

Development of Relational Reasoning with Semantically Similar Labels

Sheela Ramesh (sheelar@andrew.cmu.edu)

Department of Psychology, Baker Hall 342c, 5000 Forbes Ave
Pittsburgh, PA 15213 USA

Bryan J. Matlen (bmatlen@andrew.cmu.edu)

Department of Psychology, Baker Hall 445b, 5000 Forbes Ave
Pittsburgh, PA 15213 USA

Anna V. Fisher (fisher49@andrew.cmu.edu)

Department of Psychology, Baker Hall 342c, 5000 Forbes Ave
Pittsburgh, PA 15213 USA

Abstract

The present study explored children's ability to utilize synonymous labels during relational reasoning. In Experiment 1, 4- to 5-year-old children and adults were presented with a base pair of related words (e.g., Castle:Rock) and then were given a partially completed target word-pair (Castle:?) that they could complete with a label that made the target word-pair relationally identical to the base word-pair (e.g., Stone). Additional response options included a label thematically related to the first word in the target pair (e.g., King) or an unrelated word (e.g., Milk). Results indicated that adults and 5-year-olds successfully completed the task, whereas 4-year-olds exhibited difficulty. In Experiment 2, 4-year-old children were presented with the same task, however relations were conveyed by identical rather than synonymous labels. Under these conditions, 4-year-old children exhibited no difficulty in either lure condition. These findings are discussed with regards to the theories of learning early in development.

Keywords: Synonyms; Language Acquisition; Word Learning; Relational Reasoning; Cognitive Development.

Introduction

Many objects in the world can be referred to by more than one label, a phenomenon called polyonymy. For example, one could accurately refer to a pet as *Fluffy*, *kitty*, *cat*, and *animal*. It is well-documented that in the beginning stages of language acquisition children struggle with this phenomenon; however by three years of age children are able to learn multiple labels in reference to the same object, both in the form of taxonomically-related labels (such as *cat-animal*) and semantically similar labels at the same level of taxonomic hierarchy (i.e. synonyms, such as *kitty-cat*) (Banigan & Mervis, 1988; Blewitt, 1994; Deák & Maratsos, 1998; Haryu & Imai, 1999; Johnson, Scott, & Mervis, 1997; Liitschwager & Markman, 1994; Mervis et al., 1994). Learning to refer to an object by more than one label may signify development of understanding that labels denote categories rather than individual objects. However, mature understanding of labels as category markers requires that one not only can use multiple labels in reference to the same object, but is also able to rely on related labels to perform a variety of reasoning tasks, such as categorization, inductive reasoning, and analogical reasoning.

Research investigating development of the ability to use hierarchically-related labels in reasoning tasks indicates that this ability does not mature until 7- to 8-years of age (Gelman & O'Reilly, 1988; Johnson, et. al., 1997). However, research into children's ability to use semantically similar labels at the same hierarchical level – or synonyms – in reasoning tasks has produced mixed results. In particular, Gelman and Markman (1986) observed that 4-year-old children can perform inductive reasoning tasks at above chance level with identical labels (e.g., generalizing a property from one rabbit to another rabbit, rather than from a squirrel to a rabbit) as well as semantically similar labels (e.g., generalizing a property from a bunny to a rabbit, rather than from a squirrel to a rabbit).

However, while Gelman and Markman's study provides valuable insight into children's reasoning with semantically similar labels, several factors warrant further investigation of this phenomenon. First, some stimuli used in this study included taxonomically-related labels (e.g. *rose-flower* and *cobra-snake*) rather than synonyms. Second, some of the semantically similar labels used in this study are likely to co-occur in the speech of children and their caregivers as compound noun-phrases (e.g., *bunny-rabbit* and *puppy-dog*) according to the CHILDES database (MacWhinney, 2000). For example, the word "bunny" occurred in CHILDES 803 times, the word "rabbit" occurred 579 times, and these words co-occurred 103 times. At the same time, other semantically similar labels used in the Gelman and Markman (1986) study (e.g., *rock-stone*) never co-occurred in the CHILDES database. It is possible that effects of semantic similarity of labels on inductive reasoning in 4-year-old children were amplified by co-occurrence probability of some of the label pairs used in this research. In support of this hypothesis, Matlen and Fisher (2008) found that 4-year-old children successfully relied on semantically similar labels in a property induction task only if these labels were likely to co-occur in child-directed speech, whereas children's performance with non-co-occurring semantically similar labels was not different from chance. Therefore, the extent to which young children can rely on semantically similar labels in the course of reasoning tasks remains unclear.

