JOURNAL OF MEMORY AND LANGUAGE 25, 492–505 (1986)

Semantic Reference and Phrasal Grouping in the Acquisition of a Miniature Phrase Structure Language

RICHARD P. MEIER AND GORDON H. BOWER

Stanford University

In this study we examine the roles of semantic reference and of grammatical morphology in the learning of an artificial syntax. Subjects assigned to one of three training conditions viewed sentences from a miniature phrase structure language. In the *reference field* condition, subjects saw sentences which each referred to an array of geometric figures. In the *morphology* condition no reference field was present, but inflectional suffixes marked each sentence's constituent structure. *Control* condition subjects studied sentences lacking semantic reference and inflectional morphology. Unlike control subjects, subjects in both the reference field and morphology conditions learned the miniature syntax, as evidenced by successful discrimination of novel grammatical versus ungrammatical sentences. Therefore, when surface features mark constituents, adult learning of complex syntactic regularities proceeds even in the absence of semantic reference.

One method for experimentally examining the language acquisition process involves the exposure of adult subjects to sentences from miniature artificial languages (Braine, 1963; Esper, 1925; Miller, 1967; Reber, 1967). Use of such artificial languages in the laboratory enables the rigorous—and ethical—manipulation of the input to the language learner. Although the analogy between these experiments and first language learning is necessarily imperfect (Bever, Fodor, & Weksel, 1965), we suggest that these experiments can address hypotheses about the language acquisition process which are difficult, if not impossible, to test in the natural language-learning environment. In the experiment reported here, we examine the role of semantic reference in the learning of the syntax of a miniature phrase structure language. In particular, we will argue that the syntax of a referenceless language can be learned if the sentences of that language contain markers of constituent structure.

Part of acquiring a language is learning what the words and sentences of that language mean. Part of knowing the meaning of a sentence is knowing whether that sentence refers to objects, actions, or properties in the world. But language learning demands more than learning word-to-object and sentence-to-scene reference: meaningful sentences in human languages are not just unorganized lists of topically related words. Rather, acquisition of a language entails the learning of a syntax, a system of grammatical rules by which words are organized into sentences.

What are some of these syntactic rules? As one example, the speaker must learn rules which govern the distribution of coherent classes of words, called *form classes* (e.g., nouns, verbs, and adjectives). The members of form classes exhibit similar syntactic behavior. For example, English verbs, unlike nouns, may be modified by an adverb, inflected for person, number, and tense, and preceded by a modal such as *will* or *can*. As a second example, the sentences of natural languages exhibit not simply the left-to-right organization of a

Preparation of this article was supported in part by NIMH Training Grant MH 15157-08 and by NIMH Grant MH 13950-16 to Gordon H. Bower. Requests for reprints should be addressed to Richard P. Meier, Department of Linguistics, 501 Calhoun Hall, University of Texas, Austin, TX 78712.

0749-596X/86 $3.00

word string, but also the hierarchical organization of words into *constituents*, or *phrases*, such as noun phrases and verb phrases. These constituents can themselves be nested within other constituents. Third, an element in a constituent may require another dependent element; for example, an adjective or determiner requires a noun. Such *dependencies* typically, but not always, obtain within syntactic constituents. The artificial language we will be studying has these three syntactic properties, as well as others.

A crucial issue in recent work on first language acquisition (see Anderson, 1983; Maratsos & Chalkley, 1980; Pinker, 1984) and in experimental research into the learning of artificial languages has been the extent to which semantics is implicated in the learning of syntax. In artificial language learning experiments, the learner is presented many sentences from a miniature language, each sentence typically being paired with a reference field. The reference field is a set of pictures to which the sentence refers. Researchers working within this paradigm have asked whether successful syntax learning, as demonstrated by subjects' abilities to judge the grammaticality of novel sentences, requires strong parallelisms between the organization of the syntax and the organization of the reference field.

Moeser and Bregman (1972, 1973) and Anderson (1975) have argued that such parallelisms between the syntax and the reference field are indeed necessary for the learning of complex syntactic rules. Moeser and Bregman (1972) assigned subjects to syntax-learning conditions which differed only in reference field organization. In the condition which enabled successful syntax learning, dependencies in the artificial syntax were mirrored by properties of the geometric figures which formed the reference field. For example, the syntactic relationship between a noun-like word and a modifier word encoded the relationship between a geometric figure and a border variation upon that figure. Only the subjects who received such input successfully learned syntactic dependencies. In contrast, subjects who viewed reference fields which lacked these simple reference–syntax parallelisms did not learn the syntax, even after viewing 3200 sentence–reference field pairs (Moeser & Bregman, 1973). Using a different phrase structure language in which the vocabulary items were English words, Anderson replicated Moeser and Bregman's (1972) striking results. He argued that the crucial parallelism which must obtain between syntax and reference field is that the constituent structure of the miniature language must preserve the pattern of conceptual (node) linkages in a network representation (as in Anderson & Bower, 1973) of the reference field.

