J. Child Lang. **29** (2002), 357–377. © 2002 Cambridge University Press
DOI: 10.1017/S030500090200507X Printed in the United Kingdom

NOTE

Basic-level nouns: first learned but misunderstood*

LARAINE McDONOUGH
Brooklyn College, The City University of New York Graduate Center

(*Received* 26 *July* 1999. *Revised* 26 *February* 2001)

ABSTRACT

It is commonly believed that first-learned words correspond with first-learned categories (both described as 'basic level') leading to the belief that language acquisition is a reasonably good indicator of early cognition. However, toddlers often overextend their first words. Do these errors reflect their comprehension? Two experiments were conducted in order to examine two-year-olds' production and comprehension of basic-level terms. The results showed overextensions both in production (e.g. children labelled a rocket 'airplane') and comprehension (e.g. they pointed to a rocket when airplane was requested). One reason toddlers extend labels to a wider conceptual domain is because they have not clearly differentiated basic-level concepts from related conceptual categories.

INTRODUCTION

The goal of the present research is to uncover what young children understand about the meanings of the nouns they learn. It is traditionally thought that children's first words reflect early-learned meanings/concepts. After all, language does have a purpose. It is a rather elegant tool one can use to engage in social activities and share ideas or perhaps coerce others to share their toys. However, a review of the literature on early children's extensions of the words they learn would have us believe that children extend words based on overall shape similarity or other perceptual cues among objects. At

[*] Preparation of this article was supported by NSF Research Grants SBR-970895, SBR 9973399. I would like to thank Lisa Hoag, Laura Holmes, Shannon Prudden, and Pamela Smith for assisting with data collection and coding. Address for correspondence: Laraine McDonough, Department of Psychology, Brooklyn College, Brooklyn, NY 11210, USA. e-mail: larainem@brooklyn.cuny.edu

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

some point in development, similarity reflects not simply perceptual cues (such as those used by 0;3 in habituation/preferential looking tasks) but also an understanding of kind.

The challenge to figuring out the meanings children hold for the words they use comes from a blending of two theories, one derived from language acquisition, in which the focus has been on the meanings of first words; the other concerns categorization of objects, in which the traditional focus has been on the similarity among objects. This blending was made explicit by several researchers who have investigated what is known as the taxonomic assumption (e.g. Markman, 1989). The taxonomic assumption is thought to be a learning constraint or bias that children use when learning language. That is, children are thought to assume that nouns refer to a category of objects. Given that basic-level terms are used almost exclusively in child-directed conversations, it also makes sense that children's first words are largely comprised of basic-level nouns. Thus, one might think that the earliest meanings children have for objects are those best described as basic-level meanings/concepts. Indeed, our traditional views hold that first learned categories are basic-level categories (e.g. Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976).

However, evidence for the taxonomic assumption has not been based on research using the words children learn; instead, it has been tested using nonsense or nonce terms. Herein is where the uncertainty lies. The results of research with infants younger than 2;0 show that the extensions of nonce terms are not limited to the boundaries of basic-level categories but instead are extended to superordinate categories. The kind of task used in this research is typically administered by showing children one object (e.g. a toy dog), labelling it using a nonsense label (e.g. 'a fep') and asking them to find another object using the same label (e.g. 'Can you find another fep?'). One might expect that when children are given basic-level choices such as another dog, their performance would be significantly better than when given superordinate choices, such as other kinds of animals, but this has not been found (e.g. Bauer & Mandler, 1989). It would appear that the boundaries they use for their extensions are more accurately described as those distinguishing superordinate categories.

If children extend novel nouns to both basic-level and superordinate exemplars, then which boundaries do they use to constrain their extensions of the words typically used to refer to these objects? The evidence cited above shows that on hearing a dog labelled 'fep' children would as readily extend that label to other animals as to other dogs. One unresolved issue in this research concerns a 'translation hypothesis.' Just because a nonce term is presented along with a previously unnamed object does not mean that the child has no idea what to call the object in the first place. On hearing an item labelled 'fep' for which the child perhaps already has a label (i.e. 'dog'), does

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

the child simply translate the term fep to mean dog? According to the interpretation that 'fep' is readily extended to include other animals (not just dogs), and the assumption that 'dog' is only extended to include dogs, then fep is not simply translated to mean 'dog' but something more like animal, a term that does not enter the child's lexicon until much later (Waxman, Senghas & Benveniste, 1997). However, it remains possible that the meaning young children initially hold for the term dog is also more widely extended to other animals as well. In light of other research showing that young children (before the vocabulary spurt) are apparently comfortable having two labels for the same object (e.g. Mervis, 1987), the term 'fep' may in fact translate to mean 'dog.'

What about extensions of real words, words young children already understand? If children are treating the nonsense words in the same way they treat real words, then one would expect that early-learned nouns would also be extended throughout superordinate categories. Indeed, evidence for overextensions of real words has been reported by several researchers. For example, Rescorla (1980) found that overextensions in young children's productive vocabularies are most often made with early-acquired words and are made within superordinate boundaries.

Yet, overextensions in production can occur for several reasons. One is that young children have retrieval or memory difficulties. For example, a child sees a skunk in the yard and says 'kitty-cat.' The child may not really think that this striped creature with a large fluffy tail and short legs is a cat, but the term 'skunk' is not a common noun and, as such, retrieval of it may be difficult. Alternatively, the child may not know the name of this unusual animal, but notes that it somewhat resembles a cat, so calls it a cat for communication purposes. It is not that the child really thinks the skunk is a cat, but the word is simply missing from her/his lexicon. These production problems are based on the assumption that the child knows that cats and skunks are not the same kind of thing, but belong to two contrasting basic-level categories.

