JOURNAL OF EXPERIMENTAL CHILD PSYCHOLOGY 27, 153–165 (1979)

Development of Voice Recognition: Parallels with Face Recognition

V. A. MANN, R. DIAMOND, AND S. CAREY

Massachusetts Institute of Technology

Encoding and subsequent recognition of unfamiliar voices was examined in children, aged 6 to 16, and in adults. Performance changed markedly with age, improving sharply between ages 6 and 10, with 10-year-olds approaching adult levels. After age 10 accuracy declined significantly but returned to the adult level by age 14. The course of development closely matches that recently documented for the ability to encode and recognize unfamiliar faces. Several parallels between voice recognition and face recognition are discussed.

The acoustic patterns of speech not only transmit information about speakers' beliefs and intentions, but about their identity as well. Laboratory studies have explored the capacity for recognition of individuals from their voices. Adult listeners easily recognize the voices of familiar people (Bricker & Pruzansky, 1966; Compton, 1963; Stevens, Williams, Carbonell, & Woods, 1968). Bartholomeus (1973) reported that some nursery school children identified the voices of their classmates almost as accurately as did their teachers. On the basis of this observation, she suggested tentatively that only a minimal change in voice recognition capacity occurs after age 4. Indeed, even infants as young as 9 months of age can distinguish the voice of their mother from that of an unfamiliar female (Friedlander, 1970).

However, voice recognition has two aspects: (1) the ability to recognize the voice of an already known person and (2) the ability to make a voice familiar, that is, to encode a voice and store some representation of it in long-term memory. Adults certainly can encode unfamiliar voices from

This research was supported by NIH Grant 1-R01-HDO9179 and by a grant from the Spencer Foundation to H.-L. Teuber. We are grateful to the children who participated, their parents, and their teachers. We thank Dr. J. Lawson, Mr. R. Barnes, Mr. C. Como, Mr. E. Sullivan, and Dr. D. Clune of the Lexington School System for their cooperation. We would also like to thank Dr. F. Ganong for helpful advice and Mr. A. Cahn for assistance in collecting data. Reprint requests should be addressed to Dr. Rhea Diamond, Department of Psychology, Massachusetts Institute of Technology, Cambridge, MA 02139.

0022-0965/79/010153-13$02.00/0

short presentations, as is shown by their ability to distinguish these targets from other voices during subsequent recognition trials. Performance in this task is highly dependent upon such variables as sample duration, set size, and interval between inspection and recognition trials (Bricker & Pruzansky, 1966; Carterette & Barneby, 1975; Pollack, Pickett, & Sumby, 1954; Stevens *et al.*, 1968). For example, adults exposed to short sentences spoken by as few as four novel speakers with an interval of 15 sec before attempted recognition, performed at a level only slightly better than chance (Carterette & Barneby, 1975). Development of the capacity to make unfamiliar voices familiar has not yet been systematically addressed; it may well differ from that of the capacity to recognize already known voices.

Face recognition offers an instructive parallel. Several studies suggest that children under age 10 represent highly familiar faces as do adults, recognizing them as well as adults in simple disguises (Diamond & Carey, 1977) or as badly as adults when the photographs are inverted (Brooks & Goldstein, 1963; Goldstein, 1975). In contrast, children under 10 are markedly impaired in tasks which require the encoding of unfamiliar faces (Benton & Van Allen, 1973; Carey & Diamond, 1977; Carey, Diamond, & Woods, in preparation; Diamond & Carey, 1977; Goldstein, 1965; Saltz & Sigel, 1967). Several different studies of children between ages 6 and 16 have revealed a similar developmental pattern: marked improvement between age 6 and ages 9–10, a plateau or decline until age 14, and a subsequent gain to reach or regain the adult level by age 16 (Carey *et al.*, in preparation).

