ELSEVIER

COGNITION

Cognition 80 (2001) B11–B20

www.elsevier.com/locate/cognit

Brief article

Linguistic and cognitive abilities in infancy: when does language become a tool for categorization?

Thierry Nazzi[a,*], Alison Gopnik[b]

[a]*Laboraroite Cognition et Développement, Centre Universitaire de Boulogne, Institut de Psychologie - Université Paris 5, 71, Avenue Edouard Vaillant, 92774 Boulogne Billancourt Cedex, France*
[b]*Institute of Human Development and Department of Psychology, University of California at Berkeley, Berkeley, CA, USA*

Received 28 July 2000; received in revised form 25 November 2000; accepted 20 December 2000

Abstract

Infants' ability to form new object categories based on either visual or naming information alone was evaluated at two different ages (16 and 20 months) using an object manipulation task. Estimates of productive vocabulary size were also collected. Infants at both ages showed evidence of using visual information to categorize the objects, while only the older ones used naming information. Moreover, there was a correlation between vocabulary size and name-based categorization among the 20-month-olds. The present results establish that infants as young as 20 months can use the non-obvious cue of naming to categorize objects. The possibility of a link between this ability and lexical development is discussed. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Categorization; Naming; Word-learning

1. Introduction

There are a number of cues that infants and young children could use to categorize objects. On one hand, many studies have established that obvious (i.e. easy to observe and stable in time) cues such as color and shape are used for categorization in infancy and early childhood (see Quinn & Eimas, 1996, for a review). On the other hand,

* Corresponding author. Fax: +33-1-55-20-59-85.
E-mail address: thierry.nazzi@psycho.univ-paris5.fr (T. Nazzi).

PII: S0010-0277(01)00112-3

several studies have shown that children from at least 2–3 years of age can use different non-obvious (i.e. more transiently observable and/or more arbitrary) cues to categorize objects and make inferences about their properties. These non-obvious properties include names (Gelman & Coley, 1990; Gelman & Markman, 1986, 1987; Gopnik & Sobel, 2000), functional attributes (Kemler Nelson, 1995; Kemler Nelson, Frankenfield, Morris, & Blair, 2000; Kemler Nelson, Russell, Duke, & Jones, 2000), and causal properties (Gopnik & Sobel, 2000; Nazzi & Gopnik, 2000). In particular, children as young as 30 months old were found to use a common name to categorize objects even when those objects were completely dissimilar at the perceptual level (Gelman & Coley, 1990; Gelman & Markman, 1986, 1987; Gopnik & Sobel, 2000). There is, however, little evidence showing that infants less than 2 years old can readily use non-obvious properties, and particularly names, to form new categories of objects. The present study explores infants' use of naming to categorize objects, or more precisely, their use of naming to form new categories of objects.

Only a few experimental studies have investigated the effect of naming on infants' categorization behavior. Such an effect was first suggested by a study establishing that naming increases 10–14-month-old infants' attention to objects (Baldwin & Markman, 1989). Building on this initial result, Waxman and Markow (1995) showed that naming actually enhances object categorization by 13-month-old infants. In their study, infants were sequentially presented with different exemplars of a category of objects, at either the basic level (cars, airplanes, cows and dinosaurs) or the superordinate level (animals, vehicles and tools). To determine whether naming influences categorization, the objects were either labeled or not during their presentation. A comparison of the infants' behavior in the label and no-label conditions revealed that naming focuses infants' attention on object categories. Infants who heard a label habituated more quickly to the category, and showed a better recognition of new exemplars of that category.

