

Attentional Highlighting as a Mechanism behind Early Word Learning

Hanako Yoshida (yoshida@uh.edu)

Department of Psychology, University of Houston, 126 Heyne Building
Houston, TX 77204-5022 USA

Rima Hanania (rhanania@indiana.edu)

Psychological and Brain Sciences and Program in Cognitive Science, Indiana University
Bloomington, IN 47405 USA

Abstract

This paper investigates the role of general attention-shifting mechanisms in children's early adjective learning. A novel version of an adjective-learning paradigm was used to probe whether attentional mechanisms could account for early adjective learning without recourse to high-level syntactic knowledge. Two- and 3-year-olds participated. One condition removed all syntactic information by presenting words in incorrect syntactic order, but with explicit naming of known properties. According to the attention-shifting account, activation of past learned associations through the mention of known words should focus attention on the right association: novel adjective to novel property. Younger children with less linguistic experience successfully learned novel adjectives in this condition, supporting the attentional account. We discuss the attention-shifting mechanism by analogy to attentional highlighting, which accounts for the phenomenon of highlighting in the adult associative learning literature.

Keywords: Attentional learning; highlighting phenomena; early word learning; selective attention.

Attentional mechanisms in early word learning

Attention to relevant information is central to all learning. Selectively attending to appropriate information enables quick learning, generalization to new situations, and successful decision making; attending to irrelevant information, on the other hand, can lead to error and to a failure to learn. In language learning, mapping words to referents is a key process of learning words, and selective attention to the appropriate referent for a given label is crucial for successful word learning. For nouns such as the word "cup", for instance, attention to the correct referent object, a cup, and the relevant feature of that object, its shape, facilitates learning the noun. For adjectives such as the word "red", attention to the correct feature of an object, its color, facilitates learning the adjective. Selective attention to incorrect information, on the other hand, e.g. the shape of a cup on hearing the word "red", would lead to error in interpreting the meaning of "cup" or the meaning of "red" in future settings. Given the importance of selective attention in word learning, it is essential to understand the role of selective attention mechanisms in

children's early language learning. How might attentional mechanisms work to guide selection to relevant information, and what role do these processes play in children's early word learning?

Research suggests that part of the answer to the question of how selective attention is guided lies in understanding that what we learn to attend to in one moment and context affects what we attend to in a later moment with an overlapping context. There is strong evidence in the adult literature to support the idea that attention is guided by past learning, and by the relationship between the current moment and what was learned in the past. One phenomenon where this is apparent is Highlighting (Kruschke, 1996.) Of particular interest to this paper is the attention-shifting mechanism that accounts for Highlighting and which, broadly described, operates by reallocating attention to new information if old information is paired with a novel outcome. This type of attention-shifting mechanism could be very relevant to early word learning in children, and might provide a key ingredient that allows children to bootstrap themselves into language learning without the need for higher-level syntactic knowledge. Could general mechanisms of selective attention and appropriate attention shifting account for language learning phenomena without recourse to assumptions about children's knowledge of language structure? We pursue this question in the domain of adjective learning.

First we discuss adjective learning and present one particular finding. We then briefly describe the adult phenomenon, Highlighting, drawing parallels between it and adjective learning. This leads us to consider the attentional explanation for Highlighting to suggest a type of mechanism that could also underlie word learning. We test the plausibility of such a mechanism in two experiments and discuss the results.

Adjective Learning

Different kinds of words refer to different kinds of meanings. It has been suggested that early-learned noun categories are most likely organized by shape (Samuelson & Smith, 1999); adjectives typically refer to other properties, those that can vary within an object category (Klibanoff & Waxman, 2000). Relative to common nouns,

children learn adjectives slowly (see Gasser and Smith, 1998; for a review) and, indeed, often misinterpret adjectives as referring to an object category rather than a property (Clark, 1997; Golinkoff, Mervis, and Hirsh-Pasek, 1994; Soja, Carey, and Spelke, 1991; Markman, 1990; Landau, Smith, and Jones, 1988; Macnamara, 1982; Keil, 1991). However, children’s adjective-like interpretations of novel words increase (Klibanoff and Waxman, 2000; Waxman and Klibanoff, 2000; Waxman, 1998) in contexts in which the object category is named. These phenomena have been interpreted by some as indicating early knowledge of syntax-meaning correspondences (Waxman, 1998; Mintz and Gleitman, 2002).

