The Role of Reference in the Acquisition of a Miniature Artificial Language

SHANNON D. MOESER and ALBERT S. BREGMAN

McGill University, Montreal 101, Quebec

College students learned a miniature phrase-structural language employing nonsense syllables as words in one of four conditions: in the first, the language had no reference; in the second, each sentence referred to a picture created out of geometric forms, but there was an arbitrary association of words with pictorial features; in the third, there was a similar reference field and words belonging to the same syntactic class referred to items belonging to the same visual class; the fourth was like the third, but in addition, syntactic constraints in sentences mirrored logical constraints in the pictures. Learning of complex features of syntax was possible only when these reflected properties of, or constraints in, the reference field.

Although the theory of language has traditionally distinguished between syntax and semantics, it is possible that an interaction between the two may affect language acquisition. Using an artificial language that formed a miniature linguistic system, Moeser (1969) found that subjects presented with a language containing semantic referents appeared to learn the language syntax in a different way from subjects presented with the same language without the semantic referents.

A miniature linguistic system attempts to experimentally reproduce some of the processes of language learning employing the older child or adult as a subject. The subject’s task is similar to that of the child learning his first language; he is presented with only a limited subset of all grammatically correct utterances from the language and he must somehow discover how to create new grammatically correct utterances. Using semantically empty systems, it has been found that subjects can learn that a correct utterance consists of one word from one word class and one word from a second word class in a specific word-class order (for instance, Braine, 1963; Smith, 1966). Attempts to teach subjects that the occurrence of one word class in a sentence depends on the presence or absence of another word class (selection restrictions) have not been so successful. Braine (1965, 1966) has demonstrated such learning when the contingent word classes consist of one word each, or when they involve a contingency between a class and a single item. However, Smith (1966), and Segal and Halwes (1965, 1966) failed to find any learning of selection restrictions with a semantically empty system.

The system used by Moeser (1969) employed word classes and selection restrictions defined not only by sentence position and privileges of occurrence, but also by a set of semantic referents (visual forms or relations that were correlated to the words of the language). Subjects exposed to sentences and their semantic referents performed significantly better on tests of syntax than subjects exposed only to the sentences; a few of the subjects exposed to semantic referents were able to
show above-chance learning of the selection restrictions but none of the subjects presented with the semantically empty language did likewise. It was thus suggested that the learning of syntactic aspects of language might be dependent upon the interaction of the semantic and syntactic aspects.

From conversations with subjects and tape recordings taken while learning was in progress it appeared that subjects in the semantic referent condition approached the task of learning the system in a different way from those subjects given only the words. They concentrated on learning to associate each word with its referent and showed no interest in learning the grammatical relations of the words (for instance, in learning the positions that words occupied in a sentence). These subjects appeared to use the semantic referents to mediate their learning of the language syntax. On the basis of these results, it was decided to undertake further investigation of this question using a modification of the Moeser (1969) experimental design.

For the present experiment, four reference conditions have been designed so that increasing amounts of information about syntax are supplied by relations among the referent objects. If, as proposed, subjects learn language syntax via mediation through the semantic referents, it can be predicted that the more syntactical information that is supplied in the semantic references, the better the language rules will be learned. Furthermore, it is suggested that if rules regarding selection restrictions are not reflected in the semantic referents these rules will not be learned.

In addition, three systems will be utilized, designed so that each language contains the same number of words but differs in syntactic complexity. It is predicted that subjects who are presented with the language in which relations among semantic referents are not reflected in the syntactic rules will find it increasingly more difficult to learn the systems as they increase in complexity. However, when the semantic referents mirror the syntax rules there should be little difference in performance on the languages. This is because it is assumed that before learning to map the words on the sentence positions, subjects in semantic reference conditions must first learn (a) to associate each word with its visual referent and (b) the ways in which the visual referents are organized (that is, learn that visual stimuli can be grouped into classes on the basis of their physical attributes and privileges of occurrence, that certain classes of stimuli always occur, that certain classes must occur or can only occur when a specific visual arrangement takes place, that only certain visual arrangements can take place). If (a) and (b) are held nearly constant, the subjects should experience almost equal difficulty even when the syntax differs in complexity.