The results at both category levels, but especially at the superordinate level, for which object similarity is lowest, suggest that 13-month-olds use naming to categorize objects. However, the superordinate categories that were tested were probably already available to the infants. Even young infants will spontaneously produce such categories (see Mandler & McDonough, 1993, 1996). Although there was no direct evidence of spontaneous superordinate categorization in the Waxman and Markow (1995) study, that is probably because infants were not given enough time to habituate, a possibility suggested by Waxman and Markow themselves. Hence, Waxman and Markow's finding might show that names make it easier for infants to display their knowledge of *known* object categories, rather than showing that infants are capable of forming completely *new* categories based on naming information. This interpretation also applies to the Waxman and Hall (1993) study, in which labeling increased the likelihood that infants would form taxonomic categories rather than thematic categories, an effect that also appeared to be more robust at 21 months than at 15 months. This interpretation is further supported by the result of Waxman and Markow's second experiment. They found no effect of naming on infants' behavior during the habituation phase when arbitrary, that is, entirely new, categories were presented.

Hence, categorization studies fail to establish whether infants can form new object categories based on naming information alone.

This question is not solved either by recent studies on the related issue of fast-mapping, defined as the ability to learn links between sound patterns and concepts after only a few repetitions of their association (Hollich, Hirsh-Pasek, & Golinkoff, 2000; Schafer & Plunkett, 1998; Werker, Cohen, Lloyd, Casacola, & Stager, 1998; Woodward, Markman, & Fitzsimmons, 1994). These studies show that infants as young as 12–15 months can establish the link between a sound pattern and an object after hearing as little as six to nine repetitions of the sound pattern. But these studies do not provide strong evidence that the infants might be linking sound patterns to object categories as opposed to particular objects because the test for label generalization was limited to objects differing from the original objects in color only (Woodward et al., 1994 for 13- and 18-month-olds; Mervis & Bertrand, 1994, and Ross, Nelson, Wetstone, & Tanouye, 1986, for 20-month-olds).

In the present study, we investigate infants' ability to use naming cues to form completely new object categories in the absence of any other cue at two different ages: 16 and 20 months. In order to assess these name-based categorizations, infants were presented with triads of unfamiliar, unrelated, and perceptually dissimilar objects, two of which received the same label, and the third a different label. Each label was repeated six times, as in Werker et al. (1998). The goal was to determine whether infants would pair the two objects that had received the same label, a behavior that would signal the use of naming to form new object categories. To ensure that infants understood the task of pairing the objects, the naming trials were preceded by visual trials in which two of the objects in the triad were identical.

In addition to the categorization data, we collected estimates of the infants' productive vocabulary using the Communicative Development Inventory (CDI) questionnaire (Fenson et al., 1993). These vocabulary measures were used to evaluate the level of lexical development of the infants, and to examine whether it might be related to their categorization. Earlier studies have suggested that there might be a specific relation between vocabulary size and manual sorting. In particular, children who use more names are more likely to exhaustively sort objects into different categories (Gopnik & Meltzoff, 1987, 1992; Mervis & Bertrand, 1994). There might be a similar relation between vocabulary size and the ability to use names to form categories.

2. Method

2.1. Participants

Twenty-four 16-month-old infants ($M = 16$ months, 6 days; range 15 months, 24 days to 16 months, 25 days) and 24 20-month-old infants ($M = 20$ months, 2 days; range 19 months, 14 days to 20 months, 19 days) from monolingual English-speaking American families were recruited through the infant listing of the Institute of Human Development, University of California at Berkeley. An equal number of boys and girls of each age participated in the study. Most infants came from white, middle-class

backgrounds, although infants from other ethnic backgrounds were also represented. Three additional infants of each age were tested, but failed to complete the session.

2.2. Materials

Six triads of small objects were used during the testing session. All objects were selected so that the infants would be unfamiliar with them and would not have a name for them. Three of the sets were made up of a pair of identical objects and a third object that differed from the other two in shape, color, and texture (e.g. two identical gray, plastic, tubes vs. one red, play-doh, disk). The remaining three sets were made up of three very distinct objects, that all differed in shape, color, and texture in an effort to equalize their perceptual distance (e.g. silvery, metal, door-knob vs. yellow, wood, trapezoid vs. gray, plastic, bar).