Mintz and Gleitman (2002) reported the specific result that motivates the experiments in this paper. They contrasted two kinds of sentences in which a novel adjective might be embedded. In one of these, the novel adjective modified a pronoun, “This is a stoof one”. In the other, the novel adjective modified a basic level noun, “This is a stoof elephant”. They reasoned, on linguistic grounds, that children should more successfully map the novel adjective to a property given the noun rather than the pronoun because the syntactic role of the adjective as a modifier of a noun was more obvious. This prediction was supported by the performances of 3-year-olds, but not by that of younger children who performed at chance. Mintz and Gleitman interpreted these results in terms of emerging knowledge of syntax and linkages between syntactic categories and meaning.

Mintz and Gleitman’s interpretation suggests that children younger than 3 years old cannot learn the meaning of adjectives since they do not have sufficient knowledge of syntax - that knowledge of syntax is necessary for learning adjectives. We suggest, in contrast, that young children can learn adjectives even without knowledge of language structure because general attentional mechanisms automatically guide selective attention to the appropriate object feature. We explain this view by first introducing a phenomenon and explanation from the adult associative learning literature.

Attentional Highlighting

Highlighting is a robust phenomenon in adult associative learning. The phenomenon emerges in tasks, as illustrated in Figure 1, in which learners first learn a conjunctive cue (A+B) that predicts an outcome (X) and then are presented with a new conjunctive cue that contains one old component (A) plus a new one (C) and predicts a new outcome (Y). During this second phase, learners associate the new cue with the new outcome more than the old cue with the new outcome. In brief, novel cues are associated with novel outcomes. By one explanation, this derives from the rapid shifting of attention away from the previously learned cue in the context of the new outcome – the new cue thus becoming attentionally highlighted and strongly associated with the new outcome (e.g., Kruschke, 1996, 2005). Such attentional mechanisms preserve

previous learning and constrain new learning in novel situations.

<p>Highlighting:</p> <p><u>Early trained (n trials):</u> Cue A + Cue B -> Outcome X</p> <p><u>Later trained (n trials):</u> Cue A + Cue C -> Outcome Y</p> <p><u>Result:</u> A associated with X more than with Y C associated with Y more than A is</p>	<p>Adjective learning:</p> <p><u>Previous world training:</u> Elephant shape + standard texture -> “elephant”</p> <p><u>In experiment:</u> Elephant shape + novel <i>stoof</i> texture -> “<i>stoof</i>”</p> <p><u>Result:</u> Elephant shape associated with word “elephant” <i>Stoof</i> texture associated with word “<i>stoof</i>”</p>
---	---

Figure 1: Illustration of a parallel structure between highlighting and adjective learning

There are important parallels between the pattern of learning in Highlighting tasks and the pattern for learning adjectives, particularly in Mintz and Gleitman’s 2002 adjective learning task. Figure 1 illustrates the parallel structure of the two phenomena. Children learn nouns before adjectives, and they learn that shape is associated with nouns. That is, they learn early in language development that “elephant” refers to animals of a particular shape and particular texture, having seen various examples of elephants that have been referred to as “elephant”. This is analogous to having learned that A+B predicts X in the first phase of Highlighting tasks, where A is elephant shape, B is elephant texture, and X is the word “elephant”. (There may be other associations as well, but these are the ones relevant to our explanation). This knowledge is in place before children come to the adjective-learning experiment. In an adjective-learning task such as Mintz and Gleitman’s, a novel feature is introduced along with a novel label for a known object. This is analogous to the second phase of Highlighting tasks, where A (elephant shape) is now paired with C (a novel texture) and the outcome is Y (a novel word). In Highlighting tasks, this results in associating C (the novel texture) with Y (the novel word), while continuing to associate A (elephant shape) with X (“elephant”), which is equivalent to learning the meaning of the adjective in the adjective learning task.

