THE INDUCTION OF MENTAL STRUCTURES WHILE LEARNING TO USE SYMBOLIC SYSTEMS

T. G. Bever and Ralph E. Hansen
University of Rochester

ABSTRACT

Subjects learned to map phrase-structure-defined strings onto geometric figure arrays. "String-generation" subjects produced symbol strings corresponding to arrays; "String-interpretation" subjects constructed arrays corresponding to strings. "Mixed" subjects alternated between these tasks. Subjects' knowledge of symbol sequence acceptability was periodically probed. Mixed subjects learned the structure dramatically faster than other subjects. This suggests that natural acquisition of structure underlying symbol-world mapping systems like language depends on learning multi-directional mappings.

* * * * * * * *

Humans inevitably construct implicit mental representations to guide their concrete actions and percepts. The research in this paper investigates a theory of the conditions which elicit such formal mental structures. The essence of the theory is that these structures are mental devices which integrate superficially distinct perspectives on a situation. Consider the intuitive formation of a map between home and work. If you always walk from home to work, but get a ride home after dark, there is little basis for developing an intuitive map of the relevant neighborhood. Rather, you can memorize a unilinear series of turns and distances. Contrast this with a situation in which you walk sometimes in each direction: it is intuitively clear (though the relevant research remains largely undone), that you are then much more likely to construct an intuitive map of the area in an abstract representation of the streets and crucial landmarks. The differing perspectives gained from walking in both directions stimulate the instinctive need for a mental representation which is neutral concerning the direction of travel - namely, a map.

The functional value of such inner mental representations is unquestionable, but the fact that the structures are often functional does not explain their existence, form, or the dynamics of their discovery. In fact, it is frequently the case that humans ascribe unnecessarily elaborate internal structures to superficially regular phenomena. For example, people often intuitively invoke a complex causal schema as the internal structure of a series of events, which might in fact be unrelated; similarly, people may develop overly elaborate hypotheses about the structure of machines, as they learn to control them; finally, people acquire complex grammars with independent interlocking levels of representation to account for the structure of sentences, which might have simpler behavioral descriptions. In each case, the more superficial analysis might be correct or at least more functional - the person who attributes an unnecessary underlying structure is guilty of a cognitive illusion, the intuitive formation of an incorrect mental representation. As in the study of perception, the importance of such illusions is that they demonstrate the presence of an active set of mental processes which automatically form mental representations during the organization of behavior, regardless of their specific functional role. The puzzle about why people induce complex abstract structures resolves into several component questions:

(1) Why do humans hypothesize the particular structures they do?
(2) What environmental conditions elicit the structures?
(3) What mental conditions elicit the structures?
(4) What motivates invoking the structure without direct reinforcement?

Questions (1) and (4) are usually taken to be the most profound: the former bears on hypotheses about innate constraints, the latter on the motives for active learning of abstract structures. It is difficult, however, to answer either question without a better understanding of the dynamics of the learning process. Accordingly, our research concentrates on the second and third questions: our theoretical goal is to understand the interactions between the environment and the learner's mental state which result in the formation of mental structures. Our practical goal is to develop some insights into the conditions that best elicit spontaneous formation of an appropriate mental representation for a situation.

Our theory of structure induction is rooted in the dynamic role of abstract representational schemata, as systems that resolve inconsistencies between superficial systems of representation (Bever, 1986). On this view, structure induction has some formal similarities to problem solving. It involves several phases; first, the formulation of distinct representations which seem to be inconsistent (the real mental 'problem'); then evocation of a more abstract representational schema which allows for the integration of the conflicting representations. An example of this is Duncker's (1945) classic explication of the solution for the use of x-rays to kill an internal tumor: at first, the subjects oscillate between postulating an x-ray 'gun' which shoots the tumor (but destroys the intervening tissue), and an x-ray 'bomb' which explodes only at the tumor site (but cannot get there because of the intervening tissue). The solution lies in an integration of features of both the 'gun' and the 'bomb'; a focussing lens disperses the x-rays, like light, harmlessly through the intervening tissue, and focusses them lethally only on the tumor. The concept of a lens which manipulates x-rays as though they were light provides a new schema in which to integrate the initially inconsistent representations of the problem.

