Conceptual and Lexical Hierarchies in Young Children

Susan A. Gelman
Sharon A. Wilcox
University of Michigan

Eve V. Clark
Stanford University

Linguistic form and conceptual level both play a role in the structure of adult lexical hierarchies. The present studies examined how these factors might affect acquisition. In their linguistic form, labels can be single nouns (e.g., *oak*) or compound nouns (e.g., *oak-tree*). In conceptual level, categories can be structured at the basic (e.g., *tree*), superordinate (e.g., *plant*), or subordinate (e.g., *oak*) levels. Both of these factors were varied in two experiments, in which 133 children, aged 2;11 to 5;11, were taught novel lexical hierarchies. As predicted, compound nouns were easier to learn than single nouns, especially at the subordinate level. Age of child and category level influenced the kinds of errors children made. The most common error was to treat hierarchically related words as labeling mutually exclusive subsets (analogous to *oak* and *elm*, for example), so that terms from different levels contrasted at the same level. Both factors—linguistic form and conceptual level—influenced children's learning of new lexical hierarchies.

INTRODUCTION

Language allows speakers to label categories at more than one level of abstraction. These categories may be detailed (e.g., *buffy tufted-ear marmoset*) or broader in scope (e.g., for the same animal, *monkey, primate,* or *animal*).

Furthermore, the terms at each level in the hierarchy are included in the level above, so all *buffy tufted-ear marmosets* are *monkeys*, and all *monkeys* are *animals*. Taxonomies like this are found in all languages (Berlin, Breedlove, & Raven, 1973), for as Frake (1969, p. 34) put it, "The use of taxonomic systems is not confined to librarians and biologists; it is a fundamental principle of human thinking." Taxonomies, in brief, offer both flexibility and power in organizing categories (Markman & Callanan, 1984).

Although taxonomic hierarchies are taken for granted in adult cognition, they pose difficulties for young children. For instance, children often seem reluctant to accept more than one name for an object (Macnamara, 1982). They may deny that the same thing can be both "a pig" and "an animal." Children also appear to learn basic-level words (e.g., *dog, horse*) more readily than superordinate or subordinate words (e.g., *animal* or *beagle;* Anglin, 1977; Mervis & Rosch, 1981). And even in tasks that equate exposure to each concept, novel superordinate terms are more difficult for children to learn than novel basic-level terms (Horton & Markman, 1980). Finally, children have difficulty with inclusion itself because they appear not to realize that lower-level instances are included within the next level up (Inhelder & Piaget, 1964; Markman & Callanan, 1984).

Little is known about which of these factors causes children most difficulty, or what kinds of errors children make as a result. But consistent errors often reflect the kinds of principles children observe when acquiring words. The present research was therefore designed to look more closely at how linguistic form and category level affect children's acquisition of more than one term for an object.

Linguistic Form

Categories included within other categories, in many languages, can be labeled with compound nouns, as in the English *oak-tree* or *water-spaniel* (Adams, 1973; Berlin, Breedlove, & Raven, 1966, 1973; Conklin, 1962). Compounds make the relation between levels explicit: *Oak-tree* designates a kind of *tree*, just as *sewing-machine* is a kind of *machine*. Consequently, the compound form may help speakers realize that objects may be labeled at more than one level. So it may be easier to understand that an *oak-tree* is a tree than that an *oak* is a tree. And indeed, children's usage is consistent with this view. Two-year-olds can both produce and interpret novel compounds for subcategories (Clark, Gelman, & Lane, 1985). They produce compounds like *beach-car* and *taxi-car* for different kinds of toy cars, and they do this even younger than age 2 (e.g., Rescorla, 1981). Furthermore, they treat the first, stressed noun as the modifier and the second noun as the head that labels the kind of category being talked about, so they understand that a *mouse-hat* is a kind of *hat*. Compounds, then, might be helpful in the learning of lexical hierarchies in taxonomies.

Category Level

When learning words for objects at different levels within a hierarchy, children have to go beyond their early basic-level terms. They have to work up to terms at higher levels, and down to terms at lower levels. Their use of familiar category labels suggests that superordinate level terms (e.g., *animal*) and subordinate-level terms (e.g., *beagle*) are equally difficult (Anglin, 1977). But evidence from familiar terms does not control for frequency of exposure. When this is controlled, superordinates appear easier for children to learn than subordinates (Mervis & Crisafi, 1982). However, other investigators have questioned how well the geometric figures used in controlled experiments correspond to actual subordinate and superordinate concepts typically encoded in language (Blewitt, 1983a). The subordinate-level distinctions in Mervis and Crisafi's task were very subtle (differences in the placement of a shape and two protrusions on a geometric figure), and so possibly were more difficult than many subordinate-level distinctions otherwise. The importance of category level for the acquisition of labels remained unresolved. Furthermore, category level may interact with linguistic form. Since a major function of compounds is to refer to subordinates (Downing, 1977), they may be more helpful at lower than at higher levels in a linguistic hierarchy.