Just as the mechanisms underlying Highlighting involve attentionally highlighting the novel cue given a novel outcome, we propose similar mechanisms might be at work in the adjective learning task, with attention shifting away from known features to unknown features when a new label is given. Connecting this idea to Mintz and Gleitman’s findings, the presence of the basic level noun in

constructions such as “This is a stoof elephant” may have aided adjective learning because the noun activated a previously established association between the noun and relevant object properties (e.g., shape). In the context of this noun, the novel label may have shifted attention from these properties to an object feature, such as texture, that was not associated with the noun.

Highlighting and adjective learning are roughly analogous, and no claim is being made here that the phenomena are exactly the same. Given the similarity in their structure, however, it is plausible to consider that similar attentional mechanisms might be at work in both phenomena – mechanisms of rapid attention shifting to novel cues in the presence of novel outcomes (or words).

Determining the viability of an attention shifting account of adjective learning is important for two reasons. First, it may unify a robust laboratory phenomenon in adult learning with a significant class of real world learning, showing that the mechanisms of attention shifting are capable of doing real developmental work. Second, although older children and adults surely do acquire the syntactic categories and the syntax-semantic linkages that Mintz and Gleitman propose, children cannot know them at the start and must acquire them. General attentional mechanisms that shift attention to relevant information based on previous associations and learned attention may help language learning get started, prior to children’s learning of syntactic categories.

Experiment 1

In the Mintz and Gleitman experiment, 3-year-olds successfully mapped the novel adjective to the property of the named object when the sentence contained a basic level noun rather than a pronoun, for example, in sentences of the form “This is a stoof elephant”. The syntactic role of stoof as an adjective is specified by its position before the noun. In Experiment 1, we take this information away by using sentences of the form “This is an elephant red stoof”. If children are using syntactic information to map the novel word to a property, they should fail given this sentence since the syntactic information would not specify “stoof” as an adjective (but perhaps as a noun). However, by an attention-shifting account that operates analogously to attentional highlighting, this order of information should not matter since naming the noun (and color) would activate prior associations resulting in attention shifting rapidly to an unassociated property (e.g., texture) with the novel word “stoof”.

Method

Participants Twenty 2-year-olds (mean=25.6 months) and 28 3-year-olds (mean=36.1 months) were randomly assigned to the Noun or the Pronoun condition.

Stimulus Materials The stimuli replicate those of Mintz and Gleitman (2002). For each set, the 3 training objects were instances of 3 different basic level categories, all the same color and surface texture (see Figure 2). The test objects included one category present in the training set but with a different texture and one novel category instance of the same target texture as the training set. Both test objects were the same color as the training objects. All objects were approximately 10 cm³. Texture - the intended target property - was chosen to be highly novel and included: a star pattern, a wire surface, bumpy striped, a felt surface, a Velcro surface, and holes. These properties were named by novel labels such as *Brickish*, *Stoof*, *Zav*, *Afe*, *Vap*, and *Toop*, respectively.

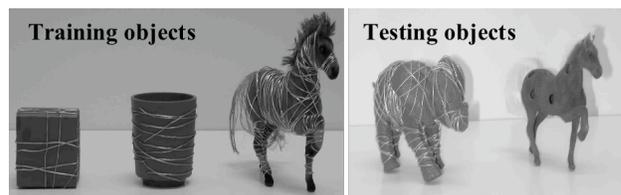


Figure 2: The photographs of a set of the stimuli objects used in one trial.

Procedure On each trial, the child heard each of the 3 training objects labeled by the same novel adjective. The 2 test objects were then placed on the table and the child was allowed to play with them to reduce choices based on toy preference. The 2 test objects were then placed equally distant from the child and the child was asked to select the named one. This procedure was repeated for the 6 trials. Children in the Noun condition heard the training objects labeled in a sentence containing the novel adjective, a color word, and a basic-level noun, but in a scrambled order - e.g., “This is an elephant red stoof”. Children in the Pronoun condition heard each training object labeled with the novel word embedded in a correctly ordered English sentence but with the adjective modifying a pronoun - e.g., “This is a stoof one”. Children in both conditions were queried with the very same carrier phrase, “Can you get me the stoof one?”

