



Title	Children construct novel word meaning ad-hoc based on known words : Computational model of shape and material biases
Author(s)	Kurosaki, Kosuke; Omori, Takashi
Citation	Cognitive Systems Research, 8(2), 110-130 <a href="https://doi.org/10.1016/j.cogsys.2006.06.002">https://doi.org/10.1016/j.cogsys.2006.06.002</a>
Issue Date	2007-06
Doc URL	<a href="http://hdl.handle.net/2115/20565">http://hdl.handle.net/2115/20565</a>
Type	article (author version)
File Information	COGSYS_kurosaki.pdf



[Instructions for use](#)

# Children Construct Novel Word Meaning Ad-hoc Based on Known Words: Computational Model of Shape and Material Biases

Kosuke Kurosaki<sup>a\*</sup>, Takashi Omori<sup>b</sup>

<sup>a</sup>Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido, JAPAN

<sup>b</sup>Tamagawa University Research Institute, Tamagawa University, Tokyo, JAPAN

Taking the stance that two well-known word learning biases (shape and material bias) are formed through learning (learned bias account, LBA), we illustrated a concrete computational mechanism with “ad-hoc meaning substitution (AMS)” hypothesis, and verified it by two computer simulations. AMS represents that when given a novel word and a corresponding instance, children create novel word meaning by using the known word meaning and the instance as an ad-hoc template. The AMS function enables fast mapping and vocabulary spurt. To describe the AMS process computationally, we introduced “word distributional prototype (WDP),” which is the explicit representation of word meaning with an inductive learning function. Simulation 1 revealed that when a network with WDP and AMS was given a biased vocabulary reflecting young children, it demonstrated shape, material, and overgeneralized shape biases. This result suggested that a triad of word meaning induction, ad-hoc meaning substitution, and early biased vocabulary is essential for the emergence of biases. Simulation 2 introduced the notion of *maturity* that denoted a degree of learning convergence for each word meaning, and then the network showed neither shape nor material bias during an early small vocabulary. This result indicated that the period at which each bias emerges is decided by maturity. Though AMS consists of simpler and valider mechanisms than those proposed in previous studies, it could reproduce behavior of shape and material biases and explain their emergence process clearly. These results suggest that phenomena concerning shape and material biases are explicable with a simple ad-hoc learning instead of meta-learning among LBA or innate language-specific ones.

## 1. Introduction

When we encounter a novel word such as *Gavagai* and guess its meaning, too many logically possible meanings exist (Quine, 1960). Nevertheless, children as well as adults can estimate the meaning of words relatively well. Although children have little knowledge of the world, their learning of words is so fast that they can achieve estimation even with a single experience (Carey & Bartlett, 1978). Such *fast mapping* can’t be explained by existing machine learning algorithms based on trial and error. However, children in the first stage of word acquisition don’t show such intelligence. They need to hear a word repeatedly before producing it, and they still often use words incorrectly. And the pace with which they acquire

meanings is slow. But after a children’s productive noun vocabulary exceeds 50 words, which happens around eighteen months, their vocabulary begins to grow quite rapidly. This is called *vocabulary spurt*.

To explain these phenomena, developmental psychologists have suggested *word learning biases*, which made children focus on particular features when applying a novel name to an object instead of other possible features. They also greatly narrow the infinite number of possible correspondences between a word and objects and enable children to estimate a word’s meaning more accurately (Markman, 1989; Markman & Hutchinson, 1984; Markman & Wachtel, 1988; Landau, Smith, & Jones, 1988; Clark, 1987; Soja, Carey, & Spelke, 1991). The problem is that these biases are just phenomenological explanations that can’t explain why they exist or how they are pro-

---

\*Corresponding author.

cessed in the human brain. We have to describe a more detailed computational mechanism of these biases not only to explain them but also to apply these findings to studies of education and language related disorders. In this paper, we try to closely illustrate a computational mechanism of two well-known biases: *shape bias* (Landau, Smith, & Jones, 1988) and *material bias* (Dickinson, 1988; Soja et al., 1991) especially in their early stages.

Shape bias is the behavior that when people hear a novel *solid* object named with a novel noun, they tend to extend the noun to objects similar in *shape* to the named object (Landau, Smith, & Jones, 1988). Material bias is the behavior that when people hear a novel *nonsolid* substance named with a novel noun, they tend to generalize the noun to the substance similar in *material* (Dickinson, 1988; Soja et al., 1991). Some studies have argued against “learned bias account (LBA),” which proposes that these biases result from simple learning. But we discuss the counterargument in section 6.2. We support LBA for the following three reasons. First, shape bias doesn’t appear in early stages of development. It only appears after children have acquired a certain number of words (Smith, 1995; Samuelson & Smith, 1999). Material bias also appears in the same manner, but later than shape bias (Soja et al., 1991; Colunga & Smith, 2005). This reason is necessary but not sufficient for supporting LBA; we can explain that specific modules for shape/material biases exist in human brains and that they begin to function along with development of the brain area concerned with word acquisition. The other reasons increase the persuasiveness of LBA. The second reason is existence of overgeneralization of shape bias to material bias. Samuelson (2002) reported that children extended novel nouns assigned to nonsolid substances to objects similar in shape to the named substance. The overgeneralization is not easy to explain if one assumes specific modules for those biases. Meanwhile, LBA may be able to explain it because such overgeneralization is just a behavior actually observed in machine learning. The third reason is that shape/material bias is behavior for solid/nonsolid objects and both arise in

almost identical situations except for their target stimuli’s *solidity*. This similarity indicates that they may possess the same internal mechanism. Based on these reasons, we assume in this paper that shape and material biases emerge from the same mechanism of simple learning; hence material bias is overgeneralized by shape bias and both appear after the learning of a certain amount of vocabulary.

Here, we propose a simpler computational hypothesis that can account for the behavior of biases and explain their emergence process from the LBA framework. Specifically, we hypothesize that shape and material biases arise from two internal general learning abilities that are available from infancy and one external factor - biased early vocabulary. The first internal general learning ability is the one to learn the meanings of words by induction. Though it is fundamental to word meaning learning, it isn’t enough to explain fast mapping. Consequently, we hypothesize another internal general learning ability: the ability to instantly estimate novel word meaning based on already learned word meanings and given encounters with the word. This substitution of word meaning from a known noun to a novel noun enables us to show fast mapping. We call the second ability “ad-hoc meaning substitution (AMS).” When children with the two abilities are exposed to the early biased vocabulary, which contain among solid/nonsolid words a high percentage of nouns organized by shape/material similarity, they show shape and material bias because AMS accounts for the substitution of meaning among such biased vocabulary (solid-shape, nonsolid-material) for meanings of novel solid/nonsolid nouns. In our model, this triad of the word meaning induction, the ad-hoc meaning substitution, and the learning of early biased vocabulary is a necessary and sufficient factor for the emergence of biases. But this triad can cause different behavior from children’s no biases: stable biases from the beginning of word learning. So we introduce a notion of *maturity* to inhibit the early stable bias. Here maturity means a degree of learning convergence for each word meaning, and relates to the developmental level of the nervous system, the experience of each word’s learn-

ing, and other factors. We posit that immaturity prevents effects of inductive learning and AMS and causes no bias at the beginning of word learning.

The format of this paper is as follows. In section 2, we summarize findings about word learning biases and present behavior our simulation should reproduce. And we introduce previous LBA studies and clarify distinction between their hypotheses and ours. In section 3, we explain AMS and discuss the learning process with it, and in sections 4 and 5 we explain the design of simulations and discuss their results. In section 6, we give a general discussion and conclude this study in section 7.

## 2. Previous Word Learning Bias Studies

### 2.1. Novel Noun Generalization Task

Novel noun generalization task investigates subject naming behavior when novel objects/substances and corresponding novel nouns are presented. Samuelson (2002), Samuelson & Smith (1999) used the following procedures to examine shape and material biases. Experimenter prepares two stimulus sets: solid and nonsolid. Each set consists of three stimuli: a novel nonsense target stimulus, two shape-match stimuli that have the same *shape* as the target stimulus, and two material-match stimuli that have the same *material*. First, the experimenter assigns a novel noun to the target stimulus in front of a subject. Next, the experimenter presents a pair of a shape-match stimulus and a material-match stimulus, asks children to select the one that can be called by the same noun as the target stimulus, and repeats the selection with 4 pair (2 shape-match stimuli  $\times$  2 material-match stimuli). When either is selected significantly more often by subjects, we conclude that subjects have bias to generalize novel nouns based on similarity in shape/material. We replicate this task in sections 4 and 5.

### 2.2. Summary of Findings about Word Learning Biases

In this subsection we summarize shape and material bias findings and show the behavior LBA

models should reproduce in simulation experiments. The emergence of shape bias are stable from 24-month-olds to adults (Soja et al., 1991; Dickinson, 1988; Imai & Gentner, 1997). But some studies reported that toddlers less than two years old didn't show shape bias (Samuelson & Smith, 1999). In contrast, material bias results aren't so stable: two- and two-and-a-half-year-olds showed material bias (Soja et al., 1991); so did three-, four-, and five-year-olds (Dickinson, 1988); meanwhile, toddlers ranging from 17 to 33 months old never showed material bias (Samuelson & Smith, 1999); two-, two-and-a-half-, and four-year-olds and adults who spoke Japanese and two-year-olds who spoke English showed material bias, but two-and-a-half-, four-year-olds, and adults who spoke English didn't (Imai & Gentner, 1997); three-, four-, and five-year-olds didn't show material bias when given such syntax information as determiners (Subrahmanyam, Landau, & Gelman, 1999). And surprisingly, Samuelson (2002) reported that children showed extraordinary shape bias for nonsolid stimulus, which is the exact opposite case of material bias. Here, we call such opposite bias "overgeneralized shape bias" because it looks like the overgeneralization of shape bias to material bias. If material bias is derived from an innate module independent of shape bias, such overgeneralization from shape bias is improbable. That kind of overgeneralized bias has also been reported as to rigidity, too (Samuelson & Smith, 2000; Samuelson et al., 2006). Existence of such overgeneralized biases strongly supports that they are phenomena neither specific to solidity nor poorly-reproducible, and indicates that they can be a key to clarification of bias mechanisms.

Before summarizing the findings above, we consider some related issues here. First, we discuss the mechanism of these biases under neutral syntax conditions. It is because we want to reveal the mechanism of children's naming bias to solid objects and nonsolid substances but the addition of syntax information (e.g., articles or determiners) complicates it (Dickinson, 1988; Subrahmanyam et al., 1999).

Next, only Imai & Gentner's results that English-speaking adults didn't show the material

bias aren't consistent with other findings. But they shouldn't be considered exact counterexamples of material bias because the results are odd, even for native English speakers (Bloom, 2000, p. 103). Besides, Hall (1996) reported that English-speaking adults (and children) concentrated on the *material* of novel *nonsolid* substances in situations in which they were asked to explain the substances. Though Hall's study wasn't word learning task, this material attention in language-related task supports existence of material bias for adults. These suggest that although Imai & Gentner's results are controversial, adults generally show material bias in language learning situation.

Finally, we discuss the correlation between vocabulary size and age in the previous word learning bias studies. We introduced the studies above by referring to the age of children because most of them had investigated biases from the aspect of age. However, the latest studies investigated them from the aspect of vocabulary size (Samuelson & Smith, 1999; Smith et al., 2002; Colunga & Smith, 2005), and we also consider that vocabulary size is essential for bias emergence. Here we note a report which showed high correlation between age and productive vocabulary size (Samuelson & Smith, 1999). Based on the correlation, we apply the findings about age to vocabulary size in this study.

So, combining these findings, we summarize behavior that models explaining shape and material biases to reproduce as below:

### (2.2.1) Solid stimuli

- (a) No shape bias during a small vocabulary (Samuelson & Smith, 1999)
- (b) Stable shape bias after a middle vocabulary (Samuelson & Smith, 1999)

### (2.2.2) Nonsolid stimuli

- (a) No material bias during a small vocabulary (Samuelson & Smith, 1999)
- (b) Overgeneralized shape bias during a small vocabulary (Samuelson, 2002)

- (c) Robust material bias after a sufficiently large vocabulary (Soja et al., 1991).

## 2.3. Previous LBA Studies

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 have explained word learning biases by associative learning. First, they proposed that when exposed to word learning experience, a "nonlinear attentional system" organized attention to particular dimensions by associative learning and that selective attention resulted in word learning biases (Smith, 1995). Recently, they explained the associative learning hypothesis by a 4-step model (Smith et al., 2002); A child experiences a pair of an instance and a word and learns mapping between them (step 1); he learns the meaning of each word through adequate experience of the pairs by general statistical learning (step 2); he abstracts higher-order knowledge from the learned word meanings, and such knowledge represents the meta meaning nouns generally have (step 3); and he uses the higher-order knowledge to estimate the meanings of novel nouns and to facilitate vocabulary acquisition (step 4). So far there is no specific mechanism that causes shape bias. However, when children experience early vocabulary which is dominated by nouns organized by similarity in shape (Samuelson & Smith, 1999), the learned higher-order knowledge becomes shape-based, which subsequently encourages the acquisition of shape-based meanings (i.e., shape bias). We call the multi-step abstracting function "higher-order abstraction (HOA)." The 4-step model/HOA has also been materialized by a neural network, a kind of associative learning machine (Samuelson, 2002; Colunga & Smith, 2005).