However, it is possible the child really thinks that a skunk is the same kind of thing as a cat and considers them both as belonging to the same category. To answer this question, one needs to evaluate the child's comprehension of basic-level terms. Previous research on comprehension of basic-level nouns has used tasks that require children to look longer or point to referents of basic-level nouns with either inappropriate contrasts (e.g. Reznick & Goldfield, 1992) or anomalous trials in which the correct referents were not available (Behrend, 1987; Naigles & Gelman, 1995). For example, in a study of lexical development in which comprehension of 75 terms was assessed, Reznick & Goldfield contrasted the correct referent with a distractor from the same superordinate category on only 3 occasions. In their examination of whether or not children know the meaning of words such as 'shoe' (a

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

clothing item) or 'bee' (an insect/animal), the correct referents were contrasted with 'banana' (a food item) or 'motorcycle' (a vehicle). Such contrasts do not tell us whether or not basic-level categories provide the boundaries for the early meanings of basic-level terms because the contrasts cross superordinate category boundaries (e.g. contrasting a bee and a banana and a motorcycle is a contrast between an insect, a food item and a vehicle). Appropriate contrasts should be selected from the same superordinate category. For example, to test whether or not a child comprehends the word 'bee', the correct referent should be contrasted with another insect or animal such as a 'butterfly.' Otherwise, it is unknown how broadly the term will be extended within a given domain.

Naigles & Gelman (1995) tested overextensions using trials in which the referent was not available (i.e. anomalous trials). They used a preferential looking paradigm to examine comprehension of three animal terms: dog, cat, and cow. On hearing the instructions to look at the cat during anomalous trials (i.e. the dog and cow were shown but no cat was available), participants showed no looking preferences: looking to the dog and cow was the same. However, on the anomalous trials in which children were shown a dog and cat and were directed to find a cow, they showed a preference to look at the dog. Naigles & Gelman explained this finding in terms of similarity but not in terms of categorical immaturity (as proposed by Mervis, 1987). But given that it is typically thought that early categories are based on perceptual similarity, it remains an open question as to whether or not such an overextension is similarity- or category-based because both bases should theoretically lead to the same prediction. After all, the defining characteristics proposed of basic-level categories was similarity (e.g. Rosch *et al.*, 1976). It is possible that infants knew that the correct referent was not available and simply looked at the next best exemplar when one was available (e.g. Kuczaj, 1982) or that they thought that the cow was a kind of dog.

Thus, the previous research is ambiguous in terms of young children's comprehension of nouns. The goal of the present research is to examine more extensively children's overextensions of basic-level nouns in both production and comprehension. In the following two experiments, anomalous trials were not used: overextension was tested in the presence of correct referents, which were selected from the animal, vehicle, clothing and food domains. Distractor items were drawn from either the same domain as the target (a low contrast condition) or from a contrasting domain (a high contrast condition). Triad tasks were used in which three pictures were presented together on a single page for each trial, and the child was asked to label or point to one of the pictures. Labels tested were those which are typically produced early in development (i.e. early items tested in Experiment 1) and those which are not produced until later in development (i.e. late items tested in Experiment 2). Based on Rescorla's (1980) research, it was predicted that children would

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

overextend the labels for the early items to the late items. If children's production errors are comprehension errors as well, then overextensions in comprehension would mirror the production data, particularly in the low contrast trials in which an item from the same superordinate category was available. However, if production errors are due to retrieval problems then overextensions would not be found in the comprehension data. If errors in production are made because children did not know the labels for the late items, then they may make more correct selections in the comprehension task by inferring that the unknown word refers to the item for which they do not have a label.

A control condition was conducted in Experiment 2 to evaluate whether or not children actually comprehend the late nouns or use a 'fast mapping' strategy (Carey, 1982) by putting together the noun they did not comprehend with the picture for which they did not know the label. Children were asked to point to the referent of a nonsense term (e.g. 'Point to the fep.'). It was predicted that they would choose the late item if they did not know the correct label for it.

EXPERIMENT 1

PARTICIPANTS

Twenty participants (12 females and 8 males) aged 1;10 to 2;0 (mean = 1;11) were tested. Participants were recruited through advertisements in local newspapers. All were learning English as their first language. The MacArthur Communicative Development Inventory for toddlers aged 1;4 to 2;6 (Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994) was filled out by parents indicating their children's productive vocabularies. (Note that the comprehension vocabularies of children this age are generally too large and therefore are difficult to obtain using a standardized assessment scale.) Sixteen inventories were filled out before testing took place and the remaining 4 were filled out after testing. The mean number of words produced were 360 with a standard deviation of 184 words. A comparison of each child's vocabulary was made to the norms reported in Fenson *et al*. One child fell in the lower 10th percentile, 2 children in the lower 25th percentile, 8 children were at the median point, and 9 children were at or above the 90th percentile for children of this age. Thus, most of the participants in this study fit within the normal to upper range of productive vocabulary in comparison with other children their age. Given that one of the purposes of the study was to examine comprehension, none of the children was excluded from the final analyses.

At the time this study was conducted, no records were kept as to ethnicity or SES because such requirements were not in effect. However, the participant pool is representative of the ethnic background of the local San

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

Diego neighborhoods (75 % white, 6 % black, 8 % Asian, 10 % Hispanic, and 1 % American Indian or other). Although participants were not formally asked for this information, a review of the tapes indicates (as near as possible) a similar distribution. One additional toddler was tested but is not included in the analyses due to an equipment problem with the video recorder. A small toy was given to each for their participation.

TEST STIMULI

Sixteen cards were constructed, each with three realistically coloured pictures of objects from the animal, vehicle, food and clothing domains. Eight objects were selected as test items within each domain. Four of the 8 objects in each domain were chosen because they had labels that are among the earliest comprehended and produced by toddlers (i.e. typically comprehended around 1;3 and produced at 1;8; Fenson *et al.*, 1994). These are referred to as 'early items'. The remaining 4 objects had labels that are typically learned somewhat later in development (i.e. 'late items'). It was expected that most children aged 2;0 would not know many of the labels for the late items. This expectation was based on the research by Fenson *et al.*, (1994) and the fact that these items are low frequency items in children's books (Thorndike & Lorge, 1944). All but one of the pictures (early and late) were selected from *Pictures Please*, a language book used to test children's language production and comprehension (Abbate & LaChappelle, 1979). Each of the pictures was coloured in by hand using coloured pencils. The colours were selected by how natural they were for each pictured object.

