Acoustic similarity in long-term paired-associate learning[1]

H. C. A. DALE, University of Hull, England, and A. D. BADDELEY, University of Sussex, Falmer, Brighton, England

Baddeley & Dale (1966) using a classical A-B, A'-C interference paradigm showed an adverse effect of semantic similarity on the long-term retention of paired associates. The present study applied the same procedure to the study of acoustic similarity. No significant effect was found.

Interitem acoustic similarity has a strong adverse effect on short-term serial learning (SL). This is demonstrable by the conventional memory-span technique (Conrad & Hull, 1964; Baddeley, 1966a), and also by using an interpolated interference task (Dale, 1964; Wickelgren, 1965). In contrast to this, there appears to be little evidence for any marked effect of acoustic similarity in long-term memory. Woodworth (1938, p. 37) observed that when trying to recall a word from long-term memory he would occasionally produce an acoustically similar item, and examples of this phenomenon have been produced under laboratory conditions by Brown & McNeil (1966). However, the size of this effect is relatively small, and an attempt by Baddeley (1966b) to produce an effect of acoustic similarity in long-term serial learning proved unsuccessful.

In the case of semantic similarity, however, the reverse is true, for while it may affect short-term serial memory (Dale & Gregory, 1966), its effect tends to be smaller than that of acoustic similarity (Baddeley, 1966a). It also differs from acoustic similarity in producing an effect on long-term learning (Baddeley, 1966b; Underwood & Goad, 1951). For serial learning, then, the picture seems fairly clear: STM shows a marked effect of acoustic similarity and a small effect of semantic similarity; LTM shows little effect of acoustic similarity but a pronounced effect of semantic similarity.

The question then arises as to whether this reciprocal relationship also holds for paired-associate learning (PAL). In the case of semantic similarity, Baddeley & Dale (1966) have shown that it does. They used a classical interference paradigm in which S learns A-B, then learns A'-C and subsequently tries to recall A-B. The Ss, A and A', were semantically similar adjectives, and the Rs, B and C, were unrelated adjectives. This showed that semantic similarity had a marked deleterious effect on retention. An STM adaptation of the paradigm, however, revealed no comparable effect of semantic similarity (Baddeley & Dale, 1966; Dale, 1967). Thus far, then, the results from SL and PAL are in agreement. To complete the picture, it is necessary to examine the effect of acoustic similarity on the long- and short-term retention of paired associates. If the parallel with serial learning results continues, we should find it has no effect on LTM but has a marked effect on STM. Two experiments using the LTM paradigm are to be reported here.

A study of Bruce & Murdock (1968), suggests that acoustic similarity does not effect RI in LTM PAL. They have no positive results against which this may be contrasted, however. Thus their negative findings could conceivably be attributed to peculiarities of their Ss' strategies, or to other aspects of their technique. The present paper begins by reporting an experiment which closely parallels that of Baddeley & Dale (1966). The Ss were drawn from the same source; the design was the same, and as far as possible the OL stimuli were the same. We can therefore assert with confidence that these conditions will yield substantial RI effects.

EXPERIMENT 1

Method

As in the previous study of semantic similarity, Baddeley & Dale (1966), four lists of eight pairs of disyllabic adjectives were employed: A, B, X, and Y. The stimuli in Lists A and B were matched so that for each stimulus in A there was an acoustically similar (rhyming) stimulus in B (see Table 1). Stimuli in X and Y were similarly matched, while care was taken to avoid similarity between stimuli in the different matched pairs of lists, i.e., between A and Y and B and X. Twelve of the 16 stimuli of Lists A and Y and 28 of the 32 response words were taken from the previous study.

Four groups were run, two experimental and two control. All the Ss, who were tested in groups, learned two lists, both for eight trials. Those in the experimental groups learned either A and B or X and Y, while the control groups learned either A and Y or X and B. Each trial consisted of successive presentation and test phases. Both presentation and test rates were 4 sec per pair. The interval between presentation and test was 10 sec, as was the interval between trials. Both presentation and test orders were freshly randomized on every trial. The interval between OL and IL lists was approximately 3 min and OL retention was tested immediately after IL. The material was written freehand with marker pens on large cards (10 x 24 in.) which were held up by E. The responses were written in specially prepared booklets using one page per trial. Between trials the Ss were required to turn over to the next page so that their previous responses were no longer visible. An invigilator checked that they obeyed instructions. The Ss were all young, newly enlisted men.