**METHOD**

*Description of the Miniature Linguistic Systems*

Three systems were employed, all utilizing the same fourteen CVCs (having a moderately high association value of between seventy and eighty according to Archer, 1960). These fourteen "words" (CVCs) were grouped into four classes, with four words each occurring in classes A, B, and C, and two words occurring in class D.

In the least complex system, system 1, the syntactic structure of the language is described as follows:

\[ S \rightarrow AP + BP + (CP); \]
\[ AP \rightarrow A + (D); \]
\[ BP \rightarrow B + (CP); \]
\[ CP \rightarrow C. \]

In other words, a grammatical sentence in system 1 has to contain one word each from classes A and B; the class A word can be followed by one word from class D and the class B word can be followed by one or two words from class C. Thus, six acceptable sentence structures are possible in system 1.

In system 2 the syntactic structure of the language can be described as follows:

\[ S \rightarrow AP + BP + (CP); \]
\[ AP \rightarrow A + (D); \]
\[ BP \rightarrow [B_1 + CP]; \]
\[ CP \rightarrow C. \]

This language is identical to system 1 except that the B-phrase (BP) contains an option. Either a B_1 word can
be used alone or a B₁ word can be used in combination with a C phrase (CP). Thus, the number of C words used in a sentence depends on the type of B word used. If a B₁ word is used, zero or one C word can appear; if a B₂ word is used, one or two C words can appear. There are two B₁ words and two B₂ words. There are eight acceptable sentence structures possible in system 2.

In the most complex system, system 3 the syntactic structure of the language is described as:

\[ S \rightarrow AP + BP + (CP); \]
\[ AP \rightarrow A + (D); \]
\[ BP \rightarrow \{B₁ \}
\[ \{B₁ + B₂ \}
\[ CP \rightarrow C + (D). \]

This language is identical to system 2 except that the C phrase (CP) contains an option, either a C word can be used alone or it can be followed by a D word. There are eighteen acceptable sentence structures possible in system 3.

Thus, the three systems show increasing syntactic complexity: all three have the same S rule; system 2 differs from system 1 in that it has a more complex BP rule; system 3 keeps the BP rule complexity of system 2 and complicates the CP rule as well. Another way of viewing the differences between these three languages is to see them as adding progressively more nonterminal elements. System 1 has no nonterminal elements (CP could simply be called C, and BP simply B + (C), if we wished); system 2 requires BP but not CP; and system 3 requires both BP and CP.

It has been pointed out to us that since the above languages are finite and nonrecursive, they could all be written as simple finite-state grammars. Thus the phrase-structure grammars we have written above are not unique representations of the restrictions on the selection and ordering of the words. However these grammars will be seen to be a particularly appropriate way of describing the languages when their semantic reference is discussed below. Thus our presentation of the grammars in phrase-structure format was motivated by semantic as well as syntactic considerations.

**Description of the Reference Conditions**

There were four reference conditions in each language under which subjects were tested: (a) *Words-only*, in which only the words of the language are presented; (b) *Arbitrary figures*, in which figures are presented along with the words of the language but the correspondence between words and figures is arbitrary—the reference figures act simply as a restatement of the information carried in the words; (c) *Class correspondence*, in which the figures presented along with the words of the language incorporate into their design visual features defining class membership; and (d) *Syntax correlation*, in which the figures presented along with the words of the language incorporate into their design both visual features defining class membership and visual reasons for the selection restrictions. In the three reference conditions in which figures are presented, the spatial order of the figures is identical to the order of the words that represent them.