A complete listing of the pictures and how they were grouped onto the sixteen test cards is shown in Table 1. The first page in Booklet A on which is displayed a pig, train and bus will be used as an example to describe how the cards were constructed. Two of the 3 pictures on each card were selected from the same domain (e.g. train and bus), one of which was an early item (i.e. train) and the other a late item (i.e. bus). The third object, also an early item, was from a different domain (i.e. pig). Participants in this experiment were asked on different trials to point to each of the two early items. For example, on one trial a subject would be asked to point to the train and later in the experiment would be asked to point to the pig (or vice versa). Trials in which participants were asked to label or point to the target that was contrasted with two items from another domain (i.e. the pig was requested) were called high-contrast trials. Trials in which participants were asked to label or point to the target that was paired with the distractor from the same domain (i.e. the train was requested) were called low-contrast trials.

As can be seen in Table 1, the early items listed in the far left column were presented in a high contrast context (e.g. the pig is paired with two vehicles); in the next column are early items that are paired in a low contrast context

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

TABLE 1. *Test booklets used in Experiments 1 & 2*

Early noun	Early noun	Late noun
(High contrast) Comprehension tested in Experiments 1 & 2	(Low contrast) Comprehension tested in Experiment 1	(Low contrast) Comprehension tested in Experiment 2
Booklet A		
pig	train	bus
cow	pants	shorts
carrot	shoe	sandal
orange	bicycle	motorcycle
dog	car	truck
apple	shirt	vest
cat	dress	sweater
egg	airplane	rocket
Booklet B		
shirt	pig	hippo
bicycle	cow	moose
boat	carrot	celery
pants	orange	beet
dress	dog	fox
car	apple	strawberry
train	cat	raccoon
airplane	cake	pie

(e.g. train and bus are both vehicles). Placement of the target items on each card was balanced for left, right and centre positions. The sixteen cards were divided into two booklets (booklet A and B). In Booklet A, the high contrast items were animals and food and the low contrast items were vehicles and clothing. In Booklet B, the high contrast targets were vehicles and clothing and the low contrast targets were animals and food. Half the participants were tested using book A and then book B. The remaining half were tested in the opposite order. High and low contrast trials were interleaved; e.g. one participant received a high contrast query on page 1 and a low contrast trial on page 2 of book A (and so on until both books were completed) whereas another received a low contrast query on page 1 and a high contrast query on page 2 of book A (and so on until both books were completed). Booklet order and contrast trial order were counterbalanced among participants.

PROCEDURE

Participants were invited into a laboratory set up as a playroom. After a brief warm-up period they were seated in a child's seat with their parent sitting in a chair next to them across the table from the experimenter. Parents were

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

asked not to assist their toddlers in any way throughout the sessions. Two instruction tasks, the first used to accustom the toddlers to the procedure and the second used to help keep them on task, were administered prior to and half way through the task. The tasks were cards showing three objects, two of which had the same label (e.g. two birds) and one that had a different label (e.g. a hat).

Twelve of the 20 participants were first given a production task, after which their comprehension was tested. The remaining 8 children were reluctant to label any of the objects, so they were tested only on the comprehension task. The opposite order (i.e. comprehension then production) was not given because of the probable influence the experimenter's use of the words in the comprehension task would have on later production. For example, on hearing the experimenter ask the child to point to a dog and then asking if there is another dog, the child may be more willing to produce the word 'dog' for one or two of the items later on.

The production task was administered as follows. The experimenter showed the first card and allowed the child to inspect it ('Look at these pictures!'). After the child had ample time to look at all three pictured objects on a card, the experimenter then pointed to each picture and asked for its label ('What is this one?'). If the child did not respond, the experimenter would pretend to guess incorrectly (e.g. 'Is it a cup? No? What is it?') thus encouraging participants to state the correct label. All responses, correct or not, were praised with comments from the experimenter ('So that is what that one is called! Good job.'). If an item was mislabelled by the child, no correction was made. If the child said nothing or seemed puzzled by what to say, the experimenter paused a bit longer and then went on to the next item ('Okay, let's look at this one now. What is this one called?'). If, after a few items, the child had offered no responses, the production part of the experiment was abandoned and comprehension was then tested.

The comprehension task was given to all 20 participants. The comprehension task was administered by asking children to place a sticker (a small post-it note with a squiggle on it) on one of the pictured objects. One instruction trial was administered before the test items. For example, children were asked to 'Put the sticker on the bird.' After the sticker was placed on one of the birds, the experimenter praised the child and asked if there was another object of the same name (e.g. 'Good job! Is there another bird?'). The experimenter then gave the child another sticker to place on the other bird. After the stickers were removed, the experimenter asked the child to put the sticker on the hat. After the child did so, the experimenter once again asked if there was another hat and handed the child another sticker. If no response was made or the sticker was placed on top of the other sticker on the hat, the experimenter said, 'That's right! This is the only hat! Good job.' If the child responded incorrectly (which seldom happened), the ex-

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

perimenter gently made a correction and went on with the session. All participants passed at least one of the instruction trials. A couple of children seemed more interested in getting a sticker than in the task itself. For these children, the experimenter did not hand out another sticker; instead, the first sticker was removed from the page and handed back to the child with each request (i.e. the same sticker was used throughout the entire task). Half the test trials were then administered. A second instruction trial was administered in the same manner half way through the testing session after which the remaining half of the test trials were given. If a participant became frustrated or bored with this technique, the experimenter ended the session (see results section for average number of trials completed). Since the procedure for the high- and low-contrast trials was the same, a comparison of participants' selections could be made between the trial types.

