Cognition 136 (2015) 99–122

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Pigeons acquire multiple categories in parallel via associative learning: A parallel to human word learning?

Edward A. Wasserman [a,b], Daniel I. Brooks [c], Bob McMurray [a,b,d,*]

[a] Dept. of Psychology, University of Iowa, United States
[b] Delta Center, University of Iowa, United States
[c] Dept. of Psychology, Tufts University, United States
[d] Dept. of Communication Sciences and Disorders, University of Iowa, United States

ARTICLE INFO

Article history:
Received 19 May 2014
Revised 23 September 2014
Accepted 17 November 2014

Keywords:
Animal behavior
Associative learning
Categorization
Language
Word learning
Comparative cognition

ABSTRACT

Might there be parallels between category learning in animals and word learning in children? To examine this possibility, we devised a new associative learning technique for teaching pigeons to sort 128 photographs of objects into 16 human language categories. We found that pigeons learned all 16 categories in parallel, they perceived the perceptual coherence of the different object categories, and they generalized their categorization behavior to novel photographs from the training categories. More detailed analyses of the factors that predict trial-by-trial learning implicated a number of factors that may shape learning. First, we found considerable trial-by-trial dependency of pigeons' categorization responses, consistent with several recent studies that invoke this dependency to claim that humans acquire words via symbolic or inferential mechanisms; this finding suggests that such dependencies may also arise in associative systems. Second, our trial-by-trial analyses divulged seemingly irrelevant aspects of the categorization task, like the spatial location of the report responses, which influenced learning. Third, those trial-by-trial analyses also supported the possibility that learning may be determined *both* by strengthening correct stimulus–response associations and by weakening incorrect stimulus–response associations. The parallel between all these findings and important aspects of human word learning suggests that associative learning mechanisms may play a much stronger part in complex human behavior than is commonly believed.

1. Introduction

Who was correct: Chomsky or Skinner? This question captures in stark terms one of the oldest debates in cognitive science. Are complex—perhaps uniquely human—behaviors like language acquired via specialized and perhaps innately constrained cognitive mechanisms? Or can such behaviors emerge from more basic and general mechanisms like associative learning (e.g., Fitch, 2010)? This debate spans virtually every level of language: from the acoustic signal (Liberman & Whalen, 2000; McMurray & Jongman, 2011), to word learning (Golinkoff & Hirsh-Pasek, 2006; Ramscar, Dye, & Klein, 2013; Xu & Tenenbaum, 2007; Yu & Smith, 2012), to syntax and grammar (Christiansen & Chater, 2008; Hsu & Chater, 2010; McClelland & Patterson, 2002; Pinker & Ullman, 2002), and to basic learning mechanisms that may underlie language acquisition (Marcus, Vijayan, Bandi Rao, & Vishton, 1999; Saffran & Thiessen, 2007).

Such debates have resulted in considerable theoretical progress. Behaviorist notions of association have developed into more complex emergentist views (e.g., connectionism, dynamical systems); and, language-specific

* Corresponding author at: Dept. of Psychology, E11 SSH, University of Iowa, Iowa City, IA 52242, United States.

http://dx.doi.org/10.1016/j.cognition.2014.11.020
0010-0277/

accounts have adopted more sophisticated symbolic and probabilistic algorithms. However, harkening back to Skinner (1957) and (Chomsky, 1958; but see, MacCorquodale, 1970), many of these debates rest on assumptions about what simple mechanisms like associative learning can or cannot do in a domain like language. Yet, there is little empirical basis for appreciating what associative learning can actually do. Studies of human learning can be ambiguous in this regard because the same learning problem may be solved via associative or inferential routes (cf., Medina, Snedeker, Trueswell, & Gleitman, 2011; Yu & Smith, 2007) (and see, Ramscar et al., 2013). Computational models have helped clarify what associative mechanisms can do (Mayor & Plunkett, 2010; Samuelson, 2002), but they also rely on controversial or simplifying assumptions that make them less than definitive. In view of these difficulties, animal models, particularly those that employ less cognitively sophisticated animals like pigeons or rats, may offer a complementary perspective by revealing which learning mechanisms are species-general and by distilling a purely associative learning paradigm through the careful control of inputs, outputs, and reinforcement schedules. Adopting an animal-model approach may prove to be crucial for understanding the potential of purely associative components to contribute to higher-level cognitive abilities like language learning.