Results

Children’s performance in terms of their choices of the texture matching test objects is shown in Figure 3. A 2(Condition) X 2(Age) analysis of variances of these choices revealed a main effect of Age; $F(1, 38)=5.544$, $p<.05$ and an interaction between Age and Condition, $F(1, 38)=4.38$, $p<.05$. In the Noun condition, younger children performed better than older children and above chance, $t(9)=2.88$, $p<.05$, whereas older children did not. Both older and younger children performed comparably and at chance level in the Pronoun condition. Thus, younger children were able to successfully map the novel word to

the target property despite the scrambled order and lack of syntactic information.

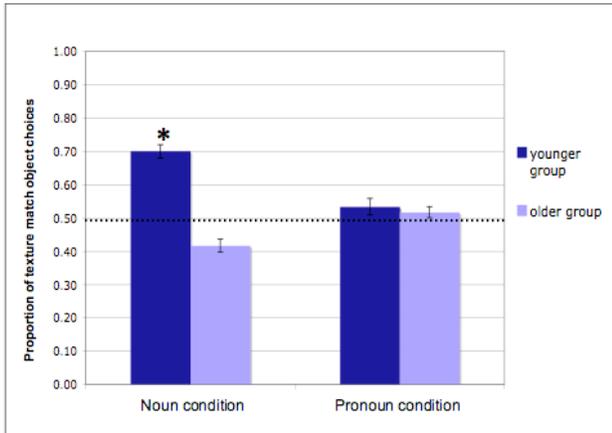


Figure 3: The proportion of texture match object choices (property matched test object choices) made by younger children and older children in the Noun and Pronoun conditions.

This pattern is consistent with the attentional hypothesis. If activated past associations induce attention shifting to novel cues given novel outcomes, this condition contains just the right information to shift attention away from the shape and parts of the named test objects, away from color, and toward the novel texture. That the younger children did not succeed in the syntactically well ordered pronoun condition---that lacked mention of the noun or color---supports the idea that the activation of these past associations contributed to the shift in attention to texture. The 2-year-olds' pattern of performance thus fits the hypothesis, and shows how a domain general mechanism of attention shifting could lead to the learning of adjectives without any necessary knowledge of the syntactic category of adjective or its link to semantic categories.

The older children's pattern of performance does not fit this prediction. Given sentence frames with scrambled word order or with only a modified pronoun, these children do not map the novel word to a property. Given older children's successful performance in Mintz and Gleitman's original study which used proper English word order, it would seem that for these children syntactic cues and not the mere activation of prior labeled associations (and the attention shifting that engenders) matters.

Experiment 2

Two-year-olds in Experiment 1 succeeded in mapping the novel word to texture given a scrambled word order in which the object and its color were explicitly labeled. By our attention-shifting account, the mention of color should matter. By activating any past association between a color word and COLOR along with past association of the noun

and object, attention should be shifted away from these competitors and thus toward the novel texture properties. Experiment 2 tests the role of explicit labeling of color.

Method

Participants The participants were 12 2-year-olds (mean=26.7 months, ranging from 22.6 to 30.4 months); none had participated in Experiment 1.

Stimuli Materials and Procedure The stimuli replicate those of Mintz and Gleitman (2002). Two conditions varied within subject. On half of the trials, the +Color trials, these were identical to the Noun condition in Experiment 1 (e.g., "This is an elephant red stool") and on half, the -Color trials, the color word was dropped (e.g., "This is an elephant stool.") Order of the +Color and -Color trials was randomized across children. All other aspects of the procedure were the same as Experiment 1.

