Learning Words from Context

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
Center for Cognitive Science and Department of Psychology
The Ohio State University
208 Ohio Stadium East, 1961 Tuttle Park Place
Columbus, OH 43210, USA

Xin Yao (yao.64@osu.edu)
Department of Psychology and Center for Cognitive Science
The Ohio State University
208 Ohio Stadium East, 1961 Tuttle Park Place
Columbus, OH 43210, USA

Abstract

Word learning is a notoriously difficult induction problem because word meaning is severely underdetermined by positive examples. One way of solving this problem early in development is to use prior knowledge in a word learning situation. Some have argued that young word learners have a number of "mapping" assumptions that reduce the overall hypothesis space, whereas others have argued that young children have a combination of "sampling" and "mapping" assumptions. In this research, we demonstrate that children may rely on a different mechanism -- word learning is directed by associations among the context and a new word. These are novel findings pointing to a powerful associative mechanism underlying word learning.

Keywords: Cognitive Development, Attention, Language Acquisition, Psychology, Human Experimentation.

Introduction

The ability to extend knowledge from known to novel, or inductive generalization, and the development of this ability are interesting and controversial issues in human cognition. People deploy this ability every time they (1) extend a known word to a novel entity (i.e., in naming or label extension), (2) treat a novel entity as a member of a familiar class (i.e., categorization), and (3) extend a property from a familiar entity to a novel entity (i.e., projective induction). Inductive generalization is critically important because by some accounts it is "the only process… by which new knowledge comes into the world" (Fisher, 1935/1951, p. 7). Furthermore, some researchers have concluded that an understanding of this ability and its development can elucidate "a single main theme to cognitive science – the question of how people come to have knowledge" (Murphy, 2002, p. 272).

One difficult induction problem that has to be solved early in development is word learning. For example when an object (say a cat) is named, the word cat is compatible with an infinitely large number of possibilities: this particular cat, some cats, all cats, cats and dogs, the cat's texture, the speaker's attitude towards the cat, and many others. How do young children solve this challenging problem? One possible solution is that word learning is akin to hypothesis testing with young word learners using their knowledge about the language and the world to rule out the implausible hypotheses and to weigh the probability of the plausible ones (Markman, 1989; Xu & Tenenbaum, 2007).

Some proponents of this idea suggested that young word learners hold assumptions about word-world mappings that constrain the number of possible mappings. For example, if a young word learner assumes that words denote objects rather than parts and that words refer to taxonomic kinds rather than to thematic groupings (Markman & Hutchinson, 1984), these assumptions would substantially reduce the hypothesis space, thus facilitating word learning. Others have suggested that young learners have "sampling" assumptions as well as assumptions about the structure of categories in the word (Xu & Tenenbaum, 2007). In this case, if a word dog was introduced multiple times, each time referring to a terrier, the probability that this is a random sample from all animals (and hence the word refers to all animals) is low, whereas the probability that the word refers to only terriers is high. Therefore, according to this view, word learning is a variant of Bayesian statistical inference that capitalizes on the young word learner's intuitions about event probabilities, the sampling process, and the structure of categories.

While we agree that prior knowledge may affect future learning, we propose a different idea of how this knowledge may affect word learning. In particular, it has been well established that young word learners are sensitive to statistical regularities in the input, including linguistic input (Saffran, Aslin, & Newport, 1996) and language-world correspondences (Colunga & Smith, 2005; Yu & Smith, 2007). Statistical regularities in linguistic input are often reflected in associations between words: early in development words that have high transitional probabilities tend to be highly associated (Brown & Berko, 1960). At the same associations among taxonomically related words (e.g., *cat* and *dog*) are not present until later in development (Brown & Berko, 1960).

Therefore, a context, in which a novel word is introduced, may provide additional guidance as to what this word might mean by triggering semantic associations. For example, in a sentence "look at this dax, he is very cute", the words *cute* and *he* could activate properties associated with animacy, thus guiding learning of the word *dax*. This possibility was tested in the two reported experiments with 4-year-olds and adults.

Experiment 1

Method

Participants Participants were 30 4-year-old children (*M* = 53.1 months, *SD* = 3.6 months) recruited from day-care centers located in middle class suburbs of Columbus, Ohio and 30 undergraduates from the Ohio State University participated in the current experiment for a course credit. There were two between-subjects conditions, with 15 participants in each condition.

