1-1-1975

A test for short term memory coding preference in deaf subjects.

Anne Chartier Steele

A TEST FOR SHORT-TERM MEMORY CODING PREFERENCE IN DEAF SUBJECTS

by

Anne Chartier Steele

A Thesis

Presented to the Graduate Committee

of Lehigh University

in Candidacy for the Degree of

Master of Science

in the

School of Education

Lehigh University

1975

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

Dec 8, 1975
(date)

Acknowledgments

The writer would like to thank Drs. J. Gary Lutz, Paul Van R. Miller, and Herbert Rubenstein for their help throughout the conducting of this investigation. Appreciation is also extended to the administration, faculty, and students of the Pennsylvania State Oral School in Scranton, Pennsylvania, for their assistance and cooperation.

This manuscript was typed by Mrs. Marie Young. Thanks are due to her.

TABLE OF CONTENTS

ABSTRACT

It has been found that hearing subjects use an auditory/articulatory coding process in the short term memory recall of letter strings. The results of two previous experiments designed to determine what type of coding process is used by deaf subjects were not in agreement. The present study attempted to determine whether these differences were due to sampling differences or to differences in the methods used.

In the two previous experiments the stimuli were exposed for different durations of time, the stimuli that were used were different, and the methods of analysis used in the two studies were different. The two experiments were also conducted on different samples of subjects. In the present study a single sample of subjects was used and all subjects were tested on two occasions. The stimuli used on one occasion were those used in one of the previous experiments and those used on the other occasion were those from the other previous experiment. On both occasions the stimuli were exposed for the same time duration and were analyzed by the same method. The results of each testing occasion classified each subject as either an auditory/articulatory coder, a non-auditory/ articulatory coder, or as uncertain. A subject should be

classified in the same way by each test. If that is not the case, then the differing results of the two previous experiments were not due to sampling but were due to differences in the methods themselves.

Few of the subjects in this experiment were classified in the same way by both tests. This indicates that the differences in the results of the two previous experiments were not due to sampling but to differences in the methods themselves.

INTRODUCTION

Researchers (Conrad, 1964; Wickelgren, 1965; Hintzman, 1965) have found that most people use an auditory/articulatory coding method during tasks that require the immediate recall of letter sequences. This means that once the visual input stimulus has been scanned, the subject will code this input into an auditory/articulatory form. It is in this auditory/articulatory form that the input is rehearsed for memory and retrieved for immediate recall. For example, when the letter "B" is shown, the subject's visual system recognizes this stimulus and immediately the subject transforms this into the sound of the name of the letter. It is this sound that is rehearsed to facilitate memory. The sound of the letter rather than the visual form of the letter is used in the short term memory process. (The above example loosely illustrates the brain process. The writer does not intend it to be a model for visual processing. Rather, hopefully it enables the reader to understand the general concept of auditory/articulatory coding.) The term auditory/articulatory is used to cover either the possibility that it is the sound of the letter name that is coded or that it is the motor activity connected with forming that sound that is used. It is very difficult to separate the two experimentally.

Knowing that most people use this auditory/articulatory coding method, psychologists asked the interesting question "What coding method would a deaf person employ?" In England, Conrad tackled that question experimentally. He has published two articles offering answers to this question (Conrad, 1970; Conrad, 1972). One article appeared in 1970 and one in 1972. The two articles do not support the same answer. It is the purpose of this project to partially replicate Conrad's two experiments in hopes of clearing up some of the possible causes of his differing results. The basic paradigm of Conrad's experiments involves error analysis. Each subject is shown a sequence of letters for a very brief period of time. When the sequence is removed from view, the subject writes down the letters in the sequence as he remembers them. If the subject is using an auditory/articulatory coding method he will frequently confuse a letter in the sequence with another letter which sounds similar. For example, if the visual input sequence was B C X Z T the auditory/articulatory coder might recall B T X Z T. If the subject is not using an auditory/articulatory coding method, it is unlikely that he will make such an error.

In working with deaf subjects Conrad hypothesized that they might use a visual code in short term memory.

In other words, the letters would be rehearsed in the brain in their visual form rather than in their auditory/articulatory form. If this hypothesis is correct, this subject would make errors between letters that looked alike rather than sounded alike. Assuming that X and Z look alike, a visual coder might recall the previous sequence B C X Z T as B C Z Z T. If the subject were not an auditory/articulatory coder, he would not have trouble recalling the B, the C and the T. (It is analogous to the problem of differentiating identical twins. If two stimuli are very similar, they are much more difficult to differentiate than two dissimilar stimuli. To the auditory/articulatory coder, B and T are identical twins, while to the non-auditory/articulatory coder they are dissimilar.)

In 1970, Conrad found that over half (twenty-one out of thirty-six) of his deaf subjects did use an auditory/articulatory coding method just as hearing subjects do. However, in 1972, only nine out of ninety-six deaf subjects used an auditory/articulatory coding method. There are several factors that might explain these differing results. First, the amount of time a subject was allowed to look at the visual input was different. In the 1970 experiment, a subject was allowed a total of five seconds to look at a sequence of five letters; while in 1972, a subject was only allowed one second to look at a sequence of five letters. Second, the visual input in the two Conrad experiments differed. In 1970, Conrad constructed letter sequences from the vocabulary B C H K L T X Y Z. Some of the letters in this vocabulary sound alike (e.g. B and C) while others do not (e.g. X and Y). Conrad believed that some of the letters, namely K X Y Z, looked alike on the basis of the presence of diagonal elements. Therefore, within each input sequence there were some letters that sounded alike and some that hopefully looked alike but at any rate did not sound like the others. In 1972, he used two vocabularies: B C D P T V and K N V X Y Z. All elements of the first vocabulary sound alike while all elements of the second do not sound alike and probably do look alike. A sequence was constructed from one vocabulary or the other. Therefore no one sequence contained both letters that sounded alike and those that did not. Third, the methods of analysis used by Conrad in the two experiments were different. In 1970, he determined whether a given subject was an auditory/articulatory coder on the basis of some "indefensible calculations" (Conrad, 1970, p.183) of the observed number of confusions between letters that sounded alike as compared to the expected number.