LBA seems to be plausible based on the three reasons in section 1. In the LBA framework, Smith et al.'s hypotheses are meaningful because they intended to illustrate the mechanism that had been compressed into the words "innate" or that hadn't been discussed. However, there may be other computational mechanisms under the LBA framework along with HOA. HOA is one

of the hypotheses that can reproduce and explain the biases, but the existence of the biases can't prove HOA per se. Which LBA-related hypotheses are more plausible should be discussed in terms of the following points:

- (2.3a) Reproductivity and illustrativity of bias findings
- (2.3b) Cognitive adequacy
- (2.3c) Prediction and its confirmation

The terms reproductivity and illustrativity of bias findings (2.3a) refer to whether a hypothesis can reproduce experimental results of solid stimuli (2.2.1) and nonsolid stimuli (2.2.2) and illustrate how the results arise. The term cognitive adequacy (2.3b) refers to whether cognitive faculties necessary for a hypothesis are available at 24 months of age, when shape bias has appeared. The terms prediction and its confirmation (2.3c) refer to whether a hypothesis can predict new phenomena and whether they can be confirmed by further experiments. Prediction and confirmation (2.3c) is the most effective in proving a hypothesis. However the aim of this paper is to propose another LBA hypothesis and its verification by computer simulation. Consequently, we focus on (2.3a) and (2.3b) in this paper, referring to (2.3c) in section 6.2 only.

In discussing cognitive adequacy (2.3b), we first consider the age at which children can attend to particular sensory dimensions in nonnaming categorization tasks. Smith (1989) demonstrated that children start to present selective attention after becoming five years old. From this result, we consider it difficult for 24-month-olds to acquire association between words and particular dimensions even in naming tasks (Soja et al., 1991). Although experiencing novel nouns may encourage the organization of selective attention (Smith, 1995), it is unlikely to bridge the cognitive gap between 24-month-old and five-year-old children. And we also feel the necessity to discuss the age of concept acquisition. Children generally begin to use words of superordinate-level categories at about four years of age. If the knowledge of a superordinate-level category word is ab-

stracted from the knowledge of basic-level category words<sup>2</sup>, the abstraction process should resemble the HOA proposed by Smith et al. (2002) and both available ages can hardly be expected to be so different. From the viewpoint of availability, especially for 24 months of age or earlier, we raise concerns about the cognitive adequacy of HOA.

As for reproductivity and illustrativity (2.3a), we attended to simulation results of HOA studies. Samuelson (2002) reproduced shape and overgeneralized shape bias but not material bias, while Colunga & Smith (2005) reproduced shape and material biases but not overgeneralized shape bias. The reason for their difference is not clear. Their different parameter setting might result in the different results. However, at least there is no study reproducing the three biases integratively. Moreover, they used at most two dozens of nouns in their computer simulation (Samuelson, 2002; Colunga & Smith, 2005), while Smith (1995) and Samuelson & Smith (1999) reported that children who had larger vocabularies started to show shape and material biases. An explanation of the gap between their vocabulary size is necessary in terms of reproductivity<sup>3</sup>. As to illustrativity (2.3a), though HOA proponents argued that correlation in vocabulary was necessary for bias

<sup>2</sup>Some researchers have claimed that superordinate categories exist from early infancy (Mandler, Bauer, & McDonough, 1991; Mandler & McDonough, 1993) and we agree with them. However, the issue here is not innate superordinate knowledge, but a function of newly abstracting knowledge through experience. Such knowledge and function are two different things which should be discussed separately.

<sup>3</sup>They may argue that using only a dozen words in their computer simulations has no problem if the small vocabulary has the same statistics as a larger vocabulary which Samuelson & Smith (1999) revealed and that children with a small vocabulary don't show biases because their small vocabulary doesn't have the statistics. But there are a few problems with the small vocabulary. First, there is no evidence that the statistical bias doesn't exist in such a small vocabulary. The statistical bias is rather likely to be strong, if children pick up a dozen words from strongly biased vocabulary. Second, according to our study (see section 4, 5), not only structural proportion within vocabulary but also vocabulary size is important for emergence/disappearance of biases. So there is some doubt about the validity of substituting a dozen words for a larger vocabulary in computer simulation.

emergence, they haven't clarified how the correlation and HOA's associative learning caused the developmental course of the three biases and why the difference of their simulation results arose. From these discussions, it seems that at present, reproductivity and illustrativity of bias findings is incomplete.

Summarizing the discussions above, we believe that LBA is a plausible explanation for bias emergence and that Smith et al. have made a large contribution to its investigation by proposing LBA and discussing the tangible mechanism of biases: HOA. However, there is no evidence that HOA are causing shape and material biases. Furthermore, HOA is not fully supported in terms of reproductivity and illustrativity (2.3a) and cognitive adequacy (2.3b). In the next section, we introduce another hypothesis which has higher reproductivity and illustrativity (2.3a) and cognitive adequacy (2.3b) for the biases emergence. We propose ad-hoc meaning substitution (AMS) as a simpler but more concrete hypothesis of shape and material biases. AMS must show and explain the behavior for solid stimuli (2.2.1) and nonsolid stimuli (2.2.2). To describe AMS's representation and computational process, we introduce "word distributional prototype (WDP)."

### 3. Proposed Hypothesis and its Rationale

#### 3.1. Word Representation in This Study

First, we explain input and category representation used in this study. Following Samuelson (2002), input representation consisted of three attributes: SOLIDITY, SYNTAX, and FEATURE. SOLIDITY, which denotes the solidness of objects/substances presented to learners, has three discrete attribute values: SOLID, NONSOLID, and AMBIGUOUS. AMBIGUOUS denotes borderline cases in which SOLIDITY is neutral or both SOLID and NONSOLID are possible. SYNTAX is the contextual attribute in a sentence given in parallel with the named object. It has three discrete attribute values: COUNT, MASS, and AMBIGUOUS. AMBIGUOUS within SYNTAX denotes borderline cases similar to SOLIDITY. For example, COUNT appears with indefinite article *a* or a determiner such as *many*, etc.

Such syntactic information was introduced because children receive it in a natural word learning situation. FEATURE is the attribute that denoted other perceptual information and consists of three attributes: SHAPE, MATERIAL, and OTHER. OTHER includes miscellaneous information other than SHAPE and MATERIAL: color, odor, fun, and so on.

We assumed children would experience pairs of a new noun and the above information, and would construct a category knowledge for the noun. Based on the above 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 miscellaneous information, respectively. These categories may overlap for each noun. For example, *crayon* is considered SOLID, a COUNT noun, and SHAPE-BASED, but it can also be simultaneously considered MATERIAL-BASED; *glue* is NONSOLID, a MASS noun, and MATERIAL-BASED.

#### 3.2. Multidimensional input and category representation

In this paper, we defined input representation based on Samuelson (2002). The input information that constituted a novel object was represented as a 36-dimensional vector in which the first three dimensions denoted SOLIDITY, the next three dimensions denoted SYNTAX, and the remaining 30 dimensions denoted FEATURE. In FEATURE, 10 dimensions were assigned to SHAPE, MATERIAL, and OTHER. Each dimension was represented as a real number of  $[-1, 1]$ . When the SOLIDITY of an input was SOLID, NONSOLID, or AMBIGUOUS, each was represented as  $(0.95, 0, 0)$ ,  $(0, 0, 0.95)$ , or  $(0, 0.95, 0)$ , respectively, and contained the noise between  $[-0.05, 0.05]$ . When the SYNTAX of an input was COUNT, MASS, or AMBIGUOUS, each was represented as  $(0.95, 0, 0)$ ,  $(0, 0, 0.95)$ , or  $(0, 0.95, 0)$ , respectively, and contained the same noise as SOLIDITY.

Each input belonging to the SHAPE-BASED category was assigned arbitrary but constant values to the dimensions of SHAPE attributes and random and changeable values to the dimensions

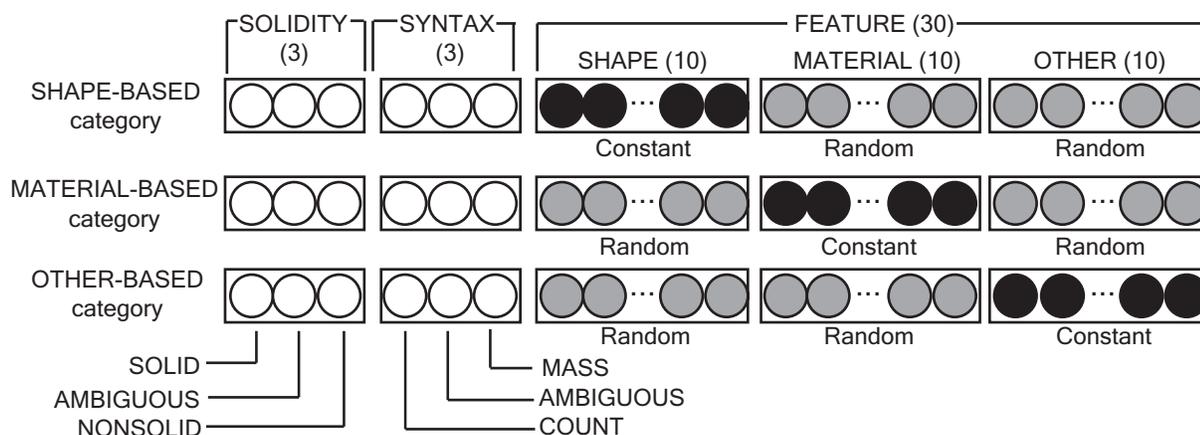


Figure 1. Representations of input and categories. Black circles denote constant values in each type of word category. Gray circles denote random values in the category. White circles denote that they aren't related to the category.

of other attributes. The same relation applied to MATERIAL- and OTHER-BASED categories; that is, MATERIAL and OTHER attributes were assigned arbitrary but constant values and random and changeable values to other attributes respectively. See Fig. 1 for more concrete understanding. To each of fixed feature input, noise between  $[-0.05, 0.05]$  was added.

### 3.3. Word Distributional Prototype (WDP)

We introduced explicit representation of word meaning, which enable us to describe a computational process of AMS concretely. There have been few studies using word meaning explicitly. But we need a concrete model of meaning representation to discuss on the computational mechanism. In psychological studies, word meanings have generally been considered in relation to categories and concepts and defined as prototypes, which were the most typical and central members among category members (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Some studies have sought vocabulary learning by associations between sensory feature vectors and discrete categories (see Roy,

2005). As pointed out by Roy, not all properties concerning word meaning can be covered by such associative learning models. But at least, they are consistent with prototype models in psychology. Therefore, in this paper, we propose a prototype-like model of word meaning derived from associative learning.

In our model, we assume that word meaning is defined as the prototype that consists of the distribution formed by input information co-occurring with the word (Kurosaki & Omori, 2005). Input information isn't limited to perceptual information but it also contains other multimodal and hidden information such as memory, emotion, time series information, and so on. In this paper, we assume that the input contains SOLIDITY, SYNTAX, and FEATURE information. Here it is notable to use input information distribution as word meaning. For example, *banana* is generally used for crescent-shaped yellow fruit, but not for red or globular things. It's so sensitive to shape and usage information that even slight discrepancy 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 characteris-

tics of word meaning can be expressed by two factors: the separated mean value in each information dimension and the allowable deviation from the mean value. Children learn them based on experience.

We used multidimensional normal distribution, one of the simplest possible representations for the meaning of a word, 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. However, non-diagonal elements of the variance matrices were fixed to zero for simplification, meaning that input dimensions had no correlation to each other. This simplification might have some problems, but we thought that at least for young children, WDP was sufficient for the prototype-like representation of word meaning.

In this model, each word meaning was described as below.  $\vec{x} \in \mathfrak{R}^M$ ,  $\vec{x} = (x_1 \ x_2 \ \dots \ x_M)^T$  is an input vector in  $M$ -dimensional input space,  $j$  is an ID number for a WDP <sub>$j$</sub> ,  $\vec{\mu}_j \in \mathfrak{R}^M$ ,  $\vec{\mu}_j = (\mu_{j1} \ \mu_{j2} \ \dots \ \mu_{jM})^T$  is a mean vector of WDP <sub>$j$</sub> ,  $\mu_{ji} \in \mathfrak{R}$  is a mean value of input unit  $i$  of WDP <sub>$j$</sub> ,  $\Sigma_j \in \mathfrak{R}^M \times \mathfrak{R}^M$  is a diagonal variance matrix of WDP <sub>$j$</sub> , and  $\sigma_{ji} \in \mathfrak{R}$  is a standard deviation of input unit  $i$  of WDP <sub>$j$</sub> . Then, likelihood  $p_j(\vec{x})$  of WDP <sub>$j$</sub>  is calculated as:

$$p_j(\vec{x}) = \frac{1}{(2\pi)^{\frac{M}{2}} |\Sigma_j|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu}_j)^T \Sigma_j^{-1} (\vec{x}-\vec{\mu}_j)\right) \quad (1)$$

WDPs were trained by the following algorithm:

- (3.3a)** Initially, each word is paired with an individual WDP.
- (3.3b)** Given input vector  $\vec{x}$  and a corresponding word, all WDPs calculate likelihood  $p_j(\vec{x})$  for the input, and the winning WDP <sub>$c$</sub>  that outputs the highest likelihood is chosen.
- (3.3c)** If WDP <sub>$c$</sub>  is the correct WDP paired with the given word, then it learns to increase its likelihood  $p_j(\vec{x})$  for the input, and the others don’t. Loss function  $\varepsilon_c(\vec{x})$  and update rules for each parameter are defined as:

$$\varepsilon_c(\vec{x}) = -\log(p_c(\vec{x})) \quad (2)$$

$$\Delta\mu_{ci} = -\alpha \frac{\partial \varepsilon_c(\vec{x})}{\partial \mu_{ci}} = -\alpha \frac{\mu_{ci} - x_i}{\sigma_{ci}^2} \quad (3)$$

$$\Delta\sigma_{ci} = -\beta \frac{\partial \varepsilon_c(\vec{x})}{\partial \sigma_{ci}} = -\beta \left( \frac{1}{\sigma_{ci}} - \frac{(\mu_{ci} - x_i)^2}{\sigma_{ci}^3} \right) \quad (4)$$

- (3.3c’)** If WDP <sub>$c$</sub>  is incorrect WDP for the given word, then it learns to decrease its likelihood  $p_j(\vec{x})$  for the input. Its update rules correspond to those of the opposite direction of (3.3c). The correct WDP, paired with a given word and which should give the highest likelihood, simultaneously learns by the update rule (3.3c).