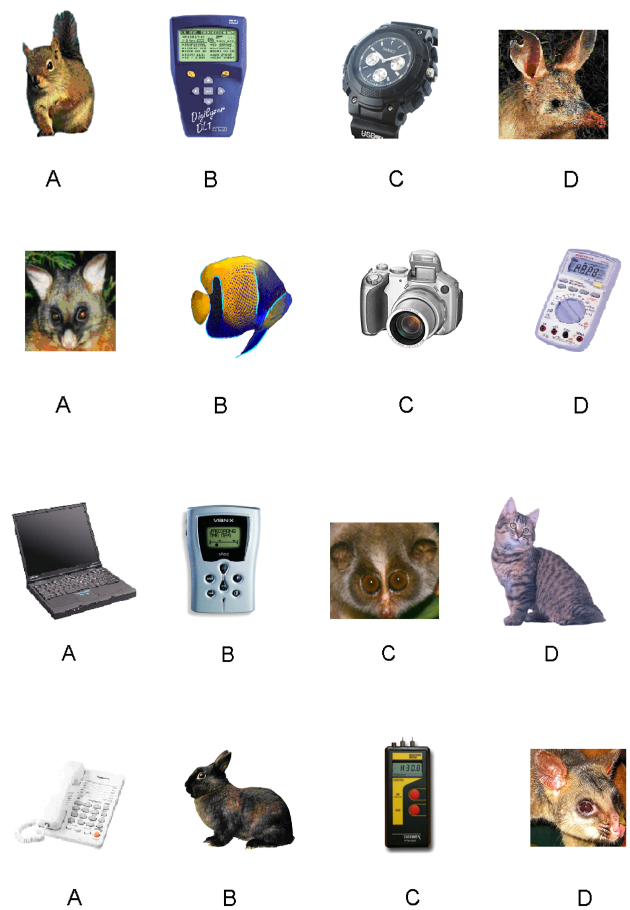
Materials and Design The experiment included two phases: study and test. During the study phase, participants were presented with an implicit word learning task, whereas the test phase included a lexical extension task. The experiment has a 2 by 2 between-subjects design, with Age (Young Children vs. Adults) and Study Phase Condition (Biological Kind vs. Baseline).

Study Phase. During the study phase participants were given an inferred word learning task. Materials included an 8cm by 13cm picture of a woman holding an ambiguous item.

The item was made ambiguous by blurring it using Adobe Photoshop software. Materials also included two novel count nouns "a dax" and "a fep". Participants were told that they will play a game in which they will learn a new word. They were then presented with the picture of the woman and the item was named. For half of the participants, the object was named "fep" and for another half, it was named "dax." There were two Study Phase conditions. In the Baseline condition, participants were told that the woman was playing with a fep/dax, whereas in the Biological Kind condition they were told that the woman was feeding a fep/dax. It was expected that in the latter condition the novel word would be interpreted as referring to a biological kind, whereas in the former condition, the novel word could be referred to a biological kind or an artifact.

Test Phase. Test phase consisted of a label extension task, in which participants were presented with a set of four 3cm by 3cm pictures. Each set consisted of two animal and two artifact pictures. There were four different sets of pictures (see Figure 1) which were used in the four label extension trials. The positions of the four items were pseudorandom and the four trials were randomized across participants.

Figure 1. Test stimuli used in Experiments 1-2.

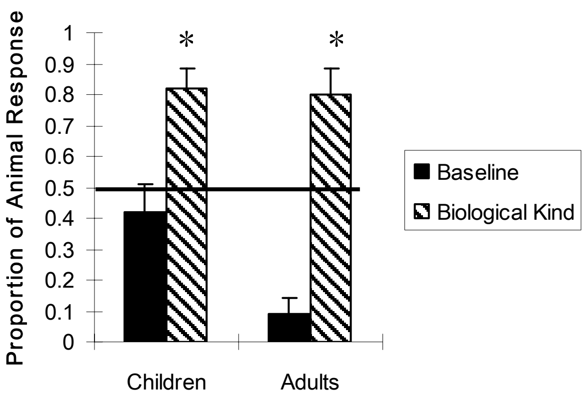


Procedure Young children were interviewed individually by a female researcher in a quiet room in their day-care center. Undergraduate students were interviewed in testing rooms on campus. For all participants, the experiment was administered using Superlab 2.0 software. No feedback was given in any of the tasks.

Results and Discussion

In this section we focus on the probability that participants interpreted the novel word as referring to a biological kind. As expected, for both children and adults, there were differences between the Baseline and Biological Kind conditions. As can be seen in Figure 2 both adults and children were significantly more likely to interpret the novel word as denoting an animal when it could be fed (i.e., in the Biological Kind condition) than when it could be played with (i.e., in the Baseline condition), for both children and adults, independent sample *ts* > 3.263, *ps* < 0.01.

