BRAIN AND LANGUAGE 34, 246–252 (1988)

Hemispheric Specialization for Voice Recognition: Evidence from Dichotic Listening

JODY KREIMAN AND DIANA VAN LANCKER

Phonetics Laboratory, Department of Linguistics, University of California at Los Angeles, Los Angeles, California 90024

To measure lateralization of voice recognition abilities in normal subjects, listeners identified both the speaker (a famous male) and the word spoken on each trial in a dichotic listening paradigm. The voice identification task resulted in a zero ear advantage, which differed significantly from the significant right ear advantage found for word identification. This suggests that voice and word information, although carried in the same auditory signal, engage different cerebral mechanisms. © 1988 Academic Press, Inc.

Dichotic listening studies of voice recognition have failed to produce clear evidence for or against lateralization of this function to either hemisphere. A right ear advantage (REA) (Doehring & Bartholomeus, 1971), a left ear advantage (LEA) (Riley & Sackeim, 1982), and zero ear advantages (Bartholomeus, 1974a, 1974b; Tartter, 1984) have been reported, as has a sex-of-speaker by ear interaction effect (Buttet, Thuillard, Assal, & Aubert, 1983).

The literature on voice recognition in brain-damaged patients provides a possible explanation for these varying results. Studies of familiar voice recognition and unfamiliar voice discrimination in this population suggest that recognition is lateralized to the right hemisphere (Van Lancker & Canter, 1982), while successful discrimination requires both hemispheres (Assal, Zander, Kremin, & Buttet, 1976; Van Lancker & Kreiman, 1987; Van Lancker, Kreiman, & Cummings, 1985). The inconsistent results produced by dichotic listening studies may thus be due to their exclusive use of *unfamiliar* voices: If both cerebral hemispheres are required for processing unfamiliar voice information, as data from clinical subjects suggest, then one would not expect a clear ear advantage to emerge

We thank Karen Emmorey for her help in preparing the stimuli used in this experiment.
Address correspondence and reprint requests to Jody Kreiman, Phonetics Laboratory, Department of Linguistics, University of California, Los Angeles, California 90024.

0093-934X/88 $3.00

TABLE 1
STIMULUS VOICES

Jack Benny	John F. Kennedy[a]
George Burns	Jack Klugman
Johnny Carson	James Mason
Maurice Chevalier[a]	Bob Newhart
Walter Cronkite	Richard Nixon
W. C. Fields	Vincent Price
Henry Fonda	Tony Randall
Alec Guinness	Ronald Reagan

[a] Practice voices.

when these are used in dichotic listening. In fact, the data on brain-damaged patients strongly suggest that a right hemisphere specialization exists for familiar voice recognition only (Van Lancker & Kreiman, 1987).

To examine hemispheric specialization for voice identification in normal subjects, we have therefore used a set of familiar voices as stimuli in the present study. To ensure the stimuli were uniformly familiar to an adequate number of subjects, famous male voices were selected. Our hypothesis was that individual voice quality is an auditory pattern, and that voices would therefore be better recognized at the left ear in a dichotic listening paradigm, even though carried in a speech stimulus. This hypothesis is based on a model of brain function that associates superior pattern recognition abilities with the right hemisphere (Bogen, 1969; Bradshaw & Nettleton, 1983; Levine & Koch-Weser, 1982), in contrast to the sequencing and analytic functions of the left hemisphere (Bever, 1975; Levy, 1974). This model further interprets superior performance at the left ear as a reflection of an underlying right hemisphere specialization for the genre of stimuli (e.g., Bryden, 1982; Kimura, 1967).

To test our hypothesis most stringently, listeners were required to identify both the linguistic and voice information that occurred at each stimulus presentation: Subjects monitored a single ear on each trial, and always identified both the word and the speaker. We hypothesized that the word identity of each stimulus would be better identified at the right ear (e.g., Kimura, 1967), while the specific identity of the speaker of the same stimulus would be better recognized at the left ear of the same subjects.

METHOD

Stimuli. From a set of 50 famous male voices previously studied (Van Lancker, Kreiman, & Emmorey, 1985; Van Lancker, Kreiman, & Wickens, 1985), the 16 voices familiar to the most listeners were selected for use in this experiment (Table 1). Two of the voices were designated practice voices and were used only in training trials.