An illustration of the words and referents used in the four reference conditions of system 1 is shown in Figure 1 and an example of how one sentence would appear in each of the four reference conditions is given in Figure 2. From these it can be seen that in the *syntax correlation* condition all A-class words refer to colored rectangles, all B-class words refer to orientations of these colored rectangles, all D-class words refer to line variations of the colored rectangles, and all C-class words refer to separate geometrical figures. The A, D, B group of referents form a single perceptual entity and because of the very nature of these referents neither a B referent nor a D referent can appear unless an A referent is also present (a fact that mirrors the syntactic constraints of the system). In the *class correspondence* condition, words are illustrated by figures similar in design to those used in the syntax correspondence condition, but the referents for A words, B words, and D words are presented as separate units, instead of as one perceptual entity. Thus there is nothing in the design of the visual field that would prevent a D referent or a B referent from appearing without an A referent. In the *arbitrary figures* condition the same separate figures are used as are found in the class correspondence condition but there is no systematic pairing of a particular class of words with a particular class of figures. All pairings are random.

The three conditions of systems 2 and 3 are related in an analogous way to one another. In the *words-only* condition, obviously only the syntax of the permissible sentences can be different from system 1 (since there are no referents). In the *arbitrary figures* condition, the same assignment of word to figure was used as in system 1, the choice and order of these now being determined, of course, by the more complex syntactic rules of systems 2 and 3. In the *class correspondence* condition, exactly as in system 1, the class of figures referred to by any syntactic class is marked by a definite perceptual feature. The same assignments are used as in system 1. It is in the *syntax correlation* condition that the assignment of figures to words differs in the three systems. In systems 2 and 3, the B class subdivides into B₁ and B₂ word classes, B₂ requiring a CP phrase following it. In the domain of figures this constraint is reflected as follows: B₂ referents now become relationships between A and C referents. The B₂ word BIF now refers to the presence of a geometric figure (C referent) above a colored rectangle (A referent) and the B₂ word ZOR now refers to the presence of a geometric figure...
Fig. 1. Illustration of the four reference conditions in system 1.

Fig. 2. Example of how one sentence appeared in each of the reference conditions in system 1.
(C referent) below a colored rectangle (A referent). Thus the very nature of the $B_2$ referent requires the C referent to be present. On the other hand, the $B_1$ referent, an orientation of the colored rectangles (A referents), does not imply a C referent. Thus the reference properties of $B_1$ and $B_2$ words in system mirror the more complex BP rule.

The more complex CP rule of system 3 introduces a second D word, conditional upon a C word's presence. By making the D referent be a border variation in a figure, it becomes obvious that a second border variation requires a second border, that is, it requires a C referent (uncolored geometric figure).

Materials

The stimulus materials consisted of eighty colored Kodachrome II slides projected on a screen by a Kodak 850 projector that automatically changed the slides, projecting each slide for 6.5 seconds with a .8-second interval between slides. For the three reference conditions, the pictures occupied the upper two-thirds of the slide and the words the lower one-third; for the words-only condition, the sentences were centered.

For each system, 40 different sentences were used, ranging from two to five words in length. Each word was presented ten times in this group of forty and all sentences were grammatically correct. In each presentation of eighty sentences there were two different series of these forty sentences, with the slides arranged within each series such that each word appeared on two consecutive slides exactly two times. All subjects given the same system received the same order of slides.

Subjects were given four presentations of the 80-slide series. After each presentation of eighty slides a test was given to measure the subject's progress in learning the language. The tests were presented on paper to the subject and no figures appeared in the tests, only sentences. The four tests were identical in form but differed in terms of the sentences used. No sentence used in a test had previously been seen by the subjects.

Each test in system 1 consisted of fifteen pairs of multiple-choice questions from which the subjects were requested to choose the correct sentence in each pair. Only one error was inserted in each incorrect sentence. In each test there were three questions testing each of the following rules:

1. An A word must appear in a sentence.
   Example: (a) GAV SIV ZOR RUD*;
   (b) LIM SIV ZOR RUD.

2. Only one A word can appear in a sentence.
   Example: (a) GAV PAX LIM*;
   (b) GAV KUS PAX LIM.

3. A B word must appear in a sentence.
   Example: (a) DEP PAX NAK*;
   (b) DEP NAK LIM.

4. Only one B word can appear in a sentence.
   Example: (a) KUS SIV BIF PAX RUD;
   (b) KUS SIV BIF RUD NAK*.

5. A D word can only follow an A word.
   Example: (a) KUS YOW ZOR LIM*;
   (b) KUS ZOR YOW LIM.

The sentences followed by asterisks are the correct choices.

