

KASUMI YAMAMOTO AND FRANK KEIL

THE ACQUISITION OF JAPANESE NUMERAL
CLASSIFIERS: LINKAGE BETWEEN GRAMMATICAL
FORMS AND CONCEPTUAL CATEGORIES*

This study examines the acquisition of Japanese numeral classifiers in Japanese preschool children, ages 3 to 6, with a primary emphasis on developing comprehension ability. Numeral classifiers, widely distributed in languages of East and Southeast Asia and the New World, are a group of morphemes that usually occur adjacent to quantity expressions. The selection of numeral classifiers is determined by the inherent semantic properties of the noun whose quantity is being specified, suggesting that developing patterns of comprehension should be linked to underlying patterns of semantic and conceptual development. Previous research claims that children acquire certain distributional patterns very early but that the acquisition of the semantic system is a very slow process. We argue instead that different techniques and stimulus contrast sets reveal a much greater sensitivity to semantic relations in young children than was previously considered possible. Reasons for the apparent slowness in classifier acquisition are also discussed.

1. INTRODUCTION

Many languages have grammatical systems that require a linkage between abstract, generative conceptual categories and grammatical forms. These abstract categories can have an indefinitely large number of subcategories; thus the category of flat two-dimensional objects includes not only familiar subcategories like sheets of paper and leaves, but also a potentially indefinite number of novel categories that conform to the abstract properties of the supercategory. Relations between such categories and grammatical forms can often be elaborate and multi-layered in the numeral classifier systems of some languages. In the studies reported here, we ask how developing patterns of language comprehension and production for numeral classifiers in Japanese might help us understand links to patterns of semantic and conceptual development.

Numeral classifiers are a grammatical system that reflect how speakers categorize objects that they count or quantify. Classifiers usually occur adjacent to quantity expressions including numbers. The selection of numeral classifiers is determined by the inherent semantic features of the noun being quantified or counted. In Japanese, the structures¹ in (1a–d) are typical numeral classifier constructions, with the classifier in the examples, *-mai*, referring to flat and two-dimensional objects such as paper, cloth, leaves, and so forth.



- (1) a. Kami-o **ni-mai** katta.
 paper-Acc **two-Cl** bought
 ‘(I) bought two sheets of paper.’
- b. Kami **ni-mai-o** katta.
 paper **two-Cl-Acc** bought
 ‘(I) bought two sheets of paper.’
- c. **Ni-mai-no** kami-o katta.
two-Cl-Gen paper-Acc bought
 ‘(I) bought two sheets of paper.’
- d. **Ni-mai** kami-o katta.
two-Cl paper-Acc bought
 ‘(I) bought two sheets of paper.’

Although there are approximately 150 Japanese numeral classifiers, only about 30 are found in frequent daily use (Downing, 1984, 1996), in contrast to other East and Southeast Asian languages, whose numeral classifier inventory is more extensive. Adams and Conklin (1973) find that 37 East and Southeast Asian languages have numeral classifier systems, and argue that animacy, shape, and function are critical semantic features in most of them. Subsequent research has found that languages with classifier systems show similar patterns (Croft, 1994). There has been, however, far less work that systematically investigates which aspects of shape, animacy, and function play a role in categorizing classifiers across languages. Thus, we do not yet know if there are universal constraints on the sorts of categories so employed. Studies on classifiers in child language are therefore of special interest, as the earliest forms across languages might converge to yield common patterns, while developmental patterns may reveal constraints on the acquisition of categories represented by numeral classifiers.

In Japanese, classifiers are divided into two major categories, animate and inanimate (see Figure 1), with animate classifiers being further divided into two subcategories, human and animal. Inanimate classifiers are also divided into two categories, with classifiers for concrete objects and abstract objects. Finally, concrete object classifiers are divided into two categories consisting of shape-specific classifiers and functional (non-shape specific) classifiers.² With shape-specific classifiers the semantic range is quite broad, referring not only to inanimate concrete objects, but also to plants and natural substances when they have a solid shape like an icicle. Actually these major categories do not necessarily have specific classifiers; instead,

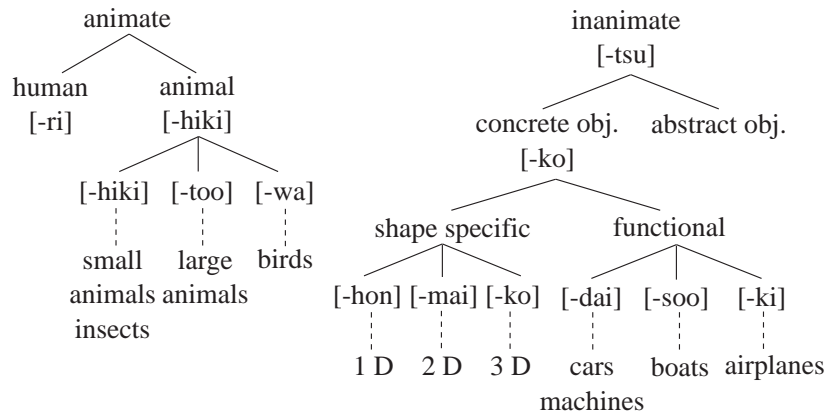


Figure 1. Japanese numeral classifier system (classifiers tested in the comprehension experiments are indicated in the brackets).

they are determined based on the semantic features that group classifiers into maximally general categories.

1.1. *Previous Research*

Since classifiers are a system that classifies nouns on the basis of their referents' attributes, patterns of acquisition of classifiers have often been discussed as providing a potentially important source of information about underlying patterns of semantic and conceptual development (Adams and Conklin, 1973; Clark, 1977; Muraishi, 1983; Craig, 1986; Matsumoto, 1987; Carpenter, 1987, 1991). Adams and Conklin (1973) stress the possibility that classifier systems reflect basic cognitive categories despite their cultural or language-family specificity. On an alternative view, since the classifier system reflects the categorization of objects, the acquisition of numeral classifiers might influence conceptual development itself in a Whorfian manner, rather than merely being a product of such patterns of development (Muraishi, 1983). Yet, despite the strong interest in the numeral classifier system and its relation to cognitive development, relatively few empirical studies have been carried out to investigate the acquisition process of numeral classifier. Among East and Southeast Asian languages, only four have been studied in terms of the acquisition of numeral classifiers: Japanese (Sanches, 1977; Muraishi, 1983; Matsumoto, 1985a, 1985b, 1987; Uchida and Imai, 1996); Mandarin (Erbaugh, 1986; Fang, 1985; Hu, 1993); Thai (Gandour et al., 1984; Carpenter, 1991) and Korean (Lee, 1994).

Each of the four Japanese studies has found essentially similar patterns characterizing the process of numeral classifier acquisition. Sanches (1977) focused on differences in the numeral classifier systems from different generations of speakers, the relation of these differences to the acquisition process, and their relation to a context of social meanings. Sanches administered elicitation tests in a written format to adults ranging in ages from 18 to 73 and children from ages 9 to 12, using a verbal version of the test and other tasks for children from ages 2 to 9. Her results showed the mean number of classifiers in the repertoires of people over age 30 to be 36 and those under age 30 was 28. Central to the findings of her experimental evidence was that particular classifier domains such as 'living things' were structurally maintained quite well across generations, while some domains such as 'furniture and implements' were undergoing a structural change.

As other child acquisition research confirm (Muraishi, 1983; Matsumoto, 1985a, 1985b, 1987; Uchida and Imai, 1996; Gandour et al., 1984; Carpenter, 1991; Lee, 1994), children prefer to use unmarked, more general classifiers rather than specific classifiers. Sanches suggested that to acquire numeral classifiers, children first master the number concept along with a set of lexemes representing concepts to which they could be applied. Developmentally then, the acquisition of numeral classifiers would start at around 2.5 years, after children acquire cardinal number forms.³ By age 5 or 6, children learn the basic six forms of classifiers, and these six forms correspond to the structure of the forms surviving in adult models. These forms were *-hon* for one-dimensional objects, *-mai* for two-dimensional objects, *-ko* for three-dimensional objects, *-nin* for human, *-hiki* for animals in general, and *-dai* for machinery and machine-like objects. Six-year-olds had an average of 7 classifiers, 7-year-olds an average of 9, 8-year-olds an average of 12, and by age 12 children had an average of 22.3 classifiers, while adults had 28 classifiers. As compelling as Sanches' evidence appears, it must be interpreted with caution: Sanches does not explicate what happened before age five nor how acquisition took place. Moreover, Sanches' data for children under age 9 is not reliable because her sample size and data collection method were both unspecified.

In a later study, Muraishi (1983) conducted a series of experiments with children from the first grade to the fourth grade (ages 6 to 9), using college students as a control group, to investigate children's classifier acquisition process. To test children's knowledge about the subtle differences among semantically similar classifiers, the experimenter contrasted objects that required classifiers from the same category and asked children which classifier they would use to count the object. The experimenter encouraged

children to use more specific classifiers and also asked the rationale for the selection of the classifiers.

Even though there was a gradual increase in the use of correct classifiers with increasing age, there was an enormous difference between the performance of elementary school children and that of the adult control group. The typical adult explanation cited features such as size, shape, function, structure, material, and category names to explain the selection of classifiers, while the children tended to use general classifiers and provided features only for prototypical objects. By contrast, many children avoided responding to the experimenter because they could not come up with specific classifiers. In many cases it was too hard for children to explain why they chose particular classifiers (Muraishi, 1983).