To prepare the stimulus tapes, four words were excised from each speaker's recorded voice sample. Care was taken to avoid words which could cue the speaker's identity. Editing was done on the PDP-11/23 computer in the UCLA Phonetics Laboratory. Speech samples were low-pass filtered at 4 kHz and sampled at 10 kHz. Expanded displays were used so that word onsets and offsets could be accurately located. All words had two to four syllables; durations ranged from 250 to 650 msec, with a mean duration of 454.77 msec (SD = 109.33).

In order to avoid effects of specific pairings of words and voices, each speaker was paired with four other speakers (one for each excised word), for a total of 56 stimulus pairs. Words were matched to within 40 msec in duration. Individual stimulus words were "merged" to produce dichotic pairs with one word on each channel. Stimulus onsets were exactly aligned; alignments were verified on two-channel oscillograms. After merging, dichotic pairs were played out onto stereo tapes. Recording levels were equalized during play-out. Stimulus pairs were separated by 13 sec; each pair was preceded by its consecutive number, spoken in a female voice.

Subject selection. In order to ensure that subjects were genuinely familiar with all test voices, a listening pretest was prepared, made up of 4-sec samples of continuous speech for each target speaker. No test stimulus words appeared in these samples, nor did samples contain speech trademarks or typical topics which could cue the speaker's identity. The pretest samples were randomly ordered and presented to candidate subjects, who were asked to match each voice sample to a list of 25 names (the 14 target voices, 2 practice voices, and 9 distractors). Only listeners who correctly identified all of the target voices participated in the dichotic listening study.

The first 15 listeners to successfully complete the pretest were given the dichotic listening test. All subjects were right-handed with no familial sinistrality. All reported normal hearing in both ears.

Design and procedure. Dichotic stimuli were arranged into four blocks of 28 trials for a total of 112 trials. Headphone left–right orientation was switched after the first and third blocks, and initial orientation was alternated across subjects. Subjects were instructed to monitor one ear at a time and to ignore the competing stimulus in the other ear. To discourage subjects from developing a response set, the ear monitored was alternated every 14 trials. Within a block of 28 trials, each voice was monitored twice and ignored twice; a different word was used each time a voice occurred. In the next block, the two ignored words were monitored, and the monitored words ignored. Thus, across the four blocks of trials, each word was monitored twice and ignored twice, once through each side of the headphones.

Eight training trials preceded the test trials. For the first four trials, subjects were told in advance what voices and words they would hear; they then heard four trials for which post-trial feedback was provided. They were given the opportunity to repeat these training trials until they felt comfortable with the dichotic task.

For each test item, subjects reported both the word and the voice heard at the monitored ear by selecting one each of six words and six names provided on answer sheets. Names included the target voice, the voice in the unattended ear, and two foils for each (all from the target set). Care was taken to select foils which were similar enough in vocal quality to be plausible choices. Words included the target and the word in the unattended ear; word foils were selected by playing the tape to four naive listeners who wrote down the words they thought they heard in each ear. Their misperceptions were used as foils. No attempt was made to control order of report; however, for half the items, words were listed first on the answer sheet, and for the other half of the items, voices were listed first, alternated every other test item.

All instructions to subjects were recorded on the stimulus tape. The test was administered using an Aiwa HS-J02 tape player and Sony MDR-80 headphones. This apparatus was calibrated by observing the output (on an oscilloscope) when a 1-kHz tone was played

TABLE 2
SCORES ON VOICE AND WORD RECOGNITION TASKS

	Voices		Words	
	Left ear	Right ear	Left ear	Right ear
Percentage correct	45.83	50.25	65.11	76.78
SD	7.96	7.05	9.48	9.06

through each channel; relative output at the two ears was equal. Testing took approximately 40 min.

RESULTS

For each listener, the percentage correct recognition for each task (word and voice recognition) at each ear, as well as the laterality index λ (lambda; Bryden, 1982), was calculated for each task.[1] Mean percentage correct results are given in Table 2[2]; chance for both tasks equals 16.7% correct.

