

# Hemispheric contributions to the integration of visual and auditory information in speech perception

K. BAYNES

*University of California, Davis, California*

M. G. FUNNELL

*Dartmouth College, Hanover, New Hampshire*

and

C. A. FOWLER

*Haskins Laboratories, New Haven, Connecticut*

Differential hemispheric contributions to the perceptual phenomenon known as the *McGurk effect* were examined in normal subjects, 1 callosotomy patient, and 4 patients with intractable epilepsy. Twenty-five right-handed subjects were more likely to demonstrate an influence of a mouthed word on identification of a dubbed acoustic word when the speaker's face was lateralized to the LVF as compared with the RVF. In contrast, display of printed response alternatives in the RVF elicited a greater percentage of McGurk responses than display in the LVF. Visual field differences were absent in a group of 15 left-handed subjects. These results suggest that in right-handers, the two hemispheres may make distinct contributions to the McGurk effect. The callosotomy patient demonstrated reliable McGurk effects, but at a lower rate than the normal subjects and the epileptic control subjects. These data support the view that both the right and left hemisphere can make significant contributions to the McGurk effect.

The modular organization of language is a common presumption of much work in the psychology and neuropsychology of language. Numerous researchers have examined the modular properties of input and output processes, visual and auditory processes, orthographic, phonological, and semantic processes, and so on (Besner & Humphreys, 1991; Caramazza, 1990; Coltheart, Sartori, & Job, 1987; Denes, Semenza, & Bisiacchi, 1988). As we better define these component processes of language performance, it is equally important that we understand how different forms of information are integrated.

To examine the integration of the visual and auditory processing streams in understanding spoken language, we have considered a phenomenon called the *McGurk effect* (MacDonald & McGurk, 1978; McGurk & MacDonald, 1976), which can occur when a syllable or word that is heard does not agree with a syllable or word that is seen. Seeing the vocal gestures that are used to produce /ga/, while hearing /ba/, yields perception of /da/. MacDonald and McGurk argue that their results rule out a simple sum-

mation of acoustic and optical information. Frequently, responses are fusions that reflect input from both modalities, and fusions differ depending on which information is presented optically and which is presented acoustically. Optical /ga/ with acoustic /ba/ is most likely to be identified as /da/, whereas optical /ba/ with acoustic /ga/ yields /bga/ or /ga/ most frequently. The McGurk effect is robust in that it is maintained even when the perceiver is aware of the dubbing (see Liberman, 1982). It would therefore appear that this effect documents some significant integration of phonological or phonetic information presented optically and acoustically that occurs below a level of conscious reflection.

The fact that the McGurk effect involves integration of phonetic and phonological information both from the acoustic and the optic consequences of speech production raises a question as to whether the integration involves specializations of just one or both cerebral hemispheres. Evidence from a variety of subject groups, including aphasics and dyslexics (Paterson & Besner, 1984; Paterson & Marcel, 1977), split-brain subjects (Sidtis, Volpe, Wilson, Rayport, & Gazzaniga, 1981; Zaidel & Peters, 1981), and normals under tachistoscopic presentation of stimuli (Geffen, Bradshaw, & Nettleton, 1972; Klatzky & Atkinson, 1971), suggests that the right hemisphere (RH) has limited phonological competence (see Baynes, 1990, for a review), whereas the intact left hemisphere (LH) handles a variety of phonological tasks with ease.

This research was supported in part by NIDCD Grant R29 DC00811 to K.B. and by NICHD Grant HD-01994 to Haskins Laboratories. We would like to thank Peter Williamson and Vijay Thadani for encouraging their patients to participate in this study and Carol Bruzewicz for scheduling them to see us. We would also like to thank William C. Lofus for help in recording and programming the video display. Address correspondence to K. Baynes, Center for Neuroscience, University of California, Davis, CA 95616 (e-mail: fzbaynes@bullwinkle.ucdavis.edu).

On this basis, one might expect a left-hemisphere locus for integration of phonological information presented optically and acoustically. However, a case might be made for greater right-hemisphere participation in the processing of any facially signaled information, including the extraction of phonological information from an optic signal, because the right hemisphere so frequently has been shown to have the advantage in processing faces, particularly when attention must be placed on a particular feature, such as the lips (Hillger & Koenig, 1991). Ellis (1989) has speculated that there are multiple processing channels for facial and voice recognition. He suggests that the left hemisphere is responsible for the processing of phonetic information provided by the voice and the face, whereas the right hemisphere specializes in the extraction of information about identity and affect from faces as well as from voices. However, there are many indications that both hemispheres make some contribution to the processing of faces. In normal subjects, some studies have documented a left-hemisphere advantage for tasks that require the processing of faces (Patterson & Bradshaw, 1975). Although prosopagnosia is generally reported following right-hemisphere lesions, postmortem studies (Damasio & Damasio, 1986) have discovered bilateral damage in all patients who were prosopagnosic.

Given these different empirical bases on which to make predictions, it is of some interest to resolve the question of whether there are hemispheric differences in extracting phonological information from the face. If the operative consideration is that the optical signal provides *phonological* information, then one might (following Ellis, 1989) expect a left-hemisphere advantage in extracting that information from the face and using it to influence identification in a McGurk procedure. This might be the expectation of motor theorists of speech perception (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985), who hold that speech perception is accomplished by a special-to-speech brain center that accepts information from either the auditory or the visual modality. Alternatively, if the operative consideration is that the phonological information is being conveyed by a *face*, it is possible that a right-hemisphere advantage for optically presented phonological information will be found. Campbell (1986; Campbell, Landis, & Regard, 1986; Campbell et al., 1990) have produced data ostensibly supporting both alternatives.

