Published in final edited form as:

Dev Sci. 2008 March ; 11(2): 219–222. doi:10.1111/j.1467-7687.2007.00669.x.

The shape bias: an important piece in a bigger puzzle

Jeffrey L. Elman

Department of Cognitive Science, University of California, San Diego, USA

Introduction

No theory remains fixed for ever, at least not until it is abandoned. Theories develop over time. Some changes reflect the discovery of new facts that prompt modification of the earlier version. And often, especially in the case of theories that are tied to mechanisms, changes arise because with time comes a deeper understanding of the mechanism's characteristics. This understanding in turn generates greater precision in the predictions that can be made by the theory. Obviously, a critical issue is whether the changes are post hoc and unprincipled – in which case the theory is at risk of losing credibility – or whether the developments reflect true maturation and greater insight.

I believe that changes in the evolution of the Attentional Learning Account (ALA) play a large role in the apparent disagreements that motivate the current set of papers. In fact, there is considerably more agreement than appears on a casual reading of these papers. But I also believe there remain significant points of disagreement. In this commentary, I will try to address what I see as the issues on which there is convergence, and those for which there remain important differences.

Apparent disagreements

Let me address what seem to be the major disagreements. These revolve mainly around the following claims:

1. The shape bias is language/naming specific. It does not appear in non-linguistic domains.
2. The shape bias only emerges when language learning is well on its way, typically at a stage where a child has a productive vocabulary of 50–150 nouns.
3. The shape bias is sensitive only to perceptual (not conceptual) factors.

For each of these, there appear to be two camps. One camp says Yes to each of these propositions, and is represented in the papers by Colunga and Smith (2008; CS), and by Samuelson and Horst (2008; SH). This is the ALA group. The other camp says No, and is represented by Booth and Waxman (2008; BW), and by Markson, Diesendruck and Bloom (2008; MDB). This is the anti-ALA group. To be clear: I do not claim that any of these authors would be happy with this heavy-handed parsing of positions. Indeed, I am quite sure that several of them would object, particularly given the simplistic way I've phrased things. But it provides a start.

Confusion quickly arises because many of the data provided by the No folks are seen by the Yes folks as supportive of the ALA, particularly demonstrations that the shape bias can be seen

in non-naming contexts (data cited by MDB) and that apparently conceptual information affects the strength of the shape bias (data cited by BW). Even more confusing, some of the pro-ALA folks themselves report data that the anti-ALA folks claim undermine the ALA.

How can this be?

I suggest that two things are going on here. The first arises from developments over time in the ALA. The second, to which I turn in the second portion of this paper, is that the true differences of opinion have not to do with the empirical phenomena, but how they are understood.

In its early form, the ALA account of the shape bias led to at least five predictions (e.g. Smith, 1999). One of these (Hypothesis 2) was that the shape bias does not preexist word learning. A second hypothesis (Hypothesis 3) was that the shape bias is lexically specific when it first emerges.

Eighteen years ago, these hypotheses were compatible with the empirical data that were then available, including studies suggesting that ‘higher level’ information was not significant in naming generalization. And the hypotheses seemed reasonable, given the central issue which the ALA was addressing. What was that issue?

At the time, the existence of the shape bias was undisputed (and remains so). The question is where it came from. One possibility is that the shape bias is innate and domain-specific. The ALA suggested another answer to this question. According to the ALA, the shape bias might be specific to the domain of language in its operation (because that's what the extant data indicated), but the bias itself ‘emerges from very general learning processes, processes that in and of themselves have no domain-specific content’ (Smith, 1999, p. 282). I understand ‘general’ here to mean ‘not specific to the domain of language’, ‘not innate’.

Thus, the force of early ALA was to address the origins of shape bias, with the specific proposal that the bias could be learned. The proposal provided an elegant explanation of how a bias might both be the *cause* of language learning and at the same time the *result* of language learning. By the way, it is important to note that the data supporting the contribution of linguistic experience to the emergence of the shape bias stand uncontested.