Results

Mentioning the color clearly mattered as shown in Figure 4. Children mapped the novel word to the texture more in the +Color than -Color condition; $t(11)=2.016$, $p<.05$. Children performed above chance in the +Color trials; $t(11)=2.016$, $p<.05$ but not in -Color condition; $t(11)=-1.2$.

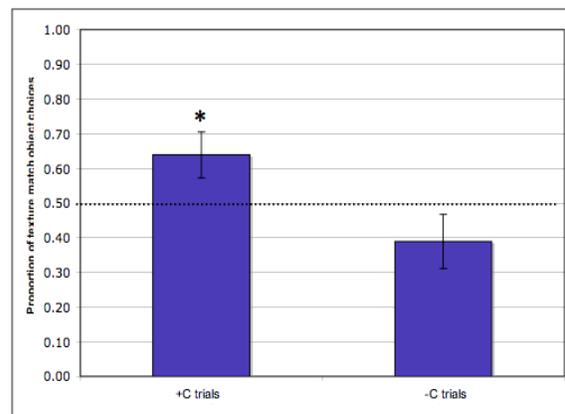


Figure 4: The proportion of texture match object choices (property matched test object choices) made by children on the +Color and -Color trials.

This pattern is consistent with the attention-shifting hypothesis, demonstrating the role of additional previously formed associations (in this case, color words) in reducing competition for attention.

General Discussion

Highlighting along with phenomena such as blocking, learned irrelevance and categorical perception (Goldstone 1998; Kruschke & Blair, 2000; Kruschke 2006), reveal a

fundamental truth about associative learning: that the contingencies present in the moment are not all that matters; rather it is the relationship among these contingencies and past learned associations that determines what is attended to and learned in the moment. The present results show such mechanisms may benefit early word learning. When a prior learned association is activated, for example elephant shape cueing the word “elephant” and the color red cueing the word “red”, this activation effectively shifts attention to the novel property (texture) with the novel word (“stooof”), such that they are selectively linked.

Given this learning mechanism, there is no need to presume higher-level explanations of the process, such as word learning constraints. One could argue, in particular, that these types of processes of learned attention and attention shifting could be the mechanisms responsible for the word-learning phenomenon of Mutual Exclusivity. This phenomenon is often explained in terms of a constraint, suggesting that children constrain their early word learning by assuming names of things are mutually exclusive, and this would explain the tendency children have of using only one label for one instance – a phenomenon which has long been reported and shown by a number of studies even across different languages (e.g., Markman & Wachtel, 1988; Markman, 1989; Merriman & Bowman, 1989). However, there have been very few attempts to explain how this tendency could be generated. Many researchers in child language have proposed that word learning builds on itself, with learners using past knowledge to guide subsequent learning (Smith, Jones, Landau, Gershkoff-Stowe, and Samuelson, 2002; Graf Estes, Evans, Alibali, and Saffran, in press), but there has been little specification of the precise mechanisms involved, neither how they operate at the moment of learning nor precisely how prior learning directs current learning (see Merriman, 1999, for one exception). The present results thus provide evidence for one such explanation, a mechanism that is well known and well studied in adult human learning. In addition, the present study demonstrates the importance of such a mechanism in the developmental process, in solving a real developmental problem such as how to get word learning started when one does not yet know much (if anything) about words and their syntactic categories or linkages from syntactic categories to meanings.

Thus, the main contribution of our experimental results with respect to children’s learning of adjectives is centered on the performances of the youngest children, 2-year-olds. Past work using methods similar to those used here (by Mintz and Gleitman, 2002) suggests that children this young have considerable difficulty in mapping an adjective to a property such as texture rather than the whole object or its shape (Clark, 1997; Golinkoff, Mervis, and Hirsh-Pasek, 1994; Soja, Carey, and Spelke, 1991; Markman, 1990; Landau, Smith, and Jones, 1988; Macnamara, 1982; Keil, 1991). Nonetheless, children this age succeeded in making

this mapping in Experiment 1 which provided two apparently key ingredients: the explicit labeling of the noun and the explicit labeling of a competing property (color). By the attentional hypothesis, the explicit mention of these words did this by activating specific past associations and thus shifting attention to the novel cue and outcomes presented in the learning context. These children’s success however did not depend on using English word order, which is critical to the syntactic definition of a word as an adjective. This implies that these younger children were not using knowledge about the syntactic category of the novel word to guide their learning. Given the general learning mechanism of attentional highlighting, they do not need it. Past learning of other words, and the activation of that learning by the explicit mention of these words can work to appropriately direct attention and learning.