Campbell et al. (1986) approached the localization issue by using 2 subjects with complementary right- and left-hemisphere lesions. These 2 patients showed a clear behavioral dissociation between phonological and facial processing. The first patient, D., with a right posterior cerebral artery (PCA) infarct, had difficulty recognizing faces and deriving information pertaining to sexual identity and emotional valence from pictures of faces. However, she was able to recognize speech gestures and to lip-read accurately. In contrast, Patient T., with a left PCA infarct, recognized faces and had no difficulty making judgments about sex or emotion. She was unable to lip-read. It is of particular interest that Patient D. showed

the normal influence of visual words on auditory words under McGurk conditions, but Patient T. did not. This pattern of results seemed to argue quite clearly that lipreading skills are lateralized to the left hemisphere, and facial processing is a right-hemisphere process. The assumption of interest is that lipreading involves the translation of optical information about the face into a phonological code, or at least a code that can be compared to a phonological code.

The clarity of this result was marred by Campbell's (1986) own observations with lateralized presentations to normal subjects. In an experiment that tachistoscopically lateralized still photographs of a female speaker mouthing speech sounds that were congruent or incongruent to heard sounds, Campbell found a significant LVF/RH (left visual field/right hemisphere) reaction time advantage in deciding whether the visual and audible syllables matched. In a second experiment, in which subjects named consonants or vowels being mouthed by each half of a chimeric face, the LVF/RH advantage was reduced, but still present. An LVF/RH advantage was obtained when accuracy on an LVF presentation of the right side of a face was compared with accuracy to its mirror presentation in the RVF (right visual field). No overall visual field advantage was obtained when the normal left side of a face presented in the RVF was compared with its mirror reflection in the LVF. This difference may reflect the fact that the right side of a speaker's face provides more expressive speech information (Graves, Goodglass, & Landis, 1982). Across the two experiments, however, the main finding with regard to perception was one of overall LVF/RH superiority in performance. To explain her findings, Campbell considered, but discarded, the possibilities that stimulus conditions, task difficulty, or task novelty could account for the right-hemisphere advantage she had observed. She raised the possibility that the diffuse nature of right-hemisphere representation might favor cross-modal matching, as had been suggested by Semmes (1968). Campbell speculated that the processor that gave rise to the effects that she observed might be similar to that responsible for the McGurk effect.

With the localization issue unresolved, Campbell et al. (1990) studied the McGurk effect in 4 patients with neurological disorders. Cases 1 and 2 had left-hemisphere damage. Case 1, who had word-meaning deafness, could lip-read a little and showed a McGurk effect for consonants, but demonstrated visual dominance for vowels (i.e., in audiovisual trials, made the visually compatible choice). Case 2, the pure alexic studied in 1986, was still unable to lip-read and made choices dominated by the auditory stimulus. Both of the right-hemisphere-damaged patients were prosopagnosic. Case 3 developed the disorder in childhood; she could lip-read, but did not use the visual information in audiovisual trials and was dominated by the auditory stimulus. Case 4, also studied in 1986, continued to have difficulty extracting information about gender or age from the face. However, she was able to lip-read, and she performed normally on the McGurk task.

Campbell et al. (1990) argued from these results that the left hemisphere has a special role in the integration of seen and heard speech. They suggested that there is a single phonological processor that is indifferent to the modality of the input it receives. For Case 1, auditory input is blocked, and for Case 2, visual input is blocked. Both Cases 3 and 4 can lip-read, but Case 3 doesn't show the effect, possibly because of a more pervasive linguistic deficit due to her developmental disorder, or possibly because the visual input into the phonological processor is blocked for Case 3. Campbell et al. (1990) account for the results of Campbell (1986) by suggesting that input from right-hemisphere visual areas may have some privileged input into the speech processor.

This account is not totally satisfying. Although both of the studies involving patients (Campbell et al., 1990; Campbell et al., 1986) directly examined the McGurk phenomenon, in the studies of normal subjects (MacDonald & McGurk, 1978; McGurk & MacDonald, 1976) a distinctly different task was employed. In the Campbell task, still photographs were used and a comparison between heard speech and the pictured lip position was explicitly required. Campbell speculated that these skills might be the same ones that underlie the McGurk effect. Compatibly, some associations between lipreading skills and the strength of the McGurk effect were observed in the patients (except for Case 3). However, the hemispheric contributions to the visual and auditory integration that yield the McGurk percept remain unclear.

We hoped to provide additional information on lateralization of extraction of phonological information from the face. Rather than present still pictures tachistoscopically to the left and right visual fields of normal viewers, as Campbell (1986) had done, we lateralized video presentations of a speaker mouthing words. We also examined the possibility of eliciting the McGurk effect in the right hemisphere of a callosotomized subject with documented right-hemisphere language, but no speech.

We investigated two questions with regard to the McGurk effect: (1) Is there a distinctive pattern of lateralization associated with the McGurk effect? and (2) Is access to the left-hemisphere speech system necessary for the McGurk effect to occur? The first question is addressed in Experiment 1. The second question was approached in Experiment 2 with J.W., a patient with a complete callosotomy and right-hemisphere reading ability, but no access to speech mechanisms.