What has happened in the interim is that a number of important empirical discoveries have been made. These include:

- data (reviewed in BW) that show that shape bias is not automatic, but that it is modulated by additional information provided: Information suggesting an object is an artifact promotes the shape bias; info suggesting it is animate decreases it;
- data (reviewed in BW and MDB) that suggest that there is probably a broader developmental window; infants, for example, show awareness of shape in a variety of tasks;
- data (reviewed by MDB and CS) that suggest that shape operates as a cue in non-naming (categorization) tasks;
- data (reviewed by CS and SH) that show that in training studies, the shape bias can be made stronger or weaker, depending on the nature of relatively brief training; in the age groups tested, however, it does not appear possible – or at least easy – to develop a material bias;
- data (reviewed by SH) from simulations studies that show that the apparent strength of the shape bias may depend on task demands. One consequence is that whether or

not there appear to be developmental trends in shape bias (as predicted by the ALA, but not found in several studies) may depend on differential sensitivity of different tasks.

SH and CS argue that these data are compatible with ALA. BW and MDB disagree. We ask, once again, what's going on?

First, it appears to me that ALA has evolved over time. In its early form, the ALA did suggest that that window started with language and that language was the driving force. Are these claims central to ALA? In my view, they are not. ALA maintains that language is itself a powerful driving force for the emergence of the shape bias, but leaves open the possibility (mentioned even in early papers) that non-linguistic information may play a very important role in the shape bias. This now seems to be the case.

In addition to responding to a growing body of empirical knowledge regarding the shape bias, the ALA has also evolved in response to a large work in the computational literature about learning, including results from connectionist models and dynamical systems theory. There are two major findings from this literature. We now know that so-called 'simple learning' algorithms can produce outcomes that are far from simple. The limits of pair-wise associations are well known and severely limiting. The power of higher order associations (correlations over correlations) is considerably greater. Indeed, a number of analyses now suggest that these systems possess something approaching Turing computability, at least in their idealized form. Second, we have come to understand that such mechanisms also permit a very rich range of behaviors that on the one hand allow for exquisite sensitivity to context and task demands, while on the other hand demonstrating sensitivity to overarching generalizations that appear behaviorally to reflect very abstract knowledge.

Real disagreements

However, while these changes in the ALA also bring it into alignment with the kinds of data that are presented in BW and MDB, they also expose another set of issues around which I believe there remains more substantive disagreement. This is the question – raised very explicitly in CS – about the nature of cognition. More specifically, the questions now have to do with the *form* of knowledge, and with questions about how that knowledge can be *acquired*. In the context of the shape bias, the debate appears to revolve around whether the knowledge underlying the shape bias is *conceptual* or *perceptual*.

MDB argue that shape bias 'is not an autonomous aspect of cognition . . . it emerges from the interactions of other capacities that children possess, some having to do with language . . . some having to do with categorization' (p. 214). Implicit in this characterization is that the shape bias reflects conceptual information.

BW are even clearer on this point. They argue that 'one goal of the ALA appears to be to explain early word learning without appealing to conceptual factors' and they question whether 'simple associative processes [are] at the heart of the development of the lexicon in general, and the shape bias in particular' (p. 193). Referring to work that suggests that causality and inference is not based on correlational mechanisms, they come down heavily on the side that maintains that – whatever role might also be played by perception – conceptual information is critical to the shape bias.

Here, as CS argue (and BW acknowledge) everything depends on how one defines conception and perception. This is an issue that is quite controversial and has been the subject of lively debate for almost two decades. It in turn is embedded in the broader question of cognition, as

‘rule-like amodal propositions, on the one hand, or as embodied, modal, and dynamic processes on the other’ (CS).