In 1972, he calculated something which he termed AI (Articulatory Index). This index is a ratio of the number of errors in the acoustically similar sequences to the total number of errors in all sequences. Fourth, the sampling of subjects differed. In 1970, Conrad used subjects from a private school for the deaf requiring high standards of educability as well as severe hearing loss for entry. In 1972, Conrad used two samples of deaf subjects. One was from a state school requiring only severe hearing loss and the absence of mental subnormality or other handicap for entry. The second sample was selected from a private school where entry was determined by educational attainment, assessed ability to benefit from oral teaching, and severe hearing loss.

The writer hopes this experiment will clear up some of these ambiguities. A single sample was used and sequences in both the 1970 form and the 1972 form were presented to each subject. Every sequence has been exposed for the same time duration. The writer has tried to use the same method of analysis on all sequences - although the nature of the sequences of Conrad's 1970 format make this still somewhat questionable. This experiment is an attempt to determine, when all other factors are held constant, whether the 1970 testing format and the 1972 format are parallel forms.

REVIEW OF THE LITERATURE

It was Conrad's (Conrad, 1964) work in 1964 that first clearly demonstrated that hearing people do use an auditory/articulatory coding method in the short term memory recall of letter strings. In that experiment he presented letter sequences both visually and aurally. By comparing the errors made in the immediate recall of these two types of presentation, he was able to conclude that of the letters presented visually, subjects largely confused those that were acoustically similar.

The visual letter sequences were constructed from the vocabulary B C P T V F M N S X. Conrad believed that these ten letters represented two distinct groups: B C P T V and F M N S X. Each group was believed to have high within-group acoustic confusability while the between-group acoustic confusability was believed to be low. A set of 120 six-letter sequences was prepared according to the following restraints: no letter occurred more than once in any sequence, within each block of twenty sequences each letter occurred equally often in each serial position, and within each block every possible successive letter pair occurred at least once in each possible serial position. A test comprised forty sequences, i.e. two blocks of twenty. Six such tests were formed. The letters of a sequence were exposed one at a time for a duration of .75 seconds per letter. After each sequence the subjects wrote down the sequence as they recalled it. They were allowed as much time as they needed and were instructed to guess rather than leave a blank. The ten-letter vocabulary was written on a blackboard and was visible to the subjects at all times. The subjects were 387 male and female telephone trainees between the ages of sixteen and fifty. They were tested in groups of about ten.

In scoring, all sequences were ignored except those in which only a single substitution error occurred. Conrad writes, "If the data had not been restricted in this way, it is unlikely that anything other than noise would have been added." (Conrad, 1964, p. 77)

The aural letter sequences were recorded by ten untrained speakers. Using the same vocabulary as in the visual sequences, the randomized letters were recorded at the rate of one letter every five seconds. Equal amounts of speech signal and noise were mixed. Two tapes were made. Each contained five randomized vocabularies - three spoken by male and two by female voices. The subjects were 300 male and female Post Office employees between the ages of sixteen and sixty. They were instructed to listen carefully for a letter once every five seconds and to write down the letter they thought they had heard. They were instructed to guess rather than leave a blank.

The results are presented in Tables 1 and 2.

TABLE 1. Listening Confusions

Response Letter	Stimulus Letter: B	C	P	T	V	F	M	N	S	X
B	.	171	75	84	168	2	11	10	2	2
C	32	.	35	42	20	4	4	5	2	5
P	162	350	.	505	91	11	31	23	5	5
T	143	232	281	.	50	14	12	11	8	5
V	122	61	34	22	.	1	8	11	1	0
F	6	4	2	4	3	.	13	8	336	238
M	10	14	2	3	4	22	.	334	21	9
N	13	21	6	9	20	32	512	.	38	14
S	2	18	2	7	3	488	23	11	.	391
X	1	6	2	2	1	245	2	1	184	.

TABLE 2. Recall Confusions

Response Letter	Stimulus Letter: B	C	P	T	V	F	M	N	S	X
B	.	18	62	5	83	12	9	3	2	0
C	13	.	27	18	55	15	3	12	35	7
P	102	18	.	24	40	15	8	8	7	7
T	30	46	79	.	38	18	14	14	8	10
V	56	32	30	14	.	21	15	11	11	5
F	6	8	14	5	31	.	12	13	131	16
M	12	6	8	5	20	16	.	146	15	5
N	11	7	5	1	19	28	167	.	24	5
S	7	21	11	2	9	37	4	12	.	16
X	3	7	2	2	11	30	10	11	59	.

It can be seen in Table 2 (the results of the visually presented letters) that the confusability within groups is much higher than between groups. ($p < .001$) In other words, letters which sounded alike were more likely to be confused. Table 1, which presents the results for the aurally presented letters, shows that during aural presentation confusions are also much more common within groups than between groups. Moreover, using the Spearman coefficient of rank correlation (Conrad, 1964, p. 77) to measure the association between the two confusion matrices (i.e. Tables 1 and 2) it was found that these two matrices are highly related. ($p < .0001$)

These results led Conrad to conclude that errors in the recall of visually presented letters were most likely to occur between acoustically similar letters. This implies that subjects verbalize the stimuli rather than trying to store them in their visual form.

These results then led to an interesting question: Would a child who had been deaf from birth also tend to confuse letters that were acoustically similar? As was mentioned in the introduction, Conrad attempted to answer this experimentally. Since his two experiments did not lead to the same conclusion, the present experiment was designed to begin to locate the factor that led to the differing conclusions. Before describing the present experiment it is necessary to understand in depth the two experiments designed by Conrad.

CONRAD'S 1970 EXPERIMENT

Method

The vocabulary used was B C H K L T X Y Z. These letters were chosen in the hope that they would lead to both auditory/articulatory and shape confusions. Five and six-letter sequences were constructed such that no letter occurred twice in any single sequence, each letter occurred equally often in each serial position of the entire test, and self-evidently easy runs of letters were excluded. A test consisted of forty-five sequences. Sequences were exposed for a duration equal to one second per letter. Immediately following the concealment of the sequence the subjects wrote the letters that they could recall on prepared answer sheets.