- (3.3d)** Repeat (3.3b), (3.3c), and (3.3c’) depending on word input.

To maintain stable learning, we set the lower limit of  $\sigma_{ji}$  to 0.1 and the range of  $\mu_{ji}$  to  $[-1, 1]$ . The above algorithm is relevant to the extension of “learning vector quantization (LVQ)” (Kohonen, 1995). The initial value of every  $\sigma_{ji}$  is set to a sufficiently large value so that all WDPs don’t output higher likelihood to particular inputs in the initial state.

### 3.4. Ad-hoc Meaning Substitution (AMS) and Nearest Neighbor Hypothesis (NNH)

Consider a situation again in which children are given a pair of novel objects/substances and the corresponding noun. In this situation, nothing teaches them whether they should expand the noun based on similarity of SHAPE, MATERIAL, or OTHER. In other words from the perspective of WDP, just one experience of the set can reveal the mean vector of the word, but not either standard deviation. Nevertheless, as seen in word learning bias, children tend to generalize word meaning based on similarity of certain properties. As discussed above, such behavior is learned, not innate. Children must have the ability to complement the missing information of standard deviations by plausible methods nonspecific to vocabulary learning.

Therefore, we proposed the following hypothesis on a mechanism to realize novel noun generalization. When given a pair of input information and a novel noun, children determine the noun’s meaning based on the meaning of a particular known noun: The variance matrix of the most similar known noun WDP, which output the highest likelihood to the input information, is copied as the variance matrix of the novel noun. We call the copying process “nearest neighbor hypothesis (NNH)” (Fig. 2). Since we used multidimensional normal distribution as WDP, NNH denotes that the WDP nearest to the input vector from the perspective of Mahalanobis distance was chosen, and its variance matrix was used as that of the novel noun (Eqs. (5) and (6)). Meanwhile, the input vector presented with a novel noun was copied and pasted as the mean vector of the novel noun (Eq. (7))<sup>4</sup>.

$$c = \arg \max_j p_j(\vec{x}) \quad (5)$$

$$\Sigma_{\text{new}} = \Sigma_c \quad (6)$$

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

In a WDP framework, a single presentation of a novel word corresponds to one point in input space. The fact that children show word learning biases indicates a process in which they construct certain form of distribution (Eqs. (5) and (6)) and locate it on certain position in input space (Eq. (7)) based on a single datum. To complement the missing information of the variance matrix and the mean vector, one of probable strategies is to substitute the nearest known word and input vector for the information, even if it may not work always correctly. We call this ad-hoc construction of novel noun meaning ad-hoc meaning substitution (AMS) (Eqs. (5), (6), and (7)).

<sup>4</sup>Suppose that a child, who has known the word *dog* and its corresponding concept but hasn’t seen rabbits, meets a rabbit and its corresponding name *diff*. In the framework of our hypothesis, he accepts the instance of the rabbit as a prototype of *diff*; if he thinks the rabbit resembles dogs, he substitutes his knowledge of *dog* for his new knowledge of *diff*. (e.g., *diff* can be a noun organized by similarity of shape and texture, but not by similarity of color and size, etc.)

For the variance matrix, such competitive learning is often used in studies of neural networks and is thought to actually exist in the brain (Kohonen, 1995).

### 3.5. Predicted Process of Word Learning under AMS

Here, we explain in detail the process of how AMS causes fast emergence of shape bias for solid stimuli (2.2.1) and delayed emergence of material bias for nonsolid stimuli (2.2.2). We have argued that the triad of inductive learning, AMS, and biased early vocabulary is fundamental to bias emergence and that no bias behavior in the very first stage of word acquisition is caused by maturity, which will be explained in the next section. To understand how AMS works more clearly, we attend to the roles of maturity and two aspects of vocabulary:

- (3.5a) Vocabulary structure: statistical bias of word attributes among early-acquired vocabulary
- (3.5b) Vocabulary size (sparseness): density of WDPs in the SOLID and NONSOLID fields
- (3.5c) Maturity of word meanings and related cognitive faculties

For a simplified and essential discussion, here we consider SHAPE FEATURE and MATERIAL FEATURE as one-dimensional information and omit the other attributes (SYNTAX, AMBIGUOUS SOLIDITY, and OTHER FEATURE) from the input vector. Then, we have three remaining attributes: SOLIDITY, SHAPE, and MATERIAL. We discuss relationship between AMS and the three factors (3.5a-c) in the three-dimensional space. In that case, we can consider a SOLID and a NONSOLID fields, where SOLIDITY is fixed to SOLID or NONSOLID (Fig. 3). Each SOLID and NONSOLID stimulus used in novel noun generalization tasks is represented as one point in the SOLID and NONSOLID fields, respectively. WDP corresponding to a word is represented as an oval on either field (though it actually forms a hyper-oval-sphere in multidimensional input space). Assuming that the horizontal and vertical axes denote SHAPE and MATE-

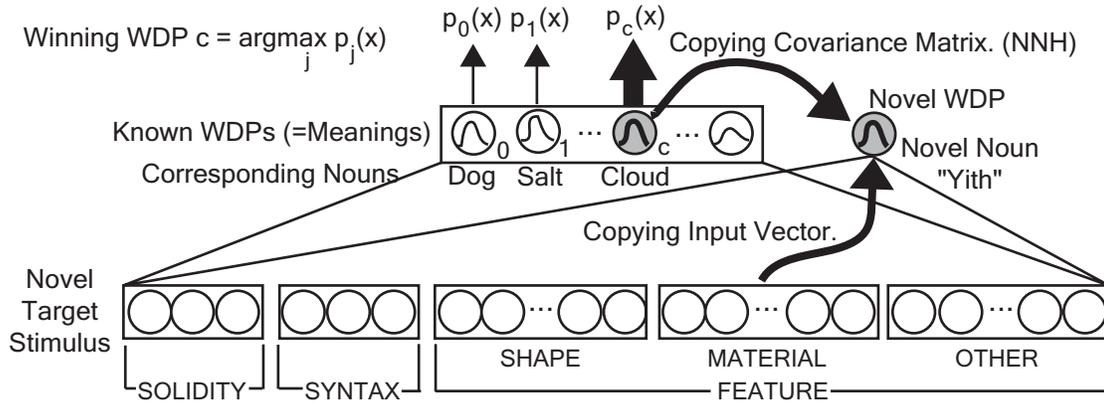


Figure 2. Ad-hoc meaning substitution (AMS). For a novel noun, new WDP copies its variance matrix from the nearest neighbor word which outputs the highest likelihood to the target stimulus (NNH) and copies its mean vector from the input vector.

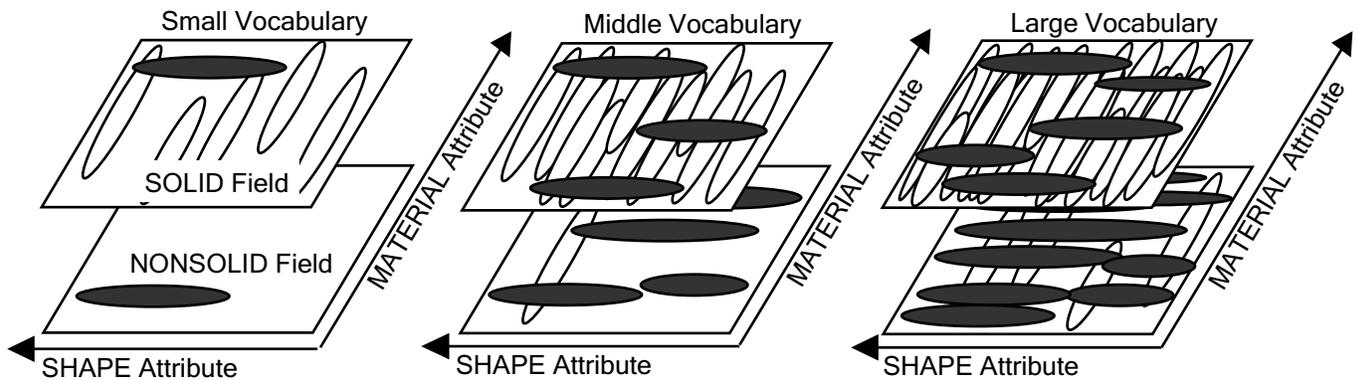


Figure 3. Predicted transition of vocabulary structure. WDP of SHAPE-BASED nouns is represented as white vertical ovals and MATERIAL-BASED nouns as filled horizontal ovals. Upper planes depict SOLID fields and lower planes do NONSOLID fields.

Table 1

Structural ratio in early vocabulary. Upper table shows ratio of each attribute value in SOLIDITY and FEATURE. Lower table shows conditional ratio of each attribute value in FEATURE when SOLIDITY is SOLID, NONSOLID, or AMBIGUOUS, respectively. (see Samuelson & Smith, 1999, for whole ratio)

	SOLIDITY			FEATURE		
	SOLID	NONSOLID	AMBIGUOUS	SHAPE	MATERIAL	OTHER
	.63	.04	.32	.48	.16	.39
SOLID				.61	.11	.31
NONSOLID				.00	.51	.39
AMBIGUOUS				.30	.15	.50

RIAL values respectively, we can represent the WDP of SHAPE-BASED as an oval that has a minor axis in the SHAPE direction and a major axis in the MATERIAL direction. Then the WDP of MATERIAL-BASED is represented as orthogonal to SHAPE-BASED.

Samuelson & Smith (1999) reported that the ratio of SOLIDITY, SYNTAX, and FEATURE attributes in an early vocabulary is not equal but quite biased. We call it the statistical bias (3.5a). We especially focus on the following three points in the bias: The number of SOLID nouns is much larger than NONSOLID nouns; the number of SHAPE-BASED nouns is much larger than MATERIAL-BASED nouns when SOLIDITY is SOLID; meanwhile, the number of MATERIAL-BASED nouns is much larger than SHAPE-BASED nouns when SOLIDITY is NONSOLID (Table 1).

We also focus on the transition of WDP density in the SOLID and NONSOLID fields along with vocabulary size (Fig. 3), which we call *sparseness* (3.5b). Becoming dense in a field, the WDPs in each field are more likely to be chosen as the nearest one to novel input in the field.

From the perspective of these three points (3.5a-c), we consider the transition of shape and material bias. First, consider cases with a small vocabulary, such as ten words. WDPs are sparse in both SOLID and NONSOLID fields. The NONSOLID field is especially sparse because an early vocabulary contains only 4% of NONSOLID nouns (Table 1). Therefore, when NNH is applied to the novel inputs in the SOLID field, the

WDPs in the SOLID field are almost chosen as the nearest for the input. Since SHAPE-BASED nouns are dominant in the SOLID field (Table 1), a significant shape choice can emerge (2.2.1b). Meanwhile, when NNH is applied to the novel inputs in the NONSOLID field, WDPs in the NONSOLID field are rarely chosen as the nearest; instead those in the SOLID or AMBIGUOUS fields are chosen because the NONSOLID field is too sparse to choose the nearest WDP from among the field. In that case, significant shape choice takes place as with the case of a novel SOLID input, and it leads to shape bias to NONSOLIDS, i.e. (2.2.2b) overgeneralized shape bias.

Second, consider large vocabulary cases. Applied to novel input in the SOLID field, NNH can make more significant shape choices (2.2.1b) because the density of WDPs in the SOLID field is more increasing than small vocabulary cases. Meanwhile, applied to novel input in the NONSOLID field, NNH can cause significant material choice (2.2.2c) because the sparseness of WDPs in the NONSOLID field has been cleared; i.e. NONSOLID field has become dense. Since WDPs in the NONSOLID field are chosen as the nearest and MATERIAL-BASED nouns are dominant in the NONSOLID field (Table 1), it leads to a significant material choice (2.2.2c).

Finally, consider middle vocabulary cases. Applied to novel input in the SOLID field, NNH can cause significant shape choices (2.2.1b) because it is middle cases of small and large vocabulary, which cause shape choice. Meanwhile, applied to novel input in the NONSOLID field, NNH may

cause mixed results depending on the degree of the sparseness. The complicated results for material bias in some studies may have arisen not from the difference of experimental conditions but from the subjects' vocabulary size - sparseness.

