Linguistic knowledge and language acquisition

VIRGINIA VALIAN*

Columbia University

This paper presents several hypotheses about knowledge and knowledge acquisition that are relevant to problems of language acquisition, and in terms of them assesses one aspect of the study of language acquisition and makes suggestions about future progress.

Knowledge is a mental state, and may be explicit or tacit. Knowledge may be of, about, or that something. Knowledge is organized in terms of propositions whose elements are concepts. (The discussion excludes knowledge how, even though such knowledge plays an important role in language acquisition.)

We have and acquire knowledge, rather than knowing and learning knowledge. Learning is one method whereby we acquire knowledge. Why is it incorrect to say that we know or learn knowledge? It can be more easily seen with other propositional attitudes. Take desire. Say that one could characterize a rule system for desire that was mentally represented and via which people determined their desires. (Whether such a system exists is immaterial.) One would not say that one desires the rules: the rules allow one to derive desires; the rules are mentally represented; a person has the rules. Similarly for every propositional attitude: one can have the rules that constitute the attitude but one can not bear that attitude toward the rules.

There is good reason to think that knowing is like every other propositional attitude. Knowledge represents a relation between a person and something which can be known of, about, or that. Knowledge has a content which states what is known. Grammatical rules comprise one's knowledge of language. One has the knowledge, and therefore also has the rules which constitute the knowledge. If, however, one knew the rules, there would have to be another set of propositions comprising that knowledge. Whenever one attributes knowledge one must also attribute a set of propositions of which the knowledge consists. If one knows a rule then that knowledge must be characterized, and it cannot be done by using the same content as the rule. Only if one is said to have knowledge (in the form of having rules) rather than knowing knowledge (in the form of knowing rules) can the problem be avoided.

*I thank J. J. Katz for a discussion of these issues. Reprint requests should be sent to V. Valian, Psychology Department, Columbia University, New York, NY 10027, U. S. A.
The word 'learn' operates similarly. Just as the expression 'knowing knowledge is ill-formed, so is the expression 'learning knowledge'. The same sorts of things that can be known can be learned.

The state of having knowledge can arise either as part of initial mental structure, or as the consequence of acquiring knowledge. The controversial problems begin in characterizing how knowledge can be acquired. One hypothesis put forward here is that knowledge can only be acquired via learning or via derivation from prior knowledge. Another hypothesis is that there are two methods of learning, one being hypothesis-testing, the other being what I shall call fitting into or elaborating an already existent knowledge structure. Other candidates can either be reduced to one of these methods or ruled out. A third hypothesis is that there are different types of propositional knowledge, and different sources may be linked with different types. Taken together, the hypotheses suggest that each phenomenon of language acquisition is likely to be a mixed one. Thus, interpreting them as if they were unitary can result in incorrect rejection of certain learning methods.

The first hypothesis, that knowledge can only be acquired via learning or derivation from prior knowledge, contrasts with suggestions that knowledge can be acquired by taking a pill, re-wiring a brain, being hit on the head, or the like. It may also contrast with suggestions that knowledge can be 'grown' (Chomsky, 1980), depending on how that metaphor is interpreted. Nothing definitive can be said in favor of or against any of the positions, but there is an argument that shows the difficulty of being confident that there are no boundaries on possible methods of acquisition of knowledge.

There are two preliminary considerations. First, all conceivable methods require change in a mental state, and all methods outside of learning or derivation, both of which are mental processes, are methods involving change in a physical brain state which must be assumed either to cause or to be the same as change in a mental state.

Second, not all changes in brain states correspond to changes in mental states, so that proponents of being hit on the head as an acquisition method must show what kinds of head hits will work. Thus, even if one denied that knowledge was a mental state and claimed it was only a special kind of physical state, it would still be necessary to specify the conditions under which being hit on the head could induce this special kind of physical state. Taking a pill and neural re-wiring connote more precise and localized brain changes than concussion, but there is no difference in principle. Physical methods are also problematic in that we know of no cases where purely physical alterations have resulted in knowledge acquisition, or anything that looks like knowledge acquisition.
An apparent parallel exists in that not all mental processes result in knowledge acquisition, and thus proponents of learning must provide theories showing what kinds of mental changes result in knowledge acquisition. There is a difference, however, since 'learning' is, by definition, a method of knowledge acquisition.

The preliminary considerations to one side, here is the argument. One can theoretically mimic a process or state to any degree of precision. If there is a difference between knowing something and acting in such a way that is behaviorally indistinguishable from knowing something, then it must be demonstrated that a proposed method of acquisition delivers a knowledge state and not a state mimicking a knowledge state. Learning does so by definition, so that with learning the problem is to come up with a correct theory. Other proposed methods do not necessarily lead to knowledge.