The subjects were thirty-six boys, aged twelve through seventeen, who were pupils at the Burwood Park School. This is a private school for the deaf requiring high standards of educability for entry as well as severe hearing loss (i.e. less than 75 db. in the better ear). School records showed the subjects to all have IQ's well above average. All teaching in the school was carried out orally. The school's teachers rated the speech production of the subjects to generally be above average for deaf boys of that age.

The subjects were tested in groups that never exceeded three. Test instructions were given orally by the school Head. A decision as to whether a five or six-letter sequence length test should be used with each group was made at the time of testing on the basis of the results of five practice sequences of the five-letter length.

Two test conditions were used. In one (S) the subjects were instructed to read the sequences silently, while in the other (L) the subjects were instructed to read aloud. (Conrad believed that this latter condition was feasible since the subjects were unable to hear each other.) The condition was changed after every nine sequences. Half of the groups started with the S condition and half with the L.

The subjects were instructed that letters recalled should be recorded on the answer sheet from left to right and that no writing was to be done during the sequence exposure. Instructions for forced guessing were given.

A letter was scored as correct only if it was in the correct serial position.

A control group of seventy-five Cambridge housewives was used. Only the S condition was appropriate for this group. These subjects received the letter sequences letter by letter at a rate of two letters per second.

Results

Error matrices (the correct letter on one axis, the letter reported on the other) were obtained for both the S and the L conditions separately. Inspection indicated that, in both, the same cells contained the large number of errors, so the two matrices were combined into one for further analysis. Inspection of this matrix indicated three main groups of confusions: B C T; and X H

K X Y Z. (There were also many random confusions.) Both of the groups B C T and X H seemed to be auditory/articulatory based confusions (X and H have similar articulatory patterns), while K X Y Z seemed to possibly be confusions based on shape but at least not on an auditory/articulatory basis.

The error matrix for each subject was inspected. Based upon some "indefensible calculations" (Conrad, 1970) of observed and expected probabilities of particular confusions, each subject was placed into either an auditory/articulatory (A) coding group or a non-auditory/articulatory (non-A) coding group. Those subjects who were placed in the A group had made more confusions between letters that were similar on an auditory/articulatory basis while the non-A group members had made more confusions among letters that were not similar on this basis. Table 3 shows the tabulation of confusion errors for the A group and the non-A group for each of the three letter clusters - B C T, X H, and K X Y Z. The writer does not know what method Conrad used to calculate the expected values. The auditory/articulatory group shows significantly more errors between B C T and between X H than would be expected by chance. ($P < .001$ and $p < .01$ respectively) This group does not show significantly more errors between K X Y Z.

TABLE 3. Frequency of confusion by hearing controls, articulators and non-articulators; showing observed, expected and X^2 values, and significance levels

		O	E	χ^2	P
Controls (n=75)	BCT	676	408	99·0	0·001
	XH	245	149	28·1	0·001
	KXYZ	866	825	1·0	n.s.
A Group (n=21)	BCT	193	95	53·6	0·001
	XH	62	38	6·9	0·01
	KXYZ	242	233	1·0	n.s.
Non-A Group (n=15)	BCT	83	78	1·0	n.s.
	XH	40	38	1·0	n.s.
	KXYZ	310	248	11·9	0·001

This is the expected result since the auditory/articulatory group should confuse letters that sound alike. The opposite confusion pattern was found for the non-auditory/articulatory group. This group did not make significantly more confusions between B C T or X H than would have been expected by chance. However, they did make more confusions between K X Y Z ($p < .001$) than would have been expected. These results substantiate Conrad's division of the subjects into two coding groups. Twenty-one subjects were placed in the A group and 15 into the non-A group. Most of these deaf subjects were classified by this format as using an auditory/articulatory coding method for the processing of the letter sequences.

In addition, the subjects' teachers were asked to rate each boy for his speech quality. A three point scale was used: below average, average, or above average. (Only one subject had been rated as "below average" so he was combined with the "average group".) The Yates' corrected χ^2 value of 13.03 ($p < .001$) indicates that these two classifications are not independent. This highly significant degree of association between the two grouping schemes indicates that subjects who were rated as "above average" speakers tended to be classified as auditory/articulatory coders and those who were rated as "average" speakers tended to be classified as non-auditory/articulatory coders.

CONRAD'S 1972 EXPERIMENT

Method

Two vocabularies, each of six letters, were used in this experiment. One (B C D P T V) was designed to have intra-letter similarity on an auditory/articulatory dimension. This vocabulary will be referred to as AS. These letters were chosen on the basis of the results of the 1964 experiment. The other vocabulary (K N V X Y Z) was designed to have visual similarity. This vocabulary will be referred to as VS. These letters were also chosen on the basis of the results of the 1964 experiment, although there was still a great deal of intuition in the selection. Even if this vocabulary should turn out not to be visually similar, it certainly is not similar on an auditory/articulatory dimension. Therefore, these two vocabularies should serve to distinguish between those subjects who have trouble with an auditory/articulatory vocabulary and those who do not.

From each vocabulary, letter sequences were constructed of four to six-letter length. No letter was repeated within a sequence and each letter occurred equally often in each serial position. For each letter sequence length, a test comprised eighteen AS sequences and eighteen VS sequences alternately presented. The letters themselves were 1.6 inches high and printed in black on white cards. Sequences were presented so that all letters were simultaneously available for a period of one second. After one second the sequence was concealed and the subjects immediately wrote, on a prepared answer sheet, the sequence as they remembered it. Forced guessing instructions were given.