This paper begins to address these issues in the context of word learning. As we argue below, no animal model currently exists that captures three critical aspects of word learning. First, existing animal models do not isolate a purely associative framework for word learning, but also include social interaction and an enriched learning environment. Although a purely associative animal model surely represents a pale replica of human word learning, it can also afford a great theoretical benefit by allowing us to isolate and refine our understanding of the associative mechanisms that, although embedded in the more sophisticated cognitive processes of humans, are likely to subserve at least part of word learning. Second, few animal models include tasks in which many inputs are mapped to many outputs. This is a critical problem that limits our ability to extend our knowledge of associative learning primarily in animals to those aspects of human word learning where it might be relevant. Third, most animal models progressively add associations over training. However, children are not successively taught their expanding vocabularies; rather, children are simultaneously barraged with thousands of words only a fraction of which are added as language learning proceeds.

Thus, the primary goal of our study was to develop such an associative learning model that meets these three criteria, using pigeons as experimental subjects. We specifically deployed this model to clarify a recent debate concerning the mechanisms of observational word learning in humans. We wish to stress at the outset that an animal model meeting the above criteria is insufficient to capture the full richness of human word learning. However, an animal model that satisfies these criteria is nevertheless a significant step toward capturing the types of associative processes that may form a crucial component of human word learning; an effective animal model can therefore help us understand these foundational associative processes. We start our article with a broad review of association in word learning and the motivation for our animal model; we then detail this recent debate on observational word learning; finally, we present our experiment.

1.1. Association as a component of human word learning

In recent years, human word learning has offered an important domain in which the debate between language-specific and emergent associative approaches has played out. The prime challenge of word learning is to map phonological word-forms onto concepts and categories. This mapping seems straightforward, but it requires children to solve a range of problems. When children are confronted with a novel name, infinite interpretations are available involving: any available object, its properties, and so forth (Quine, 1960). Even if the learner can identify the referent of a word, they must generalize the word to new exemplars, with numerous available dimensions over which to generalize (color, shape, function, etc.). And, of course, they must ultimately do both of these things for tens of thousands of words.

Given the sheer magnitude of these problems, many theorists argue that word learning cannot be based on associative processes. Mechanisms such as constrained Bayesian inference (Xu & Tenenbaum, 2007), some form of social inference (Akhtar & Martinez-Sussman, 2007; Golinkoff & Hirsh-Pasek, 2006; Tomasello, 2001), or both may be needed (Frank, Goodman, & Tenenbaum, 2009). Other authors insist that word learning is largely a logical, symbolic problem (Halberda, 2006; Medina et al., 2011; Trueswell, Medina, Hafri, & Gleitman, 2013).

A rich discussion of these issues is underway. For example, empirical work is asking if complex social abilities may derive from attentional and/or associative factors (Akhtar, Carpenter, & Tomasello, 1996; Samuelson & Smith, 1998). Researchers are also asking if the way that words influence the formation of categories is best described by conceptual theorizing (Waxman & Gelman, 2009) or if labels simply guide attention (Robinson & Sloutsky, 2007). Further inquiry is exploring the learning mechanism itself, revealing that when we strip away many of the social or inferential cues, infants and adults can still master word–object mappings using only statistical co-occurrence (Smith & Yu, 2008; Yu & Smith, 2007), although it remains to be seen whether this type of learning by observation is associative or inferential in nature (McMurray, Zhao, Kucker, & Samuelson, 2013; Medina et al., 2011; Ramscar et al., 2013).

Similarly, computational work has explored the ability of simple associative systems to exhibit some of the hallmarks of more inferential processes, including: acceleration in the rate of learning (McMurray, 2007; Regier, 2005), abstraction of generalizable dimensions or selective attention (Colunga & Smith, 2005; Samuelson, 2002), and rapid inference of new names (McMurray, Horst, & Samuelson, 2012; Regier, 2005). There is also evidence that the componential structure of words can even emerge from simple, Rescorla–Wagner associative systems which