Developmental Science 4:1 (2001), pp 109–118

REPORT

Labelling patterns and object naming

Barbara Landau[1] and Elizabeth Shipley[2]

1. *University of Delaware, USA*
2. *University of Pennsylvania, USA*

Abstract

We examined the effects of different labelling patterns on the generalization of object names. Two-year-olds, three-year-olds and adults were shown two 'standard' objects, which were named with the same label, or with two different labels, or with no label at all. Participants were then asked whether objects morphed to be intermediate to the standards belonged to one of the labelled categories or, in the No Label condition, were 'like' one of the standards. The Same Label condition showed generalization to all intermediates, whereas the Different Label and No Label conditions showed division of the intermediates into two separate categories, with somewhat sharper division under Different Label. These results suggest two possible mechanisms of lexical learning: 'boosting' the equivalence of different exemplars through label identity, and 'differentiating' the exemplars through differences in labelling. The studies provided strong evidence for boosting. Learners are sensitive to the distribution of labels across exemplars, and they hold powerful assumptions about the relationship between these distributions and the underlying naming space. These findings have implications for the early emergence of cross-linguistic differences in lexical learning.

One of the most striking phenomena of early development is children's capacity to generalize object names in a way that closely resembles that of their speaking community. Over the past 25 years, a great deal of research has been devoted to discovering the bases for these generalization patterns (see Woodward & Markman, 1998, for a review). Theoretical contenders have included ontological and taxonomic categories (Markman & Hutchinson, 1984; Soja, Carey & Spelke, 1992; Waxman & Markow, 1995), perceptual representations of objects (Landau, Smith & Jones, 1998), functional properties of objects (Kemler-Nelson, 1995; Landau, Smith & Jones, 1997) and knowledge of an object's history (Bloom, 1996). Although there is controversy concerning the roles of these different similarity metrics, there is general agreement that object naming involves judging which objects are relevantly similar to each other.

In this paper, we ask whether the generalization patterns characteristic of object naming can be modulated by differences in the overt pattern of labelling across objects that is provided to the child learner. Investigators have already shown that object naming tasks engage special kinds of similarity judgments. For example, children generalize names on the basis of taxonomic category rather than thematic relationship (Markman & Hutchinson, 1984; Waxman & Markow, 1995) and on shape rather than texture, size or color (Baldwin, 1992; Smith, Jones & Landau, 1992). The syntactic context in which a novel word appears can also alter the particular dimension that becomes relevant for naming judgments (Landau, Smith & Jones, 1992).

The possibility that the basis for generalization in naming is also modulated by labelling patterns takes on particular importance in the context of renewed attention to cross-linguistic differences in the lexicon. Such differences have led investigators to recognize that any similarity metric must be capable of significant modulation by the distribution of labels over parts of a representational space (Bowerman, 1996). For example, the basic similarity metric underlying color terms is rooted in human color perception (Berlin & Kay, 1969; Ratliff, 1976). Surveys of the world's color term systems have shown that the color lexicon follows the boundaries of color perception, with no single term expressing discontinuous categories such as 'red or green', and all

Address for correspondence: Barbara Landau, Department of Psychology, Wolf Hall, University of Delaware, Newark, DE 19716, USA; e-mail: blandau@udel.edu

terms fully consistent with breaks following from human color perception. At the same time, however, the exact location of the breaks between categories is subject to cross-linguistic variability. For example, although English has different terms for green and blue, some other systems possess a single term covering shades that English speakers call green and those they call blue. Conversely, Russian has two basic terms for our single term 'blue', with separate words for the two shades we call 'light blue' and 'dark blue' (Berlin & Kay, 1969; Davidoff, Davies & Roberson, 1999).

Such cross-linguistic variation also occurs in the domain of spatial terminology. Although universal principles of spatial cognition are assumed to be reflected in much of the spatial lexicon, there is significant cross-linguistic variation in the particular spatial concepts that are expressed by a language's basic spatial terms. For example, Bowerman (1996) points out that English categorizes together the spatial relationships of 'apple IN bowl' and 'hand IN glove', considering both as instances of 'containment'. Yet Korean distinguishes between the two, using different verbs for relationships of 'loose fit' (such as 'apple IN bowl' or 'bowl ON table') versus 'tight fit' ('hand IN glove' or 'top ON pen'). Bowerman reports evidence from spontaneous and elicited language that 24-month-old children have already acquired such language-specific distinctions in their spatial lexicon. She argues that the specificity of these acquisitions proves that linguistic input must have played a strong role in learning.

These cases show that the range of psychologically natural categories receives further structuring from the number of terms in the language and their location relative to each other. It follows from this cross-linguistic variability that, regardless of how clearly the possible natural lexical categories might be served up by non-linguistic systems of knowledge, there must also be a powerful learning mechanism whereby the child comes to set up the right number of divisions based on the number and location of terms in her language. Portions of a conceptual space that are labelled by two different terms in one language may be labelled by just one in another. The learner's task is to discover where the boundary of one lexical category ends and that of another begins.