On our view, the induction of structures underlying behavior works most effectively in an analogous way. Different superficial regularities, or different modes of use, stimulate the development of incompatible representations of the behavior: the deeper representation supplies a resolution of these apparently conflicting representations. This sequence of mental stages is a standard account of how children go about discovering elaborate mental systems, such as causal reasoning, number, naive physics, and so on. We are suggesting that the same kind of processes occur during adult learning of complex systems.

The use of symbol manipulation paradigms

Symbol-sequence learning offers a rich paradigm to examine the induction of an abstract structure from specific concrete training experiences. In these experiments, subjects typically are asked to discover the principles underlying the well-formedness of sequences. The paradigm offers the possibility of experimental control over the intermediate stages of structure formation, and careful probing for the ultimate structures.

In practice, such studies are often described as investigations of 'artificial language learning', since the description of the symbol sequences is often expressed in terms of language-like rules (Anderson, 1975; Braine, 1963, 1966; Esper, 1925; Miller, 1967; Miller and Stein, 1963; Moeser & Bregman, 1972; Reber, 1967; Saporta, Blumenthal and Reiff, 1963; Segal and Halwes, 1965, 1966; Smith, 1969). The rationale for these studies has usually been taken to be that one can include or violate formal properties of natural language in the artificial mapping systems; if the selective absence of a particular formal property makes the language hard to learn, then one might conclude that the property is a critical part of the structure of any language (Esper, 1925; Chomsky, 1965). That is, artificial languages can be used to test linguistic universals, one by one.

Such learning paradigms have been used to investigate some behavioral issues, primarily contrasting the importance of structural information about where phrases begin and end (Green, 1979; Morgan and Newport, 1981; Morgan, Meier and Newport, 1986) and the relative importance of parallels between grammatical structure and its extra-symbolic reference for learning (Moeser and Bregman, 1972, 1973; Moeser, 1977; Anderson, 1975; Meier and Bower, 1986).

These studies of symbol sequence learning have had the same limitations as other paradigms one might use to explore naturalistic discovery of structure. They pose subjects directly with the problem of discovering the structures at issue, rather than placing subjects in tasks which might naturally elicit the structures as implicit components of the solution to the tasks. Below, we outline an adaptation of artificial language learning paradigms in which subjects learn to <u>use</u> the symbolic structure without being asked to learn the structure explicitly. Our initial results suggest that the paradigm can be used to answer questions about the behavioral conditions governing the discovery of implicit structures in general.

THE STUDY

In our study, subjects learn to use symbols in sequences defined by simple structural constraints, and are not given any direct training on structural well-formedness. As they learn to use the symbol-sequences, we periodically test their knowledge of the structure of the symbol system. Subjects are either asked to learn to 'produce' symbol sequences correctly or to 'perceive' them. In 'production', the subjects are given a visual array of shapes on a computer screen and must type the sequence describing the array. In 'perception', subjects are given a sequence and must construct the visual array which it describes.