   To investigate these issues, we taught children novel hierarchies in which we varied both linguistic form and category level. Half the concepts taught were new superordinate-level terms (e.g., a novel term analogous to *plant*); the other half were new subordinate-level concepts (e.g., a novel term analogous to *tulip*). Half the children were taught new compound names (e.g., *zav-flower*), and the other half were taught new single names (e.g., *zav*). After each word was taught, children were tested on their understanding (e.g., "Show me the zavs") as they were shown pictures where they were to identify appropriate exemplars. The question we hoped to answer was whether some hierarchies would be easier to acquire than others. We also analyzed children's errors to find out whether they made consistent misinterpretations of the words they were taught.

EXPERIMENT 1

Method

   Subjects.   Eighty children participated in Experiment 1: 40 three-year-olds (2;11–4;2, mean age 3;5) and 40 five-year-olds (4;11–5;11, mean age 5;4). There were 10 children at each age in each of four conditions, with approximately equal numbers of boys and girls in each. An additional 10 children were dropped from the study for failing the pretest or not completing the session.

Design.   The experiment was a 2 (age: 3 years, 5 years) × 2 (novel concept level: subordinate, superordinate) × 2 (linguistic form: single noun, compound noun) design, with all three factors as between-groups variables.

Materials.   Each child saw 10 sets of pictures. Each set contained two exemplars each of two categories that contrasted either at the subordinate level (e.g., tulips vs. daisies) or at the basic level (e.g., flowers vs. trees), as indicated in Table 1. In the latter case, both instances of each basic-level category were members of the same subordinate kind (e.g., both daisies). One exemplar of each category was located on the right-hand side of the page, with the second exemplar on the left-hand side. The two pictures on the left-hand side, for each picture set, were initially covered by a cardboard flap, so only the two pictures on the right were visible.

Table 1.   Experiment 1, Concepts Taught and Corresponding Picture Sets

Concept Taught	Pictures Used	Words Used
Subordinates		
tulip (daisy)[a]	tulips, daisies	flower, zav
dachshund (mutt)	dachshunds, mutts	dog, mef
skate [fish] (bass)	skates, bass	fish, kiv
cobra (rattlesnake)	cobras, rattlesnakes	snake, pume
parakeet (gull)	parakeets, gulls	bird, jop
tank top shirt (button-down shirt)	tank tops, button-down shirts	shirt, zon
slotted spoon (tablespoon)	slotted spoons, tablespoons	spoon, wug
director's chair (easy chair)	director's chairs, easy chairs	chair, lorse
wallabee shoe (pump)	wallabee shoes, pump shoes	shoe, cak
police car (compact)	police cars, compact cars	car, vit
Superordinates		
plant	flowers, trees	flower, tree, zav
animal	dogs, rabbits	dog, rabbit, mef
sea animal	fish, octopi	fish octopus, kiv
amphibian	snakes, frogs	snake, frog, pume
flying animal	birds, butterflies	bird, butterfly, jop
clothing	shirts, dresses	shirt, dress, zon
utensil	spoons, knives	spoon, knife, wug
furniture	chairs, couches	chair, couch, lorse
footwear	shoes, socks	shoe, sock, cak
vehicle	cars, bicycles	car, bicycle, vit

[a]Half the subjects were taught the concepts in parentheses; half were taught the concepts not in parentheses.

Procedure. The experimenter saw all the children individually. She began by introducing a puppet, and then explained that they were going to look at some pictures. She first gave children a pretest designed to check that they would be willing to point to the same picture in response to two different descriptions. For instance, children were first shown a card with three shapes on it—a circle covered in yarn, a circle covered in metal foil, and a triangle covered in metal foil. They had first to point to "the circles," and then to "the shiny ones." If children failed the first pretest, they were corrected and given a second one. Those who failed both pretest items were not included in the study.

After the pretest, children were told that some of the words they would hear were words the puppet knew. They were told this to ensure that they realized we were not teaching them actual English words. They were then given new terms to learn (e.g., *a zav*), in one of four labeling conditions, for the categories listed in Table 1. For each picture set, children saw the two right-hand pictures, one from each of two contrasting categories (e.g., a tulip and a daisy). The experimenter labeled each picture, in counterbalanced order, for example, "This [pointing to the tulip] is a flower; it's a zav. This [pointing to the daisy] is a flower; it's not a zav." For each picture, one label named the familiar basic-level category; the other was a nonsense name. Then the experimenter opened the flap to reveal two more pictures, so the child saw a total of four in each set, with two exemplars from each category (e.g., here, two daisies and two tulips). Now children were asked, in counterbalanced order, questions about both labels. For example, a child might hear, "Can you show me the zavs?" followed by "Are there any other zavs?" The latter prompt was repeated until children said "No." They were then asked about the other category, in this example the flowers.

This procedure provided enough information for children to work out the hierarchical relation between category levels. But they could not simply memorize the experimenter's labels since only two of the four pictures (those visible during naming) were explicitly labeled by her. A flap was used to conceal the other two pictures initially so as to avoid the pragmatically odd situation of not naming all the pictures in sight.

After children had seen 10 picture sets, they received a post-test of the same form as the pretest. Ninety percent of the children answered the post-test correctly.