Faces and voices are both complex stimuli which convey important social information. Encoding of both plays a part in the ability to recognize and distinguish individual persons. These similarities make it plausible that the development of voice recognition may, in some respects, parallel the development of face recognition. The purpose of this paper is to describe the development of the capacity for encoding unfamiliar voices and to compare its course with the development of the capacity for encoding faces.

The procedure of Carterette and Barneby (1975) was adapted for use with children. Either one or two speech samples were presented followed by a forced choice recognition among either two or four samples. In addition to the manipulation of set size, two levels of presumed difficulty were imposed by varying the relation between inspection and recognition utterances. Half the subjects were tested with items in which the same sentences were spoken during inspection and recognition. The remaining subjects were tested with items in which the sentences spoken during inspection and recognition differed. When utterance is held constant, discrimination should be facilitated. When utterance changes, the situation more closely resembles voice recognition in everyday life.

METHOD

Subjects

Twenty MIT students and 20 children at each of ages 6, 8, 10, 11, 12, 13, 14, and 16 participated. The MIT students were paid volunteers; the children attended the Lexington Public School System. Equal numbers of males and females were tested at each age.

Materials

Practice stimuli. We attempted to make the practice items easy by having target and foil speakers differ in sex or accent. Twenty native American speakers of English (10 males and 10 females) and 6 non-native speakers (3 males and 3 females) produced the practice items. Three instances of each of the following sentences were recorded by each speaker on ¼" magnetic tape:

Who am I?
He is tall.

Intonation was not controlled. However, the stimuli were mechanically equilibrated for intensity. Each voice was heard in only one item. That is, each item was produced by a unique set of either two or four speakers. Three of the practice items (one-target items) consisted of one inspection utterance followed by two recognition utterances. Four items (two-target items) consisted of two inspection utterances followed by four recognition utterances. Subjects who were to be tested in the condition in which the same sentence was presented during inspection and recognition (same-utterance condition) were provided practice items in which the sentence presented during inspection and recognition was the same. Subjects who were to be tested in the condition in which different sentences were presented during inspection and recognition (different-utterance condition) were presented practice items in which the two also differed. The particular voices used and the order in which they were presented was matched across the two utterance conditions. During recognition trials target and foil speakers were presented in a fixed random order.

In the same-utterance condition, exact physical matches were not presented. Different recorded instances of the same sentence comprised the inspection and recognition utterances produced by the same speaker.

Test stimuli. The corpus of speech samples from which the test items were constructed was produced by 22 native American speakers of English. The speakers were all adult females between the ages of 25 and 45; none spoke with a strong regional accent or marked idiosyncrasy of enunciation. Each speaker attempted to match her intonation and cadence to that of a recorded model of each sentence to which she listened immediately prior to speaking. Recordings were mechanically equilibrated for intensity.

Frequency of vocal cord vibration and size and shape of the vocal tract are important sources of information utilized in speaker identification (Carterette & Barneby, 1975; Flanagan, 1965). Information concerning the parameters of the vocal tract is available from the extremes of the triangle formed in F_1–F_2 space by the vowels *u*, *a*, and *ï*. The sentences used contained each of these vowels. They are given below:

(*a*) Who are we?
(*b*) We call you.
(*c*) He threw the ball.
(*d*) He grew tall.

Each speaker produced each sentence three times. Sentences *a* and *b* were used as inspection utterances (and as recognition utterances for the same-utterance condition). Sentences *c* and *d* were used as recognition utterances for the different-utterance condition. For each utterance condition, four one-target test items and three two-target test items were constructed. Each person spoke in only one item, either as target or foil. As in the practice items, the same-utterance series matched the different-utterance series in order and choice of speakers in each item.

Procedure

For the child subjects an introductory training procedure was employed to assure understanding of the task and to minimize the necessity of talking during testing. We felt that speech would interfere with processing of the auditory materials and therefore devised a system of nonverbal cues and responses. Initially, the procedure was demonstrated for each child with a visual task.