Despite the pervasiveness of this problem, there is surprisingly little evidence on the role of linguistic labelling patterns in shifting such boundaries. Apart from several studies of cross-linguistic differences in generalization to the same stimuli (Bowerman, 1996; Imai & Gentner, 1997), little is known about the mechanisms underlying the shift of lexical boundaries. However, there is a tradition within speech perception that addresses this very question, but with regard to phonemic categories. There, abundant research has shown that infants reorganize the sound space of their language as they learn about the distinctions made by their native language (Werker, 1994; Jusczyk, 1997). For example, 6-month-old infants raised in an English-speaking environment can discriminate between syllables that incorporate English phonemic distinctions as well as Hindi distinctions; but by 12 months, they can discriminate only English phonemes. This emergence of native phonemic categories (and suppression of non-native phonemic distinctions) precedes the production of lexical items, but nevertheless suggests that infants are sensitive to distributions of linguistic contrasts that are present in their native language and are pertinent to meaning. Thus it is plausible to ask whether the emergence of lexical distinctions is similarly affected by the distribution of lexical items over the possible space of generalization.

In order to examine the role of labelling patterns on the shifting of lexical boundaries, we studied children's and adults' naming generalization under a variety of labelling conditions. In all cases, we used novel objects, for which people could not have formed lexical categories prior to the experiment. In the first experiment, we introduced two quite different objects, which lay at the opposite ends of a similarity space defined by changes in object shape. In one condition, we named both objects with the same name, and in a second condition we named each object with a different name. People were then tested for their generalization of one of the labels to a range of objects lying in the similarity space. The question was whether the difference in labelling patterns would result in different patterns of generalization to the same objects.

Experiment 1

Participants

Twenty-four 2-year-olds ($M = 29$ months, range 27–31 months), 24 3-year-olds ($M = 40$ months, range 36–43 months), and 24 adults participated. The groups were balanced for gender. Children were drawn from preschools local to the University of California, Irvine, and adults were undergraduates there. Informed consent was obtained from the parents of children and from all adult participants.

Design, materials and procedures

Participants were randomly assigned to either the Same Label or the Different Label condition. They observed two standard objects from one of two object sets (see Figure 1). The two standards were placed on stands to left and right of center in front of the participant, and an empty stand was placed between these (see Figure 2). As participants in the Same Label condition were shown the first standard, they were told 'See this (pointing to the standard)? This is a blicket.' Then they were shown the second standard and told 'See this (pointing to it)? This is a blicket.' Participants then were shown each of six objects in the test set, one at a time, placed on the center stand and were asked, for each one, 'Is this a blicket?' Participants in the Different Label condition heard the same labelling event for the first standard, but when they observed the second standard they were told 'See this? This is a steb.' Then they were shown the six test objects and asked, for each one, either 'Is this a blicket?' or 'Is this a steb?' The two standards remained in full view during the procedure. All participants were asked to judge the six test objects from each of the two sets in two

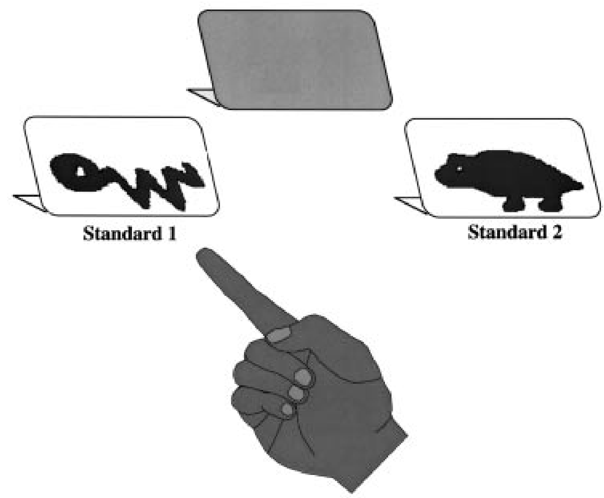

Figure 2 *Prior to test, participants were shown the two standards, each placed on a stand to the right and left, with an empty stand placed in the middle. The six test objects were each placed one at a time on the middle stand.*

