Perception & Psychophysics
1995, 57 (8), 1124-1133

Facial identity and facial speech processing: Familiar faces and voices in the McGurk effect

STEPHANIE WALKER
University of Nottingham, Nottingham, England

VICKI BRUCE
University of Stirling, Stirling, Scotland

and

CLAIRE O'MALLEY
University of Nottingham, Nottingham, England

An experiment was conducted to investigate the claims made by Bruce and Young (1986) for the independence of facial identity and facial speech processing. A well-reported phenomenon in audio-visual speech perception—the *McGurk effect* (McGurk & MacDonald, 1976), in which synchronous but conflicting auditory and visual phonetic information is presented to subjects—was utilized as a dynamic facial speech processing task. An element of facial identity processing was introduced into this task by manipulating the faces used for the creation of the McGurk-effect stimuli such that (1) they were familiar to some subjects and unfamiliar to others, and (2) the faces and voices used were either congruent (from the same person) or incongruent (from different people). A comparison was made between the different subject groups in their susceptibility to the McGurk illusion, and the results show that when the faces and voices are incongruent, subjects who are familiar with the faces are less susceptible to McGurk effects than those who are unfamiliar with the faces. The results suggest that facial identity and facial speech processing are not entirely independent, and these findings are discussed in relation to Bruce and Young's (1986) functional model of face recognition.

Over the past forty years, there have been many laboratory studies demonstrating how the perception of speech is facilitated by the sight of the speaker. Sumby and Pollack (1954) reported that there was an advantage to seeing the face of the speaker if that speaker's auditory speech signal was accompanied by background noise, and since then, numerous studies have shown that the visual information obtained from a speaker's face is influential in determining what the listener perceives (Dodd, 1977; MacDonald & McGurk, 1978; Massaro & Cohen, 1983; McGurk & MacDonald, 1976; Reisberg, McLean, & Goldfield, 1987).

One of the most striking displays of the influence of vision on speech perception was provided by the McGurk and MacDonald (1976) demonstration. During this experiment, normal hearing subjects were asked to repeat the consonant–vowel (CV) syllables they heard while watching and listening to the videotaped face of a speaker. The videotape had been created such that the seen and heard speech syllables had conflicting consonants, but were nevertheless presented in synchrony. The subjects frequently reported hearing a "blend" or a "combination" of the seen and heard utterance. For example, a seen *ga* accompanied by a heard *ba* was reported by subjects as being heard as a *da* or *tha*, and a seen *ba* accompanied by a heard *ga* was frequently reported as *bga*.

Since this initial demonstration of the *McGurk effect*, further studies have supported and extended the phenomenon and have sought to offer explanations for its existence (e.g., see Fowler & Dekle, 1991; MacDonald & McGurk, 1978; Massaro & Cohen, 1983; Summerfield, 1979). The effect appears to be extremely robust under a variety of conditions: It is not lessened by the perceiver having prior knowledge of the illusion, nor is it decreased when the perceiver has had considerable practice at selectively attending (McGurk & MacDonald, 1976); further, the effect remains when subjects are specifically requested to report *only* what they have heard (Summerfield & McGrath, 1984). While the majority of previous studies investigating the effect have used CV stimuli with an /a/ vowel context (e.g., *ba*, *ga*), Green, Kuhl, and Meltzoff (1988) found that an /i/ vowel context produced the strongest McGurk effect, with the /a/ context giving a moderate effect and a /u/ context having almost no effect. The McGurk effect is a convincing demonstration of the potency of visual information, even when the auditory signal is clear and unambiguous, and as such, many researchers working in the field of speech perception have endeavored to use the ev-

This research was supported in part by Grant R000233560 from the Economic and Social Research Council. The authors wish to thank Quentin Summerfield from the Institute for Hearing Research, Nottingham, for helpful advice during the early stages of this study, and Geoff Yarnall from the Audio/Visual Unit at Nottingham Medical School, for technical assistance with construction of the stimuli. Address correspondence to S. Walker, Department of Psychology, University of Nottingham, Nottingham NG7 2RD, England (e-mail: sww@psyc.nott.ac.uk).

idence from such demonstrations in support of their theoretical accounts of audio-visual speech perception (e.g., see Massaro, 1987; Summerfield, 1987).

In one recent paper describing a study involving the McGurk effect (Green, Kuhl, Meltzoff, & Stevens, 1991), emphasis was placed more on the particular circumstances that may affect the integration of auditory and visual information. Typically in experiments using the McGurk effect, the auditory and visual speech signals are provided by the same speaker, but Green et al. conducted an experiment in which they examined the effect of having two different speakers providing the visual and auditory components of the utterance, thereby giving rise to a *cognitive incongruency* between the auditory and visual signals. They achieved this by having a discrepancy in the gender between the seen face and heard voice of the speaker; thus, a male speaker's voice was dubbed onto a videotape containing a female speaker's face, and vice versa. These gender-incongruent videotapes were then compared with gender-congruent tapes, in which the face and voice were of the same gender. The results showed that there was no significant difference in the magnitude of the reported McGurk effect between congruent and incongruent videotapes, even though on a later stimulus-compatibility rating using new subjects, it was indicated that the gender discrepancy between the faces and voices of the speakers in the cross-gender stimuli was quite apparent. On the strength of their findings, Green et al. claimed that a reduction in cognitive congruency had no apparent influence on the strength of the McGurk effect, and this led them to conclude that during auditory-visual speech perception, "auditory and visual phonetic information is integrated even when the two inputs could not have been derived from a single, biological source" (Green et al., p. 534).