The tests used for system 2 were identical to those used in system 1 except there were now twenty-one pairs of multiple-choice questions. Each test contained three questions testing each of the above five rules plus three questions testing each of the following rules:

6. Two C words cannot appear with a B word.
   Example: (a) KUS MUL*
   (b) KUS MUL COZ RUD.

7. Zero C words cannot appear with a B word.
   Example: (a) GAV BIF RUD NAK*
   (b) GAV BIF.

The tests used for system 3 were identical to those used for system 2 except that rule 5 became "A D word can only follow an A word or a C word."

For all systems, rules 1, 2, 3, and 4 were concerned with learning that a correct utterance consists of one word of one word class and one word from a second word class. Rule 5 was essentially a syntactic rule concerned with learning the position of a word-class in a sentence in relation to other word-classes. Rules 6 and 7 were concerned with learning the selection restrictions contained in the two more-complex systems.

Subjects

Subjects were one hundred and twenty college students between the ages of twenty and twenty-six. All were native speakers of English. Five male and five female subjects were assigned randomly to each group. They were paid $2.50 to participate in the experiment.

Procedure

The subjects were tested in groups of two and three. As soon as they were comfortably seated, those subjects in the three reference conditions were told:

This is a language learning experiment using an artificial language which refers to an artificial environment. The language has a grammatical structure which is not English or French or any other real language. It is your job to discover what this grammatical structure is. I will present this language to you on slides which will be shown on the screen in front of you. Above the words of the language are pictures to which the words refer. I want you to learn what each word refers to.

In the words-only condition, subjects were told:
This is a language learning experiment using an artificial language. The language has a grammatical structure which is not English or French or any other real language. It is your job to discover what this grammatical structure is. I will present this language to you on slides which will be shown on the screen in front of you. After a while you should see a systematic pattern in the way the words are arranged. I want you to discover this pattern.

After this introduction the subjects were presented with the first set of eighty slides. Then they were given fifteen minutes in which to complete the first test. When the fifteen minutes had passed they were presented with the eighty slides again, then given the second test. This procedure was repeated for tests 3 and 4.

RESULTS

All subjects in systems 2 and 3 received twenty-one questions in each of their four tests, whereas those subjects given system 1 had only fifteen questions in each test. These fifteen questions in system 1 are, however, comparable to fifteen questions in systems 2 and 3 when the questions testing rules 6 and 7 are omitted from the two more complex languages. Thus the total $3 \times 4$ cell matrix consisting of the three systems and four reference conditions can be compared on the set of fifteen questions that test rules 1 to 5. Also a $2 \times 4$ matrix consisting of the two more complex languages, systems 2 and 3, and the four reference conditions can be compared on the set of twenty-one questions that test rules 1 to 7.

Mean Total Scores

The mean total correct scores on all four tests are shown in Table 1. Chance score for the fifteen questions on the four tests would be thirty; chance score for the twenty-one questions would be forty-two.

Several a priori hypotheses were made in the introduction regarding the ease with which the systems could be learned under the different reference conditions. First, it was predicted that the difficulty in learning the lang-