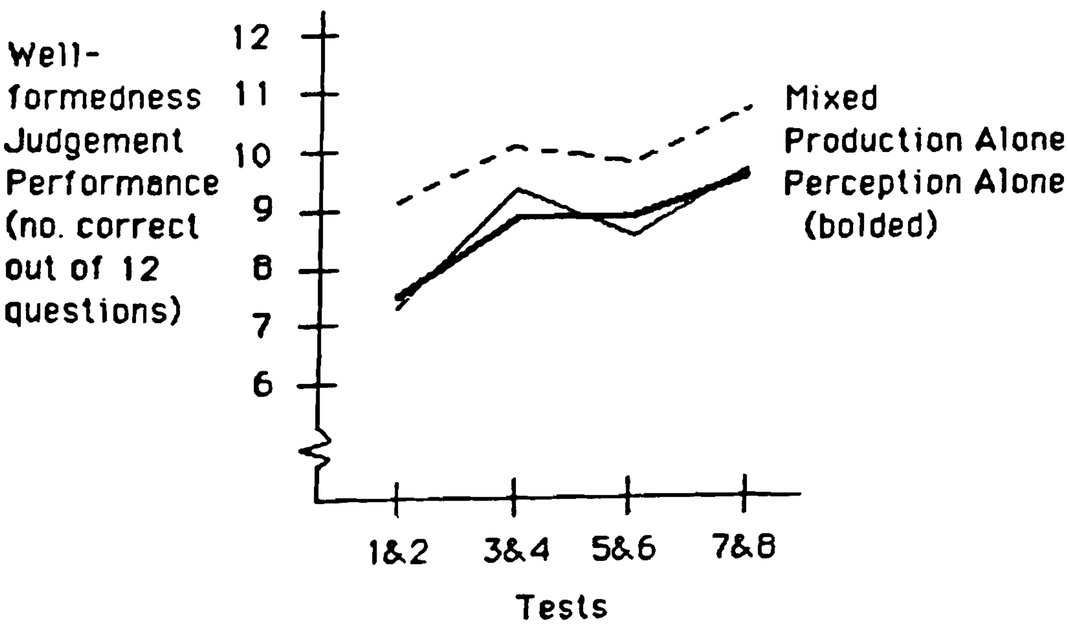
The symbol sequences are structured according to rules taken from a standard artificial language used in previous studies in the literature (e.g., Anderson, 1975; Meier and Bower, 1986). It is a phrase structure verb-final language, with embedding. Sequences ranged from 4 to 12 words long; the separate words are in English ('triangle, red, large, above' etc.) except for one grammatical function word, 'te'. The visual pattern is defined on a set of geometric shapes which can have different sizes and colors. Each shape can be located in one of four quadrants on a computer screen. Subjects can 'paint' the figure they want in each quadrant separately, using an adapted graphics package: they are provided with labelled buttons for each figure attribute (triangle, red, large, etc.). Even this simple mapping system allows for complex mappings. For example, a large striped triangle above a red circle, which is to the left of a small striped square, would be denoted by the string,

triangle large striped circle te red te square small striped left-of above

The subjects are never trained on isolated sequences; rather, they are exposed to unidirectional mapping tasks which naturally reflect the normal uses for symbol systems, comprehension and production. In all conditions, subjects are pre-trained in mapping isolated symbols, to become familiar with using the computer-controlled printing and drawing techniques.

We ran groups of 10 subjects, balanced for such variables as SAT scores, sex and age, in each of three paradigms: 1) 'perception': on each trial, subjects are given a symbol sequence and asked to construct the corresponding visual pattern; 2) 'production': on each trial, subjects are given the visual pattern and asked to construct a corresponding symbol sequence; and 3) 'mixed'; trials alternate between 'perception' and 'production'. There were 48 trials, selected to balance for various complexity variables in each 1/8 of the experimental session. After every 6 trials, subjects were presented with a test of their knowledge of the structure of well-formed sequences. Subjects judged which member of each of six pairs of sequences is structurally correct. Following previous research, the correct sequence in each pair was mated to an incorrect sequence which violated one of six kinds of structural properties characteristic of the system.

