Serial position and clustering effects in a chimpanzee's "free recall"

JAMES P. BUCHANAN

University of Scranton, Scranton, Pennsylvania 18510

TIMOTHY V. GILL

Yerkes Regional Primate Research Center, Emory University, Atlanta, Georgia 30322

and

JOHN T. BRAGGIO

University of North Carolina, Asheville, North Carolina 28801

Lana, a chimpanzee sophisticated in the language Yerkish, was tested for free recall on lists consisting of from one to eight words randomly drawn from one of three taxonomic categories or on lists consisting of nine words with every third word from a different category. Serial position effects were observed for the four- to eight-item lists, with statistically significant first-item primacy effects on the seven- and eight-word lists and last-position recency effects on the six-, seven-, and eight-word lists. Also, above-chance clustering was observed in the recall of the nine-item categorized lists. These results were compared and contrasted with the serial position and clustering effects obtained in free recall with humans. Several alternative theoretical accounts of the results were considered.

This study was designed to examine the free recall capabilities of Lana, a chimpanzee sophisticated in the language Yerkish (Rumbaugh, 1977). Lana was tested to determine if her episodic free recall performance would demonstrate the well established serial position (Murdock, 1962) and category clustering (Bousfield, Cohen, & Whitmarsh, 1958) effects that are obtained with humans. Serial position effects refer to the finding that when adult humans are given a list of, say, 20 words presented sequentially, their recall will be better both for words occupying the first few positions in the list (primacy effect) and for words occupying the last few positions in the list (recency effect) than for words occupying the middle positions of the list (Murdock, 1962). Category clustering refers to the finding that when adult humans are given a sequentially presented list of words that can be grouped together into a number of different taxonomic categories, they will tend to organize or cluster their recall by the categories even when the presentation of the list was not sequenced by category (Bousfield et al., 1958).

Serial position effects have previously been investigated with infrahuman subjects in studies employing serial probe recognition tasks (SPR) (Sands & Wright,

1980a, 1980b; R. Thompson & Herman, 1977). R. Thompson and Herman (1977) presented tone lists of one to six items to a bottle-nosed dolphin. After hearing the tone list, the dolphin had to classify a subsequent tone as either old (in the list) or new. The probability of recognizing the old sound was highest for the most recent sound presented and dropped off sigmoidally for successively earlier sounds. The dolphin's performance, then, indicated a significant recency effect without any corresponding primacy effect. On the other hand, Sands and Wright (1980a, 1980b) report both a primacy effect and a recency effect on an SPR task with a rhesus monkey. Both the monkey and a human were shown series of 10 or 20 colored pictures drawn from a set of 211 pictures. After seeing the picture lists, the subjects were shown a single picture and were required to make a "same" (in previous set) or "different" judgment. The human's and the monkey's serial position curves were similar in shape; however, the human's recognition performance exceeded the monkey's recognition performance. The monkey's primacy effect was significant only at the first serial position on both list lengths. Also, the monkey's recency effect was significant only at the last serial position on the 10-item lists. On the other hand, the monkey's recency effect was significant at the last four serial positions on the 20-item lists.

Category clustering effects have also previously been investigated with a recognition task (Premack, 1976). In this study, the chimpanzee Sarah was randomly presented objects from three conceptual categories

This research was supported by NIH Grant HD-06016, awarded to Duane M. Rumbaugh, and NIH Grant RR-00165, awarded to the Yerkes Regional Primate Research Center of Emory University, to whom the authors wish to express their gratitude for their support to the study. Requests for reprints should be sent to James P. Buchanan, Department of Psychology, University of Scranton, Scranton, Pennsylvania 18510.

(fruits, candy, and doll's clothes). Then she was required to pick out those objects from a random display of those items and other foils from the same conceptual categories. Premack (1976) reports that Sarah's order of selection of objects revealed some partial clustering along the conceptual categories.

The current study expands in two ways on the previously mentioned studies on infrahuman serial position and clustering effects. First, the current study introduces a free recall procedure by which serial position and clustering effects can be studied in an infrahuman subject. And second, the study employs a more abstract type of list item (Yerkish symbols) than has previously been employed with infrahumans.

METHOD

Subject

The subject was Lana, an 8-year-old chimpanzee (*Pan troglodytes*) verbally sophisticated in Yerkish (Rumbaugh, 1977).

Stimuli

The "word" lists were drawn from three taxonomic categories of 15 food-drink names (apple, banana, beancake, bread, cabbage, carrot, chow, M&Ms, orange, sweet potato, coffee, Coke, juice, milk, and water), 12 object names (ball, bowl, box, can, cup, door, feces, gate, money, room, shoe, and window), and 8 color names (black, blue, green, orange, purple, red, white, and yellow). These words constituted about 30%-35% of Lana's active vocabulary at the time of the study.