Results

The sizes of groups available for testing varied from 16 to 24. Initially, data from all Ss who had three or more correct items on the last OL trial were analyzed. These showed that retention was slightly better in the control condition than in the experimental. As in the previous study (Baddeley & Dale, 1966), the results were analyzed in terms of the mean percentage loss $\{[(OL - R)/OL] \times 100$, where OL = score on the final learning trial, and R = retention score$\}$. The overall mean percentage loss was 41.80 for the Ss in the experimental groups (N = 38), and 34.75 for the controls (N = 32). This difference was not statistically significant ($t = 1.00, df = 68$).

For further analysis, equal sized groups were extracted, as in Baddeley & Dale (1966). Since retention was positively correlated with OL performance (tau = 0.28; N.D. = 2.88; $p < .01$), these groups were matched on OL (see Table 3). The difference between experimental and control conditions was not significant ($t = 0.93, df = 50$).

Discussion

These results suggest that the acoustic similarity of stimuli does not differentially effect retroactive interference in this LTM

Table 1
Word Lists used in Experiment 1

List A		List B		List X		List Y	
S	R	S	R	S	R	S	R
Fearful	Mixed	Cheerful	Spicy	Bragging	Alike	Dragging	Equal
Silly	Fertile	Chilly	Rusty	Damned	Colored	Rammed	Shining
Handy	Biting	Sandy	Tuneful	Weary	Daring	Beery	Clever
Involved	Recent	Revolved	Dizzy	Spoken	Shabby	Broken	Narrow
Skillful	Lovely	Willful	Complex	Funny	Distant	Sunny	Oily
Ready	Golden	Steady	Hollow	Ailing	Flashy	Failing	Creased
Faultless	Certain	Thoughtless	Tiny	Seedy	Foamy	Speedy	Windy
Prepared	Heated	Repaired	Feeble	Afraid	Stubborn	Betrayed	Common

Table 2
Word Lists used in Experiment 2

List A		List B		List X		List Y	
S	R	S	R	S	R	S	R
Rose	Beat	Shows	Back	Flap	Bill	Clap	Use
Preys	Add	Phase	Check	Toss	Work	Boss	Wind
Make	Inch	Take	Foot	Stroke	Learn	Soak	Park
Cry	Join	Sigh	Case	Shear	Judge	Peer	Floor
Light	Free	Write	Last	Flout	Name	Shout	Bank
Store	Land	Pour	Kiss	Dare	Stick	Wear	Catch
Cuff	Milk	Snuff	Trip	Weep	Cold	Heap	Close
Fish	March	Wish	Plant	Hail	Set	Bale	Give

task. The generality of this finding may be questioned, however. The experimental material was derived from that of the previous investigation and although it possesses the virtue that it permits comparison with the earlier findings, it carries the weakness that the material has no demonstrable potency. Accordingly, this experiment has been replicated using acoustically similar words which have previously been shown by Dale & Gregory (1966) to cause a decrement in STM performance.

EXPERIMENT 2

Method

The design and procedure was closely modeled upon that of Experiment 1. The stimulus words were taken from the Dale & Gregory (1966) study in which 12 sets, each consisting of six monosyllabic, acoustically similar words, were employed. The present design required a total of 16 pairs of acoustically similar words. These were obtained by taking one pair from each set of the Dale and Gregory material and generating four additional pairs. Formal similarity was minimized (cf. Table 2). Care was taken to avoid similarity between the response words, which were high frequency monosyllables (AA in the Thorndike-Lorge count), and also between stimulus words within a list, and between stimulus and response words. Following preliminary tests, the number of OL and IL trials was reduced to four; otherwise the method was the same as in Experiment 1. Fresh Ss from the same source were employed.

Results

The groups were again of unequal size, varying from 14 to 33, and matched groups were extracted as before. These were also matched as far as possible on IL learning, and where a choice had to be made Ss were selected so as to maximize the difference between experimental and control retention scores. This method of selecting Ss biases the results in the direction of an effect (percentage loss calculated on all the data was 38.9 for experimental groups and 37.2 for controls). The resulting scores are given in Table 4. The overall difference in these selected results was still not statistically significant ($t = 1.25$, $df = 42$). Analysis of pooled data from Experiments 1 and 2 also failed to achieve statistical significance ($t = 1.3, df = 94$).

DISCUSSION

Both experiments showed differences in the overall means which suggest that some small effect of acoustic similarity upon RI might exist. In neither experiment was this difference statistically significant, even though the analysis of Experiment 2 was biased to increase its sensitivity. Since the pooled results of both experiments, when analyzed together, still fail to reveal statistical significance, we have to conclude that with this particular paradigm acoustic similarity does not significantly influence RI. This result must be contrasted with the substantial effects found for semantic similarity by Baddeley & Dale (1966) using the same paradigm with a similar experimental technique and Ss from the same source. The tentative conclusion to be drawn from this comparison of the two studies is that in normal (LTM) PAL the stimulus words are coded semantically but not acoustically. This conclusion is reinforced by Bruce & Murdock's (1968) evidence.

