cluded data from five to ten informants, questioned individually. Each informant within a given study was presented with the same set of words, although the order of presentation was generally varied from one informant to the next. Informants made up their own sentences, except that completely general sentence frames (of the type, "I saw a ________") were discouraged if they occurred with great frequency. After the data had been collected from a group of informants, the individual data matrices were coded on data-processing cards, summed over subjects, and subjected to analysis.

The first study using this technique (details are included in Appendix B) employed thirty-five terms from all parts of the *seŋ* chart. Its major features can be summarized as follows: The major split between *town* and *forest things* and most of the classes at the next level of specificity are approximately maintained. However, ambiguities masked by, but present within, the orderly presentation of the *seŋ* chart appear. For example, *working things* does not appear as a unitary category; its ambiguous status in the chart is reflected by its tendency to divide into clumps that attach to other classes on the chart. *Cooking things* and *town animals* appear in a group next to *foods,* reflecting what seem to be natural relations that are suppressed in the search for order in the dichotomy between *town* and *forest* in the chart. Likewise, *structures* appears together with *clothing, tools,* and *sleeping things,* all of which are kept in the houses that compose the village.

Additional elicitations following the same procedure were subsequently made to evaluate relations within the two major groups, *town* and *forest*. In one case the terms were all *town things,* and in a second the terms were all *forest things*. In these studies the terms used included names of minimal species as well as general class names. In all these cases we observed ordering of the stimulus terms that was generally consistent with the ordering in the *seŋ* chart but often *seŋ*-chart distinctions failed to be reflected, and occasional large discrepancies were encountered (see Appendix B).

Free Association

In order to reduce the remaining ambiguity about class membership and interclass relations, one further kind of study was introduced, the free-association experiment. This technique for eliciting information about

the structure of noun classes requires people to associate to words or other stimuli. For example, a person can be asked, "what do you think of when I say, 'cars'?" Properly analyzed, the free-association technique can provide evidence on the extent to which the set of stimuli (in our case, words from the *seŋ* chart) elicit each other and other words as associations. The strength of association among a set of stimuli, as well as class membership, can be evaluated, using both common associates (*cat* and *dog* both elicit the associate *hair*) and direct associations (*cat* and *dog* elicit each other) as indicators of the relationships among words (Deese, 1962). The details of this work as applied to the classes of the *seŋ* chart are contained in Appendix C. Once again the general result was a replication of the content of classes contained in the *seŋ* chart, although as was the case when the sentence-substitution method was used, the pattern of relationships among classes was often different, and in some cases class membership was different (for example, snake is grouped with items that fall in the class of *medicines* rather than the class of *animals*).

Discussion of Alternative Verbal-Eliciting Techniques

Before considering a more detailed exploration of a limited domain of nouns that we will use in connection with experiments to be described in later chapters, we would like to comment on the significance of the three general studies of the Kpelle domain of *seŋ*. Each of the techniques we used (formal elicitation, sentence substitution, and free association) has certain virtues and certain drawbacks. At first glance the formal inquiry carried out by John Kellemu would appear to be the most direct way to obtain information about class membership, class inclusion, and class subordination. At the same time our informants experienced some difficulty in keeping classes discrete because class membership is determined by the attributes considered important, and these central attributes shifted from time to time. Agreement about class membership once given the class name was not a problem, but items often were included in more than one class.

As the elicitation was made less constrained (using the sentence-substitution and free-association techniques), the ambiguities of class relation were more clearly expressed and the resultant structural description was less definite. The free-association task even gave rise to instances where the preferred responses to category names were members of different categories. All in all, our impression is one of considerably more variability and ambiguity than customarily appear in reports of linguistic elicitations of the type studied here.

To some extent the greater heterogeneity in our work with the *seŋ* classes is the result of our own relatively informal approach to elicitation and the overly ambitious attempt to describe the contents and relations among such a large set of classes. In addition the kinds of responses produced in each of the eliciting situations certainly depends upon subtleties of the way in which the informants interpret the aims of the elicitor. Considering all the possible sources of variation among the tasks, we have been impressed by the fact that the major contrasts and most of the specific classes of the *seŋ* chart repeatedly occurred in the different studies. However, the hierarchical relations among classes at lower levels of the chart probably represent only one of several possible ways in which various subclasses of *seŋ* can be related.

A Detailed Look at Words Used in Experimentation

For most experimental purposes the *seŋ* chart is too large a conglomerate to work with, although two studies described in Chapter 4 (pp. 94–96) are specifically concerned with organizational features of Table 3–1. In most of our memory and verbal-concept learning studies, we concentrated our attention on a relatively small subset of the things described in the *seŋ* chart.

Our experimental aims called for us to obtain two lists of twenty items each. Both lists had to consist of common nouns naming physical items that were small enough to be easily transported by our research assistants. One set of items (the categorizable list) was to consist of twenty words divisible into four closely knit classes (we chose *foods, tools, clothing,* and *utensils*), while items on the second list (hereafter referred to as the noncategorizable list) should have only minimal semantic connections with any other items on that list.

Our first objective was to obtain a set of common clusterable items. We hit upon the following informal listing procedure to generate the list. Our assistant walked around the local town recording the answers of adult informants to questions such as the following: "If you were to go to the market, what kind of things could you buy?" or "What kinds

of things can you buy in Ukatu's store?" We assumed that objects named by a majority of the people questioned were relatively common. From the most commonly named objects, the list of twenty terms presented in the left-hand column of Table 3–2 was constructed. According to the *seŋ* chart, these twenty words fall into four separate semantic categories. Three of these are subheadings of the general heading,

TABLE 3-2

List of Items Used in Kpelle Free-Association and Experimental Studies

CLUSTERABLE	NONCLUSTERABLE
plate	bottle
calabash	nickle
pot	chicken feather
pan	box
cup	battery
	animal horn
potato	stone
onion	book
banana	candle
orange	cotton
coconut	hard mat
	rope
cutlass	nail
hoe	cigarette
knife	stick
file	grass
hammer	pot
	knife
trousers	orange
singlet	shirt
headtie	
shirt	
hat	

household things (*clothes, tools,* and *utensils*), while the fourth consists of five instances from the general heading *food,* two from the subheading *root crops* (*onion* and *potato*), and three from the subheading *tree fruits* (*banana, orange,* and *coconut*). In terms of our desire to obtain clearly discrete "clusters," these classes do not appear to be optimal since three of our classes are subheadings of the general category, *household things.* But since we needed not only classes that were discrete but a list of common, familiar objects as well, our overall purpose seemed to have been fulfilled, although judgment about class discreteness had to await verification by verbal-elicitation studies.

A corresponding list of noncategorizable items was constructed in a different and slightly more haphazard manner. Working with two informants, we constructed a list of sixteen items that our informants claimed (1) were common objects known to everyone in Kpelle culture; (2) were small enough to be easily transported by our research assistants; and (3) if compared with our clusterable list and with each other, were judged to be dissimilar to the objects on the clusterable list and would themselves be dissimilar.

The list of sixteen noncategorizable items obtained in this informal manner is presented in the right-hand side of Table 3–2. For experimental purposes we added four terms from the categorizable list to this list to complete the set of twenty items.

The informal manner in which these lists were constructed was clearly not an adequate foundation for our experimental work. Not only did intuition play a large role in their construction, but what little evidence we had cast doubt on the cohesiveness of one class on the categorizable list. Consequently, the entire list of thirty-six items (the twenty categorizable and sixteen non-categorizable items) was subjected to the sentence substitution-analysis previously applied to the *seŋ* chart as a whole. The rearranged ordering of words from this analysis is presented in Table 3–3.

At the left side of Table 3–3, from top to bottom, are listed the thirty-six words in the order of their similarity to each other as defined by the technique. This list is separated at intervals defined by our a priori hypothesis that *utensils, clothing, food,* and *tools* would tend to group together in the list. Inspection of the table indicates that such groupings in fact occurred. There are three kinds of evidence for this in Table 3–3. First, in the rearranged order based on similarity scores (which reflect the extent to which two stimulus words substitute in a like manner into a variety of sentences), all of the hypothesized categorizable classes occurred in groups, separated by items from the noncategorizable list. It should be emphasized that these words were presented in different random orders to each subject. Hence the reordering was clearly not predetermined. Second, the average similarity number (which can vary from zero to one) for all the relations within a semantic group are consistently higher than the corresponding numbers between groups or between noncategorizable words. Third, the similarity among the items from the noncategorizable list is lower than among the categorizable items. These average similarity numbers are listed beneath the class name in the case of within-class scores, and between hypothetical classes in the case of between-class scores. We should note that Table

TABLE 3-3

Rearranged Experimental Terms

ITEM		SUBHEADING	HEADING
calabash			
bottle			
pot		utensils	household
pan		.814	
cup			
plate			
.773			
box			
animal horn	.797		
book			
.756			
trousers			
singlet			
shirt		household clothes	household
headtie		.817	
hat			
.693			
cotton			
rope	.702		
stick			
grass			
.689			
onion			
potato		root crops	food
banana		.893	.821
orange		tree fruit	
coconut		.830	
.684			
cigarette	.766		
nail			
.712			
file			
hammer			
hoe		tools	
knife		.810	household
cutlass			
.656			
hard mat			
candle			
stone			
battery			
chicken feather	.679		
nickle			

3–3 reflects only the relative degree of "categorizability." The absolute size of the similarity score is affected by several factors. An important determinant of the overall level of similarity is the generality of the sentence frame used. This is why we emphasize the relative nature of "categorizability."

The results contained in Table 3–3 suggest that we have indeed hit upon two lists, one of which consists of relatively cohesive, distinct classes, while the other does not.

Free Association and the Experimental Terms: The Kpelle

One further study of the properties of the classes, *tools, utensils, food,* and *clothing* was undertaken, this time using free association as the eliciting device. The results are discussed here because up to now we have presented no data for Kpelle groups other than traditional adults and no data employing similar materials with American subjects.