Muraishi argued that there was not only a strong correlation between performance on the numeral classifier test and knowledge of the names of objects used in the test, but also a tendency for a correlation between performance on the numeral classifier test and a knowledge of Chinese characters which represent numeral classifiers. Muraishi also concluded that although acquisition of the conceptual basis of the numeral classifier system was under way, children's concepts of numeral classifiers were not structured yet with the tested age groups.⁴

Among the acquisition studies of Japanese numeral classifiers, Matsumoto's (1985a, 1985b, 1987) research is probably the most conclusive. Matsumoto conducted a series of tests on children from ages 5 to 7 on the acquisition of 12 numeral classifiers in the animal, configurational (shape-specific), and non-configurational (functional) categories. In his pilot study, Matsumoto also discussed data from ages 2 to 4, but this was based on the natural speech data from Okubo (1967) and Murata (1983). In the experimental procedure, the experimenter showed photographs or pictures of objects and asked children the number of objects in each picture. Matsumoto prepared pictures with both familiar and unfamiliar objects (but not novel objects) to test the extent of the children's knowledge of classifiers. The data revealed that *-tsu* (inanimate generic classifier) and/or *-ko* (shape-specific classifier for 3D objects) emerged as general classifiers before other classifiers. Among the animal classifiers, *-hiki* (classifier for small animals and insects) was acquired first: Most 5-year-old children showed clear use of *-hiki*, but gave relatively few responses with *-wa* (classifier for birds) and *-too* (classifier for large animals). Among the configurational (shape-specific) classifiers, *-ko* appeared first, followed by *-mai* (classifier for 2D objects) and *-hon* (classifier for 1D objects); *-tsubu* (classifier for tiny objects like grains of rice) was acquired much later.

Matsumoto was unable to establish any specific acquisition order between *-mai* and *-hon*. Among the non-configurational (functional) classifiers, *-dai* (a classifier for machine-like objects) emerged early as a general classifier for vehicles, then later the more specific classifiers *-ki* (classifier for airplanes) and *-soo* (classifier for boats) were acquired. Other non-configurational classifiers, *-ken* (classifier for houses) and *-satsu* (classifier for bound objects like books) were also acquired relatively late. Overall, children showed roughly the same correct responses for both familiar and unfamiliar entities, implying that those who used correct classifiers had an abstract knowledge of the semantic categories of their referents. Here, Matsumoto especially noted that *-tsu* and *-ko* were frequently used in place of specific classifiers and that there was a change in progress among the younger generation, where *-ko* was becoming a type of generic classifier.

Finally, Uchida and Imai (1996) tested the acquisition of animate classifiers by having children from ages 4 to 6 count items in picture cards. Children acquired animate classifiers in the order of *-ri/-nin* (classifier for human beings), *-hiki* (classifier for small animals and insects), *-too* (classifier for large animals), *-wa* (classifier for birds). Children seemed to acquire numeral classifiers gradually, acquiring classifiers of high frequency first and using classifiers with prototypical items first. Uchida and Imai found that by age six, children could generate the overall pattern of the animate classifiers; thus these researchers assumed a correlation between conceptual development with respect to knowledge of living kinds and acquisition of animate classifiers.

Figure 2 shows the overall picture of the Japanese classifier acquisition process as presented by previous acquisition research (Sanches, 1977; Muraishi, 1983; Matsumoto, 1985a, 1985b, 1987; Uchida and Imai, 1996). At around one and a half years of age, Japanese children start producing numeral expressions (Matsumoto, 1985a),⁵ and until age four they produce about four classifiers: *-tsu* an 'inanimate, general classifier'; *-ko* for 'small 3D objects'; and *-ri/-nin* for 'human beings.' Early on, children use these four classifiers rather randomly. At around five, most children start acquiring the unmarked classifier, *-hiki* for non-human animate beings, and then the more specific animal classifiers, shape-specific, and functional classifiers. The acquisition of numeral classifiers continues gradually, but even older children (Matsumoto's 7-year-olds and Muraishi's 9-year-olds) do not display adult-like use of the classifier system.

Cross-linguistically children can acquire the basic syntactic patterns of numeral classifier phrases very early on (as young as 2 years old), but the acquisition of a corresponding semantic system seems to be a much slower process (Sanches, 1977; Gandour et al., 1984; Matsumoto, 1985a, 1985b,

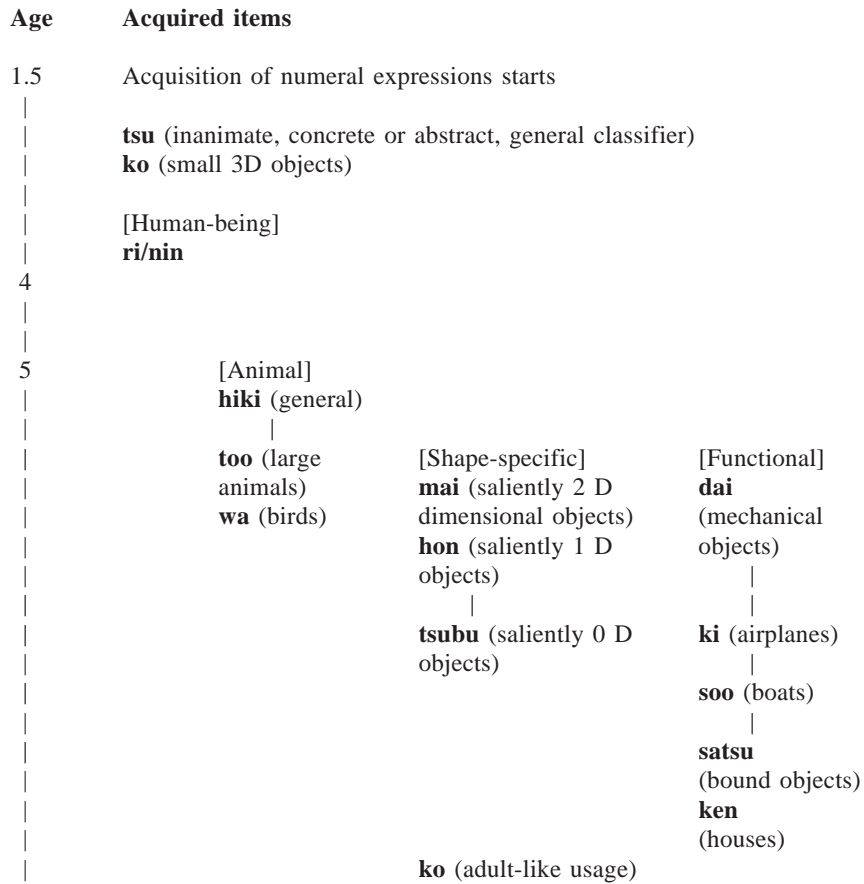


Figure 2. Acquisition order of Japanese numeral classifiers (based on production studies by Sanches, 1977; Muraishi, 1983, Matsumoto, 1985a, 1985b; Uchida and Imagi, 1996).

1987; Carpenter, 1987, 1991). Children first acquire a general classifier as a place holder and overuse it for any referent, a tendency that persists in much older children (Sanches, 1977; Muraishi, 1983; Gandour et al., 1984; Matsumoto, 1985a, 1985b, 1987; Carpenter, 1987, 1991). Studies of Thai numeral acquisition, however, reveal a slight discrepancy in this pattern: In Thai, across-the-board usage of one particular classifier occurs at a very young age along with overuse of repeaters⁶ across all age groups (Gandour et al., 1984; Carpenter, 1991). Generally speaking, pin-pointing the precise acquisition order of specific classifiers is difficult, although the assumption has been made that animate classifiers emerge relatively earlier than classifiers in other categories (Gandour et al., 1984; Matsumoto,

1985a, 1985b, 1987; Carpenter, 1987, 1991). However, in the case of animate classifiers it has been claimed that children may have the (cognitive) concept of animacy long before its use in the (linguistic) system at age 5 (Gelman and Coley, 1990; Mandler, Bauer, and McDonough, 1991). Research to date therefore leaves an unclear picture of how grammatical and conceptual coordination of specific classifiers occur in development.

1.2. *Classifier Acquisition in the Light of Recent Research on Cognitive Development and Language Acquisition*

In addition to the somewhat unclear development pattern of classifier acquisition, there appear to be certain discrepancies between recent research in cognitive development and work on classifier acquisition. In other areas of cognitive development, young children have been shown to use complex concepts in a flexible and productive manner. For example, two and a half-year-olds can overlook perceptual appearance in favor of category membership when appropriate (Gelman and Coley, 1990); and inductions about word meaning are guided by knowledge of two major ontological categories – object and substance (Soja, Carey and Spelke, 1991). Moreover, it has been shown that children who are just at the dawn of language acquisition (18 months) seem to understand such global conceptual categories as animals and vehicles (Mandler, Bauer, and McDonough, 1991).

The results of studies in a variety of languages have shown impressive early conceptual competencies in the acquisition of noun classes, categories which play a role similar to that of numeral classifiers. In Sesotho, a Bantu language spoken in South Africa, children were found to mark agreement productively with demonstratives and possessives at age 2 and mark most nouns appropriately at age 3 (Demuth, 1992). In French, children first use nouns without articles and when they start using articles, make widespread gender agreement errors, but by age 3 can assign proper articles to nouns most of the time (Clark, 1983). In Spanish, children master gender-marking and gender-agreement before they are 4 years old (Perez-Pereira, 1991). A study in German has reported that no errors in pronoun use in connection with the natural-gender rule were found in 3-year-olds' natural speech (Mills, 1986). And Icelandic provides a clear example of natural category clues to facilitate the learning of an agreement system. In Icelandic, nominal endings represent case, number, and gender, with nouns of one gender class having inconsistent phonological shapes, such that there is no phonological correspondence between nominal and pronominal forms. While Icelandic children cannot reliably use noun endings to determine pronoun reference before age 7, at age 4, they can use natural gender correctly to

interpret pronoun reference (Mulford, 1985). Thus, where formal cues are obscure, children look for semantic cues to determine agreement.