The problem with the comparison between the effects of acoustic and semantic similarity, as with any comparison across stimulus dimensions, is that the degree of similarity achieved cannot be equated. It can, therefore, be objected that a higher degree of stimulus similarity must have been employed when semantic similarity was studied. This possibility cannot be excluded. Nonetheless, we emphasize that every effort has been made to employ as high a degree of similarity as possible. The material in Experiment 2 was chosen because of its known potency in STM and we note that Bruce & Murdock (1968) formally analyzed the acoustic properties of their material, selecting pairs of S words which differed in only a single distinctive feature. Despite the logical impossibility of excluding this interpretation we are inclined to discount it and conclude that acoustic coding is not employed in LTM PAL.

REFERENCES

BADDELEY, A. D. Short term memory for word sequences as a function of acoustic, semantic and formal similarity. Quarterly Journal of Experimental Psychology, 1966a, 18, 362-365.

BADDELEY, A. D. The influence of acoustic and semantic similarity on long-term memory for word sequences. Quarterly Journal of Experimental Psychology, 1966b, 18, 302-309.

BADDELEY, A. D., & DALE, H. C. A. The effect of semantic similarity on retroactive interference in long- and short-term memory. Journal of Verbal Learning & Verbal Behavior, 1966, 5, 417-420.

BROWN, R., & McNEIL, D. The tip of the tongue phenomenon. Journal of Verbal Learning & Verbal Behavior, 1966, 5, 325-337.

BRUCE, D., & MURDOCK, B. B. Acoustic similarity effects in memory for paired associates. Journal of Verbal Learning & Verbal Behavior, 1968, 7, 627-631.

CONRAD, R., & HULL, A. Information, acoustic confusion and memory span. British Journal of Psychology, 1964, 55, 429-432.

Table 3
Percentage Learning and Retention Scores for Experiment 1

Group	N	OL	IL	Retest	% Loss
A - B	13	81.73	90.38	58.65	33.92
X - Y	13	81.73	95.19	50.00	40.46
Mean	26	81.73	92.78	54.33	37.19
A - Y	13	81.73	95.19	57.69	33.38
X - B	13	81.73	96.15	58.65	30.62
Mean	26	81.73	95.67	58.17	32.00

Table 4
Percentage Learning and Retention Scores for Experiment 2

Group	N	OL	IL	Retest	% Loss
A - B	11	90.91	75.00	61.36	33.70
X - Y	11	90.91	78.41	55.68	41.00
Mean	22	90.91	76.70	58.52	37.36
A - Y	11	90.91	78.41	63.64	31.45
X - B	11	90.91	81.82	71.59	23.36
Mean	22	90.91	80.11	67.61	28.41

DALE, H. C. A. Retroactive interference in short term memory. Nature, 1964, 203, 1408.
DALE, H. C. A. Semantic similarity and the A-B, C-D, paradigm in STM. Psychonomic Science, 1967, 9, 79-80.
DALE, H. C. A., & GREGORY, M. Evidence of semantic coding in STM. Psychonomic Science, 1966, 5, 75.
UNDERWOOD, B. J., & GOAD, D. Studies of distributed practice: I. The influence of intra-list similarity in serial learning. Journal of Experimental Psychology, 1951, 42, 125-134.
WICKELGREN, W. A. Acoustic similarity and retroactive interference in short-term memory. Journal of Verbal Learning & Verbal Behavior, 1965, 4, 53-61.
WOODWORTH, R. S. *Experimental psychology.* New York: Henry Holt, 1938.

NOTE

1. The authors gratefully acknowledge financial support from the British Medical Research Council.

Effects of rewarding children's high- and low- amplitude motor responses[1]

GORDON N. CANTOR and JOHN H. WHITELEY, Institute of Child Behavior and Development, University of Iowa, Iowa City, Iowa 52240

Forty-eight preschool-age Ss provided baseline amplitude data by striking a target with a mallet seven times. Subsequently, a series of 36 training trials was administered, with half the Ss being rewarded for response amplitudes exceeding their baseline scores, the remaining Ss being rewarded for amplitudes of lower magnitudes than their baseline scores. Analyses of raw response amplitude scores and number of rewarded responses revealed that both groups learned in comparable fashion to respond in accordance with the reward contingency under which they were run. A difference score analysis yielded similar results, except for the first trial block. Reference is made to some related child studies and to implications regarding motivational vs associative interpretations of high magnitude responding.