A picture playing card was presented and the child instructed to look at it very carefully. Then the card was turned face down and mixed with a second card. The experimenter turned both cards face up, one at a time, and as each was presented asked the child to nod "yes" if it was the one he had originally seen and "no" if he had not seen it before. Any errors were pointed out, and the task repeated with a second set of cards.

Immediately afterward, the child was fitted with binaural headphones and instructed that voices would be presented. He was told that later he would be asked to try to tell the speakers apart, in the same way that he had told the cards apart. The experimenter said, "You don't know any of these people, but if you listen carefully, you will be able to tell whether the same person or a different person is talking. I'll tell you the name of the first person who speaks. Listen to him. Then two people will speak, one at a time. One will be the person you just heard, and the other will be a mystery person. Just as you did with the cards, nod 'yes' if you think it is the same person, and 'no' if you think it is the mystery person." Then the first practice item was presented. The experimenter said, "This is Ken" and displayed a name card with "Ken" printed on it as "Ken" spoke. After a 3-sec interval, the two comparison utterances were presented, separated by a 2-sec interval. A card containing a question mark accompanied each sample. The experimenter waited for the child to respond after each utterance. If he failed to respond "yes" and "no" at the proper time, the instructions and trial were repeated. If the child followed the correct procedure and responded correctly, he was praised.

Following successful performance on the first practice item, a second item was presented with target and comparison speakers marked in the same manner. In these first two practice items, target and mystery speakers were of opposite sex. Next, a third practice item was presented in which target and mystery persons were both women, but there was a marked difference of accent between them. Before the third practice item was presented, the child was warned that "Carol" and the mystery person might sound very much alike.

On completion of these practice items, the child was advised that the next items were much more difficult and required careful attention. In addition, the experimenter would no longer tell whether or not the answer was correct. At this time the four one-target test times were presented in a fixed order. As in the practice items, target speakers were marked with name cards, and comparison speakers with a question mark. If the child gave the same response to both comparison utterances or if he requested to hear the voices again, the item was repeated.

A rest period of 5 min was given between one-target problems and two-target problems. Again, visual pretraining preceded practice items. The child was shown two playing cards in succession. Each was turned face down immediately after being shown and then both were mixed with two additional cards. The four cards were then turned face up, one at a time, and the child asked to nod "yes" to the cards he had already seen, and "no" to the ones which he had not seen previously. Errors were pointed out and the task repeated with additional cards, until the child answered quickly and correctly.

These visual practice items were followed by four voice practice items. Name cards and question mark cards were used as in the one-target case. In the first two practice items, target and foil speakers differed in sex. The third practice item consisted of speakers who differed either in sex or in accent while the fourth consisted of speakers who differed in accent alone. Children were advised that the last two practice items were of increasing difficulty. Error correction was provided and items repeated if an error occurred.

When a child had completed these practice items he was cautioned that the next items would be even more difficult. Again, the experimenter would not tell whether or not the response was correct, but the item could be repeated if desired. The three two-target items then followed in a fixed order. If the child failed to respond "yes" to two recognition samples and "no" to the remaining two, the item was repeated.

The instructions and procedure were modified appropriately for use with 16-year-olds and adults. Total length of the procedure varied from 15 to 30 min.

RESULTS

Practice Items

No errors were made by children at any age on visual practice items or on voice practice items in which speakers differed in sex. Children often remarked that these items were easily judged because "the mystery person is a lady . . .". Among the child subjects, some errors occurred on practice items produced by speakers of the same sex with different accents. Overall, the error rate for the accent practice items was 16%. These errors were made predominantly by children age 6, 8, and 12 (see Fig. 1). When a child incorrectly responded to a practice item it was repeated until he responded correctly.