Considering such recent findings in cognitive development and noun class acquisition, the literature on the acquisition of numeral classifiers does not seem to fit well with other work on semantic and conceptual development.

2. PRESENT RESEARCH

Previous research has not provided a clear view of how children acquire classifier systems. Moreover, there are discrepancies between the previous research in classifier acquisition and recent research in cognitive development and noun class acquisition. We see no reason to assume that either the conceptual complexity of the classifier categories or the grammatical complexities of their usage should cause children to be unable to link up grammar and conceptual structure until well into middle childhood. Instead, the apparent developmental delay may be occurring for very different reasons that in the end will reveal, almost paradoxically, surprisingly sophisticated and abstract linkages between grammatical and conceptual structures in very young children. The appearance of delay may be due to the faulty assumption that failures with lower level categories entail failures with higher level ones. In this study, we explore the alternative hypothesis that young children can easily conceptualize categories in numeral classifier systems and link them to grammatical forms. Previous research may have under-estimated children's ability to acquire the numeral classifier system accurately because the conclusions are all drawn from children's production data, except for two Chinese studies (Fang, 1985;⁷ Hu, 1993).⁸

Comprehension does not always precede production, but children do often seem to know much more than their speech reflects (Hirsh-Pasek and Golinkoff, 1996). Indeed Hu (1993), in a study of Chinese numeral classifier acquisition, reported a large gap between children's comprehension and production ability, which implies a much earlier onset of acquisition – though his study is limited to shape-specific classifiers. Thus, in this study, we focus on children's comprehension of numeral classifier; our results suggest that both an excessive focus on production tasks and misleading assumptions about the ease of categorization at different levels of abstraction may have created the illusion of lengthy development for classifiers.

In the studies reported below, we tested 11 numeral classifiers that are in frequent daily use in Japanese and have been identified in production research, enabling a direct comparison between production and comprehension ability of the children. The 11 classifiers are from four different

categories: the classifier *-ri* for human beings;⁹ three animal classifiers *-hiki*, *-too* and *-wa*;¹⁰ three shape-specific classifiers, *-hon*, *-mai*, and *-ko*;¹¹ three functional classifiers *-dai*, *-soo*, *-ki*;¹² and the default inanimate classifier *-tsu*.¹³ Since children had great difficulty identifying the differences among classifiers from the same category in Muraishi's (1983) experiment, we tested the comprehension of numeral classifiers under two different conditions to control the degree of difficulty. In Experiment 1, classifiers from different domains were compared, where the contrast among tested classifiers was strong (for example, animal classifiers vs. shape classifiers vs. functional classifiers) (see Table I). In Experiment 2, classifiers from the

TABLE I
List of picture stimuli for Experiment 1

Test 1		
Hiki 1 (animals)	Ri (human)	Tsu (inanimate obj)
foxes	men	chairs
dragonflies	children #1	broccoli
frogs	children #2	stones
fish	women	clocks
Test 2		
Wa (birds)	Hon (1D objects)	Soo (boats)
parrots	carrots	blue boats
seagulls	spoons	brown boats
geese	pencils	yachts
hens	trees	speed boats
Test 3		
Too (large animals)	Mai (2D objects)	Ki (airplanes)
elephants	shirts	yellow planes
lions	yellow leaves	helicopters
horses	green leaves	red planes
cows	towels	jet planes
Test 4		
Hiki 2 (animals)	Ko (3D objects)	Dai (machines)
cats	apples	telephones
fish	stones	cars
mice	candies	trucks
butterflies	cups	TVs

same domains were compared, where the contrast among tested classifiers was weak (for example, a classifier for small animals vs. a classifier for large animals vs. a classifier for birds) (see Table II).

3. EXPERIMENT 1

3.1. *Method*

3.1.1. *Subjects*

One hundred fifty-seven children from three Japanese daycare centers in a suburb of Tokyo participated in the experiment. They were broken down into four age groups. Eighty children participated in Tests 1 and 2, including: 19 children with a mean age of 3 years and 5 months (henceforth abbrev-

TABLE II
List of picture stimuli for Experiment 2

Test 1 (functional CLs)		
Ki (air planes)	Soo (boats)	Dai (machines)
red planes	blue boats	cars
jet planes	brown boats	trucks
helicopters	speed boats	telephones
yellow planes	yachts	TVs
Test 2 (shape CLs)		
Hon (1D objects)	Mai (2D objects)	Ko (3D objects)
carrots	yellow leaves	apples
pencils	towels	cups
trees	green leaves	broccoli
spoons	shirts	stones
Test 3 (animal CLs)		
Too (large animals)	Hiki (animals)	Wa (birds)
lions	mice	parrots
cows	fish	seagulls
elephants	cats	geese
horses	butterflies	hens

viated as 3;5) (range 2;11 to 3;11); 22 children with a mean age of 4;5 (range 4;0 to 4;11); 24 children with a mean age of 5;5 (range 5;0 to 5;10); and 15 children with a mean age of 6;4 (range 5;11 to 6;9). There were 9 females and 10 males in the 3-year-old group, 13 females and 9 males in the 4-year-old group, 11 females and 13 males in the 5-year-old group and 7 females and 8 males in the 6-year-old group.

Seventy-seven children participated in Tests 3 and 4, including: 17 children with a mean age of 3;4 (range 2;11 to 3;11); 22 children with a mean age of 4;5 (range 4;0 to 4;11); 20 children with a mean age of 5;6 (range 5;0 to 5;11); and 18 children with a mean age of 6;4 (range 6;0 to 6;9). There were 7 females and 10 males in the 3-year-old group, 12 females and 10 males in the 4-year-old group and an equal number of females and males in the other two age groups.¹⁴

3.1.2. *Stimuli*

Experiment 1 consisted of four tests (see Table I). For each test, we prepared 12 cards, each containing three pictures. The pictures were color photographs and colored illustrations of familiar objects selected from children's books. In each picture there were two instances of the same item. The three pictures in a card corresponded to three different numeral classifiers, respectively. The placement of the pictures was controlled so that each picture appeared three times in all possible placements, left, center, and right on the cardboard (see Figure 3). Since each numeral classifier was tested four times, there were four three-card sets. Each picture stimulus was tested by 13 to 15 adult native speakers for the "correct" or "appropriate" classifiers, and there was more than 95% of agreement on the classifier for each picture stimulus.

In Experiment 1, numeral classifiers from different domains were tested against each other. Therefore the contrasts among classifiers were maximized, that is, animal classifiers vs. shape-specific classifiers vs. functional classifiers (see Table I).

3.1.3. *Procedure*

Training sessions. Before the test trials, one or two training sessions were provided. For the training sessions, three sets of three cards (in total nine cards) were prepared. Three numeral classifiers that were not included in the test trials were selected from Downing's (1984, 1996) list of the 30 most commonly used numeral classifiers. The training sessions were carried out with an individual child in a small room in a preschool to avoid any

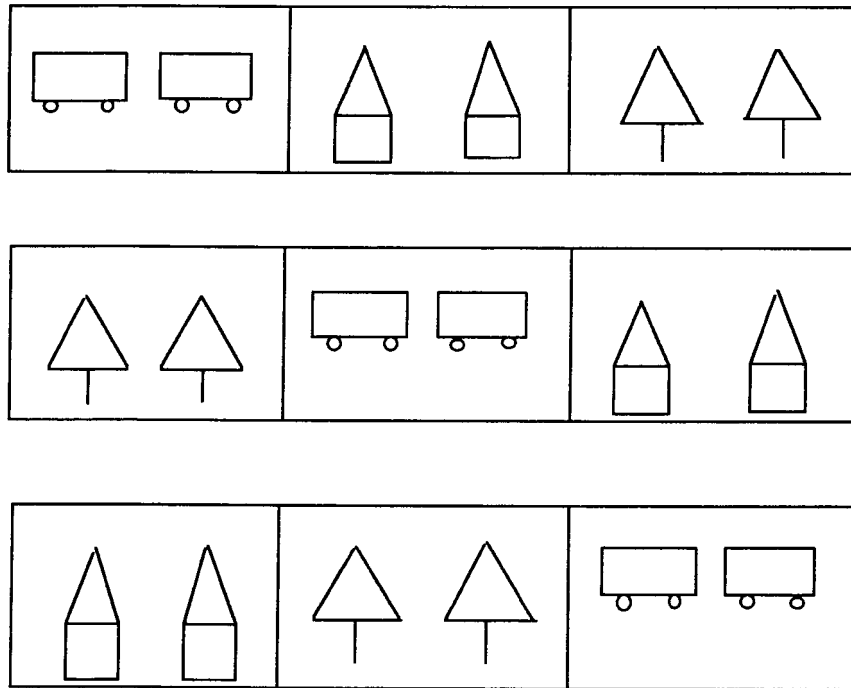


Figure 3. Drawn examples of picture stimuli (not original stimuli).

distractions from other children. The experimenter explained to the child that they were going to play a point-to-a-picture game. The experimenter would count the items in one of three pictures in a card; then the child had to decide which picture the experimenter counted and point to it. After the explanation, the experimenter placed one of the cards containing three different pictures in front of the child and counted the items in one of the three pictures. The verbal stimulus was 'one-CL, two-CL', in Japanese *is-satsu, ni-satus*, "one-(bound object), two-(bound objects)." Since the stimulus utterance did not have an actual referent mentioned, the child had to identify objects based on the classifier. Then the experimenter asked "Which picture do you think I counted?" or "Which picture do you count this way?" The classifier *-satsu* indicates bound objects such as books, so the correct response would have been to choose a picture of two books. After every trial, feedback was provided. The experimenter also encouraged the child to pay attention to numeral classifiers by asking questions like *Doo kazoeta?* "How did I count it?" or *Nante itta?* "What did I say?" All the children in the 3-year-old group and about half of the 4-year-old group

participated in two training sessions to understand the procedure, while the rest of the children participated in one training session. There was about a two to three day interval between the two training sessions.

