Preservation of Familiar Speaker Recognition but not Unfamiliar Speaker Discimination in Aphasic Patients

Diana Van Lancker
Sepulveda Veterans Administration Medical Center, Sepulveda, California
and University of California-Los Angeles, California

Jody Kreiman
University of California-Los Angeles, California

Communication involves an exchange of different kinds of information. Besides linguistic content (structured into phonological, syntactic, and semantic systems), there is the important realm of paralinguistic information, which includes attitudes, affect, mood, and personal traits (sex, age, personality). Paralinguistic information, carried mainly in the prosodic material of speech (pitch, duration, loudness, voice quality) is not made up of discrete units as the linguistic system (phonemes and words) is, but instead, involves different characteristics altogether. Gradations and degrees occur in the case of emotions and attitudes, and unique patterns make up speaker identity.

As left hemisphere specialization for language is presumably accounted for and dependent upon the structural properties of linguistic systems, an interesting question arises about the hemispheric specialization for paralinguistic information in communication. Recent studies suggest that hemispheric specialization is related to the structural properties of the stimulus. For example, when prosodic contrasts constitute a linguistic tonal system, they are processed in the left hemisphere (Van Lancker and Fromkin, 1973, 1978; Gandour and Dardaranda, 1983). On the other hand, a right hemisphere specialization has been found for prosodic material signalling emotional content (Heilman, Scholes and Watson, 1975; Ross, 1981) and personal voice identity (Van Lancker and Canter, 1982) (Table 1).

Table 1. A model for hemispheric specialization.

HEMISPHERE		LINGUISTIC STRUCTURE
LEFT ↑	Phonological (tones)	YES ↑
	Grammatical (word stress)	
	Grammatical (sentence intonation)	
	Attitudinal (disapproval, irony)	
	Emotional (happy, sad)	
RIGHT ↓	Voice information (sex, age, identity	NO ↓

Thus although all this acoustic material is carried in speech, only part--the linguistic content--is processed in the left cerebral hemisphere, whereas paralinguistic information may be processed in the right.

To further investigate the hypothesis that the right hemisphere is specialized for processing paralinguistic information in speech, we studied the abilities of left-brain damaged (LBD) and right-brain damaged (RBD) subjects to perceive personal voice identity in the speech signal. Our goal was to ascertain how well aphasic and RBD individuals could recognize familiar voices and discriminate between unfamiliar voices.

Although recognition of the speaking voices of family and friends is a common ability utilized daily, little research on neural substrates underlying this ability and its disturbance in brain damage has been done. Most experimentation in voice perception in normal subjects, in fact, has utilized unfamiliar voices in discrimination (same/different) tasks.

Information about personal identity is carried in speech, and hence it might be thought to be specialized in the left hemisphere in normal, right-handed persons. But as mentioned above, this information has none of the properties of speech and language (such as discrete, sequential units organized by rules); instead, recognizing a voice involves apperception of an overall pattern. Thus speaker recognition is analogous to face recognition, which has been demonstrated in normal subjects (Levine and Koch-Weser, 1982) and in brain-damaged patients (Benton, 1980) to crucially involve the right hemisphere.

Data from our previous studies of normal subjects support this view of familiar voice recognition as essentially a pattern recognition process (Van Lancker, Kreiman and Emmorey, 1985; Van Lancker, Kreiman, and Wickens, 1985). In those studies, listeners showed only small decrements in performance when recognizing voices presented backwards or with altered rates (compared to normal presentation). Secondly, rather than having a uniform effect, these acoustic alterations affected the recognizability of some voices but not others. In the backwards condition, for example, some voices were as easily recognized backwards as forwards, whereas others were rendered unrecognizable by the alteration (Figure 1).

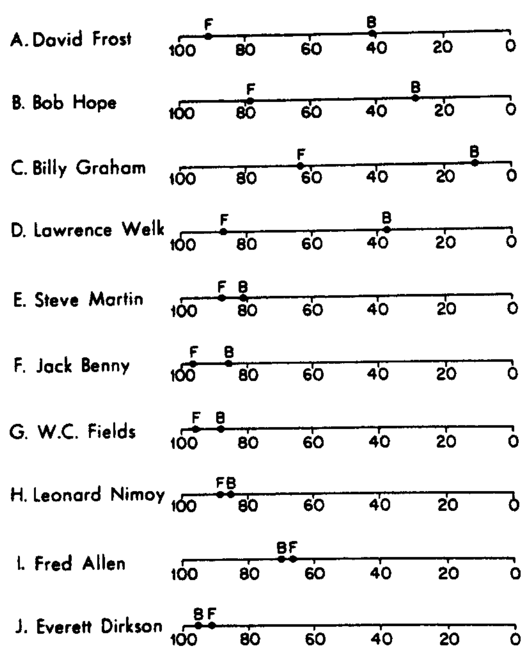

Figure 1. Scores for listeners' recognition of familiar voices presented backwards and forwards.

