Voice Perception Deficits: Neuroanatomical Correlates of Phonagnosia*

Diana Roupas Van Lancker  
Department of Neuroscience  
University of North Dakota Medical School

Jody Kreiman  
Department of Linguistics  
UCLA

Jeffrey Cummings  
Los Angeles Veterans Administration Medical Center (Brentwood Division)

ABSTRACT

Voice perception (recognition of familiar voices and discrimination of unfamiliar voices) was studied in brain-damaged patients and normal controls. Left- and right-brain-damaged subjects were tested on familiar voices (25 famous males) and 26 pairs of unfamiliar voices. Deficits in recognizing familiar voices were significantly correlated with right-hemisphere damage; discrimination of unfamiliar voices was worse in both clinical groups than in normal controls. Computerized tomographic scans indicated that an intact right parietal-lobe was present in all cases of normal voice recognition, while right parietal-lobe damage was significantly correlated with a deficit in voice recognition. Temporal-lobe damage of either hemisphere was associated with a voice discrimination deficit.

Neuroanatomical substrates of voice perception have been little studied, despite the importance of this ability. To study the perception and cognition of human voices, it is important to distinguish between voice discrimination and voice recognition (Bricker & Pruzansky, 1976). Discrimination abilities involve judgments about unfamiliar voices, whereas voice recognition involves identification of a speaker known to the listener.

We wish to report our findings on studies of voice perception deficits, or phonagnosia, in brain-injured subjects, using both kinds of processing (recognition and discrimination). Individuals in four groups--unilaterally left-brain damaged (LBD), unilaterally right-brain damaged (RBD), bilaterally brain

* We wish to express our appreciation to Bruce Dobkin, M. D., Bette Hadler, M. S., Wayne Hanson, Ph.D., E. Jeffrey Metter, M. D., Gail Monahan, M. S., and Douglas Noffsinger, Ph. D. for assistance in obtaining patient information; D. Wesley Burgess, Ph. D., helped in preparing stimulus materials. This project was supported in part by the Veterans Administration.

Address correspondence to: Dr. Diana Van lancker, Department of Neuroscience  
University of North Dakota Medical Education Center, 1919 Elm St. Fargo, ND 58102, USA.

Accepted for publication: June 20, 1988
damaged (BBD), and normal control subjects—were tested on two tasks, familiar voice recognition and unfamiliar voice discrimination. Our purpose was to compare performance on these two tasks and to establish neuroanatomic substrates of these different aspects of voice perception.

In this report, we present data extending our findings that voice discrimination and voice recognition are dissociated neurological functions (Van Lancker, Cummings, Kreiman & Dobkin, 1988; Van Lancker & Kreiman, 1988), and we explore the specific neuroanatomical correlates of observed deficits in the two abilities.

METHODS

Stimuli
The unfamiliar voice stimuli were obtained from recordings of 10 male Southern Californians selected from an original set of 22 who were recorded while making telephone survey calls (Kreiman, 1987). Speakers ranged from 20-31 years of age, and were matched for regional accent. All were free of vocal pathology.

Using a survey sheet designed for this purpose, speakers were instructed to make telephone survey calls to a prearranged party at each of two recording sessions separated by at least 1 week. Rather than obtaining the voice recording over telephone lines, the speaker was recorded directly via a high-quality dynamic microphone attached to the mouthpiece of the telephone. This made it possible to obtain good-quality recordings while allowing the speaker to carry on a normal telephone conversation; additionally, only the speaker’s voice (and not that of the interviewee) was recorded. One call was selected from each recording session for use in our studies.

Sentences were excerpted from each speaker using a wave-form editing computer program. A stimulus tape was constructed consisting of 26 pairs of voice samples. For 13 pairs, the two samples represented the same speaker, in the other 13 pairs, they represented two different speakers. Within a pair of voices, the speakers always said the same thing; when speakers were the same, one utterance was taken from two survey calls, so listeners never compared two identical stimulus tokens. Each of the 10 voices appeared equally frequently; each voice occurred as both the first and second member of a “voices different” pair and in at least one “voices same” pair. Mean performance scores were known for the voice pairs used. As far as possible, “voices different” and “voices same” trials were matched for difficulty: 20 normal listeners aged 47-72 (mean age = 58.8, SD = 7.88) made an average of 4.08 errors on the “voices different” trials and 4.15 errors on the “voices same” trials (Kreiman, 1987).