However, an early study of artificial language learning in the absence of semantic reference (Saporta, Blumenthal, & Reiff, 1963) demonstrated greatly enhanced recall of letter strings which were spaced in accordance with a phrase structure in comparison to recall of the same strings with either equal spacing between the letters or random grouping of the letters. Recently, Morgan and Newport (1981) and Morgan, Meier, and Newport (1986) have demonstrated that the syntax of a variant of the Moeser and Bregman (1972) language can be learned even if the input contains much less highly structured reference fields than those employed by Moeser and Bregman or by Anderson (1975). Morgan and Newport suggested that the crucial property of the reference fields in Moeser and Bregman's successful syntax-learning condition was that they cued the constituent structure of the accompanying word strings. Both Morgan and Newport and Morgan et al. employed reference fields in which each figure was uniquely paired with a single word of the accompanying string. In Morgan and Newport, spatial grouping of those figures in accordance with the constituent structure of the string enabled successful syntax learning, including the learning of dependencies (even though the reference fields did not directly cue those dependencies). Morgan et al. noted that this spatial manipulation of the reference field has no obvious counterpart in natural languages. Rather, natural languages as spoken (or signed) contain *grouping cues*, such as prosodic intonation, function words, and concord morphology, which cluster words into syntactic constituents.[1]

Although not an important characteristic of English, many languages have concord inflections which mark certain words as agreeing in such grammatical categories as case, person, number, and gender. These inflections are bound morphemes which are prefixed or suffixed to the agreeing words. In Spanish, nouns and their modifiers agree in gender and number. In (1), the Spanish concord markers are in boldface:

(1) L**os** gring**os** ric**os** vieron l**as** ruin**as** incaic**as**.

"The rich Gringos saw the Incan ruins." Although not all Spanish sentences are so transparent as (1), concord identifies the members of the two noun phrases *los gringos ricos* and *las ruinas incaicas* by affixing identical markers to the words which fall in a particular noun phrase.[2]

In lieu of either explicit cueing of dependencies in the reference fields, as in Moeser and Bregman (1972), or spatial clustering in the reference fields, as in Morgan and Newport (1981), Morgan et al. (1986) manipulated the presence or absence of three types of grouping cues: prosodic intonation, function words, and concord morphology. In three experiments examining whether these string-internal cues could subserve syntax learning, subjects learned the miniature language only if the stimuli contained grouping cues which demarcated syntactic constituents.

The Morgan et al. (1986) experiments raise a question as to the exact role played by reference fields in the acquisition of a miniature syntax. In those experiments, the reference fields indicated the meaning and form class of lexical items, redundantly displayed the adjacencies found within the accompanying sentence, and perhaps served as a useful mnemonic for subjects. The reference fields did not cue either syntactic dependencies or constituent structure. Subjects in all conditions viewed identical reference fields, thereby demonstrating that reference fields having the above properties are not sufficient for successful syntax learning. In the experiment reported here, we will examine the further issue of whether reference plays any necessary role in the acquisition of an artificial syntax, leaving aside the question of semantic/referential mediation in the acquisition of form classes (see Billman, 1985, and Mori & Moeser, 1983, on the learning of form classes in artificial languages). In doing this, we will further examine the efficacy of one type of grouping cue, concord inflections, in the acquisition of an artificial syntax.

One prior study examined the adequacy, in the absence of any reference field, of another type of grouping cue in the acquisition of constituent structure. Green (1979) demonstrated that subjects could learn a completely meaningless phrase structure language if the input contained function words at the beginnings of constituents. The design of his study differs from ours here in several respects, specifically, in the availability to subjects of a corpus of up to 11 sentences at any one time during

[1] We are aware of the class of *nonconfigurational languages* in which the surface structure of sentences is relatively nonhierarchical and in which word order is virtually free (Hale, 1983). However, even in such languages, certain constituent types, for example embedded sentences, are preserved intact. Interestingly, nonconfigurational languages typically display rich inflectional morphology which marks words between which dependencies exist.

[2] A second type of concord governs the form of the verb in example (1): Spanish verbs agree in person and number with the sentential subject. Thus, *vieron* is the third person plural preterite form of the verb *ver*, meaning "to see."

training and of feedback after training trials in which subjects attempted to produce sentences from the miniature language.[3] Also, Green had no condition in which subjects viewed sentences paired with reference fields and thus he did not contrast syntax learning resulting from a reference field condition versus his function word condition.

In the current experiment we adopt Anderson's (1975) basic methodology. By its use of English vocabulary, his method allows us to distinguish two aspects of semantics: reference and word meaning. Using a variant of his artificial language, we examine syntax learning in three training conditions which differ only in the properties of the linguistic input to subjects: (1) a *reference field* condition in which sentences are paired with a highly structured reference field, (2) a *morphology* condition, presented without reference fields, in which concord inflections are incorporated into the language, thereby providing substantial information within the sentences themselves as to their constituent structure, and (3) a *control* condition lacking both concord morphology and reference fields. The results will, we believe, permit a considerable refinement of our understanding of the roles of semantics and of surface markers of constituency in the acquisition of miniature languages in the laboratory.

METHOD

Subjects

Thirty undergraduates from the Introductory Psychology class at Stanford University received class credit for participating in this experiment. All were native speakers of English. All had also had at least some experience with a second language. Ten subjects were assigned at random to each of the three input conditions described below.