This indicated that since a given parameter is irrelevant to the characteristic "quality" of one voice but central to others, each voice is perceived as a unique pattern, made up of an integrated constellation of parameters leading to correct recognition. This notion of the individual voice as auditory pattern is borne out by clinical studies.

Two such studies suggested that speaker recognition is indeed mediated by the right hemisphere. One, using famous voices in a recognition task (Van Lancker and Center, 1982), found a significant difference in the performance of LBD and RBD subjects, with LBD preservation of familiar speaker recognition. The other, using unfamiliar voice stimuli in a discrimination paradigm, found a nonsignificant trend for RBD subjects to be worse in discriminating between speakers (Assal, Zander, Kreimin, and Buttet, 1976).

The questions posed in the present study were these: (1) Is familiar speaker recognition mediated primarily by the right hemisphere, such that the aphasic patient would retain the ability to recognize the voices of family and friends despite a severe deficit in speech and language function? (2) Are recognition and discrimination abilities, with regard to voice perception, mediated by the same, or by different neural substrates?

METHOD

Two test protocols were used to assess and compare speaker recognition and discrimination abilities in brain-damaged subjects. From our previous studies of voice recognition in normal subjects (Van Lancker et al., 1985a) the voices of 25 male entertainers and politicians (e.g., Johnny Carson and John F. Kennedy) familiar to most Americans were selected. The voice samples were low-pass filtered at 4 kHz and sampled at 10 Hz on a PDP-11/23 computer, and then edited to create 4-second stimuli free of pauses, background noises, and identifying content. A written pretest ensured that the correct answer was not available from the linguistic content of the sample. Response sheets consisted of vertically aligned photographs of the target speaker and three foils, randomly ordered, with the typed names next to each photograph. Foils were carefully selected to challenge the listener to recognize the target voice.

The examiner presented subjects with the 4-choice response sheet; next she read the four names aloud and then played the voice. In this fashion, response alternatives were made available in visual, written, and spoken forms, compensating for specific neurological deficits. Subjects responded by pointing. Little or no verbal instruction was necessary. Three practice items, repeated until the subject understood the task, preceded the testing.

The unfamiliar voice stimuli were prepared using the voices of 10 male Southern Californians, 20-31 years of age, free of vocal pathology, who were recorded using a high-quality dynamic microphone (attached to the mouthpiece of the telephone) making telephone survey calls (Kreiman and Papcun, 1985). Sentences were excerpted from each call and a stimulus tape was constructed consisting of 26 pairs of voices. Subjects listened to a pair of voice samples, and were asked to indicate whether the samples were said by the same person or by two different people.

Experimental subjects were 45 normal and 26 unilaterally brain-damaged listeners, all were right-handed, native speakers of English and raised in the United States. Most had normal hearing. For the few with mild to moderate high frequency hearing losses, audiograms were examined, and no relation between performance on either listening task and patterns of hearing loss was observed. In the few patients with neglect syndrome, the tendency to ignore the left or right side was compensated for by vertical display on the response cards. Clinical subjects ranged in age from 45-82, with a mean age of 61.5. Normal listeners, all native Americans reporting normal hearing and vision, ranged in age from 50-85 with a mean age of 64.1.

RESULTS

The measure used for the familiar voice recognition task was the percentage of those voices known to a given subject (established by a questionnaire after the test) which were correctly recognized. For the unfamiliar voice discrimination task, the measure was the probability of a correct response. Because of listeners' varying familiarity with the foils in the recognition tasks, chance was set at 50% for both tasks. Results are significant at $p < .01$.