Response sheets for this task consisted of a column of 26 same/different choices. Clinical subjects were asked to say “same” or “different” (or to point to the written words “same” or “different” if a speech difficulty interfered with a spoken response) on hearing each stimulus pair. Normal subjects circled “S” or “D” on an answer sheet.

The 25 familiar voice stimuli consisted of brief excerpts of texts spoken by famous males politicians and entertainers (e.g., John F. Kennedy, Johnny Carson). Using a computer wave-form editing program, the voice samples were then edited to create 4-s stimuli free of pauses, background noises, and identifying content. A written pretest administered to a group of normal subjects demonstrated that no targets were identified at above chance from linguistic content alone.
tested on two tasks, imitation. Our purpose was to abolish neuroanatomic deficits in the two male Southern Californian subjects while making telephone years of age, and were instructed to make reading sessions separated over telephone lines, the ophone attached to the good-quality recordings version; additionally, recorded. One call was for editing computer of voice samples. For 13·13 pairs, they representers always said the same from two survey calls, so if the 10 voices appeared id member of a “voices performance scores were ent” and “voices same” (mean age = 58.8, SD = 14l and 4.15 errors on the same/different choices. to point to the written h a spoken response) on on an answer sheet. ects spoken by famous nny Carson). Using a hen edited to create 4-targets. A written pretest of targets were identified

For the clinical group, response sheets consisted of vertically aligned photographs of the target speaker and three foils, randomly ordered, with typed names next to each photograph. Foils were matched to the target speaker for speaking style and perceived voice quality to challenge the listener to actually recognize the target voice and not to use cues from content, rhetorical style (i.e., politician vs. comedian) or other deductive strategies. For each test item, brain-damaged subjects were presented with the 4-choice response sheet; the four names were then read aloud and the stimulus voice played. Response alternatives were thus made available to these subjects in visual, written, and spoken forms (to compensate for any specific impairments in language comprehension, facial recognition or reading). Normal subjects circled one of 5 written names on an answer sheet.

At the end of this test session, for each target speaker, subjects were asked whether they felt they would normally recognize the voice. Responses were scored only for those voices a given listener claimed were familiar. This ensured that performance scores measured familiar voice recognition, and not merely familiarity with the set of test voices.

Subjects In all, 56 brain-damaged patients were tested. One subgroup consisted of 23 LBD, 15 RBD patients, and 6 BBD subjects who were tested on both the voice recognition and the discrimination protocols, extending the experimental group from the previous study, which reported on 15 LBD, 11 RBD, and 6 BBD subjects (Van Lancker & Kreiman, 1988). The other subgroup consisted of 12 unilaterally brain-damaged patients (2 LBD, 10 RBD) who were tested with familiar voice recognition only.

All the BD subjects had cerebral infarctions; all except two were right-handed; all, without exception, were native speakers of American English, and were educated in the United States. Site of lesion was determined by computerized tomography (CT), supported by clinical and neurological data. CT scans were unenhanced; slices were obtained at 10 mm intervals with the patient’s head in a standardized position (15 degrees from the canthomeatal line).

The LBD patients were tested using standardized language assessment instruments (Boston Diagnostic Aphasia Examination and Western Aphasia Battery). All were aphasic (7 fluent, 12 nonfluent, 6 anomic). In the LBD group, ages ranged from 41 to 80, with a mean of 61 years (SD = 8.67). RBD subjects ranged in age from 34 to 82, with a mean of 62.5 years (SD = 10.54). The BBD subjects were ages 57 to 82, with a mean of 71 years (SD = 10.38). In all three clinical groups, years of schooling ranged from 2 years of high school to postgraduate work. The normal control group consisted of 48 subjects aged 51 to 83 years, with a mean of 64 years (SD = 8.73). Their education ranged from 4 years of high school to postgraduate work.