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEANS AND STANDARD DEVIATIONS OF THE TOTAL TEST SCORES</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>Fifteen questions</th>
<th>Twenty-one questions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>Syntax correlation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 1</td>
<td>55.1</td>
<td>3.67</td>
</tr>
<tr>
<td>System 2</td>
<td>56.6</td>
<td>3.57</td>
</tr>
<tr>
<td>System 3</td>
<td>55.4</td>
<td>3.17</td>
</tr>
<tr>
<td>Class correspondence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 1</td>
<td>53.5</td>
<td>4.17</td>
</tr>
<tr>
<td>System 2</td>
<td>53.6</td>
<td>2.95</td>
</tr>
<tr>
<td>System 3</td>
<td>47.7</td>
<td>7.48</td>
</tr>
<tr>
<td>Arbitrary figures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 1</td>
<td>43.0</td>
<td>3.59</td>
</tr>
<tr>
<td>System 2</td>
<td>45.0</td>
<td>6.29</td>
</tr>
<tr>
<td>System 3</td>
<td>35.3</td>
<td>4.95</td>
</tr>
<tr>
<td>Words-only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 1</td>
<td>47.2</td>
<td>4.37</td>
</tr>
<tr>
<td>System 2</td>
<td>44.1</td>
<td>6.35</td>
</tr>
<tr>
<td>System 3</td>
<td>41.2</td>
<td>7.10</td>
</tr>
</tbody>
</table>
The language would be affected both by the reference conditions and by the complexity of the system. A two-way analysis of variance comparing the four reference conditions and the three levels of complexity on the fifteen questions indicated significant main effects due to reference conditions, \( F(3, 108) = 52.91, p < .001 \), and levels of complexity, \( F(2, 108) = 12.41, p < .001 \). The overall interaction between conditions and complexity was not significant, \( F(6, 108) = 2.06, p > .05 \). Similarly, a two-way analysis of variance comparing the four reference conditions and two levels of complexity (systems 2 and 3) on the twenty-one questions indicated significant main effects due to reference conditions, \( F(3, 72) = 77.96, p < .001 \), and levels of complexity, \( F(1, 72) = 21.55, p < .001 \). In addition, the overall interaction between conditions and complexity was significant, \( F(3, 72) = 3.09, p < .05 \). With regard to the two more complex languages at least, the difficulty in learning involved an interaction between language complexity and type of semantic reference.

The analysis of variance shows that the independent variables significantly affected the results. However, our a priori hypotheses not only predicted that there would be a difference among the different reference conditions and levels of complexity but predicted a complex pattern of results including the directions in which these differences would occur, and, in addition, the nonoccurrence of differences among certain conditions. The analysis of variance is a clumsy tool for looking at definite patterns in data. Yet we need some way of screening the data to see whether a particular difference between two means is within the “noise” range. We chose to apply a \( t \) test to each pair of means we were interested in, treating the resultant two-tailed \( p \) value as a readily understandable signal-to-noise index. This detailed analysis was applied only to the fifteen-question data of Table 1.

First, it was predicted that if syntax is learned via mediation through semantic reference, the more closely the semantic reference mirrored the syntactic constraints, the better would be the subjects’ performance. Thus the expected ordering of performance is syntax correlation best, class correspondence next, and then (in no predicted order) words-only and arbitrary figures. This predicted pattern leads to fifteen specific predictions of differences between means in the three languages. Of these, twelve yielded \( p \) values less than .001, one less than .01, one less than .05, and one nonsignificant but in the right direction.

Second, it was predicted that in the syntax correlation condition, where all of the syntax rules were mirrored in the semantic referents, there should be no difference in performance from the least to the most complex language. This prediction was supported by three nonsignificant differences between the systems (with \( t \) values of less than 1.00).

Third, the words-only and arbitrary figures should show a monotonic decrease in performance as the language grows more complex, since there was no mediation of syntax learning. This yields six predictions of difference. For the six tests, one gave a \( p \) value of less than .001, two gave values of less than .01, two gave nonsignificant \( p \) values with differences in the right direction, and one difference was in the wrong direction (a reversal of the predicted superiority of system 1 over system 2 in the arbitrary figures condition), but not significantly. In the class correspondence condition partial semantic mediation was possible, so we do not expect as sharp a drop in performance with increased system complexity. While no \( t \) tests were done, it appears that system 2 was learned as easily as system 1, and only in system 3 was there a drop in performance. Generally speaking, the interaction between language complexity and reference condition was not as strong as expected, the syntax correlation condition being the only case where language complexity seemed not to matter.