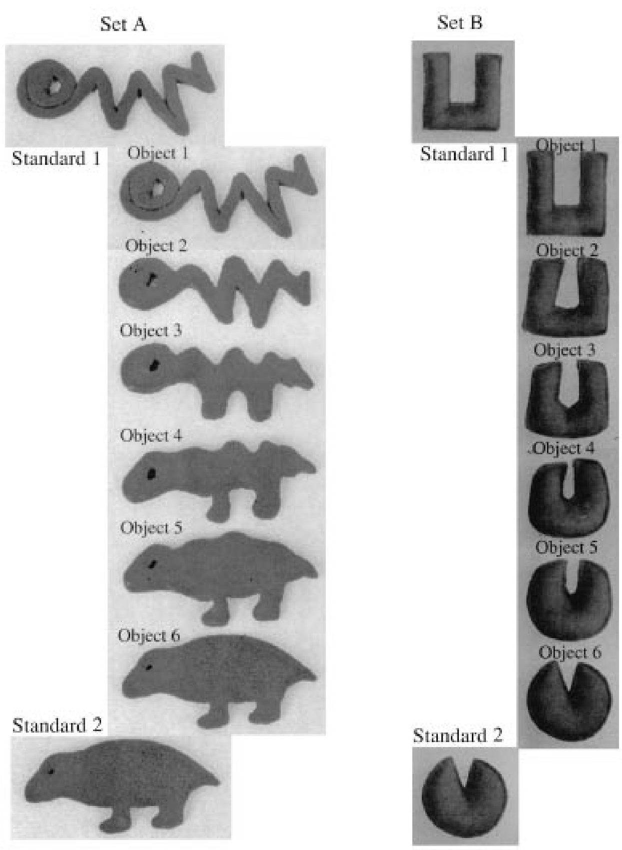

Figure 1 *Objects used in Experiments 1 and 2. Participants were shown the two standards of each set, and were asked to compare these to the six test objects, which included a replica of each standard plus four intermediate objects. All participants saw both sets of objects, one set at a time.*

successive rounds, for a total of 12 test trials per set and participant. In addition, two distractor items were presented twice each at random points during the procedure, and were also queried. These distractor items were completely different from any of the items in the test series, were regularly rejected by most children and all adults, and will not be discussed further.[1]

The six objects of a test set included replicas of the two standards (Objects 1 and 6) as well as four objects of the same size and material as the standards but having different shapes that lay in the similarity space between the two standards (see Figure 1). The four intermediate shapes (hereafter Objects 2, 3, 4, 5) were created by morphing the two standards, using a commercially available morphing package, and rendering each successive shape equidistant from the next. One of the object sets was made of Fimo clay and ranged from a shape similar to a deformed lorgnette to that of an animal. The second set was made of stuffed blue felt and included geometric shapes ranging from a U-shaped object to a pie-shaped object. Different nonce names were used for each of the two sets ('blicket' and 'steb' for one set,

[1] A number of the children in this experiment did accept the distractor items, including nine of the 24 2-year-olds and two of the 24 3-year-olds. This did not occur in Experiment 2, using the same methods, and hence was probably a random effect. However, in order to examine whether any effects in Experiment I were biased because of the children who accepted the distractors, we conducted all analyses both with and without these children. All results remained the same for both analyses. We therefore report the findings based on all children.

'fendel' and 'stad' for the other), and order of presentation of the two object sets was counterbalanced over participants. In the Different Label condition, assignment of the two nonce names to Objects 1 and 6 was counterbalanced over subjects, as was whether they were queried to the name of Object 1 or Object 6. The series of six test objects was presented in one of two randomized orders.

Results and discussion

Preliminary analyses showed no effects of object set in this or further experiments, so these were collapsed for further consideration. Figure 3 shows the mean proportion of 'yes' responses to the question 'Is this a (nonce label)?' over labelling condition for each object and age group. The key issue is the extent to which participants generalize the label to each of the objects, relative to the standard, whose name is being used in the test query. Therefore, the data have been organized so that Object 1 always refers to the standard object whose name was queried, and Objects 2–6 refer to the test objects in descending order of similarity to that object. Thus, data for participants in the Same Label condition and for those in the Different Label condition who were queried to Object 1 are reported for Objects 1–6 as they correspond to the objects in Figure 1. Data for participants in the Different Label condition who were queried to Object 6 have been 'folded over' so that the means for Object 6 are reported as Object 1, means for Object 5 are reported as Object 2, and means for Object 4 are reported as Object 3, and so forth.

As Figure 3 shows, there were large differences in generalization patterns under the two labelling conditions. Participants in the Same Label condition accepted all test objects at ceiling levels. Participants in the Different Label condition showed a sharp decline in acceptance of test objects starting by Object 3 and continuing thereafter. The pattern was the same within each age group, indicating that the distribution of labels affects children as young as 2 years of age.

A 2 (age) × 2 (labelling condition) × 6 (object) analysis of variance performed on the number of 'yes' responses revealed main effects of age ($F(2, 66) = 4.5$, $p < 0.01$), labelling condition ($F(1, 66) = 46.7$, $p < 0.01$) and object ($F(5, 330) = 57.9$, $p < 0.01$), as well as interactions between age and object ($F(10, 330) = 2.5$, $p < 0.01$) and between labelling condition and object ($F(5, 330) = 49.5$, $p < 0.01$). Tukey's *post hoc* comparisons showed that 2-year-olds accepted more objects than adults (Tukey's critical difference 0.69, $p < 0.05$, for this and all subsequent comparisons), participants in the Same Label condition accepted more objects than those in the Different Label condition, and the objects ordered themselves according to similarity, with Objects 1, 2 and 3 reliably different from all others, but Objects 4, 5 and 6 not different among themselves (critical difference 0.37). The interaction between age and object was due to more acceptance by 2-year-olds than 3-year-olds for Objects 2 and 4 (critical difference 0.79).