Two kinds of analyses were conducted. First, a statistical analysis of group performance data compared normals (n = 48) and the unilaterally damaged patients (n = 38) who had performed both the recognition and discrimination protocols. Group means were obtained and an analysis of variance (ANOVA) performed to determine the effect of hemispheric side on lesion of discrimination and recognition abilities. Second, a slightly larger group on whom CT-scans were available (n = 43) was used to determine intrahemispheric lesion sites correlated with deficits in voice recognition and discrimination. A quantitative analysis of observed clinicopathological correlations is presented.
RESULTS

Statistical Analyses
Mean scores for each group (LBD, RBD and normal) on the recognition and discrimination tasks are given in Table 1. A two-way (group x task) ANOVA with repeated measures on task showed a significant effect of group on score (F(2,83) = 24.40, p < .001), a significant effect of task (F(1,83) = 4.58, p < .04), and a significant group x task interaction (F(2,83) = 3.64, p < .03). Post-hoc comparisons showed that, on the recognition task, LBD subjects did not differ from normals (F(1,69) = .42, n.s.), while RBD subjects performed significantly worse than either normals (F(1,61) = 28.58, p < .001) or LBD subjects (F(1,36) = 13.07, p < .001). Both LBD and RBD subjects performed significantly worse than normals on the discrimination task (LBD: F(1,69) = 12.84, p < .01; RBD: F(1,61) = 28.01, p < .01). The two clinical groups did not differ in performance on this task. Not surprisingly, findings are similar to those reported in Van Lancker and Kreiman (1988).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Means and standard deviations of percent correct on recognition and discrimination tasks for left-brain damaged (LBD), right-brain damaged (RBD), and normal control subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task</td>
<td>LBD</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Recognition</td>
<td>80.10 (13.8)</td>
</tr>
<tr>
<td>Discrimination</td>
<td>77.07 (13.3)</td>
</tr>
</tbody>
</table>

In summary, an impairment in familiar voice recognition was again found in association with RBD, whereas a deficit in unfamiliar voice discrimination was observed to be associated with damage to either hemisphere. LBD patients were impaired relative to normal subjects only in the voice discrimination task; RBD patients were impaired in both familiar voice recognition and unfamiliar voice discrimination. This extends the preliminary findings (Van Lancker & Kreiman, 1988), that recognition and discrimination of voices are separate and independent abilities with different neuroanatomical substrates. We sought to further investigate the precise nature of these neuroanatomical differences by correlating anatomical lesions with clinical and radiographic data in these patients, on whom voice perception abilities had been established.

1 Note that in the recognition task, chance performance floats between 25% correct and 50% correct, due to listeners' varying familiarity with the foils. Chance in the discrimination task is fixed at 50% correct. We conservatively assume that chance is set at 50% for both tasks.
normal) on the recognition and two-way (group x task) ANOVA significant effect of group on score (F ask (F(1,83) = 4.58, p < .04), and .83) = 3.64, p < .03). Post-hoc task, LBD subjects did not differ subjects performed significantly < .001) or LBD subjects (F(1,36) cts performed significantly worse: F(1,69) = 12.84, p < .01; RBD: ups did not differ in performance similar to those reported in Van

Correct on recognition amaged (LBD), right-brain control subjects

<table>
<thead>
<tr>
<th></th>
<th>RBD</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pup</td>
<td>58.57 (23.0)</td>
<td>82.10 (11.39)</td>
</tr>
<tr>
<td></td>
<td>68.21 (17.5)</td>
<td>87.19 (10.0)</td>
</tr>
</tbody>
</table>

e recognition was again found in familiar voice discrimination was er hemisphere. LBD patients were voice discrimination task; RBD recognition and unfamiliar voice indings (Van Lancker & Kreiman, f voices are separate and indepenent substrates. We sought to further atomical differences by correlating raphic data in these patients, on ablished.

nance floats between 25% correct and ith the foils. Chance in the discriminay assume that chance is set at 50% for

Clinical and Radiographic Analyses

More precise neuroanatomic substrates of phonagnosia were sought. Two types of analysis were performed. First, a descriptive study was undertaken, to establish which lesions sites were associated with voice perception deficits. Second, a quantification of the data was performed, testing hypotheses derived from the descriptive analysis.