Labeling Conditions. Children were tested in one of four conditions: (a) subordinate-level concepts, single nouns; (b) subordinate-level concepts, compound nouns; (c) superordinate-level concepts, single nouns; or (d) superordinate-level concepts, compound nouns. That is, two of the labeling conditions focused on the subordinate level, the level below the basic level, and two focused on the superordinate level, the level above the basic level. Furthermore, within both the Subordinate and Superordinate conditions, half the children received single words as labels, and half compound words, as shown in Table 2. The

Table 2.   Experiments 1 and 2, Phrases Used to Label Each
Type of Picture, as a Function of Labeling Condition
(with examples drawn from Experiment 1)

Subordinate Level

Single word	[tulip]	This is a flower; it's a zav.
	[daisy]	This is a flower; it's *not* a zav.
Compound word	[tulip]	This is a flower; it's a zav-flower.
	[daisy]	This is a flower; it's *not* a zav-flower.

Superordinate Level

Single word	[flower]	This is a flower; it's a zav.
	[tree]	This is a tree; it's a zav, *too.*
Compound	[flower]	This is a zav; it's a flower-zav.
	[tree]	This is a zav, *too;* it's a tree-zav.

order in which labels were introduced was determined by what appears typical or natural in adult speech to children. The experimenter began with the familiar basic-level term, because previous research indicates that parents generally begin with basic-level terms even when going on to talk about other hierarchical levels (Blewitt, 1983b; Callanan, 1985). The experimenter then gave the novel subordinate or superordinate term.

The only condition in which this order could not be followed was the Superordinate Compound condition, since there the basic-level item received a novel name (e.g., *flower-zav*) and no familiar basic-level term was available. To determine the order of labeling in this condition, we conducted a small pretest with 12 adults. Each subject was asked to read 12 pairs of sentences, each pair varying the order of two clauses. For example, subjects might read "This is a snake-pume; it's a pume" and "This is a pume; it's a snake-pume." (The order of sentences was counterbalanced within each subject's booklet as well as across subjects.) Subjects then were asked to judge which sentence in the pair sounded more "natural." We found a clear preference for the compound in last place in the sentence (94% of the trials with superordinate compounds; 88% of the trials with subordinate compounds). For example, adults judged "This is a pume; it's a snake-pume" to be more natural than "This is a snake-pume; it's a pume." They also judged "This is a fish; it's a kiv-fish" to be more natural than "This is a kiv-fish; it's a fish." On the basis of these results, in the Superordinate Compound condition of the main experiment (with children), the first label given was the more inclusive superordinate label (as in the Subordinate Compound condition). While this ordering makes each condition as natural as possible, it does confound condition and order of naming.

Results

Scoring.   For each set of pictures, children's responses to both requests (e.g., "Can you show me the zavs," "Can you show me the flowers") were

combined and given a single score. Children could give four different types of response, as illustrated in Table 3: (a) *Correct*. The child selects all four pictures when asked for the higher-level category, and two pictures of one subset when asked for the lower-level category. (b) *Mutually exclusive subset*. The child chooses two pictures of one subset when asked for the higher-level category, and two pictures of the other subset when asked for the lower-level category. (c) *Same subset*. The child chooses two pictures of one subset when asked for the higher-level category, and two pictures of the same subset when asked for the lower-level category. (d) *Other*. This included all other responses. In scoring, children were allowed the wrong subset (e.g., daisies instead of tulips) when asked for the lower-level category, so as not to be penalized for forgetting which picture had received the novel label.

Each child received four scores, indexing the number of correct responses, the number of mutually exclusive subset errors, the number of same subset errors, and the number of all other errors. These results are shown in Table 4. The scores were transformed by an arcsine square-root transformation (Rao, 1973, p. 427) and entered into four separate analyses (one per response type). Except where otherwise stated, each analysis was a 2 (age) × 2 (category level) × 2 (linguistic form) ANOVA.

Correct Responses. Children gave more correct answers when they learned compound nouns (43%) than when they learned single nouns (29%), $F(1,72) = 3.84$, $p = .054$. Although linguistic form did not interact with category level, we conducted separate 2 (age) × 2 (linguistic form) ANOVAs on the data from each category level, to determine more precisely the locus of the naming effects. For subordinate categories, compounds elicited significantly more correct answers (41%) than did single words (20%), $F(1,36) = 5.00$, $p < .05$; however, for superordinate categories, there was no difference attributable to compounds (46%) versus single words (38%), $F(1,36) = 0.50$, n.s. That is, children were more likely to learn the name for a new car when it was called *a fep-car* than

Table 3.   Scoring in Experiment 1: Sample Choices

	Subordinate		Superordinate	
Correct	(flowers?)	4 flowers	(flowers?)	2 flowers
	(zavs?)	2 tulips	(zavs?)	2 flowers, 2 trees
Mutually exclusive subset	(flowers?)	2 daisies	(flowers?)	2 flowers
	(zavs?)	2 tulips	(zavs?)	2 trees
Same subset	(flowers?)	2 tulips	(flowers?)	2 flowers
	(zavs?)	2 tulips	(zavs?)	2 flowers
Other	(flowers?)	4 flowers	(flowers?)	2 flowers, 2 trees
	(zavs?)	4 flowers	(zavs?)	2 flowers, 2 trees

Note: Words in parentheses indicate those prompted by the experimenter.