## EXPERIMENT 1

### Method

**Subjects.** The subjects were 40 undergraduates at Dartmouth College who participated in this study for course credit in an introductory psychology course. Twenty-five of the subjects were right-handed and fifteen were left-handed. All had normal hearing, normal or corrected vision, and were native speakers of English.

**Stimulus materials.** Stimulus materials were derived from those used by Dekle, Fowler, and Funnell (1992). The nine trial types listed in Table 1 were taken from the videotapes recorded by Dekle et al. Of these, four had word-initial /b/ on the sound track, paired with word-initial /v/ on the videotape, with /v/ the expected

(McGurk) consonant. The other five trial types had /m/ on the sound track, paired with either /d/ or /g/ on the videotape, with /n/ the expected McGurk consonant. All the audio, video, and expected McGurk stimuli were real English words.

A female model was videotaped as she spoke each video and audio word at least once. One token of each of the nine auditory words was filtered at 10 kHz, digitized at a sampling rate of 20 kHz, and stored on a computer disk. Words beginning with nasals were edited to shorten the nasal murmur in order, after dubbing, to improve their apparent synchrony with the stop consonants on the video display. To dub the words onto a videotape, the video signal from the original tape, played on one VCR, was recorded onto a new videotape on a second VCR. At the same time, the audio signal from the original tape triggered a voice key interfaced with the computer. When the audio track of the original videotape triggered the voice key, the computer output a designated sound file to the second videotape, thereby dubbing it. The latency between voice key triggering and sound-file output was estimated to be 9 msec by Rosenblum and Fowler (1991). The nine dubbed trial types were recorded on video disk. The frame-by-frame control provided by this methodology permitted editing of each video word to yield a segment of 50 frames in length, with the initial frame occurring as closely as possible to the onset of articulation. Each video frame was displayed for 33.3 msec.

The response choices for each of the dubbed items consisted of two words—one corresponded to the auditory dub and the other corresponded to the expected integrated word (McGurk response). The words were printed vertically in capital letters, and each response pair was presented side by side. The right or left relative position of the printed word pairs was counterbalanced so that, whether the response pairs were presented to the left or right of the fixation point, each response word was sometimes closer and sometimes farther from the fixation mark. In this way, both response words should have been equally visible in each visual field over trials. As a result of the counterbalancing, there were two tokens of each of the 9 word pairs for a total of 18 word pair items. These 18 word pair items were printed out by computer and made into slides, which were then recorded onto video disk. Slides of left-pointing and right-pointing arrows were also made and recorded onto video disk.

A practice set of trials consisted of 18 trials, with each of the nine dubbed face tokens appearing twice (once in each visual field) and each of the 18 word pair slides appearing once. The test set consisted of 12 appearances of each of the nine dubbed face tokens. These 12 appearances consisted of 6 presentations of each dubbed face token in the right visual field and 6 presentations in the left visual field. The visual field of the face presentation was crossed with that of the response words, so that for half of the trials the face and words were presented to the same visual field, and for the other half they were presented to different visual fields. This set of 108 trials was divided into two test blocks of 54 trials each.

Table 1  
Word Pairs Presented Auditorily and Visually in  
Experiments 1 and 2 With the Expected McGurk Integration;  
Response Options Consisted of the Auditory Word  
and the Expected McGurk Response

Auditory	Visual	Expected McGurk Response
bat	vet	vat
bet	vat	vet
bent	vest	vent
boat	vow	vote
might	die	night
mail	deal	nail
mat	dead	gnat
moo	goo	new
met	gal	net

Each trial in the practice and test sets consisted of the following sequence: an arrow indicated direction of fixation, a dubbed audiovisual word was presented, a second arrow indicated direction of fixation, and then a printed word pair was presented. The arrows appeared for 1,833 msec, the face mouthing the word appeared for 1,667 msec,<sup>2</sup> and the word pairs appeared for 167 msec. The subject was given up to 2 sec to respond by using a joystick. If the subject did not respond in that interval, the next trial was initiated.

A television monitor was positioned 45 in. from the end of the table at which the subjects sat. At this viewing distance, the speaker's mouth subtended 5.5° of visual angle horizontally and 2.5° of visual angle vertically. Maximally, words subtended 3.5° vertically and .75° horizontally. Letters were never closer than 1.5° to the fixation point. Black paper was placed on the monitor so that only a 5.25 × 6.5 in. area of the monitor could be seen. All stimuli appeared in this rectangle. Fixation points were placed on the left and right sides of this rectangle, centered vertically. When the subjects looked at the left fixation point, the face or response words appeared in the right visual field; when they looked at the right fixation point, stimuli appeared in the left visual field. The fixation points were colored a bright yellow in order to maximize their visibility. The response device was a joystick that could be moved to the left or to the right. The subjects' joystick responses were recorded on a Macintosh computer interfaced with a MacPac.

**Procedure.** The subjects were tested individually in a room that was quiet, but not soundproof. They sat facing the television monitor. A chinrest was used in order to ensure that they maintained a 45-in. distance from the monitor. For each trial, they were instructed to fixate the fixation mark indicated by each arrow—that is, to look to the extreme right or the extreme left of the exposed screen. They were told that fixation might be difficult to maintain, but that it was important for the purposes of the experiment that they maintain fixation to the best of their ability. Their task was to indicate, using the joystick, which of the two printed words was closest to the word that they had heard the model say. Each subject used his/her dominant hand to manipulate the joystick. They were told to move the joystick to the left if the word they had heard the model say was the left word of the pair, and to the right if the word they had heard was the word on the right of the word pair. The subjects then completed the practice set of 18 items. At the conclusion of the practice set, they completed the two test blocks of 54 trials. Each subject was debriefed at the conclusion of the test session.