FIGURE 1



There are two measures of performance: the acquisition of structural knowledge, and the acquisition of the unidirectional mapping skills. Figure 1 shows how the structural knowledge increased with training in the three different training conditions. Training in production resulted in slightly faster acquisition of structural knowledge than did training in perception, but the difference is not significant. Most striking is the fact that the mixed condition resulted in superior mastery of the structure of the symbol system ($p<.025$, by Fisher exact test on subjects, both comparing mixed against perception alone and against production alone ($p<.03$ by a Wilcoxon matched-pairs signed-ranks across trials). This finding is not obvious: for example, one might have predicted that structural learning in the mixed condition would be the average of that in the two separate conditions. Furthermore, correlations of subjects' performance on the second half of the session shows that structural knowledge correlated strongly with production and perception in the mixed condition, but less strongly with production or perception alone (see Table 1, below). It seems clear that the mixed condition elicited a more unified representation of the structure with the behavioral skill.

TABLE 1

Correlations across subjects between mapping skill and well-formedness judgments (second half of sessions).

Production alone X judgment	.64 ($p<.02$)
Production mixed X judgment	.93 ($p<.001$)
Perception alone X judgment	.61 ($p<.03$)
Perception mixed X judgment	.80 ($p<.001$)

FIGURE 2

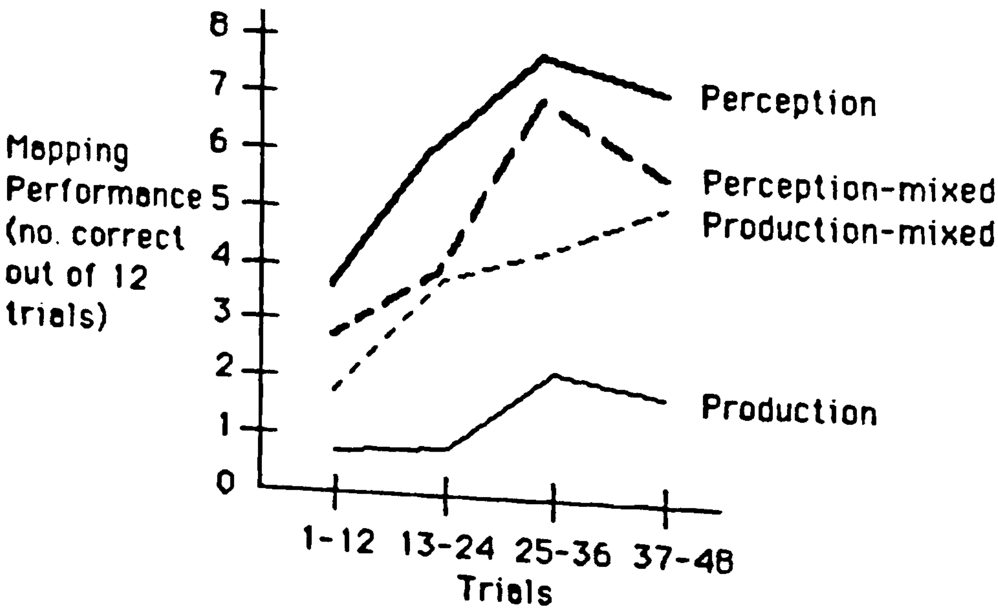


Figure 2 presents the acquisition of just the mapping skills in the different conditions. Production alone is clearly more difficult to master than perception alone; but production in the mixed condition is almost as easy as perception (we took performance on the last 12 trials as a measure: on this, production alone was more difficult than each of the other three conditions, $p<.025$, by a Fisher exact test across subjects; $p<.02$ by a Wilcoxon matched-pairs signed-ranks test on trials). Finally, perception in the mixed condition is no easier than perception alone. Several aspects of the results suggest that the acquisition of production is more directly related to the ability to make structural judgments than the acquisition of perception. First, the correlation across subjects between mapping performance and structural judgments is higher for production than perception in the mixed condition; second, the correlation across structural properties in structural judgement is higher for production and mixed, than for perception and mixed. These trends were not statistically significant given the current number of subjects and constraints, but they are suggestive as the basis for further research.