### 3.6. Predicted Process of Word Learning after introduction of Maturity

Though some of the phenomena for solid stimuli (2.2.1) and nonsolid stimuli (2.2.2) can be explained in the previous section, some points remained unexplained. We can't explain the mechanism of no shape bias during small vocabulary (2.2.1a) at all. The overgeneralized shape bias in section 3.5 is also inadequate because it was likely to emerge stably during small vocabulary, although it doesn't appear so frequently in the natural course of vocabulary development (2.2.2a). Nor we can explain no material bias during small vocabulary (2.2.2a). So we introduce the notion of *maturity* (3.5c) here.

We assume that children who have a small vocabulary have encountered fewer instances of each known word, and thus their learning of each word meaning hasn't converged. Along with vocabulary development, after they have experienced more instances of each known word, their learning of word meaning has converged. We call the degree of learning convergence for each word meaning maturity. In our WDP model, an immature situation corresponds to the state in which learning of mean vector and variance matrix for each WDP hasn't progressed beyond initial values. Each variance matrix element of WDP is assigned a large enough initial value (see section 3.3). Since WDP almost forms hyperspherical distribution in feature space, novel WDP corresponding to a novel word also copies its variance matrix and thus will output almost equal likelihood for every input. It leads to no shape bias (2.2.1a) and no material bias (2.2.2a) during a small vocabulary. Through resolving the immature state with further word learning, novel WDPs that copy the variance matrixes of SHAPE-BASED or MATERIAL-BASED nouns comes to show shape or material bias respectively, even immediately after producing a novel WDP (Fig. 4).

The entire process of word learning bias formation with maturity is summarized as follows. During a large vocabulary, learners have enough word learning experience. Therefore their biases to generalize novel nouns are same as the no maturity case (section 3.5): They will show significant shape bias to the solid set (2.2.1b) and significant material bias to the nonsolid set (2.2.2c). Meanwhile, during a small vocabulary, novel WDPs copy hyperspherical distribution because the learning of known WDPs is immature. It causes different results from the no maturity case: no shape bias (2.2.1a) and no material bias (2.2.2a). Finally, during a middle vocabulary, the resulting bias varies sensitively depending on the extent of maturity. Learners show shape bias to the solid set (2.2.1b) if their learning has somewhat converged. Meanwhile, they probably won't show material bias to the nonsolid set because its appearance may have been affected by maturity in addition to sparseness. In this case we predict that they will show the mixed results of no material bias (2.2.2a), or overgeneralized shape bias (2.2.2a).

The effect of AMS and maturity is summarized as follows. The biased vocabulary structure (3.5a) causes choice of known SOLID/NONSOLID nouns by NNH to SOLID/NONSOLID stimuli; since most of known SOLID/NONSOLID nouns are SHAPE-BASED/MATERIAL-BASED, it causes substitution of SHAPE-BASED/MATERIAL-BASED meaning to SOLID/NONSOLID stimuli; and it results in shape/material bias. This is the default process in children's novel noun generalization. However, sparseness (3.5b) and maturity (3.5c) affect the default process. Sparseness of vocabulary prevents NNH from choosing known NONSOLID nouns for novel NONSOLID stimuli and causes overgeneralized shape bias. Maturity has no effect on the choice, but it creates a situation where known nouns' WDPs are hyperspherical and uninformative for novel nouns and causes no bias behavior. The overgeneralized shape bias caused by vocabulary sparseness will not be salient because it is canceled by the maturity factor. In short, vocabulary structure creates shape and material biases, sparseness distorts it,

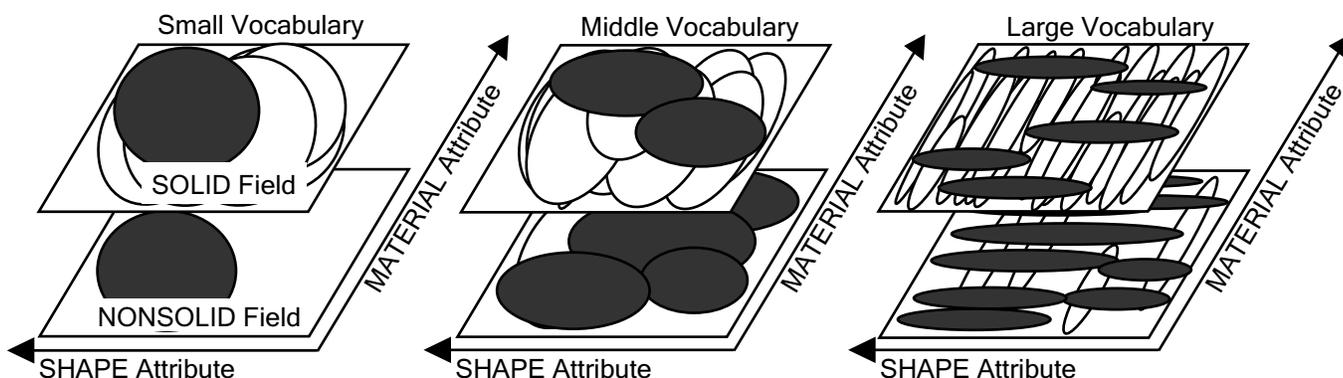


Figure 4. Predicted transition of vocabulary structure with maturity. WDPs of SHAPE-BASED and MATERIAL-BASED nouns are represented as white and filled ovals, respectively. Upper planes depict SOLID fields and lower planes do NONSOLID fields.

and maturity delays their effects. AMS consists of processing sensitive to these three effects, thus biases emerge.

In the following section, we tested whether our model based on WDP and AMS could show behavior consistent with children. In simulation 1, we didn't introduce maturity into the model and showed the behavior predicted in section 3.5. In simulation 2, we introduced maturity into the model and showed behavior predicted in section 3.6.

## 4. Simulation 1

### 4.1. Simulation course: inductive learning phase and generalizing phase

We divided the word learning process into two phases: "inductive learning phase" and "generalizing phase." In the inductive learning phase, learners received sets of words and corresponding input information and regulated WDPs (word distributional prototypes) using Eqs. (3) and (4), that is, without applying AMS (ad-hoc meaning substitution). Then in the generalizing phase, they realized fast mapping with AMS and the already learned knowledge of WDPs without updating the WDPs. Though children aren't likely to actually have such divided learning phases, this

division is useful to clarify the computational process included in word learning and to easily verify the model.

To compare the results of experiments by Samuelson & Smith (1999), we prepared six groups whose learners experienced a constant number of words in each group and verified their novel noun generalization tendencies after learning each group. Learners in each of the six groups, called groups 1, 2, 3, 4, 5, and 6, were given 18, 50, 102, 213, 281, and 312 words, respectively (see Samuelson & Smith, 1999). For each group, we prepared 21 learners who learned different words each other.

### 4.2. Inductive Learning Phase

#### 4.2.1. Used Vocabulary for Inductive Learning Phase

The vocabulary used in this phase must be appropriate words generally heard and produced by young children. Hence, we cited MCDI (Fenson et al., 1994), which include a typical vocabulary for 16- to 30-month-olds. In according with previous studies (Samuelson & Smith, 1999; Samuelson, 2002), we picked up 312 words from nine categories of the MCDI: "animals," "vehicles," "toys," "food and drink," "clothing," "body parts," "small household items," "furni-

Table 2

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	AMBIG- UOUS	COUNT	MASS	AMBIG- UOUS	SHAPE	MATE- RIAL	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
COUNT	.75	.00	.25				.66	.08	.34
MASS	.19	.19	.61				.03	.84	.16
AMBIGUOUS	.35	.14	.51				.20	.41	.45
SHAPE	.85	.01	.15	.85	.01	.15			
MATERIAL	.26	.18	.55	.26	.18	.55			
AMBIGUOUS	.55	.02	.43	.75	.05	.21			

ture and rooms,” and “outside things.”

As preparation for making learning data for computer simulations, we evaluated the 312 words from a perspective of the attributes in section 3.2. Some studies (Samuelson, 2002; Hidaka & Saiki, 2004b) asked adults to evaluate the words in terms of SOLIDITY, SYNTAX, and FEATURE attributes. But one important factor for this kind of study is the consistency of the structural ratio in the vocabulary compared to Samuelson & Smith’s one (1999) (Table 1). Therefore we evaluated the words ourselves from a perspective of the attributes in section 3.1 (‘glue’ is NONSOLID, MATERIAL-BASED, and MASS. ‘Bicycle’ is ...). The result is shown in Table 2. There may be some criticism of our evaluation procedure. However, in this study, we don’t aim to clarify vocabulary structure in early vocabulary, but attempt to reproduce bias findings and illustrate their mechanisms with AMS. Then what is important is to preserve the vocabulary structure (Samuelson & Smith, 1999) in the learning data for simulation. This method was also adopted by previous simulation studies (Samuelson, 2002; Colunga & Smith, 2005; Hidaka & Saiki, 2004b), which picked up small number of nouns from a whole early vocabulary as the structure was preserved.

To convert the evaluated noun data into a set of concrete learning data, we used a method described in section 3.2. Consider the case of ‘glue.’ Glue has some consistent properties of ‘glue’ such as nonsolidness, while there are diverse kinds of glue in the world. The consistency and diversity is represented in the inductive learning phase as follows. Its SOLIDITY dimensions are fixed in NONSOLID (0, 0, 0.95), its SYNTAX dimensions are fixed in MASS (0, 0, 0.95), and its MATERIAL dimensions among FEATURE are fixed in arbitrary 10-dimensional values. For each value in the 16-dimensional vector, noise between  $[-0.05, 0.05]$  is added. Each value in remaining 20 dimensions (SHAPE and OTHER) is set at random to generate the diversity.

The vocabulary for all learners was chosen randomly from the 312 words, and the number of words depended on the group to which learners belonged. Every learner in a group thereby learned a different set, but the same number of words. All learners experienced the inductive learning phase and formed their own WDPs with the update rules Eqs. (3) and (4).

#### 4.2.2. Results of Inductive Learning Phase

We set  $\alpha$  and  $\beta$  to 0.001, the initial values of  $\mu_{ji}$  to  $0.500 + [-0.001, 0.001]$  and  $\sigma_{ji}$  to 1.0 for the simulation. To represent a word’s distribu-

tion, we prepared 50 instances for the word, and the corresponding WDP learned the instances. In an *epoch*, learners experience all instances of all the words prepared for the learner. The learning iterated the epoch 30 times. We confirmed that the learning of all groups almost converged and that the learning parameters in each WDP were correctly estimated to form the distribution of the corresponding noun/category.

### 4.3. Generalizing Phase

#### 4.3.1. Experimental Design

After the inductive learning phase was finished, a novel noun generalization task was applied to the learners of each group. In this phase, we withheld SYNTAX information from learners because we intended to compare our results to Samuelson & Smith (1999) and Samuelson (2002), who conducted experiments with children and computer simulation without SYNTAX to examine tendencies to generalize novel nouns independent of SYNTAX information. Specifically, we calculated the likelihood of WDPs (Eq. (1)) without SYNTAX.

We used the stimuli of solid and nonsolid sets and separately conducted novel noun generalization tasks with them. Each set had 21 stimulus sets<sup>5</sup>, and a set consisted of one target stimulus, corresponding shape-match stimuli, and corresponding material-match stimuli. Twenty pairs of shape-match and material-match stimuli<sup>6</sup> were prepared for each target stimulus. Shape-match stimuli were made as input vectors that had the same SOLIDITY and SHAPE values as the target stimulus, random MATERIAL and OTHER values, and no SYNTAX value. The material-match stimuli were the same except that the roles

<sup>5</sup>Ideally, target stimuli should be absolutely unknown objects/substances for children/learners. In the case of an experiment with humans, experimenter can prepare such appropriately *novel* target stimuli based on her common sense. But it was difficult for computer simulation studies to prepare such appropriate target stimuli. Therefore, we resolved this problem by preparing a comparatively large number of target stimuli randomly. The large number reduces the effect of a few accidental uses of familiar target stimuli.

<sup>6</sup>This comparatively large number of shape- and material-match stimuli was also prepared for the same reason as the number of target stimuli.

of SHAPE and MATERIAL were switched with shape-match stimuli. Noise between  $[-0.05, 0.05]$  was added to the input dimensions that had the same attribute values as the target stimulus. Both stimulus sets were initially prepared and shared by all learners in all groups.

In our model, shape choice probabilities for each group were calculated below. Given the following pair, a novel target stimulus and a novel noun,  $Learner_{g_i}$  of  $Group_g$  made a novel WDP for the noun applying AMS. Then we gave  $Learner_{g_i}$  a pair of corresponding shape-match and material-match stimuli and compared which evoked higher likelihood. We considered that the stimulus raising higher likelihood would correspond to human choice in psychological studies. Learners made these choices for 21 pairs of shape-match and material-match stimuli. Shape choice probability  $p(g, i, j)$  to *Target stimulus<sub>j</sub>* by  $Learner_{g_i}$  in  $Group_g$  was calculated by Eq. (8):

$$p(g, i, j) = \frac{\text{winning number of shape-match stimuli}}{\text{total number of pairs}}. \quad (8)$$

Shape choice probability  $p_L(g, i)$  to all target stimuli by  $Learner_{g_i}$  in  $Group_g$  was calculated by Eq. (9):

$$p_L(g, i) = \frac{\sum_j p(g, i, j)}{\text{total number of target stimuli}}. \quad (9)$$

The mean probability of shape choice  $p(g)$  by all learners in  $Group_g$  is calculated by Eq. (10). We tested the significance of shape and material biases in  $Group_g$  by *t*-test on the basis of mean probability  $p(g)$  and standard deviations of  $p_L(g, i)$ .

$$p(g) = \frac{\sum_i p_L(g, i)}{\text{total number of learners}}. \quad (10)$$

#### 4.3.2. Generalizing Phase Results

First, we show the solid set results. *t*-test confirmed that the shape choice probability for each group was significantly larger than chance: They showed shape bias;  $t(20) = 12.95$ ,  $p < .001$ ;  $t(20) = 23.02$ ,  $p < .001$ ;  $t(20) = 26.77$ ,  $p < .001$ ;  $t(20) = 27.98$ ,  $p < .001$ ;  $t(20) = 27.89$ ,  $p < .001$ ; and  $t(20) = 32.45$ ,  $p < .001$ , respectively (Fig. 5).