## Results

For each subject we computed the percentage of response choices that reflected an influence of the visibly mouthed word. We refer to these choices as the "McGurk" (as contrasted to the "auditory") responses.

**Right-handed subjects.** On average, the right-handed subjects selected the McGurk alternative on 79% of trials. This percentage is identical to that reported by Dekle et al. (1992), who used the same dubbed words, but without lateralizing them.<sup>3</sup>

Figure 1 (top) shows the percentages of McGurk responses among our right-handed subjects, plotted separately for the two visual fields of face and word presentation. As the figure shows, the right-handed subjects chose the McGurk response alternative more frequently when the face mouthing the stimulus words was presented to the left rather than to the right visual field, and more frequently when the response alternatives were presented to the right than to the left visual field. The magnitude of the LVF "advantage" for the face was 3%, and the

magnitude of the RVF advantage for the response words was 4.8%.

An analysis of variance (ANOVA), with factors, visual field of face, and visual field of response alternatives, revealed significance for both main effects [face,  $F(1,24) = 5.33$ ,  $p = .03$ ; response words,  $F(1,24) = 6.09$ ,  $p = .02$ ]. The interaction did not approach significance ( $F < 1$ ). Analyses using items rather than subjects as the random factor showed marginally significant main effects [face,  $F(1,8) = 3.52$ ,  $p = .097$ ; response words,  $F(1,8) = 4.56$ ,  $p = .065$ ]. With just nine items, these latter tests have little power. However, seven of the nine items showed an overall difference in favor of the left visual field for presentations of the face; two showed a reverse direction of difference. On response words, likewise, seven words showed an RVF advantage in selections of the McGurk response alternative; one item showed a reverse difference and one showed no numerical difference between the visual fields.

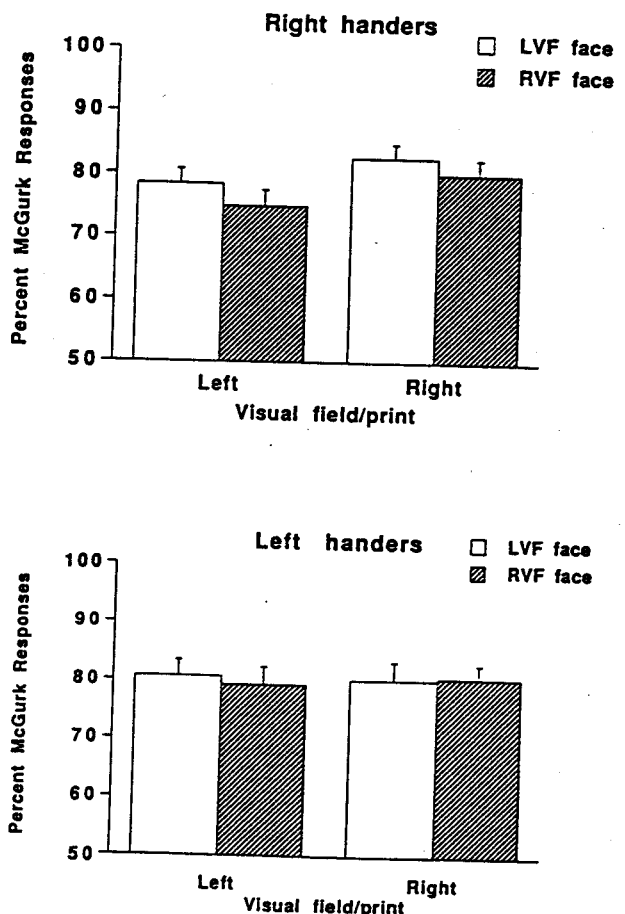


Figure 1. Mean percent selections of McGurk responses by right- (top) and left- (bottom) handed subjects with response words and face presented either to the left (LVF) or the right (RVF) visual field. Standard error bars are shown.

**Left-handed subjects.** Results from the 15 left-handers are presented in Figure 1 (bottom). On average, the left-handers chose the McGurk response alternative on 80% of the trials—a value very close to the findings with right-handers. However, unlike the right-handers, the left-handers showed no evidence of a difference in the tendency to choose the McGurk response alternative depending on the visual field in which either the face or printed words was presented. In ANOVAs such as those performed on the data from the right-handers, no effect approached significance in either the subjects or the items analyses.

### Discussion

Results of the first experiment are clear. The right-handed subjects were more likely to choose the McGurk response alternative when the face mouthing the word was presented so as to have privileged access to the right hemisphere and when the printed response words had privileged access to the left hemisphere. The latter finding is, perhaps, comparable to others that have shown an advantage in accuracy of identifying words presented to the left hemisphere, at least when confounding left-to-right scanning factors are eliminated, as they were here by vertical presentation of the words (see, e.g., Dimond, 1972; Levy, 1974, for reviews of the literature). The former finding is similar to that of Campbell (1986), who had normal right-handed subjects listen to words and look at still photographs of articulation. The left-handers, as a group, showed no lateralization effects. Bub and Lewine (1988) have argued that the less frequent occurrence of the normal pattern of lateralization in left-handers serves to attenuate the effects that rely upon it.