2.3. Procedure

Each infant was tested individually for 20 min in a quiet room of the Institute. The infant was seated on a parent's lap across a table from the experimenter. The parent was instructed not to talk, and not to influence the infant's interest in, and manipulation of, the objects. The session started with a 5 min warm-up period, during which the experimenter showed the infant plastic rings, made them spin, and encouraged the infant both to play with them and to give them back to him.

The testing session was made up of two blocks of three different trials. The object triads made up of a pair of identical objects and a third different object were used in the first three "visual" trials, while the triads of three different objects were used in the last three "naming" trials.

Each trial started with a period during which the experimenter presented the infant with the three objects, one at a time. The infant was encouraged to manipulate each object for a few seconds, before the object was placed on the table. Within each trial, the objects were arranged on the table on a left-to-right sequence (from the infant's perspective) determined by their order of presentation, in order to facilitate infants' memory of the link between the labels and the objects in the naming trials. The experimenter spoke to the infant during the presentation of each object, saying for the [visual/naming] trials: "Look! [Look at this one/A 'tib'. This is a 'tib']. Do you want to play with [this one/the 'tib']? Yes, play with [this one/the 'tib']. All right! Let's put [this one/the 'tib'] on the table. Here." On the naming trials, each object was named exactly six times. Moreover, two of the objects were labeled 'tib' and the last one 'dap' (or vice versa).

During the second period of each trial, the experimenter put one object of the visual or named pair in his hand placed at equal distance from the remaining two

[1] Although the infants might not understand the literal meaning of the question, the task was designed in such a way that the infants would be able to infer that they were expected to provide the matching object. Moreover, the main reason to always present the visual phase first was to facilitate the infants' understanding of the question by performing a visual task that previous research had shown they were able to perform.

objects, and asked the infant to give him "the object that goes with this one".[1] While waiting for the infant's response, the experimenter was looking either at the infant's face or at the object in his hand (usually going back and forth), in order to avoid biasing the infant toward one of the two objects. The parent was instructed not to influence the infant in any way. After the infant had chosen one of the objects, positive feedback was provided regardless of the choice made. The order of presentation of the trials within each block, the position of the paired objects on the table, and the side of the object picked up by the experimenter were counterbalanced across subjects. Moreover, in addition to the precautions against potential perceptual biases used in selecting the objects themselves, the pairs defined by the names in the naming trials were counterbalanced across subjects.

Finally, before the testing session, the parent filled out the vocabulary part of the MacArthur Communicative Development Inventory: Toddlers (CDI; Fenson et al., 1993) in order to determine the size of the infants' productive vocabulary.

3. Results

3.1. Vocabulary measures

The analysis of the CDI results showed that the 16-month-olds had a mean productive vocabulary of 28 words (SD = 25; range 4–80), and the 20-month-olds produced a mean of 141 words (SD = 110; range 17–397). Two-tailed *t*-tests revealed that the 20-month-olds had significantly larger vocabularies than the 16-month-olds ($t(23) = 4.71$, $P < 0.001$).

3.2. Categorization measures

The results of the visual and naming conditions were analyzed separately, given that the visual condition was used as a way of training the infants on the procedure (and was hence always presented first). However, for both conditions, the responses were coded in the same way: for each trial, infants were given a score of 1 when the object chosen was the second of the (visual or named) pair, and a score of 0 when they chose the other object. The results were coded twice, first during the session itself, then by an assistant blind to the experimental hypotheses, leading to no discrepancies. Total scores per infant and condition ranged from 0 to 3. Mean total scores per condition and age are given in Table 1.