Sense classifier tests. Four to seven days after the training session, the children were tested individually in the same room where the training sessions had taken place. The experimenter explained to the children that she wanted to play the point-to-a-picture game again and so she had prepared many new pictures. The trials were carried out in the same manner as the training session. Three practice trials preceded the actual trials using the cards from the training session. In the trials, to prevent any learning strategies in the twelve-trial blocks, the twelve cards were presented randomly. Each child participated in two tests at a time, either Tests 1 and 2 or Tests 3 and 4. The order of the two tests was counter-balanced for each child, and no feedback was provided. The experimenter recorded the child's response on a record sheet.

Nonsense classifier tests. Four to seven days after the sense classifier tests, the same children participated in the nonsense classifier tests to measure any picture biases. The procedure was identical to the sense classifier tests except that there were no practice trials. Three classifiers in each test were replaced with three nonsense words. The children were told that they were going to play the point-to-a-picture game again with the same set of pictures, but this time the picture would be counted in a foreign language, for example, *ichi-gie, ni-gie* 'one-gie, two-gie' or *ichi-pekku, ni-pekku* 'one-peck, two-peck', where *gie* and *pekku* are not actual Japanese classifiers.

3.2. Results

The number of correct responses for each classifier from the four trials was aggregated in each sense classifier test. Since it was extremely difficult to control so that every picture in the picture sets would be equally salient for all trials, we used the nonsense classifier test results as the base line rather than chance. In the nonsense classifier tests, each classifier was replaced with a nonsense classifier in the stimulus utterance; therefore children had to completely guess to select the pictures. Thus children's performance with the sense classifier tests was tested against nonsense classifier tests to examine whether performance actually reflected an understanding of the classifiers. The mean percentages of correct responses from the sense and nonsense classifier tests are shown in Table III.

TABLE III
Mean percentage of correct responses for Experiment 1

	3 years old	4 years old	5 years old	6 years old
Test 1.	(N = 76)	(N = 88)	(N = 96)	(N = 60)
Hiki 1 (animal)	0.55 [0.51]	0.77*** [0.52]	0.85*** [0.37]	0.83*** [0.44]
Ri (human being)	0.58*** [0.32]	0.67*** [0.18]	0.89*** [0.43]	0.80*** [0.39]
Tsu (inanimate obj)	0.26 [0.22]	0.34* [0.24]	0.57*** [0.22]	0.75*** [0.46]
Test 2.	(N = 76)	(N = 88)	(N = 96)	(N = 60)
Wa (bird)	0.32 [0.29]	0.59*** [0.27]	0.63** [0.38]	0.78*** [0.25]
Hon (1D object)	0.48** [0.28]	0.56*** [0.31]	0.77*** [0.33]	0.93*** [0.23]
Soo (boat)	0.34 [0.31]	0.48* [0.35]	0.65*** [0.18]	0.68*** [0.25]
Test 3.	(N = 68)	(N = 88)	(N = 80)	(N = 72)
Too (large animal)	0.50*** [0.28]	0.44*** [0.28]	0.58*** [0.31]	0.72*** [0.21]
Mai (2D objects)	0.27 [0.30]	0.58*** [0.22]	0.88*** [0.44]	0.79*** [0.31]
Ki (airplanes)	0.40 [0.27]	0.36 [0.36]	0.51 [0.50]	0.56* [43]
Test 4.	(N = 68)	(N = 88)	(N = 96)	(N = 60)
Hiki 2 (animal)	0.44 [0.31]	0.77*** [0.36]	0.93*** [0.28]	0.90*** [0.39]
Ko (3D object)	0.34 [0.38]	0.47*** [0.28]	0.64*** [0.26]	0.86*** [0.25]
Dai (machine)	0.43 [0.35]	0.63*** [0.30]	0.81*** [0.29]	0.88*** [0.44]

* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$.

[] Base line from the nonsense classifier test.

At age 6, the interpretation of all the tested classifiers was strongly significant compared to the nonsense classifiers at $p < 0.0005$ (one tailed t test), except for the classifier *-ki* for airplanes, which was significant only at $p < 0.005$. At age 5, the interpretation of the classifier *-ki* was not significant, while the rest of the classifiers were strongly significant at $p < 0.0005$ for 9 classifiers and at $p < 0.005$ for the bird classifier *-wa*. At age 4,

except for *-ki*, all the classifiers were strongly significant at $p < 0.0005$ for 8 classifiers and at $p < 0.005$ for 2 classifiers. At age 3, only three classifiers were strongly significant: *-too* for large animals ($p < 0.0005$), *-hon* for one-dimensional objects ($p < 0.005$), and *-ri* for human beings ($p < 0.0005$). The rest of the classifiers were not significant.

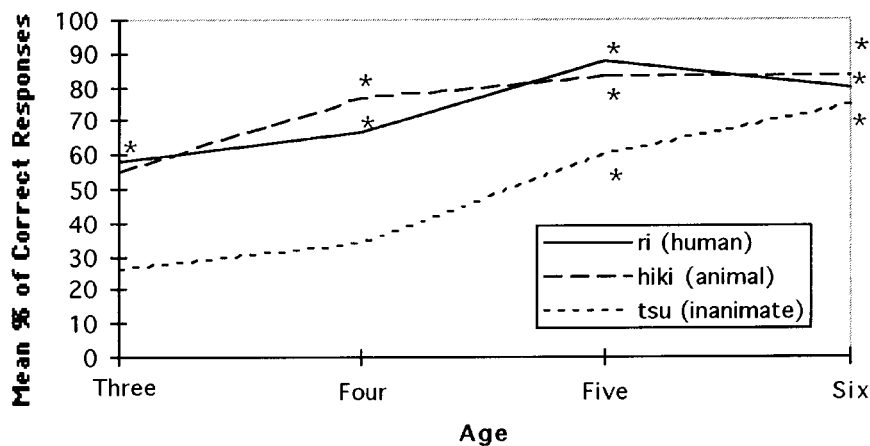
The results reveal that at age 4, children start comprehending almost all the tested numeral classifiers, which included three general classifiers and eight specific classifiers. But the classifier *-ki* for airplanes seems to be the last acquired of these classifiers; adult-like interpretation (correct responses in over 90% of the test trails) of this specific classifier does not seem to occur until some time after age 6.

Even at age 3, children start differentiating the general classifier *-ri* for human beings from the other general classifiers, *-hiki* for animals and *-tsu* for inanimate objects. They also differentiate the classifier *-too* for large animals and *-hon* for one-dimensional long objects from the other specific classifiers tested together. We tested the classifier *-hiki* twice: in Test 1 as a general classifier for animals (*-hiki 1*), and in Test 4 as a specific classifier for small animals and insects (*-hiki 2*). At age 5 and 6, the mean percentage of correct responses for *-hiki 2* in Test 4 was higher than for *-hiki 1* in Test 1. The objective in Experiment 1 was to maximize the contrast among the three classifiers, but it is likely that having two animate categories, animal and human in Test 1, made the contrast less strong, and this might have affected the children's performance with the classifier *-hiki*.

In Figures 4a, 5a, 6a, and 7a, the mean percentage of correct responses from the sense classifier tests are grouped together based on the classifier categories. A separate 4 (Age: 3, 4, 5, or 6 years old) \times 3 (Classifiers: word 1, word 2, or word 3) \times 2 (Gender: male or female) analysis of variance (ANOVA) was conducted for each of the four categories of classifiers (general classifiers, animal classifiers, shape-specific classifiers, and functional classifiers) to see if there were significant differences in the comprehension of classifiers among the four age groups. All four ANOVAs yielded significant main effects ($p < 0.0001$) for age, and all but the ANOVA for shape-specific classifiers also yielded significant main effects ($p < 0.0001$) for classifiers. There were no effects for gender and no interactions.

Scheffé post hoc analyses for general classifiers revealed that there was a significant difference in comprehension among 3-, 4-, and 5-year-olds but no difference between 5- and 6-year-olds. Children comprehended the classifiers *-ri* for human beings and *-hiki* for animals better than the classifier *-tsu* for inanimate objects.¹⁵ There was a gradual development of understanding for *-ri* and *-hiki* from age 3 to 6, and 5- and 6-year-olds

(a) Comprehension



* Test result was significantly higher than base rate.