Test Items

Fewer than 15% of all test items had to be repeated because of failure to follow instructions. The majority of repetitions occurred on two-target items; in these cases only the response to the final presentation was scored. Within each age group mean percent correct was computed separately for each type of test item. An ANOVA performed on these results indicated a main effect for age (df = 8,28; $F = 8.09$; $p \leq .0001$), for number of targets (df = 1,28; $F = 22.3$; $p \leq .0001$), and for utterance condition (df = 1,28; $F = 52.2$; $p \leq .0001$). There was no main effect for sex and no interaction of sex with any other variable. The interaction of utterance condition and number of targets was also significant (df = 1,28; $F = 4.1$; $p \leq .05$).

Figure 1 presents performance with number of targets and utterance condition collapsed.

Six-year-olds did not perform at a level beyond chance while all other subjects did ($p \leq .01$). Between ages 6 and 8 and also between ages 8 and 10, accuracy increased (df = 38; $t_{2\text{-tailed}} = 3.9$; $p \leq .0001$ and $t_{2\text{-tailed}} = 2.4$; $p \leq .02$, respectively). The 10-year-olds did not differ from the adults. Performance declined between ages 10 and 13 (df = 38; $t_{2\text{-tailed}} = 2.02$; $p \leq .05$) and returned to adult levels by age 14 (the performance of 13-year-olds was significantly inferior to that of 14-year-olds at df = 38; $t_{2\text{-tailed}} = 2.18$; $p \leq .03$).

Figure 2 shows performance on each kind of item. One-target items (mean correct 75%) were easier than two-target items (mean correct 65%).

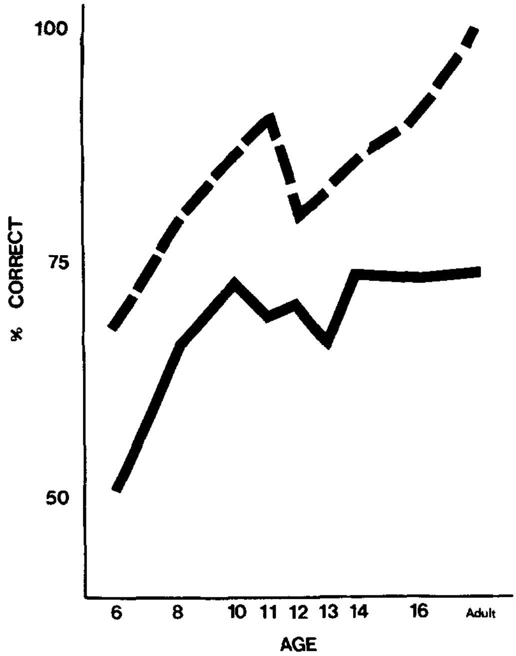

FIG. 1. Performance on test items (solid line) and on practice items in which speakers differed in accent (dashed line).

The same-utterance condition (mean correct 78%) was easier than the different-utterance condition (mean correct 60%). The significant interaction of set size and utterance condition is associated with a greater effect of utterance condition on one-target items than on two-target items. Finally, a marginally significant three-way interaction between age, number of targets and utterance condition (df = 8,28; $F = 2.12; p \leq .09$) is associated with absence of the set size-utterance condition interaction among adults.

Inspection of Fig. 2 also indicates that the effect of age on overall performance (Fig. 1) is generally true for each of the four conditions. The major exception to this uniformity is in one-target/same-utterance items, the easiest condition. In this condition, 6-year-olds performed better than chance (df = 18; $t_{2\text{-tailed}} = 2.45; p \leq .05$) and 10-year-olds did not reach the adult level (df = 18; $t_{2\text{-tailed}} = 2.5; p \leq .05$). Improvement continued until age 12 and the dip in performance was confined to age 13 (13-year-olds performed worse than 12-year-olds; df = 18; $t_{2\text{-tailed}} = 2.61; p \leq .02$). For all three of the other conditions there was significant improvement between ages 6 and 10 with children reaching the adult level by age 10. Between ages 12 and 13 accuracy declined, to return to adult level by age 14. However, performance on the two-target/different-utterance condition, the hardest condition, remained low even for adults. The curve for