The goals of the research reported below were two-fold. The first goal was to explore to what extent young children are capable of using semantically similar labels that span the same level of taxonomic hierarchy as a basis for reasoning. Prior research on this topic has primarily been concerned with children's reasoning within property induction tasks (Gelman & Markman, 1986; Matlen & Fisher, 2008). To assess the robustness of children's reasoning with semantically similar labels, we employed a relational reasoning task where relations were conveyed by synonyms. In contrast to property induction tasks, relational reasoning tasks have typically been utilized in research aimed at assessing children's analogical thinking (see Goswami 1991 for review). These tasks tend to follow the format of A:B::C?:. For example, Goswami and Brown (1990) assessed 4- and 5-year-olds' ability to perform analogical reasoning tasks with familiar relations by presenting them with a base word-pair (e.g., Spider-Web) and an incomplete target word-pair (e.g., Bee-?). Children could complete the target word-pair with a relational choice (e.g., Hive), or with a word that did not preserve the relation specified in the base word-pair: a thematic lure (e.g., Honey). Goswami and Brown found that by four years of age children could correctly complete these analogies based on the relational choice, even in the presence of a thematic lure. It follows then that if young children possess the ability to reason with semantically similar labels, then they should be able to correctly complete relational reasoning tasks when semantically similar labels convey the relations.

The second goal of the present research was to examine children's understanding of linguistic labels as markers of category membership. It has been suggested that understanding of labels as category markers develops as early as two years of age (Gelman & Coley, 1991; Welder & Graham, 2001). Therefore, children realize that objects referred to by the same label – or by semantically similar labels – refer to objects of the same kind, and that objects of the same kind are likely to have many properties in common. However, it has recently been argued that understanding of labels as referents to categories may have a more protracted developmental course than previously believed (Sloutsky, Lo, & Fisher, 2001; Sloutsky & Fisher, 2004). In particular, children may rely on identical labels in reasoning tasks because labels are features contributing to the overall similarity of presented entities, and identical labels increase the *perceived* similarity of compared objects. At the same time, children may not rely on semantically similar labels in the course of reasoning tasks, unless there are factors other than shared meaning (such as co-occurrence probability) that promote label-based inference.

To achieve the goals outlined above, the present study utilized a modified analogical reasoning task of the type employed by Goswami and Brown (1990). Specifically, participants were presented with a base word-pair relation (e.g. Castle:Rock), and a partially completed target word-pair relation (Castle: ?). Participants could complete the target word-pair with a label that preserved the relationship

specified in the base word-pair (i.e., a label semantically similar to the second term in the base word-pair, such as "Stone") or with a word that did not preserve the relation specified in the base word-pair: a word thematically related to the first term in the target word-pair (e.g., King; in the Thematic Lure condition) or an unrelated word (e.g., Milk; in the Unrelated Lure condition). Thus, this task followed an A:B::A:B' format (where B'-term was semantically similar to the B-term). Children's ability to perform the relational reasoning A:B::A:B' task using semantically similar labels was compared to their ability to perform this task using identical labels.

If children have acquired mature understanding that identical as well as semantically similar labels refer to objects of the same kind (Gelman & Markman, 1986; Gleman & Coley, 1991; Jaswal, 2004), then children should have little difficulty in completing the relational reasoning A:B::A:B' task using identical as well as semantically similar labels. However, if understanding that labels refer to categories has a protracted developmental course and is not yet complete by four years of age (Fisher, in press; Matlen & Fisher, 2008; Sloutsky & Fisher, 2004) then 4-year-old children, unlike adults, may have difficulty in completing relational reasoning tasks using semantically similar labels. At the same time, children should succeed in completing relational reasoning tasks using identical labels.

