Labels: Category Markers or Objects Features? Or How Rocks and Stones are Different from Bunnies and Rabbits

Anna V. Fisher (fisher49@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University
5000 Forbes Ave, 335-I Baker Hall
Pittsburgh, 15213 USA

Abstract

The ability to generalize knowledge is crucial for humans, and it has long been known that labels play an important role in inductive generalization. However, the mechanism by which labels contribute to induction remains contested. According to one position, labels contribute to induction because even young children understand that labels denote category membership, and members of the same category have much in common. Support for this position comes from the finding that even very young children rely on semantically related labels (such *bunny* and *rabbit* or *puppy* and *dog*) when performing induction. However, experiments presented below demonstrate that unlike adults, children rely on semantically related labels during induction only when these labels are also strongly associated in the lexicon (and label pairs such as *bunny-rabbit* and *puppy-dog* happen to be not only semantically related but also strongly associated lexically); in the absence of strong lexical association, even well-known semantically-related words do not promote generalization.

Keywords: cognitive development, inductive generalization, language and cognition, lexical development.

Introduction

The ability to generalize from known to novel is a critical component thinking, and even early in development, words influence the way people perform inductive generalizations. For example, when two entities are referred to as "a rabbit" and one of the entities is described as having a particular property (e.g., "it eats grass"), even young children are more likely to generalize this property to other entities with the same name than to entities referred to by a different name (Gelman & Markman, 1986; Sloutsky & Fisher, 2004; Welder & Graham, 2001). Why do people rely on words when performing induction? There are several theoretical explanations of this phenomenon.

One possibility is that people, including children as young as 2,5 years of age, treat linguistic labels presented as count nouns as special properties that differ from other properties of objects: even young children realize that count nouns denote categories, communicating what things are. Based on (1) the knowledge of what categories objects belong to (communicated by shared labels), and (2) the assumption that things belonging to the same category share many important properties, people conclude that objects referred to by the same name have many properties in common (Jaswal, 2004; Welder & Graham, 2001; Keil, et al, 1998; Gelman & Coley, 1991; Gelman & Markman, 1986).

The main support for this position comes from the innovative research by Gelman and Markman (1986). They presented 4-year-olds with a triad task, designed to put appearance similarity in conflict with category membership information. Specifically, children were presented with a Target item and two Test items, such that Test 1 belonged to the same category as the Target but looked dissimilar from the Target, whereas Test 2 looked similar to the Target but belonged to a different category. Category membership in this task was communicated by shared labels: Test 2 and the Target were referred to by the same name (e.g., "a rabbit") while Test 1 had a different name (e.g., "a squirrel"). Participants were informed that Test 1 had a particular property (e.g., "this squirrel eats bugs"), while Test 2 had a different property (e.g., "this rabbit eats grass"). Children were then asked to whether the Target item was more likely to share the property with Test 1 or Test 2. Gelman and Markman (1986) found that children were more likely to generalize the property of the test item that shared the target's category membership than the property of the test item that shared the target's appearance. This finding was interpreted as evidence that children's induction is category-based rather than similarity-based.

According to the alternative approach, shared labels influence generalization in young children by contributing to the overall perceived similarity of presented entities. In particular, under many conditions auditory input, including labels, overshadows (or attenuates processing of) corresponding visual input in infants and young children (Sloutsky & Napolitano, 2003; Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004). As a result of attenuated visual processing in the presence of auditory information, entities that share the same label may be perceived as looking more similar than the same entities presented without a label. Auditory overshadowing effects have been demonstrated to decrease with age, therefore labels should make no contribution to the perceived similarity of compared entities in adults. Both of these predictions have received empirical support (Sloutsky & Fisher, 2004; Sloutsky, Lo, & Fisher, 2001; Sloutsky & Lo, 1999).