In working with young subjects neither the formal eliciting procedure employed by Kellemu nor the sentence-substitution technique are easy to use. They both require periods of data collection considerably longer and more arduous than children are likely to cope with. The free-association technique, on the other hand, is quickly and easily administered. With proper analysis free-association data can yield information about category grouping that is comparable to the information yielded by the sentence-substitution technique.

Based on these considerations, free associations to the experimental words were studied in samples of subjects drawn from each of the following three Kpelle groups: (1) eighteen- to twenty-year-old students enrolled in the ninth to eleventh grades; (2) eighteen- to twenty-year-old nonliterate adults, who spoke little if any English; (3) ten- to fourteen-year-old children enrolled in the second to fourth grades. Each group was presented a list of twenty-four words one at a time (the twenty clusterable terms detailed above, and the four appropriate category names, *food, tools, clothing,* and *utensils.*) The lists were presented in a random order, and each subject was required to give at least four responses to each stimulus word if possible. All responses were tape-recorded for later transcription; then the data were analyzed according to the technique described in Appendix C. Because the results for the three groups

were quite similar, with an average correlation between group performance of .89, we present only the data for the group most similar in age and educational level to the American groups to be described, the ten- to fourteen-year-old schoolchildren (Table 3–4).

TABLE 3-4

Kpelle Free-Association Overlap Scores for Twenty Experimental Words

	FOOD	CLOTHING	TOOLS	UTENSILS
Food	.468	.010	.016	.010
Clothing	XXX	.268	.020	.033
Tools	XXX	XXX	.620	.182
Utensils	XXX	XXX	XXX	.731

NOTE: Entries on the diagonal represent within-class overlap scores, while entries above diagonal represent between-class overlap scores.

The data in Table 3–4 indicate the degree of similarity in the associational responses to stimuli within a given category and between categories. The actual numbers are average "overlap" scores reflecting the extent to which the various stimulus words elicit common responses (see Appendix C for a discussion of this measure of similarity).

As can be seen from Table 3–4, the four categories from the clusterable list emerge as readily recognizable groups; the average similarity (overlap) among scores within each of these groups is considerably higher than the overlap between groups. *Tools* and *utensils* produce quite high similarity scores and appear to be relatively tight, compact clusters, while *food* and *clothing* appear to represent somewhat looser groupings. In addition, a relatively high interrelation appears to exist between *tools* and *utensils* when they are compared with each other, while there is virtually no interrelation between any other of the possible pairs of classes.

A convenient graphic method for representing the way in which these items group themselves according to their overlap scores is presented in Figure 3–1. Using a technique introduced by S. C. Johnson (1967), Figure 3–1 represents the hierarchical grouping of the twenty four words according to their similarity scores. The greater the similarity between items, the closer to the right-hand side of the figure is the point where they are connected by a line. The numbers at the top of the figure represent the degree of overlap represented by items that intersect at that

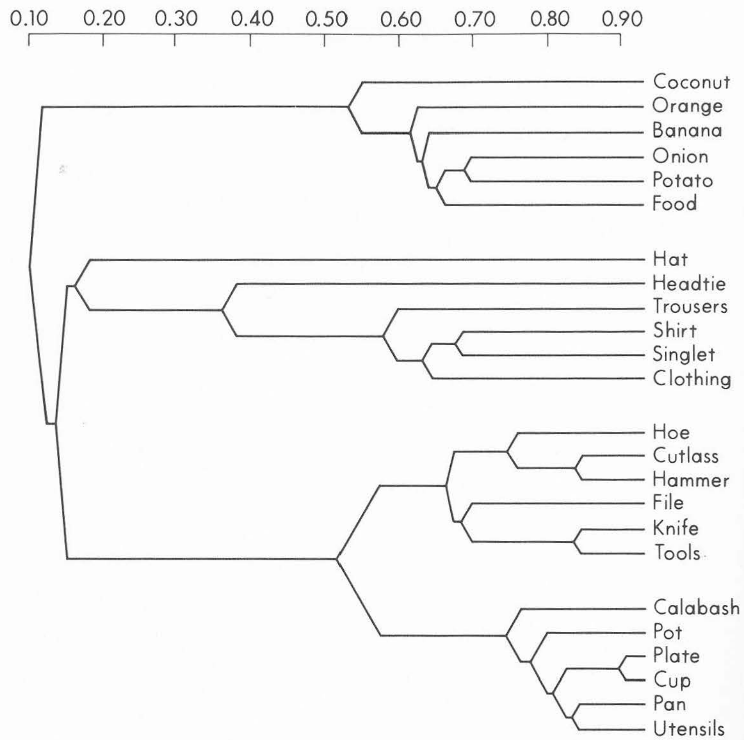

FIGURE 3–1 Results of Johnson Hierarchical-Clustering Program Applied to Free-Associations of Clusterable Stimuli; Kpelle Subjects

point. Both the distinct nature of the groups and the way in which *tools* and *utensils* come together to form a higher-order class are represented in the graph; each individual class clusters prior to the point where any items between classes meet, and the *tool* and *utensil* classes meet at a relatively high level of overlap.

Free Association and the Experimental Terms: The Americans

Since much of the experimental work to be reported later involved American comparison groups, we felt it necessary to study American free associations to these classes. In conjunction with this phase of our work, we studied three groups of American schoolchildren (grades one

and two, three and four, and five and six) from the Newport-Costa Mesa, California school system.

Each subject was seated in front of the experimenter and given the following instructions:

> ________ (child's name) I am going to say several words. Each time I say a word I would like you to tell me the words that come into your mind. For instance, if I say "cat," you could say "dog," or "mouse" or "fur" or anything else you think of. Do you understand?

The stimuli used in the American studies are listed in the left-hand column of Table 3–5 and appear comparable to those used in the African studies both in the nature and the membership of the classes. But such "face validity" is less compelling than hard evidence that the stimuli are behaved toward in similar ways by the two cultural groups. Children

TABLE 3-5

List of Words Used in American Free-Association and Experimental Studies

CLUSTERABLE ITEMS	NONCLUSTERABLE ITEMS
glass	candle
pot	book
pan	pot
cup	bottle
plate	cotton
	cigarette
hammer	box
knife	feather
ax	stone
saw	mat
file	battery
	nickle
banana	knife
orange	shirt
lemon	stick
potato	nail
onion	orange
	rope
sox	horn
shoes	grass
shirt	
hat	
pants	

NOTE: All are high frequency items from the Thorndike-Lorge (1943) tables.

were read the stimulus words one at a time and were asked to give at least four responses to each. The responses were recorded in order on separate data sheets and analyzed in the same way as the Kpelle free-association data.

Results of the hierarchical grouping analysis are presented for the combined age groups in Figure 3–2 and the average similarity scores are presented in Table 3–6. The results presented in Figure 3–2 and Table 3–6 resemble those for the Kpelle subjects as reported in Figure 3–1 and Table 3–5 above. All four categories emerge as readily identifiable units with relatively high within-class overlap scores. As shown in Table 3–6 between-class similarity scores are generally low. However, all the intergroup scores are not equal, and there appears to be some relation between *foods, utensils,* and *tools,* unlike the Kpelle data. This interrelation between categories is reflected not only in the average sim-

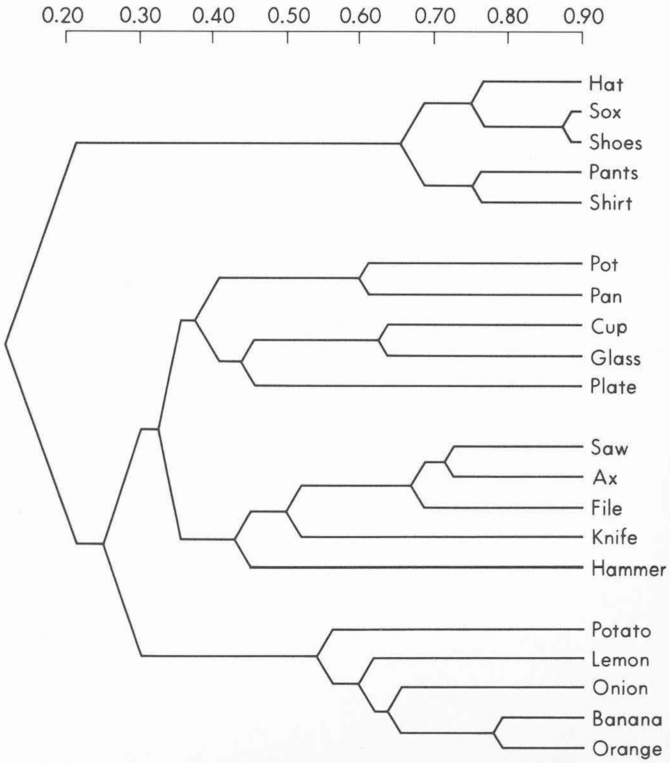

FIGURE 3–2 Johnson Hierarchical Output for All Three American Groups Combined for Word Stimuli

TABLE 3-6

American Free-Association Overlap Scores for Experimental Words

	FOOD	CLOTHING	TOOLS	UTENSILS
Food	.547	.014	.048	.104
Clothing	XXX	.626	.021	.025
Tools		XXX	.466	.100
Utensils			XXX	.394

ilarity numbers of Table 3–6, but also in the order in which the categories join in the tree diagram of Figure 3–2 (first *tools* joins *utensils,* shortly after, that pair of classes is joined by the *food* items).

Kpelle and American Free-Association Data Compared

Perhaps the major generalization to emerge from a comparison of the group performances is one of cross-cultural similarity: in both cases the four basic categories which we have labeled *tools, clothing, food,* and *utensils* emerge. However, the similarity is not an identity because the relation among categories seems to differ: *tools* and *utensils* form a higher-order grouping for the Kpelle, while *foods* joins *tools* and *utensils* in the American data.

