

Inhibition of the automatic storage of speaker's voice

RALPH E. GEISELMAN

University of California, Los Angeles, California 90024

The voice-connotation hypothesis of Geiselman and Bellezza (1976, 1977) states that a speaker's voice is sometimes remembered without intent because the connotation of the voice automatically influences the meaning of what is said. Results from the present experiment suggest that subjects have the option to prevent the speaker's-voice attribute from being stored with the contents of what is said when such processing would interfere with other cognitive operations.

Pear (1931) wrote in his book, *Voice and Personality*, that "whatever the physical sounds produced by a voice, the effect upon the hearer depends largely upon his own past experience . . . In judging a voice, we may—usually unconsciously—be reminded of another earlier voice, significant to us in the past, and our judgment may thus be powerfully influenced" (pp. 34, 75). It has been contended that this connotation of a voice influences the perceived meaning of what is said and that that is why a speaker's voice is sometimes remembered without intent (Geiselman & Bellezza, 1976, 1977). For example, a sentence spoken by a male may have a somewhat different meaning than the same sentence spoken by a female. Thus, the sex of the speaker of the sentence can be recalled at a later time. In the absence of such semantic recoding of voice by the hearer, the characteristics of the voice are lost rapidly, in accordance with Crowder and Morton's (1969) concept of precategory acoustic storage (PAS).

In support of the voice-connotation hypothesis, Geiselman and Bellezza (1976, 1977) gathered evidence from five experiments demonstrating that: (1) the retention of the sex of a speaker requires no cognitive processing beyond that needed to encode the content of what is said; that is, the speaker's voice seems to be processed automatically; (2) the sex of a speaker is not remembered above chance level if what is said contains a competing masculine or feminine denotative meaning; and (3) the sex of a speaker influences, slightly, the judged "potency" of what is said. The empirical support for each of these contentions, as well as their implications for the voice-connotation hypothesis, are discussed in turn.

First, Geiselman and Bellezza (1976) presented 20 unrelated sentences to subjects with a random 10 of the sentences spoken by a male and the remaining sentences

spoken by a female. The subjects were told that they would be expected to write down as many of the sentences as they could remember, but no mention was made of the different voices. After a free recall test for the sentences, the subjects were given a list of the 20 sentences and were asked to indicate the sex of the speaker for each sentence. On average, the subjects could identify the sex of the speaker above chance level (.68), and their performance was not improved significantly by prior instructions to intentionally remember the voice (.72). In addition, the level of sentence recall for the subjects in the incidental voice-retention condition was virtually identical to that for other subjects who participated in a condition where the sentences were spoken by a single speaker (.30 vs. .32). Therefore, the retention of speaker's voice appears to occur automatically and without requiring any cognitive processing beyond that needed to encode the sentences themselves. In contrast, other subjects in this experiment could not remember whether a sentence was presented from the left or right side of the room unless they had intent to remember the source location, and only then at the expense of sentence recall.

Second, Geiselman and Bellezza (1977) found that, when each sentence in a list of unrelated sentences contained a male or female agent (e.g., The gentleman entered the house), the sex of the speaker for each sentence was correctly identified at a level that would be expected on the basis of chance (.51). When these agents were replaced by neutral agents (e.g., youngster, employee), the level of voice retention increased to .69. This suggests that the connotation of a voice is less likely to influence the meaning of what is said when the type of information contained in the voice is already present in the content of the message. (In this case, the type of information is a nonneutral value on the masculine-feminine dimension.) In support of this view, the subjects were found to have guessed the sex of the speaker in accordance with their recollection of the sex of the agent in a sentence.

Third, Geiselman and Bellezza (1977) asked subjects

Requests for reprints may be sent to Ralph E. Geiselman, Department of Psychology, University of California, Los Angeles, California 90024.

to rate 20 sentences containing neutral agents on four semantic-differential scales; each scale was associated with the "potency" factor (Osgood, Suci, & Tannenbaum, 1957). In agreement with the voice-connotation hypothesis, the subjects rated sentences that were previously spoken by a male to be slightly more potent than sentences spoken by a female. The speaker's voice did appear to influence the meaning of the sentences.