**Total Rule Scores**

In the three systems, there were five common rules that were tested, rules 1 through 5 in the
Materials section. It was considered possible that some of these rules might be more difficult to learn than others. Thus a three-way analysis of variance comparing the three levels of complexity (systems), four reference conditions, and five types of rules indicated a significant main effect due to the rules, \( F(1, 108) = 71.94, p < .001 \), a significant interaction between the levels (systems) and rules, \( F(1, 108) = 7.21, p < .01 \), and a significant interaction between the reference conditions and rules, \( F(1, 108) = 5.16, p < .05 \). The overall interaction among levels, conditions, and rules was not significant, \( F(1, 108) = 1.61, p > .05 \). Thus it appears that the case with which the five syntax rules were learned is not equal and that the possibilities of learning these various rules was affected both by the reference condition and by the level of language complexity.

For systems 2 and 3 there were seven rules tested: Rules 1 through 7 in the Materials section. A three-way analysis of variance comparing the two levels of complexity (systems 2 and 3), four reference conditions, and seven types of rules indicated a significant main effect due to the rules, \( F(1, 72) = 46.17, p < .001 \), a significant interaction between the levels and rules, \( F(1, 72) = 4.56, p < .05 \), and a significant interaction between the reference conditions and rules, \( F(1, 72) = 5.36, p < .05 \). The overall interaction among levels, conditions, and rules was not significant, \( F(1, 72) = 1.16, p > .05 \).

Table 2 shows the mean correct scores for all seven rules under the four reference conditions. From this table it can be seen that in the syntax correlation condition the scores were only slightly related to the type of error tested or to the level of language complexity; the mean

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MEAN CORRECT SCORES FOR THE SEVEN RULES</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 1</td>
<td>11.5</td>
<td>11.9</td>
<td>11.3</td>
<td>10.0</td>
<td>10.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 2</td>
<td>11.5</td>
<td>11.6</td>
<td>11.0</td>
<td>11.0</td>
<td>11.5</td>
<td>11.7</td>
<td>10.6</td>
</tr>
<tr>
<td>System 3</td>
<td>11.9</td>
<td>11.6</td>
<td>11.3</td>
<td>10.4</td>
<td>10.2</td>
<td>11.7</td>
<td>11.0</td>
</tr>
<tr>
<td>Class correspondence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 1</td>
<td>11.5</td>
<td>11.9</td>
<td>8.9</td>
<td>9.7</td>
<td>11.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 2</td>
<td>11.5</td>
<td>11.6</td>
<td>9.6</td>
<td>10.1</td>
<td>10.8</td>
<td>6.8</td>
<td>6.8</td>
</tr>
<tr>
<td>System 3</td>
<td>10.8</td>
<td>11.6</td>
<td>7.8</td>
<td>9.0</td>
<td>7.3</td>
<td>8.6</td>
<td>4.6</td>
</tr>
<tr>
<td>Arbitary figures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 1</td>
<td>9.4</td>
<td>10.8</td>
<td>6.2</td>
<td>7.4</td>
<td>9.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 2</td>
<td>9.3</td>
<td>10.8</td>
<td>7.7</td>
<td>8.4</td>
<td>8.8</td>
<td>7.5</td>
<td>6.1</td>
</tr>
<tr>
<td>System 3</td>
<td>7.3</td>
<td>8.6</td>
<td>6.7</td>
<td>7.1</td>
<td>5.6</td>
<td>6.8</td>
<td>5.2</td>
</tr>
<tr>
<td>Words-only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 1</td>
<td>10.6</td>
<td>11.0</td>
<td>7.9</td>
<td>7.9</td>
<td>9.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System 2</td>
<td>9.2</td>
<td>10.9</td>
<td>6.7</td>
<td>8.8</td>
<td>8.5</td>
<td>8.0</td>
<td>5.9</td>
</tr>
<tr>
<td>System 3</td>
<td>9.3</td>
<td>11.3</td>
<td>6.0</td>
<td>8.3</td>
<td>6.3</td>
<td>7.3</td>
<td>5.3</td>
</tr>
</tbody>
</table>