(b) Production (Matsumoto, 1985a)

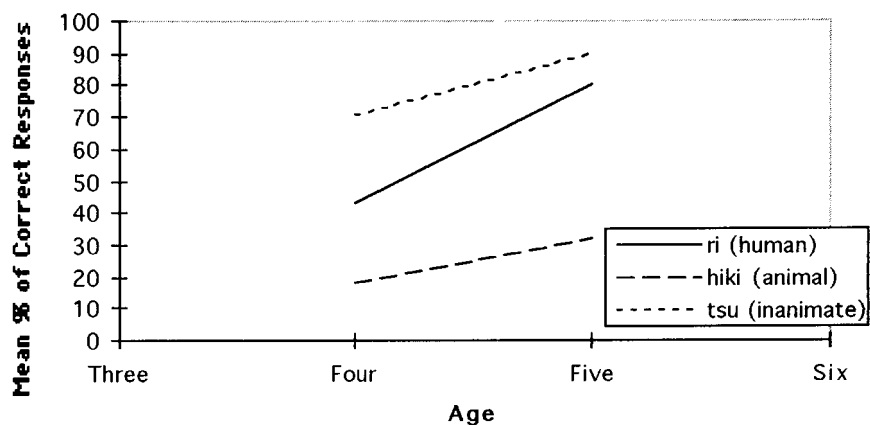
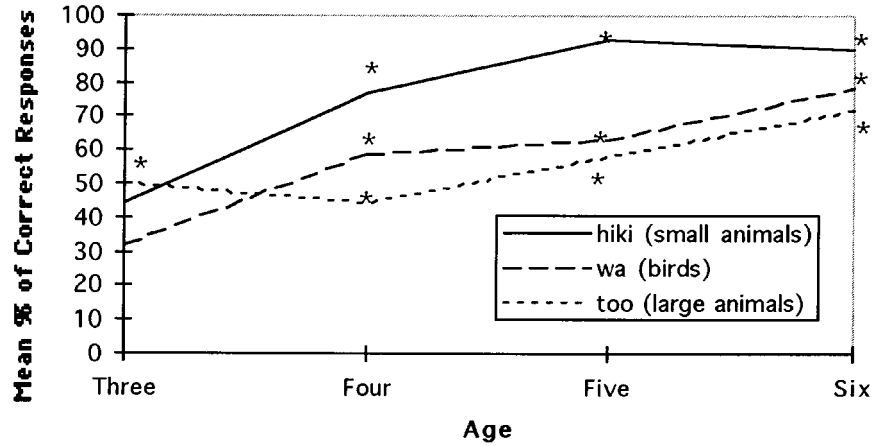


Figure 4. General classifiers.

understood *-tsu* better than 3- and 4-year-olds. For animal classifiers, besides age and classifier effects, there was a two-way interaction between age and classifier ($F(6, 219) = 2.42, p < 0.03$). Scheffé post hoc analyses indicated that 4-year-olds comprehended animal classifiers significantly better than 3-year-olds, and 6-year-olds comprehended them significantly better than 4-year-olds. Children comprehended the classifier *-hiki* significantly better than *-wa* and *-too*, and there was no significant difference between *-wa* and *-too*. For shape-specific classifiers, there was a significant devel-

(a) Comprehension



* Test result was significantly higher than base rate.

(b) Production (Matsumoto, 1985a)

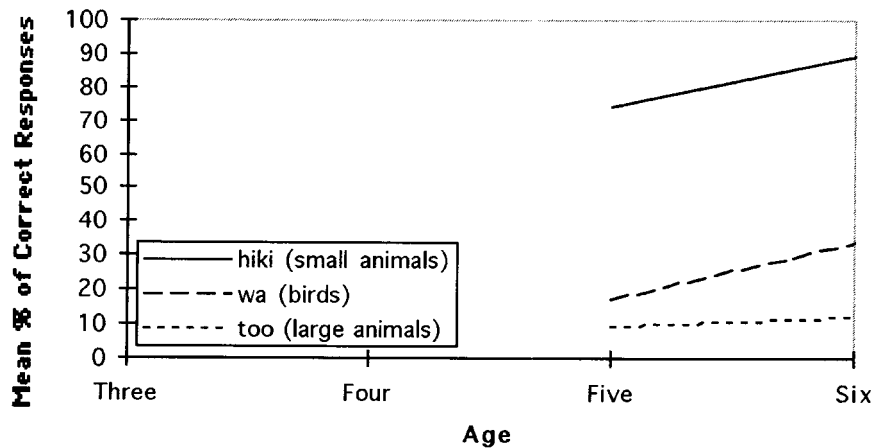
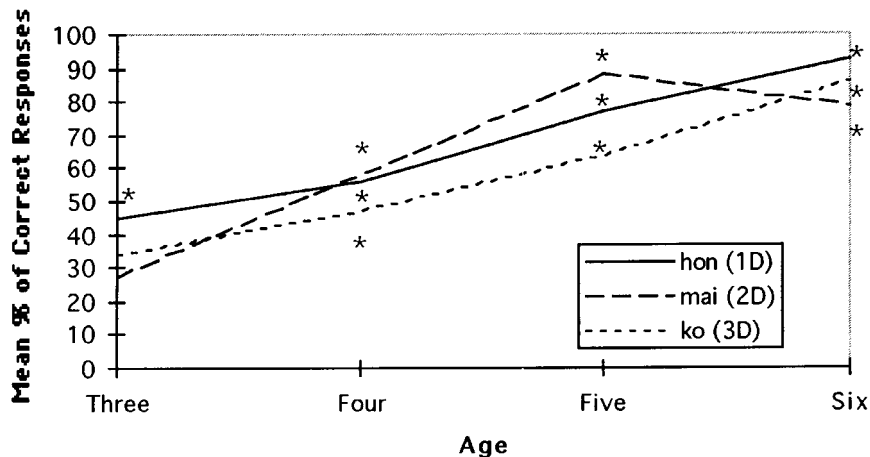


Figure 5. Animal classifiers.

omprehension from age 3 to 6, but there was no significant difference among comprehension of the three shape-specific classifiers. By age 5 children seemed to have close to adult-like understanding (correct responses in over 75% of the test trials) of all three classifiers. For functional classifiers, there was a significant difference between the comprehension of 3- and 4-year-olds and 5- and 6-year-olds. Children's comprehension of the classifier *-dai* for land vehicles and machines was significantly better than the classifiers *-soo* for boats and *-ki* for airplanes. The results indicate that by age

(a) Comprehension



* Test result was significantly higher than base rate.

(b) Production (Matsumoto, 1985a)

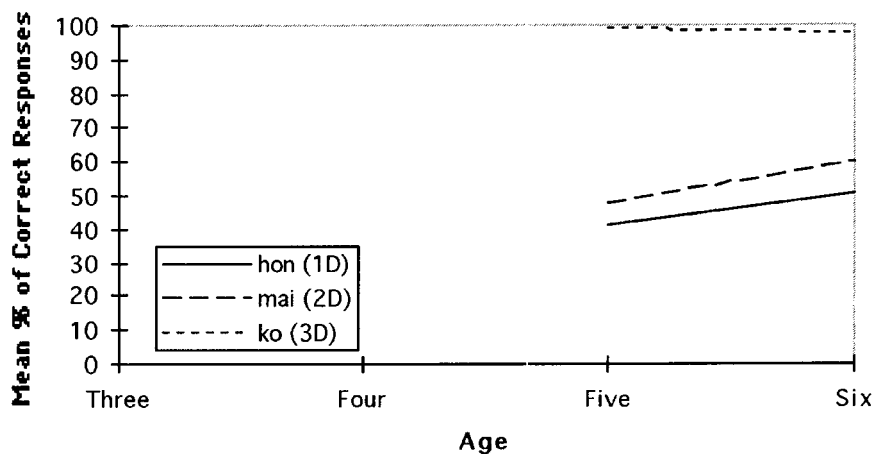
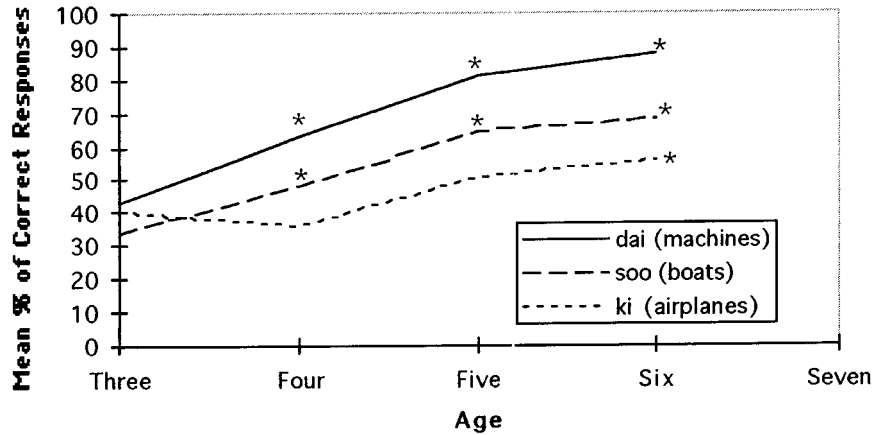


Figure 6. Shape-specific classifiers.

5 children have adult-like understanding of *-dai*, with comprehension of *-soo* following shortly thereafter, while *-ki* does not reach an adult-like level of comprehension until beyond age 6.

There is clearly a gap between children's comprehension and production abilities, and the pattern of difference is intriguing (see Figures 4–7). Since Matsumoto's (1985a, 1985b) experiments are the most comprehensive studies of production and 11 classifiers tested in our study were also

(a) Comprehension



* Test result was significantly higher than base rate.