RESULTS

Production measure

The first question asked was whether data collected in a laboratory setting would replicate the naturalistic data collection by Rescorla (1980). Based on Rescorla's findings, it was predicted that children would produce correct labels for the early items but make overextensions for the late items because these labels were not yet being produced.

Production was obtained from 12 children who labelled an average 17·7 pictured objects (range: 9–24). A label was considered correct if it was the basic-level term commonly used for the object or the correct sound associated with the object (e.g. cat, kitty, and meow were all considered correct labels for the cat; train and choo-choo were considered correct labels for the train). A label was considered an overextension if it was incorrect but from the same domain as the referent (e.g. the raccoon was called a 'cat'). Other incorrect labels (e.g. calling the egg 'table') or inaudible labels were scored as 'other errors.' Only rarely were superordinate terms offered (1 child labelled the moose as an animal; 1 labelled celery as salad). These responses were considered separately and were not counted as being either correct, overgeneralizations or incorrect.

The data were analysed by comparing the labels that were produced for the early and late items. Because twice as many early than late items were tested (see Table 1), proportion scores were entered as the dependent measure. The proportion of early nouns was calculated out of a maximum of 16 possible; the proportion of late nouns was calculated out of a maximum of 8 possible. Factors entered into a repeated measures analysis of variance were: item labelled (early, late) and response (correct; overextension; other error). The results showed main effects for item and response as well as a significant

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

interaction between them, $F\,(2, 22) = 64{\cdot}63$, $p < 0{\cdot}001$. The Hyuhn-Feldt correction was used for this and all subsequent repeated measures analyses reported. This interaction is displayed in the left portion of Figure 1.

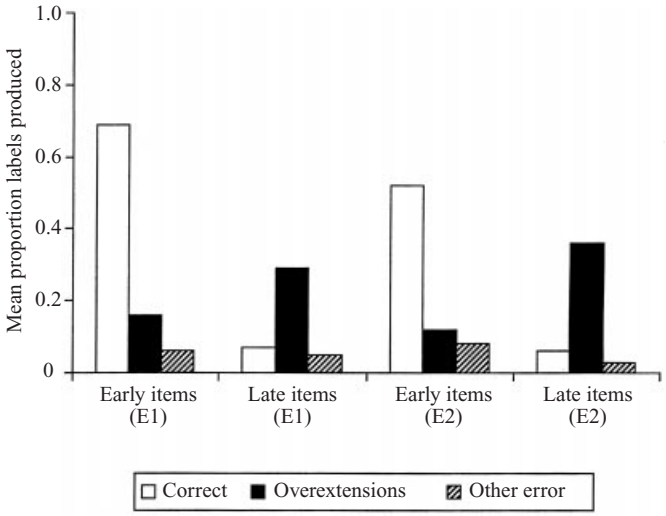


Fig. 1. Mean proportion of correct labels, overextensions, and other errors produced in Experiments 1 and 2.

As shown in Figure 1, correct labels were generated more often for the early (0·69) than late items (0·07; $F\,(1, 11) = 119{\cdot}14$, $p < 0{\cdot}001$). This finding is consistent with data obtained from parent inventories by Fenson and his colleagues (1994) and verifies that the division of the pictures into the early versus late items is correct. More overextensions were made for the late (0·29) than the early (0·16) items, $F\,(1, 11) = 5{\cdot}33$, $p < 0{\cdot}05$. (Note that proportions for each response type were calculated for each subject and the average proportion scores are presented. No credit was given for items not labelled.) Other errors made for the early (0·06) and late (0·05) items did not significantly differ. Thus, overextensions are the most probable type of label children will offer in situations in which the correct label is not part of their production vocabularies, a replication of Rescorla (1980).

Comprehension

The next questions concern the relationship between comprehension and production. In the present experiment, only the early items were tested for comprehension. It was reasoned that if errors in production are primarily due to retrieval failures, then few errors in comprehension would be found. If,

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

however, comprehension and production follow the same general acquisition pattern (Mervis & Pani, 1980; Reznick & Goldfield, 1992), then errors in comprehension can be expected to approximate those found in production.

In the first analyses, the item participants first pointed to or placed a sticker on was counted. Out of 16 test items, participants completed an average of 14·8 comprehension trials (range: 9–16 trials; high-contrast trials: $M = 7{\cdot}4$; low-contrast trials: $M = 7{\cdot}6$). As can be seen in Figure 2, the pattern of these

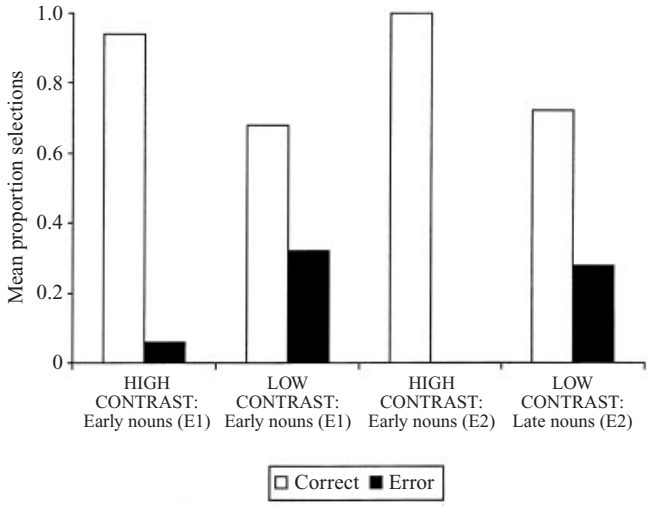


Fig. 2. Comprehension of early nouns tested in Experiments 1 and 2 is reflected in the percentage of selections to correct and incorrect items in the high and low contrast trials (out of a possible of 8). In this figure, only first choice data are displayed.

data mirror the pattern of the production data. Significantly more correct selections were made in the high ($M = 0{\cdot}94$) than the low contrast trials ($M = 0{\cdot}68$; t (19) = 5·06, $p < 0{\cdot}001$).