Each of the 240 lists for measuring serial position effects was constructed by separately and randomly drawing the items in the list exclusively from one of the three semantic categories. In all, 30 different lists were drawn up for each of the eight list lengths of from one to eight items. Color names were used for all 30 one-word and four-word lists; food-drink names and object names were each used for 15 of the 30 two- to three-word lists and five- to eight-word lists.

The 30 different nine-word category clustering lists were drawn from nine food names (apple, banana, beancake, bread, carrot, chow, M&Ms, orange, and sweet potato), five drink names (coffee, Coke, juice, milk, and water), four life-space object names (door, gate, windows, and room), seven small object names (ball, bowl, box, can, cup, money, and shoe), and eight color names (black, blue, green, orange, purple, red, white, and yellow). Each of the 30 nine-word lists was constructed by drawing three words each from three of the word groups. The 30 nine-item lists were drawn from food, life space, and color names on the odd-numbered lists and from drinks, small objects, and color names on the even-numbered lists. The sequence of words in a list was organized so that the words from a given category appeared every third word. Also, the order of categories was varied across the lists.

With respect to the category clustering lists, it should be noted that all the food-drink symbols had a red background, all the life-space and small-object symbols had an orange background, and all the color symbols had a black background. Consequently, the words could be organized or clustered either on a semantic basis or on a totally confounded perceptual basis. This, of course, provides a problem in interpreting the basis of any clustering found in Lana's recall protocols. It should also be noted that the current form of Yerkish has eliminated this confound.

Procedure

Figure 1 presents Lana at her keyboard. The keyboard consists of two 7 x 8 in. matrices of Yerkish symbols. At the time of the study, 90%-95% of the matrix positions held active keys. Consequently, this study made use of only about one-third

of Lana's total active vocabulary. Keys for words from the same semantic category (in the case of this study, food and drink, object, and color names) are not spatially located together but are instead haphazardly scattered across both matrices. Consequently, Lana could not gain any physical advantage, such as minimizing recall time or physical effort, by sequencing her recall by category.

Each word list was presented sequentially to Lana, with only one word appearing at a time for a 2- to 3-sec duration on her keyboard. As Figure 1 indicates, the display board is located directly above the keyboard and has space for projecting 10 Yerkish symbols at a time. In this study, however, only the far left position was employed to present the symbols to Lana, with each successive word in a list replacing the previous word on the display board.

Lana was trained to watch her display board and to begin her recall after the last word had been deleted from her board and the experimenter (T. Gill) had pointed to her keyboard. Lana then recalled as many words as she could and then indicated the end of her recall efforts by pushing her "period key." Lana was pretrained to this procedure in the following manner. Initially, the experimenter placed single words on her display board for 2-3 sec. At the termination of the word on the display board, the experimenter pointed to Lana's keyboard. Lana was rewarded when she duplicated with her keyboard the word that had just previously been on her display board. Lana quickly learned to duplicate single words with her keyboard. The speed at which she learned probably reflects the readiness with which chimpanzees imitate. Once she was reliably duplicating single words, she was then shaped with the outline procedure to duplicating two- and then three-word lists, she was started on the serial position phase of the study.

During the study, Lana was reinforced after each recall trial regardless of her recall pattern. Throughout the study, Lana was free to adopt any recall strategy (e.g., clustering or sequential recall). She was rewarded simply for the act of engaging in free recall. At no time during the study was she provided any feedback. The word lists were presented over a series of days, with the pacing dependent on Lana's interest in the recall task. Lana indicated disinterest in the task by removing herself from the keyboard and engaging in some other activity. On a given day, the time between trials ranged from 15-30 sec to 45-60 sec. The time interval between was filled with reinforcement activities.

After the pretraining period, Lana was tested first on the serial position task, followed by the clustering task. During the serial position phase of the experiment, Lana was tested once on each of the 240 serial position lists. She was tested on 30 trials each for the one- to eight-item lists, in the following order. Lana was first given 20 replications of a six-trial sequence. In each six-trial sequence, Lana was tested on a one-item list on the first trial, a two-item list on the second trial, a three-item list on the third trial, a four-item list on the fourth trial, a five-item list on the fifth trial, and a six-item list on the sixth trial. After receiving 20 trials each on the one- to six-item lists, Lana was tested on 20 trials with seven-item lists. Next she was given 10 more replications of the six-trial sequence followed by 10 more trials with seven-item lists. Finally, she was tested on 30 trials with the eight-item lists. During the entire series of trials, each successive word list was from a different semantic category in order to reduce proactive interference from one trial to another. During the clustering task, Lana was given two replications of the 30 nine-item lists. All of the 30 lists were presented once before the second replication began. The word order of the lists was reversed during the second replication.

RESULTS

The standard statistical procedures for analyzing serial position effects and clustering involve the use of one-way ANOVAs and t tests. However, in single-case

Figure 1. Lana at her response keyboard.

designs, successive observations in a time series are likely to be correlated. As Kazdin (1975) has pointed out, any such serial correlations or serial dependence in the data obtained in a single-case design would violate the basic ANOVA assumption of independence of the error term and, therefore, make an F test or a t test inappropriate.