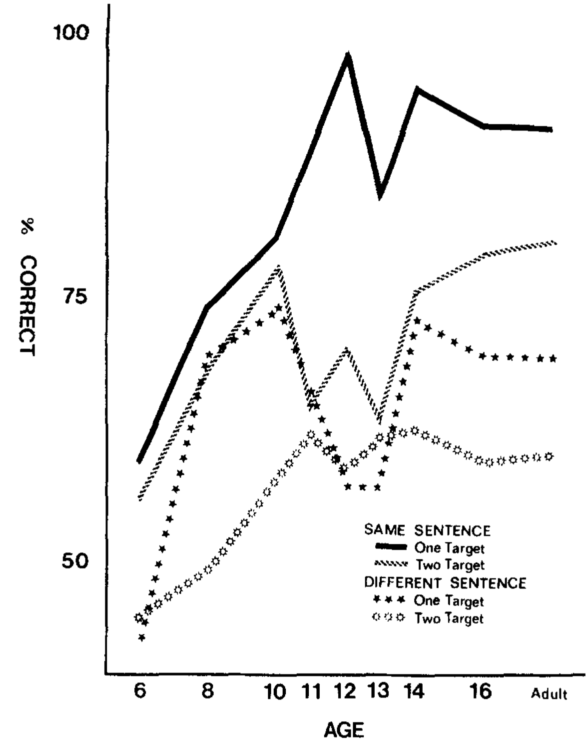

Fig. 2. Performance on each type of test item.

this condition is relatively flat with children age 6, 8, and 12 all performing at chance levels.

DISCUSSION

The high levels of performance on practice items indicate that children of all ages understood the task. Even for the youngest children, difficulty arose only when speakers of the same age and sex had to be discriminated. Both 6- and 8-year-olds appear less sensitive than adults to differences in accent, an observation which warrants further research.

Contrary to implications which might be drawn from Bartholomeus' study of recognition of familiar voices, we observed marked changes in performance with age. Apparently, tasks requiring encoding of unfamiliar voices produce a developmental pattern distinct from that found in tasks of recognition of voices already represented in long-term memory. This suggests that it is in the encoding process, the making of a voice familiar, that developmental changes take place.

As expected, performance was facilitated when target and foil speakers produced the same utterance. Also in agreement with Carterette and Barneby (1975), set size influenced accuracy; items in which one sample was presented for inspection were easier than items in which two samples were presented.

The developmental course of voice recognition resembles that which Carey *et al.* (in preparation) have shown to hold for face recognition. Before about age 10 the ability to encode an unfamiliar voice from a small speech sample was deficient, as was the ability to encode an unfamiliar face from a single still photograph. Moreover, with both materials at age 10 children approached adult levels of performance. Beyond age 10 performance declined so that 12- and 13-year-olds (in the case of voices) and 12-, 13-, and 14-year-olds (in the case of faces) were inferior to 10-year-olds and to older children and adults. In neither domain was the dip in performance earlier in girls than in boys. However, larger samples would be needed to evaluate the influence of onset of puberty.

In view of the many parallels between face recognition and voice recognition, their common developmental course is perhaps not surprising. The same biologically significant pruposes are served by both: the identification of particular individuals and the determination of their intentions or moods. While levels of performance in encoding new voices are inferior to those demonstrated for faces (in terms of sheer efficiency), adults eventually encode hundreds of voices and maintain those representations in memory over long periods of time. The basis on which voices are recognized has not been completely specified although fundamental frequency of the voice, intensity, and intonation pattern, all are known to contribute (Carterette & Barneby, 1975; Clarke & Becker, 1966; Compton, 1963; Holmgren, 1967; Kaplan, Note 2; Matsumoto, Hike, Sone, & Nimura, 1973; Voiers, 1964). The ability to recognize a large number of individual voices suggests that complex patterns of voice parameters must be encoded. In the case of faces, the ability to encode virtually unlimited numbers of instances suggests, similarly, that complex relations among what are physically similar parts must underlie individuation. The improvement in encoding of unfamiliar faces between ages 6 and 9–10 has been attributed to a shift away from reliance on relatively isolated distinguishing features toward those that are more configurational (Carey & Diamond, 1977; Diamond & Carey, 1977). Whether the type of information utilized in encoding unfamiliar voices also changes during development is, of course, still an open question.