Grammar

We have adopted a variant of the artificial grammar described in Anderson (1975). Table 1 lists the phrase structure rules of this grammar. Although this language is obviously very simple compared to any natural language, it does have certain interesting properties. Note that this language is recursive: that is, a sentence-like structure, a *clause*, can be embedded within the larger sentence. Inasmuch as these clauses are embedded within noun phrases, they are analogous to relative clauses in English and other languages. In this artificial language, clauses are introduced by a function word, *te*. *Te* must appear immediately prior to the noun located within a clause; this, therefore, is a syntactic dependency of the miniature language. *Te* also introduces the semantically diffuse adjective class consisting of the words *red* and *broken*. The word order of this language is subject–object–verb (SOV) and thus contrasts with the SVO word order of English. However, SOV order is very common among natural languages; Japanese is one example. The

TABLE 1
THE MINIATURE PHRASE STRUCTURE GRAMMAR

S → NP PRED

$$\text{NP} \rightarrow \text{N} \begin{Bmatrix} \text{(SIZE) (PATTERN) (CLAUSE)} \\ \textit{te}\ \text{ADJ} \end{Bmatrix}$$

CLAUSE → *te* PRED

PRED → NP REL

N → {*square, circle, diamond, triangle*}

SIZE → {*large, small*}

PATTERN → {*striped, dotted*}

ADJ → {*red, broken*}

REL → {*above, below, right-of, left-of*}

Note. The following notational conventions are employed here: nonterminal categories are represented entirely in capital letters: S = sentence, NP = noun phrase, N = noun, PRED = predicate, and ADJ = adjective. Lexical items are typed in lower case. The arrow should be read as "is rewritten as." Categories enclosed within parentheses are optional. The braces indicate a disjunction.

[3] Based upon results reported by Brown and Hanlon (1970), most theories of first language acquisition have assumed that children do not receive feedback as to the grammaticality of their utterances. In formal learnability theory, Gold (1967) has shown that the availability of such feedback profoundly eases the task of language induction.

miniature language also resembles Spanish and French in that modifiers (i.e., size and pattern words) follow the noun.

While remaining essentially faithful to Anderson's (1975) grammar, we have adopted this variant of it so as to reduce the maximum length of the stimuli and the number of different sentence types which must be presented among the training stimuli. For example, Anderson's grammar has 12 syntactically distinct expansions of the noun phrase, whereas ours has only 9. Our grammar has three basic sentence types with zero or one embedding: a sentence with no embedded clause, a sentence with a clause embedded in the subject noun phrase, and a sentence with a clause embedded in the object noun phrase.

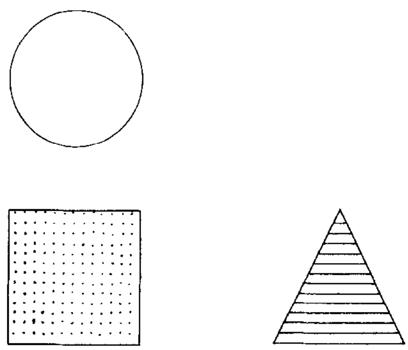
With the exception of the function word *te*, the vocabulary of this language, and of Anderson's (1975), is English. Using English vocabulary eliminates the need for subjects to learn a set of word-referent pairs. The use of English vocabulary will allow us to distinguish two types of semantic information which are available to our subjects: (1) the semantics of the English words which appeared in all three training conditions of this experiment, and (2) the information available from the reference fields which was present in the stimuli shown subjects in one training condition. Each English word conveys several types of semantic information: its definition, its form class (which can invariably be defined semantically in this artificial language; e.g., the four nouns are the names of geometric figures), and the number of arguments it requires (zero for nouns, one for noun modifiers—i.e., size words, pattern words, and adjectives—and two for the relational terms). Interestingly, the semantic information which is present in the stimuli by virtue of our use of English vocabulary is essentially the same as that which many recent theories of language acquisition have supposed that children possess prior to the onset of syntax learning (see Berwick & Weinberg, 1984; Pinker, 1984; Wexler & Culicover, 1980).

Training Conditions

Each subject was assigned to one of three training conditions. In designing these conditions, we manipulated two characteristics of the stimuli: the presence or absence of a reference field and the presence or absence of inflectional morphology in the word string.

Reference field condition. In the reference field condition, subjects viewed 48 sentences which were each paired with a semantically appropriate reference field. Each reference field contained either two or three geometric figures, depending on the number mentioned in the paired sentence. The figures were arranged in a spatial array appropriate to the relational terms included in the sentence. The function word *te* was not associated with any feature in the reference fields. An example of a sentence-picture pair from this condition appears in Fig. 1.

Sentences were each typed in 10-point Courier along the long dimension of an 8½ × 11-in. card and were centered approximately 1½ in. above the edge. The appropriate reference field was centered approximately 1⅛ in. above the sentence. The figures were regularly situated within the reference field. The set of possible figures included large and small versions of the four basic geometric forms. If a noun



Square large dotted te triangle
striped left-of circle large below.

FIG. 1. A stimulus item from the reference field condition.

phrase did not specify the size of its referent, a large or small figure was randomly assigned to it.

Morphology condition. In the morphology condition, subjects viewed the same set of 48 sentences that was shown to the subjects in the reference field condition, except that inflectional morphemes indicating concord were suffixed to the words of the sentences. Furthermore, the sentences were not accompanied by a reference field. Each sentence therefore appeared alone on a card. Sentences were centered at the same locations on those cards as in the reference field condition.