A group of thirty-two hearing control subjects was drawn from a state primary school. They were ten or eleven years old. Deaf subjects were drawn from two schools. School one was a small private school to which entry at age eleven was governed by educational attainment, profound hearing loss, and assessed ability to benefit from oral teaching. The entire school population of forty students, aged eleven through sixteen, was tested. School two was a state school with students of ages five through sixteen. Entry requirements included only profound deafness and the absence of mental subnormality or other known handicap. Fifty-six subjects, aged nine through sixteen, were tested. All subjects knew the names of the letters used in the test.

The subjects were tested in groups of up to eight with the deaf subjects and of up to sixteen with the hearing controls. Instructions were given by whatever means seemed appropriate and enough practice trials were given to insure that all subjects fully understood the directions. The practice trials also served the purpose of allowing the experimenter to determine which sequence length should be used with each group. The groups formed for testing purposes were faily homogeneous with regard to educational attainment. Therefore, one sequence length was chosen for each group on the basis that that length would provide enough errors from each subject to make comparisons feasible between the two vocabularies but that it would not make the recall task impossible. For each group the sequence length was the same for both vocabularies.

Results

An Articulatory Index was calculated for each subject. This index indicated the proportion of all errors that were AS (acoustically similar) errors. Thus a high Articulatory Index (AI) score indicated that a subject had found the AS vocabulary difficult to recall. This implies that the subject is using an auditory/articulatory code in the memorizing of these letter strings. Tables 4 and 5 show the results of this experiment.

TABLE 4. Percentage Wrong Letters for Various Sequence Lengths

	Four letters			Five letters			Six letters		
	n	AS	VS	n	AS	VS	n	AS	VS
Hearing	-	--	--	32	20·3	8·8	-	--	--
Deaf 1	-	--	--	25	9·8	20·0	15	13·3	17·9
Deaf 2	24	16·4	34·4	27	17·2	38·7	5	22·6	40·0

TABLE 5. Number of Subjects in Each Bracket of the Articulatory Index (AI)

	n	0–0·20	0·21–0·40	0·41–0·60	0·61–0·80	0·81–1	Mean AI
Hearing	32	0	1	9	9	13	0·70
Deaf 1	40	9	12	14	4	1	0·37
Deaf 2	56	13	35	7	1	0	0·32

A Mann-Whitney U Test shows the AI distributions for the two deaf samples to be significantly different at the .001 level. However, the mean AI for either of the deaf groups as far lower than that of the hearing controls. In the Deaf 1 Group, where all teaching was oral, thirty-two out of forty subjects had Articulatory Indices below fifty. In other words thirty-two out of forty subjects had been classified as using a non-auditory/articulatory coding process. However, in the Deaf 2 group all but one of the subjects had Articulatory Indices below fifty.

SUMMARY

The results of Conrad's 1970 experiment indicated that over half (twenty-one out of thirty-nine) of the deaf subjects were using an auditory/articulatory coding process. The results of his 1972 experiment indicated that only nine out of ninety-six subjects were using an auditory/articulatory coding process. The results of these two experiments do not agree as to the proportion of deaf children who are using an auditory/articulatory coding process. Perhaps sampling differences can account for this. However, one of the two samples of deaf subjects used in the 1972 experiment was chosen from a private school similar in admission requirements to the school from which the 1970 sample was chosen and yet the proportions found in these two also differed. For this sub-sample of the 1972 study that was from a school similar to the one used in the 1970 sampling, eight out of forty subjects were classified as auditory/articulatory coders. Although this is more similar to the 1970 proportion than that found in the total 1972 sample (nine out of ninety-six) it is still very different than the proportion (twenty-one out of thirty-six) which Conrad had found in the 1970 sample.

Other factors differed in the two experiments and one or more of these might account for the different results. These factors were exposure time of the stimulus, the stimulus itself, and the method used to analyze the results. By controlling these factors the writer hopes to determine if sampling differences could explain the differing results of Conrad's two experiments.

METHOD

Two separate tests were constructed. Test One used the vocabulary from the 1970 Conrad study: B C H K L T X Y Z. Thirty-six four, five, or six-letter sequences were formed according to the same restraints that Conrad had used: no letter was repeated within a sequence, each letter occurred equally often in each serial position, and self-evidently easy runs of letters were excluded. Test Two used the two vocabularies from the 1972 Conrad study: B C D P T V and K N V X Y Z. From each vocabulary eighteen sequences of four, five, or six letter length were formed following the same restraints as above. This test was compiled by alternating a sequence from one vocabulary with a sequence from the other vocabulary.

The subjects were fifty-five students enrolled at the Pennsylvania State Oral School in Scranton, Pennsylvania. All subjects had severe hearing loss since birth and were of normal mentality. None had multiple handicaps and none came from homes with deaf parents. The subjects ranged in age from nine through sixteen. Many kinds of instruction had been used at the school. All of the subjects were able to use the deaf sign language and most could lip read at least to some extent.

The subjects were tested in groups of eight or less. The groups were formed so as to be relatively homogeneous with respect to age and academic ability. In both of Conrad's studies the appropriate sequence length for each group was determined on the basis of the results from the practice trials. That subjective element is, of course, impossible to replicate. Also, mechanically it was impossible for the writer to leave that decision to the last minute. Therefore, it was predetermined that all nine and ten year olds would receive four-letter sequences; eleven, twelve, and thirteen year olds would receive five-letter sequences; and fourteen through sixteen year olds would receive the six-letter sequences. The writer hoped that these would approximate Conrad's decisions. Conrad had introduced the use of four-letter sequences in the 1972 experiment. The subjects in that study were aged nine through sixteen. In the 1970 experiment, in which only five and six-letter sequences were used, the subjects were aged twelve through seventeen. Based upon that, the writer concluded that the younger subjects needed the shorter sequences. Beyond following this general pattern, that sequence length increased as age increased, the choice of the cut-off points were based upon intuition.

A slide projector was used to display the letter sequences on a screen in front of the group to be tested. A tachistoscope was used to regulate the exposure duration of one second per sequence with all letters of the sequence exposed simultaneously. That was the exposure duration and method that Conrad had used in the 1972 experiment. In 1970, he had used a duration equivalent to one second per letter. Based upon other short-term memory experiments, this latter time length seems extraordinarily long. The 1972 duration of one second per sequence seems the more reasonable of the two.