There is also reason to believe that the neural substrate of face and voice recognition may have aspects in common. Studies of normal adults (e.g., Geffen, Bradshaw, & Wallace, 1971; Hilliard, 1973; Klein, Moscovitch, & Vigna, 1976; Rizzolatti, Umiltà, & Berlucchi, 1971), brain-injured adults (e.g., Benton & Van Allen, 1968; De Renzi, Faglioni, & Spinnler, 1968; Milner, 1958; Tzavaras, Hécaen, & Le Bras, 1970; Warrington & James, 1967; Yin, 1970), and commisurectomized patients (Levy, Trevarthen, & Sperry, 1972; Sperry, 1974) have provided strong evidence that the right cerebral hemisphere plays an essential role in the ability to encode faces. While not as firmly fixed, the role of the right hemisphere in voice recognition is supported by evidence from several sources. Tonal memory and the discrimination of timbre have been shown to be differentially impaired in persons with right hemisphere damage (Milner, 1962). Dichotic listening studies, which have consistently supported a left hemisphere advantage for verbal materials in general have yielded a right hemisphere advantage for intonation contour (Blumstein & Cooper, 1974). The same paradigm has also revealed a right hemisphere advantage for recognition of melody (Kimura, 1967) and environmental sounds (Knox & Kimura, 1970). Evidence that the right hemisphere is differentially involved in processing of mechanically produced sounds has been provided in a study of auditory evoked potential (Molfese, Freeman, & Palermo, 1975). These findings are consistent with a right hemisphere role in the encoding of voices. Direct evidence for such a role has come from studies of adults with focal lesions of the brain. Using brief sentences as stimuli, Grenier (Note 1) and Assal, Zander, Kremin, and Buttet (1976) demonstrated that patients with damage to the right hemisphere were differentially impaired on a task involving encoding of unfamiliar voices. On the other hand, Doehring and Bartholomeus (1971) in a voice recognition task using consonants, vowels, and single words, found a left hemisphere advantage for normal adults. It seems likely that what is involved in matching these abbreviated stimuli differs somewhat from what is involved in encoding voices when whole sentences are presented.

Evidence for a direct relation between development and the role of the right hemisphere has been provided in the case of faces. Two different tasks involving face encoding have been presented to both children and brain-injured adults. In both instances, the performance of young children matched that of patients who had suffered damage to the right hemisphere while that of 9- or 10-year-olds was in the normal adult range (Benton & Van Allen, 1973; Carey *et al.*, in preparation). In another study, development of the ability to encode unfamiliar faces was directly linked to changes in the functioning of the right hemisphere; a left visual field advantage did not emerge until age 9 and diminished at ages 12 and 14 (Leehey, 1977). The decline in lateral asymmetry in this procedure appeared at the same ages at which a dip in overall performance on faces was observed by Carey *et al.* Whether there is similarly a direct tie between the development of voice recognition and changes in functioning of the right hemisphere has not yet been evaluated.

A curious aspect of the developmental history of both face recognition and voice recognition is the decline in performance after age 10. It is possible that these dips result from maturational changes, presumably in the right hemisphere, associated with the onset of puberty. It is also possible that these dips reflect temporary shifts in general information processing strategies or in reliance on certain nonveridical cues for identifying faces and voices. Changes of this kind might not implicate maturational factors at all. The dips in performance on the two types of material may also, of course, reflect the operation of different factors in each case.