Four inflectional suffixes were employed in a regular manner in this condition. The suffixes were typed in uppercase letters (in contrast to the lowercase letters of the vocabulary items) and were separated from the vocabulary items to which they were suffixed by a hyphen. The inflectional morphemes and their distribution are listed below:

-O. Suffixed to the subject noun of the main clause and its modifiers (see examples 2a–c below). *-O* also marks agreement on a main clause relation (examples 2a–c) or on an embedded relation (example 2b) whose subject is the main clause subject.

-A. Suffixed to the object noun of the main clause and its modifiers (examples 2a–c). *-A* also marks subject agreement on an embedded relation whose subject is the main clause object (example 2c).

-AO. Suffixed to the object noun, and its modifiers, of a clause which modifies the main clause subject (example 2b).

-AA. Suffixed to the object noun, and its modifiers, of a clause which modifies the main clause object (example 2c).

The inflectional morphemes were never suffixed to the function word *te*. Three examples of stimulus sentences from the morphology condition appear below.

(2a) *Diamond-O triangle-A small-A dotted-A left-of-O.*

(b) *Square-O large-O dotted-O te triangle-AO striped-AO left-of-O circle-A large-A below-O.*

(c) *Diamond-O small-O dotted-O circle-A small-A dotted-A te diamond-AA large-AA striped-AA left-of-A below-O.*

Note that the suffixes *-AO* and *-AA* can be further analyzed as sequences of the suffixes *-A* and *-O*. By this analysis, *-A* is understood as being suffixed to any object noun, whether in the main clause or in an embedded clause. If the object noun phrase is in a clause which modifies the object of the main clause, the words are further suffixed by *-A*, yielding the sequence *-AA*. If, however, the embedded object noun phrase is in a clause modifying the subject, the members of the object noun phrase are further suffixed by *-O*, thereby forming *-AO*.

Importantly, the inflectional morphemes signal the constituent structure of the stimulus sentences. Sequences of adjacent words which share the same suffix belong to the same syntactic constituent. Thus, the inflectional morphemes group the words of each string into phrases.[4]

Control condition. Subjects assigned to the control condition studied stimuli which contained neither reference fields nor inflectional morphology. The set of stimulus sentences was identical to that presented to subjects in the reference field condition.

Stimuli

Subjects viewed a total of 48 sentences from the miniature language. This set comprised 24 two-noun and 24 three-noun sentences. Three-noun sentences necessarily contained an embedded clause. Stimulus sentences ranged in length from 4 to 12 words. Unlike Anderson (1975), we did not include any four-noun sentences in the stimuli because such sentences could be extremely long. By excluding them, we did not omit any syntactic patterns (with the

[4] This morphological system is idealized vis-à-vis one which is likely to be found in any single natural language. However, many features of the artificial system are amply attested in natural languages. For instance, in many Bantu languages, identical concord inflections are prefixed to a noun, its modifiers, and to the verb for which the noun is the subject; see examples cited in Lyons (1968, pp. 284–285). Our intent is not to exactly replicate any one natural system, but rather to examine in isolation the effect of one cue for constituency.

exception of doubly embedded clauses, that is, a clause embedded within a clause). The total set of possible stimuli includes over two million sentences.

All nouns were presented equally often in the subject, main clause object, and embedded positions. Likewise, the four relation words occurred equally often in the main clause and embedded clause positions. The two-noun and three-noun sentences were constructed separately. In both the two-noun and three-noun sentence sets, each of the four possible nouns appeared in the subject position at least once, and no more than twice, with every noun in the main clause object position and with every main clause relation word. The same was true with respect to the embedded nouns and relations of the three-noun stimuli.

In the 24 three-noun stimuli, embedded clauses modifying the main clause subject occurred equally often as those modifying the main clause object both within the entire set and within each six-sentence subset sharing a specific subject noun.

The syntax generated four noun modifier patterns: (i) no modifier, (ii) one size or pattern modifier, (iii) a size + pattern sequence, and (iv) *te* + adjective. Within the two-noun stimuli, each pattern followed each subject noun at least once, and no more than twice. Within the six two-noun stimuli sharing a specific subject noun, each pattern again appeared once, and no more than twice, in the main clause object noun phrase. The three-noun stimuli were constructed in identical fashion, with the modifiers of the embedded noun phrase being constructed according to the same constraints as the modifiers of the main clause direct object. No stimulus sentence consisted entirely of nouns with null modifiers. Individual modifier words were randomly assigned to appropriate modifier patterns.

Finally, all stimuli met two semantic constraints. Within any sentence, all noun phrases were referentially distinct. That is, no sentences like *Square large square large left-of*, which means "The large square is to the left of the large square," appeared in the stimulus set. Also, some three-noun sentences were ruled out because they were vague descriptions, as in (3):

(3) *Square large te circle below diamond below*. "The large square which is below the circle is below the diamond."

All subjects viewed the same set of 48 sentences, except that in the morphology condition the words of these sentences were augmented by the suffixes described previously. By virtue of the affixation of the inflectional morphemes, the morphology subjects viewed input sentences which were substantially longer than those presented in the other two conditions.