Other aspects of the free-association data also indicate that there may be important differences in the associations evoked by the stimuli, which are not reflected in general measures of similarity. With the exception of the category *clothing,* overlap scores are higher among the Kpelle. The greater average similarity among responses evoked by these stimuli among the Kpelle arises from several more specific characteristics of the associated responses. For one thing there is better agreement (more stereotypy) among Kpelle subjects in the particular words that they choose as associates; the first four most common responses to each word constitute 55 percent of the Kpelle data but 44 percent of the American data. Another indication of higher stereotopy in the Kpelle responses was that 89 percent of their responses fell within the same semantic class. The American subjects, on the other hand, generally gave fewer responses within the same semantic classes, and there was a clear difference between the children in grades one to four and grades five to six in this regard (no major differences among Kpelle groups were detected). The younger American subjects gave within-class responses about 73 percent of the time, while the fifth to sixth graders did so only 35 percent of the time.

On the basis of verbal-classification data presented thus far, it would not be profitable to speculate on the significance of the second-order differences in response patterns. We will return to these data, however, after we have studied how the various groups respond in other classification situations.

For the moment several conclusions are suggested by our examination of the materials to be used in experiments. First, the categories of *tools, utensils, food,* and *clothing* emerge in both the Kpelle and American settings. Second, there seem to be differences in the way in which these categories are produced by different groups: the Kpelle responses are much less variable, tending strongly to be "other examples" of the class to which the stimulus word belongs. The American responses show greater variability both in the particular words chosen and semantic class to which they belong. Finally, the three Kpelle groups yield very similar associational patterns, while there is an increase in response variability with increasing grade (age) among the American subjects.

Looking ahead to experimental studies using these stimuli in learning problems, we can be confident that we are dealing with categories that are recognized and used by both cultural groups. However, the secondary differences in the way the items are grouped into categories may present problems of interpretation.

Nonverbal Measures of Classification

We have thus far discussed techniques for discerning categories implicit in the organization of verbal responses. We found that each eliciting technique revealed a generally orderly set of categories, yet with subareas of ambiguity and cross-classification. We now consider the relation between semantic classes as elicited by verbal techniques and classes manifested in nonverbal behavior involving sorting of the objects themselves. Does semantic organization as measured verbally describe the actual manipulation of the objects?

When we consider the question of the relation between linguistic categories and nonverbal behavior, we are moving into an area of great current interest for the study of the relation between language and thought. For example, J. B. Carroll and J. B. Casagrande's (1958) studies referred to earlier, involved sorting objects that are categorized differently in Hopi and English; Greenfield (Bruner, Olver, and Greenfield, 1966) used a sorting task with Wolof children in order to assess the role of linguistic and educational factors on classification behaviors.

Our emphasis in this chapter differs from that of the authors cited. We know, on the basis of our linguistic elicitations, that certain stable semantic classes exist in the Kpelle language and are expressed in various verbal contexts. We wish to determine the ways in which these categories control the nonverbal behavior of our subjects.

Sorting

Our first study in which we asked subjects to classify objects was conducted before we had settled on a set of coherently, semantically classified items. In order to test subject's responses to the request to classify objects and to work out instructions, we gathered together a set of items that were potentially classifiable in a variety of ways—according to function, semantic class, length, size, color, and so forth. These objects were laid out on the floor in front of the subject, who was asked to sort them into piles that made sense to him.

The dominant mode of classification in this pilot work was what we have called "functional entailment." A pair of objects was selected so that the first went with, or operated on, the second. For example, a potato and a knife were put together because "you take the knife and cut the potato." Very rarely was a large group formed, and we virtually never had a classification justified in terms of the way things look or their common membership in a taxonomic category. This work is discussed in J. Glick (1968).

When we had settled on the set of clusterable items listed in Table 3–2, we used them to pursue the question: what control does category membership exert over the classification of objects? Two features of this work differed from our pilot studies. First, the entire set of objects could be classified according to membership in a semantic category (in our pilot studies, no single criterion of classification could exhaust the list). Second, we constrained the number of classes that the subject could make.

The task of sorting these twenty objects was given to three groups of

Kpelle subjects: ten- to fourteen-year-old schoolchildren in grades two to five, ten- to fourteen-year-old children who had never attended school, and nonliterate, Kpelle-speaking adults aged eighteen to fifty years.

When the subject entered the room where the experiment was conducted, he saw the twenty experimental items arranged on a table before him in a manner that was intended to be haphazard. In addition there were chairs (two for half the subjects and four for the remainder) arranged against one wall of the room with a two-foot distance between chairs. Subjects were then instructed as follows (in Kpelle):

> These things that are on the table, we are going to divide them into four (two) groups. You should find some sense to divide these things. Here are four (two) chairs around the table. Each chair is for one of the groups. (After the subject is finished you say): What sense did you use to divide these things?

Ten subjects from each of the basic population groups were included in the two-class and four-class conditions, a total of sixty subjects in all.

It seemed to us reasonable to assume that the provision of four chairs (classes), in the presence of objects belonging to four linguistic categories, might be a powerful cue to sort according to these categories. The data did not confirm these expectations. The subjects frequently put objects from one class on more than one chair. Using as our measure the number of different chairs that members of a given category were placed on, Table 3–7 shows the distribution of items from each category for each subject group.

TABLE 3-7

Chairs per Category – Four-Chair Sorting Experiment

GROUP	CATEGORY				
	CLOTHES	UTENSILS	TOOLS	FOOD	AVERAGE
Ten- to fourteen-year-old schoolchildren	2.2	2.4	2.6	2.8	2.50
Ten- to fourteen-year-old nonliterates	2.9	2.2	2.8	2.3	2.55
Adult nonliterates	2.8	2.8	3.1	2.7	2.85
Average	2.63	2.47	2.83	2.60	

It is clear from Table 3–7 that each category occurs on an average of two or more chairs, and hence we are not observing perfect, semantically based sorting. As was the case in the free-association work among the Kpelle, there were no statistically reliable differences among the three groups in the way they sorted the objects.

If the subjects were not performing perfectly according to the previously elicited category system, there remains the possibility that some alternative category system better describes the data. In order to test this possibility, we computed a score for each item based on the proportion of times it occurs with each other item on a chair. This co-occurrence matrix was then analyzed using Johnson's hierarchical clustering program described earlier in connection with our free-association work (Appendix C).

The Johnson method forces the co-occurrence data into a hierarchy, possibly even one that we might consider inappropriate. If groups and hierarchies different from those predictable from categorical membership are produced, we have evidence for an alternative organizational scheme or perhaps the absence of a consistent scheme.

The hierarchical organization of these twenty items is presented in Figure 3–3 for all three groups combined. This figure should be interpreted as follows (the interpretation is similar to that applied to the free-association data). Two items have co-occurred to a degree indicated by the number on the top of the figure at the point where they are joined. For example, *trouser* and *cap* were sorted together nineteen times. Since the maximum is thirty (the number of subjects in the combined groups), this would be a relatively strong association between these two items.

Inspection of Figure 3–3 indicates that the items are, in general, organized with respect to category membership—although inspection of the strength of the association suggests that this mode of organization is not strong. A reasonable conclusion seems to be that there is an absence of any general violation of categorical expectations because there is no strong alternative way to categorize the items. The linguistic categories are weakly expressed, but are stronger than any other system when the items are sorted onto four chairs.

In the two-chair condition, sorting according to categorical membership appeared at first glance to be much stronger than in the four-chair condition: most commonly two complete categories were placed on each chair. In interpreting this finding it should be remembered that with only two chairs on which to place the objects, there is a greater chance possibility of perfect sorting. Taking the four-chair condition as a basis

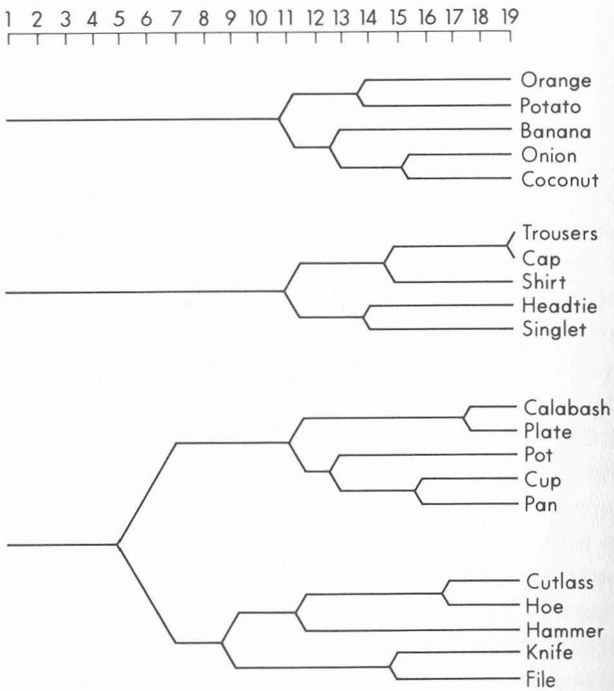

FIGURE 3–3 Hierarchy Generated by Sorting Clusterable Items onto Four Chairs

for estimation, we might expect a dispersion measure for the two chairs that is half that presented in Table 3–7. The expected score would be approximately 1.25 chairs per category, with an ideal of 1.0 chair per category, a difference that is too small to allow statistical comparison.

Accordingly, some other means was needed in order to study the apparently better categorical performance encountered in the two-chair condition. To do this we shifted our attention to the justifications given by subjects after they had sorted the objects in either of the two conditions. Two kinds of reasons were frequently offered by our subjects: (1) justification in terms of categorical membership, and (2) justification in diffuse, noncategorical, terms. Typical noncategorical classifications were "I divided them into groups" or "Because my Kpelle sense told me." We calculated the distribution of types of reasons for each of the conditions (since there are not between-group differences of any consequence, the groups are combined for this analysis).

In the two-chair condition, twenty-eight of a total of thirty reasons were categorical in nature, while in the four-chair condition, less than

half (twelve of thirty) are of a categorical nature. This difference between the two-chair and four-chair conditions is statistically reliable ($\chi^2 = 16.2$, $df = 1$, $p < .01$).

We have no way of explaining this difference at the present time, but it raises a point to which we will return repeatedly: in trying to assess people's competence with respect to a particular task, one has to be very careful to consider more than one form of the task. There is no general answer to the question: can the Kpelle describe their categorizations? The answer, as we see, depends upon the particular conditions of categorization.