In conclusion, several parallels between voices and faces have been noted. Both serve as bases for person recognition, suggesting that in both cases complex relations among stimulus features are represented in memory. Deficiencies in encoding both voices and faces are associated with damage to the right hemisphere. In the present paper, evidence has been presented that the capacity for making new voices familiar shows the same developmental course as the capacity for making new faces familiar. Experience with persons, during the first decade of life, could account for acquisition of schemata supporting efficient encoding in both domains. In addition, maturational changes in the right hemisphere either in areas common to visual and auditory processing or occurring in parallel in areas specific to each modality, could contribute to the developmental similarities.

REFERENCES

Assal, G., Zander, E., Kramin, H., & Buttet, J. Discrimination des voix lors lesions du cortex cérébral. *Archives Suisse de Neurologie, Neurochiurgie et Psychiatry,* 1976, **119,** 307–315.

Bartholomeus, B. Voice identification by nursery school children. *Canadian Jorunal of Psychology,* 1973, **27,** 464–472.

Benton, A. L., & Van Allen, M. W. Impairment in facial recognition in patients with cerebral disease. *Cortex,* 1968, **4,** 344–358.

Benton, A. L., & Van Allen, M. W. *Manual: Test of facial recognition.* Neurosensory Center Publication No. 287, Department of Neurology, University Hospitals, Iowa City, IO, 1973.

Blumstein, S., & Cooper, W. Hemispheric processing of intonation contours. *Cortex,* 1974, **10,** 146–158.

Bricker, P. D., & Pruzansky, S. Effects of stimulus content and duration on talker identification. *Journal of the Acoustical Society of America,* 1966, **40,** 1441–1449.

Brooks, R. M., & Goldstein, A. G. Recognition by children of inverted photographs of faces. *Child Development,* 1963, **34,** 1033–1040.

Carey, S., & Diamond, R. From piecemeal to configurational representation of faces. *Science,* 1977, **195,** 312–314.

Carey, S., Diamond, R., & Woods, B. Development of face perception—A maturational component? In preparation.

Carterette, E. C., & Barneby, A. Recognition memory for voices. In E. Cohen & G. Nottebohn (Eds.), *Structures and processes in speech perception.* New York: Springer-Verlag, 1975.

Clarke, F. R., & Becker, R. W. Comparison of techniques for discriminating among talkers. *Journal of Speech and Hearing Research,* 1966, **12,** 747–761.

Compton, A. J. Effects of filtering and vocal duration upon the identification of speakers aurally. *Journal of the Acoustical Society of America,* 1963, **35,** 1748–1752.

De Renzi, E., Faglioni, P., & Spinnler, H. The performance of patients with unilateral brain damage on face recognition tasks. *Cortex,* 1968, **4,** 17–34.

Diamond, R., & Carey, S. Developmental changes in the representation of faces. *Journal of Experimental Child Psychology,* 1977, **23,** 1–22.

Doehring, D. G., & Bartholomeus, B. N. Laterality effects in voice recognition. *Neuropsychologia*, 1971, **9**, 425–430.

Flanagan, J. L. *Speech analysis, synthesis and perception*, New York: Academic Press, 1965.

Friedlander, B. Z. Receptive language development in infancy: Issues and problems. *Merrill-Palmer Quarterly of Behavior and Development*, 1970, **16**, 7–51.

Geffen, G., Bradshaw, J. L., & Wallace, G. Interhemispheric effects on reaction times to verbal and non-verbal stimuli. *Journal of Experimental Psychology*, 1971, **87**, 415–422.

Goldstein, A. G. Learning of inverted and normally oriented faces in children and adults. *Psychonomic Science*, 1965, **3**, 447–448.

Goldstein, A. G. Recognition of inverted photographs of faces by children and adults. *Journal of Genetic Psychology*, 1975, **127**, 109–123.

Hilliard, R. D. Hemispheric laterality effects on a facial recognition task in normal subjects. *Cortex*, 1973, **9**, 246–258.