Further evidence suggesting that early in development labels are features contributing to the overall perceived similarity rather than markers denoting membership in a common category stems from a series of studies on the

flexibility of reliance on labeling and appearance attributes. It has been argued that even young children realize that category labels are theoretically central for induction while appearances are peripheral (Jaswal, 2004; Keil, Smith, Simons, & Levin, 1998; Gelman & Coley, 1991). In this case, however, it should be difficult if not impossible to change the pattern of reliance on labels and appearances unless there is a theoretical shift. However, it has been demonstrated that reliance on labels and appearances in the course of induction can be flexibly changed by changing predictiveness of labels and appearances in the course of associative training, such that children's reliance on the attributes predictive during training increases, whereas reliance on non-predictive attribute decreases (Fisher & Sloutsky, 2006; Sloutsky & Spino, 2004). Furthermore, these training effects can be obtained only in the course of associative training but not explicit instructions to respond based on a single attribute. Finally, there is evidence suggesting that young children perceive labels as subjectively continuous properties rather than symbols. In particular, it has been demonstrated that inductive generalization is promoted not only by identical but also by phonologically similar labels (Fisher & Sloutsky, 2004; Sloutsky & Fisher, under review), which should not be the case if labels are treated as symbols denoting category membership.

However, one of the findings obtained in the modified version of the triad task has remained a challenge for the researchers arguing that early in development labels are not category markers, but features contributing to the perceived similarity of objects. In particular, in one of their studies Gelman and Markman (1986) substituted identical labels with semantically related labels, such *bunny-rabbit* and *puppy-dog*. They reasoned that if children treat linguistic labels as markers of category membership, then not only identical labels, but also semantically related labels should promote induction, and this prediction was confirmed empirically.

However, examples of synonyms provided in Gelamn and Markman's 1986 paper, consist mostly of pairs of labels that are not only semantically related, but also have a strong lexical association to each other (e.g., *bunny-rabbit*, *puppy-dog*, or *cobra-snake*). The strength of association for a pair of words is typically estimated using a free association task in which people are presented with a word and asked to generate the first word that comes to mind in response (Jenkins & Palermo, 1964). The strength of association between a pair of words is measured in terms of the probability of the second word in a pair being produced in response to the first word in a pair. Word association strength has not been extensively studied in pre-school age children, however several databases have been created to document lexical association strength in adults (Nelson, McEvoy, & Schreiber, 1998; Wilson, 1988). According to these databases, some words in English are strongly associated (for example, when presented with words *puppy*, *bunny*, and *cobra*, the probability of obtaining the words *dog*, *rabbit*, and *snake* in response is 71%, 74%, and 83% respectively), whereas other words are associated only weakly or not at all (for example, when presented with the word *child* only 4% of adults will produce the word *kid* in response). Therefore, it is possible that semantic similarity in Gelman and Markman's 1986 study was confounded with lexical association strength, and observed effects could be driven primarily by the strength of lexical association rather than by semantic similarity.

It has been demonstrated that children experience considerable difficulty when initially acquiring synonyms, but these difficulties diminish by 24 months of age (Liittschwager & Markman, 1994). However, it is unclear whether this behavioral change corresponds to the change in the mechanisms by which words are learned and ways in which words are represented in memory. One possibility is that this behavioral change is indicative of the development of the insight into symbolic nature of language and a change in the learning mechanism; another possibility is that this behavioral change arises from the same associative learning mechanisms responsible for jump-starting word learning (Smith, Jones, & Landau, 1996; for discussion see Regier, in press). If by four years of age children realize that labels are symbolic in nature and treat labels as category markers, then removing high association strength from the task (by presenting children with familiar semantically similar labels that are not strongly associated in the lexicon) should produce results consistent with those obtained by Gelman and Markman (1986). However, if four year-olds have not yet fully mastered the symbolic nature of labels, then high degree of association within a pair of semantically similar labels may be necessary to promote generalization.