Similarity Mediation: Constrained Category Construction

The pattern of results we obtained when subjects were asked to categorize the full set of twenty items seems to indicate that greater constraints placed on the mode of categorization produced results more consistent with the underlying semantic category structure. The procedure followed in the present experiment imposed a maximum of constraint under conditions where, it was hoped, the actual basis for grouping objects could be assessed.

We began the experiment by selecting fifteen of the twenty items listed in Table 3–2, the *tools, foods,* and *utensils.* These items were spread haphazardly on a table at which the subject was seated. According to a prearranged list, pairs of items (the "constraint pairs") were placed before the subject about twelve inches apart. The subject was then instructed as follows:

> Do you see these things on the table? I want you to place one of these things between the (Item 1) and the (Item 2) so that the thing you choose and the (Item 1) are alike in some way and at the same time the thing and (Item 2) are also the same thing in some way.

When the subject had selected an object for placement, he was asked why that object belonged with the constraint pair.

These instructions and paraphrases thereof were difficult for the subjects to understand. Moreover, several subjects were unwilling to give full justification of their selections. Nevertheless, we felt that sufficient understanding was achieved for the experiment to continue, and as we shall see, consistent categorizing behavior was observed.

The pairs of objects were presented one at a time and ample opportunity was given for subjects to respond. Different subjects were presented constraint pairs in different orders, the only restriction being that the categorical makeup of the pair changed on each trial. The selection of constraint pairs (presented in Table 3–8) represented all possible combinations of the categories *food, tools,* and *utensils* (food-food, food-tool, food-utensil, tool-tool, tool-utensil, utensil-utensil). Three pairs were chosen to represent each combination of classes by a random selection procedure.

TABLE 3-8

Object Pairs by Category Used in Similarity-Mediation Experiment

CATEGORY PAIRS	OBJECT PAIRS
food-food	onion-orange potato-orange banana-coconut
tool-tool	hoe-cutlass hammer-knife cutlass-hammer
utensil-utensil	pan-cup plate-calabash plate-pot
food-utensil	coconut-calabash orange-cup potato-pot
food-tool	orange-hoe banana-cutlass coconut-hammer
utensil-tool	calabash-file pot-cutlass calabash-hoe

Approximately thirty subjects were chosen from each of four groups. Three populations were the same as those in the free-sorting study: ten- to fourteen-year-old schoolchildren in grades two to five, ten- to fourteen-year-old nonliterate children, and nonliterate traditional adults. In addition a group of high-school students, ranging in age from sixteen to twenty, was selected from the Lutheran Training Institute and the Zorzor Rural Teacher Training Institute.

RESULTS

We looked first at the kinds of items subjects used to mediate between the two given objects. Two major conditions need to be considered separately: the cases where the constraint items were from the same class (intraclass groups), and those where the constraint items represented two different classes (interclass groups).

Table 3–9 shows the category of the item chosen to complete each trial when the constraint objects belonged to the same class. According to this table, all groups except the high-school students made choices in very similar ways. The dominant choice of an object placed between two food items was a tool of some sort. In 94 percent of these cases, the tool was a cutting implement, either a knife or cutlass. In most cases these groups chose food items to mediate between two utensils. Where

TABLE 3-9

Category Membership of Mediating Objects for Like-Class Constraining Objects

SUBJECT GROUP	CONSTRAINING OBJECTS: FOOD-FOOD MEDIATING OBJECT		
	FOOD	TOOL	UTENSIL
NC[a]	1	30	2
A[b]	6	22	2
SC[c]	6	27	0
HS[d]	28	4	1
	CONSTRAINING OBJECTS: TOOL-TOOL MEDIATING OBJECT		
	FOOD	TOOL	UTENSIL
NC[a]	1	32	0
A[b]	1	29	0
SC[c]	1	32	0
	CONSTRAINING OBJECTS: UTENSIL-UTENSIL MEDIATING OBJECT		
	FOOD	TOOL	UTENSIL
NC[a]	12	0	20
A[b]	16	1	13
SC[c]	7	1	25
HS[d]	4	0	28

[a] Ten- to fourteen-year-old nonliterate children.
[b] Nonliterate traditional adults.
[c] Ten- to fourteen-year-old schoolchildren, grades two to five.
[d] High-school students, age sixteen to twenty.

both constraint objects were tools, all groups overwhelmingly chose another tool to complete the set.

On the whole it appears that only high-school students used consistent categorical grouping. The other groups choose mediating objects in different ways, depending on the constraining pairs. Even in the case of tools where the grouping appears to be categorical, as we shall see, the high-school and other groups chose the mediating tool for different reasons.

The nonliterate children, nonliterate adults, and schoolchildren also tended to make the same specific choice of an object to mediate between two tools. The set of objects that constituted the choice set consisted, after these tools had been removed, of thirteen objects, three of which were tools. Any of the three remaining tools satisfied the criterion of category membership. We might expect then that subjects' choices of mediating object would be randomly distributed over the remaining tools. In fact, in 93 percent of the cases, the non-high-school groups chose the file to mediate between all pairs of tools. In contrast the high-school group chose the file only 34 percent of the time—thus accurately reflecting its probability of occurrence if category members were treated in an equivalent manner.

Similar, though less dramatic, deviations from randomness appeared for other cases where the object was chosen from the same category as the constraint pair. Since, however, in other categories the percentage of intracategorical choices was relatively low for the non-high-school groups, analysis comparable to that offered for tool-tool relationships was not possible.

The reasons subjects gave for their choices were important in understanding this pattern of categorization. There are many ways to describe the subject's choices, but we found that two main categories of justification include most of the data. First, there are static responses which group the items in some fixed class. Typical are statements of a common category ("They are all food") or a common function ("They are all used in rice farming"). Second, there are dynamic responses which link the items in an action sequence that we earlier labeled functional entailment. In some cases the item chosen may act on the constraining items ("The knife can cut both the banana and the orange"); in others one constraining item may be said to act on the mediator and the other constraint item ("Soup from the calabash goes in the pan and the cup"); in still others the three items may be linked in a sequence of actions ("I can take the knife and cut the orange and drink the juice from a cup").

This type of response seems most natural when the constraining items belong to different categories.

In virtually all cases the high-school students justified their choices in static, categorical terms, while the choices by the non-high-school groups were not justified according to static rules of classification even in those cases where the object chosen was a member of the same class. An examination of the actual objects chosen for the tool-tool pair suggests the process involved. It will be recalled that a file mediated between two other tools in 93 percent of the cases for these groups. In virtually all these cases, the reason given was of the form "the file and the hammer sharpen the cutlass (knife)." In other words, the pragmatic use of the tools rather than their class membership was the basis for selection.

In general the same pattern of results appears in an analysis of the responses to the intercategory constraint pairs (see Appendix D). The high-school students handled an ambiguous situation by choosing a mediating item that shared one category relation with one of the constraint items and a different category relation with the other constraint item. For instance, one high-school student said that he placed a knife between a cup and an orange because the cup and knife are made of iron, while the knife (handle) and orange both come from trees. The other groups used dynamic, noncategorical justifications in virtually all cases. (The data are presented in Appendix D, Table D–3.)

Replication of the Similarity-Mediation Experiment with American Schoolchildren

These results obtained among the Kpelle naturally raise the question of how American schoolchildren would respond to the same task. Unfortunately, we only have data collected from children six to nine years of age (kindergarten, first and third grades) so that comparisons are only legitimate with the Kpelle schoolchildren. However, certain patterns in the American data are sufficiently striking to warrant mention.

Two major contrasts between the American groups and the Kpelle groups stand out. First, the American children, especially those in kindergarten and first grade, often justified their responses in terms of only *one* of the constraint pairs, even though prompted to relate their choices to both objects. Moreover, static categorical reasons dominate when only two of the objects are related.

Second, younger American subjects have a strong tendency to give static categorical justifications in terms of a visual attribute of the stimulus, such as color or form. On the other hand, the African groups gave, at most, 1 to 2 percent of their responses in terms of such attributes, a fact in strong contrast to results we will report in later chapters. These modes of justification, particularly the frequent restriction of a justification to pairs of objects instead of the triplet, make the American subjects appear to be considerably more adept at static classification than all but the high-school Kpelle. But such a conclusion would overlook the major point that the two groups, faced with the same task, *respond to it quite differently*. It might be, for instance, that our young American subjects, by responding only to pairs, reflect their implicit recognition of a "demand" for a static category. Conversely, the Kpelle subjects, for whom no special mode of classification seems "right," comply better with the instructions to consider all three objects in their choices. This raises an additional question: are there situations where the Kpelle do respond as if there was a "right" kind of classification?

The following section provides an illustration of one traditional Kpelle activity which seems to foster a particular kind of categorizing activity.

Sorting Leaves

The Kpelle play a game in which twenty to thirty leaves are tied to a rope. The object of the game is to name and describe the function of each leaf without hesitation—a long pause or an error and the player is "out." Since we remembered the importance of leaves in Kpelle medicine, this game suggested to us an experiment on classification in which both the materials and the procedure would be relevant to traditional Kpelle culture. We asked Akki Kulah, a Kpelle-speaking college student, to select twelve forest leaves, six from vines and six from trees. These twelve leaves were presented to ten nonliterate adult Kpelle farmers and ten American adults working in the Cuttington College area in the following manner. The leaves were spread out on a table in front of the subject, who was told: "I have some leaves here. I want you to sort them into two piles according to which ones you think go together. There should be six leaves in each pile." The experimenter then recorded which leaves were placed in each of the two piles.

As an index of the degree of category separation, the six vine leaves were assigned the numbers one to six and the six tree leaves the numbers seven to twelve. Then the averages of the numbers of items in piles

were computed. If a perfect score was obtained, the averages would be 3.5 and 9.5 for the vines and the trees, respectively, yielding a difference between the scores for the two piles of 6.0. The average difference between the two piles for the Kpelle subjects was 4.90, 84 percent of the maximum. On the other hand, the average difference for the American subjects was 2.48, roughly half that for the Kpelle.

Here is an instance where the Kpelle were asked to sort objects that have previously been sorted under similar circumstances. The task closely resembles the sorting tasks discussed previously in this chapter. Kpelle performance would be difficult to improve upon.