Our evidence favors a hypothesis that the factor of more importance in determining the hemispheric advantage for audiovisual integration of phonological information is that the information is conveyed by a face, whereas the factor of less importance is that the face provides phonological information. Of course this finding does not in itself challenge the motor theory of speech perception, which claims that extraction of phonetic or phonological information from either modality requires the special-to-speech processing of a speech module. The right-hemisphere advantage may accrue from its provision, to the left hemisphere, of better information on the visible actions of the vocal tract of the speaker than the left hemisphere itself could extract. It is possible that extraction of phonological information from the face still requires the left hemisphere. We attempted to test this idea by looking for a McGurk effect in a split-brain subject believed to have right-hemisphere visual and auditory language comprehension, but no control of motor speech. If the abstraction of phonological information requires the left-hemisphere motor speech system, this subject should not show a McGurk effect when visual phonological information is presented to the left visual field.

### EXPERIMENT 2A

With a callosotomy subject, two directions of significant effect might be expected in an experiment with the design of Experiment 1. If information about the face is available to and is usable by the hemisphere that receives the response words, McGurk selections should significantly exceed 50%. If the face information is unavailable to or not usable by the hemisphere receiving the response choices, McGurk selections should fall significantly below 50% (i.e., auditory selections should exceed chance). We expected to see below-chance McGurk selections when the face was presented to one visual field and the response words were presented to the other. When face and words were presented to the same visual field, the results would depend on whether or not the hemisphere receiving the acoustic and optical phonetic information could integrate it.

### Method

**Subjects.** The subject was a 39-year-old male, J.W., who underwent a complete callosotomy in 1979. For further information about this subject, see Gazzaniga, Nass, Reeves, and Roberts (1984).

**Materials.** The materials used were the same as in Experiment 1.

**Procedure.** The procedure used was the same as in Experiment 1, with some exceptions. Subject J.W. was tested on two occasions, with a 5-week interval between testing sessions. In order to control for any hemispheric effects in the motor response, he was asked to respond with both his right and left hands. During each testing session, J.W. completed one practice block and the two test blocks by responding with one hand, and then he completed the same two test sets by responding with the other hand. The order in which the test blocks were presented and the order of hand response were both counterbalanced across sessions. The subject was not debriefed at the conclusion of each test session, because he was tested on repeated occasions.

During each session, one of the experimenters stood behind the TV facing J.W. and attempted to guess, on the basis of J.W.'s direction of gaze, which fixation point the arrow had pointed to. The experimenter could not see the arrows on the TV screen. This was a difficult task, and some errors were almost certainly made simply because the experimenter could not see where J.W. was looking. Even so, on 85% of occasions, the experimenter's guesses matched the arrow direction; accordingly, we were confident that most of the time J.W. was fixating the appropriate position.

### Results and Discussion

Table 2 provides the results separately for each hand and session. Percentages greater than or equal to 59% and less than or equal to 41% differ significantly from 50%. The majority of cells in the table reveal chance responding. Two cells in which responding falls significantly below 50% are, as expected, cells in which the face and word were presented to different visual fields. However, in the remainder of such cells (cells off the main diagonals in the table), performance either does not differ from 50% or, in one case, significantly exceeds chance. This latter outcome should not be possible unless the subject failed to fixate as instructed.

Table 2  
Percentages of McGurk Response Selections by Subject J.W.  
in the Different Conditions of Experiment 2A

Visual Field of Response Words	Visual Field of Face			
	Left Hand		Right Hand	
	LVF	RVF	LVF	RVF
First Session				
LVF	61*	44	44	44
RVF	50	50	22*	58
Second Session				
LVF	58	39*	50	67*
RVF	50	64*	44	72*

\* $p < .05$ .

Of the on-diagonal cells, in which face and response words were presented to the same visual field, some cells do reveal significant McGurk effects. However, the pattern is not consistent. In the first session, significant McGurk selections occurred only when face and words were presented to the left visual field with left-hand responding. In the second session, responses in that cell did not exceed chance statistically, although they did so numerically. Instead, with either hand responding, only RVF presentations led to significant tendencies to select McGurk response words.

Overall, 10 of 16 cells show evidence of chance responding. Further, responses in 1 cell suggest a failure to fixate as instructed. Accordingly, we concluded that the task demands may have been too heavy for this subject, and we endeavored to make the task easier.

In the new task, neither faces nor words were lateralized. Therefore, we could only ask whether J.W. would show a McGurk effect under these conditions. If he did, we could not localize the integration of acoustic and optical phonetic information in either the left or right hemisphere. If he could do the task, but did not choose McGurk response words, we would know that both hemispheres (presumably the right for the facial information and the left for phonetic information) are required for McGurk integration. Using this new task, we tested J.W. and 4 control subjects—3 with a left-hemisphere epileptic focus and 1 with a right focus.

## EXPERIMENT 2B

### Method

**Subjects.** The subjects were J.W. (from Experiment 2A) and 4 female patients who were neurosurgery candidates due to intractable epilepsy. Three of the surgical candidates were right-handed and 1 was left-handed. All were left-hemisphere dominant for language, as indicated by a Wada examination that was part of the surgical workup (Wada & Rasmussen, 1960). Further psychometric data for these subjects appear in Table 3.

**Materials.** The dubbed face items from Experiments 1 and 2A were used in this experiment. As in those experiments, the response choices for each of the dubbed face items consisted of two words—one corresponded to the auditory dub and the other corresponded to the expected integrated word (McGurk response). In this experiment, however, the words were printed horizontally in capital letters, one above the other. The up-down position of the words in

the auditory McGurk word pairs was counterbalanced. As a result, again there were two tokens of each of the 9 word pairs for a total of 18 word pair items. As in Experiment 1, these 18 word pair items were printed out by computer and made into slides, which were then recorded onto video disk. The slides of left-pointing and right-pointing arrows were not included in this experiment because fixation points were not used.