These hypotheses were tested with adults and four year-old children in two experiments using a label extension task. Label extension task is a generalization task in which instead of generalizing a hidden property participants are asked to generalize object labels. In the Experiments described below participants were presented with a set of novel objects consisting of a Target and several Test items varying in the degree of similarity to the Target. Participants were then presented with a pair of labels (consisting of either strongly or weakly associated semantically similar labels), told that the first label in a pair refers to the Target, and asked to which Test item the second label in a pair was most likely to refer. In essence, participants were asked to select a test item that is of the same kind as the Target item. If participants treat labels as symbols denoting category membership, then they should be likely to generalize semantically similar labels to perceptually similar items regardless of the strength of lexical association within a pair of labels.

Experiment 1A

It is commonly believed that adults appreciate the symbolic nature of language and treat labels as markers of category membership (Yamauchi & Markman, 2000; Osherson et. al., 1990; Gelman & Markman, 1986). Therefore, performance

of adult participants in the Label Extension task should be guided by semantic similarity of word pairs, regardless of the association strength between the words in a pair.

Method

Participants Participants in Experiment 1A were 27 undergraduate students (M = 19.85, SD = 1.05; 13 women and 14 men) recruited from the introductory psychology courses at Carnegie Mellon University.

Materials Materials in Experiment 1A consisted of eight picture sets and sixteen label pairs. The picture sets were comprised of four novel objects, one of which was designated to be a Target and the rest were Test items, such that Test 1 was highly similar to the Target, Test 2 was less similar, and Test 3 was dissimilar from the Target. Half of the picture sets were comprised of novel artificial objects and the other half of novel natural kind-like stimuli (see Figure 1 for examples of picture sets). The position of the Test items relative to the Target item (directly below, below and to the left, or below and to the right) was randomized for each trial.

Label pairs used in Experiment 1A consisted of eight pairs of semantically similar labels with mean association strength of 0.41 (in the Associated Semantically Similar Labels condition) and eight pairs with mean association strength of 0.03 (in the Non-associated Semantically Similar Labels condition). The list of synonym pairs used in the experiment is presented in Table 1.

Table 1: Label pairs used in Experiment 1A.

Associated Semantically Similar Labels	Association Strength	Non-Associated Semantically Similar Labels	Association Strength
Bunny-Rabbit	0.74	Couch-Sofa	0.08
Puppy-Dog	0.71	Child-Kid	0.04
Street-Road	0.35	Father-Dad	0.05
Toad-Frog	0.34	Rock-Stone	0.03
Pony-Horse	0.32	Hat-Cap	0.02
Kitty-Cat	0.31	Cup-Mug	0.02
Jacket-Coat	0.30	Jelly-Jam	0.02
Ship-Boat	0.20	Cake-Pie	0.01

Design and Procedure There were two between-subject conditions in Experiment 1: Associated Semantically Similar Labels and Non-Associated Semantically Similar Labels conditions. Participants were randomly assigned to these conditions and label pairs were randomly assigned to the pictures sets. The order of trials was randomized for each participant.

Participants were interviewed individually in a laboratory on campus. Picture sets were presented on a computer screen and labels were supplied verbally by the experimenter. Participants were told they would be asked a set of questions about objects from "far away places". The experimenter then informed participants, for example, that "in a far away place" the Target object is called a "bunny", and asked which of the Test items would be called a "rabbit" in the far away place.