To discuss the effects of sparseness in each

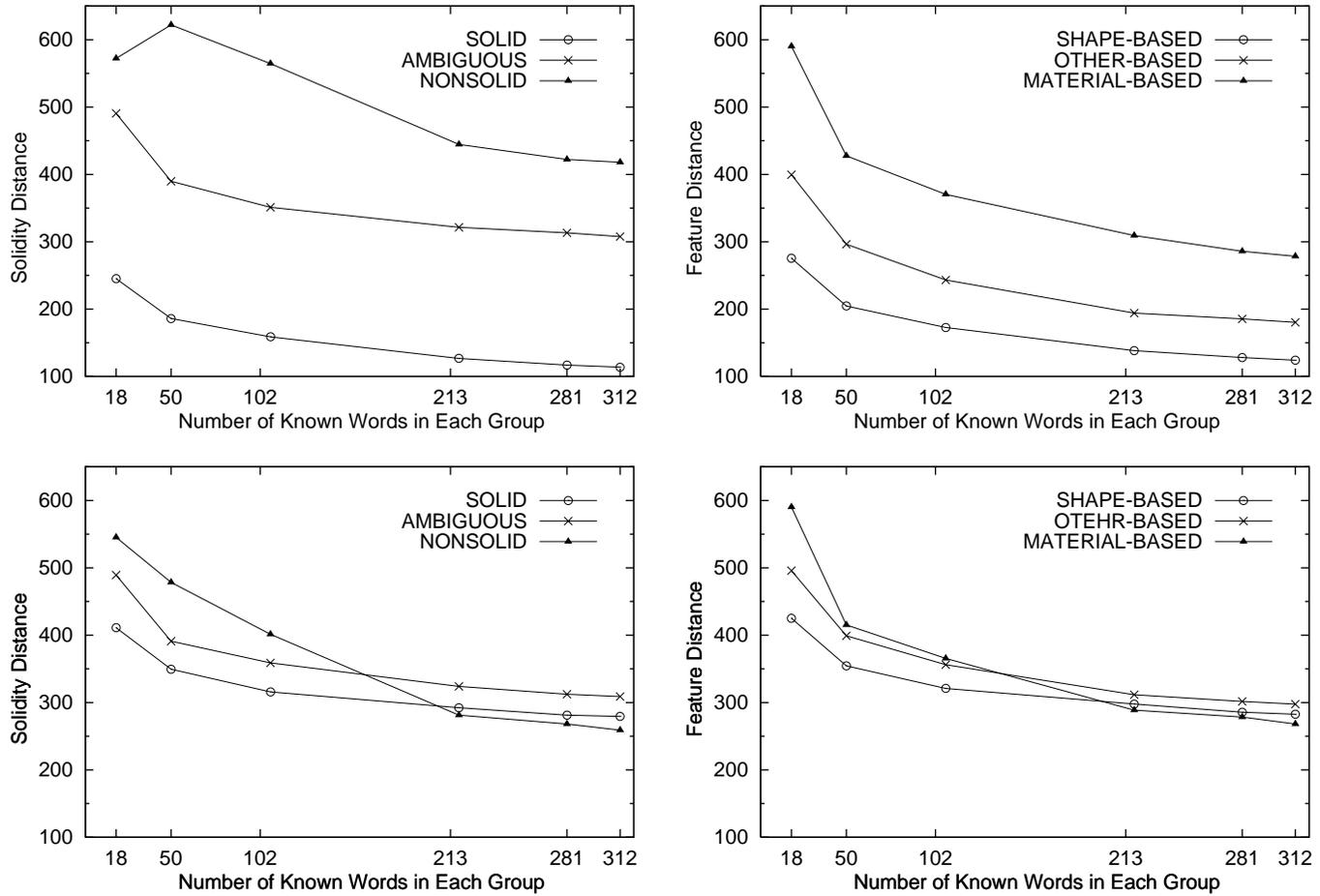


Figure 6. Solidity and feature distances in simulation 1. (Upper left: solidity distance in solid set; upper right: feature distance in solid set; lower left: solidity distance in nonsolid set; lower right: feature distance in nonsolid set.)

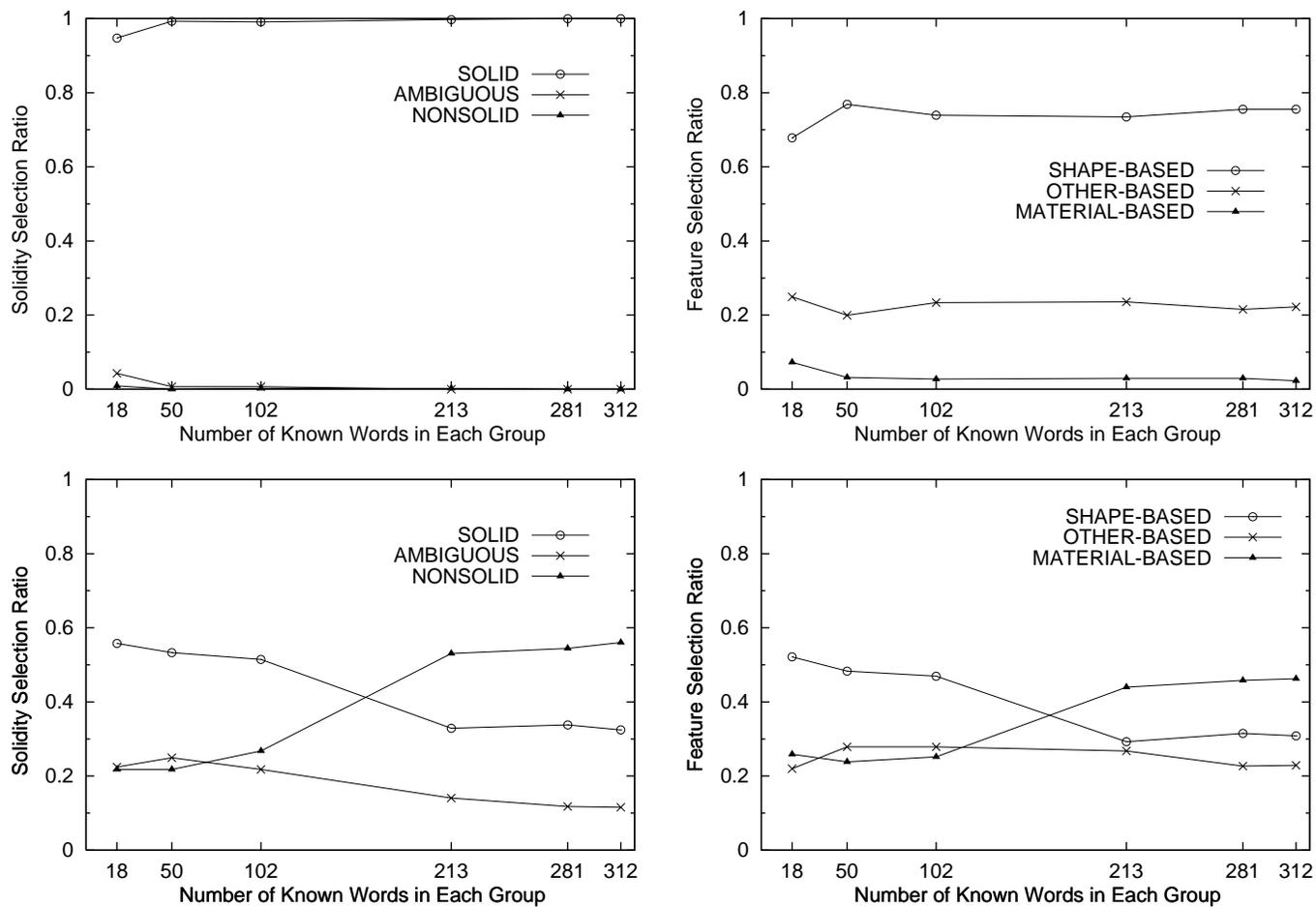


Figure 7. Solidity and feature selection ratios in simulation 1. (Upper left: solidity selection ratio in solid set; upper right: feature selection ratio in solid set; lower left: solidity selection ratio in nonsolid set; lower right: feature selection ratio in nonsolid set.)

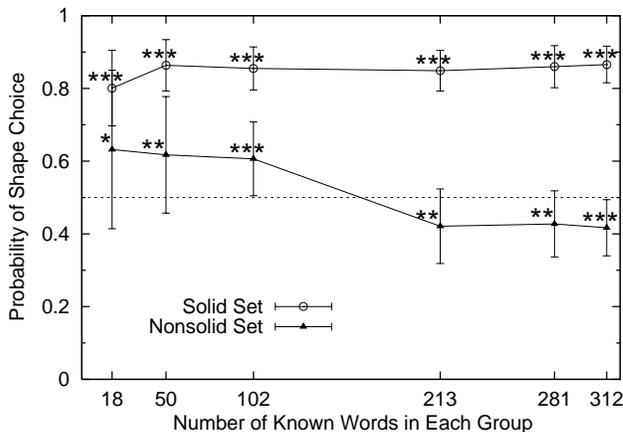


Figure 5. Probabilities of shape choices in generalizing phase of simulation 1. Vertical lines depict standard errors. Horizontal broken line depicts chance level ( $=.5$ ).  $*p < .05$ .  $**p < .01$ .  $***p < .001$ .

SOLID and NONSOLID field, we measured the Mahalanobis distance between the target stimulus and the nearest WDP among each WDP set respectively assigned SOLID, NONSOLID, or AMBIGUOUS values to SOLIDITY attributes. Each Mahalanobis distance was identical to the value calculated by Eq. (1) without the normalizing term. We averaged the Mahalanobis distances for 441 target stimuli in each group, which consisted of 21 target stimuli of 21 learners. We called this mean distance the “solidity distance” and used it as an index to represent sparseness in the SOLID and NONSOLID fields. We plotted solidity distances for each WDP group (Fig. 6, upper left).

To discuss the effects of choice in NNH (nearest neighbor hypothesis), we also measured the Mahalanobis distance between the target stimulus and the nearest WDPs among each WDP set that belonged to SHAPE, MATERIAL, or OTHER-BASED categories. We averaged the distances for all stimuli as with the solidity distance and called it the “feature distance” and plotted it for

each group (Fig. 6, upper right). The feature distance directly represented the effect of sparseness to category of chosen WDP by NNH.

Both solidity and feature distances almost monotonically decreased along with vocabulary development. The solidity distance to SOLID and the feature distance to SHAPE-BASED were consistently nearest to the novel inputs among SOLIDITY and FEATURE.

We measured which SOLIDITY value a WDP had when selected as the nearest one for a target stimulus. Then, we calculated the selection rates at which SOLID, NONSOLID, or AMBIGUOUS WDPs were selected as the nearest to the inputs in each group (Fig. 7, upper left). In the same way, we calculated the selection rates for the SHAPE-, MATERIAL-, or OTHER-BASED WDPs (Fig. 7, upper right). They were also investigated by using 441 target stimuli. We call them the “solidity selection ratio” and the “feature selection ratio,” respectively.

The solidity selection ratio of SOLID was dominant from group 1, the small vocabulary group, and this tendency was strengthened with an increase of vocabulary (Fig. 7, upper left). The feature selection ratio of SHAPE was also the highest from group 1, and this tendency was stable independent of vocabulary development (Fig. 7, upper right). We discuss the computational interpretation of these results in the next section.

Next, we show the nonsolid set results. Each bias was confirmed by  $t$ -test as below. 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;  $t(20) = 2.71$ ,  $p < .05$ ;  $t(20) = 3.27$ ,  $p < .01$ ; and  $t(20) = 4.68$ ,  $p < .001$ , respectively. But those for groups 4, 5, and 6 were significantly smaller than chance: They showed material bias;  $t(20) = -3.54$ ,  $p < .01$ ;  $t(20) = -3.57$ ,  $p < .01$ ; and  $t(20) = -4.82$ ,  $p < .001$ , respectively (Fig. 5).

As with the solid set, all solidity and feature distances monotonically decreased with vocabulary development (Fig. 6, lower row). The difference from the solid set results was that, though the solidity distance to SOLID and the feature distance to SHAPE-BASED in the earlier two

groups were closer to novel inputs than NON-SOLID and MATERIAL-BASED, relative orders gradually switched along with vocabulary development. This was caused by the drastic decreases of solidity distance to NONSOLID and the feature distance to MATERIAL-BASED.

The same mechanism also applied to the solidity and feature selection ratios (Fig. 7, lower row). Originally, the solidity selection ratio of SOLID and feature selection ratio of SHAPE-BASED categories were higher than NONSOLID and MATERIAL-BASED, respectively. But the relative orders gradually switched with vocabulary development (Fig. 7, lower row).

#### 4.4. Discussion

The solid set results (Fig. 5) show that our model reproduced stable shape bias during a somewhat larger vocabulary (2.2.1b). We discuss the adequacy of the prediction in section 3.5 here and the appearance of shape bias during a small vocabulary in the next section.