As in Experiments 1 and 2A, the test set consisted of 12 appearances of each of the nine dubbed face tokens. This set of 108 trials was divided into two test blocks of 54 trials each. The practice set consisted of 18 trials, with each of the nine dubbed face tokens appearing twice and each of the 18 word pair slides appearing once.

Each trial in the practice and test sets consisted of the model saying a word, followed by a printed word pair. The face mouthing the word appeared for 1,667 msec, and the word pairs appeared for 500 msec. The subject was given up to 2 sec to respond via the joystick. If the subject did not respond in that interval, the program continued on to the next item.

As in the previous two experiments, the stimulus items on the video disk were presented on a color television monitor that was controlled by a Macintosh SE computer. Black paper was again placed on the monitor so that only a  $5.25 \times 6.5$  in. area of the monitor could be seen. All stimuli appeared in this square. The response device was a joystick that could be moved up or down. The subjects' joystick responses were recorded on the computer via a MacPac connection.

**Procedure.** The subjects were tested individually in a room that was quiet, but not soundproof. They sat facing the television monitor at a comfortable viewing distance. There were two experimental conditions, and 4 of the 5 subjects participated in both conditions. (We were unable to run the second condition with Subject C.K.) In one condition (view), the subjects were instructed to watch and listen to the model say the word and then indicate which word they had heard the model say. In the second condition (no view), the television screen was entirely covered by black paper, and the subjects were instructed to simply listen to the model say the word and then indicate which word they had heard the model say. After the procedure was explained, the subjects completed a practice set and were then tested under both conditions. One set of 54 items was presented in the view condition and the other in the no-view condition. The 4 subjects who were neurosurgery candidates all responded by using their dominant hands. Subject J.W. completed the practice set and then was tested under each condition twice—once responding with his right hand and once responding with his left.

### Results and Discussion

The results for J.W. and the 4 control subjects are presented in Table 4. All the subjects chose significantly more McGurk response options than auditory options in

Table 3  
Comparison of Age, IQ, and Education (in Years)  
of Epileptic Controls and Subject J.W.

Subject	Age	FSIQ*	VIQ*	PIQ*	Education
Controls					
M	35	94.4	94.2	97	12.2
SD	12.7	7.7	10.0	5.3	2.9
J.W.	39	97	95	96	12

\*These scores are from the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981), an individually administered test of intelligence for individuals from 16 to 74 years of age. The Full Scale Intelligence Quotient (FSIQ) is derived so that a score of 100 represents the mean with a standard deviation of 15. The Verbal Intelligence Quotient (VIQ) and Performance Intelligence Quotient (PIQ) combine to form the FSIQ and are similarly distributed.

the view condition and significantly more auditory options in the no-view condition. Chance performance is exceeded ( $p < .05$ , two-tailed, 54 trials per condition) when performance exceeds or falls below 50% by 13% or more. Subject J.W. showed a smaller separation between the conditions than did the control patients. However, his performance, like theirs, differed from chance in both conditions. Therefore, we can infer that cross talk between the hemispheres is not required for the McGurk effect to occur.

The performance of the 4 control subjects is remarkable for its consistency in the view condition. Their performance exceeded that of normal undergraduates tested under similar conditions (Dekle et al., 1992; see below). We are not sure what to make of this direction of difference between epileptic and normal subjects. We chose to run epileptic subjects because they appeared to provide the most adequate controls for J.W., due to their common history of long-standing seizure disorder. If J.W. had been unable to make McGurk responses, we had reasoned that it might be partially due to a deficit related to his long history of epilepsy. We did not expect that a neurological disorder would encourage the formation of phonological fusions. It may be relevant, however, that the subjects were instructed to select the word they heard; that is, they were urged not to let what they saw affect their responses. In this way of looking at McGurk selections, the resultant responses reflect a lesser ability to prevent integration of sight and sound. Although patients are more consistent than normal subjects in making response selections, it may also be said that they are less able to prevent McGurk integrations.

It may also be relevant that findings from several sources suggest that the poorer the quality of the acoustic in relation to the visual information (Massaro, 1987; Sekiyama & Tohkura, 1991), the more that McGurk integrations increase. For example, Sekiyama and Tohkura found that integrations increased when noise was added to the acoustic signal. Two of the 3 control patients in the present study, on whom we had acoustic measures, showed a rather high percentage of McGurk selections in the no-view condition. This may suggest that the acoustic syllables were less intelligible to them than they would be to the average normal listener (who, in comparable conditions of Dekle et al., 1992, made 3% McGurk selections in the no-view condition). In contrast to this account of the high percentage of McGurk selections in the view condition, however, is the right-hemisphere-damaged subject, E.E., whose accuracy in the no-view condition was almost perfect, yet who was perfectly consistent in choosing the McGurk selection in the view condition.

As for J.W.'s consistency, which was considerably less than that for the control subjects, several accounts can be offered. An interesting one is that audiovisual integration of phonological information normally involves integration of information between the hemispheres; J.W. cannot benefit from that integration. This account is consistent with our findings that right-handers showed larger

Table 4  
Percentages of McGurk Response Selections for Subject J.W.  
(Responding With the Left or Right Hand)  
and Control Subjects in Experiment 2B

Condition	Subjects					
	J.W. (Left)	J.W. (Right)	C.K.	P.M.	J.M.	E.E.
View	72	65	98	100	98	100
No view	33	19		14	27	02

Note—Subjects C.K., P.M., and J.M. each had a left-hemisphere epileptic focus; E.E.'s focus was in the right hemisphere.