* Chance score is 7.5.
scores in all cells are roughly the same. (The maximum possible score for any rule was fifteen.) However, in the other three conditions, both the type of rule and the level of language complexity affected performance. In the class correspondence condition, scores were relatively stable for all three systems for rules 1, 2, and 4; but rules 6 and 7 suffered considerably in system 2, and rules 3, 5, and 7 similarly showed lower performance levels in system 3. In the arbitrary figures condition, Rules 3 and 4 showed lower performance levels in system 1, Rules 3, 6, and 7 showed lower performance levels in system 2, and performance on all of the rules in system 3 was only at about chance level. In the words-only condition, rules 3 and 4 showed lower performance levels in system 1, rules 3 and 7 showed lower performance levels in system 2, and rules 3, 5, 6, and 7 showed lower performance levels in system 3.

Discussion

The results of this experiment support the hypothesis that the learning strategies employed when semantic referents are present differ from the learning strategies employed when semantic referents are not present.

Chomsky (1965, page 33) has written:

Thus it has been found that semantic reference may greatly facilitate performance in a syntax-learning experiment, even though it does not, apparently, affect the manner in which the acquisition of syntax proceeds; that is, it plays no role in determining which hypotheses are selected by the learner.

From this, Chomsky concludes that the language acquisition device (LAD) may be put into operation in the child by certain kinds of situational information, but that its manner of functioning is not affected. Chomsky supports his argument by reference to the Miller and Norman (1964) experiment in which subjects were required to discover rules for generating admissible strings of letters that formed an artificial language. The subject was required to type on a computer what he believed to be a grammatical string of letters; if the string was grammatical the typewriter performed a specified set of indicated procedures corresponding to the letters that were used (this set of procedures being the semantic reference of the language); if the string was ungrammatical, it typed the word WRONG. Subjects learning with semantic reference learned the language at about the same speed and in the same way as did subjects without semantic reference. However, semantic reference was incorporated into the Miller and Norman experiment only after the subject had already formed a perfectly syntactically correct sentence; only then did the teaching machine treat the sentence as an instruction to do something (and only then did the subject have any chance to discover the referent of each word). So naturally the hypotheses of subjects were unaffected by semantic referents since the semantic referents were mostly absent while subjects were forming their hypotheses.

The present experiment allows subjects to make correspondence between words and referents right from the start (as in real language learning) and the hypotheses are affected by the referents.

It appears that when semantic referents are present the learning strategy consists of (a) learning to associate each word with its referent; (b) learning the specific rules of the reference field (the ways in which these referents can be organized); and (c) learning to map words referring to relevant aspects of the visual field onto the sentence positions. Support for this hypothesis comes from: (1) the fact that subjects in the syntax correlation condition learned more about the language syntax than the subjects in the words-only condition, even though their instructions stressed vocabulary learning; (2) the patterns in which the rules were learned in the different conditions (there was a significant interaction between rules and conditions); and (3) the pattern in which the different languages were learned in the different conditions (there was a significant interaction between conditions and language complexity when the two more complex languages were compared).
Strategy in the Words-only Condition

In the words-only condition the semantic learning strategy could not be employed because there were no referents to associate with each word nor a field to perceptually organize, so words had to be mapped directly onto sentence positions without mediation through a semantic system. In this condition it is most likely that the learning strategy employed consisted of: (1) using the positions that words occupied in a sentence to learn their syntactic class; and (2) learning that a correct sentence consists of one word from one class, one word from a second class, and so on. This strategy is effective only in very elementary language systems; as syntax complexity increases the sentence position of a word class becomes increasingly variable. In the present experiment it was found that the more difficult the language, the lower the overall performance in syntax learning. The pattern of rule learning also supported the position-learning hypothesis. The only position constancy in all three languages was the fact that the initial word in every sentence consisted of one of four words (belonging to the A class). Thus subjects performed at considerably above chance levels in all three languages only on rule 1 (an A word must appear) and rule 2 (only one A word can appear). Performance was slightly above chance on rule 4 (only one B word can appear) in all three systems and on rule 5 (D word position rule) in systems 1 and 2. There was a marked decrement in performance on rule 5 in system 3 when the D word became able to considerably vary its position. Performance on the other rules approximately averaged chance level.