Experiment 1

Method

Participants

Participants were 35 four-year-old children ($M = 4.43$ years, $SD = .30$ years; 19 females and 16 males), 30 five-year-old children ($M = 5.5$ years, $SD = .32$ years; 12 females and 18 males), and 45 adults ($M = 19.94$ years, $SD = 1.14$ years; 26 females and 19 males).

Design

For the purpose of brevity, semantically similar labels will be referred to as "synonyms" henceforth. Experiment 1 had a two (Lure type: Thematic vs. Unrelated) by four (Age: 4-year-old, 5-year-olds, and adults) between-subjects design. Participants were randomly assigned to one of the two experimental conditions: the Thematic Lure or the Unrelated Lure condition.

Materials

Materials consisted of 12 picture sets presented on a computer screen, accompanied by 12 label sets provided by the experimenter (see Table 1 for the list of labels used in the experiment). Each picture set consisted of a series of four pictures: the first picture contained four doors in two rows of two. One by one, the first three doors disappeared to reveal objects hidden behind them (see Figure 1 for a schematic depiction of the task). As the doors disappeared to reveal hidden objects, the experimenter labeled each object. The objects behind the first and second doors had a

clear relationship, and the object behind the third door was identical in picture and label to the object behind the first door. The fourth door revealed no hidden object, but instead the experimenter provided two response options for the participant to guess the final object.

Table 1: Labels provided during task.

| | | B'-term: | | |
|--------|--------|----------------|----------------|----------------|
| A-term | B-term | Synonym Choice | Thematic Lure | Unrelated Lure |
| Bread | Jam | Jelly | Crumbs | Foot |
| Hand | Mitten | Glove | Foot | Ant |
| Castle | Stone | Rock | King | Milk |
| Cat | Couch | Sofa | Milk | Banana |
| Fly | Toad | Frog | Ant | Phone |
| Cheese | Rat | Mouse | Cracker | Dress |
| Apple | Belly | Tummy | Banana | Puppy |
| Duck | Lake | Pond | Feathers | TV |
| Vacuum | Carpet | Rug | Mop | Lion |
| Water | Ship | Boat | Fish | Cookie |
| Beach | Ocean | Sea | Sand | Chair |
| Car | Road | Street | Steering Wheel | Clock |

In both the Unrelated Lure and the Thematic Lure conditions the relational choice was communicated by a label that was synonymous to the B-term of the base word-pair. For example, in the trial where the base word-pair consisted of the words “Castle:Rock”, the relational choice consisted of the word “Stone”. Note that half of the participants received the base word-pair of “Castle-Rock” with the word “Stone” being the relational response option, whereas the other half of the participants received the base word-pair of “Castle-Stone” with the word “Rock” being the relational choice (the B- and B'-terms alternated in this manner for all of the trials in this and other experiments described in this paper). To avoid the potential confound of co-occurrence probability influencing children’s responses, only synonyms that never co-occurred in child-directed speech according to the CHILDES database (MacWhinney, 2000) were chosen for this study. Thus, common synonym pairs used in prior research, such as puppy-dog and bunny-rabbit, were not utilized. The outcome measure was the proportion of relational responses (i.e. choosing synonymous labels over thematic and unrelated lures) across the 11 experimental trials.

Calibration of Experimental Materials Experimental materials used in this research were calibrated in several separate studies to establish that (1) 4-year-old children were familiar with all of the labels used in this research and were willing to apply synonymous labels to the same object, (2) 4-year-old children were familiar and could identify the relationship specified by the A-term and the B-term labels, and (3) 4-year-olds children perceived labels chosen as thematic lures to be thematically related to the A-term labels.

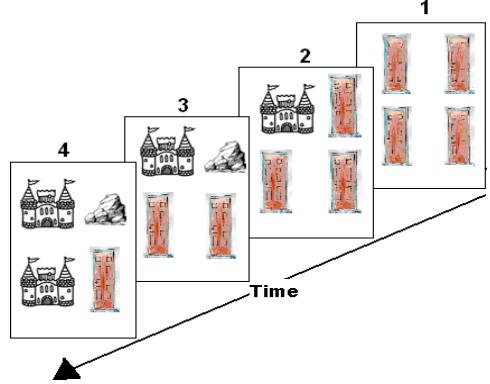


Figure 1: Schematic depiction of the A:B::A:B' task.