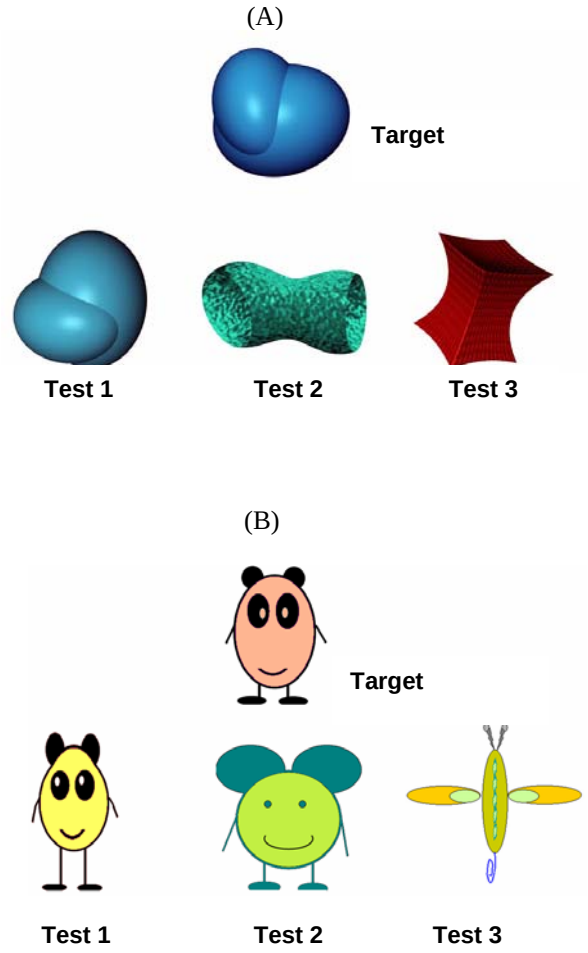

Figure 1: Examples of object sets used in Experiment 1: (A) novel artifacts and (B) novel natural kind-like items.

Results

Proportion of choices of each Test item was calculated for each participant and averaged across participants. Data presented in Figure 2 indicate that in both labeling conditions participants were highly likely to generalize semantically related labels to the most similar Test item in the set, Test 1 (91% and 81% of generalizations in the Associated and Non-Associated Synonyms conditions, respectively, above chance, both one-sample ts > 7.4, Bonferroni adjusted ps < .001), and unlikely to generalize semantically related labels to either Test 2 (8% and 15% respectively, below chance, both one-sample ts > 3.9,

Bonferroni adjusted *p*s < .001) or Test 3 (1% and 4% respectively, below chance, both one-sample *t*s > 7.5, Bonferroni adjusted *p*s < .001). There was no difference in the proportion of Test 1 choices for novel artifacts and novel natural kind-like items (both types of stimuli averaged 91% and 81% of Test 1 choices respectively in the Associated and Non-Associated Semantically Similar labels conditions, both paired-samples *t*s < 1).

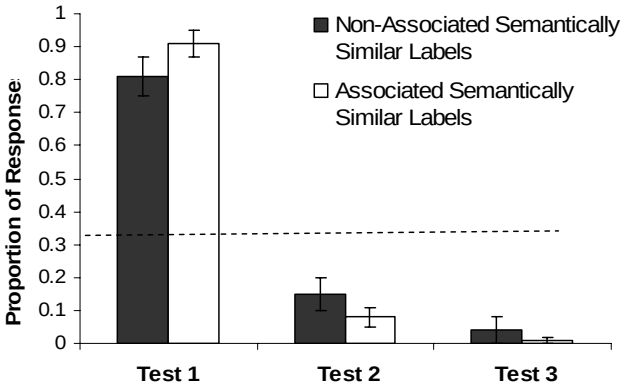

Figure 2: Proportion of responses in Experiment 1A by test item type and labeling condition. The dashed line represents the chance level.

It is possible however, that adult participants failed to take the task seriously and simply matched the Target to the most similar test item in the test set regardless of the semantic similarity of labels. This possibility was addressed in Experiment 1B.

Experiment 1B

To control for this possibility that in Experiment 1A adults matched similar pictures regardless of presented labels, in Experiment 1B participants were presented with the same Label Extension task as in Experiment 1A, however semantically related label pairs were substituted for the pairs of labels that were not related either lexically or semantically (mean association strength between labels used in Experiment 1B was 0.01; the list of labels used in this experiment is presented in Table 2).

Method

Participants Participants in Experiment 1B were 16 undergraduate students (*M* = 19.8, *SD* = .9) recruited from the introductory psychology courses at Carnegie Mellon University.