In sum, this collection of results seems to form a clear picture of the incidental retention of speaker's voice in terms of the voice-connotation hypothesis. Unfortunately, a few subsequent (unpublished) experiments have failed to show incidental retention of voice above chance level, even for sentences containing neutral agents. All of these subsequent experiments had one factor in common: The list of sentences, or words, contained twice as many items with prior masculine or feminine meanings than items that were neutral. In previous experiments, the items in a list were either all neutral or all had masculine or feminine values, or the list contained an equal number of neutral and masculine-feminine items. One explanation of the new results is as follows. When a sufficient number of the sentences in a list contain male or female agents, the subject must not become confused as to whether the agent associated with a particular predicate was in fact male or female. In addition, when the sex of the speaker is varied orthogonally to the sex of the agent, the connotation of the voice may further interfere with the subject's efficiency in encoding the sex of the agent. To combat the latter source of interference, the subject may have the option to block most, if not all, of the "automatic" associative analyses of the speaker's voice from being carried out. Consequently, when the content of the majority of sentences is nonneutral, the speaker's-voice attribute may not be stored specifically in episodic memory with the semantic content of any of the sentences; therefore, the voice attribute for each sentence is forgotten within a few seconds.¹

To test this hypothesis in the present experiment, three groups of subjects were presented a list of sentences with either $\frac{3}{4}$, $\frac{1}{2}$, or $\frac{1}{4}$ of the sentences containing nonneutral (male or female) agents. The remaining sentences in each list contained agents that were relatively neutral with respect to gender. In addition to the prediction that incidental voice retention and the proportion of nonneutral agents in the list should be inversely related, the inhibition hypothesis makes a further prediction involving the input position of a sentence in the list. Since the male-female dimension within the list should not be detected until a few sentences have been presented, the speaker's voice for sentences that are presented early in the list (and contain neutral agents) should be remembered better than chance, regardless of the proportion of sentences containing nonneutral agents in the list. Of course, the voice-connotation hypothesis predicts that speaker's

voice should not be remembered for sentences with nonneutral agents, regardless of the input position. This is because the representation of the sentence on the masculine-feminine dimension would be determined by the sex of the agent, which is more central to the meaning of the sentence than the connotation of the speaker's voice (Geiselman & Bellezza, 1977).

METHOD

Subjects

The subjects were 30 male and 30 female undergraduate volunteers from the introductory psychology course at the University of California, Los Angeles.

Materials and Apparatus

All subjects were presented 16 simple active sentences auditorily, each of which was constructed in the past tense. A random eight of the sentences were spoken by a male and the remaining eight sentences were spoken by a female. The sentence predicates were constructed to be relatively neutral with respect to the masculine-feminine dimension, such as "spent the money" and "broke the pencil." The sentences were presented with 5-sec intersentence intervals using a Sony Quadradial tape recorder in conjunction with two AR-5 loudspeakers positioned at the front of the experimental room.

For one group of subjects, 12 of the sentences contained male or female agents such as "husband," "actress," and "lady"; the other 4 sentences, which were intermixed, contained relatively neutral agents such as "citizen," "employee," and "winner." For a second group of subjects, a random eight of the sentences contained male or female agents and the other eight sentences contained neutral agents. For a third group, 12 of the sentences contained neutral agents and the other 4 sentences, which were intermixed, contained male or female agents. For each of the three groups of subjects, half of the nonneutral agents were male and half were female, and half of the sentences containing each type of agent were spoken by the male and half were spoken by the female. As counterbalancing measures, each sentence was spoken by each speaker for an equal number of subjects in each group, and each sentence predicate appeared with either a neutral or nonneutral agent for some subjects in each group.

Procedure

The subjects were tested in groups of approximately five subjects each. All subjects were told that they would be presented a list of unrelated simple sentences and that they would be expected to write down these sentences at the end of the experiment in any order that they wished. No mention was made of the different speakers or the ensuing voice-identification test. Twenty of the subjects heard a tape where 12 of the sentences contained neutral agents, 20 heard a tape where 8 of the sentences contained neutral agents, and 20 of the subjects heard a tape where only 4 of the sentences contained neutral agents. Immediately after the presentation of the sentences, all subjects were asked to solve some deductive-reasoning problems for 2 min, followed by 3 min of sentence free recall. Then the subjects were provided a sheet of paper containing the 16 sentences that they had heard, randomized with respect to input serial position. They were asked to simply put a check next to each of the sentences that they remembered hearing. Finally, the subjects were to indicate whether the speaker for each sentence had been male or female. This voice-identification test was conducted only after the sentence-recognition task was completed, and the sex of the speaker was to be indicated even for those sentences that the subject did not remember hearing.

Design and Analysis

There were three dependent variables: the proportion of sentences recalled, the proportion of sentences recognized, and the proportion of sentences for which the sex of the speaker was correctly identified. Except for changes of tense, the recall test was a verbatim recall test. The data corresponding to each of the dependent variables were analyzed using a 3 by 2 by 4 design, with the factors being list composition (12 nonneutral agents, 4 neutral agents; 8 nonneutral agents, 8 neutral agents; 4 nonneutral agents, 12 neutral agents), type of agent (nonneutral, neutral), and list quarter. The only between-subjects factor was list composition.