There should be little surprise at the American performance except insofar as the American subjects manifest any ability at all to distinguish the two unnamed categories. Their relative inferiority simply underscores the necessity of relevance and familiarity for successful performance of an act of classification.

Summary

We have now completed the presentation of our data on the classification of Kpelle nouns as reflected both in verbal tasks and in tasks that require subjects to classify objects. It should be apparent that even restricting our analysis to Kpelle nouns, the pattern of results is quite complex.

On the surface, at least, it appears that we have successfully identified rather stable class groupings in the way our Kpelle groups use nouns in the verbal tasks. When explicitly asked to identify class membership and to indicate superordinate labels for particular words in the eliciting technique that generated the *seŋ* chart (Table 3–1), we obtained a fairly orderly set of classes with seemingly well-defined relations of subordination and superordination. We also noted some conflicts in the chart caused by changes in class membership stemming from consideration of different attributes of objects at different times. Evidence from the other two verbal-eliciting procedures reflected both the class structure of the *seŋ* chart and the fact that alternative groupings were possible, depending on context. Using the sentence-substitution technique, we found that in some cases (for example, subclasses of *town things* and *working things*) similarity of the way the words were used in sentences better reflected everyday similarity in the use of the objects than the category relations of the *seŋ* chart. The same kind of relationships were to be found in the free-association data, although semantic category relations dominated.

Our interpretation of these data is that semantic classes can serve as a means of organizing verbal behavior, but the extent to which this happens in naturally occurring contexts is very much open to question. It is quite possible that in our desire to find "the" classification of Kpelle nouns, we overlooked situations where quite different kinds of classification would dominate.

For example, in conducting interviews prior to beginning the work that generated the *seŋ* chart, John Kellemu tried out a number of informal procedures designed to elicit ideas of category membership. The first few of these procedures produced interesting results but not a taxonomic classification system. In one, the men were given an example of a general class and asked to group all the examples they had named into subgroups. This technique elicited what Kellemu called "informal classification"; the men organized the objects according to *function* and *use* rather than according to a formal semantic system of classes.

Three other techniques also resulted in functional schemes of classification. In one, persons were asked to name all the things they had seen in a given day and then to group these things. A second technique was to have the men name all the things visible in a given scene and then to group them. The third asked them to name all the things similar to a given thing. In each case the subject drew on immediate experience.

The use of alternative classification principles emerges from our studies of the way in which objects are sorted. In pilot work using a large array of objects bearing no salient class relations to each other, functional pairing was a dominant means of classification. But when objects bore a class relation to each other, and when only two classes were permissible, semantic class relations were strongly expressed. Finally, where taxonomic relationships are habitually used as a basis for classification (such as was the case for the leaves that we asked our subjects to sort), the recognized taxonomic class was the dominant basis for classification.

In short, we have demonstrated to our own satisfaction that the Kpelle know and use taxonomic class relationships to structure their verbal behavior. But we have also established that the use of this kind of structuring is neither universal nor obligatory for the situations we have studied. The question then becomes: how do cultures lead people to adopt different kinds of classification systems under different circumstances? This is a very broad and difficult question to which we will turn in the concluding chapter of our book. In the remaining chapters we will pursue a closely related question. Under what circumstances do classification schemes enter into various situations where the subject is required to learn something new?

FOUR : Classification, Learning, and Memory

> To study the mind as a transformer, you have to ask subjects to do more than denote and define.
>
> ROGER BROWN, 1964, p. 251

The techniques described in Chapter 3 can be viewed as alternative ways to express habitual, well-learned relationships among things. The data reported there should reflect linguistic usage and its relation to the referents of words. Subjects are not asked to form any new categories but only to use familiar categories. We tacitly assumed that any learning that occurs in the elicitation or categorization situations does not change the nature of the subject's longstanding semantic habits.

In the present chapter we will reverse the direction of our attention. Based on what we have discovered about the Kpelle organization of nouns, we can now ask, "Does such organization affect the way in which subjects learn tasks whose successful performance is aided by the use of their lexicon, their store of 'things'?" In pursuing this question, we must be aware of such issues as the degree of structuring in the task, the nature of the stimulus materials, and the experiences of the subject populations. We have shown in Chapter 3 that such factors materially affect habitual categorization.

Verbal-Concept Discrimination

The question asked in this set of experiments is whether subjects use semantic categories when asked to learn a particular classification scheme. An alternative to the use of categorical information would be to learn the system by memorizing each item rather than using information

about the class to which it belongs. An example will clarify what we mean by these distinctions.

In each case we told the subject that we would name a series of pairs of things and that he should indicate each time which one of the two "we are thinking of." We would then tell him if he were right or wrong. Suppose that the kinds of things that appear in these pairs are either items of clothing or items of food, but the subject is not told this. One example of each class is presented on a single trial. When the first pair is presented (say, "orange-shirt") the subject has no way of knowing which item we have in mind. But if he learns that we had shirt in mind on the first trial, and if he is presented with "lemon-hat" on the second trial, he might infer that we always select clothing and, therefore, choose the hat on the second trial. On the other hand, he might merely memorize the correct item for each pair (shirt instead of orange, hat instead of lemon) and never recognize that it is always members of one particular semantic class that are correct. In this way he can only guess an item until it appears for the second time. The experiments discussed below continued a given series of pairs until the subject was correct ten times in a row or until forty-eight trials were completed. We then asked the subject how he knew which word we had in mind.

Using the basic technique outlined above, a series of experiments was designed to investigate two questions. First, how are these sets of conceptually orderable materials learned? Are they learned by rote or by a rule-governed process? Second, will the processes of learning or the speed of learning vary with features of the concepts, such as their interrelation as measured by the *seŋ* chart or their grammatical class?

Each subject had six different problems, selected from a large set of problems. The order for presenting the six problems was randomized across subjects so that we could independently determine whether categories in the material influence the rate of learning, whether or not subjects learn to learn this type of discrimination (by comparing performance across successive problems), and whether or not different types of concepts differ in their difficulty (by comparing performance across problem types regardless of presentation order).

Experiments Dealing with *Seŋ* Chart Organization

The first experiments use the concept-discrimination technique described above to investigate some possible implications of the *seŋ* chart (Table 3–1). Any conceptual hierarchy can be looked at as having both a vertical and horizontal dimension of organization. The vertical dimension refers to the fact that classes "higher up" on the chart are quite general and include classes lower on the chart. Looking at the *seŋ* chart, for example, we see that *town things* subsume such subclasses as *playing things, people,* and *town works,* which in turn subsume their own subspecies. The conceptual classes that we used in our experiments represented different levels of generality. If there is strong dependence of conceptual learning on the specificity or generality of the classes to be contrasted, then an experiment using the vertical organizational feature of the *seŋ* chart would indicate the influence of this particular aspect of class structure on learning.

The horizontal organization of the *seŋ* chart provides different evidence on interclass relationships. Here the question is not one of subordination but rather the relationship between specific classes according to the general class to which they both belong. In the *seŋ* chart, for example, the classes of *children* and *adults* are equally specific and are both subsumed under the category of *people*. Since they are subsumed under the same relatively low-order category, one might expect them to be rather closely related classes. The categories of *children* and *houses* are also equally specific, but they are related through the more general class of *town things,* since children belong to the class of *people,* houses belong to the class of *town works*. The "distance" between *children* and *houses* might be expected to be greater than the distance between *children* and *adults* because the former pair is related at a higher level on the *seŋ* chart. Note, however, that the classes involved are all equally specific, in that all are minimal species.

If the discriminability of classes is related to their distance from one another and learning is affected by discriminability, then an experiment making use of different levels of horizontal distance might provide independent evidence of the organization posited by the chart. Two experiments in which to-be-discriminated classes varied in their vertical and horizontal distance from each other were performed as a means of

learning about the situations in which class relations, such as those reflected in the *seŋ* chart, influence performance.

The first experiment investigated the identification of twelve classes: *food, clothing, utensils, tools, town animals, forest animals, insects, birds, trees, mats, town things,* and *forest things*. This group of items can be divided grossly into two general levels of specificity, as defined by vertical organization. Ten of the classes are relatively unitary and specific, while the two remaining categories (*town things* and *forest things*) are composed of a general selection of items that might be found within these major classes. Each class was represented by five examples which on a given problem were randomly paired with the five examples from one of the other classes.

The primary measure of performance was the number of trials required to attain a criterion of ten successive correct responses. In addition, we noted whether or not the subjects (nonliterate adults) could correctly verbalize the basis of their solution once they had attained it.

The learning we observed in this first experiment was extremely rapid, ranging from two to five trials required to obtain the criterion for different classes. More interesting than the bare fact of rapid learning are certain more detailed results.

For one thing, we found that the more specific classes were learned significantly more rapidly (an average of 2.8 trials) than the general classes (4.9 trials). In addition to supporting our characterization of these classes as "specific" and "general" in the Kpelle lexicon, these results suggest that the learning process needed to describe these results is *concept*-based and in some way related to the vertical organization of the *seŋ* chart. We can infer this because the number of items to be learned is the same in each problem. The fact that the more specific classes are learned more quickly suggests that the subjects are able to make prompt use of the more specific concepts. The fact that the general classes are learned in approximately five trials indicates that one presentation per pair is sufficient for learning. The fact that less than three trials are required to learn five pairs is rather convincing evidence that in the case of the specific categories the category rule governs learning.

Another aspect of the data warrants mention here. Not only were the general classes more difficult to learn, but also they were more difficult to describe following solution. Subjects were able to describe the basis of solution only 10 percent of the time for the two most general classes, but about 25 percent of the specific classes were described successfully.

The second experiment concerned with organizational features of the *seŋ* chart studied additional classes and investigated questions related to *horizontal* organization. Three major groupings of classes, representing different degrees of horizontal distance, were used. The least distant relationship (grouping 1) selected classes from within the same relatively low-order category (for example, *animals that drag themselves* versus *animals that crawl,* both being subclasses of *animals*). An intermediate distance was represented by classes that were related one step higher on the *seŋ* chart (for example, *children* versus *dancing things*). The greatest distance is represented by pairs of classes that fall on the one hand into *town things* and on the other into *forest things* (for example, *town animals* versus *shrubs*) and hence only share the common class of *things.* In all cases classes were at the same vertical level, low on the *seŋ* chart.