One alternative is to deny the difference between knowing and acting in a manner behaviorally indistinguishable from knowing, with the consequence that a chess computer, say, and a chess player have the same relation to the rules of chess. If one denies the difference, however, there seems no point in talking about knowledge. One may as well talk merely of mechanisms which govern behavior.

An analogy can be made with ESP. When magicians show that by clever sleight of hand feats of telekinesis can be understood as trickery we do not say, 'Oh, that's what telekinesis is, clever sleight of hand'. Rather, we say, 'There's no such thing as telekinesis, it's all clever sleight of hand'. Clever sleight of hand is not an exposition of telekinesis. Thus, if we want to retain the concept of knowledge we must also maintain a difference between having knowledge and mimicking having knowledge. Given the reality of that distinction, an appropriate concern about a proposed acquisition model is whether it can in principle arrive at knowledge. Again, for a learning theory or a derivation theory, the in-principle question does not arise; we know knowledge can be acquired by learning or by derivation from prior knowledge. The only question is whether we have specified the method properly.

There seem, then, grounds for provisionally accepting the hypothesis that only learning and derivation from prior knowledge will allow knowledge acquisition. The second hypothesis, that there are only two methods of language learning, hypothesis-testing and fitting into a knowledge structure, cannot be

---

1 Chess rules are different from linguistic rules. One can know the rules of chess, unlike the rules of grammar, because chess rules are like facts about the language. Knowledge of the rules of chess is a trivial component of chess-playing: one does not derive moves from chess rules unless one does not know how to play. A chess player has another set of rules, implicit ones which are more similar to linguistic rules, that determine moves; that set the player does not know but has.
argued for here. Rather, the obvious methods which do not fit either model (e.g., conditioning, reinforcement, and imitation) are dismissed as already ruled out, and most remaining methods are taken as variants of hypothesis-testing or knowledge structure elaboration. Is ‘organ growth’ (Chomsky, 1980) a possible exception?

Organ growth may or may not be a learning theory. If it is not, it is subject to the problems mentioned above. Chomsky seems to suggest it is not by likening the language system to the visual system. The analogy is not helpful, however, because it is not clear that any knowledge is involved in either the development of the visual system, or as a product of that development. To task a simple example, in recognizing an object, and in knowing it is a chair, the knowledge is not in the visual system. One has learned that a certain visual configuration represents a chair, and made a link between cognition and perception. Thus, the knowledge situation is sufficiently confused in perception that making analogies with it does not clarify the status of linguistic knowledge and its acquisition.

Chomsky also, however, likens language learning to Peircean abduction which is a hypothesis-testing model with a heavily constrained initial hypothesis space. If that is a correct characterization, then an organ growth model would seem to be a learning theory.

All other language learning models appear to be variants of a hypothesis model, a knowledge-structure elaboration model, or a composite. Therefore, it will be useful to briefly characterize such models. A hypothesis-testing model has five components; it is in how the components are specified that the theories differ. The first component states what is acquired, the second what is innate, the third what the content of the hypotheses is, the fourth the role of experience. The fifth component delimits the constraints on the hypothesis-testing mechanism (e.g., how many rules it can change at one time, whether it can store unanalyzed strings, and so on) and the procedures the mechanism uses in testing hypotheses. The first two components highly constrain the hypothesis space; experience serves as confirming or disconfirming evidence.

\(^{2}\)Strategies (Bever, 1970) and operating principles (Slobin, 1973) have not been offered as learning theories. As Cromer (1976) points out, strategies might explain how children behave when they do not understand a structure but say nothing about how children learn a structure. Operating principles are based either on commonly-encountered properties of languages (e.g., 'avoid exceptions' reflects the fact that languages are systematic) or on commonly-encountered features of learners (e.g., 'avoid exceptions' reflects the fact that learners tend to form rules), or both. They are compatible with any type of learning theory. Piagetian principles have also not been offered as a language learning theory, although cognitive phenomena predicted by Piagetian theory often occur just prior to or concomitant with language phenomena, the significance of which is unclear.
Erreich, Valian, and Winzemer (1980) and Valian, Winzemer, and Erreich (1981) describe and illustrate hypothesis-testing models in some detail. They will therefore not be discussed further here, except to note that 'generalization' models seem to fall under the hypothesis-testing rubric, as do the models proposed within learnability theory (Hamburger, 1980; Wexler and Culicover, 1980). (These models have tended to emphasize the constraints on the hypothesis space and the constraints on the mechanism.)

Models which propose an elaboration of a knowledge structure are common in cognitive psychology (see Anderson, 1980 for review), but not in language acquisition. The existence of such models in part suggests the third hypothesis, that there are different types of propositional knowledge and different sources for the different types. We can roughly characterize two types: Type I includes laws and principles, and theories and explanations, and tends to be innate or acquired via hypothesis-testing; Type II includes single facts and systems of facts (like taxonomies) and tends to be acquired by elaborating a knowledge structure. The basic outline of such a model is that the concepts in which the new knowledge will be encoded are already present; learning involves hooking up a new combination, or new sub-combinations, and wherever possible connecting the new fact with prior knowledge to make it less isolated.