RESULTS AND DISCUSSION

Sentence Recall

The mean proportion of sentences recalled as a function of type of agent and list quarter is shown in Figure 1. The results of an analysis of variance confirmed that both the main effect of type of agent and the main effect of list quarter were significant [$F(1,57) = 12.5$, $p < .001$, $MSe = .004$ and $F(3,171) = 16.9$, $p < .001$, $MSe = .003$, respectively]. This pattern of results was unaffected by the composition of the list. The difference between the recall of sentences with neutral and nonneutral agents has been observed before (Geiselman & Bellezza, 1977), but its cause is not yet understood. It is not the case that more attention was allocated to sentences with nonneutral agents, because the difference is observed even when type of agent is manipulated as a between-subjects factor. Geiselman and Bellezza (1977) have also shown that nonneutral agents do not, in general, have higher imagery ratings than neutral agents. The present results suggest that the effect is a reliable one, but no new light has been shed upon the cause of the effect.

Sentence Recognition

Although the analysis of variance conducted on the recognition data showed no significant effects (all $F_s < 1$), the proportion of sentences checked as "remembered" ranged from .96 to 1.00 across cells in the design. Therefore, the recognition results are of little

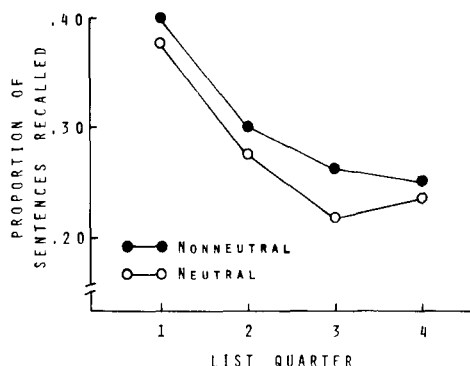


Figure 1. The proportion of sentences recalled as a function of type of agent (neutral or nonneutral) and list quarter.

Table 1
Probability of Voice Retention as a Function
of Type of Agent and List Composition

Type of Agent	List Composition (Neutral-Nonneutral)		
	4-12	8-8	12-4
Neutral	.54	.67	.71
Nonneutral	.51	.52	.54

value, except to note that most subjects felt that they remembered hearing nearly all of the sentences.

Voice Retention

The voice-retention data collapsed across the list-quarter factor are shown in Table 1 as a function of type of agent and list composition. As predicted by the voice-notation hypothesis, the speaker's voice was remembered more often for sentences containing neutral agents than for sentences containing nonneutral agents [$F(1,57) = 32.3$, $p < .001$, $MSe = .005$]; as predicted by the inhibition hypothesis, the difference between voice retention for neutral and nonneutral sentences was greater when fewer sentences in the list contained nonneutral agents [$F(2,57) = 49.5$, $p < .001$, $MSe = .005$]. With three-fourths of the sentences containing nonneutral agents, voice retention for the sentences containing neutral agents did not exceed that for the sentences containing nonneutral agents. It is also instructive to note that, as in previous studies (Geiselman & Bellezza, 1976, 1977), voice retention for nonneutral sentences did not significantly exceed chance performance (.50), regardless of the list composition. Thus, if the content of a sentence contains a nonneutral masculine or feminine value, then the connotation of the speaker's voice is unlikely to influence the perceived meaning of the sentence, and the voice attribute is therefore not remembered.

The voice-retention data are also shown in Figure 2, but as a function of list quarter. As was just mentioned, voice retention for sentences containing nonneutral agents is always near chance, and Figure 2 indicates that this result is independent of the position of such a sentence in the list. However, voice retention for neutral sentences did vary as a function of both list quarter and list composition [the three-way interaction was significant, with $F(6,171) = 2.80$, $p < .025$, $MSe = .005$]. In brief, when there were several nonneutral sentences in the list, voice retention for the neutral sentences decreased with input position. Further, the decline in voice retention appears to have begun earlier in the list as the proportion of sentences with nonneutral agents increased. These results provide strong support for the inhibition hypothesis.