Next, the kinds of errors made were examined. In the low-contrast trials participants could err in two ways. They could choose the distractor item from the same domain as the referent (i.e. they could make an overextension) or they could choose the distractor item from the contrasting domain as the referent. The majority of errors in the low contrast trials were overextensions ($M = 0{\cdot}29$ out of 0·32; t (19) = 6·01, $p < 0{\cdot}01$). For example, on trials in which the train was requested, participants were more likely to err by selecting the bus than the pig. On the high contrast trials, both distractors constituted errors from the contrasting domain as the referent. The incorrect selections made on the high-contrast trials ($M = 0{\cdot}07$) were also significantly less than the overextensions made on the low contrast trials, t (19) = 4·74, $p < 0{\cdot}001$.

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

Second choices were examined next. Of interest was whether children who made incorrect selections would make a correct second selection. The data were also examined to see if children who made a correct selection would then make a second selection to another item. Conditional probabilities were calculated. On the low contrast trials, children who made a first selection to the overgeneralization item, made more second selections to the correct than the incorrect item, (0·74 v. 0·04, t (19) = 13·7, $p < 0·001$). Children who made correct first selections went on to make second selections to the overgeneralization item significantly more often than to the item that contrasted to the domain of the referent (0·42 v. 0·02, t (19) = 5·87, $p < 0·01$). On the high contrast trials, the few incorrect first selections (9 total) were followed by correct second choices on 2 occasions. Correct first selections were followed by incorrect second selections only 12% of the time, a rate significantly less than the second selections found on the low contrast trials, t (19) = 4·66, $p < 0·01$.

The next analyses showed that the above findings are consistent across the four domains of objects. An analysis of variance for repeated measures was conducted with domain (animal, vehicle, food, clothing) and selection (correct, incorrect) as within subjects measures. No significant results were found for domain (F (3, 57) = 0·49, $p = 0·69$) nor for the interaction involving domain and selection (F (3, 57) = 1·08, $p = 0·36$).

DISCUSSION

In the present experiment, overextensions were often made to the late items, items whose labels are not typically in children's productive vocabularies. One reason for the overextensions may be that children do not know, either in terms of production or comprehension, the correct labels. That is, they may categorize the late items as different from the early items in terms of what kinds of things they are (because they belong to contrasting basic-level categories), but they are willing to consider a known label for them. Alternatively, children may actually think that the same label is appropriate for two items from contrasting basic-level categories (e.g. bus and train) because they do not distinguish them as different kinds of things. In the latter case, overextensions may occur for the late nouns as well (e.g. the child may point to the train when asked to point to the bus).

It was also reasoned that correct selections of the referents to the late nouns can occur for two reasons. One, it is plausible that many children may already comprehend these terms even if they do not produce them. Alternatively, children might be unfamiliar with these nouns. In this latter case, a correct response may be attributable to a 'fast mapping' strategy (e.g. Carey, 1982). That is, participants may point to the bus when it is requested not because they know that 'bus' is the correct term for that item, but because they know

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

the correct terms for the other items (e.g. train and pig) and they then can use mutual exclusivity to infer that the correct item (i.e. bus) is the one requested.

In order to test a fast mapping hypothesis, the experimenter used nonce labels to test the remaining half of the participants (e.g. fep: Where is the fep?). A comparison between the selections based on requests using nonce terms to those based on the late nouns should indicate the extent to which the late nouns are comprehended. If the late nouns are unknown and mutual exclusivity guides responses then no significant differences between the nonsense and late item trials would be expected. In the next experiment, production was tested for all the items in order to replicate the findings from Experiment 1.

EXPERIMENT 2

METHOD

Participants

Sixteen participants, 8 males and 8 females, participated in the experiment. Their mean age was 2;1 (range: 2;0–2;1). For this experiment, only the first two pages of the MacArthur Communicative Development Inventory were filled out by parents indicating their children's productive vocabularies on nouns from 5 domains: animal labels and sounds, vehicles, toys, food and drink, and clothing items. It was decided to use only the first two pages of the inventory because these pages contained the words in the domains most relevant to the experiments and, since it was less time-consuming for the parents, we could get the information from them in a more timely manner and always before the experiment was conducted. Of the 165 items, participants were reported to be producing an average of 125 items (S.D. = 45). No significant differences were found between the group tested on the late learned nouns ($M = 138$; S.D. = 38) and the group tested on using the nonsense labels ($M = 114$; S.D. = 50). The data on these items were then compared to the participants in Experiment 1 who produced an average of 117 items (S.D. = 55). Thus, the participants in this experiment were similar to those tested in Experiment 1.

One additional subject was tested but dropped from the data due to experimental error. Participants were obtained from the same source as in Experiment 1 and were given a small toy for their participation.

Procedure

The same test booklets used in Experiment 1 were used to assess production and comprehension. Production was elicited in the same manner as in Experiment 1 with the exception that the experimenter asked participants to label the referent of the late item before asking for labels to the two early

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

items on each page. (In experiment 1, the experimenter tended to move from left to right fashion across each page when asking for labels.) For the comprehension task, all participants were asked to point to the referent of the early items in the high contrast trials (i.e. the items listed in the left column in Table 1). On the low contrast trials, half of the participants were requested to indicate the referent of the basic-level term that is commonly learned later in development (i.e. the items listed in the right column in Table 1). For the remaining participants, nonce terms were used by the experimenter (e.g. 'Put the sticker on the fep.'). As shown in the production data in Experiment 1 the real word names for these referents are acquired much later in development.