For the descriptive analysis, we examined records of 56 patients (25 LBD, 24 RBD, and 6 BBD patients) tested on one or both voice protocols. Radiographic data were used to determine where in the right hemisphere (RH) damage was most often associated with a familiar voice recognition deficit, and which sites in either hemisphere were correlated with deficient unfamiliar voice discrimination.

CT scans of the head were obtained for 43 of the patients (22 LBD, 16 RBD, 5 BBD). The regions of infarction were mapped onto anatomical templates (Damasio, 1983; Matsui & Hirano, 1978). Lesion sites were classified by determining the lobe with the largest extent of the lesion (e.g., parietal, temporal, frontal, occipital). If the lesion involved two or three lobes equally, a combined lobar assignment was made (e.g., fronto-parietal, temporo-parieto-occipital). Subcortical lesions were indicated according to whether the lesion involved primarily the basal ganglia or the thalamus.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Numbers of patients (n = 43) distributed by lesion location and voice discrimination or recognition deficit. (R = right-sided lesion; L = left-sided lesion; B = bilateral lesion.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Deficit</td>
</tr>
<tr>
<td>Parietal</td>
<td>R 0</td>
</tr>
<tr>
<td>Lesion</td>
<td>L 1</td>
</tr>
<tr>
<td></td>
<td>B 1</td>
</tr>
<tr>
<td>Temporal</td>
<td>R 0</td>
</tr>
<tr>
<td>Lesion</td>
<td>L 1</td>
</tr>
<tr>
<td></td>
<td>B 0</td>
</tr>
<tr>
<td>Parietal &amp;</td>
<td>R 0</td>
</tr>
<tr>
<td>Temporal</td>
<td>L 3,2*</td>
</tr>
<tr>
<td>Lesions</td>
<td>B 0</td>
</tr>
<tr>
<td>Lesion</td>
<td>R 1,3*</td>
</tr>
<tr>
<td>Elsewhere</td>
<td>L 5</td>
</tr>
<tr>
<td>Totals</td>
<td>17</td>
</tr>
</tbody>
</table>

* took recognition test only.
Patients were considered deficient in voice recognition and discrimination if their performance differed by 1 SD or more from the normal mean. Using these criteria, patient information regarding lesion location and performance on the two voice protocols (Table 2) was used to obtain a Chi-Square statistic. From the available neurological data we could reliably classify patients into cells to test two hypotheses.

The first hypothesis is that abnormal performance in familiar voice recognition is associated with injury to the right parietal lobe; and the second hypothesis is that abnormal performance on unfamiliar voice discrimination is associated with damage to either right or left temporal lobes.

Subjects were sorted according to whether they attained high scores or low scores (more than 1 SD away from the normal mean) on either task, and whether or not they had documented damage to the right parietal (for the recognition task) or to either temporal lobe (for the discrimination task). A separate matrix was prepared for each task. That is, for the recognition task, each subject was identified as “high” or “low” on performance, and as having “right parietal lobe damage” or “damage elsewhere”; for the discrimination task, each subject was identified as “high” or “low” on performance and as having “left or right temporal damage” or “damage elsewhere.”

Table 3 shows the results for voice recognition. As was seen in the descriptive study, this analysis also shows that no subjects having right parietal lobe damage performed normally on voice recognition. Twenty-seven patients with lesions elsewhere than the right parietal lobe (LBD = 19, RBD 6, BBD = 2) did perform normally on the voice recognition task. Nine patients (7 RBD and 2 BBD) having damage to the right parietal lobe showed impaired performance on the task. Exceptions to the pattern are six patients with damage elsewhere than the right parietal lobe who also showed impaired performance on familiar voice recognition. The Chi-square (with 1 degree of freedom, with Yates’ correction for continuity) = 15.96 (p < .01).