A two-way (group by task) repeated measures analysis of variance comparing the three groups on the two experimental tasks produced significant main effects of task ($F(1,71) = 1558.88$), and group ($F(2,71) = 12.98$) as well as a significant task by group interaction ($F(2,71) = 12.78$). The task by group interaction reflects the fact that both brain-damaged groups showed impaired discrimination abilities relative to normal performance, while only subjects with RBD were significantly impaired in recognizing familiar voices. On the discrimination task, the normal group performed significantly better than either the LBD or the RBD group (LBD: $F(1,61) = 11.09$; RBD: $F(1,57) = 20.85$); the LBD and RBD groups did not differ significantly on this task. In contrast, an effect of hemispheric side of lesion was observed for the familiar voice recognition test: whle LBD subjects did not differ from normals on the familiar voice recognition task, RBD subjects performed significantly less well than either LBD ($F(1,24) = 11.14$) or normal subjects ($F(1,57) = 21.53$) (Figure 2). This outcome was further supported by a second one-way ANOVA, which also showed a significant main effect of group ($F(1,37) = 12.66$) for a larger group (n = 39) of LBD and RBD subjects who performed the recognition task.

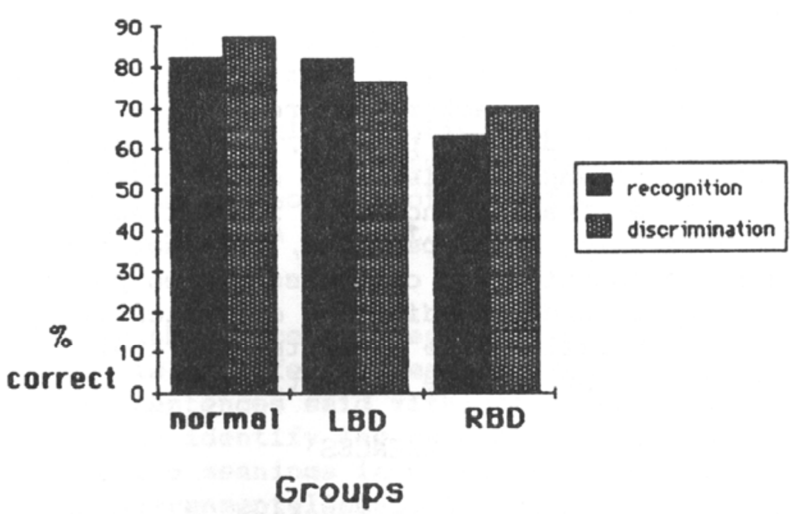

Figure 2. Mean performance scores of left-brain damaged, right-brain damaged, and normal-control groups on voice recognition and voice discrimination.

DISCUSSION

Both LBD and RBD subjects were impaired in unfamiliar voice discrimination, whereas only patients with RBD were impaired in familiar voice recognition. In individual subjects, the dissociation of these two abilities occurred in both directions--preservation of recognition with discrimination impairment, and preserved ability to discriminate alongside a deficit in recognition. This observation indicates that these two are independent and unordered abilities with different neuroanatomic substrates.

How is this dissociation between voice recognition and voice discrimination to be explained? We propose that the two tasks involve different psychological abilities. Recognizing a familiar voice engages a pattern-recognizing mechanism, whereby a holistic Gestalt (the unique voice pattern) is matched to a person. This ability is specialized in the right hemisphere--hence the finding that RBD patients are significantly more likely than LBD patients to be impaired in familiar voice recognition. On the other hand, unfamiliar voice discrimination requires featural analysis as well as overall pattern recognition. To tell whether unfamiliar voices are the same or different, details of the voice samples must be isolated and compared. Since featural analysis is specialized in the left hemisphere, we would expect voice discrimination performance to be depressed following LBD, which did occur. However, auditory pattern-matching is doubtless also used in the discrimination task. Thus we would expect voice discrimination performance to be affected by RBD, which also did occur. Anatomical studies using data from CT scans have indicated that brain damage near the auditory receiving areas in the temporal lobes of either hemisphere will affect discrimination ability, whereas voice recognition ability is affected only in cases of damage to the lateral aspect of the right parietal lobe.

The preservation of familiar voice recognition abilities was observed in LBD patients with severe language comprehension deficits, including those with global aphasia. That a severely aphasic patient is able to recognize nonverbal information in speech may be of value to the patient and to his or her family and clinician. The personal identity of the speaker is only one small part of the paralinguistic information embedded in the prosodic material of speech. Other information, such as personal identity, sex, age, personality, mood, emotion, and background are also available as paralinguistic information. Thus the aphasic patient, who cannot produce spontaneous speech and does not understand _what_ is being said, may know quite well _who_ is saying it and _how_. This capacity should be identified and nurtured in these patients. Furthermore, for RBD patients, the detection of a selective deficit in speaker recognition and other associated parameters of prosody will be of importance in understanding the communicative problems of this clinical group, in counselling the patients and their families, and in designing a course of therapy for the RBD patient.