Although vigorous responding is frequently interpreted as reflecting a state of heightened drive in the organism, writers have for many years now cautioned that intense responding may, in fact, be attributable in some instances to the organism's prior learning history (e.g., Skinner, 1938; Brown, 1961). Thus, the individual who fails to achieve a desired effect by means of a low-amplitude response may resort to intense responding—not because of a heightened motivational state (e.g., "frustration"), but because he has learned in the past that vigorous responding frequently "works" when relatively mild responding does not.

The proposition that differential *speeds* of responding may reflect habit variations has been given formal recognition and support in the theorizing and empirical rat data of Logan (1960). In a recent paper, Cairns & Proctor (1968) reported that 8- to 12-year-old children's response *latencies* in a button-pressing task tended to be shortened when relatively rapid responses were rewarded, whereas the latencies tended to be lengthened when relatively slow responses were rewarded. Using a differential conditioning paradigm, Longstreth & Gilbert (1969) have demonstrated that 7- to 9-year-old children can learn to push a joystick with relatively strong *amplitudes* in response to one CS and with relatively weak amplitudes in response to a second CS. The purpose of the present study was to determine if comparable between-Ss effects could be demonstrated using preschool-aged children and response amplitude as the dependent variable.

SUBJECTS

Forty-eight children enrolled in the University of Iowa Preschools served as Ss. They ranged in age from 4 years 0 months to 5 years 8 months, with a mean of 4 years 11 months. Twelve males and 12 females were assigned to each of two treatment groups: (a) Hi Group—rewarded for vigorous responding; and (b) Lo Group—rewarded for weak responding.

APPARATUS

The apparatus and equipment included a wooden mallet, a padded octagonal metal plate that served as a response target, an aluminum marble delivery chute feeding into a transparent receptacle, and a supply of white marbles. The metal plate, padded with foam rubber and covered by a black rubber surface, was mounted on the top surface of a gray wooden box. Two strips of ¾-in. white masking tape formed an "X" across the center of the target. The target was mounted toward the front end of the box on a Statham Model UC3 transducer connected with a Model UL4-100 load cell accessory. An Eico-2 buzzer, located on the top of the box and 7½ in. from the target, served as S's signal to respond. Response amplitudes were recorded on an Offner Dynograph Amplifier Recorder.

PROCEDURE

Each S, run individually, was first seated in front of the target with the mallet in his preferred hand. He was then told to listen for a buzzer to come on, at which time he was to strike the "X" with the mallet. S was informed he could hit the target "hard" or "soft" (E demonstrated both). He was then asked to "hit the 'X' *hard*" and then to "hit the 'X' *soft*" for practice. Seven nonrewarded practice trials were subsequently administered, with S responding to buzzer onset. The median amplitude obtained from the resulting seven scores was designated as S's basal level.

Following the seven baseline responses, S was told that he could win marbles by striking the target and, if he won "many" marbles, he could trade them in for a toy at the end of the session. He was told to check the marble receptacle (located above the rear end of the box) after each response to see if he was obtaining a marble on that trial. Finally, he was told, "Remember, you can hit the 'X' as hard or as soft as you like." Following the first rewarded response, E said, "You got a marble that time!" Following the first nonrewarded response, E said, "You didn't get a marble that time!"

The E initiated each trial by throwing a switch that activated the buzzer. Response amplitude, based on amount of pen deflection on the polygraph record, was measured to the nearest millimeter. A 3-mm pen deflection automatically terminated the buzzer, but deflections of less than 3 mm necessitated E's terminating the buzzer manually. Since deflections of less than 1 mm were not perceptible to E, it was possible for S to make one or more extreme low-magnitude responses on a given trial before responding with sufficient magnitude to effect buzzer offset.

Each S received 36 training trials, the intertrial interval (buzzer offset to onset) being approximately 10 sec. For Ss in the Hi group, a marble was delivered whenever the amplitude exceeded S's basal amplitude; for those in the Lo group, marble rewards occurred only when the amplitude was less than S's basal amplitude. A variable interval not exceeding 4 sec occurred between S's response and marble delivery. The marbles were delivered manually by E. The E, the upper end of the marble chute, and all of the remaining equipment (save for the target, the supporting box, and the remainder of the marble delivery apparatus) were hidden from S's view by a cloth screen.

Each S was given a toy reward at the close of the session, regardless of his performance.