3.2.1. Visual condition

The total scores were first submitted to a 2×2 ANOVA with the main between-subject factors of age (16 vs. 20 months) and sex. These analyses revealed no significant effects (age: $F(1, 22) = 1.49$, $P = 0.23$; sex: $F(1, 22) < 1$), and no interactions ($F(1, 22) = 1.49$, $P = 0.23$), indicating that infants behaved similarly in all conditions. Further analyses were conducted to determine whether infants were choosing the second object of the visual pair significantly more than the 1.5 chance score (given three forced-choice trials). Results showed that it was the case for both

Table 1
Mean numbers of correct responses (and SD) in the visual and naming conditions according to age[a]

	Visual	Naming
16-month-olds	2.29 (0.75)**	1.42 (0.72)
20-month-olds	2.54 (0.66)**	1.88 (0.61)*

[a] Two-tailed *t*-tests: $*P < 0.01$, $**P < 0.001$.

the 16-month-olds ($M = 2.29$, $t(23) = 5.17$, $P < 0.001$) and the 20-month-olds ($M = 2.54$, $t(23) = 7.76$, $P < 0.001$).

We also conducted χ^2 analyses to assess that the results from the ANOVA did not result from a few outliers, but that the distribution of infants' performance (Table 2) was different from that predicted by chance. Given the validity constraints on the use of χ^2 analyses, the infants were assigned to two categories: total score of 0 or 1 vs. total score of 2 or 3. If the 24 infants within each age group were performing at chance, 12 infants should fall into each of these two categories. The χ^2 tests revealed that the distributions of both the 16- and the 20-month-olds differed from chance ($\chi^2(1, N = 24) = 10.7$, $P = 0.001$, and $\chi^2(1, N = 24) = 16.7$, $P < 0.001$, respectively).

3.2.2. Naming condition

The results in the naming condition were analyzed as in the visual condition. An ANOVA with the main between-subject factors of age and sex revealed a significant effect of age ($F(1, 22) = 5.48$, $P = 0.024$), indicating that 20-month-old infants were choosing the second object in the naming pair more often than 16-month-old infants. On the other hand, neither the main effect of sex nor its interaction with age reached significance (both $F(1, 22) < 1$), indicating that boys and girls performed similarly. Infants' scores were then tested against the chance score of 1.5. These analyses revealed that 16-month-old infants were responding at chance ($M = 1.42$, $\mathrm{SD} = 0.72$, $t(23) = -0.57$, $P = 0.58$), while 20-month-old infants were responding above chance level ($M = 1.88$, $\mathrm{SD} = 0.61$, $t(23) = 3.00$, $P = 0.006$), confirming the age effect found in the ANOVA.

Table 2
Distribution of the infants in both conditions according to performance

	Total score			
	0	1	2	3
Visual				
16-month-olds	0	4	9	11
20-month-olds	0	2	7	15
Naming				
16-month-olds	2	11	10	1
20-month-olds	0	6	15	3

Table 3
Pearson correlation coefficients (and significance level) between visual performance, naming performance and productive vocabulary size at 16 and 20 months

	16-month-olds			20-month-olds		
	Visual	Naming	Words	Visual	Naming	Words
Visual						
Naming	0.25 (0.24)			− 0.36 (0.08)		
Words	− 0.01 (0.96)	− 0.46 (0.02)		− 0.24 (0.26)	0.64 (0.001)	

Finally, the distribution of infants' performance (see Table 2, lower panel) was tested against chance after collapsing the results into two categories (total score of 0 or 1 vs. 2 or 3). These analyses revealed that the distribution of the 16-month-olds was not different from chance ($\chi^2(1, N = 24) = 0.2, P = 0.68$), while that of the 20-month-olds differed from chance ($\chi^2(1, N = 24) = 6.0,\ P = 0.014$). Again, these results confirm those obtained by the ANOVA.