Figure 2. Proportions of label extension responses by age group and study phase condition in Experiment 1. Note: * -- Above chance, p < .01.



These results clearly indicate that the presence of a feature (i.e., "being able to take food") signals to both children and adults that an item in question is an animal. This inference however could be arrived at by at least two different routes. First, it is possible that the inference is performed conceptually, with participants making a sequence of reasoning steps. (1) "Dax/Fep" can be fed. (2) If something could be fed, then it is likely to be an animal. (3) Therefore "dax/fep" is likely to be an animal. Another possibility is that the inference is performed associatively such that the feature "being fed" activates other features associated with animacy (e.g., has eyes, looks like an animal, etc.), which in turn guides extension of the word to animals. If the first possibility is the case, then presenting the participants with taxonomic information directly indicating that the "dax/fep" is an animal should result in participants interpreting the "dax/fep" as an animal. Conversely, if the second possibility is the case, then presenting participants with information associated with animacy should result in participants interpreting the "dax/fep" as an animal. This issue was addressed in Experiment 2.

Experiment 2

The goal of Experiment 2 was to examine whether performance in Experiment 1 was driven by conceptual knowledge or triggered by a set of associations. To achieve this goal, the word learning task used in Experiment 1 was modified. The word learning task was described as memory task, with participants being presented with a list of words, with a novel word ("fep" or "dax") embedded in the list. In one condition (hereafter "Taxonomic"), all words denoted familiar animals, such as *cat*, *dog*, etc. In the other condition (hereafter "Associative"), all words denoted semantic associates of the word "animal" (e.g., *fuzzy*, *farm*, *zoo*, etc.). The word learning task was followed by the label extension task that was identical to those used in Experiment 1. In addition, participants were asked to recall words presented during word learning. If participants make inference in a reasoning-type manner, they should consider novel words as referring to an animal in the Conceptual condition, whereas if they perform it in associative manner, they should consider novel words as referring to an animal in the Associative condition.

Method

Participants Participants were 58 4-year-old children (28 boys and 30 girls, *M* = 4.4 years old, *SD* = 3.6 months) recruited from day-care centers located in middle class suburbs of Columbus, Ohio and 45 undergraduates (22 women and 23 men, *M* = 19.3 years old, *SD* = 1.7 years) from the Ohio State University participated in the current experiment for a course credit. There were three between-subjects conditions, with comparable numbers of participants in each condition.

Materials and Design. Similar to Experiment 1, Experiment 2 included two phases: study and test. During the study phase, participants were presented with an implicit word learning task, whereas the test phase included a lexical extension task.

Study Phase. Word Learning was introduced as a memory task and it included three between-subjects conditions: Taxonomic, Associative, and Control. The experimenter read a word list to the participants twice and participants were told explicitly to recall the words later. There were nine words in the list, eight of which familiar words (for young children familiarity of each word established by consulting the MRC Psycholinguistic Database -- http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). There was also a novel word (either fep or dax) added to each list. In the Taxonomic condition, the familiar words belonged to the category of animal, such as *cat*, *dog*, *fish*, *bird*, etc. In the Associative condition the familiar words had forward semantic association with the word *animal*, such as, *zoo*, *farm*, *furry*, etc. Finally, in the Control condition, the familiar words denoted color terms. For adults, the task was introduced as a memory task, whereas for young children, it was introduced as a memory game.

Taxonomic, Associative, and Control lists used in the word learning phase (Semantic associations with the word "Animal" is in parentheses for children and adults, respectively.

Taxonomic List: *Cat* (0; 0.02), *Dog* (0; 0.01), *Fish* (0; 0.01), *Bird* (0; 0), *Horse* (0; 0.04), *Squirrel* (0; 0.04), *Cow* (0; 0.02), *Rabbit* (0; 0.01).

Associative List: *Creature* (0.09; 0.26), *Farm* (0.20; 0.19), *Giraffe* (0; 0.09), *Hamster* (0; 0.13), *Bear* (0; 0.09), *Furry* (0.07; 0.21), *Feeding* (0.09; 0.02), *Zoo* (0.39; 0.51).

Control List: *White* (0, 0), *Red* (0, 0), *Blue* (0, 0), Purple (0, 0), *Pink* (0, 0), *Black* (0, 0), *Green* (0, 0), *Yellow* (0, 0).