Consequently, serial correlation tests of randomness (Bennett & Franklin, 1961) were run on the data from the serial position and clustering lists to determine if the assumption of independence of error term was met or violated. Serial correlations test for randomness by correlating the data with themselves a set distance or lag apart. A significant correlation indicates serial dependency and nonrandomness in the data and, consequently, indicates that the independence-of-error-term assumption has been violated.

Serial Position Effects

Lana's recall was perfect for the one-item list and ranged from 90% to 93% across the serial positions on the two- and three-item lists. Figure 2 presents, for

the four- to eight-item lists, the percentage of items correctly recalled at each serial position and the average percentage of recall across serial positions for the four- to eight-item lists. Lana's recall on the four- to eight-item lists was clearly higher for items presented at the beginning and end of the list than for items presented in the middle of the list. Clear primacy and recency effects were observed at the first and last positions, respectively, for the four- to eight-item lists. Also, a primacy effect was present at the second position on the four- and five-item lists but not on the longer lists, on which recall fell greatly from the first to the second position. A recency effect also seemed to be present at the next-to-last position on the five- through eight-item lists. Averaged across the four- to eight-item lists, Lana recalled 85% of the first serial position items, 67% of the second serial position items, 95% of the last serial position items, 76% of the next to last serial position items, and 61% of the remaining intermediate serial position items.

Serial correlations of a lag of one and of a lag equal to the list length were calculated for each list. A lag of

Figure 2. Serial position by percentage of recall for four- to eight-word lists.

one tests for serial dependency between successive serial positions within and then across trials; a lag equal to the list length tests for serial dependency at the same serial position on successive trials. These particular lags were chosen because they would be the two lags at which serial dependency would logically be expected to be the greatest.

Both tests of serial dependence proved to be nonsignificant for the six-, seven-, and eight-item lists, with a range of $R_h = +.01$ to $-.11$. Consequently, for the three longest lists, the assumption of independence of error term was met. Under these circumstances, Kazdin (1975) has suggested that ANOVAs and *t* tests are appropriate.

Following procedures outlined by Gentile, Roden, and Klein (1972), one-way ANOVAs were run for the six-, seven-, and eight-word lists. The main effect for serial position was significant for the six-, seven-, and eight-word lists [respectively: $F(5,174) = 2.95$, $p < .05$; $F(6,203) = 6.53$, $p < .01$; and $F(7,232) = 4.00$, $p < .01$].

The presence of recency and primacy effects in the data for these two list lengths was evaluated in two different manners. First, since a serial position curve, especially for short lists, tends to approximate a quadratic curve, quadratic trend analyses were carried out for the six-, seven-, and eight-word lists. The six-, seven-, and eight-word lists all proved to have significant quadratic trends present [respectively: $F(1,174) = 8.42$, $p < .01$; $F(1,203) = 22.88$, $p < .01$; and $F(1,232) = 23.43$, $p < .01$]. Consequently, all three list lengths generated the standard U-shaped curve associated with recency and primacy serial position effects.

The second analysis carried out to test for recency and primacy effects on the three longest lists was done

with multiple *t* tests comparing the first, second, last, and next-to-last serial positions with the remaining combined intermediate serial positions (e.g., for the eight-word lists, Serial Positions 1, 2, 7, and 8 were individually compared with Serial Positions 3, 4, 5, and 6 combined).

These multiple *t* tests revealed that recall was significantly higher for first-position items than for the combined intermediate-position items on the seven- ($p < .01$) and eight- ($p < .01$) item lists, indicating a significant primacy effect for the two longest lists. Likewise, recall was significantly higher for the last serial position items than for the combined intermediate-position items for the six- ($p < .05$), seven- ($p < .01$), and eight- ($p < .01$) item lists, indicating a significant recency effect at the last serial position on the three longest lists. Finally, neither the second nor the next-to-last serial positions proved to be significant, indicating that the primacy and recency effects were limited, respectively, to the first and last word presented in each list.

Table 1 presents an analysis of Lana's order of recall by serial position for the last-, next-to-last, and first serial positions collapsed across the six-, seven-, and eight-word lists. A standard finding with adult humans is that they tend to sequence their recall by recalling the items from the end of the list first and recalling the words at the beginning of the list in the middle or later portion of their recall (Waugh, 1961). Goodwin (1976), however, has shown that the human tendency to begin recall with the recency items is a function of practice. Goodwin found that subjects initially tend to begin recall with primacy items, but with a few trials of

Table 1
Order of Recall by Serial Position in Percentages, Collapsed
Across Six-, Seven-, and Eight-Word Lists

Input Serial Position	Recall Position					
	1	2	3	4	5	6
Last	90.7	4.7	1.2	.0	2.3	1.2
Next to Last	5.8	55.1	20.2	15.9	.0	2.9
First	5.0	18.8	30.0	25.0	18.8	1.2

practice they shift to an increasing tendency to begin recall with recency items.