Holmgren, G. L. Physical and psychological correlates of speaker recognition. *Journal of Speech and Hearing Research*, 1967, **10**, 57–66.

Kimura, D. Functional asymmetry of the brain in dichotic listening. *Cortex*, 1967, **3**, 163–178.

Klein, D., Moscovitch, M., & Vigna, C. Perceptual asymmetries and attentional mechanisms in tachistoscopic recognition of words and faces. *Neuropsychologia*, 1976, **14**, 227–237.

Knox, C., & Kimura, D. Cerebral processing of nonverbal sounds in boys and girls. *Neuropsychologia*, 1970, **8**, 227–237.

Leehey, S. A developmental change in right hemisphere specialization. *Abstracts of Individual Papers*, Society for Research in Child Development, Biennial Convention, 1977. P. 91.

Levy, J., Trevarthen, C., & Sperry, R. W. Perception of bilateral chimeric figures following hemisphere deconnexion. *Brain*, 1972, **95**, 61–68.

Matsumoto, H., Hike, S., Sone, T., & Nimura, T. Multidimensional representation of personal quality of vowels and its acoustical correlates. *IEEE Transactions on Audio and Electroacoustics*, 1973, **AU-21**, 428–436.

Milner, B. Psychological deficits produced by temporal lobe excision. *Proceedings of the Association for Research in Nervous and Mental Disease*, 1958, **36**, 244–257.

Milner, B. Laterality effects in audition. In V. B. Mountcastle (Ed.), *Interhemispheric Relations and Cerebral Dominance*. Baltimore: Johns Hopkins Press, 1962.

Molfese, D. L., Freeman, R. B., & Palermo, D. S. The ontogeny of brain lateralization for speech and nonspeech stimuli. *Brain and Language*, 1975, **2**, 356–368.

Pollack, R. K., Pickett, J. M., & Sumby, W. H. On the identification of speakers by voice. *Journal of the Acoustical Society of America*, 1954, **26**, 403–406.

Rizzolatti, G., Umiltà, C., & Berlucchi, G. Opposite superiorities of the right and left hemispheres in discriminative reaction time to physiognomic and alphabetical material. *Brain*, 1971, **94**, 431–442.

Saltz, E., & Sigel, I. E. Concept overdiscrimination in children. *Journal of Experimental Psychology*, 1967, **73**, 82–93.

Sperry, R. W. Lateral specialization in the surgically separated hemispheres. In F. O. Schmitt & F. G. Worden (Eds.), *The neurosciences: Third study program*. Cambridge, MA: The MIT Press, 1974.

Stevens, K. N., Williams, C. E., Carbonell, J. R., & Woods, B. Speaker authentification and identification, a comparison of spectrographic and auditory presentations of speech material. *Journal of the Acoustical Society of America*, 1968, **44**, 1596–1607.

Tzavaras, A., Hécaen, H., & Le Bras, H. Le problème de la spécificité du déficit de la reconnaissance du visage humain lors des lésions hémisphériques unilatérales. *Neuropsychologia*, 1970, **8**, 403–416.
Voiers, W. D. Perceptual bases of speaker identification. *Journal of the Acoustical Society of America*, 1964, **36**, 1065–1073.
Warrington, E. K., & James, M. An experimental investigation of facial recognition in patients with unilateral cerebral lesions. *Cortex*, 1967, **3**, 317–326.
Yin, R. K. Face recognition by brain injured patients: A dissociable ability? *Neuropsychologia*, 1970, **8**, 395–402.

REFERENCE NOTES

1. Grenier, D. *La prosopagnosia et la agnosie des voix*. Unpublished M.A. thesis, University of Montreal, 1969.
2. Kaplan, E. L. *The role of intonation in the acquisition of language*. Unpublished doctoral dissertation, Cornell University, 1969.

RECEIVED: January 17, 1978; REVISED: April 21, 1978.