(b) Production (Matsumoto, 1985a)

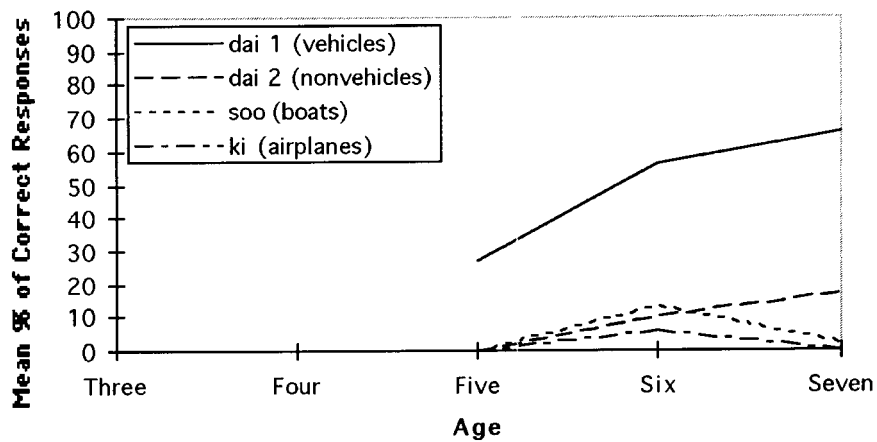


Figure 7. Functional classifiers.

tested in his experiments, we discuss our comprehension data in comparison to Matsumoto's (1985a, 1985b) production data.¹⁶

Production data (Matsumoto, 1985a, 1985b) have shown that children ages 4 to 7 are capable of using general classifiers such as *-tsu* (inanimate general classifier) and *-ri/-nin* (human general classifier), but are incapable of using specific classifiers such as *-wa* (birds), *-too* (large animals), *-hon* (1D objects), *-mai* (2D objects), *-dai* (vehicles and machines), *-soo* (boats), and *-ki* (airplanes). In addition to the general classifiers *-tsu*

and *-ri/-nin*, both *-hiki* (small animals and insects) and *-ko* (3D objects) have high mean percentages in production in this age range. This indicates that children use both of them as a general default classifier. Indeed, the animal classifier *-hiki* was originally used in Japanese for smaller sized animals and insects but is currently used more and more as a general default animal classifier. Matsumoto (1985a) tested *-hiki* twice, first with *-tsu* (inanimate general classifier) and *-ri/-nin* (human general classifier) plus a few inanimate classifiers as distracters, and second with other animal classifiers, *-wa* (birds) and *-too* (large animals) plus shape-specific classifiers and functional classifiers. The mean percentage of correct responses for *-hiki* was much higher in the second experiment (see Figure 5b) than in the first experiment (see Figure 4b). *-Ko* as a shape-specific classifier was originally used for three-dimensional objects whose size is manipulable by hand; it is now also used increasingly as a general default inanimate classifier like *-tsu*. Matsumoto treated both *-tsu* and *-ko* equally as correct responses when they were used to count four tested inanimate items: apples, paper clips, mountains, and clouds.¹⁷ In our comprehension experiments, we tested *-tsu* and *-ko* separately, and we used items such as chairs, broccoli, stones, and small clocks to test *-tsu*, and apples, stones, candies, and cups to test *-ko*. In Matsumoto's data, the functional classifier *-dai* was analyzed separately for two classes of objects, vehicles and nonvehicles (see Figure 7b). Children used *-dai* not only for trucks and cars but also planes and boats in place of *-ki* and *-soo*, but rarely used it for nonvehicles such as TV sets and typewriters. In the comprehension experiment, there was no prominent difference between the use of *-dai* for vehicles and the use of *-dai* for nonvehicles; therefore, we did not separate the data.

In comprehension, 6-year-olds showed adult-like interpretation (correct responses in over 90% of the test trials) of the classifier *-hiki 2* (small animals in Test 4) and the classifier *-hon* (1D objects); close to an adult-like interpretation (correct responses in over 75% of the test trials) for *-hiki 1* (animals in general in Test 1), *-ri* (human beings), *-wa* (birds), *-too* (large animals), *-tsu* (inanimate objects), *-hon* (1D objects), *-mai* (2D objects), *-ko* (3D objects), and *-dai* (cars and machines). In production, even 7-year-olds hardly ever produced *-ki* (airplanes) and *-soo* (boats), while in comprehension 6-year-olds responded correctly on 56% and 68% of the trials, respectively. Five-year-olds showed adult-like comprehension of the classifier *-hiki 2* (in Test 4); close to adult-like interpretation of *-hiki 1* (in Test 1), *-ri*, *-hon*, *-mai* and *-dai*, and responded correctly at least 57% of the time for the rest of the classifiers except *-ki*. Four-year-olds showed close to adult-like interpretation of *-hiki* and responded correctly at least

44% of the time for the rest of the classifiers except *-tsu* and *-ki*. Thus, in Experiment 1, with the strong contrast comprehension task, children showed a much more fine-grained categorical ability than had been observed in their production.

There was also a gap between the comprehension of the three-dimensional classifier *-ko* and the general default inanimate classifier *-tsu* and the production of these two classifiers. In Matsumoto's (1985a) first production experiment, 4- and 5-year-olds used *-tsu* and *-ko* correctly 70% to 90% of the time, depending on the item they counted. However, in our comprehension experiment, children interpreted *-tsu* correctly only 34% of the time at age 4, and 57% of the time at age 5. At age 6 children showed close to adult-like interpretation of *-tsu* (interpreted correctly 75% of the time). The comprehension of *-ko* was slightly better: Children interpreted this classifier correctly 47% of the time at age 4, and 64% of the time at age 5, and 86% of the time at age 6. Actually, in production, Matsumoto pointed out that errors in specific classifiers primarily involve the substitution of *-tsu/-ko* for another classifier. In the *inanimate* category, children substituted *-tsu/-ko* for specific classifiers 60 to 80% of the time, depending on the items. Even in the *animate* category, 4-year-olds substituted *-tsu/-ko* for *-ri/-nin* and *-hiki* 20 to 50% of the time and 5-year-olds 10 to 50% of the time. This indicates that children, especially at younger ages, classify these two classifiers as across the board default classifiers. Thus the comprehension data show that it is not simply the default general classifiers that children first start acquiring. In Experiment 1, given the strong contrast, the acquisition of specific classifiers is taking place even before that of the general classifier *-tsu*, which has the largest semantic range.

In Experiment 1, 11 classifiers were tested under conditions in which the contrast among the three classifiers was maximized, because classifiers from three different categories were tested against each other. In Experiment 2, we addressed the follow-up question of whether children have a much more precise categorical knowledge of numeral classifiers than previously recognized, by comparing classifiers within categories.

4. EXPERIMENT 2

4.1. Method

4.1.1. Subjects

Thirty-two children from a Japanese daycare center in a suburb of Tokyo participated in the experiment.¹⁸ All the children participated in three sense

classifier tests and three nonsense classifier tests. The younger group of 16 children had a mean age of 4;9 (range 4;6 to 5;3); the older group of 16 children had a mean age of 5;6 (range 5;4 to 6;0). There were equal numbers of females and males in both the 4-year-old group and the 5-year-old group.

4.1.2. *Stimuli*

Experiment 2 consisted of three weak contrast tests. In each test, three classifiers from the same category were tested against each other (see Table II). Most of the pictures from Experiment 1 were used in Experiment 2. However, the combinations of the pictures were arranged to present three within-category classifiers, thereby minimizing the contrast among the three classifiers.

4.1.3. *Procedure*

The procedures were identical to those in Experiment 1. All the children participated in one or two training sessions, three sense classifier tests, and three nonsense classifier tests.

4.2. *Results and Discussion*

The data were analyzed in the same manner as in Experiment 1. The numbers of correct responses for each classifier from the four trials were aggregated, with the base rate being derived from the nonsense classifier test results. The mean percentages of correct responses from the sense and nonsense classifier tests are shown in Table IV in comparison to the results from Experiment 1.

At ages 4 and 5, the children did not comprehend all three functional classifiers in Experiment 2 (no significant difference between sense and nonsense classifier tests). At age 4, the shape-specific classifier *-ko* for 3D objects and the classifier *-too* for large animals were interpreted incorrectly, but children started interpreting the rest of the tested classifiers correctly ($p < 0.005$). At age 5, children started interpreting both shape-specific classifiers and animal classifiers correctly in Experiment 2 ($p < 0.0005$).

To see if there was any significant difference between Experiment 1 and Experiment 2, a separate 2 (Age: 4 or 5 years olds) \times 3 (Classifiers: word 1, word 2, or word 3) \times 2 (Gender: male or female) \times 2 (Condition: strong contrast or weak contrast) analysis of variance (ANOVA) was con-

TABLE IV
Mean percentage of correct responses for Experiments 1 and 2

	Four years old		Five years old	
	Experiment 1 (strong contrast)	Experiment 2 (weak contrast)	Experiment 1 (strong contrast)	Experiment 2 (weak contrast)
Test 1	(N = 88)	(N = 64)	(N = 80)	(N = 76)
Ki (air planes)	0.36 [0.36]	0.38 [0.36]	0.51 [0.50]	0.38 [0.38]
Soo (boats)	0.48* [0.35]	0.42 [0.41]	0.65*** [0.18]	0.39 [0.39]
Dai (machines)	0.63*** [0.30]	0.36 [0.34]	0.81*** [0.29]	0.33 [0.33]
Test 2	(N = 88)	(N = 64)	(N = 80)	(N = 76)
Hon (1D objects)	0.56*** [0.31]	0.53** [0.39]	0.77*** [0.33]	0.59*** [0.38]
Mai (2D objects)	0.58*** [0.23]	0.52*** [0.28]	0.88*** [0.44]	0.77*** [0.23]
Ko (3D objects)	0.47*** [0.28]	0.42 [0.38]	0.64*** [0.26]	0.58** [0.30]
Test 3	(N = 88)	(N = 64)	(N = 80)	(N = 76)
Too (large animals)	0.44*** [0.28]	0.36 [0.36]	0.58*** [0.31]	0.53*** [0.28]
Wa (birds)	0.59*** [0.27]	0.50** [0.30]	0.63*** [0.38]	0.63*** [0.36]
Hiki (small animals)	0.77*** [0.36]	0.63*** [0.36]	0.93*** [0.28]	0.63*** [0.42]

* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$.