Table 4. Experiment 1, Mean Percent Responses of Each Type as a Function of Age and Condition

	C	MES	SS	Other
Age 3				
Subordinate				
Single	17	53	1	29
Compound	39	43	0	18
Superordinate				
Single	31	52	11	6
Compound	34	47	8	11
Age 5				
Subordinate				
Single	23	72	0	5
Compound	42	55	1	2
Superordinate				
Single[a]	46	32	8	13
Compound[a]	58	22	5	12
Means	36	47	4	12

C = correct.
MES = mutually exclusive subset errors.
SS = same-subset errors.
[a]Numbers do not add up to 100 because of missing data.

when it was called *a fep*. But they were not more likely to learn the superordinate-level name for a vehicle when it was called *a car-fep* rather than *a fep*.

One possible reason that compounds were no more effective than single words for children learning superordinate categories, could be that the Superordinate Compound condition was the only condition in which labeling did not begin at the basic level (see Table 3). To check this possibility, we tested 10 children (mean age 3;9) in a small follow-up study that varied order of labeling. All children learned compound labels for superordinate-level categories, using the same materials as in the main experiment. For 6 of the children, pictures were labeled exactly as in the main experiment (e.g., "This is a zav; it's a flower-zav"). For the remaining children, the basic-level terms were given first (e.g., "This is a flower-zav; it's a zav"). Performance was somewhat better with the original wording (e.g., zav/flower-zav) than when the basic-level term was heard first (e.g., flower-zav/zav) (51% vs. 30% correct), though this difference was not significant. Thus, the relatively worse performance in the Superordinate Compound condition cannot be attributed to the order of the wording.

Errors.   When children made errors, they typically treated the words they heard as labeling mutually exclusive subsets (47% of all responses). Mutually

exclusive subset errors, in fact, accounted for 74% of erroneous responses. Children made more of these errors at the subordinate level than at the superordinate one, $F(1,72) = 5.91$, $p < .02$. However, whereas 3-year-olds made mutually exclusive subset errors equally for both category levels (48% and 50%), the older children made more such errors on subordinates (64%) than on superordinates (27%).

Children also occasionally made same-subset errors, where they chose the same subset in response to both requests, as though the labels were treated as possible synonyms. Although these errors were rare (4% of responses overall, Table 4), they appeared significantly more often with superordinates (8%) than subordinates (0.5%), $F(1,72) = 17.97$, $p < .001$.

Finally, all other errors (random errors) accounted for just 12% of the responses. Most of these errors were made by the 3-year-olds in learning subordinate-level labels. This made for a significant age-by-category-level interaction, $F(1,72) = 5.80$, $p < .02$.

Discussion

Experiment 1 demonstrates that linguistic form influences how readily children learn category hierarchies. In particular, compound nouns are more effective than single nouns, particularly at the subordinate level. It is noteworthy that the compound form was effective despite only minimal training. However, it is also important to note that the task was quite difficult for children, even when they were given compound nouns. Presumably children would perform better if given more extensive training on a smaller set of items.

Despite the straightforward results, there was one limitation of Experiment 1 that we wished to address in another experiment. In particular, we had included subordinate and superordinate categories for which children may already have learned names. For example, some of our subjects may already have known the words *vehicle* or *tulip*, which could have influenced their performance (either positively or negatively) when learning *vit* and *zav* for these same concepts in Experiment 1. In order to address this issue, we carried out a small control study with 27 children (3;7 to 5;11, mean age 4;9, none of whom had participated in Experiment 1) to determine whether they already knew the concepts we were teaching. The materials and procedure were the same as those in Experiment 1, except that no teaching was involved, and children were tested on only the actual English labels. Thirteen children were tested on subordinate-level items only (e.g., *tulip*); 14 were tested on superordinate-level items only (e.g., *plant*). Results when children were tested on real words (40% correct overall) were highly comparable to results when they were tested on novel words in Experiment 1 (36% correct overall).

Thus, it is possible that children's knowledge of the true subordinate and superordinate category labels for these items influenced their performance in Experiment 1. It seems unlikely that the effects of labeling (i.e., the relative

advantage of compound nouns over single nouns) could be attributed to children's prior knowledge. However, the interaction of naming and category level, and the relative advantage of different levels at different ages, could be due to prior knowledge of these concepts.

We conducted Experiment 2 in order to eliminate this possibility. We pretested all items to make sure they would be unfamiliar to young children. And we also simplified the task to make it easier than the task we used in Experiment 1.

