

Voice Recognition: Effects of Orienting Task, and a Test of Blind versus Sighted Listeners



Eugene Winograd; Nancy H. Kerr; Melanie J. Spence

The American Journal of Psychology, Vol. 97, No. 1. (Spring, 1984), pp. 57-70.

Stable URL:

<http://links.jstor.org/sici?sici=0002-9556%28198421%2997%3A1%3C57%3AVREOOT%3E2.0.CO%3B2-B>

The American Journal of Psychology is currently published by University of Illinois Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/illinois.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

Voice recognition: Effects of orienting task, and a test of blind versus sighted listeners

EUGENE WINOGRAD
Emory University

NANCY H. KERR
Oglethorpe University

MELANIE J. SPENCE.
University of North Carolina at Greensboro

Three experiments on memory for unfamiliar voices are reported. In the first two experiments, the type of encoding was manipulated by orienting tasks. In each study, one group rated the pleasantness of the voices, another group made judgments of pitch, and a third group performed no orienting task. Recognition memory was not correlated with encoding condition in either experiment. In Experiment 2, voice memory was better when the message remained constant from study to test than when it was changed. In Experiment 3, a group of blind adults was tested with no orienting task. The blind subjects' memory was equivalent to that of the appropriate sighted group of Experiment 2. It is proposed that voices represent a psychologically unitary stimulus that cannot be encoded with differing degrees of elaboration; therefore, voices are not sensitive to orienting tasks in the way that other classes of events, such as words and faces, are.

How do we remember voices? Like faces and gaits, voices are peculiar to their owners. Ordinary experience suggests that we store information about voices independently of what the voices say. For example, upon hearing the voice of a well-known actor in a commercial extolling the virtues of a particular airline or beer, we often recognize it as familiar and search our memory for information as to the speaker's identity. Clearly, just as a clarinet is a clarinet whether it is playing Mozart or Schubert, my voice is recognizably mine whether I am lecturing or telling a joke. From another room, or on the telephone, we quickly identify the voice of a familiar speaker.

Before going farther, we must distinguish between memory for an already familiar voice, as in the examples just given, and the study of how an unfamiliar voice becomes familiar. The present research concerns the latter process, the encoding of information about new voices. It is an inquiry into the origins of familiarity and asks how factors known to affect memory for many other kinds of events affect memory for voices.

AMERICAN JOURNAL OF PSYCHOLOGY
Spring 1984, Vol. 97, No. 1, pp. 57-70

© 1984 by the Board of Trustees of the University of Illinois

One of the most potent manipulations memory researchers have found in studying encoding processes is the nature of the cognitive activity engaged in, or the kind of stimulus analysis carried out, by subjects when they encounter material. In a word-memory experiment, for example, Craik and Tulving (1975) reported recognition memory scores four to five times greater following judgments about category membership (e.g., "Is it a fish?") than judgments about the appearance of the word (e.g., "Is it in capital letters?"). Substantial effects of the nature of the orienting task on retention have also been found for visual events such as pictures of unfamiliar faces (Bower & Karlin, 1974; Winograd, 1978) and landscapes (Miller, Mueller, Goldstein, & Potter, 1978). In all this work, orienting tasks requiring judgments about physical, or intrinsic, aspects of the stimulus (e.g., "Does he have a big nose?") have been associated with poorer retention than tasks focusing on extrinsic aspects of the stimulus (e.g., "Does he look friendly?"). In the only experiment we know of that examined the effects of orienting tasks on auditory memory, Ferrara, Puff, Gioia, and Richards (1978) found that naturalistic sounds such as the mooing of a cow and the ringing of a telephone were more likely to be recalled following judgments of pleasantness than judgments of whether the sound was louder in the right or left ear. Although this was the expected outcome, because loudness judgments concern a physical property, the result can as readily be interpreted as pertaining to verbal memory as to auditory memory. The stimuli used by Ferrara et al. (1978) were purposely chosen so that they could be named; indeed, the use of a recall measure is predicated on the availability of names for the stimuli.