<table>
<thead>
<tr>
<th>Table 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrix for voice recognition scores ($\chi^2 = 15.96, p &lt; .01$).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RIGHT</th>
<th>ELSEWHERE</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI</td>
<td>0</td>
</tr>
<tr>
<td>LO</td>
<td>9</td>
</tr>
</tbody>
</table>

Several con...
The matrix for unfamiliar voice discrimination is shown in Table 4. Patients were sorted by their high or low scores on the task and by the presence or absence of a lesion in either left or right temporal lobe, or elsewhere in the brain. CT scans or radiologists’ reports were available for 25 patients on whom voice discrimination scores had been obtained. The presence of a temporal lesion on either side was significantly associated with low performance on voice discrimination. Of the 25 patients examined, 13 with temporal-lobe damage had deficient performance on the voice discrimination task, while 9 without temporal-lobe damage (damage elsewhere) performed normally on the task. Eight patients (of 25) did not conform to this pattern: Four patients with temporal-lobe damage performed in the normal range; all had left-hemisphere lesions. In addition, four patients with damage elsewhere performed poorly on the task. Two had large, right-sided subcortical damage undercutting the temporal lobe, a third had a large left frontal lesion apparently sparing the temporal lobe, and the fourth had anterior-parietal damage on both sides, possibly involving superior-posterior portions of the temporal lobes. Chi-square (with 1 degree of freedom, including Yates’ correction for continuity) = 4.54, p < .05.

Table 4

<table>
<thead>
<tr>
<th>Damage</th>
<th>LEFT TEMPORAL LOBE</th>
<th>ELSEWHERE</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>LO</td>
<td>13</td>
<td>4</td>
</tr>
</tbody>
</table>

DISCUSSION

Several conclusions can be drawn from the current study. A RH specialization for familiar voice recognition abilities was found, supporting similar findings of Van Lancker and Canter (1982) and Van Lancker and Kreiman (1988). Assal, Zander, Kremin, and Buttet (1976) reported only a trend toward an association of RH damage with voice perception deficits, but the lack of a strong RH finding in that study may reflect the use of unfamiliar voices in a discrimination task. This study gave further evidence of a previously reported double dissociation for the two aspects of voice perception, recognition, and discrimination (Van Lancker et al., 1988). Of 43 brain-damaged patients performing both tasks, there were 7 who could discriminate well but recognized voices very
poorly (below 2 SD from the normal mean), and 9 patients who could recognize well but not discriminate voices (below 2 SD). The double dissociation, along with the different neuroanatomical substrates described in this paper, indicates that the two skills are independent. Dissociation of recognition and discrimination abilities has been reported for a similar faculty in the visual modality, face perception (Benton & Van Allen, 1972; Malone, Morris, Kay, & Levin, 1982; De Renzi, Faglioni, & Spinazzola, 1968).

The correlation of familiar voice recognition deficits with right parietal lesions is compatible with current theories about RH function. First, voice identity is a type of nonverbal information carried in speech prosody, and several studies have found nonlinguistic-prosodic comprehension to be mediated by the RH (Heilman, Shores, & Watson, 1975; Kent & Rosenbek, 1982; Ross, 1981). Further, the voice is a complex auditory pattern. A current theory of hemispheric specialization describes the right hemisphere as a holistic pattern recognizer, whereas the left hemisphere excels at sequential and featural analysis (Bever, 1975; Bogen, 1969; Bradshaw & Nettleton, 1983). This model ascribing different modes of processing to left and right cerebral hemispheres might be invoked to explain the differences seen in this study between performance of familiar-recognition and unfamiliar-discrimination voice tasks. Familiar voice recognition requires the association of an auditory pattern with a mental “trace” of a person, as has been postulated for face recognition by Damasio, Damasio, & Van Hoeven (1982). Such cognitive processes occur in the posterior “heteromodal” association areas (Mesulam, 1985). In contrast, evaluating unfamiliar voices is more purely a “unimodal” auditory function, and requires comparison of selected features in the voice pattern without multimodal processing or reference to another stored entity. This activity engages the temporal lobe, with its specialized auditory abilities and, for optimum performance, requires both hemispheres, as both featural and holistic strategies are needed. These neuroanatomical findings give further support to the notion that recognizing a familiar voice is neuropsychologically different from distinguishing among unfamiliar voices.