[] Base line from the nonsense classifier test.

ducted for each of the three categories of classifiers (functional, shape-specific, and animal classifiers). An ANOVA for functional classifiers yielded main effects for classifier ($p < 0.0003$) and for condition ($p < 0.02$) but there were no effects for age and gender. Scheffé post hoc analyses revealed that children comprehend the classifier *-dai* for land vehicles and machines significantly better than the classifiers *-ki* for airplanes and *-soo* for boats, but there was no significant difference between 4- and 5-year-olds' comprehension ability. An ANOVA for shape-specific classifiers yielded main

effects for age ($p < 0.0001$) and for classifier ($p < 0.004$). There was a trend for condition ($p < 0.06$) but no effects for gender. Scheffé post hoc analyses revealed that children comprehend *-hon* for 1D objects and *-mai* for 2D objects significantly better than *-ko* for 3D objects, and the 5-year-olds' comprehension was significantly better than that of the 4-year-olds'. An ANOVA for animal classifiers yielded main effects for age ($p < 0.002$), for classifier ($p < 0.0005$), and for condition ($p < 0.0032$), and there was an interaction between classifier and condition ($F(2, 225) = 7.09, p < 0.001$). Scheffé post hoc analyses revealed that children comprehend the classifier *-hiki* significantly better than the classifiers *-wa* and *-too*, but there was no significant difference between *-wa* and *-too*. The comprehension of *-hiki* in Experiment 1 was significantly better than in Experiment 2.

The results from the strong contrast (Experiment 1) and weak contrast conditions (Experiment 2) indicate that comprehension of the tested classifiers from three different categories proceeds through a differentiation of broader categories (animal classifiers vs. shape-specific classifiers vs. functional classifiers) to much finer distinctions (small animal classifier vs. large animal classifier vs. bird classifier). Although children could not differentiate specific classifiers from the same categories (weak contrast), children could differentiate them when they were tested against specific classifiers from different categories (strong contrast). It is likely that children start conceptualizing the classifier system at around age 3 at a broader and higher level in a hierarchical categorical structure, and then gradually master more specific subclasses until, by age 6, they have an adult-like semantic system for commonly used classifiers.

5. GENERAL DISCUSSION

5.1. *Conceptual Categories in Classifier Acquisition*

The strong contrast comprehension tests revealed a much greater sensitivity to the semantic relations underlying numeral classifiers in young children than was previously considered possible. Previous production research (Sanches, 1977; Muraiishi, 1983; Matsumoto, 1985a, 1985b, 1987; Uchida and Imai, 1996) has suggested a slow acquisition process in which first the inanimate general classifier *-tsu* and the 3D object classifier *-ko* appear as generic classifiers before other specific classifiers are used. Then in each category domain, acquisition proceeds from more general to specific classifiers. However, even at age seven, the generic inanimate classifier *-tsu* and the 3D classifier *-ko* were still over used with a large number of nouns. Yet in our comprehension studies, children as young as 4 years

showed a good grasp of 10 out of 11 tested numeral classifiers and even at age 3, children started to differentiate not only the generic classifier *-ri* but also the specific classifier *-too* from the other classifiers tested together. By age 6, children displayed a close to adult-like understanding with almost all 11 classifiers. The acquisition of classifiers proceeds, according to our comprehension experiments, not from general to specific classifiers, but through a differentiation of broader categories (animal classifiers vs. shape-specific classifiers vs. functional classifiers) to much finer distinctions (small animal classifier vs. large animal classifier vs. bird classifier). There is increasing evidence in the recent cognitive development literature that very young children often grasp broad categories before lower level ones, sometimes developing from abstract knowledge to more concrete knowledge (Simons and Keil, 1995).

One of the most striking demonstrations of a pattern of downwards category differentiation is seen in an object examination task in which infants' patterns of inspection of objects can be used to infer their intuitive categories. In such tasks, younger infants will show clear knowledge of the categories of animals, vehicles, and furniture but not knowledge of lower level categories (for example, type of furniture) until several months later (Mandler, Bauer, and McDonough, 1991). Contrary to earlier claims, basic-level categories (Rosch et al., 1976) do not always seem to form the preferred entry level of initial categorization in the development of hierarchical categorical systems. Similarly, in object sorting tasks, children could differentiate dogs from fish at 18 months, but could not differentiate dogs from rabbits or dogs from horses either. The latter discriminations came in several months later (Mandler, Bauer, and McDonough, 1991).

More broadly, children do not always first notice similarities in terms of lower level perceptual features (Mandler, 1999). Similarities in function, or causal powers can also be highly salient in an early age (Wellman and Gelman, 1988). For example, in the realm of biological knowledge, a series of demonstrations have shown that younger children often grasp a variety of the abstract properties of living things before they can fill in concrete details (Hatano and Inagaki, 1999; Keil et al., 1999). Similarly, knowledge of ontological categories, as revealed through judgments of what predicates can sensibly be applied to various categories, appears to differentiate downwards with age (Keil, 1979).

These studies of numeral classifiers provide yet another converging source of evidence that young children often come to initially grasp the structure of the world in ways that are better understood in cognitive than perceptual terms.

5.2. *Comprehension vs. Production*

We can only speculate as to what causes the discrepancy in the onset of acquisition between children's comprehension and production. Below we outline some of the factors which may complicate the acquisition of numeral classifiers in Japanese.

First, there are two series of numerals in Japanese, the native series and the Sino-Japanese series. There are also two kinds of numeral classifiers, native numeral classifiers such as *-tsu* and *-ri* and Sino-Japanese numeral classifiers such as *-ko* and *-nin*. Generally speaking, the Sino-Japanese numeral series co-occurs with the Sino-Japanese classifiers, while the native numeral series co-occurs with the native classifiers. In some cases, the two systems must be mixed; this is the case with the human classifiers *-ri* and *-nin* (see note 9). Even though a majority of the numeral classifiers are Sino-Japanese, it would be a considerable burden for children to use two series of numerals with the combination of numeral classifiers. Second, some classifiers alternate their phonological shape depending on the preceding numeral.¹⁹ These sound alternations do not reflect regular allophonic alternations in Japanese; instead they are highly lexicalized and specific to the classifier system. Thus, children may avoid using classifiers with sound alternations. Third, classifiers are communicatively marginal items. Even though a classifier indicates the inherent semantic features of the referent, with the presence of a noun in the construction, the information provided by the classifier is somewhat redundant and the use of an incorrect classifier does not necessarily break down communication.

All of these factors in combination could provide a plausible explanation for the relatively slow development in the production of numeral classifiers. The greater precocity in comprehension can be partly explained by the fact that many of these issues are specific to production or more dominant in production.

6. GENERAL CONCLUSION

This study has provided evidence that different techniques and stimulus materials can reveal a much greater sensitivity in young children to the conceptual underpinnings of the numeral classifier system than was previously considered possible. Given a strong contrast condition in comprehension tasks, very young children demonstrated they are capable of conceptualizing the highly abstract semantic system represented by the numeral classifier system. Contrary to the previous claims that acquisition proceeds from

general to specific numeral classifiers, even young children can identify specific numeral classifiers. However, the category differentiation takes place at a higher level in the beginning of acquisition, then proceeds downwards in the hierarchical categorical system.

NOTES

* We are very grateful to John Whitman for his tremendous help and advice from the beginning of this project and to the following people for their help with data and various versions of this work: Andrew Doyle, Giyoo Hatano, Yo Matsumoto, Joan Sereno, Yasuhiro Shirai, and Nobuko Uchida. We also thank the children and the teachers at the daycare centers for their indispensable participation.

¹ A number of researchers have reported on fine-grained semantic distinctions between these patterns. Discussion of such distinctions is beyond the scope of this paper, but see Miyamoto (1991, 1998).

² The subclassification of concrete object classifiers into "functional" and "shape-specific" is generally accepted and "the primes of function can themselves be either bases for a class or they may interact with shape primes to form classes" (Adams and Conklin, 1973).

³ Sanches (1977, 51–62) defines the inanimate generic classifier *-tsu* as a "cardinal number marker" and explains that "this particular alternative specifies nothing at all about the semantic properties of the thing being numerated."

⁴ We would suggest that Muraishi's (1983) result might be due to a method requiring advanced linguistic and cognitive skill, resulting in a large number of avoidance responses of "I do not know."

⁵ Matsumoto (1985a) analyzed the natural speech data from Okubo (1967) and Murata (1983) for children under age 4.

⁶ Repeaters are reduplications of the head noun or a part of the head noun, which occur in the syntactic slot for classifiers. Unlike true numeral classifiers, repeaters do not add any more information to that contributed by the head noun. (See Thai example in (i))

- (i) **kham si-kham**
word four-Repeater
'four words'

⁷ Fang (1985) tested the production of 12 frequently used Mandarin numeral classifiers by an elicitation task and also tested the comprehension of four Mandarin shape-specific classifiers by having children (4;0 to 6;0) choose the objects created from play-doh which matched the specific shape represented by each classifier.

⁸ Hu (1993) investigated both the production and comprehension of Mandarin numeral classifiers with children (3;0 to 6;0). Children's production was tested by a sentence elicitation test, and comprehension was tested by a sentence completion test in which children had to complete a numeral phrase by selecting an appropriate referent name from several pictures of referents.

⁹ *-Ri* is used to count up to two people; if the number is more than three, *-nin* is used. There are other human classifiers but *-ri/-nin* is used most commonly and can be considered as a default human classifier. In our study, there are two instances of items to count in each picture stimulus, thus only *-ri* was tested.

¹⁰ *-Wa* is used for birds, *-too* is used for animals which are larger than human beings, and *-hiki* is used for animals which are not referred to with *-wa* and *-too* and is also used to count animals in general.

¹¹ *-Hon* is used for one-dimensional concrete objects, *-mai* is used for two-dimensional

objects, and *-ko* is used for three-dimensional objects; however *-ko* is increasingly used as a default concrete inanimate classifier.

¹² *-Dai* is used for machines including cars, TVs and such, *-soo* is used for boats, and *-ki* is used for airplanes.