Results

When semantically similar labels were substituted with unrelated labels, the pattern of performance changed dramatically. In particular, in Experiment 1B adults were unlikely to generalize unrelated labels to similar items (averaging 9% of generalizations to Test 1, below chance, one-sample *t* (15) > 7.5, Bonferroni adjusted *p* < .001), not different from chance in generalizing unrelated items to less similar items (averaging 39% of generalizations to Test 2, *p* > .18), and likely to generalize unrelated labels to dissimilar items (averaging 51% of generalizations to Test 3, above chance, one-sample *t* (15) > 4.09, Bonferroni adjusted *p* < .001).

Table 2: Label pairs used in Experiment 1B.

Label Pair	Association Strength
Horse-Sky	0.01
House-Shark	0.01
Key-Skirt	0.01
Lamp-Cup	0.01
Purse-Ship	0.01
Rug-Baby	0.01
Bag-Bird	0.01
Spoon-Cat	0.01

These results indicate that performance of adults in Experiment 1A was driven by the semantic relatedness of labels rather than by a tendency to merely match similar pictures. Overall, results of Experiment 1A and 1B indicate that label generalization in adults is influenced by the semantic similarity of labels, with lexical association strength having no effect on performance.

Experiment 2

In Experiment 2, 4-year-old children were presented with the same task of Label Extension as adults in Experiment 1A[1]. If by four years of age children have appreciation of the symbolic nature of labels, then their pattern of performance should be similar to that of adults: in this case children should be likely to generalize semantically related labels to perceptually similar items regardless of the lexical association strength within a pair of labels. Alternatively, semantic relatedness alone may be insufficient and a strong lexical association within a pair of labels may be necessary to promote generalization of semantically similar items.

Method

Participants Participants in Experiment 2 were 34 4- to 5-year-old children (*M* = 4.6, *SD* = .44; 13 girls and 21 boys) recruited from day care centers in Pittsburgh, Pennsylvania.

[1] Experiment 2 used stimuli selected from adult databases. Collecting lexical association strength data from young children and basing stimuli selection on these data rather than adult databases, will likely improve precision of the reported findings.

Materials, Design, and Procedure Materials, design, and procedure of Experiment 2 were identical to those of Experiment 1A. A calibration experiment with a separate group of 20 4-year-old children established that all words used in Experiments 1 and 2 were familiar to children in this age group. In particular, results of a test analogous to the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997) indicated that words used as stimuli in the current experiments were correctly identified by children with the mean accuracy of 97%.

Children were randomly assigned to Associated Semantically Similar Labels or Non-Associated Semantically Similar Labels condition and presented with the Label Extension task. Participants were interviewed individually in a quiet room in their day care centers by hypothesis-blind experimenters.

Results

Results of Experiment 2 are presented in Figure 3. As can be seen in the Figure, in the Non-Associated Semantically Similar Labels condition children were equally likely to generalize labels to any item in the test set (41%, 28%, and 30% of generalizations to Test 1, Test 2, and Test 3 respectively, not different from chance, all one-tailed *t*s < 1.3, *p*s > .233). However, the pattern of responses was different in the Associated Semantically Similar Labels condition. In particular, 4 year-old children were likely to generalize semantically related associated labels to Test 1 (65%, above chance, one-sample *t* (15) > 4.5, Bonferroni adjusted p < .001), and unlikely to generalize these labels to either Test 2 or Test 3 (17% and 18% respectively, below chance, both one-sample *t*s > 3.9, Bonferroni adjusted *p*s < .01). The difference in proportion of Test 1 choices for novel artifacts and novel natural kind-like items was not statistically reliable (37% vs. 47% in the Non-Associated Semantically Similar Labels condition, and 67% vs. 63% in the Associated Semantically Similar Labels condition respectively, both paired-sample *t*s < 1.5, *p*s > .16).