RESULTS

Production

The production results were highly similar to those in Experiment 1. Production was obtained from 14 of the 16 participants who labelled an average of 15·1 pictured objects (range: 4–26). As in experiment 1, a label was considered correct if it was the basic-level term commonly used for the object or if the correct sound associated with the object was used. On only 1 occasion was a superordinate term given (i.e. the beet was called a plant).

Proportions were calculated in the same manner as in Experiment 1. The proportion data were analysed with Item labelled (early, late) and Response (correct, overextension, other error) as within-subject factors. As can be seen in the right portion of Figure 1, a significant interaction between Item and Response was found, $F(2, 26) = 25{\cdot}14$, $p < 0{\cdot}001$. Significantly more correct labels were given for the early (0·52) than late items (0·06), $F(1, 13) = 53{\cdot}74$, $p < 0{\cdot}001$. Significantly more overextensions were made for the late (0·36) than early (0·12) items, $F(1, 13) = 9{\cdot}42$, $p < 0{\cdot}01$. No significant difference was found for the other errors between the early (0·08) and late (0·03) items ($p > 0{\cdot}10$).

Comprehension

All participants were given the same instructions on the high-contrast trials (e.g. on the trial in which a shirt, pig and hippo were depicted, participants were asked to put the sticker on the shirt). However, in the low contrast trials, half of the participants were tested on the late item label (e.g. put the sticker on the hippo) and the remaining half were tested using nonsense labels (e.g. put the sticker on the fep). To make sure the two groups were equivalent, the data on the high contrast trials were examined and the two groups were compared. No significant differences were found (100% v. 95% correct selections on the high contrast trials). An examination of the second choice data on the high contrast trials was also made between the two groups.

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

Second choices were made only 5% of the time, all of which followed correct selections.

The first analysis to be presented and discussed was conducted on the data from the participants ($N = 8$) who were asked to point to the late items (e.g. hippo). As will be shown, overextensions were frequently made on the low contrast trials, a result mirroring that of Experiment 1. This finding was somewhat surprising in that the pattern of overgeneralization between the early and late items is symmetrical. That is, not only are early learned terms overextended to items whose labels are learned later in development (e.g. the term pig is extended to include a hippo) but later learned terms were also overextended to the items whose labels were learned early in development (e.g. the term hippo is extended to include a pig).

As can be seen in the right portion of Figure 2, participants also made significantly more correct selections in the high ($M = 1{\cdot}0$) than the low-contrast trials ($M = 0{\cdot}72$), $t\ (7) = 4{\cdot}69$, $p < 0{\cdot}01$. An examination of the incorrect selections made in the low-contrast trials indicated that participants made significantly more overextensions ($M = 0{\cdot}22$) than other errors ($M = 0{\cdot}06$, $t\ (7) = 2{\cdot}72$, $p < 0{\cdot}05$). No errors were made on the high contrast trials by this group.

The pattern of first and second choices was examined next. Children who first chose the overextension item on the low contrast trials went on to make a second selection to the target item more often than to the item from the contrasting domain (0·13 v. 0·01; $t\ (7) = 4{\cdot}06$, $p < 0{\cdot}01$). Children who made correct first choices on the low contrast trials then went on to make more second choices to the overextension items than to the contrasting error items (0·26 v. 0·03; $t\ (7) = 3{\cdot}39$, $p < 0{\cdot}05$).

As in Experiment 1, the data were analysed to see if differences were found among the four domains. No significant differences were found for domain, $F\ (3, 5) = 0{\cdot}85$, $p = 0{\cdot}52$, nor the domain by selection interaction, $F\ (3, 5) = 1{\cdot}14$, $p = 0{\cdot}42$.

The data from the group tested using the nonce terms ($N = 8$) were examined next. It was predicted that if children were using a 'fast mapping' strategy to point to correct referents for the late items, then they may also select the late items when asked to point to the referent of a nonce term. However, these participants frequently refused to make selections at the experimenter's request. Out of 58 trials using the nonsense words, participants said 'no' to 23 of them (40%). This is considerably greater than the percentage of 'no' responses made on the high contrast trials (16%: $t\ (7) = 2{\cdot}41$, $p < 0{\cdot}05$). It would appear that children frequently did not know how to interpret these terms and were unwilling to guess. However, the data for the remaining 35 trials (in which participants were willing to guess) shows that some degree of 'fast mapping' was helpful. These results are consistent with other research showing that children will typically choose an item for which

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

they do not have a name when presented with a novel term such as 'dax' (Golinkoff, Shuff-Bailey, Olguin & Ruan, 1995). Selections were scored as late item (e.g. hippo; the item that would be predicted given a fast mapping strategy), early item, (e.g. pig) and contrast (e.g. shirt).

Results showed that participants tended to make more selections to the late items ($M = 0{\cdot}48$) than either the early ($M = 0{\cdot}20$; $t\ (7) = 2{\cdot}12$, $p = 0{\cdot}07$) or contrast items ($M = 0{\cdot}20$, $t\ (7) = 2{\cdot}39$, $p < 0{\cdot}05$). However, second choices were made to all three items equally (M's $= 0{\cdot}08$). Note that this is a considerably smaller proportion of second choices than made in previous conditions. Whereas participants were willing to generalize the labels for the late items to the early items and vice versa (thus making overgeneralizations), they did not treat the nonce terms the same way. It is possible that the children in this study did not generalize the nonce terms because they were treating them as if it they were either basic or subordinate labels (see also Taylor & Gelman, 1989).

The data were examined next for any differences that may have occurred among the four domains. No significant difference was found, $F\ (3, 5) = 0{\cdot}82$, $p = 0{\cdot}54$.

GENERAL DISCUSSION

Two experiments were conducted in order to examine children's production and comprehension of basic-level nouns selected from the domains of animals, vehicles, clothing and food. The goal of these experiments was to explore further the taxonomic assumption that proposes that children assume that nouns refer to a category of objects (Markman, 1989). Overall, the results suggest that the boundaries between domains guide children's extensions of the nouns they learn early in development (nouns produced at around age 1;4 to 1;6) as well as those they learn slightly later in development (nouns comprehended but seldom produced at 2;0).