The performance of the older children (3-year-olds) in Experiment 1 does not support the attentional hypothesis. The original motivation of these experiments was, in a sense, to “explain away” Mintz and Gleitman’s results by showing that the success they reported for 3-year-olds in their study did not require knowledge of syntactic categories or their linkage to semantic categories. We did not do this, and Mintz and Gleitman’s original conclusion – about the role of the noun in defining the syntactic category of adjective – still stands for 3-year-olds. The fact that 3-year-olds in the present experiment performed worse than 2-year-olds when English word order was taken away shows that these children are using syntactic knowledge. What we add to Mintz and Gleitman prior results, then, is new insights into an earlier origin for adjective learning, one that does not depend on having yet acquired that syntactic knowledge.

In conclusion, the power of associative learning is the effect of systems of interacting associations that both build over time and exert direct influence on attention and learning in the moment. Early adjective learning, like Highlighting, may be an example of the effects of known words and referents on the learning of new ones, and also a demonstration of how associative learning, through its effects on attention, contributes to smart word learning.

Acknowledgments

This research was supported by NIMH grant (R01-MH60200), NIH-NICHHD grant (R01-HD28675.) We thank the children and parents who participated in this study.

References

- Clark, E. V. (1997) Conceptual perspective and lexical choice in acquisition. *Cognition* 64, 1-37.
- Gasser, M. & Smith, L. B. (1998). Learning nouns and adjectives: a connectionist account. *Language and Cognitive Processes*, 13, 269-306

- Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, 49, 585-612.
- Golinkoff, R., Mervis, C., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, 21, 125-155.
- Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (in press). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*.
- Keil, F. C. (1991). Theories, concepts, and the acquisition of word meaning. In J. P. Byrnes & S. A. Gelman (Eds.), *Perspectives on language and cognition: interrelations in development* (pp. 197-224). Cambridge: Cambridge University Press.
- Klibanoff, R. S. & Waxman, S. (2000). Basic Level Object Categories Support the Acquisition of Novel Adjectives: Evidence from Preschool-Aged Children. *Child Development*, 71(3), 649-659.
- Kruschke, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(1), 3-26.
- Kruschke, J. K. & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin & Review*, 7, 636-645.
- Kruschke, J. K., Kappenman, E. S., & Hetrick, W. P. (2005). Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(5), 830-845.
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, 3, 299-321.
- Macnamara, J. (1982). *Names for things: A study of human learning*. Cambridge, MA: MIT Press.
- Markman, E. M. (1989). *Categorization and naming in children*. Cambridge, MA: Bradford Books
- Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, 14, 57-77.
- Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S. A. Gelman & J. P. Byrnes (eds), *Perspectives on language and thought : interrelations in development*. New York, NY: Cambridge University Press.
- Markman, E. M. & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology* 20, 121-57.
- Mintz, T. H. & Gleitman, L. (2002). Adjectives really do modify nouns: the incremental and restricted nature of early adjective acquisition. *Cognition*, 84, 267-293.
- Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category organization and syntax correspond? *Cognition*, 73(1), 1-33.
- Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L. & Samuelson, L. (2002). Object Name Learning Provides On-the-Job Training for Attention. *Psychological Science*, 13, 13-19.
- Soja, N. N., Carey, S., & Spelke, E. S. (1991) Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition*. 38, 179-211.
- Waxman, S. R. (1998). Linking object categorization and naming: Early expectations and the shaping role of language. In D. L. Medin (Ed.), *The Psychology of Learning and Motivation*, Vol. 38 (pp. 249-291). San Diego: Academic Press.
- Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology*, 36 (5), 571-581.