Overall, the pattern of label extensions observed in children in Experiment 2 is dramatically different from that of adults in Experiment 1A. In particular, adult participants were equally likely to generalize semantically similar labels to perceptually similar items regardless of the lexical association strength within label pairs. Unlike adults, 4-year-old children were likely to generalize semantically similar labels to perceptually similar items only if the labels were also strongly associated in the lexicon. Therefore, it appears that appreciation of the symbolic nature of labels continues to develop beyond the four years of age.

General Discussion

Overall, results presented above indicate that adults overwhelmingly generalize semantically similar labels to perceptually similar objects regardless of the lexical association strength (Experiment 1). Unlike adults, 4- to 5-year-old children are unlikely to generalize semantically similar labels to perceptually similar labels unless these labels are not only semantically related but also strongly associated in the lexicon; however, when lexical association strength within a pair of synonyms is low, generalization of semantically similar labels to perceptually similar objects does not exceed chance level (Experiment 2). These results suggest that the pattern of performance observed by Gelman and Markman (1986) may have been driven to a large degree by high association strength of used label pairs, rather than by semantic similarity alone.

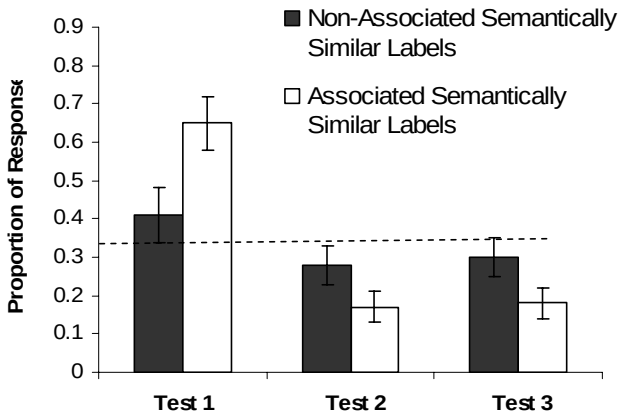

Figure 3: Proportion of responses in Experiment 2 by test item type and labeling condition. The dashed line represents the chance level.

Recall that several mechanisms have been proposed to account for the effects of labels on generalization processes. One possibility is that labels influence generalization by communicating semantic information, such as category membership (e.g., Gelman & Coley, 1990; Gelman & Markman, 1986, Gelman, 2003, Waxman, 2003). According to this proposal semantically similar labels promote generalization because children as well as adults realize the symbolic nature of labels. Alternatively, early in development linguistic labels may be perceived as object features rather than symbols denoting category membership information (Fisher & Sloutsky, 2004; Sloutsky & Fisher, 2004; Sloutsky, Lo, & Fisher 2001). According to this proposal semantic similarity alone may be insufficient to promote generalization early in development. Results of the two experiments reported above are inconsistent with the former proposal, however they lend further support to the view that insight into the symbolic nature of language is a product of learning rather than a developmental default.

Traditionally, it has been argued that word knowledge resides in a *mental lexicon* which, akin to a dictionary, contains information about word meaning (as well as other properties, such as pronunciation). However, more recently it has been suggested that rather than *having* meaning words provide *cues* to meaning, since word meaning is often context-dependent in ways that can not be easily explained by homonymy (Elman, 2004). For example, it is difficult to articulate a context-free meaning of the word “baking”,

since "baking a potato" means inducing a change of state to an existing entity by means of heating, whereas "baking a cake" means creating a completely new entity by means of heating (Elman, 2004). Therefore, word knowledge, particularly in the early stages of language acquisition, maybe influenced by the context in which words are used, and words that are used in the same context may become strongly associated. Thus, it is possible that early in development the strength of lexical association between words rather than semantic similarity is primarily responsible for promoting generalization. If this is the case, then words that are used in the same context and become strongly associated even though they do not share meaning (i.e., bread-butter, thread-needle, doctor-hospital, etc.) should promote generalization in young children (who are yet to develop an understanding of the symbolic nature of language) but not in adults (who have already developed this understanding). In other words, unlike adults, young children may judge that bread shares more in common with butter than with bagels. Therefore, while reported results demonstrate that strong lexical association is *necessary*, it may be the case that it is *sufficient* to promote generalization early in development. This possibility remains to be addressed in future research.