The initial findings from this study have a number of implications. First, the fact that structural knowledge is arrived at much more quickly when learning to map in both directions suggests that structural knowledge may be discovered as an integrated solution to multiple representational constraints. This follows as a special case of our original hypothesis that structure induction is facilitated if it provides a framework for incompatible systems of representation - clearly, perception and production in our paradigm involve distinct input/output relations. This contrast is emphasized by the fact that structural properties that are hard to master in one mode tend to be easier in the other. The mixed condition may be effective for independent reasons, for example, because the conflicting generalizations are in different modalities. The study of this has exciting implications for theories of the induction of structure not just of toy rule systems, but of such mental abstractions as causal reasoning, and such concrete objects as complex machines, computer algorithms and so on.

References

Anderson, J.R. (1975). Computer simulation of a language acquisition system: A first report. In R. L. Solso (Ed.), Information processing and cognition: The Loyola Symposium (pp. 295-349).

Bever, T.B. (1987). The aesthetic constraint on cognitive structures. In M. Brand and R. Harnish (Eds.), The representation of knowledge and belief. Tuscon, AZ: University of Arizona Press.

Bever, T.G., Fodor, J.A., & Weksel, W. (1965). On the acquisition of syntax: A critique of "contextual generalization." Psychological Review, 72, 467-482.

Braine, M.D.S. (1963). On learning the grammatical order of words. Psychological Review, 70, 323-348.

Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, MA: MIT Press.

Duncker, K. (1945). On problem-soilving (translated by L.S. Lees). Psychological Monographs, 58, No. 270.

Esper, E.A. (1925). A technique for the experimental investigation of associative interference in artificial linguistic material. Language Monographs, No. 1.

Green, T.R.G. (1979). The necessity of syntax markers: Two experiments with artificial languages. Journal of Verbal Learning and Verbal Behavior, 18, 481-496.

Lenneberg, E.H. (1967). Biological foundations of language. New York, NY: Wiley.

Meier, R.P., & Bower, G.H. (1986). Semantic reference in phrasal grouping in the acquisition of a miniature phrase structure language. Journal of Memory and Language.

Miller, G.A., & Stein, M. (1963). Grammarama I: Preliminary studies and analysis of protocols. Scientific Report #CS-2, Harvard University Center for Cognitive Studies.

Miller, G.A. (1967). The psychology of communication. New York, NY: Basic Books

Moeser, S.D. (1977). Semantics and miniature artificial languages. In J. Macnamara (Ed.), Language learning and thought. New York, NY: Academic Press.

Moeser, S.D., & Bregman, A.S. (1972). The role of reference in the acquisition of a miniature artificial language. Journal of Verbal Learning and Verbal Behavior, 11, 759-769.

Morgan, J.L., Meier, R.P., and Newport, E.L. (1986). Structural packaging in the input to language learning: Contributions of intonational and morphological marking of phrases to the acquisition of language. (Under Review)

Morgan, J.L., & Newport, E.L. (1981). The role of constituent structure in the induction of an artificial language. Journal of Verbal Learning and Verbal Behavior, 20, 67-85.

Newport, E.L. (1982). Task specificity in language learning? Evidence from speech perception and American Sign Language. In E. Wanner and L.R. Gleitman (Eds.), Lanaguage acquisition: The state of the art (pp. 450-486). Cambridge: Cambridge University Press.

Reber, A.S. (1967). Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior, 5, 855-863.

Saporta, S., Blumenthal, A.L., & Reiff, D.G. (1963). Grammatical models and language learning. In R.J. DiPietro (Ed.), Report of the Fourteenth Annual Round Table Meeting on Linguistics and Language Studies, (pp. 133-142). Georgetown University Monograph Series on Languages and Linguistics.

Segal, E.M., & Halwes, T.G. (1966). The influence of frequency of exposure on the learning of phrase structure grammar. Psychonomic Science, 4, 157-158.

Smith, K.H. (1969). Learning co-occurrence restrictions: Rule learning or rote learning? Journal of Verbal Learning and Verbal Behavior, 8, 319-321.