EXPERIMENT 2

Method

Subjects. Fifty-three children participated in the main experiment: 27 three-year-olds (3;2–3;11, mean age 3;7) and 26 five-year-olds (4;11–5;11, mean age 5;4). Fourteen three-year-olds and 13 five-year-olds were in the Single Word condition; 13 children of each age were in the Compound Noun condition. There were approximately equal numbers of boys and girls in each condition. One additional child did not pass the pretest and so was not included in the study. An additional 24 children (4;11–6;9, mean age 5;7) participated in a pretest of the materials (see "Pretest," below).

Design. The experiment was a 2 (age: 3 years, 5 years) × 2 (novel concept level: subordinate, superordinate) × 2 (linguistic form: single noun, compound noun) design, with age and linguistic form as between-groups variables, and novel concept level as a within-group variable.

Materials. Each child saw 12 sets of color photographs of animals arranged in a large photograph album (see Table 5). We included only animals in order to lower potential variation across items due to domain, and because it simplified the task of finding sufficient numbers of unfamiliar superordinate-level categories.

Each picture set consisted of five pictures: two training pictures and three test pictures. For the subordinate-level items, each set had the following structure: The training pictures were instances of two distinctly different subtypes of a familiar basic-level category (e.g., two kinds of frogs, one of which was an *amalops*); the test pictures included both training pictures plus a distractor (e.g., a lion). For the superordinate level items, each set was structured as follows: The training pictures were instances of two familiar basic-level categories which together formed a novel superordinate (e.g., a crab and a snail, both kinds of *crustaceans*); the test pictures included both training pictures plus a distractor (e.g., a cat). We included the unrelated distractor picture in each test set in order to identify whether children were translating the new word as *animal*. Each set was arranged with the training pictures on one page and the test pictures on the

Table 5. Experiment 2, Concepts Taught and
Corresponding Picture Sets

Concept Taught	Pictures	Distractors
Subordinates		
tellin	tellin, another clam	cow
dromid	dromid, another crab	rabbit
wapiti	wapiti, another deer	puffin
amalops	amalops, another frog	lion
wallaby	wallaby, another kangaroo	heron
dingo	dingo, another dog	whale
Superordinates		
mollusk	clam, octopus	horse
crustacean	crab, snail	cat
ungulate	deer, elephant	fish
amphibian	frog, snake	tiger
marsupial	kangaroo, koala	lizard
carnivore	dog, bear	dolphin

following page of the album. To equate the subordinate and superordinate sets as much as possible, each set at one level included the same basic-level category as a set at the other level (e.g., *wapiti* is a subordinate of *deer; ungulate* is a superordinate of *deer*).

Pretest. We pretested 10 subordinate and 10 superordinate items to select sets for which children did not already know the conventional words. Each item was tested on 12 subjects. For each item, children saw a page with the three test pictures (e.g., a crab, a snail, and a cat) and were asked to point to instances of both the basic-level name (e.g., "Can you show me the crabs? Are there any other crabs?") and the subordinate- or superordinate-level name (e.g., "Can you show me the crustaceans? Are there any other crustaceans?"), in counter-balanced order. Items retained for use in Experiment 2 were highly unfamiliar to these children: Overall performance on the subordinate level items was 3% correct; overall performance on the superordinate level items was 6% correct.[1]

Procedure. The experimenter saw all the children individually. As in Experiment 1, she began by introducing a puppet and then explained that they were going to look at some pictures. Children received the same pretest used in Experiment 1. After the pretest, children were taught a new word for each picture set. For each picture set, children saw the two training photographs on one page,

---

[1] Children were scored as "correct" in this analysis only if they answered both questions correctly (i.e., the basic-level question and either the subordinate- or superordinate-level question).

one from each of two contrasting categories (e.g., a dingo and another kind of dog). The experimenter labeled each picture, in counterbalanced order, e.g., "This [pointing to one dog] is a dog; it's a dingo. This [pointing to the other dog] is a dog; it's not a dingo." For each, one label named the familiar basic-level category; the other was an unfamiliar English word. Then the experimenter turned the page of the photograph album to reveal three photographs, including the two training exemplars as well as a distractor picture (e.g., here, a whale). Children were asked about both labels, in counterbalanced order. For example, a child might hear, "Can you show me the dingos?" followed by "Are there any other dingos?" The latter prompt was repeated until children said "No." They were then asked about the other category, in this example the dogs.

After children had seen all 12 picture sets, they received the post-test used in Experiment 1. Ninety-six percent of the children answered the post-test correctly.

Labeling Conditions. All children learned six subordinate-level names and six superordinate-level names, presented in counterbalanced blocks. Half the children received single words as labels, and half compound words, as in Experiment 1 (see Table 2).