References

Dunn, L.M., & Dunn, L.M. (1997). *Peabody Picture Vocabulary Test - Third Edition*. American Guidance Services Publishing.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Science*, 8, 301-306.

Fisher, A. V., & Sloutsky, V.M. (2006). Flexible Attention to Labels and Appearances in Early Induction. In R. Sun and N. Miyake (Eds.) *Proceedings of the XXVIII Annual Conference of the Cognitive Science Society*.

Fisher, A. V., & Sloutsky, V. M. (2004). When mats meow: Phonological similarity of labels and induction in young children. *Proceedings of the XXVI Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Gelman, S. A. (2003). *The essential child*. Oxford: University Press.

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

Gelman, S. A., & Coley, J. D. (1991). Language and categorization: The acquisition of natural kind terms. In S. Gelman, Susan, E. Byrnes, and P. James (Eds.) *Perspectives on language and thought: Interrelations in development*. New York: Cambridge University Press.

Jaswal, V. K. (2004). Don't believe everything you hear: Preschoolers' sensitivity to speaker intent in category induction. *Child Development*, 3, 279-300.

Jenkins, J. J., & Palermo, D. S. (1964). *Word Association Norms*. Minn.:University of Minnesota Press.

Keil, F. C., Smith, W. C., Simons, D. J., & Levin, D. T. (1998). Two dogmas of conceptual empiricism: Implications for hybrid models of the structure of knowledge. *Cognition*, *65*, 103-135.

Liittschwager, J. and Markman, E. (1994). Sixteen and24-month-olds' use of mutual exclusivity as a de-fault assumption in second label learning. *Developmental Psychology*, 30, 955–968.

Napolitano, A. C., & Sloutsky, V. M. (2004). Is a picture worth a thousand words? Flexible attention in young children. *Child Development*, *75*, 1850-1870.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/

Osherson, D. N., Smith, E. E., Wilkie, O, Lopez, A, & Shafir, E. (1990). Category-based induction. *Psychological Review*, 97, 185-200.

Robinson, C.W., & Sloutsky, V.M. (2004). Auditory Dominance and its change in the course of development. *Child Development*, *75*, 1387-1401.

Regier, T. (in press). Emergence of words. *Cognitive Science*.

Sloutsky, V. M., & Fisher, A. V. (under review). Linguistic Labels: Conceptual Markers or Object Features?

Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General*, *133*, 166-188.

Sloutsky, V. M., & Lo Y-F. (1999). How much does a shared name make things similar? Part 1. Linguistic labels and the development of similarity judgment. *Developmental Psychology*, 35, 1478-1492.

Sloutsky, V. & Napolitano, A. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*, *74*, 822-833.

Sloutsky, V. M., & Spino, M. (2004). Naive theory and transfer of learning: When less is more and more is less. *Psychonomic Bulletin & Review*, 11, 528-535.

Sloutsky, V. M., Lo, Y.-F., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels and the development of inductive inference. *Child Development*, *72*, 1695-1709.

Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition*, 60, 143-171.

Waxman, S. R. (2003). Links between object categorization and naming: Origins and emergence in human infants. In Rakison, D., & Oakes, L. M. (Eds.), *Early category and concept development: Making sense of the blooming, buzzing confusion* (pp. 213-241). New York: Oxford University Press.

Welder, A. N., & Graham, S. A. (2001). The influences of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development*, *72*, 1653-1673.

Wilson, M.D. (1988) The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. http://www.psych.rl.ac.uk/

Yamauchi, T., & Markman, A.B. (2000). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 776-795.