The interaction of labelling condition and object was examined using two sets of Bonferroni planned comparisons. The first set examined the objects relative to each other, within each labelling condition. There were no differences in acceptance of objects in the Same Label

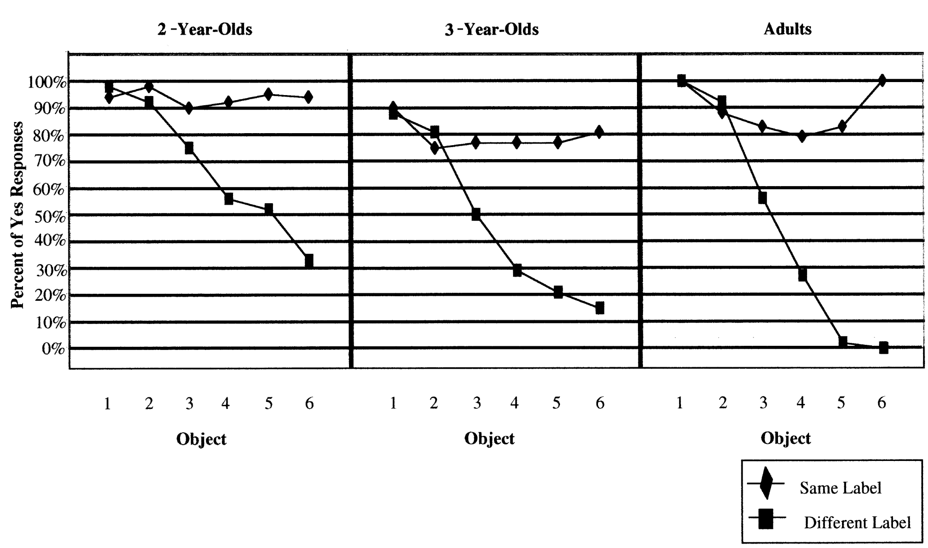

Figure 3 *Mean percentage of 'yes' responses over object in Experiment 1.*

condition. In the Different Label condition, however, there were reliable drops in acceptance with increasing dissimilarity to the standard, with all pair-wise comparisons between objects reliable, except for the comparison of Objects 1 and 2, and Objects 5 and 6 (critical $d = 0.48$). Thus objects dropped in acceptance under the Different Label condition only. The lack of a three-way interaction with age shows that the nature of this interaction did not differ over age groups.

The second set of comparisons examined the difference across labelling conditions for each object. Participants in the two labelling conditions did not reliably differ from each other for Objects 1 or 2. However, starting with Object 3, there was a reliable drop in acceptance of each test object in the Different Label group relative to the identical test object in the Same Label group (critical $d = 0.48$). This suggests powerful effects of labelling patterns on the generalization of a single label. Despite the fact that only objects at the endpoints of the similarity space were named, 2-year-olds, 3-year-olds and adults all assumed that the two very different objects that were named with the same label were part of a continuous category that included all intermediates. In contrast, they assumed that if the same two objects were labelled with different labels they were members of two separate categories whose boundary was somewhere midway between the two standards. Once again, there was no interaction with age, indicating that this pattern held across age groups. Note that the information presented – labelling of the endpoints only – could logically have led to a very different generalization pattern, in which the name of Object 1 was extended all the way through to Object 5, stopping short only of the single object – Object 6 – that was given a different name. However, participants did not organize the naming space this way. Rather, they all appeared to divide the space somewhere in the middle, as if two names require splitting the range of objects in this way. We return to this in the General Discussion.

The results of this study suggest two possible mechanisms of shifting boundaries and thereby modifying generalization. One is 'boosting' the equivalence of objects when a single label spans two different objects, even if they are quite dissimilar. A second possibility is 'differentiating' the objects when two different labels span those objects.

One way to better determine whether the work is being done by one or both of these mechanisms is to compare the two labelling conditions with one in which there is no label involved at all. Many recent studies of lexical learning have shown that children generalize differently when given a naming task than when asked to make similarity judgments in a non-naming context. For example, when asked whether one object 'goes with' another, 'is like' another or 'matches' another, children often use different criteria, putting together objects that share overall similarity, function, or thematic relatedness (Markman & Hutchinson, 1984; Smith *et al.*, 1992). Given that our stimuli were specifically designed to manipulate one kind of similarity (shape-based), a No Label condition should elicit a generalization curve following the changes in shape, but this curve might or might not be the same as the one elicited in the Same or Different Label condition. A direct comparison between the No Label curve and the Same versus Different Label curves would allow us to determine whether these two conditions exert different effects on generalization relative to some neutral pattern of generalization that might be induced in a non-linguistic context. Therefore, we conducted a modified replication of Experiment 1, incorporating a third condition in which neither of the objects was named.