Results

Scoring. As in Experiment 1, children's responses to both requests (e.g., "Can you show me the dingos," "Can you show me the dogs") were combined and given a single score for each set of pictures. Children could give four different types of response, as illustrated in Table 6: (a) *Correct.* The child selects both exemplars when asked for the higher-level category, and one picture of the subset when asked for the lower-level category. We also included as correct those trials on which children answered correctly at the basic level and selected all three pictures for the superordinate-level category (i.e., children

Table 6.   Scoring in Experiment 2: Sample Choices

	Subordinate		Superordinate	
Correct	(dogs?)	dingo, other dog	(dogs?)	dog
	(dingos?)	dingo	(carnivores?)	dog, bear
Mutually exclusive	(dogs?)	dog (not dingo)	(dogs?)	dog
	(dingos?)	dingo	(carnivores?)	bear
Same set	(dogs?)	dingo	(dogs?)	dog
	(dingos?)	dingo	(carnivores?)	dog
Other	(dogs?)	dog (not dingo)	(dogs?)	dog
	(dingos?)	dingo, whale	(carnivores?)	bear, whale

Note: Words in parentheses indicate those prompted by the experimenter.

treated the new superordinate word as if it meant *animal*), which accounted for less than 8% of all responses on the superordinate trials. (b) *Mutually exclusive error*. Two patterns are consistent with this response: (i) the child chooses the picture of one subset when asked for the higher-level category, and the picture of the other subset when asked for the lower-level category; (ii) the child chooses correctly at the basic level, and chooses the distractor when asked for the new category. (c) *Same-set error*. The child makes the same response (choosing pictures of either the subordinate, basic, or superordinate level) to both questions in a set. To be conservative in scoring, we excluded those cases in which subjects selected all three pictures on both trials, as such a response could be due to response bias. (d) *Other errors*. This included all other responses. In scoring, children were allowed the wrong subset (e.g., the other dog instead of the dingo) when asked for the lower-level category, so as not to be penalized for forgetting which picture had received the novel label.

Each child received four scores, indexing the number of correct responses, the number of mutually exclusive errors, the number of same-set errors, and the number of all other errors. These results are shown in Table 7. The scores were transformed by an arcsine square-root transformation and entered into four separate analyses (one per response type). Except where otherwise stated, each analysis was a 2 (age) × 2 (category level) × 2 (linguistic form) ANOVA.

Table 7. Experiment 2, Mean Percent Responses of Each Type as a Function of Age and Condition

	C	ME	SS	Other
Age 3				
Subordinate				
Single	32	44	14	10
Compound	62	18	19	1
Superordinate				
Single	37	43	10	10
Compound	59	23	12	6
Age 5				
Subordinate				
Single	63	33	3	1
Compound	69	27	4	0
Superordinate				
Single	74	22	4	0
Compound	74	15	3	8
Means	59	28	9	5

C = correct.
ME = mutually exclusive errors.
SS = same-set errors.

Table 1. Coefficient Alphas for IBQ Scales: Mothers

IBQ Scale	12 Months (n = 37)	18 Months (n = 34)	High Risk (n = 14)	Moderate Risk (n = 19)	Low Risk (n = 18)	12-Month Rothbart Norms[a]
Activity	.80	.84	.78	.73	.83	.84
Distress	.82	.76	.82	.75	.77	.78
Fear	.70	.77	.71	.77	.81	.81
Orienting	.56	.79	.73	.62	.77	.72
Smiling	.79	.61	.81	.58	.72	.80
Soothability	.58	.78	.72	.72	.81	.82

[a] Rothbart, 1981.

Table 2. Coefficient Alphas for IBQ Scales: Fathers

IBQ Scales	12 Months (n = 37)	18 Months (n = 34)	High Risk (n = 14)	Moderate Risk (n = 19)	Low Risk (n = 18)
Activity	.82	.81	.74	.74	.84
Distress	.86	.66	.61	.67	.93
Fear	.82	.85	.81	.79	.89
Orienting	.73	.46	.70	.34	.79
Smiling	.84	.61	.81	.67	.91
Soothability	.78	.63	.78	.77	.75

Rothbart scales. Second, the concordance between PTMs' and PTFs' responses is addressed by means of correlations. Finally, two group contrasts are tested: preterm (PT) versus full-term (FT) infants, and low-risk preterm (LR–PT) versus high-risk preterm (HR–PT) infants. The latter comparison was selected to maximize the within-group homogeneity of the preterm contrast.

Internal Consistency

The internal consistencies of the three PT risk groups' responses to the IBQ were examined to determine whether or not parents of PT infants of differing risk status perceive the constructs of temperament similarly to Rothbart's normative sample. Tables 1 and 2 present the Cronbach alphas for the mothers and fathers across the PT groups and across the two age groups.

As shown in Table 1, the general pattern of results reveals that the scales for the PT mothers show moderate to high internal consistency, with alphas ranging from .61 to .84 (median = .75). Furthermore, the alphas are similar across age groups, with medians in the mid .70s for both groups. A similar pattern for fathers' scores is observed in Table 2, with alphas ranging from .61 to .91 (median = .78). The fathers appear to show less internal consistency in the 18-month group (12 months: median α = .82; 18 months: median

Table 3. Correlations of Mothers' and Fathers' IBQ Scores Partialling Out Sex, SES, and Age of Assessment

IBQ Scale	HR–PT n = 14	MR–PT n = 19	LR–PT n = 18	FT n = 20
Activity	.33	.20	.47*	.26
Distress	.83**	.56**	.67**	.72**
Fear	.93**	.66**	.75**	.54**
Orienting	.78**	.21	.23	.26
Smiling	.43	−.05	.23	−.15
Soothability	.59*	−.10	.10	.19

* $p < .10$. ** $p < .01$.