An inspection of Table 1 reveals that Lana ordered her recall in a manner similar to that of practiced adult humans. On the trials that Lana recalled the first and/or last item, she first recalled the last item 90.7% of the time, whereas she first recalled the first item only 5% of the time. On the other hand, the first item was recalled by Lana in the third or later output position 76.2% of the time, whereas the last item was recalled in the third or later output position only 4.7% of the time. Also, Lana most often tended to recall the next-to-last item in the second output position (55.1%).

Clustering

The clustering score employed was the adjusted clustering ratio, or ARC (Roenker, Thompson, & Brown, 1971). This measure, like most clustering measures, is based on repetitions that are defined as adjacent occurrences of two words from the same experimenter-defined categories. The ARC measure indicates the ratio of actual repetitions to the total possible repetitions corrected for chance occurrences of repetitions. Zero indicates chance clustering and one equals perfect clustering. Negative values are possible, but they are uninterpretable, since the ARC is not symmetric around the zero point. Following Pellegrino's (1975) suggestion that only above-chance clustering is theoretically relevant, one analysis of Lana's clustering was done with negative ARC scores set to zero. Also, a second analysis of Lana's clustering was done with negative ARC scores not set to zero, to provide a maximally conservative test of clustering.

Table 2 presents the means and standard deviations for Lana's recall and clustering scores (with and without negative ARC scores set to zero). Since the same lists were employed for the first and second blocks of 30 trials, it is reasonable to compare Lana's performance

Table 2
Means and Standard Deviations: Clustering and Recall Scores

Variable	Trials 1-30		Trials 31-60		Trials 1-60	
	Mean	SD	Mean	SD	Mean	SD
ARC ₁	.19	.65	.41	.58	.30	.62
ARC ₂	.38	.41	.50	.44	.44	.43
Recall	5.37	.81	5.77	.94	5.57	.89

Note—ARC₁ = negative scores not set to zero; ARC₂ = negative scores set to zero.

across these two blocks, especially to look for increased clustering indicating a practice effect. Consequently, the recall and clustering scores in Table 2 are broken down into the first and second blocks of 30 trials, as well as with both blocks combined.

An inspection of Table 2 reveals that both Lana's recall and her clustering increased from the first to the second block of 30 trials. Serial correlations of a lag of one were run on the ARC scores, and recall and ARC scores for each block of 30 trials and for both blocks combined proved all to be nonsignificant, with a range from $R_h = +.01$ to $R_h = .195$. Consequently, the independence of error term was met, indicating that *t* tests could be carried out on the data (Kazdin, 1975).

In order to determine if Lana's clustering was significantly above chance (i.e., ARC = 0), single mean *t* tests were run on the first and second blocks and on both blocks combined. When negative ARC scores were not set to zero, the *t* test was not significant for the first block [mean = .19; $t(29) = 1.60$, $p < .10$]; on the other hand, the *t* tests were significant for the second block [mean = .41; $t(29) = 3.93$, $p < .001$] and for both blocks combined [mean = .30; $t(59) = 3.76$, $p < .001$]. When negative ARC scores were set to zero, the *t* tests proved to be significant for the first block [mean = .38; $t(29) = 4.95$, $p < .001$], for the second block [mean = .50; $t(29) = 6.26$, $p < .001$], and for both blocks combined [mean = .44; $t(59) = 8.02$, $p < .001$]. Consequently, Lana, when all trials are considered, clearly exhibited significantly above-chance clustering.

Dependent *t* tests were also run comparing Lana's increase in clustering and recall from the first block to the second block of 30 trials. Lana's increase in clustering on the first to the second block of 30 trials proved not to be significant both when negative ARC scores were not set to zero [mean = .22; $t = 1.33$, $p < .10$] and when negative ARC scores were set to zero [mean = .13, $t(29) = 1.13$, $p < .20$]. On the other hand, Lana's higher recall on Block 2 than on Block 1 did prove to be significant [mean = .40; $t(29) = 1.71$, $p < .05$].

Table 3 presents a comparison between the first and second blocks of 30 trials for the number of observed repetitions, the maximal number of repetitions possible, and recall broken down into the five categories of food names, drink names, small-object names, life-space object names, and color names. An inspection of Table 3 reveals that for the food, drink, and color categories, both the total observed and possible repetitions counts and the total recall count remained remarkably stable between the first and second block of 30 trials. On the other hand, for the two object categories, the total number of observed and possible repetitions increased by a total of 10, and recall increased by a total of 11 from the first to the second block of trials. Consequently, both Lana's increase in clustering and her recall from the first to the second block of trials were primarily the result of her greater tendency in Block 2 to cluster and recall object words. Also, a further inspection

Table 3
Observed and Possible Repetitions and Recall by Category Across Block 1 and Block 2

Category	Observed Repetitions		Possible Repetitions		Recall	
	Block 1	Block 2	Block 1	Block 2	Block 1	Block 2
Foods	13	15	19	20	34	35
Drinks	15	15	20	20	35	34
Small Objects	5	10	13	13	25	27
Life-Space Objects	5	10	7	17	20	31
Color	7	9	19	19	45	46

Note—Each block = 30 trials.

tion of Table 3 reveals that Lana showed equal amounts of clustering when the food and drink categories are compared. Likewise, Lana also showed equal amounts of clustering when the small-object and life-space object categories are compared.