For 4-year-olds semantic associations were established by presenting 50 participants with a free association task, whereas for adults these associations were found in the Edinburgh Word Association Thesaurus database (http://www.eat.rl.ac.uk/), which was developed by presenting adults with a free association task. In addition, a novel word was inserted in each of the lists, with the serial position of the novel word being randomized.

Test Phase. In the test phase participants were presented with a label extension task, which consisted of four trials. On each trial, participants were shown a set of four 3cm by 3cm pictures, depicting a familiar animal, a novel animal, a familiar artifact, and a novel artifact and were asked which one was a dax. There were four different sets of pictures (see Figure 1) which were used in the four label extension trials. The positions of the four items were pseudorandom and the four trials were randomized across participants.

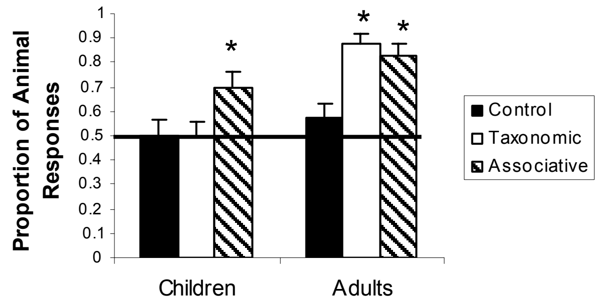
Young children were interviewed individually by a female researcher in a quiet room in their child-care center. Undergraduate students were interviewed individually in testing rooms on campus. For all participants, the experiment was administered using Superlab 2.0 software. No feedback was given during learning or testing.

Results and Discussion

Results are presented in Figure 3. In the Control condition, adults were equally likely to extend the novel word to either an animal or an artifact, thus exhibiting chance performance ($p > .65$). At the same time, in both the Associative and Taxonomic conditions, they were likely to extend the novel word to an animal, (83% and 88%, respectively, both exceeding chance performance, one-sample ts > 3.6, ps < .005.

Similar to adults, young children in the Control condition, were equally likely to extend the novel word to either an animal or an artifact, thus exhibiting chance performance ($p > .88$). However, in contrast to adults, their extension of novel words to animals in the Taxonomic condition (40%) did not exceed chance ($p > .85$). At the same time, they reliably extended the novel word to an animal only in the Associative condition (70%, above chance and above the Taxonomic condition, both ps < .05).

Figure 3. Proportions of label extension by age and study phase condition in Experiment 2. Note: * -- Above chance, p < .05. ** -- Above chance, p < .005



General Discussion

Several important regularities emerge from reported experiments. Experiment 1 indicated that similar to adults, young children extract the meaning of the word from the context in which this word is presented. Experiment 2 indicated that whereas adults can rely on associative and taxonomic information when learning new words, young children can rely only on associative information.

The fact that children learned words only in the Associative, but not in the Taxonomic condition is difficult to reconcile with word learning as hypothesis testing. Recall that according to the constraints and biases variant of hypothesis testing, children assume that words refer to categories rather than to thematic groupings (e.g., Markman & Hutchinson, 1984), whereas according to the Bayesian approach, they estimate the likelihood of observed evidence given a hypothesized word meaning (e.g., Xu & Tenenbaum, 2007). According to either approach, it is much likely that a novel word refers to an animal if it is presented in a list where every word refers to an animal (e.g., *cat*, *dog*, *tiger*, etc.) than when most of the words do not refer to an animal (e.g., *zoo*, *farm*, *furry*, *creature*, etc.). The results also point to differences in word learning in children and adults, which should not be the case if children and adults learn words by hypothesis testing guided by the same set of assumptions.

Results of the reported studies support an idea that early in development words can be learned from the context, thus suggesting a powerful associative mechanism of word learning that is not reducible to previously proposed ideas of word learning as hypothesis testing. One way of how this mechanism may work is that the context in which words are presented (e.g., familiar words co-occurring with a novel word) activates features associated with to-be-learned words, and labels are extended to items that have these activated features. This mechanism will have to be further tested in future research.

Acknowledgments

This research has been supported by grants from the NSF (REC 0208103) and from the Institute of Education Sciences, U.S. Department of Education (R305H050125) to Vladimir M. Sloutsky.

References

Brown, R., & Berko, J. (1960). Word association and the acquisition of grammar. *Child Development*, *31*, 1–14.

Fisher, R. (1951). *The design of experiments*. New York: Hafner Publishing Company. (Original work published in 1935).

Colunga, E. & Smith, L.B. (2005) From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*, *112*, 347-382.

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

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *8*, 561–577.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*, 245–272.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*, 414-420.