The production data showed that although the majority of labels were basic-level nouns, only 62% were correctly used by adult standards whereas 29% were overextensions (i.e. labelling an item with a term for another referent belonging to the same domain as the target item). The remaining 9% of the production data were other error types (associative or perceptual extensions as well as inaudible utterances). These findings closely replicate earlier work by Rescorla (1980) who examined children's spontaneous (rather than elicited) productions.

The comprehension data of the early nouns tested in Experiment 1 showed that when distractors from the same domain as the correct referent are available along with the correct referent itself, overextensions in comprehension mirror production data (29%). Thus, comprehension and production appear to follow the same general course of development (e.g. Reznick & Goldfield, 1992). Interestingly, in Experiment 2, children showed

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

the same pattern of selections when tested on their comprehension of the late nouns, nouns that were not commonly found in their productive vocabularies. This finding points to the common belief that comprehension precedes production, yet further shows that even in their comprehension of newly learned nouns (when communicative constraints are not pressing) overextensions are likely.

One question that is raised is why a child would generalize a relatively novel term to objects whose labels they are already producing. One possibility is that the tasks themselves led children to use weaker criteria for making their selections. However, children made few errors on the high contrast trials and they were comfortable telling the experimenter 'no' when queried using nonce terms (unlike most developmental research, 'no' was an acceptable response). Given that the production data replicated Rescorla's data collected under naturalistic situations, it seems unlikely that the comprehension data were compromised by an experimental artifact. The more likely explanation of the overgeneralizations in comprehension is that children use the boundaries between domains as a reliable guide in limiting their extensions of nouns, particularly of nouns for which they only have some rudimentary understanding.

The results from Experiment 2 also provided evidence for a fast-mapping strategy (Carey, 1982). Queries made using nonce terms showed that when children made selections, they tended to choose the late item. Thus, it remains possible that correct selections to the late nouns were also attributable to fast mapping. Yet, pronounced response differences were found in the data comparing the late and nonce terms. Children were more likely to make second selections to real than to nonce words even though on every trial another item from the same domain could have been selected – that is, the conditions were exactly the same for both groups. On the other hand, the data from the children who said 'no' suggest that some of the late nouns were comprehended and children were not willing to apply two labels for one object, a result that supports the mutual exclusivity hypothesis (Markman, 1989). Previous research suggests that after the vocabulary spurt, children are less willing to apply two terms to the same referent (e.g. Mervis, 1987). The 'no' responses are likely due to the fact that the children in the present experiments were rapidly acquiring new words and many of them knew at least the correct domain of many of the late learned nouns.

The question arises as to why domain boundaries are salient in children's extensions of basic-level nouns? There are likely to be at least two answers to this question, one pertaining to how children learn about the activities or properties associated with objects, the other pertaining to their learning of language itself.

The first answer is found in previous research showing that early concepts are not basic-level concepts but those more globally defined (e.g. Bauer &

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

Mandler, 1989). Recent data collected by Mandler & McDonough (2000) shows that even though two-year-olds are forming some basic-level concepts about objects, errors in their generalizations of basic-level properties are often found around 25–30% of the time. McDonough (2000) also showed that animal sounds such as meow, moo, and quack, are overextended around 30% of the time by the age of 2;0. This finding may seem surprising because sounds are often learned as early and sometimes earlier than the labels for animals. The results of these studies show that overextensions are found for properties and sounds alike, just as they were found for the labels in the present experiments.

Young children's understanding of the general domains of objects also proves to be valuable for making inferences about unknown words. For example, Goodman, McDonough & Brown (1998) showed that children aged 2;0 to 2;6 can use the semantic context of a sentence as well as their understanding of superordinate categories to make an inference about the referent of an unknown noun. On hearing 'Mommy feeds the ferret' and asked to choose the ferret from an array of four novel objects, children understand that the correct referent is an animal. They use their understanding of the verb 'feeds', a verb associated with things that eat, namely animate objects, to make this inference. It is interesting to note that many of children's first verbs such as spilling, driving, and reading are those associated with domains such as liquids (e.g. juice, milk, coffee), vehicles and library items (e.g. books, newspapers, magazines). It should be noted that children typically learn these verbs in the context of particular basic-level items. When children are asked to complete a sentence such as 'Daddy drives the _____.' , they reliably give an answer such as 'car' (Goodman, 1989). But, children do not constrain their understanding of drives as limited to the basic-level category of car. Instead, they extend the association to the domain of vehicles, a pattern that matches the results from inductive generalization tasks (Mandler & McDonough, 2000).

In the above examples, similarity can be considered in terms of the common properties among objects, such that animate objects drink and sleep and vehicles are driven and used for transportation purposes. Thus, the presence of features such as mouths, eyes, or wheels can be reliable similarity cues. However, similarity is traditionally thought to be based on the shared overall shapes among objects within a category, a view held by many who investigate language acquisition. Language acquisition itself may play a role in helping children sculpt out basic-level categories. As shown by Landau, Smith & Jones (1988), children's use of overall shape similarity in their extensions of nouns INCREASES as they learn more and more basic-level nouns. The important point here is that the ability to detect similarity (such as found with infants aged as young as 0;3; Eimas & Quinn, 1994) does not automatically determine whether or not one will actually make use of

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

perceptual similarity (of the naïve sort) in cognitive and/or linguistic tasks. The use of similarity heuristics, whether in terms of particular features or overall shape, is learned.