A further inspection of Table 3 reveals that across blocks Lana showed the most recall (138) and the greatest proportion of clustering (.73) for food-drink words, an intermediate amount of recall (103) and clustering (.60) for object words, and the least recall (91) and proportion of clustering (.42) for color words. It should be noted that these differences may reflect Lana's sensitivity to either the semantic or the totally confounded perceptual background categorical relationships for these words.

Intrusion Errors

A final important aspect of Lana's performance is her intrusion errors. Lana made only 33 intrusion errors on the 240 serial position trials and only 7 intrusion errors on the 60 category clustering trials. Also, she never made more than two intrusion errors in any single trial. Consequently, Lana's minimal number of intrusion errors suggests that she was fairly carefully searching and retrieving items from her memory and that her recall performance was not based on guessing.

DISCUSSION

The results of this study clearly indicate that Lana was capable of storing and retrieving Yerkish symbols from memory. Also, the presence of primacy and recency effects and above-chance clustering in her episodic free recall suggests that, at the surface level, Lana's recall performance was similar to what would be expected from a human subject. However, quite clearly, many qualifying statements are in order regarding the exact nature of the similarities and dissimilarities between Lana's recall performance and that typical for humans. This is especially the case with regard to the mechanisms that may have mediated the primacy and recency effects and clustering in Lana's free recall.

Also, before any direct comparisons can be made between Lana's recall performance and typical human

recall performance on serial position and clustering tasks, some discussion is required in terms of the differences between the free recall procedures employed with Lana and those procedures typically employed with humans. The single-subject methodology required that Lana be given a series of free recall trials, whereas the most common procedure with humans is to employ a single recall trial with a group of subjects (e.g., Bousfield et al., 1958; Murdock, 1962). However, both serial position effects (Craik, 1970) and clustering effects (C. Thompson, Hamlin, & Roenker, 1972) have been observed in humans when subjects are given a series of recall trials over different lists. Also, the limited size of Lana's vocabulary required that her recall lists be much shorter than the lists typically employed with adult humans (e.g., Bousfield et al., 1958; Murdock, 1962). However, lists similar in length to Lana's have been employed in developmental studies with human adults and children. These developmental studies with lists of less than 10 items have clearly demonstrated serial position and clustering effects (Hagen, Jongeward, & Krail, 1975). Lana's limited vocabulary also required that the items in a given serial position list be drawn from the same semantic category. Humans, on the other hand, are typically presented semantically unrelated word lists in serial position studies. However, again, serial position effects have been reported for humans with short lists drawn from a small pool of semantically related words (Keeley, 1971; Phillips, Shiffrin, & Atkinson, 1967). Finally, Lana's limited vocabulary also makes the taxonomic categories of words employed in the category clustering lists relatively exhaustive in nature. That is, the items in any given list represent a large portion of the items in a given category. Cohen (1966) has shown that both clustering and recall are higher in humans with exhaustive as opposed to non-exhaustive categories.

The standard information processing explanation of serial position effects revolves around the different properties of short-term memory (STM) and long-term memory (LTM) and the operation of the rehearsal buffer as it maintains information in STM and acts to build up a permanent memory trace in LTM. STM is considered to be a transient memory whose capacity in adult humans is limited by the capacity of the rehearsal

buffer to 7 ± 2 chunks, whereas LTM is considered to be a permanent memory of extremely large capacity.

At the beginning of a list, the rehearsal buffer in STM is assumed to be empty or nearly empty. As the successive items are presented, they are entered into the STM rehearsal buffer until it is full. Thereafter, as each new item enters STM, it is at the cost of replacing an old item already present in the rehearsal buffer. During free recall, the items in STM are recalled first. Since these items are most likely the last few items presented, this immediate recall from STM gives rise to the recency effect. All the remaining items not in STM then have to be recalled from LTM. Because the items presented at the beginning of the list enter an empty or partially empty rehearsal buffer, they are given more cumulative rehearsal time than later items, which enter a full rehearsal buffer. This extra rehearsal of the first few items increases the probability that they will be stored and hence retrievable from LTM, thus giving rise to the primacy effect (Atkinson & Shiffrin, 1971).

The largest discrepancy between Lana's free recall performance and that typical for adult humans comes at the primacy position. Adult humans generally show primacy effects for the first three or so positions on lists up to 40 items long (Murdock, 1962), whereas Lana showed a primacy effect only at the first position. Also, the primacy effect in adult humans is generally attributed to cumulative rehearsal. This attribution is supported by the finding that primacy effects decrease as presentation rates increase (Glanzer & Cunitz, 1966) and by monitoring humans' overt rehearsal during list presentation (Rundus & Atkinson, 1970). In Lana's case, however, there is no direct evidence that her primacy effect was due to cumulative rehearsal or any type of rehearsal strategy. Consequently, if Lana was cumulatively rehearsing, it was only for the first word in each list.