¹³ *-Tsu* is used for almost all kinds of inanimate entities, both concrete and abstract, although specific classifiers tend to be used as long as objects being counted meet the criteria of those specific classifiers.

¹⁴ The children who participated in Tests 1 and 2 also participated in Nonsense Classifier Tests 1 and 2. However, one male child from the 3-year-old group, one male child from the 4-year-old group, and 2 female children from the 6-year-old group did not participate in nonsense classifier tests. Those who participated in Tests 3 and 4 all participated in Nonsense Classifier Tests 3 and 4, but one female child from the 3-year-old group did not participate in Nonsense Classifier Test 3.

¹⁵ Production research reported the early acquisition of *-ri* and *-tsu* and the overgeneralization of *-tsu*. See additional discussion of the overgeneralization of *-tsu* and *-ko* at the end of Section 3.2.

¹⁶ Matsumoto's (1985a, 1985b) production data indicate the percentage of the correct responses in the total number of responses. In comparing production and comprehension, readers should note that the base line for the comprehension task is the result of the nonsense classifier test, which is around 33% but varies depending on the classifier (see Table III for the base line of each classifier).

¹⁷ Adult native speakers of Japanese use both *-tsu* and *-ko* for apples and paper clips but only *-tsu* for mountains and clouds.

¹⁸ This is a different center from the ones in Experiment 1; therefore, there is no overlap in the children tested.

¹⁹ For example, the voiceless fricative [h] in the 2D classifier *-hon* alternates into either a voiceless bilabial [p] or a voiced bilabial [b] such as [ip-pon] 'one-CL', [ni-hon] 'two-CL', [sam-bon] 'three-CL', [yon-hon] 'four-CL', and so forth.

REFERENCES

- Adams, Karen, and Nancy Conklin (1973) "Toward a Theory of Natural Classification," Papers from the ninth regional meeting of the Chicago Linguistic Society, University of Chicago, 1–10.
- Carpenter, Kathie (1987) *How Children Learn to Classify Nouns in Thai*, PhD dissertation, Stanford University.
- Carpenter, Kathie (1991) "Later Than Sooner: Extralinguistic Categories in the Acquisition of Thai Classifiers," *Child Language* 18, 93–113.
- Clark, Eve (1977) "Universal Categories: On the Semantics of Classifiers and Children's Early Word Meanings," in Alphonse Juilland (ed.), *Linguistic Studies Offered to Joseph Greenberg*, Anma Libri and Co., Saratoga, CA, pp. 449–462.
- Clark, Eve (1983) "Meanings and Concepts," in P. Mussen (ed.), *Handbook of Child Psychology, 3, Cognitive Development*, John Wiley and Sons, New York, pp. 787–840.
- Craig, Colette (1986) "Jacaltec Noun Classifiers," in C. Craig (ed.), *Noun Classes and Categorization*, John Benjamins, Amsterdam and Philadelphia, pp. 263–294.
- Croft, William (1994) "Semantic Universals in Classifier Systems," *Word* 45, 145–171.
- Demuth, Katherine (1992) "The Acquisition of Sesotho," in D. I. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition, 3*, Erlbaum, Hillsdale, NJ, pp. 305–321.
- Downing, Pamela (1984) *Japanese Numeral Classifiers: A Semantic, Syntactic, and Functional Profile*, PhD dissertation, University of California, Berkeley.
- Downing, Pamela (1996) *Numeral Classifier System: The Case of Japanese*, John Benjamins, Amsterdam.

- Erbaugh, S. Mary (1986) "Taking Stock: The Development of Chinese Noun Classifiers Historically and in Young Children," in Collette Craig (ed.), *Noun Classes and Categorization*, John Benjamins, Amsterdam and Philadelphia, pp. 399–436.
- Fang, Fuxi (1985) "An Experiment on the Use of Classifiers by 4 to 6 Year Olds," *Acta Psychologica Sinica*, 17(4), 384–392.
- Gandour, Jacm, Soranee Holasuit Petty, Rochana Dardarananda, Sumalee Dechongkit, and Sunee Mukngoan (1984) "The Acquisition of Numeral Classifiers in Thai," *Linguistics* 22, 455–479.
- Gelman, S. A., and J. D. Coley (1990) "The Importance of Knowing a Dodo is a Bird: Categories and Inferences in Two and a Half-year-old Children," *Developmental Psychology* 26, 796–804.
- Hatano, G., and Inagaki, K. (1999) "A Developmental Perspective on Informal Biology," in D. L. Medin and S. Stran (eds.), *Folkbiology*, MIT Press, Cambridge, Massachusetts, pp. 321–354.
- Hirsh-Pasek, K., and R. M. Golinkoff (1996) "The Intermodal Preferential Looking Paradigm: A Window onto Emerging Language Comprehension," in D. McDaniel, C. McKee, and H. S. Crains (eds.), *Methods for Assessing Children's Syntax*, MIT Press, Cambridge, Massachusetts, pp. 105–124.
- Hu, Qian (1993) *The Acquisition of Chinese Classifiers by Young Mandarin-Speaking Children*, PhD dissertation, Boston University.
- Keil, F. C. (1979) *Semantic and Conceptual Development: An Ontological Perspective*, Harvard University Press, Cambridge, Massachusetts.
- Keil, F. C., D. Levin, G. Gutheil, and B. Richman (1999) "Explanation Cause and Mechanism: The Case of Contagion," in D. Medin and S. Atran (eds.), *Folkbiology*, MIT Press, Cambridge, Massachusetts, pp. 285–320.
- Lee, Kwee-Ock (1994) "Acquisition of Korean Classifiers: Syntactic and Semantic Factors," Paper presented at the Symposium on Linguistic Theory and the Acquisition of Korean Semantics and Syntax, Seoul National University, Seoul, Korea, August 26, 1990.
- Mandler, J. M. (1999) "Seeing is Not the Same as Thinking: Commentary on 'Making Sense of Infant Categorization'," *Developmental Review* 19, 297–306.
- Mandler, J. M., P. F. Bauer, and L. McDonough (1991) "Separating the Sheep from the Goats: Differentiating Global Categories," *Cognitive Psychology* 23, 263–298.
- Matsumoto, Yo (1985a) *Japanese Numeral Classifiers: Their Structure and Acquisition*, Masters Thesis, Sophia University, Tokyo, Japan.
- Matsumoto, Yo (1985b) "Acquisition of Some Japanese Numeral Classifiers: The Search for Convention," *Stanford University Papers and Reports in Child Language Development* 24, 89–86.
- Matsumoto, Yo (1987) "Order of Acquisition in the Lexicon: Implication from Japanese Numeral Classifiers" in K. E. Nelson and A. Kleeck (eds.), *Children's Language*, 7, Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 229–260.
- Mills, A. E. (1986) "Acquisition of the Natural-gender Rule in English and German," *Linguistics* 24, 31–45.
- Miyamoto, Y. (1991) *Distributivity: A Comparative Study of English and Japanese*, ms., University of Connecticut, Storrs.
- Miyamoto, Y. (1998) "Floating Quantifiers in Japanese," Paper presented at Kyoto University (January 1998) and Yokohama National University (November 1998).
- Mulford, R. (1985) "Comprehension of Icelandic Pronoun Gender: Semantic versus Formal Factors," *Journal of Child Language* 12, 443–453.
- Muraishi, Shozo (1983) "Josuushi Tesuto [Classifier Tests]," in Kokuritsu Kokugo Kenkyuusho [The National Language Research Institute], *Yooji/Jidoo no Gengokeisei to Gengo* [Conceptual Development and Language in Children], Tokyo Shoseki, Tokyo, pp. 273–316.

- Murata, Koji (1983) *Kodomo no Kotoba to Kyoiku* [Children's Minds and Education], Kaneko Shobo, Tokyo.
- Okubo, Ai (1967) *Yoojigengo no Hattatsu* [Development of Child Language], Tokyodo, Tokyo.
- Perez-Pereira, Miguel (1991) "The Acquisition of Gender: What Spanish Children Tell Us," *Journal of Child Language* 18, 571–590.
- Rosch, E., C. B. Mervis, W. D. Gray, D. M. Johnson, and P. Boyes-Braem (1976) "Basic Objects in Natural Categories," *Cognitive Psychology* 8, 382–439.
- Sanches, Mary (1977) "Language Acquisition and Language Change: Japanese Numeral Classifiers," in M. Snaches and B. Blount (eds.), *Sociocultural Dimensions of Language Change*, Academic Press, NY, pp. 51–62.
- Simons, D. S., and F. C. Keil (1995) "An Abstract to Concrete Shift in Cognitive Development: The Inside Story," *Cognition* 56, 129–163.
- Soja, Nancy, Susan Carey, and Elizabeth Spelke (1991) "Ontological Categories Guide Young Children's Inductions of Word Meaning: Object Terms and Substance Terms," *Cognition* 38, 179–211.
- Uchida, Nobuko, and Mutsumi Imai (1996) "A Study on the Acquisition of Numeral Classifiers Among Young Children: The Development of Human-animal Categories and Generation of the Rule of Classifiers Applying," *Japanese Journal of Educational Psychology* 44, 2, 126–135.
- Wellman, H., and S. Gelman (1988) "Children's Understanding of the Nonobivious," in R. J. Sternberg (ed.), *Advances in the Psychology of Human Intelligence*, 4, Erlbaum, Hillsdale, NJ, pp. 99–135.

Received: June 30, 1998
 Revised: April 11, 2000

Kasumi Yamamoto
 Department of Asian Studies
 Williams College
 Stetson Hall
 Williamstown, MA 01267
 E-mail: Kasumi.Yamamoto@williams.edu

Frank Keil
 Department of Psychology
 Yale University
 New Haven, CT 06520
 E-mail: Frank.Keil@yale.edu