The 48 sentences appropriate to a particular condition were presented in eight study blocks of six sentences each. Three two-noun and three three-noun stimuli were randomly assigned to each study block. In turn, order of presentation within the study blocks was randomized.

Procedure

Instructions. Subjects were instructed that they were participating in a language learning experiment. They were told that they would be shown a series of cards, each of which would contain a single sentence from a language. They were also informed that because we were not interested in vocabulary learning, we had, by and large, substituted English vocabulary items for the words of the language. Subjects were asked to learn the rules by which the words were organized into sentences. Subjects in the reference field condition were also informed that above each sentence would be a set of pictures to which the sentence referred.

Presentation and testing. The experiment was administered to subjects individually. Sentences were presented on cards. Each card was shown for 30 s. After each block of six study sentences, subjects were given a test booklet containing six two-alternative forced-choice questions. Test questions were never accompanied by reference fields. In the morphology condition, test sentences contained the appropriate inflectional morphology.

Each test question appeared on a separate page of the test booklet. Subjects were instructed to answer each question in sequence without looking back or ahead to any other question. The questions required subjects to judge the grammaticality of novel sentences which had not appeared in the study set or on any prior test. Each question contained one alternative which was grammatical in the language and one which included a single syntactic violation of one or another rule. Subjects were asked to identify the correct sentence. Unlike Anderson (1975), subjects had an unlimited amount of time in which to answer each question. Subjects received no corrective feedback from the experimenter.

Six rules of the syntax of the miniature phrase structure language were tested. These rules are listed in Table 2. Each rule was tested once on each of the eight tests. Anderson's (1975) test questions were of two sorts: in questions assessing what he termed "minimal syntactic contrasts," the two alternatives were identical except, for instance, for the reordering of the size and pattern modifiers in the incorrect alternative, as in the example listed for Rule 1 in Table 2. In other questions examining "gross semantic defects" (e.g., a missing argument from the main clause), a correct sentence was paired with an unrelated, incorrect sentence. We followed Anderson in the design of our tests. The alternatives examining Rules 1–3 differ only minimally. In contrast, the alternatives in test items examining Rules 4–6 are unrelated lexically, but do share, to the extent possible, the same basic syntactic structure. This decision allowed us to equate closely the lengths of the correct and incorrect alternatives in test items examining Rules 4–6.

Within the set of eight test items (distributed across the eight test trials) examining each rule, we counterbalanced for the oc-

TABLE 2

RULES OF THE MINIATURE PHRASE STRUCTURE GRAMMAR WITH EXAMPLES OF APPROPRIATE TEST ITEMS

Rule 1: Size modifiers precede pattern modifiers.
Square large dotted diamond large right-of.
**Square dotted large diamond large right-of.*
Rule 2: *Te* must precede an embedded object.
**Triangle te red diamond triangle te broken above right-of.*
Triangle te red diamond te triangle te broken above right-of.
Rule 3: *Te* cannot occur before the main clause object.
Diamond square te broken below.
**Diamond te square te broken below.*
Rule 4: A relation must occur in the main and embedded clauses. In the morphology condition, the single relation present in incorrect test items had the suffix *-O*, indicating agreement with the main clause subject.
Square large circle small dotted te triangle above left-of.
**Diamond small striped circle large dotted te square below.*
Rule 5: A sentence cannot have three main clause arguments.
**Triangle te red circle large dotted diamond small below.*
Square te circle striped below diamond te broken left-of.
Rule 6: A sentence must have a main clause object.
**Circle large te triangle small below above.*
Square large striped circle te red above.

* Incorrect test items.

currence of specific nouns and relations in specific syntactic positions and for the syntactic role (i.e., subject or object) of the noun phrase modified by a clause. Within these constraints and the further constraint that the noun phrases within a sentence had to be referentially distinct, words were randomly assigned to sentences. To the extent possible, this counterbalancing was also carried out within the two subsets consisting of four test questions each which together formed the eight questions testing any given rule. The questions from the two subsets were randomly assigned to Tests 1–4 and 5–8, respectively. Finally, the order of the questions within each of the eight tests was randomized.

Debriefing. At the conclusion of the experiment, each subject was informally questioned about how he or she had gone about learning the language. Subjects were also asked to state whatever rules and regularities they believed they had observed.

Results

Figure 2 is a plot of the mean percentage correct for each of the subject groups over the eight trials. The data as pooled over the two halves of the experiment and from Trial 8 alone appear in Table 3. Inspection of Figure 2 and Table 3 suggests that subjects in both the morphology and the reference field conditions showed good learning of

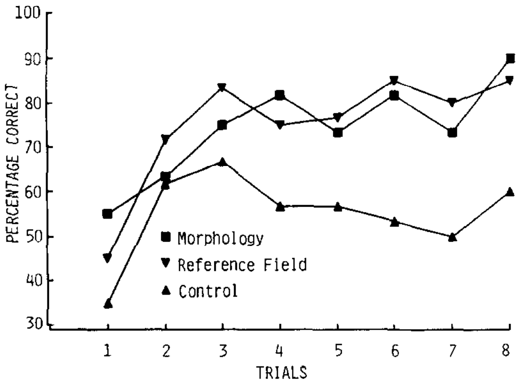


FIG. 2. Judgments of the grammaticality of novel sentences: Mean percentage correct by trials for the three training conditions (chance = 50% correct).