A potential alternative explanation for Lana's first-position-only primacy effect comes from a study by Fischler, Rundus, and Atkinson (1970). In one condition in this study, they presumably blocked cumulative rehearsal strategies by requiring their subjects to rehearse overtly at a constant rate only the currently presented item. Even with cumulative rehearsal prevented, they still found a primacy effect for the first position. This led them to suggest that rehearsal might not be the complete answer to the primacy effect for the first item in a list. They argued that the first item in a list may have a unique role in forming an "associative net" and consequently is more accessible at the point of retrieval. Bjork and Whitten (1974) also have found primacy effects when subjects are prevented from adopting a cumulative rehearsal strategy.

Another possible explanation for Lana's first-item-only primacy effect is that it may have been mediated by proactive interference (PI). Rosner (1972), working with preschoolers, who typically do not spontaneously rehearse (Kingsley & Hagen, 1969), also found a first-

item primacy effect on the trial following a shift to a new word list from a new item pool after repeated trials with lists from the same item pool. She interpreted this primacy effect found on the shift trial as the result of PI release from old item-position combinations built up over trials. In Lana's case, the word lists on successive trials differed in two respects: the semantic categories from which they were drawn and the color background of the symbols. Consequently, it is possible that shifting the item pools from which the lists were drawn from trial to trial set up a consistent PI release that resulted in Lana's first-item primacy effect.

Like her primacy effect, Lana's recency effect differed from that typical of adult humans in its extent. Adult humans generally show recency effects for the last seven or eight serial positions on lists of from 15 to 40 words (Murdock, 1962), whereas on lists of 8 items, adult humans show recency effects over the last four or five serial positions (e.g., Bernback, 1967; Phillips et al., 1967). Lana, on the other hand, showed a recency effect only at the last serial position. However, similar recency effects at only the last serial position have also been found in the free recall of 2- and 4-year-old children (Myers & Perlmutter, 1978). However, since the extent of the recency effect is presumably mediated by the capacity of STM, this difference between Lana and adult humans may in part reflect only a more limited STM capacity for Lana than for the average adult human. Lana's flat serial position curves for the two- and three-item lists would suggest that her STM capacity at best is only two or three chunks. Lana's STM capacity, then, seems much closer to a preschool child's STM capacity of 3 or 4 chunks than to an adult human's STM capacity of 7 ± 2 .

The similarity between her recall order and that of the typical adult human provides direct evidence that Lana's recency effect was mediated by recall from STM. Lana, like the typical adult human, followed a pattern of recalling the last one or two items first during her recall. Research with humans has shown that when humans are prevented from recalling the last serial positions first, their recency effects are severely attenuated or eliminated (Waugh, 1961). This is apparently the case, because the act of recall tends to prevent rehearsal while also providing interfering input into STM, which interferes with the storage of items already there (Tulving & Arbuckle, 1963; Waugh, 1961). Consequently, Lana seemed to adopt the human-like strategy of avoiding losing the items in STM by recalling them first.

In terms of the human developmental literature (Hagen et al., 1975), it has been shown that recency effects are present in recall from the preschool years on. On the other hand, primacy effects are not typically found in the recall of children until the later elementary school years (Hagen et al., 1975). Nevertheless, if young elementary school-aged children are required to rehearse overtly, their recall will show primacy effects. This

suggests that young children are capable of rehearsing but they do not spontaneously rehearse in free recall situations until the later years of elementary school (Hagen et al., 1975). Developmentally, if Lana's primacy effect reflects spontaneous rehearsal, then she is functioning at the equivalent of a middle elementary school-aged child in this regard. If her primacy effect was mediated by some other factor, then her performance more likely reflects a less developed state.

The findings of the study are also consistent with recent studies by Sands and Wright (1980a, 1980b) of serial retention of visual pictures by a rhesus monkey in a serial probe recognition task. Although the studies differed significantly in terms of procedure, they demonstrated strikingly similar serial position curves. Also, on 10-item lists, the rhesus, like Lana, showed a significant primacy effect only at the first serial position and a significant recency effect only at the last serial position.

Clustering Effects

The standard information processing explanation of category clustering suggests that the categorical relationships are employed both in the organization of the list for storage in LTM and/or in the search and retrieval of items from LTM (Tulving & Olser, 1968). Lana's performance, in terms of both her degree of clustering and her consistency in following a clustering strategy, was substantially below that which would be expected from adult humans. However, her level of clustering and her inconsistency in following a clustering strategy from trial to trial are in line with what might be expected from an early elementary school-aged child (Hagen et al., 1975).

