

Recognition: Memory for words correctly heard in noise

PATRICK RABBITT

APPLIED PSYCHOLOGY RESEARCH UNIT, CAMBRIDGE, ENGLAND

The increase in effort necessary to correctly identify words over degraded communications channels has been shown to be reflected in lowered efficiency on simultaneously-performed non-verbal secondary tasks. Two experiments show that a similar loss of efficiency may be observed for operations performed on the material shadowed. Recognition memory is poorer for words correctly shadowed over a degraded channel.

Levels of noise acceptable for speech-transmission are frequently estimated by intelligibility tests: that is, by the accuracy with which human listeners recognize words transmitted through a given communications system. Judged by these criteria, listeners can tolerate marked circuit degradation before showing performance decrements. The meaningfulness of such estimates has often been questioned, and Broadbent (1958) has shown that high intelligibility test scores may sometimes be achieved on degraded channels only at the cost of increased effort by the listener. When Ss shadowed degraded speech as a primary task, degradation which did not increase the number of shadowing errors was nevertheless reflected by reduced efficiency at simultaneously-performed secondary tasks.

In practice people do not use communications links, such as telephones, to transmit word-lists but rather to carry out some purposeful negotiation. By analogy with Broadbent's work, cognitive aspects of such negotiation may be regarded as secondary to the vital main task of hearing what is said over the circuit. It is thus possible that there are levels of circuit degradation at which listeners can accurately recognize all the words said to them, but only at the cost of extra effort which reduces the "spare capacity" (Broadbent, 1958; Brown & Poulton, 1961) available to perform cognitive operations on the material heard.

This cannot be directly argued from Broadbent's data, since the secondary tasks he used were non-verbal, and quite unrelated to Ss memory or comprehension of the words which he shadowed. A first experiment was therefore made as a direct test of the effects of circuit degradation on Listener's ability to carry out a simple operation on the words he shadowed—in this case to remember them in order to recognize them again on a subsequent presentation.

EXPERIMENT I

Method

To be remembered was a list of 20 4-letter nouns drawn from the Thorndike-Lorge (1944) AA vocabulary read at a rate of 1/1.2 sec. This list was followed by a

10 sec. pause, after which a second list of 100 4-letter AA nouns was presented at the same rate. The 20 nouns of the stimulus list appeared again in the second list at random intervals and in random order. Each S repeated aloud every word in both lists as soon as he heard it. He was told that words in the first list were to be remembered, and that he was to signal their re-appearance in the second list by tapping a microphone base with a pencil. Ss knew that all words in the first list would be repeated.

Ss' responses (in list 1 verbal repetitions and in list 2 repetitions and taps) were tape-recorded and subsequently checked against a printed protocol. Separate groups were tested under two circuit conditions: Degraded (D) and Good (G). In both conditions the same tape was played at the same intensity on a Ferrograph tape-recorder through an external amplifier and speaker unit. In the G condition fidelity was as good as this system allowed. In the D condition the signal was fed through a G.P.O. Modulated Noise Unit. This device added pulse-modulated white noise (band width 3.5 KC/sec.) to the speech signal and maintained a constant relative intensity 10 dB below the intensity of the speech.

In both conditions Ss were accepted for the experiment only if they correctly reported back every word in both lists. In the D condition this criterion was met by 29 out of 35 Ss tested, and in the G condition by 17 out of 20 Ss tested. Ss were housewives aged from 19 to 53 years (\bar{X} 39 years) drawn from the A.P.R.U. civilian Subject Panel.

Results

Table 1 gives the mean number of correct responses and false positives made in each condition. By *t*-tests no significant differences were found between the number of correct responses recorded in each condition ($p > 0.25$). The indices d' and β were derived from the group-mean hit rates and false positive rates, and are

Table 1. Mean number correct and false positive scores, with calculated d' and β for four groups of subjects experiencing different conditions of channel degradation during a recognition memory experiment.

	\bar{X} Correct	\bar{X} False +	\bar{X} d'	\bar{X} β
Experiment I				
G Condition (N=17)	12.705	2.204	2.241	5.686
D Condition (N=29)	12.172	3.924	1.930	3.765
Experiment II				
GD Condition (N=12)	11.590	2.500	2.051	5.450
DG Condition (N=12)	12.083	4.333	1.870	3.518

also presented. Five Ss in the G group made no false positives. In order to test for significant differences between conditions three different techniques were used to assign appropriate d' and β values to them: (1) They were assigned values computed from the group-mean hit-rates and false positive rates. (2) They were assigned the optimum values of d' and β consistent with the hit-rates they achieved. (3) They were each arbitrarily assigned 0.5 false positives (Broadbent, 1963). Mann-Whitney U tests applied to data computed in all three ways uniformly suggested that values for d' ($p < 0.01$) and for β ($p < 0.0025$) were larger for the G than for the D condition.