3.3. Correlation analyses

Finally, Pearson correlation analyses were performed for both age groups between the visual condition score, the naming condition score, and the productive vocabulary size (Table 3). At both ages, the only significant correlation involved naming performance and productive vocabulary size. However, the interpretation of these two correlations differs. At 16 months there was a small negative correlation between vocabulary size and performance on the naming condition. However, given that 16-month-olds were performing at chance on the naming condition as shown by the ANOVA, this correlation is most likely to be an artifact. At 20 months, both variables were strongly and positively correlated ($r = 0.64$, $P = 0.001$). Given the ANOVA showing that 20-month-olds were not performing at chance on the naming trials, this correlation suggests that naming performance and productive vocabulary size increase in a related way. Finally, it is important to note the absence of a correlation between visual and naming performance at both ages. Combined with the ANOVA results, this supports the idea that the increase in name-based categorization is not simply the result of a general improvement of infants' categorization ability.[2]

[2] Note that some of the correlation results support the idea that the precautions taken to prevent the parent, who was aware of the correct response the infant should give, from influencing the infant in spite of being asked not to were efficient (the experimenter was also aware of the correct answer, but looked at the infant's face or at his own hand while waiting for the infant's choice). First, the absence of a correlation between performance on the visual- and name-based tasks at both ages weakens the idea that the parents were indicating the correct answer to their infants in some way. Second, the positive correlation between name-based categorization and vocabulary size at 20 months couldn't be accounted for by an artifactual bias, as the parent wasn't aware of this hypothesis, and the experimenter didn't know the size of the infant's vocabulary before the testing session.

4. Discussion

The present results contribute new data regarding the kind of cues on which infants can base their object categorizations. They first confirm that both 16- and 20-month-old infants use obvious visual cues to group objects together, as already established by many studies (Quinn & Eimas, 1996). More importantly, the present results show that 20-month-old infants can use names to form completely new object categories, categories which combine objects without obvious perceptual similarities. Hence, the present study establishes that infants as young as 20 months can use a non-obvious cue, naming, to form a new object category. This investigation extends to a substantially younger age previous research establishing that 2.5–5-year-old children use non-obvious cues such as naming (Gelman & Coley, 1990; Gelman & Markman, 1986, 1987; Gopnik & Sobel, 2000), functional (Kemler Nelson, 1995; Kemler Nelson et al., 2000a,b) and causal properties (Gopnik & Sobel, 2000; Nazzi & Gopnik, 2000) to categorize objects. In particular, it shows that naming is already used as an important cue to categorization at an early stage of the child's language development.

The 16-month-olds showed above chance level performance on the visual-based categorization task, but chance level performance on the name-based categorization task. This negative result is more difficult to interpret. One interpretation is that 16-month-olds cannot use naming to form new object categories, and that this ability emerges between 16 and 20 months. However, there are other possible explanations for the 16-month-olds' difficulties. Some simple accounts seem unlikely given the 16-month-olds' good performance on the visual task. For example, it seems unlikely that the 16-month-olds simply paid less attention to the task, were less compliant, or had more difficulty with the sorting response, given their success on the very similar visual task. Other more complex explanations are more plausible. The 16-month-olds might have had difficulties remembering the new words, and the pairings between the words and the three objects, and maybe they would have succeeded if presented with more repetitions of the label. This would differentiate the naming task from the visual task. And although previous studies found that 13–15-month-olds were able to learn two word–object pairings when presented with the pairings a similar number of times (Schafer & Plunkett, 1998; Werker et al., 1998), these studies did not also involve categorization. Similarly, the infants may have had difficulty combining the demands of the basic categorization task with the demands of name learning. There may have been an interaction between the demands of learning and remembering the names, demands that were not present in the visual task, and the demands of the categorization task. While children may have been able to either remember the names or categorize on their own, they may not have been able to combine these two capacities. All these issues will have to be addressed in future research, in order to interpret the developmental pattern found here.

Finally, we discuss the possibility of a link between the development of name-based categorization and increases in productive vocabulary. Some features of our results are compatible with the idea of such a link. First, there is the evidence for an improvement, and perhaps even the emergence, of name-based categorization

between 16 and 20 months. Second, the vocabulary measures suggest changes in lexical ability over the same period. Finally, and most significantly, the correlation analyses show that name-based categorization is related to the size of the productive vocabulary at 20 months. Moreover, there was no evidence for a relation between vocabulary size and visual-based categorization: children at both ages performed equally well, and there were no correlations in either age group. Thus, the correlation between name-based categorization and vocabulary size seems to be quite specific, rather than reflecting, for example, a more general common factor of attention or intelligence. Although the above elements suggest that there may be a link between name-based categorization and lexical development, they cannot be taken as conclusive evidence for such a link, particularly given this single correlational finding. These data also cannot determine the direction of such a link (whether the acquisition of names leads to new categorization abilities or vice versa, or whether the two areas of development proceed in tandem).