The study of recognition memory for unfamiliar voices commends itself as a relatively pure auditory memory task. For the most part, research on voice recognition has been carried out within a psychoacoustic framework and, therefore, has used rather short, reasonably well-defined speech signals. Performance has been shown to be influenced by such basic parameters as set size (Carterette & Barnebey, 1975), duration of the speech sample (Pollack, Pickett, & Sumbly, 1954), and retention interval (Carterette & Barnebey, 1975; McGehee, 1973). When reading this literature, the student of memory is struck by the low recognition scores found. For example, Carterette and Barnebey reported that the mean proportion correctly recognized of a set of eight voices was only .59, with chance performance equal to .50. In a recent study showing that recognition memory for unfamiliar voices improves as children grow older, Mann, Diamond, and Carey (1979) tested immediately after presentation of either one or two target voices and found recognition proportions of .59 and .68 (interpolated from their Figure 2), respective-

ly, when the target voice was repeated saying a different message; again, the chance level was .50. It should be noted that both of these studies used brief utterances such as "a year ago" and "we call you." In the present work, longer passages were used.

The research reported here was concerned with three questions. First, is memory for voices, about which relatively little is known, sensitive to the same kinds of encoding activities as memory for other events? Second, what is the contribution of the message to voice recognition? Third, do blind people differ from the normally sighted in a purely auditory memory task? The inclusion of the blind group will be discussed when Experiment 3 is presented.

There were three encoding conditions in Experiment 1. Subjects heard samples of 20 different voices and rated each voice either on a physical (pitch) or an evaluative dimension (pleasantness), or they listened without making any judgments. With faces (see Winograd, 1978, for a review), it has consistently been shown that judgments of physical features such as size of nose or curliness of hair lead to poorer recognition than do judgments along evaluative dimensions such as friendliness or likability. Similarly, judgments of pleasantness of words have facilitated memory when contrasted with decisions about structure (e.g., Hyde & Jenkins, 1969). The group with no specific orienting task was included for comparison. With faces, Warrington and Ackroyd (1975) and Smith and Winograd (1978) have found that subjects' memory in such a condition was equivalent to that found after judgments of physical features and inferior to memory for faces judged for friendliness. With naturalistic sounds, however, Ferrara et al. (1978) found that the group with no orienting task recalled as many sounds as the group that made judgments of pleasantness.

EXPERIMENT 1

METHOD

Subjects

Subjects were 54 Emory University undergraduate students of both sexes who were enrolled in introductory psychology classes and participated to fulfill course requirements. Subjects were randomly assigned to the three encoding groups, with 18 in each group.

Materials

Samples of 40 voices, 19 male and 21 female, were used in making up the two tapes. The speakers were graduate students and staff, but not faculty, of the Emory University Psychology Department, from 22 to 50 years old, with most in their twenties. It was assumed that the subjects, mainly

freshmen and sophomores, were unfamiliar with the voices on the tape. The study tape contained 20 voices, 9 male and 11 female, each reading the following passage: "Maybe you've never considered rayon as an alternative to cotton. Yet it is. That's because rayon looks, feels, performs, and blends much like cotton." The passage, selected from a magazine advertisement, took between 5–12 s to read, depending on the speaker. The test tape contained 40 voices, each reading the above passage. There were 20 voices from the study tape plus 20 new voices, 10 male and 10 female, randomly intermixed with the old, or repeated, voices. On the test tape, the old voices were recorded a second time and, therefore, were not identical copies of the recordings on the study tape. On both the study and test tapes, each voice started about 7 s after the previous one ended.

Two of the experimental groups used 5-point rating scales during the study phase, one for pitch, running from low to high, and one for pleasantness, from unpleasant to pleasant. On the recognition memory test, subjects were instructed to use the following 5-point confidence scale: 1. Sure I heard it; 2. Think I heard it, but not sure; 3. Not sure either way; 4. Think it's new, but not sure; and 5. Sure it's new.