Experiment 2

Participants

Twenty-seven 2-year-olds ($M = 30$ months, range 26–34 months), 34 3-year-olds ($M = 40$ months, range 36–47 months), and 36 adults participated. The Same Label condition included ten 2-year-olds, ten 3-year-olds and 12 adults. The Different Label condition included nine 2-year-olds, 12 3-year-olds and 12 adults. The No Label condition included eight 2-year-olds, 12 3-year-olds and 12 adults. Mean ages and ranges within age group were comparable across experimental condition. The groups were balanced for gender. Children were drawn from preschools local to the University of Delaware, and adults were undergraduates there. Informed consent was obtained from the parents of the child participants and all adult participants.

Design and procedures

The design and procedures were identical to those of Experiment 1 except for the inclusion of a third No Label condition, to which participants were randomly assigned along with the other two labelling conditions. In the No Label condition, participants were shown the first standard and told 'See this?', followed by the second standard ('See this?'). Then they were shown each test object and asked 'Is this one (pointing to the test object) like this one (pointing to either Object 1 or Object 6)?' Half of the participants were queried to Object 1 and half to Object 6. As in Experiment 1, two distractor objects were included for each object set, and each was queried twice, along with the test objects. The distractors were completely different from any of the objects in either test set, were uniformly rejected by all participants, and will not be considered further.

Results and discussion

The numbers of 'yes' responses were submitted to a 2 (age) × 3 (labelling condition) × 6 (object) analysis of variance with the last factor within-subjects. There were main effects of labelling condition ($F(2, 88) = 75$, $p < 0.01$) and object ($F(5, 440) = 132$, $p < 0.01$), and interactions of object with age ($F(10, 440) = 3.74$, $p < 0.01$) and object with labelling condition ($F(10, 440) = 17.7$, $p < 0.01$). The interaction of object with age was due to unsystematic differences between the 2-year-olds and adults in their acceptance of the different object types.

Figure 4 shows the mean proportion of 'yes' responses over labelling condition for each object and age group. As can be seen, the Same Label and Different Label conditions replicated the results of Experiment 1 among all age groups. The No Label condition appeared quite similar in overall shape to the Different Label condition, but did not fall as sharply across the first three objects among 3-year-olds and adults. Among 2-year-olds, the curves appear quite similar, suggesting a possible developmental trend. However, the three-way interaction with age did not reach significance ($p > 0.05$).

The interaction of labelling condition with object was examined using two sets of planned Bonferroni comparisons. The first set examined the objects relative to each other, within each labelling condition. In the Same Label condition, there were no differences in how often each object was accepted. In the Different Label condition, however, Objects 1, 2 and 3 were each reliably different from all others, but there were no differences between Objects 4, 5 and 6. In the No Label condition – which was new to this experiment – the general pattern followed the Different Label condition but with a flatter distribution. Objects 1 and 2 did not differ from each other but were accepted more than all others. Object 3 was also accepted more frequently than 4, 5 or 6. Object 4 was accepted more frequently than Object 6 (critical difference 0.53, $p < 0.05$).

The second set of comparisons examined the difference across labelling conditions for each object. There were no differences for Object 1, confirming that the labelling conditions did not elicit differences for the standard comparison object. For Objects 2–6, participants in the Same Label condition accepted more than those in either the Different Label or the No Label condition. This confirms the boosting function of providing the same name to the two standards. The comparison between the Different Label and No Label conditions showed that Objects 2 and 3 were accepted at lower levels in the Different Label than the No Label condition, with Object 3 reliably lower (critical difference 0.53, $p < 0.05$). Although the interaction of label-