Unfortunately, the factors that mediated Lana's clustering are ambiguous for two reasons. The first is that the clustering could have conceivably been based either on the semantic categories of food-drink, object, and color words or on the totally confounded perceptual categories provided by the different color backgrounds of the symbols associated with each taxonomic category. Recently, perceptually based clustering has been demonstrated in preschool-aged children. These studies have shown that when children are shown chromatic pictures that can be grouped either by semantic categories or orthogonal color categories, preschool-aged children will show some perceptually based clustering (Melkman & Deutsch, 1977; Wingard, Buchanan, & Burnell, 1978). Consequently, it is possible that some or all of Lana's clustering can be accounted for by a similar perceptually based clustering category.

However, there are two lines of evidence that appear to favor a semantic clustering strategy on Lana's part. First, Premack (1976) also has reported clustering in "verbal" chimpanzees on a recognition task. He reports that Sarah's order of selection showed an overall clustering rate of 28%. Second, Lana's differential clustering

across the three taxonomic categories provides some weak evidence for semantic clustering, in that the differences between the categories would seem to be more logically explained on the basis of differences in the strength of the semantic as opposed to perceptual categories. However, a clear distinction cannot be made between the two categorization alternatives due to the nature of Lana's Yerkish symbols. Consequently, the issue of the basis upon which Lana categorized her list still remains unanswered.

The second reason that it is difficult to clearly determine the factors that mediated Lana's clustering is that relationships within a list could have been confounded with high associations between words within a category. When words within a category have high association values, for example, "table" and "chair," it is not clear whether clustering during recall is based on a semantic categorical relationship of furniture or on the high degree of association between the two words (Cofer, 1965). Most of the developmental studies have tended to confound high association value within semantic categories (Hagen et al., 1975). One exception was a study by Lange (1973) in which the items within a given category were not high associates of each other. In that study, Lange found that only 15-year-olds showed above-chance clustering and 5.5- and 11-year-olds showed only chance levels of clustering. On the basis of these findings, Lange concluded that other developmental demonstrations of clustering in young children might reflect skills of a lower level of cognition (i.e., word associations) than that assumed for adults (i.e., the use of categorical relationships in the storage and/or retrieval of words from memory).

Lana's interitem associations are unknown, and consequently, the interitem associations within the list categories could not be controlled. Quite likely, some of Lana's clustering reflects only high interitem associations within categories on some of the lists rather than a storage and retrieval strategy based on the taxonomic relationship of items within a category. However, it is also possible that some high interitem associations between categories, especially those involving color words with objects and food words, could have also existed. This would have tended to lower Lana's clustering.

In comparing Lana's performance with human performance, it should also be noted that she was at some disadvantage compared with the average human subject. The standard human clustering study is designed to maximize clustering by preventing recency effects that interfere with clustering strategies and by employing well defined categories that elicit maximal clustering. Recency strategies are typically prevented by having the subject delay his recall while performing a rehearsal-prevention task, such as counting backward by threes. It was not possible to employ a rehearsal-prevention procedure with Lana. Consequently, on some trials,

Lana tended to order her recall by recalling the last and next-to-last words first, which tended to lower her overall clustering performance.

An assessment of Lana's developmental level with respect to the human developmental literature on clustering (Hagen et al., 1975) depends on the basis on which her recall was organized. If her recall order reflected the perceptual categories defined by the three different color backgrounds of the Yerkish symbols, then her cognitive developmental level would approximate that of a preschool-aged child. If, on the other hand, her recall reflected the semantic nature of the list, be it associative or categorical, her developmental level would seem closer to that of a young elementary school-aged child.

In summary, this study can be best viewed as an initial step in the study of memory abilities in "verbal chimpanzees." Clearly, more work has to be done to clarify the mechanisms that mediated Lana's serial position effects and her tendency to partially cluster her recall. However, the results of this study clearly indicate that Lana was capable of storing and retrieving her Yerkish symbols from memory. The presence of a recency effect in her recall, along with her tendency to recall the last-presented items first, suggests the presence of an STM-LTM division in her memory processes. Also, Lana's primacy effect is suggestive of the presence of a cumulative rehearsal process. But this possibility requires a more direct confirmation by procedures that vary the presentation rate so as to enhance or attenuate cumulative rehearsal effects. Also, regardless of the factors that mediated Lana's clustering, she still did demonstrate a tendency to organize her recall. And this, in itself, is a rather significant cognitive ability. The results of the present study suggest that Lana's LTM may have been semantically organized in terms of grouping of words of similar meaning (e.g., food names) or in terms of association between words in terms of attributes (e.g., red-apple) or defining functions (milk-cup).