One letter sequence was printed on one slide. The slides were made by transferring "Quik Stik Rub-On Letters" to acetate and gluing this into a slide frame. The letters used were all gothic capitals. The subjects were instructed to look at the letter sequence while it was exposed and, as soon as the sequence was concealed, to write the letters recalled on the prepared answer sheet. They were allowed as much time as they needed to write their answers. Conrad did not say whether he had restricted this time. The subjects were instructed to guess rather than to leave a blank. The writer gave all instructions in whatever way the subjects would be able to understand as the writer is not able to use the deaf sign language. Six practice trials were given. The writer found that all subjects were able to understand the procedure by the end of the practice trials.

Each subject was tested on two separate occasions, a week apart. Each subject received both tests. Half of the groups took Test One first while the other half started with Test Two. The time needed to administer the practice trials and the test averaged fifteen minutes per group.

RESULTS

For each subject two Articulatory Indices were calculated - one for the results of the test designed according to the format of Conrad's 1970 experiment and one for the results of the test designed according to his 1972 format. Conrad had only used the Articulatory Index during his 1972 experiment. Calculating the index for the present results presented no problems for the 1972 format but did present problems for the 1970 format. In the 1972 format the sequences were either of all acoustically similar elements or of all visually similar elements. Any mistakes in the AS sequences were auditory/ articulatory mistakes and any mistakes in the VS sequences were not. Therefore, when calculating the AI, the experimenter can simply count the number of errors that occurred in the AS sequences and divide by the total number of errors. Even though the subjects were instructed to guess, omissions still occurred. In the 1972 format any omissions were counted as errors. (Conrad states that he had few omissions but does not say how he treated those he did have.)

Calculating an Articulatory Index for the results of the test based upon Conrad's 1970 experiment was not as straightforward. In his 1970 experiment, Conrad had found three clusters of confusion errors: B C T, X H, and K X Y Z. Whenever a subject from the current experiment made such an error it was easy to classify it. However, many confusion errors occurred that did not belong to any of the above clusters. Conrad had given no evidence as to how other errors should be classified. Nevertheless, the problem existed as to how to score these other errors. The writer asked five judges to rate every possible letter pair confusion as to whether they felt that it was a sound confusion, shape confusion, neither, or both. Unfortunately, sufficient agreement was not found to use this as a scoring guide. Therefore, the writer decided to only score those errors that corresponded to the three clusters Conrad had found. The results may have been affected by this procedure. Quite a bit of information was lost this way, but it did seem to be the only way to begin to locate differences, if any, that might exist between Conrad's 1970 test and his 1972 test. Omissions could not be counted at all. Any confusions among the letters B C T or X H were counted as auditory/articulatory confusions and any confusions among the letters K X Y Z were counted as non-auditory/articulatory confusions. The AI was simply a ratio between the auditory/articulatory confusions and the total confusions of any of the three clusters.

Once these Articulatory Indices were calculated, the subject could be classified as either an auditory/ articulatory coder or a non-auditory/articulatory coder. Each subject was classified on the basis of the AI from the 1970 format and from the 1972 format. For each subject these classifications were either in agreement or they were not. It was on the basis of this agreement/ disagreement that the results of this experiment were analyzed. If the 1970 format and the 1972 format are actually "parallel forms" then each subject should be classified the same way by both tests. If no such agreement exists then the two formats must be measuring two different things since it is unlikely that one subject's coding preference would change from one week to the next.

In this experiment each subject was classified in one of three ways: a non-auditory/articulatory coder, uncertain, or as an auditory/articulatory coder. A non-auditory/articulatory coder had an AI between 0 and .40, the uncertain subject had an AI between .41 and .60, and the auditory/articulatory coder had an AI between .61 and 1.00. In the 1972 experiment Conrad used five categories: 0-.20, .21-.40, .41-.60, .61-.80, and .81-1.00. However, the writer was willing to tolerate agreement with less precision. Conrad's testing procedure is still in its early stages as evidenced by the need for the present study. Agreement on even a three category scale would be an accomplishment. The writer knows of no reason to suspect that a person with an AI of .15 is a different sort of coder than a person with an AI of .28. The "uncertain" category was established because the writer does not believe that at this stage of testing it can be said that a subject with an AI of .49 is very different from a subject with an AI of .51. At this stage the AI is not that precise a measure. Therefore, the writer felt that the three category scale was optimal.

The χ^2 One-Sample Test (Siegel, 1956, p. 42-47) was used to test the degree of agreement under the null hypothesis $f_A = f_D$. If the null hypothesis is accepted then there is just a chance relationship between the two tests. If the two tests are actually parallel forms we would reject the null hypothesis and accept the alternative hypothesis $f_A > f_D$. However, the possibility exists that the two tests actually tend to disagree so that the appropriate alternative hypothesis is $f_A \neq f_D$.

Table 6 shows the results of the analysis. The row entitled "total" refers to the total sample of subjects. This is further broken down into sub-samples of those who saw four-letter sequences, five-letter sequences, and six-letter sequences.

TABLE 6

	n	Agree	Disagree	χ^2	p
Total	55	20	35	4.09	$.05 > p > .025$
4-letter	18	13	5	3.55	$.10 > p > .05$
5-letter	10	3	7	1.60	$.30 > p > .20$
6-letter	27	4	23	13.37	$p < .001$

For the total sample of subjects the null hypothesis can be rejected at the .05 level and the alternative hypothesis $f_A \neq f_D$ can be accepted. However, it is apparent from the agreement/disagreement distribution that the $f_A < f_D$. In other words the two forms of the test tended to classify a person in two different ways. This pattern is not true for all sub-samples. Those subjects who saw six-letter sequences do tend to be classified differently but even more significantly so ($p < .001$) than the total sample. The agreement/disagreement distribution for those subjects who saw five-letter sequences does not differ significantly from the chance distribution. Those subjects who saw the four-letter sequences tended to be classified in the same way. The null hypothesis for that group can be rejected at the

.10 level. (It is marginally significant at the .05 level.) The alternative hypothesis $f_A \neq f_D$ may be accepted. However, it is apparent from the agreement/ disagreement distribution that the $f_A > f_D$.