The monotonic decrease of solidity distance along vocabulary development (Fig. 6, upper left) suggested that both the SOLID and NONSOLID fields were becoming dense and that the SOLID field was consistently the densest from the small vocabulary group to the largest one. Correspondingly, the solidity selection ratio of SOLID was consistently the highest (Fig. 7, upper left) because the WDPs in the SOLID field were more likely to be nearest to the target stimuli in solid set by NNH, and the SOLID field was dense enough to support selection in the field. The results of feature distance to solid set (Fig. 6, upper right) can be explained by the above process and the structural ratio in Table 2. Since the WDPs in the SOLID field were dominated by SHAPE-BASED ones (Table 2), SHAPE-BASED WDPs were consistently nearest the solid set (Fig. 7, upper right). To sum up, we found that all these mechanisms caused shape bias (2.2.1b) and supported the explanation in section 3.5.

Meanwhile, the results of the nonsolid set (Fig. 5) showed that our model could reproduce the overgeneralized shape bias during a small vocabulary (2.2.2b) and material bias during a large vocabulary (2.2.2a). We discuss the adequacy of

our prediction in section 3.5 here and the consistent emergence of overgeneralized shape bias during a small vocabulary in the next section.

All solidity distances for nonsolid sets monotonically decreased (Fig. 6 lower left), identical to the solid set, which indirectly suggests the sparseness disappearance in all SOLIDITY fields. The NONSOLID field was drastically becoming dense compared to its initial sparseness, and thus the order of the solidity distance of SOLID and NONSOLID changed with vocabulary development. In turn, although one advantage was that NONSOLID WDPs were more likely to be the nearest ones for the inputs of the nonsolid set, learners couldn't exploit the advantage during a small vocabulary: They managed to choose the nearest WDPs from the SOLID field (Fig. 7, lower left, group 1-3). But increases of vocabulary and density in the NONSOLID field allowed them to take advantage: They came to choose the nearest WDPs from the NONSOLID field (Fig. 7, lower left, groups 4-6).

The results of feature distance (Fig. 6, lower right) were explained by the above reasons and the structural ratio in Table 2. During a small vocabulary, since SOLID WDPs were chosen as the nearest ones by NNH and most SOLID WDPs were simultaneously SHAPE-BASED (Table 2), SHAPE-BASED WDP was likely to become the nearest for the input of the nonsolid set (Fig. 7, lower right). They resulted in the overgeneralized shape bias (2.2.2b). But just the opposite situation occurred after a somewhat large vocabulary. The WDPs in the NONSOLID field were chosen by NNH, and most were MATERIAL-BASED (Table 2). Then the MATERIAL-BASED WDPs became the nearest (Fig. 7, lower right) and caused material bias (2.2.2c). To sum up, the results of the nonsolid set supported the explanation in section 3.5.

Finally, we conclude the results of simulation 1 as follows. Depending on just statistical bias in early vocabulary, AMS could explain not only the shape bias from the early period (2.2.1b) and the material bias after becoming a somewhat larger vocabulary (2.2.2c), but also the overgeneralized shape bias during small vocabulary (2.2.2b), which has not been reproduced in-

tegratedly by previous model studies. However, unexplained phenomena remain: no shape bias (2.2.1a), and no material bias during a small vocabulary (2.2.2a). The inevitable overgeneralized shape bias during a small vocabulary was also contrary to the previous findings, which haven't show such a robust overgeneralized shape bias (2.2.2b). To explain these points, we introduce the notion of *maturity* in the next section.

## 5. Simulation 2

### 5.1. Inductive Learning Phase

The simulation procedures and used vocabulary were identical to simulation 1. We introduced maturity to produce a situation in which the convergence of word meaning learning wasn't sufficient for learners who had smaller vocabularies compared to learners who had large vocabularies. Here, we realized an insufficiently learned condition by reducing the number of instances for each word, depending on the vocabulary size of the learner. (In simulation 1, the number of instances was fixed to 50 in each group.) In simulation 2, we determined the number of instances of each group as 50 multiplied by the maturity coefficient, which meant the degree of maturity in each group. Since there was no reliable evidence about the maturity coefficient, we tentatively used the ratio of the number of words in each group to 281 (the number of words for group 5) as the maturity coefficient for each group: respectively, 0.1, 0.2, 0.4, 0.8, 1.0, and 1.1 in ascending order of group number.

Fig. 8 shows the progress of learning in the inductive learning phase. The line for each group represents the transition in the mean value of the log likelihood of the nearest WDPs to all inputs, which included all instances of all words for 21 learners in each group. The graphs show that learning hadn't converged even at the end of epochs in groups 1 and 2. Their parameters had also not been estimated correctly, compared with simulation 1. We also confirmed that 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

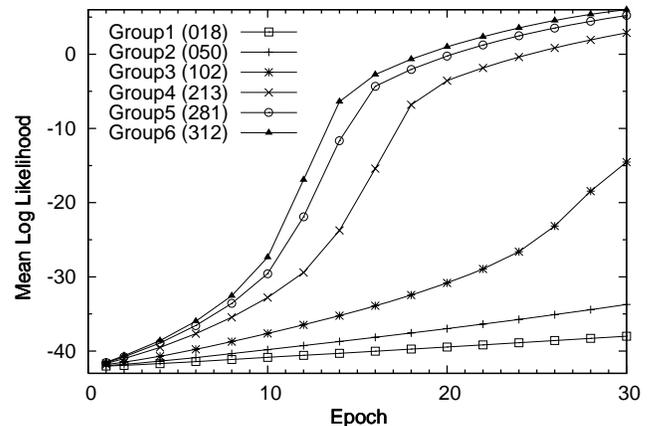


Figure 8. Results of inductive learning phase of simulation 2. Lines denote transition of mean log-likelihood of winning WDPs in each group along with learning epochs.

their parameters had been estimated well enough. From these results, the expected effects from the introduction of maturity were realized.

### 5.2. Generalizing Phase

The generalizing phase procedure was identical to simulation 1. First, we show the solid set results. Each bias was confirmed by  $t$ -test as below. The probability of shape choice for group 1 was significantly lower than chance even for the solid set, demonstrating material bias:  $t(20) = -20.39$ ,  $p < .001$ . But in group 2, there was no significant differences to chance, that is, it showed no bias:  $t(20) = 0.081$ ,  $p > .05$ . After that, the probabilities for groups 3, 4, 5, and 6 were significantly higher than chance, that is, they showed shape bias:  $t(20) = 15.01$ ,  $p < .001$ ;  $t(20) = 33.59$ ,  $p < .001$ ;  $t(20) = 26.63$ ,  $p < .001$ ; and  $t(20) = 37.49$ ,  $p < .001$ , respectively (Fig. 9).

All solidity and feature distances monotonically increased along with vocabulary development except in the first two groups (Fig. 10, upper row). The solidity distance to SOLID and the

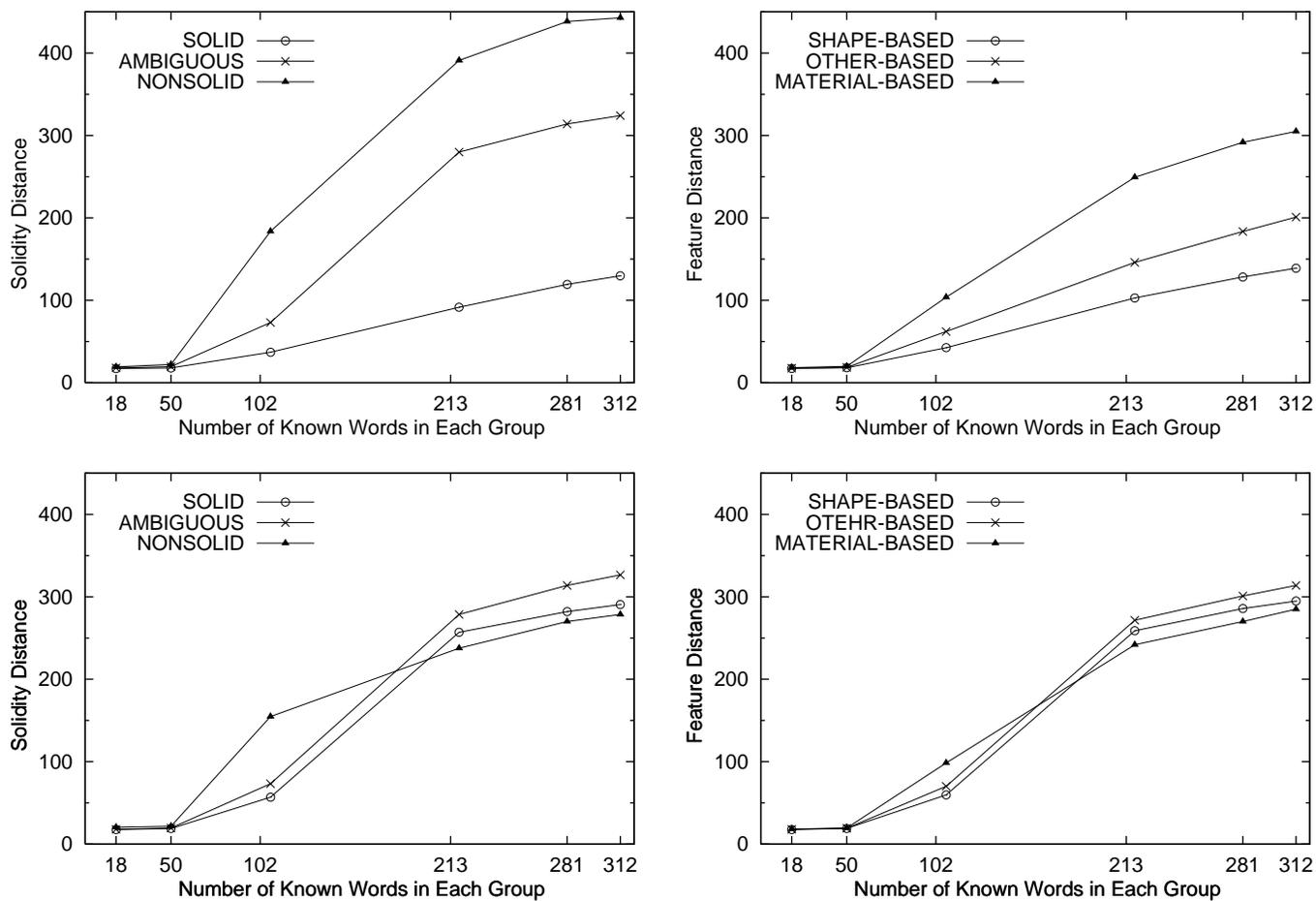


Figure 10. Solidity and feature distances in simulation 2. (Upper left: solidity distance in solid set; upper right: feature distance in solid set; lower left: solidity distance in nonsolid set; lower right: feature distance in nonsolid set.)

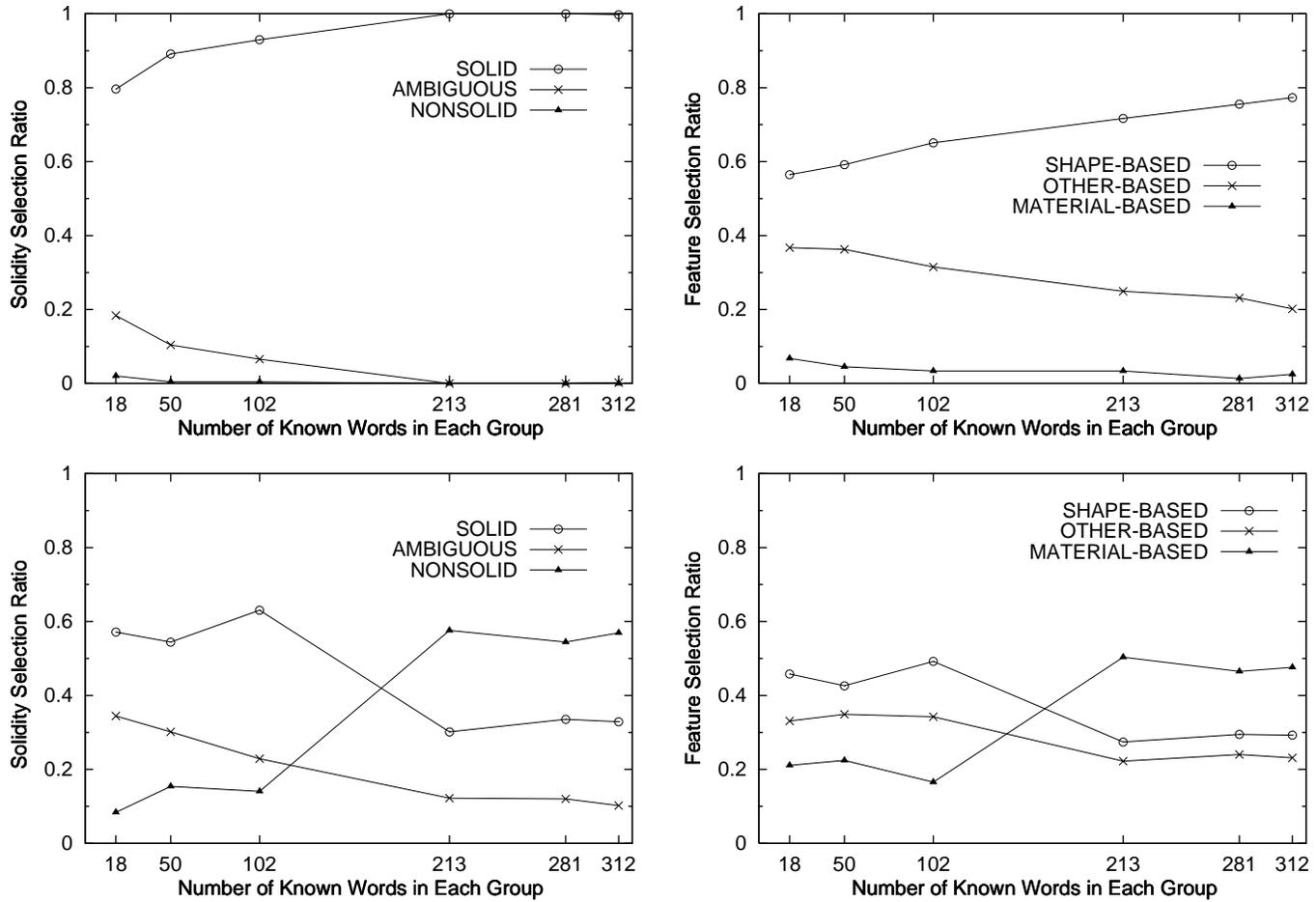


Figure 11. Solidity selection and feature selection ratios in simulation 2. (Upper left: solidity selection ratio in solid set; upper right: feature selection ratio in solid set; lower left: solidity selection ratio in nonsolid set; lower right: feature selection ratio in nonsolid set.)