Procedure

The experiment was conducted in small groups. Subjects in the Pitch group were told: "We are interested in how people perceive certain characteristics of voices, and we are particularly interested in the perception of pitch; that is, whether a voice is high or low in pitch." The instructions went on to note that there were no right or wrong answers, to emphasize that pitch should be rated relative to other voices of the same sex, to explain the 5-point rating scale, and to state that there would be a later memory test for the voices. The Pleasantness group was instructed: "We often form impressions of people from their voices alone as in phone calls and when hearing voices on the radio. This experiment is concerned with impressions we form of people by hearing them, and we are particularly interested in your impressions of pleasantness; that is, whether a person is perceived as pleasant or unpleasant by hearing his/her voice." The instructions went on, as with pitch, to explain the rating scale and to mention that memory would be tested. The Standard, or control, group made no ratings during study. They were asked to listen carefully and try to remember each voice.

The voice recognition test followed as soon as test instructions were read and the tape was advanced to the test voices. The test instructions described the 5-point confidence rating scale and stated that half of the voices were old and half were new. In summary, a given subject heard 20 voices and rated each of them for pitch or pleasantness, or did not rate at all, depending on the group the subject was in; this was followed by a test tape of 40 voices containing the 20 voices from the study tape intermixed with 20 new voices. All the speakers on both tapes read the same passage extolling the virtues of rayon.

RESULTS

The first column in Table 1 shows the mean confidence rating difference score for each encoding condition. This was obtained by calculating for each subject the mean rating for the 20 old voices on the test tape and subtracting it from the mean rating for the new voices. The larger the difference, the greater the discriminability of the old from the new voices; the maximum difference score possible was 4. Another way to show the outcome is to convert the rating scale into hit and false-alarm proportions. To this end, ratings of 1 and 2 were considered old judgments, and ratings of 3, 4, and 5 were considered new judgments. The mean proportions of hits and false positives defined this way are shown in columns 2 and 3 of Table 1. Analysis of variance yielded no significant outcome for either the rating difference measure or mean corrected recognition scores (hits minus false alarms). Indeed, the corrected recognition scores for the Pitch and Pleasantness conditions are identical at .20. Obviously, this is a difficult task.

The outcome of this study, that orienting task had no effect on voice memory, will be discussed following presentation of the results of Experiment 2, an extension of Experiment 1.

EXPERIMENT 2

In designing an experiment on voice recognition, a major decision facing the experimenter is whether the old voices on the test should repeat the same words or say something different. Although real life, with the obvious exception of radio and television commercials, infrequently presents the situation where a voice repeats the same message, we chose to keep the message invariant in Experiment 1 because our focus was on memory for voices per se. In Experiment 2, the role of the message in voice recognition was systematically included in a design that retained the three encoding conditions of the first experiment.

Table 1. Mean recognition scores: Experiment 1

Encoding condition	Rating difference	Hit proportion	False-alarm proportion
Pitch	.71	.56	.36
Pleasant	.73	.57	.37
Standard	.82	.56	.33

METHOD

Materials and subjects

We extracted 40 passages concerning a variety of topical subjects from copies of *Time* magazine. Each passage was a sentence or two in length and took about 10 s to read for almost all speakers. The study tape contained 10 male and 10 female voices reading 20 different messages in a random order. The test tape contained the voices of 40 individual speakers, half of each sex, including the 20 old, or repeated, voices from the study tape. Of the 20 old voices, 10 read the same message as on the study tape, and the other 10 old voices read new messages. All the speakers were adults chosen to be unfamiliar to the 72 undergraduates who served as subjects. In summary, on the test tape, half of the old voices read the same message that they had read earlier and half of the old speakers read new messages; similarly, half of the new speakers read old messages, and half read new messages. Two different test tapes were constructed so that messages and voices were not confounded. Each voice appeared in the same randomly determined serial position on both test tapes, but read an old message on one tape and a new one on the other.