Overall very few subjects were classified as auditory/articulatory coders by either format. The distributions for the 1970 format and the 1972 format are presented in Table 7.

TABLE 7.

	~A	U	A
1970	40	13	2
1972	15	39	1

During the administration of this experiment a possible confounding factor came to the writer's attention. In the 1972 format a letter sequence was either of all acoustically similar letters or all visually similar letters. Any subject who is, in fact, an auditory/articulatory coder would be pronouncing these letter sequences in the process of trying to remember them. Because of the nature of the letters, the acoustically similar sequences could be pronounced much faster than the visually similar letters. It seemed that one could say C D P T V B much faster than one could say X K N Z V Y. Therefore, in this short tachistoscopic presentation, the auditory/articulatory coder would be able to code more of the acoustically similar sequence than the visually similar sequences. However, this was all based on intuition. In order to try to determine if there was any truth to this, another AI for each subject who saw the five or six-letter sequences based only upon the errors in the first three positions was calculated.

The choice of the third position as a cut-off point was largely intuitive but was based in part upon a serial position analysis. This graph is shown in Table 8. The horizontal axis represents the letter position and the vertical axis the percentage of correct responses.

The slope for the four-letter sequences remained fairly constant throughout. For both the five-letter sequences and the six-letter sequences, the slope becomes steeper after the third letter position. Actually, the slope patterns for the five and six-letter sequences are typical in tachistoscopic recall experiments. The letters in the early position are best recalled and the letters in the final positions are better recalled than those in the middle positions. (Conrad used a serial position

TABLE 8

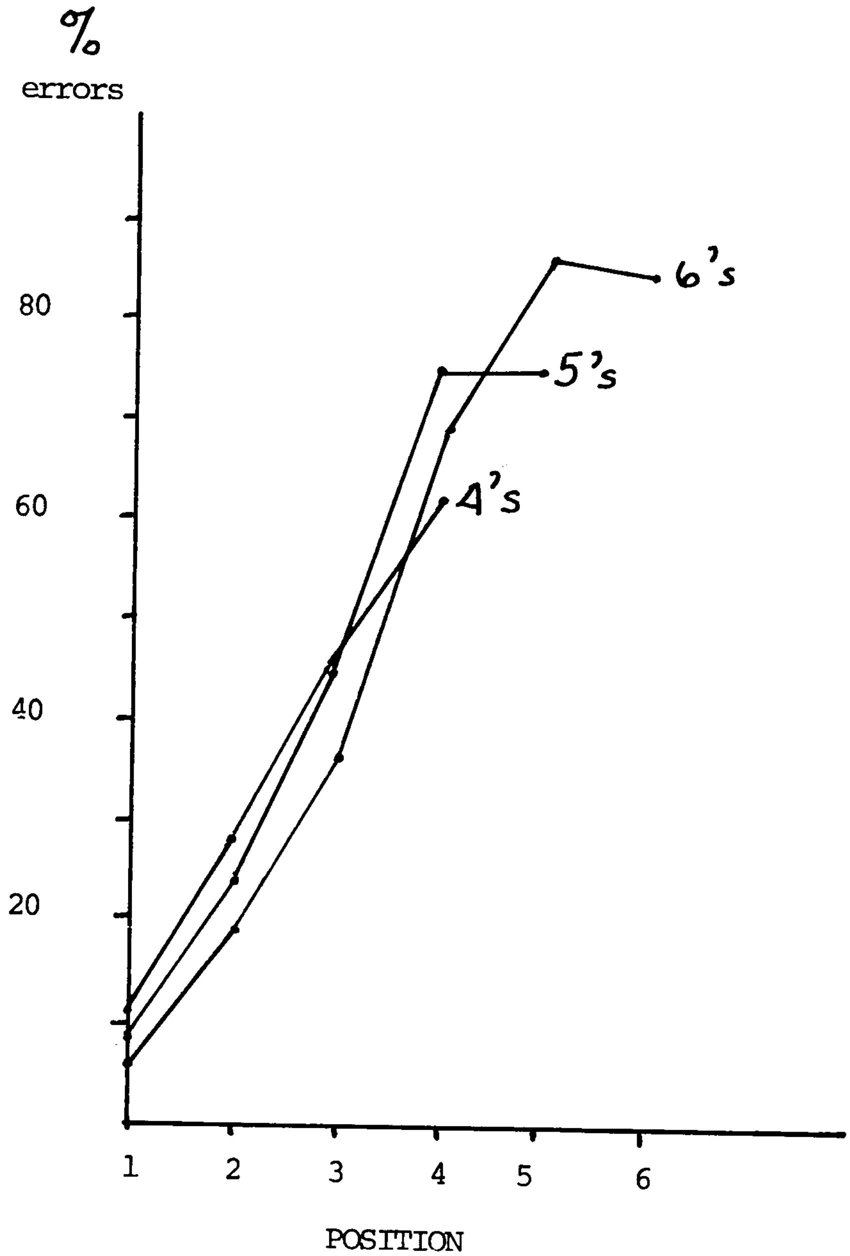
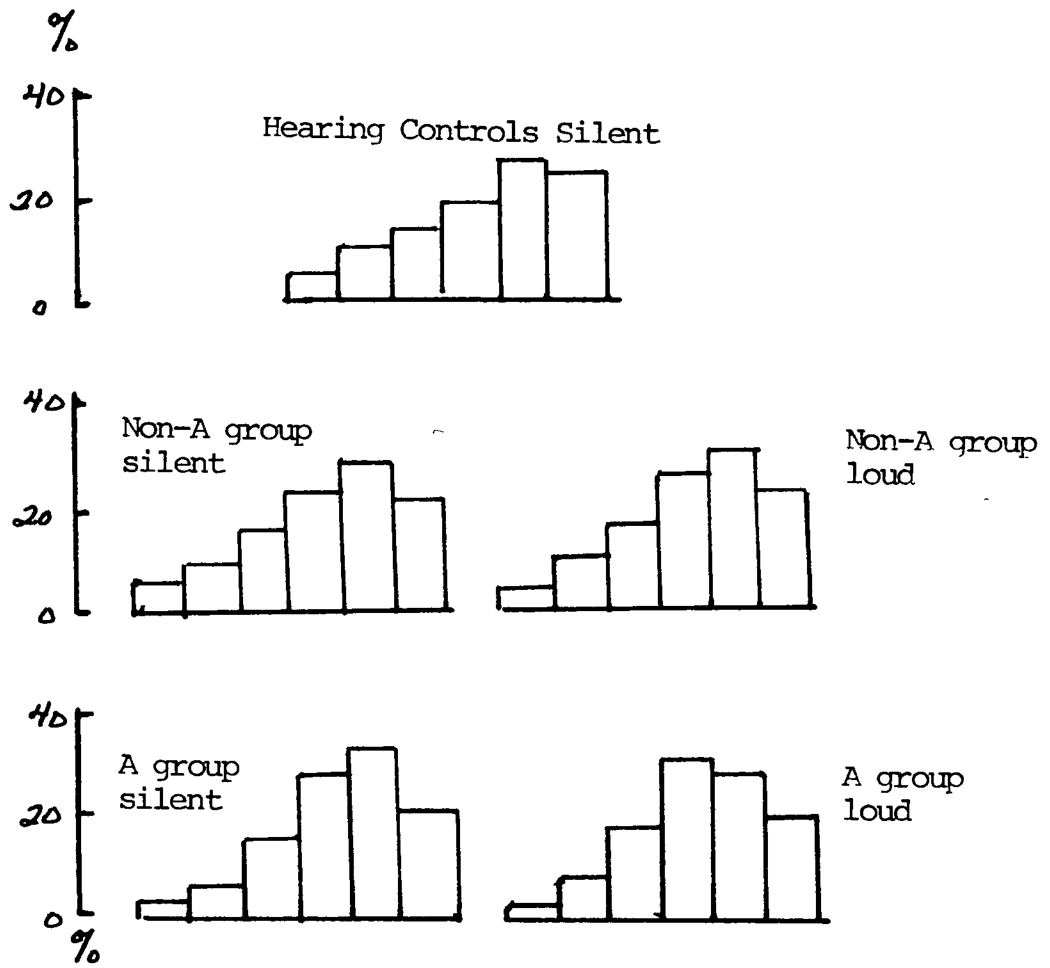