REFERENCES

- ATKINSON, B. C., & SHIFFRIN, R. M. The control of short-term memory. *Scientific American*, 1971, **225**, 82-90.
- BENNETT, C. A., & FRANKLIN, N. L. *Statistical analysis in chemistry and the chemical industry*. New York: Wiley, 1961.
- BERNBACK, H. A. The effects of labels on short-term memory for colors with nursery school children. *Psychonomic Science*, 1967, **7**, 145-150.
- BJORK, R. A., & WHITTEN, W. B. Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, 1974, **6**, 173-189.
- BOUSFIELD, W. A., COHEN, B. H., & WHITMARSH, G. A. Associative clustering in the recall of words of different taxonomic frequency of occurrence. *Psychological Reports*, 1958, **4**, 39-44.
- COFER, C. N. On some factors in the organizational characteristics of free recall. *American Psychologist*, 1965, **20**, 261-272.
- COHEN, B. H. Some all-or-none characteristics of coding behavior. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 182-187.
- CRAIK, F. I. M. The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 143-148.
- FISCHLER, I., RUNDUS, D., & ATKINSON, R. C. Effects of overt rehearsal procedures on free recall. *Psychonomic Science*, 1970, **19**, 240-250.
- GENTILE, J. R., RODEN, A. H., & KLEIN, R. D. An analysis-of-variance model for the intra-subject replication design. *Journal of Applied Behavioral Analysis*, 1972, **5**, 193-198.
- GLANZER, M., & CUNITZ, W. H. Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 351-360.
- GOODWIN, C. J. Changes in primacy and recency with practice on single-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, 1976, **15**, 119-132.
- HAGEN, J. W., JONGEWARD, R. H., JR., & KRAIL, R. V., JR. Cognitive perspective on the development of memory. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 10). New York: Academic Press, 1975.
- KAZDIN, A. E. Statistical analyses for single-case experimental designs. In M. Hersen & D. H. Barlow (Eds.), *Single-case experimental designs: Strategies for studying behavioral change*. New York: Pergamon Press, 1975.
- KEELEY, K. Age and task effects in short-term memory of children. *Perception & Psychophysics*, 1971, **9**, 480-481.
- KINGSLEY, P. R., & HAGEN, J. W. Induced versus spontaneous rehearsal in short-term memory in nursery school children. *Developmental Psychology*, 1969, **1**, 40-46.
- LANGE, G. W. The development of conceptual and rote recall skills among school children. *Journal of Experimental Child Psychology*, 1973, **15**, 394-406.
- MELKMAN, R., & DEUTSCH, C. Memory function as related to changes in the basis of organization. *Journal of Experimental Child Psychology*, 1977, **24**, 84-97.
- MYERS, N. A., & PERLMUTTER, M. Memory in the years two to five. In P. A. Ornstein (Ed.), *Memory in children*. Hillsdale, N.J.: Erlbaum, 1978.
- MURDOCK, B. B., JR. The serial position effect in free recall. *Journal of Experimental Psychology*, 1962, **64**, 482-488.
- PELLEGRINO, J. W. A reply to Fender and Boubilet on the measurement of clustering. *Psychological Bulletin*, 1962, **64**, 482-488.
- PHILLIPS, J. L., SHIFFRIN, R. M., & ATKINSON, R. C. The effects of list length on short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1967, **6**, 303-311.
- PREMACK, D. *Intelligence in ape and man*. Hillsdale, N.J.: Erlbaum, 1976.
- ROENKER, D. L., THOMPSON, C. P., & BROWN, S. G. Comparison of measures of the estimation of clustering in free recall. *Psychological Bulletin*, 1971, **75**, 46-48.
- ROBNER, S. R. Primacy in preschoolers' short-term memory: The effects of repeated tests and shift-trials. *Journal of Experimental Child Psychology*, 1972, **13**, 220-230.
- RUMBAUGH, D. M. (Ed.). *Language learning by a chimpanzee: The Lana project*. New York: Academic Press, 1977.
- RUNDUS, D., & ATKINSON, R. C. Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 99-105.
- SANDS, S. R., & WRIGHT, A. A. Primate memory: Retention of serial list items by a rhesus monkey. *Science*, 1980, **209**, 938-940. (a)
- SANDS, S. R., & WRIGHT, A. A. Serial probe recognition performance by a rhesus monkey and a human with 10- and 20-item lists. *Journal of Experimental Psychology: Animal Behavior Process*, 1980, **6**, 386-396. (b)
- THOMPSON, C. P., HAMLIN, V. J., & ROENKER, D. L. A comment on the role of clustering in free recall. *Journal of Experimental Psychology*, 1972, **94**, 108-109.
- THOMPSON, R. K. R., & HERMAN, L. M. Memory for lists of sounds by the bottle-nosed dolphin: Convergence of memory processes with humans? *Science*, 1977, **195**, 501-502.

- TULVING, E., & ARBUCKLE, T. Y. Sources of intratrial interference in immediate recall of paired associates. *Journal of Verbal Learning and Verbal Behavior*, 1963, 1, 321-334.
- TULVING, E., & OSLER, S. Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology*, 1968, 77, 593-601.
- WAUGH, N. D. Free versus serial recall. *Journal of Experimental Psychology*, 1961, 62, 496-502.
- WINGARD, J. A., BUCHANAN, J. P., & BURNELL, A. Organizational changes in the memory of young children. *Perceptual and Motor Skills*, 1978, 46, 735-742.

(Received for publication December 19, 1980;
revision accepted May 21, 1981.)