Label Calibration A group of four-year-old children, none of whom participated in the experiment proper ($N = 12$, $M = 4.28$ years) was presented with a picture-naming task analogous to the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997). Participants were presented with a series of pictures on a computer screen, four pictures at a time, with one target picture for children to identify (location of all target pictures was counterbalanced across multiple presentation of the same picture) and three distractors. Children were asked to select the target picture according to the label spoken by the experimenter (e.g., “Can you point to the *rock*?”). Each of the 12 target pictures was presented to children in two separate picture sets, resulting in a total of 24 trials. The target pictures were identified by different synonymous labels during the first and second presentation of each target picture. Children correctly identified pictures referred to by the B- and B'-terms used in Experiment 1 (see Table 1) with the overall accuracy of 97%. Therefore, the synonym labels used in Experiment 1 were familiar to four-year-old children. Importantly, four-year-old children were willing to accept the synonymous labels used in Experiment 1 as referents to the *same* objects.

A:B Relationship Calibration The same group of 4-year-old children who participated in the Label Calibration were presented with the relation familiarity check (the order of these tasks was counterbalanced across participants). Children were presented with a series of triads depicting objects that can be referred to by an A-term, an unrelated lure, and a B-term or a synonymous B'-term (see Table 1). Children were provided verbal labels for all three objects and asked to select the two objects that “go together.” Half of the participants were asked about a B-term (e.g., a *rock*) and the other half about the synonymous B'-term (e.g., a *stone*). For instance, participants could be shown a triad consisting of pictures of a ‘rock’, a ‘castle’, and ‘milk’, and asked which objects go together. Additionally, children were asked to explain why the two pictures they had chosen “go together,” and their explanations were recorded. All of the B- and B'- presented in Table 1 were judged to be identifiably related to the A-terms by 4-year-old children: children selected the B- or the B'-term over the unrelated lure and correctly specified the nature of the A:B

relationship with the average accuracy of 95% (all individual item M s were above 83%, above chance, all one-sample t s > 4.4 , p s $< .001$).

Calibration of Thematic Lures A separate group of 4-year-old children ($N = 12$, $M = 4.77$ years) was presented with a series of triads depicting objects referred to by the A-term, thematic lures, and unrelated lures (see Table 1) and asked to select the two objects that “go together.” For instance, participants could be shown a triad depicting a ‘castle’, a ‘king’, and ‘milk’, and asked which objects go together. Children selected the thematically related lure as the objects that “goes with the target” over the unrelated lure with the mean accuracy of 94% (all individual item M s were above 91%, above chance, all one-sample t s > 7.0 , p s $< .001$).

Procedure

Children were tested individually in a quiet room in their preschools and adults were tested in a laboratory on campus. Participants were presented with pictures of four doors in two rows of two on a computer screen. The experimenter explained that there were objects hiding behind all of the doors, and that after showing the objects behind the first three doors, the participant would have to guess what was hiding behind the last door. The word-pair “Bread:Jam” served as a practice trial (see Table 1) and thus was always presented first. The order of the rest of the trials was randomized for each participant. When pictures of bread and jam were revealed during the practice trial, participants were told, “bread and jam go together because jam goes on bread to make a sandwich.” Participants were then asked to guess what object was hiding behind the fourth door and presented with two response options; participants were asked to choose the option that “goes with the bread the same way that jam goes with the bread”. Upon completing the practice trial children were provided with corrective feedback. No feedback was provided after the experimental trials. At the conclusion of the practice trial children were told that they would keep playing the game, and that to solve the task they needed to think how the objects behind the first two doors go together.