TABLE 9



Percentage distribution of errors across serial positions analysis for the data from his 1970 experiment. This is presented in Table 9. It can be seen that recall is poorest for the letters in the middle positions. He published no such analysis for his 1972 data.) It is possible that the serial position analysis for the present experiment merely illustrates the usual finding, i.e. the fourth position is a middle position. However, the purpose of the analysis was to determine if the extraneous variable of speed of verbal pronunciation might have affected the results. Therefore, the writer was looking for a point at which it was likely that the letters from both kinds of sequences could have been pronounced in the time allowed. Something happened between the third and fourth positions to cause the error rate to increase. Perhaps it had something to do with the difference in pronunciation speed of the letters or perhaps it is only due to the memory process itself. At any rate the third position seemed the best choice for a cut-off point. (It is interesting to note here that this whole analysis was based upon the writer's belief that acoustically similar letter strings could be pronounced faster than those that were not acoustically similar. If that is the case it would only cause problems for those coders who were pronouncing the letter strings, i.e. using an auditory/articulatory code. However, the results of this experiment indicate that few subjects were using such a code. Therefore, pronounciation speed would present no problem. However, it is possible that the visually similar strings were able to be coded faster by these subjects than were the acoustically similar strings. In other words the VS sequences had the same advantage for non-auditory/articulatory coders as would AS sequences for auditory/articulatory coders. If that is the case then coding speed is still an extraneous factor. The differences between the sub-samples who received different sequence lengths do indicate that that is a possibility. This analysis based on only the first three positions should help determine if coding speed might be a variable. This analysis seems warranted even though few of the subjects were classified as auditory/articulatory coders.)

Table 10 presents the information from the analysis of only the first three positions. The index for that result is called AI*.

TABLE 10.

	n	Agree	Disagree	χ^2	p
Total	55	25	30	.45	$.70 > p > .50$
5-letter	27	7	20	6.25	$.02 > p > .01$
6-letter	10	4	6	.40	$.70 > p > .50$

Again the null hypothesis is $f_A = f_D$ and the alternative hypothesis $f_A \neq f_D$. For the total sample, i.e. the AI classification for the subjects who received four-letter sequences and the AI* classifications for those subjects who received five or six-letter sequences, the null hypothesis cannot be rejected. The agreement/disagreement distribution does not differ significantly from chance. Neither does the distribution for those subjects who saw the five-letter sequences. However, for those subjects who received six-letter sequences, the null hypothesis can be rejected at the .02 level of significance and the alternative hypothesis accepted. It is apparent from the agreement/disagreement distribution that the $f_A < f_D$. In other words, the two forms of the test still tended to classify the subjects differently.

As with the AI, AI* classified few subjects as auditory/articulatory coders. The distribution from the three formats (1970, 1972, and 1972-AI*) are presented in Table 11.

TABLE 11.

	~A	U	A
1970	40	13	2
1972	15	39	1
1972-AI*	24	26	5

The 1972-AI* analysis has classified more subjects as auditory/articulatory coders than either of the other two tests.

CONCLUSIONS AND DISCUSSION

The results of this experiment have shown that very few deaf children from the Pennsylvania State Oral School in Scranton, Pennsylvania use the auditory/articulatory coding form. These children were not coding in the same way that hearing children would. These findings are more in keeping with the results of Conrad's 1972 experiment in which only six out of ninety-six subjects were classified as auditory/articulatory coders than with the results of his 1970 experiment in which twenty-one out of thirty-six subjects were classified as auditory/articulatory coders.