McGurk effects with faces that were presented to the left visual field rather than to the right visual field.

A less interesting account, however, is that when J.W. fixated centrally on the face, each hemisphere may have received information about only one half of the face. It is possible that his reduced use of visual information reflects the fact that less information is available. To test this idea, we ran 24 undergraduates—12 viewed the whole face, unilateralized, and 12 viewed just one half of the face. Of the latter 12, 6 viewed the right half of the face, and 6 viewed the left half. The "half-faces" were created by simply blocking off one side of the face with construction paper. The subjects selected the McGurk response option 82.4%, 83.3%, and 83% of the time in the whole-face, right-side-only, and left-side-only conditions, respectively. Accordingly, we have no evidence that less phonological information is available from a half-face as opposed to a whole face.

Another less interesting account of J.W.'s performance, which we cannot rule out, is that he was less attentive than the other subjects in the experiment, that he found the task more taxing, or that he guessed on a greater percentage of the trials than the other subjects. This would account for his relatively high percentage of McGurk responses in the no-view condition and relatively low percentage in the view condition.

One additional issue needs to be addressed regarding J.W. Although we initiated these experiments with the assumption that J.W.'s right hemisphere lacked any ability to control motor speech, recent work has indicated that J.W. may be able to make vocal responses to LVF words and pictures (Wessinger, Fendrich, Baynes, & Gazzaniga, 1994). If the right hemisphere indeed controls motor speech, it is less damaging to the motor theory of speech that both hemispheres appear to be able to fuse auditory and optic phonological information.

## GENERAL DISCUSSION

Our findings on normal right-handed subjects paint the clearest picture we were able to obtain regarding laterality and the McGurk effect. Consistent with much of the work on visual word recognition in the right and left visual fields, we found an increase of approximately 5% in McGurk response selections when printed words were provided to the right as opposed to the left visual field.

This may not really indicate that the McGurk effect is stronger under those conditions, but that RVF presentation of the words makes the task of reading the words easier, and so performance accuracy (hearing the McGurk word, intending to choose the McGurk response option, and in fact doing so) increases. The more interesting finding was an increase in McGurk response selections when the face was presented to the left visual field as compared with the right visual field. This suggests that both hemispheres contribute to the McGurk effect.

A number of puzzles remain for further investigation. As for normal subjects, we do not yet know whether the LVF "advantage" in McGurk fusion responses implies that the right hemisphere provides better quality facial phonological information than the left, or whether it provides better quality facial movement information, with the linguistic significance of the movements still extracted by the left hemisphere. Our findings on the callosotomy subject, J.W., did not resolve that difference in interpretation. We have some evidence from Experiment 2A that either of the hemispheres can exhibit McGurk integration; however, we have one finding from that experiment that suggests that J.W. did not always maintain fixation.

Turning to the question of whether both hemispheres are required for McGurk integration to occur, Experiment 2B demonstrated that J.W. does integrate audiovisual phonetic information, although not as consistently or perhaps as unavoidably as do normal subjects. Therefore, we can rule out a strong hypothesis that the hemispheres *must* share information for the integration to occur. The fact that J.W. showed less integration than both the normal right-handed undergraduates and the epileptic control subjects may or may not signify that information-sharing increases the likelihood or the necessity of integration.

As for the epileptic control subjects, further research is required to test our speculative explanation for their extreme consistency in the integration of visible and audible phonetic information.