A two-way (task × ear) ANOVA revealed significant effects of task ($F(1, 14) = 122.74$, $p < .01$) and ear ($F(1, 14) = 10.77$, $p < .01$) on percentage correct: All subjects performed better on the word recognition task, and scores for both tasks were better overall at the right than at the left ear. A significant task × ear interaction also occurred ($F(1, 14) = 7.74$, $p < .01$): Performance was significantly better at the right ear for the word recognition task ($t\ (14) = 3.75$, $p < .01$), but no significant difference between ears occurred for the voice recognition task ($t(14) = 1.85$, n.s.). Results of analyses using λ confirmed the above finding: A significant REA emerged for the word recognition task ($z = 5.24$, $p < .01$), while no significant ear advantage was observed for voice recognition ($z = 1.81$, n.s.).

To investigate individual patterns of lateralization, following Lauter (1983) we calculated for each subject the difference in performance at the two ears for each task and plotted these to reveal individual patterns of relative ear advantages (Fig. 1). This analysis revealed that 5 of the 15 subjects showed essentially no difference in ear advantages for the

[1] λ is given by $\lambda = \ln((\text{Correct R} \times \text{Wrong L})/(\text{Wrong R} \times \text{Correct L}))$ and its significance is tested using z tables, where $z = \lambda/((1/\text{Correct L}) + (1/\text{Correct R}) + (1/\text{Wrong L}) + (1/\text{Wrong R}))$. λ has two advantages over other, more commonly used laterality indices: it does not fluctuate with differences in relative task difficulty, and significance levels may be calculated for individual subjects as well as for group data (see Bryden, 1982, for further discussion).

[2] Intrusion errors were also examined, but the small number that occurred (an average of 6.33 for voices and 1.8 for words) could have been made by chance, and no formal analyses were undertaken.

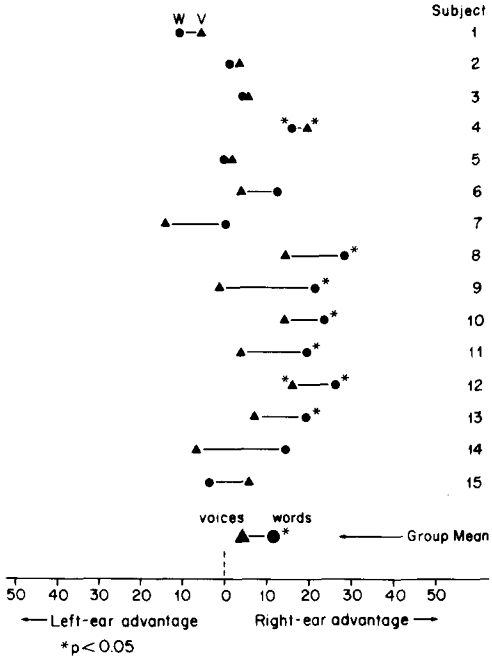

FIG. 1. Relative ear advantages for individual subjects and for the group as a whole on the voice and word recognition tasks.

two tasks, while the other 10 did show a difference in the extent to which the two tasks are lateralized: Although *absolute* ear advantages for the two tasks vary, for all but one subject in this group the word recognition task produced a large *relative* REA, and the voice task a relative LEA. The significance of individual differences in performance was tested using the λ measure; most subjects had no significant ear advantage for the voice recognition task, and about half (7/15) had a significant REA for the word recognition task ($p < .05$).

DISCUSSION

As a group, subjects in this experiment showed the expected significant REA for word recognition, coupled with a zero ear advantage for voice recognition. The "strong" version of our initial hypothesis—that a LEA for famous voice recognition (alongside a REA for word recognition in these bidimensional stimuli) would be observed—is thus not supported. However, a weaker version of this hypothesis *is* supported, in that a relative LEA for voice recognition compared to that observed for word recognition did emerge. This relative (and statistically significant) difference, found in a task where subjects were required to simultaneously process two different aspects of the same stimuli, suggests that voices and words are processed by different cerebral mechanisms.

A reasonable conclusion from these findings is that, whereas words are processed in the left hemisphere (as is known), voice information is processed differently from words, even though carried in the same auditory signal. Definitive corroboration from dichotic listening studies of a right hemisphere specialization for voice recognition, which has been observed in clinical studies (e.g., Van Lancker & Canter, 1982; Van Lancker & Kreiman, 1987), must await further experimentation.

