–19.

Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition*, *38*(2), 179–211.