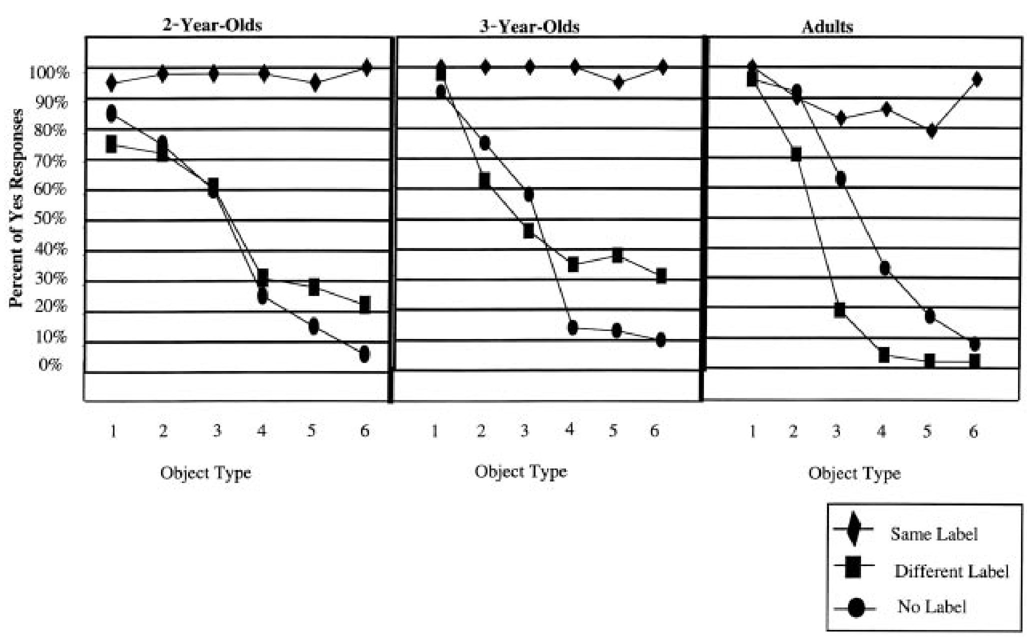

Figure 4 *Mean percentage of 'yes' responses over object in Experiment 2.*

ling condition, object and age did not reach significance, Figure 4 suggests that the difference between the Different Label and No Label conditions may have been due to the adults. *Post hoc* analyses on the differences between the Different Label and No Label conditions for each object kind and within each age group showed reliable differences among 3-year-olds for Object 5, and among adults for Objects 3 and 4 (critical difference 0.98, $p < 0.05$). In essence, the only difference between the two conditions was a decline in acceptance at an earlier point in the Different Label than the No Label condition, and this was true only among adults.

In sum, there was strong evidence supporting the idea that two identical labels boost the generalization of the label to intermediate objects. There was weaker evidence supporting the idea that providing two different labels differentiates the set of objects relative to the No Label condition. In both the Different Label and No Label conditions, participants placed the boundary between the two endpoints, showing that both conditions led to splitting of the set of objects midway into two sets.

General discussion

Our results show that different patterns of lexical labelling lead to differences in the generalization of a novel name among 2-year-olds, 3-year-olds and adults. When two very different looking objects were given the same name, children and adults generalized the name to all objects between the two, filling in the intermediate space in a way that suggests that all members became equivalent (for the purposes of naming). In contrast, when the same two objects were labelled with two different labels, people generalized the label only to the named exemplar and objects most similar to it, where the entire space of objects was divided between the endpoints. The differences in generalization patterns confirm that, by 2 years of age, children are sensitive to differing distributions of lexical labels across exemplars, and that the distribution itself can have profound consequences on the subsequent grouping of objects under a single label. The lack of age difference in this pattern is surprising, given the presence of developmental 'sharpening' effects reported in many other studies (e.g. Landau *et al.*, 1992, 1997; Keil, 1994). The stability of the pattern for the Same Label condition suggests that children may bring to word learning a deep principle specifying that two objects having the same name belong in the same lexical category.

Two mechanisms were proposed to account for the difference in generalization patterns between the Same Label and Different Label conditions. There was strong evidence that the Same Label condition 'boosted' generalization to the intermediate objects. There was only weak evidence that the Different Label 'differentiated' the objects with respect to their possible names, at least in comparison to providing no label at all; and such an effect appeared primarily among adults. The evidence for boosting was strong, since the Same Label condition was the only condition that led to accepting all intermediates as members of the lexical category. Providing the same label to the endpoints of the set rendered all objects equivalent for the purposes of naming, and strongly so, relative to the Different Label and the No Label condition. Recent evidence is consistent with this finding, in suggesting that labelling multiple objects with the same name can promote 4-year-olds' search for taxonomic similarities (Gentner & Namy, 1999). Similarly, providing a single property for two very different kinds (e.g. dogs and rats) leads people to infer that the property holds for a larger class of animals, each belonging to the same category (Smith, Shafir & Osherson, 1993; see also Carey, 1985). The 'acquired equivalence' resulting from the application of two identical labels seems analogous to the acquisition of other kinds of linguistic categories. For example, in acquiring phonemic categories, infants must learn to set boundaries for phonemic categories, and then disregard within-category variation for the purposes of language learning.

There was weaker evidence for the differentiation of objects under the Different Label condition in comparison with the No Label condition. Providing two different labels invited people to form two separate categories, divided between the endpoints. The overall shape of the generalization pattern was similar to the No Label condition, which also invited separation of the space into two categories, divided somewhere between the endpoints. The similar overall pattern for these two conditions suggests that contrast – in one case lexical, in the other case non-lexical – invited children to set up two categories. In particular, the Different Label condition provided contrast by virtue of the two different labels. Participants were asked 'Is this an X' during each test trial, and the particular label provided served as a mental pointer to one or the other lexical category. The No Label condition provided a different kind of contrast. On each test trial, the experimenter indicated the test object and target standard and asked, 'Is this (test object) like this (standard)?', pointing to each. Such *explicit* indication of the target standard was never done in either of the labelling conditions, where the label itself served as the 'pointer' to the relevant standard(s) and its associated category. The context of the No Label condition could have provided contrast by directly pointing to the standard of interest (but *not* pointing to the other standard). Other studies have shown that pragmatic contrast can facilitate the establishment of lexical sub-categories (Waxman, Lynch, Casey & Baer, 1997); hence it is plausible that this kind of pragmatic contrast also served as a mental pointer to the category of objects organized around the standard.