Many of the facts we acquire are 'direct' facts: they are presented in a manner highly similar to their ultimate mental representation. A newspaper article about the Academy Awards, for example, states who won an award for best screenplay. We know about screenplays, awards, movies, and so on, and can fit the new knowledge into our knowledge structure—about Hollywood, say—quite easily. Other facts, however, such as word meanings, are typically not presented directly: to fit them in with a knowledge structure may first require a procedure of hypothesis testing, in order to determine what the facts are.

In language acquisition both Type I and Type II knowledge must be acquired: many individual specific facts must be learned (e.g., the declensions of irregular verbs), as well as broader language-specific patterns that will be mentally accounted for in terms of rules and principles. It is thus quite likely that some knowledge is acquired via both methods of learning, and some via derivation from prior knowledge (as when a judgment can be made about, say, the ambiguity of a sentence never heard before).

If it is the case that different kinds of knowledge are being acquired simultaneously, by the operation of at least three different methods of knowledge acquisition and revealed in a performance mechanism which is also developing, each resulting phenomenon of language acquisition is likely to be a mixed bag,

---

Pinker (1981) seems to offer a mixed model, part hypothesis-testing, part knowledge structure elab-
and not susceptible to unitary interpretations. Attempts to force unitary interpretations can thus give misleading pictures of language acquisition. Salient superficial factors can appear to be the only factors. Usually, good performance requires knowledge of general rules plus specific facts, as in wh-question comprehension, which involves knowledge of \textit{wh}-movement and subject-aux inversion, and also of individual \textit{wh}-words and individual verbs (see Winzemer, 1981 for a working-out of this example). If a unitary interpretation is sought, the salience of generalizations that hold for performance on specific words can override attention to the concurrent acquisition of rules.

A second case occurs when different domains of knowledge are being acquired simultaneously. If the child is acquiring phonological, syntactic, semantic, and pragmatic knowledge simultaneously there will be numerous compatible descriptions of her behavior, all of which can be correct. If, however, the availability of several different descriptions is coupled with reductionism, then some descriptions will be incorrectly rejected. For example, syntactic and semantic descriptions are often available for the same phenomena (take Brown's 1973 discussion of syntactic and semantic cumulative complexity), which is what one would expect if both syntactic and semantic knowledge are being acquired simultaneously. But the tendency to see syntax as reducible to semantics results in syntactic descriptions being rejected if there is a semantic description available. (For further discussion see Valian, 1981.)

The current situation in the study of language acquisition is one in which the theories make few predictions about phenomena, and the phenomena radically underdetermine the theories. (Further, many of the theories are so non-specific that they cannot be evaluated on any grounds; notable exceptions include Hamburger, 1980, Wexler and Culicover, 1980, Pinker, 1981, Erreich, et al., 1980.) How can the theories and phenomena come in closer contact? The first suggestion is, paradoxically, that theorists temporarily put aside attempts to account for most current phenomena, and concentrate on deriving predictions from their theories about new phenomena. The problem with most known phenomena is that, being mixed, they frustrate efforts at theory construction.

An example is telegraphic speech, for which we tend to look for a single account. However, some word omissions may be due to one cause, while others have a different cause. Valian (1981) shows that the frequent absence of determiners in children's speech can be explained by the optionality of determiners within noun phrases. But the absence of \textit{be} as auxiliary must have a different explanation. Thus, 'telegraphic speech' may be one name covering several different processes, with one process linked to the acquisition of categories. Telegraphic speech, as a phenomenon, cannot helpfully constrain
learning theories if it is not actually a single phenomenon. Eventually the old phenomena will have to be explained or convincingly demonstrated to be outside the domain of language acquisition, but we will be in a better position to do that with stronger theories.

The second suggestion is directed to the need for phenomena that will bear more directly on learning theories, by being less mixed. There might be two ways of arriving at purer phenomena. One is to avoid clearly mixed phenomena unless guided by a theory which analyzes the types of knowledge and the methods of acquisition involved, or unless it is possible to control for all but one knowledge type and one acquisition method. An example of a clearly mixed phenomena is increasing MLU with age, which reflects all types of knowledge acquired in all ways. Thus, despite its robustness, the phenomenon is not theoretically useful.

Another way of arriving at purer phenomena is by successive approximations: one can analyze a phenomenon according to its probable knowledge types and acquisition methods, using the results to suggest future observations that will be localized to a particular knowledge and acquisition type. The phenomena will thereby be more directly relevant to the construction and confirmation of theories.

References