$\alpha = .62$). The least internally consistent ratings were given by parents of the MR infants. Nevertheless, these data demonstrate that PT mothers and fathers at 12 and 18 months TPD find Rothbart's temperament dimensions to be coherent constructs for their infants, regardless of the severity and chronicity of their postnatal illness.

Correlations Between Mothers' and Fathers' Scores

The scores of the PT and FT mothers and fathers were correlated to determine the extent of interparent agreement on ratings of their infants' temperament. The correlations of mothers' and fathers' scores within the three PT groups and the FT group, partialling out infant age, sex, and family SES, are presented in Table 3.

When the correlations across all four groups are considered, the median r is .38 (n.s.), with a range of −.15 to .93. This overall pattern is a somewhat lower interparent agreement than reported on correlations between scores of FT mothers and a second caregiver (Bates, 1980; Field & Greenberg, 1982; Goldsmith & Campos, 1982; Huitt & Ashton, 1982; Rothbart, 1981). However, interparent agreement differs widely depending on the temperament dimension being considered. Distress to limits and fear both show strongly significant interparent agreement across all four groups (median $r = .69$; range = .54–.93), whereas there is little significant agreement in the other four dimensions. It is interesting to note that the HR–PT parents show the strongest overall pattern of interparent agreement, demonstrating high concordance on distress, fear, and orienting and moderate agreement on soothability.

Comparisons of Ratings of Preterm and Full-Term Infants

The means of the mothers' and fathers' temperament ratings for the three preterm groups and the full-term group are presented in Figures 1 and 2. Two sets of planned comparisons were carried out. First, PT parents were compared with the FT parents to test for the presence of effects attributable

are often used in situations with explicit contrast (Clark, Gelman, & Lane, 1985). When learning subordinate-level terms, children need to attend to differences among subcategories—they need to realize that tulips are not daisies—and compounds help. But when learning superordinate-level terms, children need to overlook differences—flowers and trees are the same kind of thing—and there, it is not clear that compounds can help to the same extent. Finally, the relative lack of advantage of compounds at the superordinate level cannot be attributed to how pictures were labeled in that condition: A control study showed that children performed at least as well with the order of labeling used in the main experiment (e.g., "This is a zav; it's a flower-zav") as when the basic-level term is presented first (e.g., "This is a flower-zav; it's a zav").

These results are similar to those reported by Adams (1986) and Waxman (1985). Adams noted that compounds (which she calls "anchors") help 3-year-old children judge statements of class inclusion, as in "A leopard-cat is a cat." However, she did not indicate how accurate children were in their own productions of such terms. Waxman examined the ability of 3- and 4-year-old children to sort pictures of objects into contrasting groups (e.g., grapes of three different varieties). She found that children sorted at the subordinate level more successfully when given novel adjectives (e.g., "the budish ones") than when given novel single nouns (e.g., "the budips"). This effect was reversed for the superordinate level: Children sorted more successfully when given novel nouns than when given novel adjectives. The adjectives Waxman used have a function similar to that of the modifier noun in compounds. In both studies, then, noun *modifiers* functioned to highlight differences among subordinate-level concepts.

But the present study also differs significantly from Waxman's in two respects. First, although noun–noun compounds and adjective–noun combinations both consist of a modifier followed by a (head) noun, the forms are clearly distinct. Modifiers in noun compounds cannot be independently applied to other objects (e.g., *flower* from *flower-zav* or *oak* from *oak-tree*), but adjectives have an independent meaning and can combine with a wide range of nouns. Stress pattern also distinguishes noun–noun compounds from adjective–noun combinations. Compounds carry a characteristic heavy–light stress pattern on the two elements combined in rapid succession, while adjective–noun sequences simply carry slightly heavier stress on the noun (compare *hot-dog* [frankfurter] with *hot dog* [panting canine]).

Second, the skill tested in Waxman's study was the ability to identify the basis for classification, including the level at which to sort. For example, children had to decide that color was the differentiating characteristic for varieties of grapes (red vs. green vs. purple). In our task, the skill being elicited was the ability to coordinate two different levels at once in a hierarchy. The greatest difficulty for children may not lie in detecting whether distinctions are located at a superordinate or subordinate level, but in assigning dual classifications—basic level and either superordinate or subordinate—at the same time. Presumably children

would have had less difficulty if we had provided only the novel word during naming and had tested them on only the one novel word. Yet despite the differences in what we required of children, our results are compatible with Waxman's.

Children's errors may be as revealing as their correct responses. The most pervasive error in the present study was the interpretation of the two words in each set as naming mutually exclusive sets. That is, the terms were taken to be co-hyponyms rather than a basic-level and a higher- (or lower-) level term. For example, *zav* and *flower* were taken to refer to "tulips" and "flowers other than tulips," respectively. This error cannot be attributed to any reluctance to select the same picture more than once: All the children we tested performed well on the pretest and post-test, both of which required children to choose the same (familiar) picture in response to two different descriptions. The prevalence of the mutually exclusive subset error is consistent with both the Principle of Contrast (Clark, 1983, 1987, 1988) and the Principle of Mutual Exclusivity (Markman, 1987). Children may initially treat words from two different levels as being in contrast at the same level in order to avoid having two terms with apparently the same meaning.