This study represents, in part, an effort to “map” the RH, in a tradition of localizing functions, and suggests that voice recognition is disturbed by right parietal lesions, whereas voice discrimination may be disrupted by lesions of either temporal lobe. This lesion approach leads to the conclusion that the right parietal lobe is principally responsible for mediating voice recognition, and functions of both temporal lobes are necessary for competent voice discrimination. The few observed exceptions to these rules may be attributed to plastic central nervous system properties, individual differences in brain organization, alternate strategies for performing tasks, neural reorganization, and recovery of neural function after partial brain damage.

Voice recognition plays a key role in phylogenetic history, infant development, and in communicative interaction. Recent field work with nonhuman primates has shown that maternal Vervet monkeys recognize the voices of their own offspring and maintain recognition over the first 5 years of life. In macaques, peer recognition occurs in the first year of life, and it is the case that the social value of voice recognition is more pronounced in the third year of life. However, the importance of voice recognition is not limited to primate species. In nonprimate species, voice recognition is also used to identify conspecifics. For example, in the domestic cat, voice recognition is used to identify the owners of the household. Therefore, voice recognition is a fundamental aspect of communication in the animal kingdom.
ents who could recognize double dissociation, along
ed in this paper, indicates cognition and discrimin-
the visual modality, face
is, Kay, & Levin, 1982; De
icits with right parietal
function. First, voice
in speech prosody, and
prehension to be media-
. Kent & Rosenbek, 1982;
pattern. A current theory
sphere as a holistic pattern
ential and featural analy-
leton, 1983). This model
right cerebral hemispheres
this study between perfor-
ination voice tasks. Familiar
 auditory pattern with a
 for face recognition by
itive processes occur in
&mety, 1985). In contrast,
modal" auditory function,
the voice pattern without
dentity. This activity enga-
gibilities and, for optimum
ural and holistic strategies
ther support to the notion
ically different from distin-
"the RH, in a tradition of
ition is disturbed by right
be disrupted by lesions of
he conclusion that the right
ning voice recognition, and
-ent voice discrimina-
nay be attributed to plastic
ces in brain organization,
organization, and recovery of
stic history, infant develop-
field work with nonhuman
recognize the voices of their
own offspring (Cheney & Seyfarth, 1980), and that this information is used in
maintaining family contact. Human research has shown that 3-day-old infants
recognize the voices of their own mothers from a set of maternal voices (DeCasper
& Fifer, 1980), and that young children can recognize the voices of their
peers as well as can adults (Mann, Diamond, & Carey, 1979). For daily commu-
nicative interaction, it is obvious that voice recognition is crucial, and that it
occurs in tandem with language comprehension. The identity of the speaker and
the linguistic-phonetic information are extracted simultaneously from the
acoustic material of the speech signal. The data reported here indicate that,
while the left hemisphere processes the linguistic-phonetic information in
speech, the right hemisphere, from that same signal, establishes the identity of
the speaker.

REFERENCES

lésions du cortex cérébrale. Archives Suisses de Neurologie, Neurochirurgie, et de

Journal of Neurological Science, 15, 167-172.

Bever, T. C. (1975). Cerebral asymmetries in humans are due to the differentiation of two
incompatible processes: Holistic and analytic. Annals of the New York Academy of
Science, 19, 251-262.

Societies, 34, 73-105.

Cliffs, NJ: Prentice-Hall.

Contemporary issues in experimental phonetics (pp. 295-326). New York: Academic
Press.


basis and behavioral mechanisms. Neurology, 32, 331-341.


voices. Science, 208, 1174-1176.

unilateral brain damage on facial recognition tests. Cortex, 4, 17-34.


and Language, 15, 259-291.