Unfortunately results of this experiment also indicate that the 1970 format and the 1972 format of Conrad's test for the classification of the coding process of visual letter strings are not parallel forms of the same test. These two formats generally do not tend to classify the same individual in the same way. The 1970 format classified most of the subjects in this experiment as non-auditory/articulatory coders while the 1972 format classified most as uncertain.

However, this finding does not hold true for each sub-sample. Those subjects who saw four-letter sequences did tend to be classified in the same way by both formats. That was the only sub-sample, however, in which that was the case. There was no more than chance agreement for those subjects who saw the five-letter sequences and for those who saw six-letter sequences there was a definite trend for a subject to be classified in the same way by the two formats. It is interesting to note that as the sequence length increased, the two formats became less and less like parallel forms. Perhaps sequence length itself was affecting this or perhaps age itself affects the coding form used by a deaf person or perhaps both (The older subjects received the longer sequences). If it is the case that age affects the coding form, then age was interacting with one form of the test and not the other.

I think the answer to the question of the influence of sequence length could be an important one for the education of deaf children. If sequence length did not affect the results and, in fact, age is related to a coding preference this could be due to either a physiological change or to educational influence. Perhaps as a deaf student learns to speak his coding preference adapts. The hearing child uses an auditory/

articulatory code. The deaf child may learn to use this code as he learns to use a vocal language.

However, it is just as possible that that is not the case. If it turned out that sequence length itself, rather than age, brought about different classifications then perhaps psychologists and educators should investigate the relationship between coding method in a deaf child and the teaching method used with that child.

The results of the analysis based only upon the errors in the first three positions lend some support to the hypothesis that sequence length itself might affect the coding classification. The re-classification of subjects based upon that analysis still did not bring agreement. Those subjects who had seen the five-letter sequences still had no more than chance agreement and those subjects who had seen the six-letter sequences still tended to be classified differently. However, there was a trend in both groups toward improved agreement as compared with the results of the analysis of the total sequence. For the group seeing the five-letter sequences, the classifications for three subjects agreed and seven disagreed when the whole sequence was analyzed while four agreed and six disagreed when only the first three positions were analyzed. For the group seeing the six-letter sequences the change was from four and twenty-three to seven and twenty. That change was enough to reduce the probability of rejecting the null hypothesis in favor of the alternative hypothesis $f_A \neq f_D$ from less than .001 to less than .02.

It is only further research that can determine whether sequence length affects the classification of subjects. This is an extraneous variable that should be investigated. The procedure of analyzing only the first three positions was designed on the assumption that pronunciation speed might affect the results. Very few subjects in this experiment were auditory/articulatory coders. Pronunciation speed should not affect them. However, the results of this analysis suggest that sequence length might affect all subjects. Perhaps visual coders have the same effect of speed in the VS sequences that auditory/articulatory coders have in the AS sequences.

This study has shown that the difference between Conrad's two formats was not due to sampling. There is a difference in the formats themselves. This difference is not in the exposure duration since that was held constant in the present study. An attempt was also made to hold the method of analysis constant. As was discussed earlier, that may not have been successful. I would not have been successful. I would suggest that the next step in attempting to determine what has brought about the difference in these two formats would be to replicate the present study using only four-letter sequences for all subjects. That would answer the question of the effect of sequence length. Once that is determined the questions concerning the method of analysis can be tackled.

If these further controls still indicate that the two formats are not parallel, it becomes a question of just which one, if either, is valid. To answer that question researchers will have to investigate the differences between the two stimuli.

It was mentioned earlier in this paper that the writer had asked five independent judges to rate all possible letter pairs as to whether the judge believed each to sound alike, look alike, neither, or both. The writer hoped to use this as a scoring guide for the errors made in the 1970 format. The criterion for agreement was for three out of the five judges to classify a letter pair in the same way. However, this criterion was not met for a large part of the letter pairs. In many instances criterion agreement was not reached for those letter pairs which were "known" to be acoustically similar from the 1964 experiment of Conrad. For example, the distribution for the letter pair B P was one rating of "look alike," two ratings of "sound alike," and two ratings of "both." This lack of agreement raises some questions concerning the model of auditory/articulatory coding upon which these experiments with the deaf are based. None of the judges used in this study had a technical knowledge of linguistics but all were generally well educated. It is possible that a technical knowledge of linguistics is necessary. However, it is also possible that what have been considered to be auditory/articulatory confusions are actually confusions along a different dimension. Research into this seems warranted. It would be useful to compare errors in letter strings designed by technicians (such as those used in Conrad's 1970 and 1972 formats) and in letter strings based upon the ratings of educated judges who are not linguists. Such an experiment would be helpful in further understanding the short term memory processes of both the hearing and the deaf.

REFERENCES

Conrad, R. Acoustic confusions in immediate memory. _British Journal of Psychology_, 1964, _55_, 75-83.

________. Short-term memory processes in the deaf. _British Journal of Psychology_, 1970, _61_, 179-195.

________. Short-term memory in the deaf: a test for speech coding. _British Journal of Psychology_, 1972, _63_, 173-180.

Hintzman, D. L. Classification and aural coding in short-term memory. _Psychonomic Science_, 1965, _3_, 161-162.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Wickelgren, W. A. Short-term memory for phonemically similar lists. _American Journal of Psychology_, 1965, _78_, 567-574.

VITA

Anne Chartier Steele was born on November 2, 1948, in Glens Falls, New York, the only child of Donald and Louise Chartier. She received her elementary and secondary education in the public schools of Hudson Falls, New York. In 1970, she was graduated from Reed College in Portland, Oregon, with a Bachelor of Arts degree in religion. She enrolled in the Master of Science program in Educational Measurements and Research at Lehigh University in September, 1972. Presently she is living in Scranton, Pennsylvania with her husband, Rollin M. Steele.