Each class was paired against three other classes, one from each of the distance groupings. In contrast to the first experiment, each class was composed of eight items. We felt that more clear-cut results relating to the operation of categorical structures could be obtained if the number of items within class were increased from five to eight so that simple rote learning of the items would be made more difficult by the greater number of items to learn.

One hundred and forty-four nonliterate adults each solved six problems representing all possible pairings of class distances. As before, our primary measure of performance was the number of trials required for the subject to attain the criterion of ten correct successive responses.

As in the initial experiment, learning was extremely rapid, ranging from 1.9 to 5.8 trials before the last error was made. More significantly, learning varied as a function of horizontal distance between contrast classes. The means for closest, intermediate, and most distant contrasts were 4.2, 3.7, and 2.9 trials respectively, suggesting that ease of learning is inversely related to distance as measured by the *seŋ* chart. The closest and hence presumably least discriminable pairings produced relatively slower learning than the most distant and most discriminable pairings.

Both experiments reported in this section point to the potential influence of semantic organization on learning, demonstrating that learning varies with both the specificity of classes and the distance among classes represented in the material to be learned.

Experiments Contrasting Rule-Based and Random Classes

Although suggestive and instructive, the results of the previous experiments dealing with *seŋ*-chart organization could be strengthened by a stronger investigation of the operation of categorical systems. In the first two experiments there was no control over the initial ease of learning the items involved in each of the contrasts. In order to achieve this kind of information, we sought a direct comparison of semantically grouped classes with random classes, made up of the same kinds of items. By random class we mean a set of items made up in systematically random fashion of members of both classes supposedly being discriminated. Assuming that there is no accidental order in such classes, the only way that subjects can discriminate the sets is to learn each correct item individually. A comparison between the rate of learning random and rule-based classes provides a way of evaluating the control that class membership exerts over learning. If subjects are learning in a stimulus-specific manner, which we can call rote learning, we would expect the learning rates for the random classes and rule classes to be identical.

With these thoughts in mind, we undertook another study, which was intended to sample a wide range of possible bases of classification and, in each case, to compare rule-based with randomly constituted classes. The domains selected in this experiment were the following:

Animals: nocturnal-diurnal, carnivorous-herbivorous, land-water, mammal-egg laying, and harmless-dangerous

Grammar: singular-plural, adjective-verb, third person-first person, marked complement-unmarked complement, conjunction-implication

Logic: past tense-progressive tense, affirmative-negative, stressed-unstressed, question-statement, transitive-intransitive

Names: initiation-birth, men's-women's, middle-first, first-last, Western-Kpelle

Sounds: unvoiced-voiced, fricative-stop, vowel-contrast, tone-contrast, consonant-contrast

Zoology: flying animals-flying insects, jumping-crawling, claw-dragging, insects-mammals, forest animals-town animals

With each of these domains two random classes were made up, containing eight items each selected in matched fashion from each of the other classes. By making up random and rule classes from the same domains, we can control for differences in familiarity and ease of learning between domains. Any differences in rates of learning can be attributed to the fact that all the items do or do not come from a single semantic or syntactic class.

It should be noted that only the classes listed under the heading "zoology" are drawn directly from the *seŋ* chart. The other classes were chosen on the basis of our analysis of the Kpelle language and our general curiosity about the range of rule-based learning. The learning of these kinds of classes provides a useful contrast with the learning of nouns from the *seŋ* chart.

One hundred and forty-four adult subjects participated in this experiment, each receiving six subproblems, including one random problem, in a completely counterbalanced order. Each subproblem contained eight items. At the top of Table 4–1 are listed each of the dependent

TABLE 4-1

Class-Based Learning for Semantic and Syntactic Classes in the Verbal-Discrimination Experiment

COMPARISON DOMAIN	TRIALS TO CRITERION	DIFFERENCES IN TRIALS TO CRITERION: RANDOM-RULE	SUCCESSFUL VERBALIZATIONS
Animals	20.8	11.6	11
Zoology	15.5	10.7	11
Names	16.1	10.8	17
Grammar	24.4	6.6	6
Sounds	23.0	5.9	11
Logic	23.6	5.5	9

variables, trials to criterion, a measure of the difference in trials to criterion between the random conditions and the rule conditions, and the number of successful verbalizations. The new measure, the difference between the randomly constituted classes and the rule classes, was arrived at by subtracting the trial number for the average trials to criterion for the rule class from the corresponding random class.

From the column labeled Trials to Criterion, it is clear that there are differences in difficulty associated with the different domains. Moreover, difficulty of learning seems to be paralleled by the results indicating the difficulty of verbalization, although verbalization is, as in many such experiments, very poor.

One way of conceiving of the difference among the domains studied is to view grammar, logic, and sounds as all involving some linguistic feature not explicitly named in the Kpelle language, and animals, names, and zoology as embodying different taxonomic classes nameable by the Kpelle. Based on this distinction, the data can be viewed as supporting the following generalizations: (1) nameable items are easier to learn than those nonnameable items that make up our various linguistic classes; (2) verbalization of nameable discriminations is easier than verbalization based on classes that are difficult to name; (3) learning of nonnameable discriminations is much closer to learning of the corresponding random discriminations.

One disturbing feature of these data is that learning was in general slow, compared to that observed in the earlier experiments. Perhaps such slow learning occurred because the discriminations are, in general, more difficult than those presented in previous studies. However, in line with our general philosophy, we are less concerned about the differences *between experiments* than we are about the relations *between conditions within experiments.* The important point is that for a certain group of classes (the nameable nouns), learning seems to be of a concept variety, whereas for another group (the linguistic class), rule-based classes are only slightly (although consistently) easier than random classes, implying that the subjects must, at least in part, be learning particular items rather than recognizing general concepts.

We did not contrast the learning rates of various subgroups among the Kpelle in this particular set of experiments. Some of the class distinctions we asked the subjects to make are matters of general knowledge, such as the distinction between materials used in building a house and clothing items. Others are known primarily to traditional specialists, such as the distinctions among various types of root plants and mushrooms. Still others are Western distinctions, not made by the Kpelle, such as the distinction between nocturnal and diurnal animals. Some distinctions are linguistic in character. And finally some distinctions are random.

John Kellemu sorted the entire set of class pairings into these five groups. After he had sorted the class pairs, we computed the mean trials to criterion for each group of distinctions for the second and third experiments. We found in the second experiment, that the successive means for traditional-familiar, traditional-specialist, *kwii,* linguistic, and random distinctions were 5.6, 7.4, 9.0, 10.2, and 13.3, respectively. All of these differences were statistically reliable except that between *kwii* and linguistic categories.

We conducted further experiments of this type, the results of which in general confirm what we have reported thus far. An interesting feature of one of these experiments was that it included different groups of subjects: Kpelle educated adults (high-school students), Kpelle schoolchildren (ten to fourteen years old, second to fourth grade), and nonliterate adults. In general, the results support the conclusion that education is positively related to the use of classification in learning. Over a large variety of problems educated adults, schoolchildren, and nonliterate adults learned in an average of 6.0, 7.0, and 9.3 trials, respectively.

One further fact to be noted in all these experiments is an overall tendency to improve in learning from the first to the subsequent problems. There is, on the average, a three-trial improvement between the first and second problems, while thereafter learning remains relatively constant. Educated subjects show a slight tendency to improve their learning beyond the second problem, whereas nonliterate subjects seem to confine all their "learning to learn" to the step from the first to the second problem. We interpret these results as indicating that the task requires some familiarity in order for subjects to do their best, a familiarity that is generally acquired by one performance of a problem.

In summary, where the materials to be learned are organized in nameable semantic categories, discriminations are aided by the presence of those semantic categories; that is, we are dealing with true concept learning and not rote learning of particular items.

Transfer of Classes

A related concept-discrimination experiment was conducted with two purposes in mind: (1) to determine the extent to which learning the distinction between two classes would facilitate learning to distinguish between other, closely related classes according to the *seŋ* chart; (2) to link the class-identification experiment to the free-association data from the previous chapter. The procedure differed slightly from that employed previously. The subject was instructed that he would be told the names of two things, one of which belonged with one of two chairs and

the other to the other chair. Pairs of names were presented to the subject until he made correct assignments of items to chairs on ten trials in a row. The subjects were 108 nonliterate adults.

After a subject had reached criterion on this first phase of the experiment, the experimenter immediately shifted to new classes, which were also to be identified with chairs. In some cases, these new classes were very close (according to the *seŋ* chart) to the old ones. For instance, one distinction to be learned was between *big persons* and *claw animals*. After reaching criterion on this verbal discrimination, the subject was required to distinguish *children* from *jumping animals*. In other cases the classes presented in the second phase of the experiment were more distant from the original ones. For example, some subjects were required to distinguish *town animals* from *evil things* after learning the distinction between *big persons* and *claw animals*. We hoped to show that transfer of the discriminations to closely related classes would be easier than transfer to distantly related classes.

In fact, we showed nothing of the kind. There were no significant differences between the learning of related classes and unrelated classes on the second phase of the experiment. Learning of a second distinction did not seem to be influenced by learning a closely related prior distinction. In fact, learning the second discrimination was in no detectable way different from learning of the first one.

A further modification of this experiment produced results that seem initially to contradict those described above. The first half of the modified experiment proceeded as before. In the second half, instead of being asked to solve a second, related, discrimination problem, the subjects were asked to name five things that could be assigned to each of the two chairs. Under these conditions, additional members of the same class were given in a very high proportion of the cases.

Their performance on the discrimination tasks shows that the subjects had clearly learned the distinctions between the classes, and previous evidence indicates that category membership influenced that learning. Yet, when asked to learn a closely related class, even though they were able to name other members of the class, they showed no transfer of knowledge from the first task to a closely related one.