The added effect of the Different Label condition was fragile, appearing primarily among adults. Their results provided suggestive evidence that using two different labels for different objects can sharpen portions of the generalization curve (in the present case, those closest to the target standard). Further tests would be necessary to firmly establish the possible differences between the two conditions, however, and whether they change over development. Certainly, it seems likely that moving the location of the second label to one of the intermediate objects (say, Object 4) would have correspondingly moved the boundary for generalization. It is an empirical question whether a similar pattern of shift could be accomplished in a No Label condition simply by shifting the location of the two comparison objects.[2] In any case, any difference between the Different Label and No Label conditions would point to a sharpening effect consequent upon the use of lexical labels, a possibility that would be consistent with the idea that linguistic input affects categorization (see for example Gentner & Namy, 1999).

The present evidence for boosting is consistent with the idea that children and adults make certain assumptions about the structure of a category space, based on individual exemplars. This is consistent with two recent statements about category structure. Shepard (1994) suggested that one principal of basic level object categorization is the assumption that members of a category fall into a *continuous* region in representational space. That is, given two separate exemplars of a category, humans assume that the two spaces overlap, such that there exists a continuous path between the two spaces. The results of our study suggest that labelling two objects with the same name can engage such an assumption, with all objects intermediate to the two exemplars assumed to be part of the category-relevant region.

Similarly, Pothos and Chater (1999) suggest that adults make the default assumption that a given category is a roughly convex region in an internal space. They show that, when adults are shown multiple exemplars of a category, they are likely to generalize to the internal convex space bounded by the exemplars, rather than to the 'negative' space surrounding the exemplars, even though it is an equally logical possibility. Further results suggest that when adults are given a single verbal description of a set of exemplars, they assume a single convex space; but when given a disjunctive verbal description of the same exemplars, they assume two separated convex regions (Pothos & Chater, in preparation). Our results produce such uniform or disjunctive spaces among children as well as adults by manipulating the distribution of labels within the naming space.

The finding of interpolation through a space in the Same Label condition makes some interesting predictions regarding patterns of erroneous labelling that might occur in children's spontaneous language. Specifically, if two items are given the same name but are actually homonyms in the language, children might, under some circumstances, generalize inappropriately to a broad set of intermediate exemplars that do not actually belong to the adult category. A striking example of such erroneous generalization was provided by Bowerman and Choi (2000), in the context of spatial language. In Dutch, the word *uit* means 'out', and is commonly used in contexts similar to English, e.g. to take something OUT of a container. Dutch, however, also uses *uit* in the context of removing one's shoes (and other articles of clothing), e.g. 'take *out* one's shoes or coat'. When adult native Dutch speakers are asked why they use 'uit' in this context, they are often surprised, and cannot make a connection between this usage and the predominant use of 'uit' in the language, suggesting that *uit* is a homonym in Dutch, with two separate meanings. Dutch children, however, see no problem in using *uit* for removing clothing, taking objects out of containers, and a host of actions that are intermediate to these two. Their generalization pattern for *uit* is therefore quite broad. In fact, it is much broader than the corresponding pattern for *out* among children learning English – where a contrast is made between taking things *out* of containers and taking clothes *off* the body. Bowerman suggests that it might be the use of the single

[2] In addition, it is possible that the effects of providing two different labels may vary depending on the nature of the objects being considered. In the present experiments, the stimulus space varied along a single complex dimension – shape. Clearly, real-world categories vary on many dimensions at once, and it is possible that the effects of providing different labels may interact with the category being studied. For example, Waxman and Senghas (1992) taught 2-year-olds two different words (e.g. horn versus flute) for two different unfamiliar objects. The children applied the first term (horn) only to the object on which it was taught; but they extended the second term (flute) to both objects, suggesting that they interpreted it as a kind of superordinate term. Thus although the present experiments revealed a separation of a unidimensional space into two mutually exclusive categories, there are apparently circumstances in which children will interpret a second label as a term at a different hierarchical level from the first.

term *uit* in broadly different contexts that leads children to assume that generalization should be very broad.

As a whole, the findings confirm the importance of lexical labelling patterns on the modulation of lexical generalization. They are consistent with abundant research suggesting that lexical generalization can be modulated by a variety of factors, including perceptual salience (Smith *et al*., 1992), top-down knowledge (Imai, Gentner & Uchida, 1994; Keil, 1994; Kemler-Nelson, 1995) and task or learning context (Medin, Goldstone & Gentner, 1993; Schyns & Rodet, 1997). They are also consistent with early cross-linguistic differences in the acquisition of lexical categories (Bowerman, 1996; Imai & Gentner, 1997), and suggest the importance of interactions between underlying similarity metrics available for generalization and patterns of lexical distribution in guiding early lexical learning. Finally, they point to strong underlying assumptions held by young children and adults regarding the shape of underlying category spaces.