Results

Proportions of relational responses in each condition are presented in Figure 2. A 2×2 Analysis of Variance was performed on the proportions of relational responses with experimental condition (Thematic Lure vs. Unrelated Lure) and age (4-year-olds, 5-year-olds, and adults) as between-subject factors. The results indicated the main effect of condition, $F(1, 104) = 10.9$, $p = .001$, and age, $F(2, 104) = 50.7$, $p < .001$. Performance was significantly higher in the Unrelated Lure condition ($M = 89.8\%$) than in the Thematic Lure condition ($M = 80.1\%$), $F(1, 104) = 10.9$, $p = .001$. These main effects were qualified by a condition by age interaction, $F(2, 104) = 3.4$, $p < .05$.

Post hoc Tukey tests indicated that performance increased significantly from 4- to 5- years of age ($p < .001$), and again from 5-years of age to adulthood ($p < .05$). Planned

comparisons revealed that performance of adults was equivalent in the Thematic Lure and Unrelated Lure conditions (both means over 99%). However, 5-year-old children exhibited a higher level of performance in the Unrelated Lure condition compared to the Thematic Lure condition (97% and 83% of relational responses, respectively), independent samples $t(28) = 2.98$, $p < .01$. This difference in performance between the lure conditions was marginally significant in 4-year-old children (73% and 58% of relational responses, respectively), independent samples $t(33) = 1.9$, $p = .07$.

Follow-up comparisons to chance indicated that in the Unrelated Lure condition participants in all age groups responded at a level above chance (chance = 50%), all one-sample t 's > 4.5 , p 's $< .001$. However, in the Thematic Lure condition only 5-year-olds and adults responded at above chance level, both one-sample t 's > 7.6 , p 's $< .0001$, whereas responses of 4-year-old children were not different from chance, one-sample $t(16) = 1.2$, $p = .25$.

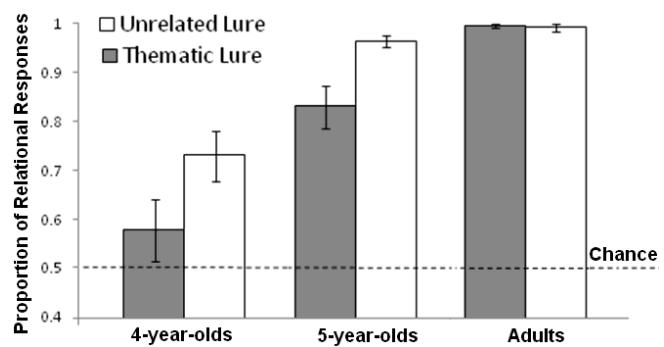


Figure 2: Proportions of relational responses in Experiment 1. Error bars represent standard errors of the mean.

To further understand the source of differences among conditions, we conducted analysis of the individual patterns of responses. Participants were judged to be relational responders if they selected the synonymous choice on at least 8 out of 11 trials (binomial $p < .05$). Proportion of relational responders in each condition is presented in Figure 3. All adult participants in both lure conditions were deemed to be relational responders. Among 5-year-old children, all participants in the Unrelated Lure condition were deemed to be relational responders, and 12 out of 15 (or 80%) participants in the Thematic Lure condition were classified as relational responders (marginally different from the pattern observed in adults, Fisher's exact $p = .058$). Among 4-year-old children, 10 out of 18 (56%) participants exhibited the relational pattern of responding in the Unrelated Lure condition, and only 5 out of 17 (29%) participants exhibited the relational pattern of responding in the Thematic Lure Condition. Proportion of relational responders in the 4-year-old group was significantly lower than that of participants in both older age groups in both experimental conditions, all Fisher's exact p s $< .006$.

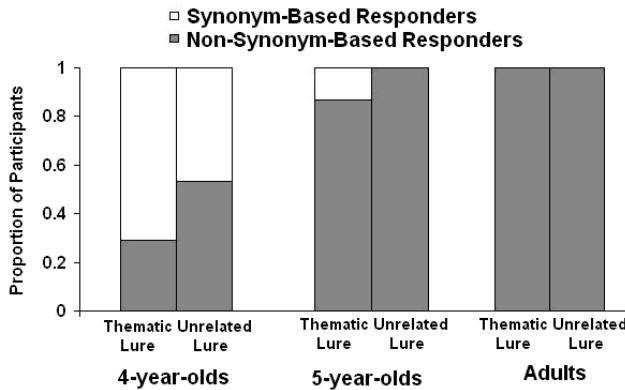


Figure 3: Individual patterns of responses in Experiment 1.