One possible cause for the lack of transfer is suggested by the results of our free-association studies. In that work subjects persisted in a strategy of giving particular instances of a class rather than names of closely related classes. For example, the responses to "farm tools" were particular tools used in farming. Our earlier analysis of free-association responding leads us to suggest that the free-association responses in this concept-discrimination experiment produce little overlap between classes. This specificity and lack of overlap may well be the cause of the lack of positive transfer. It seems likely that if we had set up problems where the second-phase problem was made up of items selected from a subclass of the classes named in the first problem, rather than a class that seemed similar on the basis of the *seŋ* chart, positive transfer could have been obtained. Unfortunately, data are lacking on this point.

Verbal Discrimination: The Twenty Clusterable Objects

For any given concept-discrimination problem of the type described in the previous section, the subject is asked to deal with two classes at a time. A central question of interest is the degree to which category structure aids learning of the particular items designated correct.

In the present experiment the subject's task is somewhat more complex with respect to the categorical structure of the learning situation. However, the essential problem remains the same: will the presence of categorical structure in the materials to be learned facilitate learning?

We adapted for this purpose the standard paired-associate experiment which has been used for many years by psychologists interested in verbal-learning processes. The textbook model for paired-associate learning is the process by which beginning language students learn the vocabulary of a foreign language. For such students each term in English serves as a stimulus term and, a corresponding term in, say, French serves as the response. Thus, a student might prepare a card for each pair, on one side of which is written the English word, on the other side the French word. In the experimental study of this learning process, the stimuli are usually presented on slides or cards and the responses are either written or verbal. All manner of stimuli and responses have been used, for example, common nouns, numbers, letters, pictures, and nonsense syllables.

Our Kpelle adaptation of the paired-associate experiment was to use the twenty clusterable terms (tools, food, containers, and clothing), contained in Table 3–2 as stimulus terms and four chairs as response "categories." The subject's task was to learn which objects were assigned to which chairs.

This modified technique, in which the stimuli are in effect "classified" according to response "categories," represents a kind of halfway point between the concept-discrimination experiments of the previous section and completely unstructured learning tasks, a variety of which we will consider in the next section. In the American experimental literature on verbal learning, this kind of problem is called a verbal discrimination. We call it a paired-associate task simply to distinguish it from the experiments in the previous section.

By a suitable arrangement of the way in which the stimuli are assigned to chairs, we can determine whether or not semantic classification exerts control over the paired-associate learning process. In this particular experiment we manipulated the possibility of semantic control by having some conditions where five items were assigned to each chair in such a way that each group had at least one item from each of the four categories. Using this kind of comparison, we can reformulate the question of the role of semantic classes in learning. When objects are assigned to chairs on a semantic basis (rule condition) do subjects learn faster than when objects are assigned to chairs at random (random condition)?

We used this experiment not only to learn about the role of semantic classification in learning, but also to study the way in which feedback about the correctness of choice affects the rate of learning. This issue arose during pilot work for the experiment, when our informants indicated that it was very unnatural to learn item by item as the paired-associate technique demands. Rather, they thought that study periods followed by evaluation periods seemed more natural. We used this suggestion to develop a series of specialized ways of presenting material for learning. Also, because some of the techniques for comparison had no counterpart in the Western literature on paired-associate learning, we ran an experiment with comparable conditions in the United States to give us some basis for comparative judgments about the effects of various presentation schemes.

Our variations in the structure of the learning opportunities were designed to evaluate a series of hypotheses (perhaps intuitions would be a better word) about conditions that our Kpelle subjects would find most congenial. In some conditions the subject was an active participant on every trial. For example, he might have to guess which chair was paired with each object before he was told the correct pairing. In other conditions all the information about pairings was given to the subject, who was merely asked to repeat the name of the object before recall was tested. The details of these conditions are somewhat complex and are not described here because the results failed to justify the elaborateness of our procedures.

The subjects were 140 nonliterate Kpelle adults (most of them men) between the ages of eighteen and fifty years. All of the subjects lived in the general area of Cuttington College and most were traditional rice farmers who spoke little or no English.

Half of the subjects were run under the rule condition, where items were assigned to chairs on the basis of their category membership as shown in Table 3–2. The remainder were run in the random condition where the same items were assigned to chairs according to a set of four randomly chosen categories. Within each of these two main groups, there were seven subgroups consisting of ten subjects each. For present purposes it suffices to say that conditions where the experimenter presents the objects for learning followed by responses on the part of the subject proved easier than conditions where the subjects were required to guess the assignments of objects to chairs and then be corrected by the experimenter.

All subjects were presented the set of items fourteen times or until they were able to place them all on the chairs correctly on a single cycle through the set. Subjects were scored according to the number of errors on each trial.

In reviewing the results obtained from our Kpelle subjects, the most striking features of the data are that learning was relatively rapid and that the rule conditions were significantly superior to the random ones.

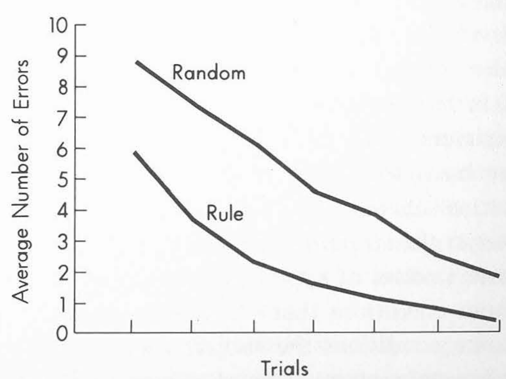

FIGURE 4–1 Average Number of Errors on Each of the First Seven Trials for Average Rule and Random Conditions. (Subjects who achieve an errorless test are credited with zero errors thereafter.)

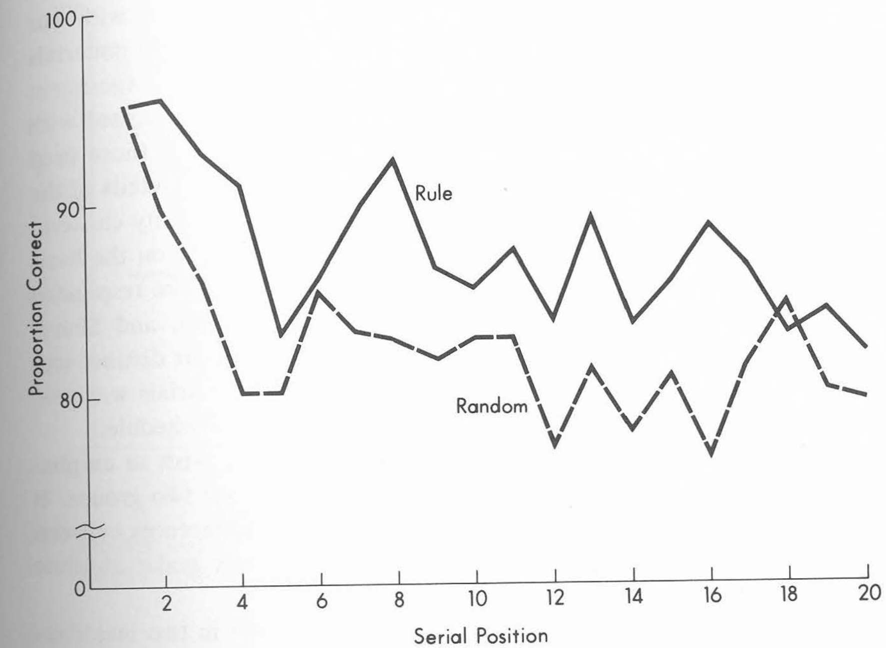

FIGURE 4–2 Proportion Correct as a Function of Serial Position for Kpelle Rule and Random Groups

The learning curves representing the relation between performance and the first seven learning trials are shown in Figure 4–1. In addition to the overall superiority of the rule conditions, it should be noted that the rule subjects were correct on fourteen out of twenty items on Trial 1, suggesting that they had learned the rule and were applying it from the outset. Statistical analysis indicated that superiority of the rule over the random procedure was constant across the various presentation conditions.

A detailed study of the results reveals an interesting relation between response accuracy and the serial position of the object-chair pairing during presentation. Figure 4–2 shows this serial-position relationship, beginning on the left-hand side of the figure with the first item in each list. It is important to remember that list order was changed from trial to trial, so that Figure 4–2 plots the first and subsequent ordinal items rather than a specific series of items. From Figure 4–2 it is clear that performance is best for early items on the list and that the decline as a function of serial position is monotonic (this is referred to as a "primacy" effect in the Western literature on verbal learning).

In order to determine if the pattern of results obtained with our Kpelle subjects is in some way peculiar to the population and materials we were working with, we undertook a replication of the experiment with a group of 140 American sixth graders. The procedures used with these American subjects were designed to be analogous to those used with the Kpelle. The following changes were made in the details of the experiment: (1) the subjects were shown pictures of the twenty clusterable objects from Table 3–2 (this change seems immaterial on the basis of our research which indicated that pictures and objects are responded to similarly by American schoolchildren [Cole, Frankel, and Sharp, 1971]); (2) a specially constructed board divided into four distinct sections replaced the four chairs; (3) a maximum of seven trials was presented because of time restrictions imposed by the school schedule.

Before describing the results of this replication, we wish to emphasize our interest in the pattern of results observed in the two groups. It is clear that the procedural changes and the multiple differences between nonliterate Kpelle adults and American sixth graders make absolute comparisons of performance in many ways specious.

With these provisos in mind, it can be reported that in two major respects the American and Kpelle data are markedly similar. First, for the American subjects the rule condition (2.8 errors per trial) was far easier to learn than the random condition (6.8 errors per trial). Second, the American subjects showed the same order of difficulty for the various presentation conditions as the Kpelle. In particular, the American subjects also found it more difficult to learn if they had to begin the learning sequence by guessing the correct response slot, instead of being shown the correct alternative and asked to recall it on a later cycle of the experiment. Similar findings reported by W. K. Estes (1969) suggest that this result is to be expected in a larger range of situations.

We chose the various presentation conditions in the first place because of our informant's intuitions about special problems that the Kpelle were likely to experience with the paired-associate task. Thus the similarity of the response patterns in this regard between Kpelle and American subjects is strong evidence of a significant *similarity* in the learning processes of the two widely divergent groups.