Even though the sex of a speaker appears to typically be coded in memory with each sentence without additional cognitive processing (Geiselman & Bellezza,

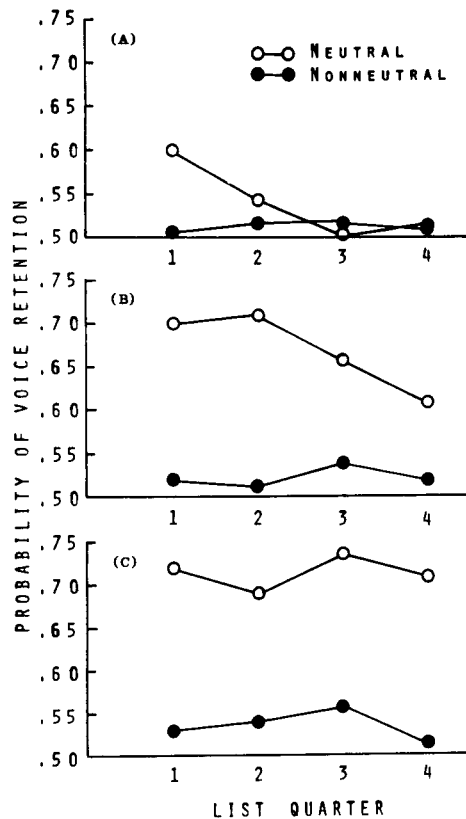


Figure 2. The probability of voice retention as a function of type of agent (neutral or nonneutral) and list quarter. Panel A corresponds to the condition in which there were 4 neutral sentences and 12 nonneutral sentences; Panel B corresponds to the condition in which there were 8 sentences of each type; Panel C corresponds to the condition in which there were 12 neutral sentences and 4 nonneutral sentences.

1976), the present results suggest that subjects have the ability to inhibit such automatic storage of voice when that coding is seen to interfere with other cognitive operations (e.g., the processing of the sex of the agents). Therefore, the proportion of sentences with nonneutral values on the masculine-feminine dimension should be considered when investigating the incidental retention of speaker's voice. In the same vein, Taylor (1977) demonstrated that letters surrounding a to-be-processed letter in a visual display are typically encoded automatically. However, subjects have considerable attentional control over whether they process the context letters when the response associated with the context letters is repeatedly inconsistent with the response required for the target letter. It is not clear whether these apparent instances of inhibition occur with effort or whether such defenses are conducted automatically by the system, but the present results tend to favor the latter possibility. If the process of blocking the associa-

tive analysis of speaker's voice requires a significant amount of processing capacity, then the probability of recalling a sentence with either type of agent should have decreased as the proportion of sentences with nonneutral agents increased. This is because the detection of a masculine-feminine dimension within the content of the list of sentences, and thus the inhibition of the storage of speaker's voice, should have been more likely with a greater proportion of nonneutral sentences. However, the levels of recall shown in Figure 1 did not change as a function of list composition ($F < 1$). Thus, the inhibition was conducted without any decrease in sentence recall, suggesting that little processing capacity was required. This is in sharp contrast to a subject's apparent inability to block potentially confusing information in other situations, as with Stroop-type tasks. Therefore, the generality of the inhibition effect needs to be explored.

The results reported here agree with the modified version of the levels-of-processing framework (Lockhart, Craik, & Jacoby, 1976), which states that items can be processed at a given depth without first completing analyses at shallower depths. When the proportion of nonneutral sentences was high, it is possible that the sentence contents were assigned meaning without a full analysis of the speaker's voice characteristics. The results are also in agreement with Treisman's (1964) conception of levels of analysis, in which a series of tests is said to be performed on the incoming sense data to allow or prevent the passage of information to deeper levels of analysis.

REFERENCES

- CROWDER, R. G., & MORTON, J. Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 1969, 5, 365-373.
- GEISELMAN, R. E., & BELLEZZA, F. S. Long-term memory for speaker's voice and source location. *Memory & Cognition*, 1976, 4, 483-489.
- GEISELMAN, R. E., & BELLEZZA, F. S. Incidental retention of speaker's voice. *Memory & Cognition*, 1977, 5, 658-665.
- LOCKHART, R. S., CRAIK, F. I. M., & JACOBY, L. L. Depth of processing, recognition and recall. In J. Brown (Ed.), *Recall and recognition*. New York: Wiley, 1976.
- OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana: University of Illinois Press, 1957.
- PEAR, T. H. *Voice and personality*. New York: Wiley, 1931.
- TAYLOR, D. A. Time course of context effects. *Journal of Experimental Psychology: General*, 1977, 106, 404-426.
- TREISMAN, A. Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior*, 1964, 3, 449-459.

NOTE

1. I wish to thank Francis S. Bellezza for his help in developing the inhibition hypothesis.

(Accepted for publication February 26, 1979.)