REFERENCES

Assal, G., Zander, E., Kremin, H. and Buttet, J. Discrimination des voix lors des lesions du cortex cerebral. _Archives Suisse de Neurologie, Neurochirurgie et de Psychiatrie_, 119, 307-315, 1976.

Benton, A. The neuropsychology of facial recognition. _American Psychologist_, 35, 176-186, 1980.

Gandour, J. and Dardaranda, R. Identification of tonal contrasts in Thai aphasic patients. _Brain and Language_, 18, 98-114, 1983.

Heilman, K., Scholes, R. and Watson, R. Auditory affective agnosia. Journal of Neurology, Neurosurgery and Psychiatry, 38, 69-72, 1975.
Kreiman, J. and Papcun, G. Voice discrimination by two listener populations. UCLA Working Papers in Phonetics, 61, 45-60, 1985.
Levine, S.C. and Koch-Weser, M.P. Right hemisphere superiority in the recognition of famous faces. Brain and Cognition, 1, 10-22, 1982.
Ross, E. The aprosodias: Functional-anatomical organization of the affective components of language in the right hemisphere. Archives of Neurology, 38, 561-569.
Van Lancker, D. and Canter, G.J. Impairment of voice and face recognition in patients with hemispheric damage. Brain and Cognition, 1, 185-195, 1982.
Van Lancker, D. and Fromkin, V.A. Hemispheric specialization for pitch and 'tone': Evidence from Thai. Journal of Phonetics, 1, 101-104, 1973.
Van Lancker, D. and Fromkin, V.A. Cerebral dominance for pitch contrasts in tone language speakers and in musically trained and untrained English speakers. Journal of Phonetics, 6, 19-23, 1978.
Van Lancker, D., Kreiman, J. and Emmorey, K. Familiar voice recognition: patterns and parameters. Part I: Recognition of backward voices. Journal of Phonetics, 13, 19-38, 1985.
Van Lancker, D., Kreiman, J. and Wickens, T.D. Familiar voice recognition: Recognition of rate-altered voices. Journal of Phonetics, 13, 39-52, 1985.

DISCUSSION

Q: Do you know if any of your RBD subjects might have had amusia?

A: That would certainly be the next thing to look at. Amusia is so difficult because of the many components of music. They may have a problem in timbre recognition, or familiar melody recognition or unfamiliar melody recognition or in the whole production aspect of music. Music has so many different parameters that we have not begun to think about how to begin making that kind of comparison. We do ask them anecdotally whether they can sing a song with us but we have not investigated this.

C: It seemed to me that the only comparison you were drawing (and you were doing that at the beginning) was with face recognition and you said that there were some peoople who did not have face recognition deficits who did have a voice recognition deficit.

A: That's right.

C: That may be because the voice problem is an auditory-specific ability.

A: What we are doing is an affective-meaning recognition task, in which a patient listens to sentences said with affective intonation (sad, happy, etc.) and they are to identify the emotions; I am comparing deficits in recognizing affective meanings in speech with voice perception deficits. We also have an environmental sound identification task. And so these three auditory tasks--voice perception, affective information in speech, and environmental sounds--are what I would like to compare in individual subjects.

Q: You mentioned the possibility that we might want to work with this problem in the RBD person. How would the person present this problem?

Would he come in and say "I can't recognize familiar voices any mor How would this problem be presented?

A: Ordinarily I believe the voice recognition problem would be camoufla because there are so many other strategies for recognizing people besi by their voices; typically the face is present for recognition of person, except while talking on the telephone. The voice recognit deficit is not really going to be a serious problem in somebody's l and I think it would have to be found out in a patient by spec testing. You may then be able to counsel them, saying that when some calls on the telephone, the patient may want to ask who it is, explain that they're having trouble recognizing their voice. There very well-defined and organized phone behaviors, and it's very import to recognize an intimate's voice, or they'll be hurt and there will trouble in the interaction. So there could be some awkwardness troubles over the phone. But beyond that, my point about vo recognition deficits implying other prosodic-recognition deficits, tha what I think it's more important to look for--the affective, attitudinal, that's what may well be missing in the RBD peerson, those will affect their communicative ability in a much bigger way.