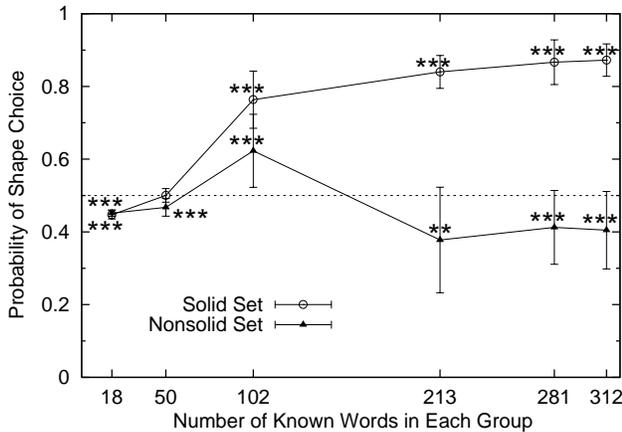


Figure 9. Probabilities of shape choices in generalizing phase of simulation 2. Vertical lines depict standard errors. Horizontal broken line depicts chance level (=0.5). \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

feature distance to SHAPE-BASED were consistently nearest to novel inputs. By introducing maturity, we got quite different results from simulation 1.

The solidity selection ratio of SOLID was dominant since the small vocabulary (group 1) and this tendency became stronger with vocabulary development (Fig. 11, upper left). The feature selection ratio of SHAPE-BASED also demonstrated the same tendency, but it was less dominant than the solidity selection rate (Fig. 11, upper right). At first both selection ratios were smaller than simulation 1, but in later groups they almost reached the same ratios as simulation 1.

Next, we show the nonsolid set results. Each bias was confirmed by  $t$ -test. The probabilities of shape choice for groups 1, 2, 4, 5, and 6 were significantly smaller than chance, showing material bias;  $t(20) = -24.54$ ,  $p < .001$ ;  $t(20) = -5.98$ ,  $p < .001$ ;  $t(20) = -3.77$ ,  $p < .01$ ;  $t(20) = -3.86$ ,  $p < .001$ ; and  $t(20) = -4.01$ ,  $p < .001$ , respectively. But in group 3, we observed a significantly larger

shape choice than chance, demonstrating shape bias even for the nonsolid set;  $t(20) = 5.48$ ,  $p < .001$  (Fig. 9).

All of the solidity and feature distances monotonically increased with vocabulary development except the first 2 groups, as the solid set (Fig. 10, lower row). After the earlier two groups, at first NONSOLID and MATERIAL-BASED distances were nearer to nonsolid inputs than SOLID and SHAPE-BASED, respectively. But the orders gradually switched with vocabulary development because of lower increasing rates for NONSOLID and MATERIAL-BASED (Fig. 10, lower row). When looking at the order among them, they were identical to simulation 1.

Concerning the solidity and feature selection ratios for the nonsolid set, at first SOLID and SHAPE-BASED were higher than NONSOLID and MATERIAL. But reflecting the solidity distance ratio and the feature distance rate, we observed the same reversing tendencies for the relative orders (Fig. 11, lower row). It also had the same order as simulation 1.

### 5.3. Discussion

First, we discuss the results of the solid set. The transition of shape choice (Fig. 9) showed no shape bias to the solid set during a small vocabulary in group 2 (2.2.1a) and stable shape bias to the solid set in subsequent groups (2.2.1b). The appearance of material bias in group 1 was inconsistent with the prediction in section 3.5, but it can be explained. In the very early group, the learning of WDPs was insufficient because of the small number of instances for each word, so it remained almost a hyperspherical WDP with wide variances. Then, calculating the likelihood of WDPs for novel noun input by Eq. (1), we didn't get large differences between the outputs of individual WDPs. This mechanism caused the problem of significant bias in group 1 (Fig. 9). Since shape choices by such learners became almost random,  $p_L(g, i)$  for all learners became around 0.5 and their deviation became very small. It resulted in statistically significant shape/material bias even though just slightly larger/smaller than chance. In other words, it was caused by setting maturity the same among learners within each

group and making the extent of learning convergence identically. But in the actual development of infants, it isn't likely that all children in the same vocabulary group simultaneously show the same developmental levels. So the problem can be resolved by modifying the current simplification of uniform maturity. These results fulfilled the prediction for solid stimuli (2.2.1).

Now, we discuss the consistency of the results with the computational process predicted in section 3.6. For the solidity and feature distances (Fig. 10), those distances in very early groups were smaller than later groups because the WDPs in the very early groups were closer to hyperspheres, and novel WDPs formed in the generalizing phase copied the variance matrixes of hyperspherical WDPs. Such WDPs specifically output comparatively high likelihood values to every input. However, the performance is the consequence of their incomprehension of each word meaning and subsequent low ability to distinguish different words. This situation corresponds to children's overextension of words ("This is *doggie*, that is *doggie* too, ..."). Meanwhile, well-learned WDP output high likelihood for inputs likely to co-occur with the word corresponding to the WDP and low likelihood for inputs that aren't likely to do so. The target stimuli presented in the novel noun generalization task were unknown to learners. They output comparatively higher likelihood if they had a lot of unlearned WDPs and very low likelihood contrary if they had a lot of well-learned WDPs. Hence, though later groups superficially seemed to show worse performances than early groups with regard to solidity/feature distance (Fig. 10), it was the consequence of their sophisticated knowledge clearly differentiating *doggie* from *cat* and so on.

As with simulation 1, the SOLID and NON-SOLID fields became dense with vocabulary development. The order among solidity and feature distances respectively decided the solidity and feature selection ratios, which subsequently affected the tendency of novel noun generalization. Hence the most important feature here was not the absolute value of those distances but the relative order among the solidity/feature distance in each group. In solid and feature distances,

SOLID and SHAPE were nearest to the inputs except in the first two groups (Fig. 10, upper row). Therefore, we found that all solidity and feature distances were equally small in the early groups because of their insufficiently learned hyperspherical WDPs; all distances became larger in the later groups because of their more learned and specialized WDPs for corresponding words, as predicted in section 3.6. Besides, the consistent orders in which SOLID and SHAPE were respectively the nearest were identical to simulation 1.

These discussions also affected the solidity and feature selection ratios (Fig. 11, upper row). As presented in Fig. 10, all the solidity and feature distances in the first two groups were almost the same to each other due to the influence of maturity. Hence, for the solidity and feature selection ratios, the rate with which each type of WDP was chosen as the nearest was almost random. In such cases, both solidity and feature selection ratios were determined depending on the structural ratio in the early vocabulary shown in the upper row of Table 2. Along with increasing vocabulary and converging learning for each word, the WDP with the same SOLIDITY as the target stimulus output higher likelihood values. Thus WDPs in the SOLID field were chosen more often as the nearest ones for novel SOLID target stimulus (Fig. 11, upper left). Following the choice, the feature selection ratio was automatically decided depending on the rate of the SOLID case in Table 2 (Fig. 11, upper right). These results illustrated the following process. WDPs in the SOLID field tend to be chosen as the nearest ones; they are likely to be SHAPE-BASED, resulting in shape bias; such a tendency increases with the development of vocabulary and maturity. Therefore, we confirmed that our model behaved as predicted in section 3.6 and explained the emergence mechanism of word learning biases for the solid set.

Next, we discuss the nonsolid set results. In groups 1 and 2, the existence of material bias was suggested from the results of shape choice probability (Fig. 9). They were inconsistent with the prediction in section 3.6, but they could be explained by the same reason as the material bias of group 1 in the solid set: Almost the same

likelihood output by insufficiently learned WDPs. Though our model couldn't so clearly show no material bias to the nonsolid set during a small vocabulary (2.2.1a), at least it inhibited the inevitable emergence of overgeneralized shape bias during a small vocabulary in simulation 1. And though the experimental evidence in subsequent groups was mixed (2.2.2a) (2.2.2c), there is no doubt that material bias finally settles in older children (Soja et al., 1991; Bloom, 2000), as was also demonstrated in our model. The results of our model for the nonsolid set were almost consistent with previous evidence.

Now we discuss whether our results arose from the same process as in section 3.6. As with the solid set, something important for NNH-based choice is the order among solidity and feature distances. We must consider the order in each group to explain the change of material bias with vocabulary development. In the early stage of groups 1 and 2, all solid and feature distances were almost equal (Fig. 10, lower row). Contrary, in the last groups 5 and 6, the solidity distance of NONSOLID became the nearest, but the NONSOLID field remained sparser compared to the SOLID field (Fig. 10, left). Hence the feature distance of MATERIAL in the nonsolid set wasn't nearer than others compared with SHAPE in the solid set (Fig. 10, right). To make material bias in the nonsolid set emerge as dominant as shape bias in the solid set, NONSOLID field must become much denser, that is, much larger NONSOLID vocabulary. In middle groups 3 and 4, the orders of SOLID and NONSOLID among solidity distance changed. But middle stage results should be unstable because they were composed of a subtle power relationship between maturity whose insufficiency led to no bias, the attracting force from WDPs in the SOLID field that produced shape bias, and from WDPs in the NONSOLID field that produced material bias. We confirmed by additional experiments that the results were unstable and changeable depending on such conditions as network parameters. Such changeable behavior might explain the mixed evidence of no material bias (2.2.2a) and stable material bias (2.2.2c) in previous behavioral studies.

This process also dealt with the solidity and

feature selection ratios (Fig. 11, lower row). In the first three groups, both selections were done depending on the structural ratio in the early vocabulary in the upper row in Table 2, as with the solid set. Then with vocabulary development, the WDPs in the NONSOLID field output higher likelihood and were chosen more as the nearest ones for the novel NONSOLID target stimulus (Fig. 11, lower left). But the solidity selection ratio of NONSOLID for the nonsolid set was much less than SOLID for the solid set, even in group 6, because the NONSOLID field resolved sparseness later than the SOLID field and thus temporarily remained affected by WDPs in the SOLID field. In that case, feature selection was done depending on the NONSOLID rate in Table 2 (Fig. 11, lower right), but was affected by SOLID. These results demonstrated the following trajectory. Affected by the WDPs in the SOLID field, WDPs in the NONSOLID field are becoming more likely to be chosen as the nearest ones; they are likely to be MATERIAL-BASED ones, which cause material bias; such a tendency increases with the development of vocabulary and maturity. Therefore, we also confirmed that our model could explain the emergence mechanism of word learning biases for the nonsolid set.

## 6. General Discussion

### 6.1. Our Position within LBA study

Here we discuss the relation between Smith et al.'s and our hypotheses. Both are LBA (learned bias account) studies and share the two fundamental factors of inductive learning and biased early vocabulary. Thus a difference exists only in the way of generalizing meanings of novel nouns, i.e. higher-order abstraction (HOA) (Smith et al., 2002) and ad-hoc meaning substitution (AMS). AMS assumes the following mechanism: children simply manipulate the nearest known word's meaning and multimodal input as ad-hoc but tolerable templates of novel noun meanings. Meanwhile, HOA can be interpreted as assuming the following mechanism: children abstract higher-order knowledge from already learned meanings and copy it as a template for novel noun meanings. In other words, HOA and AMS differ in

Table 3

Differences between higher-order abstraction (HOA) and ad-hoc meaning substitution (AMS). “O”/“X” marks denote desirable/undesirable characteristics. “\*” means they might be explained when HOA is extended well, but not currently.

Point of Comparison	HOA	AMS
Fundamental notion of process	Concept abstraction	Competitive learning and substitution
Template of novel noun meanings	Higher-order abstracted knowledge	Meaning of the nearest known noun
Shape bias	O: Explained	O: Explained
Material bias	O: Explained	O: Explained
Overgeneralized shape bias	O*: Almost reproduced	O: Explained
Cross-linguistic difference	O: Reproduced	*: Likely to reproduce
Cognitive adequacy	*: Adequate except for very early stage	O: Adequate from the period of bias emergence
Acquisition of abstract concept words	*: Likely to explain	X: Likely not to explain

how they conceptualize templates of novel word meanings.