One explanation of the lack of a significant absolute LEA for voices may come from examination of individual subjects' results. More than half our individual listeners did not produce the expected significant REA for word recognition. These results, along with the lack of a LEA for voices, may both be due to the "bidimensional stimulus" design of this experiment. That is, expected right and left ear advantages for different "levels" of the same stimulus might have been attenuated by the task, which had listeners simultaneously attend to both aspects. The lack of a LEA for voice recognition may also have been due to the salience of the linguistic content overriding the voice information. Subjects in fact reported that the word identification task was easier to focus on than was the voice identification task. This hypothesis (that the task format influenced the ear advantages) can be tested by dichotically presenting backward stimuli, which would be free of linguistic content ("one-dimensional stimuli"), and requiring only a voice identification response. If our interpretation is correct, a clear left ear advantage should emerge from such a design.

REFERENCES

Assal, G., Zander, E., Kremin, H., & Buttet, J. 1976. Discrimination des voix lors des lésions du cortex cérébral. *Archives Suisses de Neurologie, Neurochirurgie, et de Psychiatrie,* **119,** 307–315.

Bartholomeus, B. 1974a. Effects of task requirements on ear superiority for sung speech. *Cortex,* **10,** 215–223.

Bartholomeus, B. 1974b. Dichotic singer and speaker recognition. *Bulletin of the Psychonomic Society,* **2,** 407–408.

Bever, T. G. 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.

Bogen, J. 1969. The other side of the brain. *Bulletin of the Los Angeles Neurological Society,* **34,** 73–105.

Bradshaw, J. L., & Nettleton, N. C. 1983. *Human Cerebral Asymmetry.* Englewood Cliffs, NJ: Prentice–Hall.

Bryden, M. P. 1982. *Laterality: Functional asymmetry in the intact brain.* New York: Academic Press.

Buttet, J., Thuillard, F., Assal, G., & Aubert, C. 1983. *Human voice recognition: A dichotic listening test.* Presented at the International Neurological Society, Lisboa.

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

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

Lauter, J. 1983. Stimulus characteristics and relative ear advantages: A new look at old data. *Journal of the Acoustical Society of America,* **74,** 1–17.

Levine, S. C., & Koch-Weser, M. P. 1982. Right hemisphere superiority in the recognition of famous faces. *Brain and Cognition,* **1,** 10–22.

Levy, J. 1974. Psychobiological implications of bilateral asymmetry. In S. J. Dimond and J. G. Beaumont (Eds.), *Hemispheric function in the human brain.* New York: Wiley.

Riley, E. N., & Sackeim, H. A. 1982. Ear asymmetry in recognition of unfamiliar voices. *Brain and Cognition,* **1,** 245–258.

Tartter, V. C. 1984. Laterality differences in speaker and consonant identification in dichotic listening. *Brain and Language,* **23,** 74–85.

Van Lancker, D., & Canter, G. 1982. Impairment of voice and face recognition in patients with hemispheric damage. *Brain and Cognition,* **1,** 185–195.

Van Lancker, D., & Kreiman, J. 1986. Preservation of familiar speaker recognition but not unfamiliar speaker discrimination in aphasic patients. In R. Brookshire (Ed.), *Clinical aphasiology 1986.* Minneapolis, MN: BRK Publishers.

Van Lancker, D., & Kreiman, J. (1987). Voice discrimination and recognition are separate abilities. *Neuropsychologia,* **25,** 829–834.

Van Lancker, D., Kreiman, J., & Cummings, J. 1985. Voice recognition and discrimination: New evidence for a double dissociation. *Journal of Clinical and Experimental Neuropsychology,* **7,** 609.

Van Lancker, D., Kreiman, J., & Emmorey, K. 1985. Famous voice recognition: Patterns and parameters. Part I: Recognition of backward voices. *Journal of Phonetics,* **13,** 19–38.

Van Lancker, D., Kreiman, J., & Wickens, T. 1985. Famous voice recognition: Patterns and parameters. Part II: Perception of rate-altered voices. *Journal of Phonetics,* **13,** 39–52.