Overall, results of Experiment 1 suggest that the ability to utilize synonymous labels in the course of reasoning may follow a more protracted developmental course than it has been previously believed. Experiment 2 was designed to investigate 4-year-olds' performance in the A:B::A:B' task when identical labels were used. If by 4 years of age children realize that labels refer to kinds, then children's performance with identical labels should be similar to that with synonymous labels in Experiment 1 (Gelman & Markman, 1986). However, if 4-year-old children do not yet treat labels as category markers, their performance with identical labels may be superior to that with synonymous labels.

Experiment 2

Method

Participants

Participants in Experiment 2 were 27 four-year-old children ($M = 4.48$ years, $SD = .27$ years; 9 females and 18 males).

Design, Materials, and Procedure

Similar to Experiment 1, there were two between-subject conditions in Experiment 2: a Thematic Lure condition and an Unrelated Lure condition. Children were randomly assigned to one of the two experimental conditions. The order of trials was randomized for each participant.

Materials used in Experiment 2 were identical to those in Experiment 1, with the exception that relational choices were communicated by a label identical to the B-term, rather than by a synonymous label. The B- and B'-terms in Table 1 were counterbalanced across participants, such that half of the participants received the "Castle:Rock" base word-pair, whereas the other half of the participants received the "Castle:Stone" base word-pair.

Results

In both the Thematic Lure condition and the Unrelated Lure condition children averaged 94% of relational choices, above chance (chance = 50%), both one-sample $ts > 16.7$, $p < .0001$). Analysis of the individual patterns of responses

revealed that in both experimental conditions, 100% of participants successfully chose identical labels over unrelated as well as thematic lures.

Responses of 4-year-old children in Experiment 2 were compared to those in Experiment 1 in a 2 (Label condition: Identical vs. Synonymous labels) by 2 (Lure Type: Unrelated vs. Thematic Lures) ANOVA. This analysis revealed a main effect of the Label condition, $F(1, 58) = 34.17$, $p < .001$. The labeling condition by lure type interaction did not reach significance, $F(1, 58) = 2.55$, $p = .11$, possibly due to unequal variances in the Synonymous and Identical labels condition, Levene's test of equality of error variances: $F(3, 58) = 8.98$, $p < .0001$.

Overall, children performed significantly better when they could rely on identical labels rather than synonymous labels in both the Thematic Lure condition (94% and 73% of relational responses, respectively) and the Random Lure condition (94% and 58% of relational responses, respectively), both independent-sample $ts > 3.2$, $ps < .005$. This conclusion was supported by the analysis of the individual patterns of responses. In the Thematic Lure condition all 13 4-year-old children were classified as relational responders when children could rely on identical labels (Experiment 2), whereas only 5 out of 17 4-year-olds (29%) in the Thematic Lure condition exhibited this pattern of responding with synonymous labels (Experiment 1), Fisher's exact $p < .001$. Similarly, in the Unrelated Lure condition 13 out of 14 4-year-olds (93%) exhibited the relational pattern of responding with identical labels, whereas only 10 out of 18 4-year-olds (56%) exhibited this pattern Unrelated Lure condition with synonymous labels, Fisher's exact $p < .05$.

Discussion

The primary goal of the present study was to examine development of the ability to utilize semantic similarity of labels in a relational reasoning task. Results of the two experiments reported above point to several novel findings. First, adult participants successfully utilized semantic similarity of labels in the Semantic Completion (A:B::A:B') task and their performance was not affected by the type of lure. Importantly, all adult participants exhibited the same pattern of responding on this task. Second, 5-year-old children exhibited a decrease in performance in the Semantic Completion task in the presence of thematic lures compared to unrelated lures. At the same time, proportion of relational responses was well above chance level in 5-year-old children with both types of lures. Furthermore, the majority of 5-year-old children, similar to adults, were classified as relational responders in both lure type conditions. Third, 4-year-old children reliably made relational choices in the Semantic Completion task regardless of the type of lure when relational responses were communicated by identical labels; when relational choices were communicated by synonymous labels, 4-year-old children performed above chance only if correct responses were pitted against unrelated lures, but not thematic lures.