As written in section 2.3, previous studies materialized Smith et al. (2002)’s 4-step model/HOA by a neural network, a kind of associative learning (Samuelson, 2002; Colunga & Smith, 2005). In general, neural networks come to predict adequate outputs for novel inputs after enough number of inputs and outputs have been learned (step 4 of the 4-step model). It is because neural networks not only learn pairs of inputs and outputs (step 2), but learn more general functional relationships between inputs and outputs (step 3). In other words, the essence of neural networks is to learn generalized/abstracted knowledge through data. Thus, it is natural for Samuelson (2002) and Colunga & Smith (2005) to adopt associative learning to materialize the 4 steps thoroughly. Meanwhile, our model doesn’t include the generalizing/abstracting function (step 3). AMS has the meaning substitution ability and the nearest neighbor hypothesis (NNH), which also seem appropriate for explanation of early biases because the competitive processing among known nouns and the winner-takes-all learning in AMS are natural in early brain processing. In addition, AMS can be plausible from a perspective of analogous to “personification as analogy” (Inagaki & Hatano, 1987; Inagaki & Sugiyama, 1988), which argues that children reason about animals and

plants by analogy with familiar humans. It’s similar to the AMS process that children substitute knowledge of unfamiliar things based on familiar things. Consequently, both HOA and AMS are equivalently natural in processing.

Next, we compare them in terms of more detailed computational mechanisms. Their most remarkable difference is existence/nonexistence of the third step of Smith et al. (2002). It leads to difference in their hierarchical network structures: HOA has to include a hidden layer that encodes the relationship between learned word meanings, whereas AMS does not. Each of the structures was adopted not by chance, but as a logical consequence of use/disuse of higher-order knowledge. The difference in existence/nonexistence of a hidden layer subsequently results in difference between their calculation costs; a network without a hidden layer has a smaller calculation cost obviously. Therefore, it is possible that AMS’s lower-level processing and smaller cost can resolve HOA’s difficulty in cognitive adequacy (2.3b), especially for early biases around 24 months of age.

As for reproductivity and illustrativity considerations (2.3a), HOA studies currently have not fully reproduced shape, material, or overgeneralized shape biases, and have also lacked discussion about the small vocabulary size in their simulation. Meanwhile, AMS not only reproduced the bias emergence, but also demonstrated the pro-

cess of emergence and disappearance of the biases along with vocabulary development. Additionally, AMS could illustrate how they emerge, why emergence of material bias is later than shape bias, and why overgeneralized shape bias is a hard-to-observe phenomenon from the viewpoints of vocabulary structure (3.5a), vocabulary size (3.5b), and maturity (3.5c) (Fig. 3-4). However, it is true that HOA has superiority in reproductivity of cross-linguistic differences over AMS. Colunga & Smith (2005) conducted a computer simulation with English and Japanese vocabulary, and reproduced the same bias behavior found with English and Japanese children (Imai & Gentner, 1997). We think that any hypotheses within LBA are likely to reproduce a kind of cross-linguistic difference because in LBA framework, bias emergence is affected by the differences of vocabulary structure between English and Japanese. But, since other LBA hypotheses don't necessarily show the same cross-linguistic difference as children, its confirmation with English and Japanese vocabulary should be done.

In sum, we argue at least currently that AMS has an advantage in reproductivity and illustrativity (2.3a) over HOA except for some open questions, and that cognitive adequacy (2.3b) of AMS also seems more consistent than HOA. We think that especially based on cognitive adequacy findings, AMS plays a key role in the first stage of the biases (around 24 months of age). However, these findings don't necessarily prove AMS because other hypotheses which meet the conditions (2.3a) (2.3b) may exist. To prove those hypotheses, evaluations of them based on their prediction results (2.3c) are necessary. Meanwhile, we shouldn't exclude HOA because it can apparently explain well the acquisition of words belonging to superordinate categories or more abstract concepts. We think that they work together for those biases. We summarize the differences between HOA and AMS in Table 3.

## 6.2. Some predictions and comments about HOA and AMS

AMS assumes that a child forms a novel noun meaning based on a known noun meaning at single blow and the learned meaning for a novel noun depends on a known noun chosen as the nearest noun. Thus, in a state where a specific known noun is evoked in the subject's mind, the meaning of the noun will have a strong influence on the result of novel noun generalization. For example, we may be able to observe children's "overgeneralized material bias" for novel SOLID stimuli when they are evoking known NONSOLID nouns. Meanwhile, even in that situation, children won't show overgeneralized material bias under the HOA framework, which uses generalized knowledge as a template of novel noun meaning.

Similarly, under AMS framework, overgeneralized shape bias won't be observed in children whose vocabulary is over a certain size. During a small vocabulary, they choose the nearest known word for novel NONSOLID nouns from the SOLID field but not from the sparse NONSOLID field. But such odd choices disappear when sparseness in the NONSOLID field is resolved by an increase of vocabulary. Therefore, it's difficult to observe overgeneralized shape bias unless young children with small vocabularies are forced to selectively learn SOLID and SHAPE-BASED nouns.

Next, the strength of shape and material biases a child shows correlates with the level of understanding of each word, that is, maturity. If a child doesn't show biases in spite of her adequate vocabulary size and age, we speculate that her maturity is low. Maturity in this paper means a complex measure of the degree of child's comprehension of each word meaning, cerebral development, and other factors. Though it would be difficult to assess maturity, its assessment is a promising direction for future research to prove our hypothesis.

Finally, we should discuss effect of the sparseness on both SOLID and NONSOLID fields in addition to the structural ratio in the vocabulary. Structural ratio alone can't explain why material bias appears after overgeneralized shape bias, even though SHAPE-BASED nouns still remain

dominant in the vocabulary. But we can explain emergence of material bias in the dominance of SHAPE-BASED nouns based on growth of the density in NONSOLID field. After all, we might predict what kind of bias children would show or how strongly they might show it if we know their vocabulary triad: sparseness, maturity, and the structural ratio.

### 6.3. Relationship to Counterarguments

Although we support LBA, some researchers made objections to it. Booth & Waxman (2002, 2005) revealed that children behaved differently depending on whether shown a target object in an *animate* or an *artifact* context. They suggested that children generalize novel nouns based not only on similarity in SHAPE but on other conceptual information such as ANIMACY. However this isn't a crucial objection because the generalization tendency in both HOA and AMS depends on the templates of novel nouns, and the template can also be affected by ANIMACY as they were affected by SOLIDITY in our experiment.

Diesendruck & Bloom (2003) revealed that children showed shape bias when asked to choose the same kind of object as a target stimulus. They argued that shape bias is derived from guessing what kind of thing the target stimulus is, rather than from LBA's simple attention learning. However, LBA is no longer inconsistent with their view because HOA and AMS consist of knowledge-based guesses about what *kind* of meaning the noun has, which result in attentional learning or word learning biases. Critics seem to misunderstand Smith and her colleague's current stance. Their findings rather offer supporting evidence for LBA.

Bloom (2000) suggested that it was difficult for LBA alone to illustrate shape bias and that "theory of mind (ToM)" was indispensable because it is informative to know a word's referent. But as Bloom himself remarked, ToM alone is also insufficient to determine a word's meaning (Bloom, 2000, p. 87). Besides, ToM accounts also have an age problem because it is difficult to expect ToM to reach maturity at 24 months when shape bias emerges. However, it is quite conceivable that af-

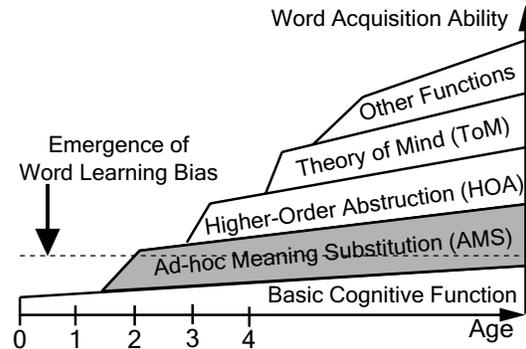


Figure 12. Probable positioning of various functions in development of word acquisition ability.

ter ToM becomes available, it first narrows the possible meanings of a word and then HOA and AMS work together.

In summary, we could find neither crucial evidence against LBA nor a more appropriate alternative to it. But it is quite possible that HOA and AMS cooperatively and competitively engage in bias emergence with ToM and other functions. We consider AMS to be the simplest function within them, and thus it takes an important role as the basis of bias emergence from an early stage (Fig. 12).

## 7. 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 simulation. AMS is a hypothesis among learned bias account (LBA) and argues that children simply use the meaning of known words and given instances as ad-hoc templates of novel noun meanings. It is a simpler hypothesis than previous studies. However, it could replicate phenomena concerning shape and material biases and explain the process clearly. Besides, to describe AMS's processing, we introduced word distributional prototype (WDP) as the explicit represen-

tation of word meaning with an inductive learning function. Few studies have explicitly represented word meaning. But we consider such representation is valid methodology for illustrating the computational mechanism of word learning bias, which itself is deeply committed to word meaning.

Simulation 1 revealed that when learners who possess WDP and AMS were exposed to early biased vocabulary, they showed shape and material biases as well as overgeneralized shape bias. This result suggested that the triad of word meaning induction, ad-hoc instantaneous estimation of meaning, and early biased vocabulary is essential for the emergence of biases. Simulation 2 revealed that when maturity was introduced, learners showed neither shape nor material bias during an early small vocabulary. This result indicated that the period of bias emergence is decided not by the triad but by maturity.

Though we have verified AMS from the perspective of reproductivity, illustrativity, and cognitive adequacy, there is no conclusive evidence that AMS causes shape and material biases. To prove AMS, we have to investigate prediction and its confirmation (2.3c). Additionally, it's also unlikely that only AMS realizes novel noun generalization throughout life because there are some functions which HOA and other models seem to be able to explain. We consider that AMS works as one mechanism of producing the biases and plays a central role, especially in their emerging stage. Future work includes investigations into how AMS cooperates with other cognitive functions (Fig. 12) and how AMS joins other word learning biases. Finally, computational modeling is one promising method for the illustration of a more concrete hypothesis of human internal processing such as word learning. We expect more advanced studies with this method and hope to integrate abundant findings from various fields including psychology, physiology, and others.

## References

- Bloom, P. (2000). *How children learn the meanings of words*. Cambridge: MIT Press.
- Booth, A. E., & Waxman, S. R. (2002). Word learning is 'smart': Evidence that conceptual information affects preschoolers' extension of novel words. *Cognition*, *84*(1), B11–B22.
- Booth, A. E., Waxman, S. R., & Huang, Y. T. (2005). Conceptual information permeates word learning in infancy. *Developmental Psychology*, *41*(3), 491–505.
- Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17–29.
- Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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.
- Diesendruck, G., & Bloom, P. (2003). How specific is the shape bias? *Child Development*, *74*(1), 168–178.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Mono-graphs of the Society for Research in Child Development*, *5*, 242.
- Hall, D. G. (1996). Naming solids and nonsolids: children's default construals. *Cognitive Development*, *11*, 229–264.
- Hidaka, S., & Saiki, J. (2004a). A mechanism of ontological boundary shifting. *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*, Chicago, IL, 565–570.
- Hidaka, S., & Saiki, J. (2004b). A connectionist account of ontological boundary shifting. *Proceedings of the 11th International Conference*

- on *Neural Information Processing*, Calcutta, 282–287.
- Hidaka, S., & Saiki, J. (2005). A model study of infants' novel word categorization [Japanese]. *Cognitive Studies*, 12(3), 235–251.
- 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.
- Inagaki, K., & Sugiyama, K. (1988). Attribution human characteristics: Developmental changes in over-and-under-attribution. *Cognitive Development*, 3, 55–70.
- Kohonen, T. (1995). *Self-organizing maps*. New York: Springer-Verlag.
- Kuroski, K., & Omori, T. (2005). 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, Canada: ACTA Press.
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, 3(3), 299–321.
- Mandler, J. M., Bauer, P. J., & McDonough, L. (1991). Separating sheep from the goats: differentiating global categories. *Cognitive Psychology*, 23, 263–298.
- Mandler, J. M., & McDonough, L. (1993). Concept formation in infancy. *Cognitive Development*, 8, 291–318.
- Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge: MIT Press.
- Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, 16(1), 1–27.
- Markman, E.M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20(2), 121–157.
- Quine, W. V. (1960). *Word and object*. Cambridge: MIT Press.
- Rosch, E., Mervis, C. Y., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive psychology*, 8, 382–439.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Roy, D. (2005). Grounding words in perception and action: Computational insights. *Trends in Cognitive Sciences*, 9(8), 389–396.
- 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.
- Samuelson, L. K., & Smith, L. B. (2000). Children's attention to rigid and deformable shape in naming and non-naming tasks. *Child Development*, 71 (6), 1555–1570.
- Samuelson, L. K., Horst, J. S., Dobbertin, B. N., & Schutte, A. R. (2006). Knowledge, performance, and task: décalage and dynamics in young children's noun generalizations. *Proceedings of the 28th Annual Meeting of the Cognitive Science Society*, Vancouver, BC, Canada, 720–725.
- 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. Basic and applied perspectives on learning, cognition, and development* (pp. 1–32). 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.

Subrahmanyam, K., Landau, B., & Gelman, R. (1999). Shape, material, and syntax: Interacting forces in children's learning in novel words for objects and substances. *Language and cognitive processes, 14*(3), 249–281.