TABLE 3
MEAN PERCENTAGE CORRECT FOR THE THREE TRAINING CONDITIONS

Condition	Tests 1–4[a]	5–8[a]	8[b]
Control			
Mean	55.00	55.00	60.00
SD	7.79	16.63	25.17
Reference field			
Mean	69.17	81.67	85.00
SD	15.25	13.63	25.33
Morphology			
Mean	68.75	79.58	90.00
SD	7.17	15.25	18.00

Note. Chance = 50% correct.
[a] 24 responses per subject.
[b] 6 responses per subject.

the miniature syntax. In contrast, the control subjects hovered marginally above chance and showed no improvement after the third test trial.

A Training Condition by Trials analysis of variance confirmed a significant effect for Training Condition, $F(2,27) = 10.12$, $p < .001$, and for Trials, $F(7,189) = 10.13$, $p < .0001$. The Training Condition versus Trials interaction was not significant, $F(14,189) = 1.18$, NS. More revealing, however, are the results of three planned comparisons which examined the efficacy of the two input manipulations. The Control versus Reference Field comparison was significant, $F(1,27) = 15.96$, $p < .001$; the interaction term in this comparison was not, $F(7,189) = 1.10$, NS. The significantly enhanced syntax learning by the subjects who viewed highly structured reference fields replicates Anderson's (1975) key finding. The Control versus Morphology comparison was also significant, $F(1,27) = 14.36$, $p < .001$; again the interaction over trials was not significant $F(7,189) = 1.58$, NS. This comparison supports the hypothesis that even in the absence of any reference field, concord morphology can provide the basis for learning a miniature syntax. Finally, the Reference Field versus

Morphology comparison was not significant, $F(1,27) < 1$, NS, nor was the interaction over trials, $F(7,189) < 1$, NS. Thus, these data indicate that concord morphology was just as useful as reference field organization in providing a basis for syntax learning.

Recall that the test items examined subjects' knowledge of specific dependency rules, that is, of rules which govern the distribution of one form class given the presence of another. Table 4 summarizes for each rule the test results from the first versus second halves of the experiment. Prior work (Morgan & Newport, 1981; Morgan et al., 1986) indicates that subjects generally show significant progress in learning dependency rules only if they view sentences in which constituent structure is signaled, either by the organization of an accompanying reference field or by grouping cues in the sentence string itself. An analysis of the learning trends in subjects' performances across Trials 1–8 confirmed this expectation. Subjects in both the reference field and morphology conditions showed significant upward linear trends over trials: for the reference field condition, $F(1,9) = 29.85$, $p < .001$; for the morphology condition, $F(1,9) = 10.16$, $p < .025$. The control group did not show a significant linear trend, $F(1,9) = 1.03$, NS. An analysis of the Linear by Linear interactions revealed a significant difference in the slopes of the linear trends for the reference field and control subjects: $F(1,18) = 6.82$, $p < .025$. The difference between the slopes of the linear components in the morphology and control condition data approached, but did not attain, significance, $F(1,18) = 3.54$, $p < .08$. The reference field and morphology conditions did not differ significantly in the slopes of their linear trends, $F(1,18) < 1$, NS.

TABLE 4

MEAN PERCENTAGE CORRECT FOR THE SIX SYNTACTIC RULES

		Syntactic rule					
Condition	Tests	1	2	3	4	5	6
Control	1–4	62.5	52.5	65.0	42.5	47.5	60.0
	5–8	77.5	45.0	50.0	40.0	45.0	72.5
Reference	1–4	80.0	57.5	70.0	62.5	65.0	80.0
	5–8	87.5	62.5	75.0	92.5	82.5	90.0
Morphology	1–4	77.5	62.5	82.5	60.0	65.0	65.0
	5–8	100.0	75.0	82.5	67.5	80.0	72.5

Note. Each rule was tested once on each test. Chance = 50% correct.

Even by the last half of the experiment (Trials 5–8), the control subjects did not perform significantly above chance overall, $t(9) = .95$, NS, one-tailed. As is shown in Table 4, their performance exceeded chance on only two of the six syntactic rules, specifically on Rules 1 and 6. Although Rule 1 states a generalization about the distribution of two form classes with respect to each other, it is not a true dependency rule. Rather, it is a rule of sequential ordering. Rule 1 states that a size word must precede a pattern word, if both modify a single noun. However, the presence of either modifier does not predict the presence of the other; each is optional. Such a rule of sequential ordering would seem to be learnable given even the limited linguistic input presented to our control subjects. This finding is consistent with results reported by Morgan and Newport (1981) for subjects who, like our control subjects, viewed sentences which lacked grouping cues.