To recapitulate, the main finding of the present study is that infants as young as 20 months use the non-obvious cue of naming to categorize objects. The study also suggests that this ability may be absent in 16-month-old infants, and that it may be linked to vocabulary size.

Acknowledgements

This research was supported by a grant from the Fondation de la Recherche Médicale to the first author. We thank Gerri Jansen and Amber Mills for help in recruiting the subjects, Wayne Ma for assistance in analyzing the data, Josiane Bertoncini, Roberta Golinkoff, Michèle Kail, Dan Slobin and Janet Werker for helpful discussions at various stages of this project, and Ruth Tincoff for useful comments on a previous version of the paper.

References

Baldwin, D. A., & Markman, E. M. (1989). Establishing word-object relations: a first step. *Child Development*, *60*, 381–398.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S., & Reilly, J. S. (1993). *Technical manual for the MacArthur Communicative Development Inventory*. San Diego, CA: Singular Press.

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: categories and inferences in 2-year-old children. *Developmental Psychology*, *26*, 796–804.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*, 183–209.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: the role of categories and appearances. *Child Development*, *58*, 1532–1541.

Gopnik, A., & Meltzoff, A. N. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development*, *58*, 1523–1531.

Gopnik, A., & Meltzoff, A. N. (1992). Categorization and naming: basic-level sorting in eighteen-month-olds and its relationship to language. *Child Development*, *63*, 1091–1103.

Gopnik, A., & Sobel, D. (2000). Detecting blickets: how young children use information about novel causal powers in categorization and induction. *Child Development*, *71*, 1205–1222.

Hollich, G. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: an emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development, 65* (3, Serial No. 262).

Kemler Nelson, D. G. (1995). Principle-based inferences in young children's categorization: revisiting the impact of function on the naming of artifacts. *Cognitive Development*, *10*, 347–380.

Kemler Nelson, D. G., Frankenfield, A., Morris, C., & Blair, E. (2000a). Young children's use of functional information to categorize artifacts: three factors that matter. *Cognition*, *77*, 133–168.

Kemler Nelson, D. G., Russell, R., Duke, N., & Jones, K. (2000b). Two-yead olds will name artifacts by their functions. *Child Development, 71*, 1271–1288.

Mandler, J. M., & McDonough, L. (1993). Concept formation in infancy. *Cognitive Development*, *8*, 291–318.

Mandler, J. M., & McDonough, L. (1996). Drinking and driving don't mix: inductive generalization in infancy. *Cognition*, *59*, 307–335.

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name-nameless category (N3C) principle. *Child Development*, *65*, 1646–1662.

Nazzi, T., & Gopnik, A. (2000). A shift in children's use of perceptual and causal cues to categorization. *Developmental Science*, *3*, 389–396.

Quinn, P. C., & Eimas, P. D. (1996). Perceptual organization and categorization in young infants. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in infancy research*, 11. Norwood, NJ: Ablex.

Ross, G., Nelson, K., Wetstone, H., & Tanouye, E. (1986). Acquisition and generalization of novel object concepts by young language learners. *Journal of Child Language*, *13*, 67–83.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development*, *69*, 309–320.

Waxman, S. R., & Hall, D. G. (1993). The development of a linkage between count nouns and object categories: evidence from fifteen- to twenty-one-month-old infants. *Child Development*, *64*, 1224–1241.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*, 257–302.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casacola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, *34*, 1289–1309.

Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*, *30*, 553–566.