I make no bones about being sympathetic to the latter view. Having begun my career working within the paradigmatic symbolic linguistic framework (generative linguistics), I have come over many years to believe that symbolic systems present inherent and ultimately unavoidable problems when one attempts to account for data that are often highly context-sensitive and content-dependent (i.e. interaction between form and content). To come to this conclusion is not to deny our ability to abstract or to generalize, nor is it to deny the importance of phenomena that more traditional symbolic approaches treat as conceptual. Rather, the framework within which the current ALA is couched understands such phenomena as emerging from a cascading set of interlocking correlations at lower levels that result in higher order and often very complex regularities. To call these either perceptual or conceptual misses the point. To be sure, knowing that an object is red is not the same as knowing that one animal is domestic and another is feral. The claim is that the latter kind of knowledge grows out of – and remains connected with – knowledge of color, shape, size, context, behavior, etc. Furthermore, this knowledge accumulates across multiple categories, supporting inferences that may appear to be entirely ‘conceptual’ (in the sense of being ungrounded in perception) but which are in fact the product of an extraordinarily rich network of connections between perceptions and action.

Clearly, BW are skeptical of such claims. MDB point out that if this is the issue truly in dispute, then ‘a few laboratory experiments are not going to settle the issue’. Skepticism is reasonable, and the stakes are high enough that many experiments indeed should be required to settle the issue. I myself am more positive about the prospect that the new non-symbolic computational models of cognition will indeed provide a unified understanding of the perception vs. conception divide. But this remains a question for the future to resolve.

In the short term, the criteria by which the competing theories of cognition may be judged will lie both in their ability to account for existing data, as well as their ability to drive research to uncover new data. Here, the (still ongoing) debate about the past tense of English verbs serves as a model. This debate began with a set of claims about how children might learn the past tense of English verbs (Rumelhart & McClelland, 1986). These claims challenged the conventional symbolic account, and were followed quickly by a forceful disagreement (Pinker & Prince, 1998). At stake were questions not dissimilar from those that arise here: What is the form of cognition, and what are the origins of knowledge? In the following two decades, an enormous body of literature has resulted from this debate. The acrimony that sometimes characterizes this debate may be unfortunate, but by any standards the wealth of data that have been generated in the course of that debate and the increase in sophistication of the theories must count as a real scientific benefit.

Already, the ALA has led to predictions (and counter-predictions) that have motivated new experiments and simulations that enrich the literature. In the long run, whether or not the ALA continues to do so with success will be one piece in the puzzle that helps us figure out what cognition is really made of.

Acknowledgements

This work was supported by a grant HD053136 from the National Institutes of Health, and by the UCSD Kavli Institute for Brain and Mind.

References

Booth AE, Waxman SR. Taking stock as theories of word learning take shape. *Developmental Science*. 2008DOI: 10.1111/j.1467-7687.2007.00664.x

- Colunga E, Smith L. Knowledge embedded in process: the self-organization of skilled noun learning. *Developmental Science*. 2008DOI: 10.1111/j.1467-7687.2007.00665.x
- Markson L, Diesendruck G, Bloom P. The shape of thought. *Developmental Science*. 2008DOI: 10.1111/j.1467-7687.2007.00666.x
- Pinker S, Prince A. On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 1988;28:73–193. [PubMed: 2450717]
- Rumelhart, DE.; McClelland, JL. Learning the past tense.. In: Rumelhart, DE.; McClelland, JL., editors. *Parallel distributed processing*. Vol. 1. MIT Press; Cambridge, MA: 1986. p. 216-271.
- Samuelson LK, Horst JS. Confronting complexity: insights from the details of behavior over multiple timescales. *Developmental Science*. 2008DOI: 10.1111/j.1467-7687.2007.00667.x
- Smith, LB. Children's noun learning: how general learning processes make specialized learning mechanisms.. In: MacWhinney, B., editor. *The emergence of language*. Lawrence Erlbaum Associates; Mahwah, NJ: 1999. p. 277-303.