Rule 6 states that a sentence must have a main clause object. This rule was the only one on which the control subjects matched the performance of one of the other training groups, in this case the morphology subjects. Examining the Rule 6 test items more closely, we note that the incorrect alternative invariably had the following structure: NP [*te* NP REL] REL. Subjects could rule out that alternative, and hence score well on Rule 6, if they had formed one simple, but partial, generalization: that is, in a sentence having only two nouns, a noun must not be immediately preceded by *te*. This partial generalization is indeed a dependency rule, albeit one between a single word, *te*, and the number of nouns in a sentence. Interestingly, *te* is a function word which marks the beginning of embedded clauses and, as such, was the only grouping cue which was available to subjects in the control condition.[5]

Debriefing. At the conclusion of the experiment, the subjects were asked to describe what, if anything, they had learned of the language. The control subjects could typically state the correct ordering of the size and pattern words, a finding which is consistent with their good performance on Rule 1. This information suffices to pull their average performance on the test trials slightly above chance. Also, these subjects often knew that the adjectives *red* and *broken* were necessarily preceded by *te*. However, their explicit knowledge of other aspects of sentence structure was minimal. Generally, their statements with respect to the distribution of the relation words and the environments in which *te* precedes a noun were limited to partial and/or incorrect generalizations about the distribution of individual words, not of classes of words. For example, one subject reported (quite erroneously) that if the sequences *diamond diamond* or *circle circle* occurred in a sentence, *te* could not appear between those words, but that *te* always intervened between *diamond* and *circle*.

Although 9 of the 10 control subjects reported such item-specific generalizations, they were almost entirely absent from the comments of the morphology and reference field subjects, even when remarking on aspects of the language which they claimed not to understand. Subjects in these conditions could often state syntactic rules of the miniature language correctly and with considerable sophistication, including rules pertaining to the structure of embedded clauses. Three morphology subjects reported that they imagined the spatial arrays which the sentences described. The remaining subjects reported that they did not adopt such a strategy, or did so only sporadically. The three subjects who did imagine arrays of figures did not show superior learning (mean percentage correct = 70.8) when compared with the remaining subjects (mean percentage correct = 75.6).

DISCUSSION

Our results demonstrate that subjects can successfully learn a miniature phrase structure grammar without semantic reference, provided that the sentence strings themselves contain cues for constituent structure. Subjects in both the reference field and concord morphology conditions showed significantly enhanced learning as compared to that of the control subjects. The success of the morphology subjects is especially striking in light of the fact that by virtue of the presence of the concord morphemes, they viewed sentences which were substantially longer than those seen by the two other subject groups. These results, along with those reported by Green (1979) and Morgan et al. (1986), affirm that grouping cues can subserve the learning of a miniature syntax.

Whether other aspects of semantics, particularly the semantics of individual lexical items, are necessary for successful syntax learning is unresolved. Note, however, that such semantic information clearly is not sufficient for successful syntax learning. Information about the meaning, form class, and required number of arguments of the individual lexical items was equally available to subjects in all three training conditions, yet the control subjects showed little or no evidence of learning the complex rules of the miniature syntax.

Grouping cues, such as concord morphology or function words, may aid syntax

[5] T. R. G. Green has kindly pointed out to us that the above chance performance of the control subjects on Rules 1 and 6 might be explained by similarities between the syntax of English and the syntax of the miniature language. In particular, the adjective order stipulated by Rule 1 is also the preferred adjective order in English. Note, that even with this positive transfer from English, the performance of the reference field and morphology subjects was superior to that of the controls on these two rules.

learning by facilitating the parsing of input sentences (Clark & Clark, 1977; Kimball, 1973). For example, the short, phonologically distinctive English function words, such as articles, auxiliaries, and prepositions, are located at the beginnings of syntactic constituents. In our control condition, no single stimulus sentence contained enough information to determine a correct parsing. Thus, given only the input *Square te red circle large right-of* a subject could not know whether *red* modifies *square* or *circle*. In contrast, subjects in the morphology and reference field conditions viewed input which permitted a correct parsing of sentences given just the information available on any single card. Thus, in the just-mentioned sentence as it was presented to the morphology subjects, the distribution of the concord morphemes suggests the following, correct constituent structure: [(*Square-O te red-O*) (*circle-A large-A*) (*right-of-O*)]. In both the morphology and reference field conditions, properties of the input signaled constituent boundaries and the correct assignment of arguments to relational words. Although we do not expect that constituents in natural languages will always be so clearly demarcated as they were in our reference field and morphology conditions, we believe that such grouping information is necessary for successful syntax learning.

We find no reason to believe that semantic reference provides a privileged source of grouping information. In fact, there may be reasons to think that it is not privileged. A learner who infers syntactic constituents from a network representation of the referent, as suggested by Anderson (1975, 1983), will encounter difficulty in acquiring syntactic constructions in which the phrase structure systematically departs from the structure of that referent. One such construction in English is exemplified by the sentence *The girl expected him to be tall*. Here the logical of *expected* is neither *him* nor *to be tall,* but rather the proposition *he be tall* which refers to a single object in the real world. Nonetheless, *him* and *to be tall* do not fall in the same constituent. This sentence may well be problematic for any learner. However, a learner who expects congruencies between surface constituency and reference will be forced to adopt a new, strictly syntactic procedure to accommodate such sentences (see the simulations reported in Anderson, 1983). In contrast, for the learner who, from the outset, used the distribution of grouping cues and other surface properties of sentence strings to induce constituent structure, such syntactic properties could once again be recruited in the acquisition of complex sentence types.