For older children, mutually exclusive subset errors were especially common at the subordinate level. Notice that children in the subordinate-level conditions were not taught a contrast for the new word (e.g., *zav*) and, when they made errors, assumed that the new word contrasted with the familiar basic-level word (e.g., *flower*). But children in the two superordinate-level conditions already knew the initial contrast (e.g., that *tree* and *flower* were contrasting terms) and so could build on that contrast in learning the new superordinate. In other words, a known contrast may effectively "block" mutually exclusive subset errors. If this interpretation is correct, then children learning subordinate-level terms should do better if taught two new words that are co-hyponyms—that contrast at the same level (e.g., "these are flowers; this is a zav and this is a wug"), because the nature of the contrast would be clear from the start.

In sum, we have examined how children learn the fundamental skill of labeling categories at different levels in a hierarchical taxonomy. The present work demonstrates the role of language form and conceptual structure in this process. Our results further suggest that the ubiquity of compounds to express subordination in languages of the world may have evolved for functional reasons, to help the language learner. Notice, however, that the relation between form and function is not always direct. Not all compounds designate subordinate categories— Pineapples are not a kind of apple, just as Eskimo pies are not a kind of pie (Frake, 1969). And some compounds may become opaque over time so that their constituent parts are no longer subject to analysis, as in "a green blackboard" or "an albino blackbird" (see Clark & Clark, 1979, for discussion of opacity). It is therefore all the more striking that young children are sensitive to compound form in interpreting (as well as producing) new words.

REFERENCES

Adams, A. (1986). Sentence frame effects on children's category judgments. *Papers and Reports on Child Language Development* (Department of Linguistics, Stanford University), *25*, 1–8.

Adams, V. (1973). *An introduction to modern English word formation.* London: Longman.

Anglin, J.M. (1977). *Word, object, and conceptual development.* New York: Norton.

Berlin, B., Breedlove, D.E., & Raven, P.H. (1966). Folk taxonomies and biological classification. *Science, 154,* 273–275.

Berlin, B., Breedlove, D.E., & Raven, P.H. (1973). General principles of classification and nomenclature in folk biology. *American Anthropologist, 75,* 214–242.

Blewitt, P. (1983a, April). *What determines order of acquisition of object categories?* Paper presented at the biennial meeting of the Society for Research in Child Development, Detroit.

Blewitt, P. (1983b). *Dog* versus *collie:* Vocabulary in speech to young children. *Developmental Psychology, 19,* 602–609.

Callanan, M. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development, 56,* 508–523.

Clark, E.V. (1983). Meanings and concepts. In J.H. Flavell & E.M. Markman (Eds.), P.H. Mussen (Series Ed.), *Handbook of child psychology: Vol. 3. Cognitive development* (pp. 787–840). New York: Wiley.

Clark, E.V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Erlbaum.

Clark, E.V. (1988). On the logic of Contrast. *Journal of Child Language, 15,* 317–335.

Clark, E.V., & Clark, H.H. (1979). When nouns surface as verbs. *Language, 55,* 767–811.

Clark, E.V., Gelman, S.A., & Lane, N.M. (1985). Compound nouns and category structure in young children. *Child Development, 56,* 84–94.

Conklin, H.C. (1962). Lexicographical treatment of folk taxonomies. *International Journal of American Linguistics, 28,* 119–141.

Downing, P.A. (1977). On the creation and use of English compound nouns. *Language, 53,* 810–842.

Frake, C.O. (1969). The ethnographic study of cognitive systems. In S.A. Tyler (Ed.), *Cognitive anthropology* (pp. 28–41). New York: Holt, Rinehart, & Winston.

Horton, M.S., & Markman, E.M. (1980). Developmental differences in the acquisition of basic and superordinate categories. *Child Development, 51,* 708–719.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child.* New York: Norton.

Macnamara, J. (1982). *Names for things.* Cambridge, MA: MIT Press.

Markman, E.M. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 255–287). Cambridge: Cambridge University Press.

Markman, E.M., & Callanan, M.A. (1984). An analysis of hierarchical classification. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 325–365). Hillsdale, NJ: Erlbaum.

Mervis, C.B., & Crisafi, M.A. (1982). Order of acquisition of subordinate, basic, and superordinate level categories. *Child Development, 53,* 258–266.

Mervis, C.B., & Rosch, E. (1981). Categorization of natural objects. In M.R. Rosenzweig & L.W. Porter (Eds.), *Annual review of psychology* (Vol. 32, pp. 89–115). Palo Alto, CA: Annual Reviews.

Rao, C. (1973). *Linear statistical inference.* New York: Wiley.

Rescorla, L.A. (1981). Category development in early language. *Journal of Child Language, 8,* 225–238.

Waxman, S.R. (1985). *Hierarchies in classification and language: Evidence from preschool children.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.