Procedure

The 72 subjects were unsystematically assigned to one of three encoding conditions, Pitch, Pleasantness, or Standard, as in Experiment 1, with 24 subjects in each condition. They were tested in small groups. Within each condition, half were tested on one test tape and half on the other. All subjects heard the same study tape. It was emphasized that the task was to learn the voices and that voice recognition would be tested. As in Experiment 1, subjects in the Pitch and Pleasantness conditions rated the voices on these dimensions, and subjects in the Standard condition merely listened to the tape after being told to try to remember the voices. Immediately following presentation of the 20 voices on the study tape, the subjects were instructed regarding the test. Again, it was emphasized that their task was to decide whether each voice had occurred earlier. It was pointed out that there were old and new messages on the tape but that the memory test concerned voices, not messages. Subjects indicated their choice by circling *Yes* or *No* on an answer sheet, and indicated their confidence on a 3-point scale. The interval between successive passages on both the study and test tapes was approximately 7 s.

RESULTS

The hit and false-alarm proportions for the three encoding conditions are listed separately for old and new messages in Table 2. A hit is saying "old" to an old voice, whether or not the voice is repeating what it said before. The hit proportions are listed separately for old and new messages. The appropriate false-alarm proportion for the

Table 2. Mean recognition scores: Experiments 2 and 3

Encoding condition	Old message			New message		
	Hit proportion	False-alarm proportion	Ag	Hit proportion	False-alarm proportion	Ag
Experiment 2						
Pitch	.73	.46	.70	.59	.33	.69
Pleasant Standard	.80	.46	.74	.52	.33	.62
Experiment 3						
Blind subjects	.84	.45	.74	.60	.34	.69
		.52	.72	.52	.31	.66

old-message hits is the probability of an "old" response to a new voice saying an old message; the appropriate false-alarm proportion for the hit proportions for new messages is the probability of saying "old" to a new voice saying a new message. The derived measure of recognition memory, A_g , shown in each case, is a measure of the area under the memory-operating characteristic free of some of the assumptions associated with d' (see Pastore & Scheirer, 1974). The A_g values have a possible range of .5 to 1.0. A two-way analysis of variance using the A_g scores, with encoding condition and type of message as the factors, found a significant effect of message, $F(1, 69) = 5.87$, $MS_e = 0.156$, $p < .05$, but no effect of encoding condition and no interaction. The lack of an effect of orienting tasks on memory for voices found in Experiment 1 was found again, both when the voice was repeating the message and when it was saying something different.

The significant effect of message means that memory for voices was improved when an old voice was repeating itself relative to when it was saying something new. A similar finding has been reported by Mann et al. (1979). That the difference is significant using the A_g measure means that the large difference in hits favoring repeated messages is not due just to a tendency to say "old" to any familiar part of an event, be it voice or message. The difference in false positives does not offset the effect, as an explanation based on bias would assert. One does not need to appeal to semantic factors, or repetition of meaning, to explain the facilitation of voice memory when the message is repeated. We suggest that voice recognition is largely dependent on recognizing the features that are distinctive to the particular voice. By this argument, a voice reading a repetition of the same message is more likely to reproduce such distinctive features than when it is reading a relatively brief, new message. For example, if a speaker has an unusually sibilant /s/, recognition of the voice will be hindered if, on the test passage, there is no /s/ to aid recognition. Feature overlap is maximal, of course, when the same message is repeated.

Inspection of the false-alarm data revealed a much greater likelihood of calling a new voice "old" when the new voice was from a reader of the same sex as that of the person originally reading the message. Collapsed across encoding conditions, the mean proportion of false positives was .55 when the new voice was from a person of the same sex as that of the person who read that passage on the study tape, compared with a mean false-positive proportion of .33 when the new voice was from a person of the opposite sex. In other words, new voices reading old messages sound more familiar when

they are from persons of the same gender as that of the original readers. This finding of retention of a gender attribute for voices agrees with the findings of other researchers (Craig & Kirsner, 1974; Geiselman & Bellezza, 1976).