Overall shape similarity correlates with basic-level categories and the more nouns children add to their lexicons, the more likely it is that children pick up on the relation between nouns, their extensions and the shape of objects. Although children made many correct selections in the present experiments, they frequently included rockets (shown in an upright launching position without wings) as acceptable airplanes, items that do not share overall shape similarity. In the same fashion, vests were considered to be shirts even though they did not have sleeves or collars; hippos were called pigs even though their faces are shaped differently and they do not have curly tails (it should be noted that hippos are also called cows); and red strawberries and green apples are both called apples even though they differ in overall texture, shape and stem configuration. One could argue that many of these items share some general degree of shape similarity. However, if this is the case, then the claim that basic-level categories are easily distinguished based on shape is incorrect. Perceptual features that allow inferences as to the function or kind of object also need to be understood (Nelson, 1985).

When examining the extensions made by older children (3;0 and older), overall shape similarity can play a role in word extensions, sometimes overwhelming their knowledge of taxonomic relations. For example, Imai, Gentner & Uchida (1994) showed that children will choose an unrelated item based on overall shape similarity rather than a taxonomically-related item when nonce terms are used (e.g. the same nonce word used to refer to a cake is extended to a top hat). However, when using the same kind of stimuli (and replicating the findings of Imai *et al*. with younger children; Poulin-Dubois, Klein, Graham & Frank, 1993), two-year-olds do not make such errors when real words are used. For example, when asked to point to the cake (or to find another cake) their errors are taxonomic rather than similarity based (McDonough, 2000). In the context of words that are meaningful, children do not simply use overall shape similarity alone for their extensions. Rather, they are able to make their extensions based on what pictured items represent, not the odd way they were drawn in order to emphasize overall shape similarity.

The findings of the present experiments along with the research on the development of conceptual categories indicate that the developmental pattern of early concepts and language acquisition is a refining process, in which children hone in on matching correctly a property, word and referent. Based on developmental data showing how an increasing knowledge base influences categorization behaviours, it is assumed that some of this honing has already been done by age two. Because of this assumption, the present experiments tested only a limited range of items within superordinate boundaries. Further

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

research examining the extent to which generalizations are made throughout a domain at various ages is clearly needed. Also needed is a direct comparison between basic level choices along with distractors from the same domain. In the present experiments basic level choices were not offered because the goal was to test the limits or boundaries of children's generalizations. Previous research has shown no significant differences between basic and superordinate level responses with this age group in their extensions of nonce terms (e.g. Bauer & Mandler, 1989); however, given that the present results indicate that nonce terms can provide slightly different results than when real terms are used, this research is underway.

REFERENCES

Abbate, M. S. & LaChappelle, N. B. (1979). *Pictures, please*! Tucson, Arizona: Communication Skill Builders, Inc.

Bauer, P. & Mandler, J. M. (1989). Taxonomies and triads: conceptual organization in one- to two-year olds. *Cognitive Psychology* **21**, 156–84.

Behrend, D. A. (1987). Overextensions in early language comprehension: evidence from a signal detection approach. *Journal of Child Language* **15**, 63–75.

Carey, S. (1982). Semantic development: the state of the art. In E. Wanner & L. R. Gleitman (eds), *Language acquisition: the state of the art*. New York: Cambridge University Press.

Eimas, P. D. & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development* **65**, 903–17.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D. J. & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development* **59**.

Golinkoff, R. M., Shuff-Bailey, M., Olguin, K. & Ruan, W. (1995). Young children extend novel words at the basic level: evidence for principle of categorical scope. *Developmental Psychology* **31**, 494–507.

Goodman, J. C. (1989). *The development of context effects in spoken word recognition*. Unpublished doctoral dissertation, University of Chicago.

Goodman, J. C., McDonough, L. & Brown, N. B. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development* **69**, 1330–44.

Imai, M., Gentner, D. & Uchida, N. (1994). Children's theories of word meaning: the role of shape similarity in early acquisition. *Cognitive Development* **9**, 45–75.

Kuczaj, S. (1982). Young children's overextension of object words in comprehension and/or production: support for a prototype theory of early object word meanings. *First Language* **3**, 93–105.

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

Mandler, J. M. & McDonough, L. (2000). Advancing downward to the basic level. *Journal of Cognition and Development*, in press.

Markman, E. M. ((1989). *Categorization and naming in children: problems of induction*. Cambridge, MA: MIT Press.

McDonough, L. (2000). Early concepts and language acquisition: what does similarity have to do with either? *Festschrift honoring Jean M. Mandler*. New Jersey: Erlbaum. in press.

Mervis, C. B. (1987). Child-basic object categories and early lexical development. In U. Neisser (ed.), *Concepts and conceptual development: ecological and intellectual factors in categorization*. New York: Cambridge University Press.

Mervis, C. B. & Pani, J. A. (1980). Acquisition of basic object categories. *Cognitive Psychology* **4**, 496–522.

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press

Naigles, L. G. & Gelman, S. A. (1995). Overextensions in comprehension and production revisited: preferential-looking in a study of dog, cat, and cow. *Journal of Child Language* **22**, 19–46.
Nelson, K. (1985). *Making sense: the acquisition of shared meaning*. San Diego: Academic Press.
Poulin-Dubois, D., Klein, B., Graham, S. A. & Frank, I. (1993). Shape and the taxonomic assumption. Poster presented at the 25th Annual Stanford Child Language Research Forum. Stanford, CA.
Rescorla, L. (1980). Overextension in early language development. *Journal of Child Language* **3**, 321–335.
Reznick, J. S. & Goldfield, B. A. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology* **28**, 406–13.
Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology* 382–439.
Taylor, M. & Gelman, S. A. (1989). Incorporating new words into the lexicon: preliminary evidence for language hierarchies in two-year-old children. *Child Development* **60**, 625–36.
Thorndike, E. L. & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Teachers College, Colombia University.
Waxman, S. R., Senghas, A. & Benveniste, S. (1997). A cross-linguistic examination of the noun-category bias: its existence and specificity in French- and Spanish-speaking preschool-aged children. *Cognitive Psychology* **32**, 183–218.

https://doi.org/10.1017/S030500090200507X Published online by Cambridge University Press