In one respect, however, the American data are unlike the Kpelle. As shown in Figure 4–3 there is no relation between serial position and accuracy for the rule condition among the American schoolchildren. For the random condition we observe both the primacy effect noted in Figure 4–2 and a corresponding recency effect in which American sub-

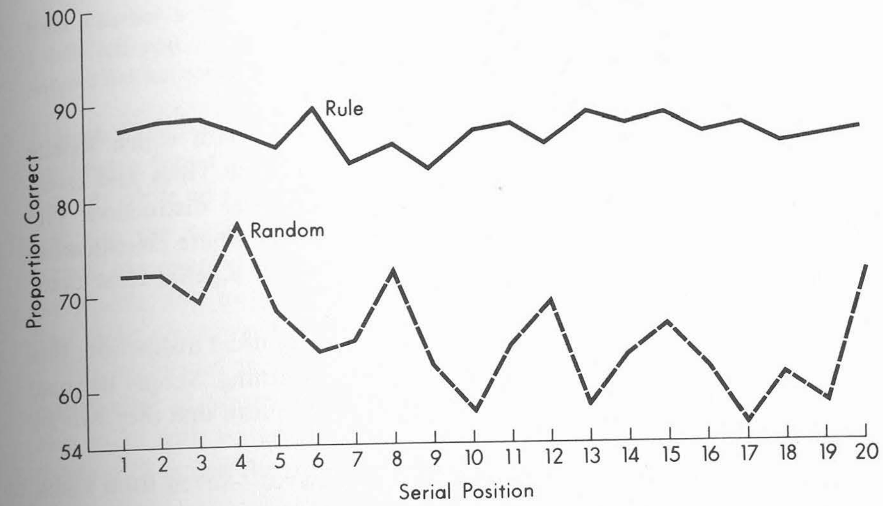

FIGURE 4–3 Proportion Correct as a Function of Serial Position for American Rule and Random Groups

jects are more accurate for the last item presented. Figure 4–3 also dramatizes the large difference between rule and random conditions for the American subjects, leading us to surmise that the lack of a relation between serial position and accuracy in the rule condition occurs because the semantic categories are so strong that their use in organizing learning completely swamps the factors that lead to serial-position differences.

Taken together, the American and Kpelle results of this experiment point to one important difference and one important similarity between the learning mechanisms of the two cultures. Both groups find it easier to learn when they are given sufficient information about stimulus-response relationships so that they can do more than guess on early trial. Both groups use the semantic structure of the list to aid learning, but the Americans make relatively more use of semantic information than do the Kpelle.

It is unfortunate that we do not have more data on this point from other populations of Kpelle people. We can only surmise on the basis of our previous results that the rule-random difference for children with a good deal of schooling (seven grades or more) would look more like the performance of the American children, while the preference for the full presentation of information would remain unchanged.

Category Discrimination: Leaves

In Chapter 3 we presented evidence on the ease with which Kpelle adults sort leaves into piles that distinguish between vines and trees, whereas American subjects fail to make the vine-tree distinction. The outcome, in that context, indicated that in situations where classification activities with common objects were familiar to the Kpelle, clear taxonomic classes emerged.

We can again ask under what conditions a categorical distinction, this time between tree and vine leaves, will control learning. Simply because the categories are applicable to learning does not mean that they will be applied.

Our materials for this study were fourteen leaves—seven from vines, seven from trees. These included the twelve leaves used in the experiment on sorting plus one additional leaf from each class. The subjects were thirty American and Canadian college students who were working in the Cuttington College vicinity and thirty nonliterate Kpelle adults ranging in age from twenty- to thirty-one years. Each group of thirty was divided haphazardly into groups of ten who were run under one of three different conditions, each designed to assess a different aspect of the relation between categories and learning. The three conditions are well described by the instructions read to the subjects who were seated at a table across from the experimenter.

Condition 1: (All the tree leaves are assigned to one name, the vine leaves to the other.) I have several leaves here. Some belong to Sumo and the others belong to Togba. I will hold up a leaf at a time, and tell you whom it belongs to. Then I will ask you to tell me as I hold up each leaf whether it belongs to Sumo or Togba. I will tell you each time whether you are right or wrong (continue until the subject has correctly named all the leaves in one trial).

Condition 2: I have here several leaves. Some are from trees and some from vines. I will tell you which leaf is from vine and which from tree. Then I will hold up each leaf after this, and I want you to tell me whether it is from tree or vine. I will tell you each time whether you are right or wrong (again continue until the subject has named correctly all the leaves in one trial).

Condition 3: (Three vine and four tree leaves for Togba and four vine and three tree leaves for Sumo.) I have several leaves here. Some belong to Sumo and some to Togba. I will tell you which belong to Sumo and which to Togba as I hold up each leaf. Then I would like for you to tell one at a time whether a leaf belongs to Sumo or not when I hold it up to you again. I will tell you whether or not you are right or wrong (continue until the subject has correctly named all the leaves in one trial).

Both Conditions 1 and 2 permit the subject to take advantage of the categories latent in the items, but in Condition 2 the categories are named, while in Condition 1 the subject has to discover them for himself. On the basis of the results from our experiment in sorting the leaves, discovering the classes should not be expected to be difficult for the Kpelle, but to be quite difficult for the Americans and Canadians. The third condition systematically violates the categories of the items.

The results of this study are summarized in Table 4–2 in terms of the number of presentations of the list required for one completely correct cycle with no errors. The really interesting point to emerge from Table 4–2 is that the tree-vine leaves distinction is helpful to the Kpelle *only* if the instructions make clear that it is this distinction that the experimenter has in mind. By contrast, nothing proves of much use to the Americans. In spite of the slight indication that Americans were able to distinguish tree and vine leaves in our sorting experiment, this sample of subjects proved unable to make use of such distinctions in the context of a learning study. There seems also to be a factor relating to discriminability of the leaves; the Kpelle learn slightly faster than the Americans even under conditions where the tree-vine distinction does not seem to influence performance.

TABLE 4-2

Number of Complete Presentations of the Leaves Required to Make One Completely Correct Recall of the Set

	CONDITION		
	1 SUMO-TOGBA RULE	2 TREES-VINES RULE	3 SUMO-TOGBA RANDOM
Kpelle	7.3	1.1	6.8
American	9.8	8.9	9.0

In our view the problems involved for the American and Canadian subjects are captured beautifully in an anecdote related by Eleanor Bowen (1954) in her book, *Return to Laughter*.

By nine o'clock that morning, I had several pages of words, and my tongue was limp from unaccustomed twisting. Unable to take in any more, I instituted a review by again naming the notables. I again got most of them right: the right man and almost the right sound. Kako looked at me with favor. Encouraged, I demanded the names of the women. They smiled, but Kako ignored my question and turned firmly back to the leaves. Rather reluctantly I began to name them. With every word Kako became more dour. I spoke more loudly; my pronunciation couldn't be that bad. Ikpoom's eyes grew sadder; the women seemed incredulous. The little boy could bear it no longer. He snatched from me the leaf I was naming and handed me another. The order had been mixed, and not once had I put the right name to the right plant.

These people are farmers: to them plants are as important and familiar as people. I'd never been on a farm and am not even sure which are begonia, dahlias or petunias. Plants, like algebra, have a habit of looking alike and being different, or looking different and being alike; consequently mathematics and botany confused me. For the first time in my life I found myself in a community where ten year-old children weren't my mathematical superiors. I also found myself in a place where every plant, wild or cultivated, had a name and a use, and where every man, woman, and child knew literally hundreds of plants. None of them could ever believe that I could not if I only would. [Bowen, 1954, pp. 15–16]

Miss Bowen is making the point that her hosts had learned to make the distinctions upon which rapid learning is based so that even in classes where they might not know the particular plant name, it is easily learned, while she could not even distinguish between the plants.

Such factors help to explain the results for the Canadians and Americans and, in fact, seem obvious. But what about the difficulty of the Kpelle subjects in "Sumo-Togba Rule" condition? Why weren't they aided by the categorical split in the response terms? One possibility is that they failed to notice that tree leaves or vine leaves could form the basis for learning; a second hypothesis would be that such labeling actually misled the Kpelle subjects. After all, why should Sumo have all the vines and Togba all the trees?

This is another instance where considerably more research is required. Looking ahead we can guess that our Kpelle subjects in fact failed to note that vine and tree could serve as tools for easier learning.

Learning and Memory: Free Recall

In the previous sections of this chapter, we dealt with a series of situations in which subjects were required to learn to distinguish taxonomic classes, or make use of particular classification schemes selected by the experimenter in learning tasks. The emphasis in our description of these studies was on *learning*—learning the class discrimination or learning the object-pair relation. But we could just as easily have described the processes involved in terms of the concept of *memory:* what conditions affect a subject's ability to remember which item from the pair was correct previously? How well does the subject remember the stimulus-response relationship? In fact, each presentation of something to be learned is evaluated in terms of what is remembered later. Little more than a choice of terms determines whether we talk about performance in terms of what was learned at the time the items were presented or remembered at the time that they were tested.

Nevertheless, it does seem that when we discuss memory in its cultural context, we probably are not thinking about the kinds of situations studied in this section. In fact, a small body of literature has grown up around the question of cultural differences in the processes of memory. The questions raised in this literature will serve as an introduction to a set of experiments designed to study the relation between various cultural factors and memory.

Introduction to the Problem

Early observers of nonliterate societies, in commenting on differences in intelligence and logical capacity, reported the existence of excellent, and in some cases extraordinary, mnemonic skills. Levy-Bruhl, for example, cites many examples of this apparent ability to memorize, claiming that "in every case in which their memorizing power, which is really excellent, could relieve them of the effort of thinking and reasoning, they did not fail to make use of it" (Levy-Bruhl, 1966, p. 25). Similar anecdotes and assertions are often reported by Westerners who have taught in Africa, the most usual form being that African students do very well with material that can be "learned by rote," but become poor or indifferent students when expected to do tasks in which brute memorization will not work.