Interestingly, recent work in first language acquisition provides converging evidence that to a surprising extent, language acquisition may proceed independently of properties of the reference field. In American Sign Language (ASL) and other naturally evolved sign languages used in deaf communities around the world, the forms of manual signs frequently, but by no means always, bear striking resemblances to objects and actions within the respective semantic extensions of those signs. To an adult observer, these *iconic* relationships between the forms of signs and the form of the referent world would seem to blaze a natural path into the linguistic system. Yet, studies of the acquisition of ASL by children who are acquiring ASL in the home from deaf, signing parents have uncovered no evidence that children either learn individual vocabulary items or formulate linguistic generalizations about the structure of morphologically complex signs on the basis of iconic relationships between form and reference (Bonvillian, Orlansky, & Novack, 1983; Meier, 1981, 1986; Newport, 1981; Supalla, in press; and, for a review, Newport & Meier, 1986). Rather, children's acquisition of ASL signs seems to be determined in large part by the relationship among linguistic forms, for example, by language-specific properties of the way in which morphemes are organized with respect to each other within morphologically complex signs.

Similarly, information available in the pattern of relationships among linguistic forms, specifically the relationship between the distribution of the suffixes and the distribution of the words, proved sufficient to enable our concord morphology subjects to learn the miniature syntax. In this fashion, adult subjects can learn syntactic rules from sentence exemplars presented without reference fields. Morgan et al. (1986) argue that marking of constituent structure by relatively meaningless linguistic units, such as concord inflections, is universal in human languages. In English, function words and prosodic intonation are markers of constituent structure and may subserve the learning of such structure. Thus, to invoke the full inductive capacity of the adult subject (and, we expect, of the child) and to avail the learner of all systematic markers of syntactic structure present in natural language, a theory of language acquisition must equip the learner to encode utterances, not merely as strings of words, but also as strings which are chunked by grouping cues for constituent structure.

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). Hillsdale, NJ: Erlbaum.

ANDERSON, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard Univ. Press.

ANDERSON, J. R., & BOWER, G. H. (1973). *Human associative memory*. Washington, D.C.: Winston.

BERWICK, R. C., & WEINBERG, A. S. (1984). *The grammatical basis of linguistic performance*. Cambridge, MA: MIT 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.

BILLMAN, D. (1985). *Increasing efficiency in acquiring syntactic categories: An adaptive attentional procedure*. Paper presented at the meeting of the Society for Research on Child Development, Toronto.

BONVILLIAN, J. D., ORLANSKY, M. D., & NOVACK, L. L. (1983). Early sign language acquisition and its relation to cognitive and motor development. In J. Kyle and B. Woll (Eds.), *Language in sign: An international perspective on sign language* (pp. 116–125). London: Groom Helm.

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

BROWN, R., & HANLON, C. (1970). Derivational complexity and the order of acquisition of child speech. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley.

CLARK, H., & CLARK, E. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.

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

GOLD, E. M. (1967). Language identification in the limit. *Information and Control*, **10**, 447–474.

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.

HALE, K. (1983). Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory*, **1**, 5–47.

KIMBALL, J. P. (1973). Seven principles of surface structure parsing in natural languages. *Cognition*, **2**, 15–47.

LYONS, J. (1968). *Introduction to theoretical linguistics*. Cambridge: Cambridge Univ. Press.

MARATSOS, M. P., & CHALKLEY, M. A. (1980). The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. E. Nelson (Ed.), *Children's language* (Vol. 2, pp. 127–214). New York: Gardner Press.

MEIER, R. P. (1981). Icons and morphemes: Models of the acquisition of verb agreement in ASL. *Papers and Reports on Child Language Development*, **20**, 92–99.

MEIER, R. P. (1986). *Iconicity and language learning: The acquisition of verb agreement in American Sign Language*. Manuscript submitted for publication.

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

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.

MOESER, S. D., & BREGMAN, A. S. (1973). Imagery and language acquisition. *Journal of Verbal Learning and Verbal Behavior*, **12**, 91–98.

MORGAN, J. L., MEIER, R. P., & NEWPORT, E. L. (1986). *Structural packaging in the input to lan-*

guage learning: Contributions of intonational and morphological marking of phrases to the acquisition of language. Manuscript submitted for publication.

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.

MORI, K., & MOESER, S. D. (1983). The role of syntax markers and semantic referents in learning an artificial language. *Journal of Verbal Learning and Verbal Behavior,* **22,** 701–718.

NEWPORT, E. L. (1981). Constraints on structure: Evidence from American Sign Language and language learning. In W. A. Collins (Ed.), *Minnesota Symposia on Child Psychology: Vol. 14. Aspects of the development of competence* (pp. 93–122). Hillsdale, NJ: Erlbaum.

NEWPORT, E. L., & MEIER, R. P. (1986). The acquisition of American Sign Language. In D. I. Slobin (Ed.), *A crosslinguistic survey of language acquisition* (Vol. 1). Hillsdale, NJ: Erlbaum.

PINKER, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard Univ. 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). Washington, DC: Georgetown University Monograph Series on Languages and Linguistics.

SUPALLA, T. (in press). *Structure and acquisition of verbs of motion and location in American Sign Language*. Cambridge, MA: Bradford Books/MIT Press.

WEXLER, K., & CULICOVER, P. W. (1980). *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

(Received November 19, 1985)
(Revision received March 25, 1986)