EXPERIMENT 3

There are two noteworthy aspects of the results of Experiments 1 and 2. First, the overall performance of subjects on the voice recognition task was rather low. Second, an orienting task, chosen for its similarity to those that have increased recognition for verbal and visual stimuli, did not enhance voice recognition. That voices are difficult stimuli to discriminate regardless of encoding strategy may explain both results, but it is also possible that the results may be related to our failure to introduce orienting tasks that call upon skills appropriate for a voice recognition task. The latter possibility raises questions about what the appropriate skills might be and who might possess them.

The most likely candidates to have developed a facility in auditory attention and perception are blind people. They are dependent on the auditory modality and are, therefore, likely to have learned to discriminate and encode auditory information with maximal efficiency. Blind subjects have, in fact, been shown to perform better than sighted subjects on auditory perceptual tasks such as signal detection in noise (Benedetti & Loeb, 1972) and the isolation and identification of simple auditory patterns (Witkin, Oltman, Chase, & Friedman, 1971). Benedetti and Loeb (1972) suggest that the enhanced performance of blind subjects is due to "perceptual compensation," which they see as a learning effect attributable to greater practice in the use of auditory information. Voice recognition is surely a skill in which blind people have a great deal of practice. Just as the audience listening to a radio drama must attend to voice to identify the characters speaking, so must blind people learn to identify people around them on the basis of voice characteristics alone.

Experiment 3 addresses the question of whether blind subjects are able to recognize voices at levels higher than those achieved by sighted subjects in Experiments 1 and 2. If they are able to do so, then the question becomes: What is the nature of the strategies and skills that lead to their better performance? If they are not, one might ask why it is that recognition of voices is relatively poor and is impervious to the effects of differing encoding tasks or habitual practice. We exposed 12 blind subjects to the same tape recording used

in Experiment 2, testing voice recognition using the materials and procedure for subjects in the Standard condition. If voice recognition proved to be more accurate for the blind, we hoped to be able to recruit enough additional blind people to test the effects of pitch and pleasantness judgments.

METHOD

The subjects were 12 blind, college-educated adults from 25 to 41 years of age, with a median age of 27. Of this group, 10 were congenitally and 2 adventitiously blind. Half of the subjects were totally blind and half had minimal form or light vision, but no subject had visual capability sufficient for visual identification of objects or people. The materials and procedure were identical to those for the Standard group in Experiment 2.

RESULTS

The hit and false-alarm proportions and A_g scores for both old and new messages for the blind subjects are presented in row 4 of Table 2. These scores may be compared with the scores for the sighted subjects in the Standard encoding condition of Experiment 2, shown in row 3 of Table 2. Performance for the sighted and blind subjects differs very little. This is confirmed by analysis of variance for the Standard conditions with type of subject and type of message as the two factors; none of the F s is significant. As in Experiment 2, voice recognition is better when voices are repeated saying the same message; unlike Experiment 2, the results are not significant here, $F(1, 34) = 2.21$.

There is no indication from these data that the blind are better than the sighted in remembering voices heard only once. Given that blind subjects in this experimental situation were no more responsive to individual differences in voices than were the sighted, there is little reason to expect that blind subjects would be any more responsive to orienting tasks.

DISCUSSION

The major finding of this research is that types of encoding activities known to have substantial effects on memory for verbal and visual events do not make a difference for voices. As with any finding of no difference, caution is advisable. Perhaps our choice of orienting tasks was ill advised. Certainly, there are other dimensions on which to judge voices besides pitch and pleasantness. The choice of pleas-