Strategy in the Arbitrary Figures Condition

When information in the reference field provides no correlation between the semantic referents and the syntax relations, as in the arbitrary figures condition, the subject can employ either the words-only position learning strategy, or the semantic learning strategy. If he employs the semantic learning strategy he will find great difficulty in attempting to organize the field. Conversations with several of the subjects after they had been tested in the arbitrary figures condition indicated that the semantic learning strategy was generally employed (although some subjects may have used a position learning strategy). Results showed that subjects in the arbitrary figures condition performed roughly identically to subjects in the words-only condition with respect to the patterns of rule learning and language learning on systems 1 and 2. In system 3, however, only rule 2 (only one A word can appear, or in respect to the visual field, only one colored rectangle can appear) showed a performance level above chance. It can be assumed that in the most complex language, the task of organizing the reference field was so difficult that practically no learning could take place.

Strategy in the Class-correspondence Condition

If information provided in the reference field corresponds to some, but not all, of the syntax rules, as in the class correspondence condition, and the semantic learning strategy is employed, the rules not mirrored in the constraints of the reference system (rules 6 and 7 in system 2 and rules 5, 6, and 7 in system 3) should not be learned. Thus subjects in the class correspondence condition showed performance levels above chance on all rules in system 1 and on all rules in system 2 except rule 6 (B₁ rule) and rule 7 (B₂ rule). In system 3 performance was above chance on all rules except rule 5 (D word must follow an A word or C word) and rule 7 (B₂ rule). Performance on rule 6 (B₁ rule) was above chance, contrary to expectation, but performance on rule 7 (B₂ rule) was considerably below chance, suggesting that subjects were simply utilizing a strategy of selecting the longer sentence as correct. If the scores on rules 6 and 7 are added together, the total for the two rules is below chance level.
Strategy in the Syntax Correlation Condition

Only when the information provided in the reference field corresponds to the language syntax, as in the syntax correlation condition, will the employment of the semantic learning strategy result in the acquisition of all grammatical rules. In the syntax correlation condition performance on all rules was at above chance levels and there was practically no difference among rules or among languages.

Conclusion

Thus evidence indicates that when semantic referents are present in an artificial language a different learning strategy is employed from that used when semantic referents are not present. In the semantically empty condition, words are mapped directly onto sentence positions (position learning strategy); under the semantic reference condition, words are associated to relevant aspects of the reference field and the field is structured into a meaningful organization, and then this semantic information is used in the mapping process. (It cannot be stated, however, that subjects in the semantic reference conditions did not use position-learning strategy at all. It seems likely that the learning of rule 5, that a D word follows rather than precedes A and C words, in the syntax correlation condition at least, must entail some form of position learning strategy that is either used in addition to or overlaid on the semantic learning strategy.)

The evidence also indicates that it is only when the elements in the reference field mirror the syntactic constraints of the language that complex grammatical relations are easily acquired. As the only major difference among the three semantic reference conditions was the way in which the referent field was organized, it appears that the information obtained from this perceptual organization plays an important part in the learning of grammatical relations.

What does all this say about the acquisition of natural languages? Only the following:

We must remember that the reference field is highly organized and that much, though not all, of the syntax in sentences reflects aspects of this organization. We have shown experimentally that people can make use of this reference-field structure in the learning of syntax. Thus it seems plausible to expect it to play a role in natural syntax acquisition. Such an expectation will lead us to try to identify the particular cognitive and perceptual organizations of the reference field that facilitate the acquisition of particular syntactic forms.

References


Braine, M. D. S. Learning the positions of words relative to a marker element. Journal of Experimental Psychology, 1966, 72, 532-540.


(Received April 17, 1972)