Finally, less than a third of 4-year-old children were classified as relational responders when response competition was strong (in the Thematic Lure condition) and only half of 4-year-olds reliably provided relational responses in the absence of response competition (in the Unrelated Lure condition).

Great care was taken to calibrate stimulus materials used in this research, therefore the reported results cannot be explained by children's unfamiliarity with the words or the relations used in the Semantic Completion and Semantic Substitution tasks. Instead, it appears that less than half of 4-year-old children spontaneously realize that synonymous labels refer to objects of the same kind and can use this knowledge in a reasoning task. This finding has important implications for different theoretical approaches to development of induction. In particular, successful performance on the Semantic Completion task with identical labels does not require that children realize that identical labels refer to objects of the same kind – children could have successfully performed the task based on simple matching of identical labels. However, successful performance on the task with synonymous labels (that do not co-occur in child-directed speech) can only be achieved if children understand that semantically similar labels refer to objects of similar kind. A large decrease in performance on the Semantic Completion task with synonymous labels compared to near ceiling performance with identical labels suggests that at four years of age many children do not yet treat labels as referents to object kind. Therefore, these results pose a challenge to the theoretical accounts of early induction that assume such understanding in very young children (Gelman & Coley, 1991; Welder & Graham, 2001) and suggest that children's understanding that labels refer to kinds continues to develop beyond toddlerhood.

Acknowledgments

We thank children, parents, and teachers for their participation in this research. This project was funded by a Small Undergraduate Research Grant at Carnegie Mellon University awarded to Sheela Ramesh.

References

Banigan, R. L. & Mervis, C. B. (1988). Role of adult input in young children's category evolution: An experimental study. *Journal of Child Language*, 15, 493-504.

Blewitt, P. (1994). Understanding categorical hierarchies: The earliest levels of skill. *Child Development*, 65, 1279-98.

Deák, G. O., & Maratsos, M. (1998). On having complex representations of things: Preschoolers use of multiple words for objects and people. *Developmental Psychology*, 34, 224-240.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test-Third Edition: Manual*. Circle Pines, MN: American Guidance Services.

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology*, 20, 65-95.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, 23, 189-209.

Gelman, S. A., & Coley, J. (1991). Language and categorization: The acquisition of natural kind terms. In S. A. Gelman, S. & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 146-196). New York: Cambridge University Press.

Gelman, S. A., & O'Reilly, A. W. (1988). Children's inductive inferences within superordinate categories: the role of language and category structure. *Child Development*, 59, 876-87.

Goswami, U. (1991). Analogical reasoning: What develops? A review of research and theory. *Child Development*, 62, 1-22.

Goswami, U. & Brown, A. L. (1990). Higher-order structure and relational reasoning: Contrasting analogical and thematic relations. *Cognition*, 36, 207-26.

Haryu, E. & Imai, M. (1999). Reorganizing the lexicon by learning a new word: Japanese children's interpretation of the meaning of a new word for a familiar artifact. *Child Development*, 73, 1378-91.

Jaswal, V. K. (2004). Don't believe everything you hear: Preschoolers' sensitivity to speaker intent in category induction. *Child Development*, 75, 1871-85.

Johnson, K. E., Scott, P. & Mervis, C. B. (1997). Development of children's understanding of basic-subordinate inclusion relations. *Developmental Psychology*, 33, 745-763.

Liitschwager, J. C., & Markman, E. M. (1994). Sixteen- and 24-month-old's use of mutual exclusivity as a default assumption in second-label learning. *Developmental Psychology*, 30, 955-968.

MacWhinney, B. (2000). The CHILDES Project: Tools for analyzing talk. Lawrence Erlbaum Associates.

Matlen, B. J., & Fisher, A. V. (2008). Development of synonym-based induction. Poster presented at the Cognitive Science Society, July 28th.

Mervis, C. B., Golinkoff, R. M. & Bertrand, J. (1994). Two-year-olds readily learn multiple labels for the same basic-level category. *Child Development*, 65, 1163-77.

Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General*, 133 (2), 166-188.

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

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.