antness was based on a large number of studies showing that memory for words and faces is enhanced by evaluative judgments, and pitch commended itself to us as a dimension with clear physical meaning. Perhaps an effect of orienting task would have been found if the overall level of performance were higher. Would not a list shorter than 20 voices be a likely point of departure? These considerations are valid but are rendered somewhat less compelling by the existence of additional relevant data of which we recently became aware. In a review of the literature on voice identification, Clifford (1980) summarizes two unpublished studies by himself and G. McCardle in which a set of 10 voices was presented to subjects who engaged in one of several different orienting tasks. For example, the four encoding tasks in Clifford and McCardle's second study were identifying the sex of the speaker of the voice, judging the speaker's age, judging the warmth of the speaker, and judging whether the speaker's voice reminded the listener of anyone. The important outcome was that there was no suggestion of a significant effect of encoding task on voice memory in either experiment. Thus, the findings of Experiments 1 and 2 are supported by the work of Clifford and McCardle, adding up to four studies with the same outcome. Another factor to be considered in evaluating the present results is that potential differential effects of the orienting tasks may have been masked by the intentional learning instructions given to all subjects. Because both of the Clifford and McCardle studies used an incidental learning procedure (B. R. Clifford, personal communication, February 1983) and found no effects of orienting task on voice memory, the use of intentional learning instructions in the present research does not appear to be critical. Furthermore, orienting tasks have been shown to have substantial effects on recognition memory for words (Craik & Tulving, 1975, Experiments 9 and 10) and faces (Wingrad, 1981) even when subjects were expecting a memory test. On the basis of the available evidence, then, it seems unlikely that orienting tasks affect memory for voices. At the same time, we note again that voice memory has been shown to be sensitive to traditional parameters of memory such as retention interval, list length, and stimulus duration. The present study adds sex of speaker and the nature of the message to the list of variables to which memory for voices is sensitive. Why voice memory should be sensitive to all of these manipulations but not to the type of cognitive activity engaged in at encoding is the question raised by the present research.

The explanation we offer for the insensitivity of voice memory to these encoding manipulations has to do with the kind of stimulus a voice is when compared with a visual stimulus such as a face. First,

though, it is necessary to present an account of why evaluative judgments usually lead to better memory than judgments about physical features. The original formulation by Craik and Lockhart (1972) in terms of depth of processing has been altered to emphasize degree of trace elaboration (Anderson & Reder, 1979; Craik & Jacoby, 1979; Craik & Tulving, 1975; Winograd, 1978) and, more recently, trace distinctiveness (Craik & Jacoby, 1979; Eysenck, 1979; Winograd, 1981). To use the example of memory for faces, it is assumed that a picture of a face can be more or less elaborately scanned, depending on the information one seeks from it. To decide if the person being judged appears friendly, an extensive scan is engaged in, resulting in more features being encoded than if one is asked whether the person has a big nose. Similarly, there is broad agreement that the degree of elaboration of a word can vary widely, depending on the cognitive analysis performed. Whether or not one hypothesizes that, ultimately, the enhanced memory for faces or words following evaluative or semantic processing is due to more features being encoded or to the associated likelihood of a distinctive feature being encoded, the relevant point here is that events such as words and faces can be encoded more or less elaborately. The present results raise the interesting possibility that voices do not allow for such flexibility. Unlike faces, voices are extended in time, not space. With a complex acoustical stimulus such as the human voice, it is probably the case that one cannot attend to one feature, such as pitch, to the exclusion of the other features. In Garner's (1976) terms, the distinction is between a stimulus with *separable* dimensions, such as a face or a word, and one with *integral* (or *configural*) dimensions, a voice. Garner emphasizes that, with integral stimuli, one cannot attend to one dimension of the stimulus to the exclusion of the others.

In brief, we are suggesting that the way in which orienting tasks affect memory is by determining the amount of featural information extracted from the stimulus. With voices, featural information is integrated so that features cannot be encoded selectively. Therefore, orienting tasks do not affect voice memory differentially. Whether a similar conclusion would be warranted with other auditory events is an interesting question. Consider listening to a symphony orchestra — one gets the impression that now one is listening to the woodwinds and now to the strings. Or, consider the well-investigated selective attention paradigm where there are simultaneous but separate messages, one for each ear. Both of these examples are presented because they seem blatantly to contradict the characterization just offered of voices. The difference is this: Both the symphony and the divided-attention paradigm are examples of acoustic events as separable

dimensions because each constitutes an ensemble of multiple voices or sources. In the experiments reported here, we always presented a single voice at a time. Although one can often pick out one voice among many in an ensemble (but not always, as in choral singing), we suggest that a single voice is psychologically a whole. Looked at in this way, the finding that memory for voices does not reflect encoding processes is attributable to the integral, unitary nature of voices. One would expect the same outcome with integral visual stimuli as well. The challenge now is to convert these speculations into experimental procedures.