#### REFERENCES

- BAYNES, K. (1990). Language and reading in the right hemisphere: Highways or byways of the brain? *Journal of Cognitive Neuroscience*, 2, 159-179.
- BESNER, D., & HUMPHREYS, G. W. (Eds.) (1991). *Basic processes in reading: Visual word recognition*. Hillsdale, NJ: Erlbaum.
- BUB, D. N., & LEWINE, J. (1988). Different modes of word recognition in the left and right visual fields. *Brain & Language*, 33, 164-188.
- CAMPBELL, R. (1986). The lateralization of lipread sounds: A first look. *Brain & Cognition*, 5, 1-21.
- CAMPBELL, R., GARWOOD, J., FRANKLIN, S., HOWARD, D., LANDIS, T., & REGARD, M. (1990). Neuropsychological studies of auditory-visual fusion illusions: Four case studies and their implications. *Neuropsychologia*, 28, 787-802.
- CAMPBELL, R., LANDIS, T., & REGARD, M. (1986). Face recognition and lipreading: A neurological dissociation. *Brain*, 109, 509-521.
- CARAMAZZA, A. (1990). *Cognitive neuropsychology and neurolinguistics: Advances in models of cognitive function and impairment*. Hillsdale, NJ: Erlbaum.
- COLTHEART, M., SARTORI, G., & JOB, R. (1987). *The cognitive neuropsychology of language*. Hillsdale, NJ: Erlbaum.
- DAMASIO, A. R., & DAMASIO, H. (1986). The anatomical substrate of prosopagnosia. In R. Bruyer (Ed.), *The neuropsychology of face perception and facial expression* (pp. 31-38). Hillsdale, NJ: Erlbaum.
- DEKLE, D. J., FOWLER, C. A., & FUNNELL, M. G. (1992). Audiovisual integration in perception of real words. *Perception & Psychophysics*, 51, 355-362.
- DENES, G., SEMENZA, C., & BISIACCHI, P. (1988). *Perspectives on cognitive neuropsychology*. Hillsdale, NJ: Erlbaum.
- DIMOND, S. (1972). *The double brain*. Edinburgh: Churchill Livingstone.
- ELLIS, A. W. (1989). Neuro-cognitive processing of faces and voices. In A. W. Young & A. W. Ellis (Eds.), *Handbook of research on face processing* (pp. 207-215). Amsterdam: Elsevier.
- GAZZANIGA, M. S., NASS, R., REEVES, A. G., & ROBERTS, D. (1984). Neurologic perspectives on right hemisphere language following surgical section of the corpus callosum. *Seminars in Neurology*, 4, 126-135.
- GEFFEN, G., BRADSHAW, J. L., & NETTLETON, N. C. (1972). Hemispheric asymmetry: Verbal and spatial encoding of visual stimuli. *Journal of Experimental Psychology*, 95, 25-31.
- GRAVES, R., GOODGLASS, H., & LANDIS, T. (1982). Mouth asymmetry during spontaneous speech. *Neuropsychologia*, 20, 371-381.
- HILLGER, L. A., & KOENIG, O. (1991). Separable mechanisms in face processing: Evidence from hemispheric specialization. *Journal of Cognitive Neuroscience*, 3, 42-58.
- KLATZKY, R. L., & ATKINSON, R. C. (1971). Specialization of the cerebral hemispheres in scanning for information in short-term memory. *Perception & Psychophysics*, 10, 335-338.
- LEVY, J. (1974). Psychobiological implications of bilateral asymmetry. In S. Dimond & J. G. Beaumont (Eds.), *Hemisphere function in the human brain* (pp. 121-183). New York: Wiley.
- LIBERMAN, A. M. (1982). On finding that speech is special. *American Psychologist*, 37, 148-167.
- LIBERMAN, A. M., COOPER, F., SHANKWEILER, D., & STUDDERT-KENNEDY, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.
- LIBERMAN, A. M., & MATTINGLY, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1-36.
- MACDONALD, J., & MCGURK, H. (1978). Visual influences on speech perception processes. *Perception & Psychophysics*, 24, 253-257.
- MASSARO, D. (1987). Speech perception by ear and by eye: A paradigm for psychological inquiry. Hillsdale, NJ: Erlbaum.
- MCGURK, H., & MACDONALD, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746-748.
- PATTERSON, K. E., & BESNER, D. (1984). Is the right hemisphere literate? *Cognitive Neuropsychology*, 1, 315-342.
- PATTERSON, K. E., & BRADSHAW, J. L. (1975). Differential hemispheric mediation of nonverbal visual stimuli. *Journal of Experimental Psychology: Human Perception & Performance*, 1, 246-252.
- PATTERSON, K. E., & MARCEL, A. J. (1977). Aphasia, dyslexia, and the phonological encoding of written words. *Quarterly Journal of Experimental Psychology*, 29, 307-318.
- ROSENBLUM, L., & FOWLER, C. A. (1991). An audio-visual investigation of the loudness/effort effect for speech and nonspeech perception. *Journal of Experimental Psychology: Human Perception & Performance*, 17, 976-985.
- SEKIYAMA, K., & TOHKURA, U. (1991). McGurk effect in non-English listeners: Few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility. *Journal of the Acoustical Society of America*, 90, 1797-1805.
- SEMMES, J. (1968). Hemispheric specialization: A possible clue to mechanism. *Neuropsychologia*, 6, 295-315.
- SIDTIS, J. J., VOLPE, B. T., WILSON, D. H., RAYPORT, M., & GAZZANIGA, M. S. (1981). Variability in right hemisphere language function after callosal section: Evidence for a continuum of generative capacity. *Journal of Neuroscience*, 1, 323-331.
- WADA, J., & RASMUSSEN, T. (1960). Intracarotid injection of sodium amyltal for the lateralization of cerebral speech dominance: Experimental and clinical observations. *Journal of Neurosurgery*, 17, 266-282.
- WECHSLER, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised*. New York: The Psychological Corporation.
- WESSINGER, C. M., FENDRICH, R., BAYNES, K., & GAZZANIGA, M. S. (1994). *The emergence of the capacity to generate speech in a dis-*



*connected right hemisphere: Implication for functional plasticity.*  
Manuscript submitted for publication.

ZAIDEL, E., & PETERS, A. M. (1981). Phonological encoding and ideographic reading by the disconnected right hemisphere: Two case studies. *Brain & Language*, 14, 205-234.

#### NOTES

1. The words were printed vertically rather than horizontally because the word pairs generally differed only in the first letter. Printed horizontally, the first letter appears on the left of the word and thus would appear in the periphery of the LVF and near the center of the RVF, possibly yielding an RVF advantage when the word pairs were presented for a short duration in alternating visual fields.

2. The length of this presentation time was sufficient to allow the subjects to shift focus to the face during utterance of the word. However, the information distinguishing the two response options was always word-initial and occurred in a few tens of milliseconds before a shift in visual fixation could take place.

3. Another difference between the present experiment and that of Dekle et al. is that, in the earlier study, subjects were offered three response options, including the present ones and the word that the model had mouthed. That last option was rarely used by the subjects in the study by Dekle et al.

(Manuscript received July 14, 1993;  
revision accepted for publication November 8, 1993.)