Thus, levels of noise which did not affect intelligibility nevertheless reduced the efficiency of recognition and the level of confidence with which judgments were made. A logical next experiment was to determine whether performance was more affected by circuit degradation during a list presented for later recognition or by circuit degradation during the presentation of the recognition list.

EXPERIMENT II

Method

The lists were presented to two further groups with identical instructions. One group experienced circuit degradation during presentation of the stimulus list, and good fidelity thereafter (DG); the second group experienced the reverse condition (GD). In the GD condition 12 out of 15 Ss and in the DG condition 12 out of 14 Ss met the criterion of repeating back correctly every word in both lists. Ss were Royal Navy ratings aged 17.5 to 25.5 years (\bar{X} 23.4).

Results

Mean correct scores (Table 1) were not significantly different by t-test ($p > 0.05$) but the DG group made more false-positives ($p < 0.01$). Only one S in the GD group made 0 false-positives, and no treatment of his score affected the direction of a trend. Mann-Whitney U tests showed that values of d' ($p < 0.05$) and of β ($p < 0.025$) were larger for the GD than for the DG group. Recognition is therefore more affected by degradation during presentation of material for subsequent recall. Because of differences in S populations it is not useful to compare the performance of these groups with those tested in the first experiment.

Discussion

Since Ss were selected for their ability to correctly repeat all words presented, the effects of degradation on recognition scores cannot simply mean that they did not hear what was said to them. At least three explanations are available:

These data may be interpreted as another demonstration of Broadbent's (1958) effect. On this paradigm the increased effort necessary to decode words presented through a degraded circuit reduces the "spare capacity"

available for the task of remembering, and perhaps retrieving, the words. The rate at which the words were presented (1/1.2 sec.) was deliberately a little faster than most Ss found comfortable. Ss might therefore be expected to be sensitive to any increase in the time required to recognize a word when it was presented through a degraded system. It will be noted that this hypothesis suggests that the act of remembering words takes up time, or channel-capacity, or both. In other words the operation of remembering is supposed to compete (for time, or channel-capacity) with the operations necessary to understand the words.

A related explanation derives from Chistovitch et al (1960) who found that Ss repeating back continuous speech could choose to shadow "with or without understanding." When shadowing "without understanding" Ss gave accurate phonetic reproductions of speech-sounds at very short latencies (150 msec. to 250 msec.) but subsequently could not recall the material they shadowed. When shadowing "with understanding" they repeated the speech at longer latencies (250+ msec.) and could better recall the material they had shadowed. It is arguable that in these experiments a combination of noise and pacing caused Ss to occasionally lapse into a "fast" strategy, with consequent decrement to their memory of what they shadowed.

Ss performing the task with added noise were unanimous, when questioned, in expressing dislike of the experiment. They invariably said that the noise made the experiment irritatingly difficult and was in some sense "unfair." This attitude might account for the differences in d' and β observed in the first experiment, but is less convincing as an explanation for the differential effects of noise on recall and recognition lists.

Although these data do not allow us to select between these explanations they have clear practical implications: Communications-links are used for a purpose, not simply to transmit lists of words. Levels of degradation which do not affect intelligibility may nevertheless affect the efficiency with which Ss carry out certain operations on words transmitted through the circuit. Memory, tested by recognition, is shown to be a case in point.

References

- Broadbent, D. E. *Perception and communication*. London: Pergamon Press, 1958. Pp. 68-80.
- Broadbent, D. E., & Gregory, Margaret. Vigilance considered as a statistical decision. *Brit. J. Psychol.*, 1963, 54, 309-323.
- Brown, I. D., & Poulton, E. C. Measuring the 'spare' capacity of car drivers by a subsidiary task. *Ergonomics*, 1961, 4, 35-40.
- Chistovitch, L. A., Alyakrinskii, V. V., & Abul'yan, V. A. Vremennye zaderzhki pri povtoreнии slishimoi rechi. *Voprosi Psikhologii*, 1960, 6, 64-70.
- Thorndike, Edward L., & Lorge, Irving. *The teachers word book of 30,000 words*. New York: Teachers College, Columbia University, 1944.