Notes

This research was supported in part by National Institutes of Mental Health Grant 32063. Requests for offprints should be sent to Eugene Winograd, Department of Psychology, Emory University, Atlanta, GA 30322. Received for publication November 8, 1982; revision received March 7, 1983.

References

- Anderson, J. R., & Reder, L. M. (1979). An elaborative processing explanation of depth of processing. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 385-403). Hillsdale, NJ: Erlbaum.
- Benedetti, L. H., & Loeb, M. (1972). A comparison of auditory monitoring performance in blind subjects with that of sighted subjects in light and dark. *Perception & Psychophysics*, *11*, 10-11.
- Bower, G. H., & Karlin, M. B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology*, *103*, 751-757.
- Carterette, E. C., & Barnebey, A. (1975). Recognition memory for voices. In A. Cohen & S. G. Nootboom (Eds.), *Structure and process in speech perception* (pp. 246-265). New York: Springer-Verlag.
- Clifford, B. R. (1980). Voice identification by human listeners: On earwitness reliability. *Law and Human Behavior*, *4*, 373-394.
- Craik, F. I. M., & Jacoby, L. L. (1979). Elaboration and distinctiveness in episodic memory. In L. Nilsson (Ed.), *Perspectives on memory research* (pp. 145-166). Hillsdale, NJ: Erlbaum.
- Craik, F. I. M., & Kirsner, K. (1974). The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology*, *26*, 274-284.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671-684.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the reten-

- tion of words in episodic memory. *Journal of Experimental Psychology: General*, 104, 268-294.
- Eysenck, M. L. (1979). Depth, elaboration, and distinctiveness. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 89-118). Hillsdale, NJ: Erlbaum.
- Ferrara, R. A., Puff, C. R., Gioia, G. A., & Richards, J. M. (1978). Effects of incidental and intentional learning instructions on the free recall of naturalistic sounds. *Bulletin of the Psychonomic Society*, 11, 353-355.
- Garner, W. R. (1976). Interaction of stimulus dimensions in concept and choice processes. *Cognitive Psychology*, 8, 98-123.
- Geiselman, R. E., & Bellezza, F. S. (1976). Long-term memory for speaker's voice and source location. *Memory & Cognition*, 4, 483-489.
- Hyde, T. S., & Jenkins, J. J. (1969). Differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology*, 82, 472-481.
- Mann, V. A., Diamond, R., & Carey, S. (1979). Development of voice recognition: Parallels with face recognition. *Journal of Experimental Child Psychology*, 27, 153-165.
- McGehee, F. (1973). The reliability of the identification of the human voice. *Journal of General Psychology*, 17, 249-271.
- Miller, D. J., Mueller, J. H., Goldstein, A. G., & Potter, T. L. (1978). Depth of processing and test anxiety in landscape recognition. *Bulletin of the Psychonomic Society*, 11, 341-343.
- Pastore, R. E., & Scheirer, C. J. (1974). Signal detection theory: Considerations for general application. *Psychological Bulletin*, 81, 945-958.
- Pollack, I., Pickett, J. M., & Sumbly, W. H. (1954). On the identification of speakers by voice. *Journal of the Acoustical Society of America*, 26, 403-406.
- Smith, A. D., & Winograd, E. (1978). Adult age differences in remembering faces. *Developmental Psychology*, 14, 443-444.
- Warrington, E. K., & Ackroyd, C. (1975). The effect of orienting tasks on recognition memory. *Memory & Cognition*, 3, 140-142.
- Winograd, E. (1978). Encoding operations which facilitate memory for faces across the life span. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 255-262). New York: Academic Press.
- Winograd, E. (1981). Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 181-190.
- Witkin, H. A., Oltman, P. K., Chase, J. B., & Friedman, F. (1971). Cognitive patterning in the blind. In J. Hellmuth (Ed.), *Cognitive Studies* (Vol. 2, pp. 16-46). New York: Brunner-Mazel.