Acknowledgements

The research reported here was supported by grants RO1 MH55240 and 2 RO1 HD28675. We thank Nicole Kurz, Jesse Yim and Andrea Zukowski for assistance in this research. We also thank the children and parents of the University of Delaware Preschool, the Irvine Montessori and the Moms Club of Delaware for their participation.

References

Baldwin, D. (1992). Clarifying the role of shape in children's taxonomic assumption. *Journal of Experimental Child Psychology*, **54**, 392–416.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley, CA: University of California Press.

Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition*, **60**, 1–29.

Bowerman, M. (1996). Learning how to structure space for language: a crosslinguistic perspective. In P. Bloom, M.A. Peterson, M. Garrett & L. Nadel (Eds), *Language and space* (pp. 385–436). Cambridge, MA: MIT Press.

Bowerman, M., & Choi, S. (2000). Shaping meanings for language: universal and language specific in the acquisition of spatial semantic categories. In M. Bowerman & S.C. Levinson (Eds), *Language acquisition and conceptual development*. Cambridge: Cambridge University Press.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a Stone-Age tribe. *Nature*, **398**, 203–204.

Gentner, D., & Namy, L. (1999). Comparison in the development of categories. *Cognitive Development*, **14**, 476–513.

Imai, M., & Gentner, E. (1997). A cross-linguistic study of early word meaning: universal ontology and linguistic influence. *Cognition*, **62** (2), 169–200.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: the role of shape similarity in early acquisition. *Cognitive Development*, **9** (1), 45–75.

Jusczyk, P. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Keil, F. (1994). Explanation, association, and the acquisition of word meaning. In L.R. Gleitman & B. Landau (Eds), *Acquisition of the lexicon* (pp. 169–198). Cambridge, MA: MIT Press.

Kemler-Nelson, D. (1995). Principle-based inferences in young children's categorization: revisiting the impact of function on the naming of artifacts. *Cognitive Development*, **10**, 347–354.

Landau, B., Smith, L., & Jones, S. (1992). Syntactic context and the shape bias in children's and adults' lexical learning. *Journal of Memory and Language*, **31**, 807–825.

Landau, B., Smith, L., & Jones, S. (1997). Object shape, object function, and object name. *Journal of Memory and Language*, **36** (1), 1–27.

Landau, B., Smith, L., & Jones, S. (1998). Object perception and object naming in early development. *Trends in Cognitive Sciences*, **2** (1), 19–24.

Markman, E., & Hutchinson, J. (1984). Children's sensitivity to constraints on word meaning: taxonomic versus thematic relations. *Cognitive Psychology*, **20**, 121–157.

Medin, D., Goldstone, R., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, **100** (2), 254–278.

Pothos, E., & Chater, N. (1999). Classification and prior assumptions about category 'shape': new evidence concerning prototype and exemplar theories of categorization. *Proceedings of the 19th Annual Conference of the Cognitive Science Society* (pp. 620–625). Mahwah, NJ: Erlbaum.

Pothos, E., & Chater, N. (in preparation). Linguistic bias in categorization.

Ratliff, F. (1976). On the psychophysiological bases of universal color terms. *Proceedings of the American Philosophical Society*, **120** (5), 311–330.

Schyns, P.G., & Rodet, L. (1997). Categorization creates functional features. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **23** (3), 681–696.

Shepard, R. (1994). Perceptual–cognitive universals as reflections of the world. *Psychonomic Bulletin and Review*, **1** (1), 2–28.

Smith, E.E., Shafir, E., & Osherson, D. (1993). Similarity, plausibility, and judgments of probability. *Cognition*, **49** (1–2), 67–96.

Smith, L.B., Jones, S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretations. *Developmental Psychology*, **28**, 273–286.

Soja, N., Carey, S., & Spelke, E. (1992). Perception, ontology and word meaning. *Cognition*, **45**, 101–107.

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

Waxman, S., & Senghas, A. (1992). Relations among word meanings in early lexical development. *Developmental Psychology*, **28** (5), 862–873.

Waxman, S., Lynch, E., Casey, K., & Baer, L. (1997). Setters and samoyeds: the emergence of subordinate level categories as a basis for inductive inference in preschool-age children. *Developmental Psychology*, **33** (6), 1074–1090.

Werker, J.F. (1994). Cross-language speech perception: development change does not involve loss. In J. Goodman & H. Nussbaum (Eds), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 95–120). Cambridge, MA: MIT Press.

Woodward, A., & Markman, E. (1998). Early word learning. In D. Kuhn & R.S. Siegler (Eds), *Cognition, perception, and language*, Vol. 2 of *Handbook of child psychology*, 5th Edn (Series Ed. W. Damon). New York: Wiley.

Received: 28 November 1999
Accepted: 21 July 2000