From the evidence reviewed so far, it appears that the McGurk effect serves as a compelling piece of evidence for the role of the face in audio-visual communication, and up to now, the findings from research using the effect have in the main been centered around theories of speech perception (see Massaro, 1987; Summerfield, 1987). One question rarely addressed however, concerns what, if anything, the McGurk illusion can tell us about how we process the information available to us in the faces themselves. Before this issue can be addressed, some attention must be given to the current theoretical models of face processing.

During the last 15 years, researchers working in the area of face processing have given serious consideration to the formulation of theoretical frameworks for face recognition, and a number of broadly comparable functional models have been put forward (e.g., Bruce, 1979, 1983; Bruce & Young, 1986; Ellis, 1981, 1983, 1986; Hay & Young, 1982). These models all describe, to a greater or lesser extent, the stages of information processing involved in face recognition, and the Bruce and Young (1986) functional model of face recognition was an attempt to pull together and expand upon earlier models.

One of the major claims made by Bruce and Young (1986) is that different types of information are extracted in parallel from the faces we see. Bruce and Young argue from their model that three of the major aspects of face perception—recognition of facial identity, recognition of facial expression, and recognition of facial speech—are all independently achieved. Thus it is proposed that, for example, one does not have to recognize a person's identity in order to be able to speech-read their lips. At a purely intuitive level, it might appear that information extracted from a face that allowed it to be classified as *familiar* would not be the same as that which needs to be extracted for making use of facial speech cues. Evidence in support of the kind of functional organization proposed by Bruce and Young comes primarily from two sources: (1) experimental manipulations carried out on normal subjects, during which one kind of performance is affected, while there is little or no effect on another; and (2) reports of brain-damaged patients, in which different patterns of selective impairment of different face-processing tasks are reported.

The independence of facial expression and facial identity processing has been supported by converging evidence from experiments with normal adults (e.g., Bruce, 1986; Young, McWeeny, Hay, & Ellis, 1986), neuropsychological dissociations (e.g., Bowers, Bauer, Coslett, & Heilman, 1985; Etcoff, 1984; Young, Newcombe, De Haan, Small, & Hay, 1993), and neurophysiological recordings (Hasselmo, Rolls, & Bayliss, 1989). There have up to now, however, been few empirical studies to support the claim made for the independence of facial speech and facial identity processing. De Gelder, Vroomen, and van der Heide (1991) reported an experiment that investigated the relationship between face recognition and lipreading skills in autistic children. The autistic subjects, individually matched for mental age with a control group of normal children, were tested on memory for unfamiliar faces and on lipreading ability. The findings from this study did provide evidence for the independence of memory for faces and facial speech—which is in accordance with the Bruce and Young (1986) model—but this independence was observed only in the autistic subjects, and was not seen in the normal controls. However, there were some methodological problems with this study, and the findings need careful interpretation.

Perhaps the strongest piece of evidence for independent processing routes of facial identity and facial speech comes from the neuropsychological case studies reported by Campbell, Landis, and Regard (1986). This report describes a dissociation between facial speech and face recognition as observed in two patients. One patient was a prosopagnosic woman who had great difficulties with the recognition of faces, but who nevertheless appeared to have no problem with judgment of mouthed phonemes when they were presented in face photographs, and she was also susceptible to the McGurk and MacDonald (1976) illusion. The second patient was an alexic woman who had no difficulty with face recognition, but who performed badly when asked to make phonetic judgments to face photographs and was not susceptible to the McGurk and MacDonald illusion. This double dissociation between facial speech and facial identity processing suggests that these two functions are indeed independent, each calling on a different cortical processing mechanism.

Findings from both the de Gelder et al. (1991) and the Campbell et al. (1986) studies offer some support for the notion of independent facial identity and facial speech processing, but neither has totally addressed the question of the parallel processing of these two functions. Evidence from the Green et al. (1991) experiment, during which video recordings of speaking male and female faces were used, suggests that facial speech is independent from facial-*gender* processing, but for this study, the question of facial *identity* processing was not addressed. During episodes of face-to-face spoken communication, listeners are able to acquire simultaneous information concerning the gender of the person speaking (*Is this a male or a female face?*), the identity of the person speaking (*Is this a familiar or an unfamiliar face?*), the facial expression worn by the speaker (*What does the emotional content of the face convey?*), the particular lip movements made during the act of speaking (*What is the information carried by facial speech cues?*), and, of course, the acoustic information given by the speaker (*What are the voice characteristics of the speaker and what is the actual meaning and intent conveyed by the utterance?*).

In view of the findings from these previous studies, and their possible limitations, the present study, which combined the facial identity and facial speech processing tasks, was conducted to investigate the independent-routes theory for these two functions. In keeping with previous research, use of facial speech cues was investigated by measuring susceptibility to the McGurk effect. However, in order to investigate the effects of facial speech and facial identity processing simultaneously, the faces used to test facial speech were manipulated such that they were either familiar or unfamiliar to subjects. Therefore, whereas the findings from the Green et al. (1991) experiment showed that the McGurk effect was not affected by the mismatching of the face and voice *gender*, the present study was set up to investigate whether this also holds true in a situation in which the face and voice *identities* mismatch. Face–voice combinations of presented familiar or unfamiliar faces were thus made up such that they were from the same person, a different person of the same sex, or a different person of the opposite sex. A comparison between subjects familiar with and subjects unfamiliar with the faces could then be made with respect to their susceptibility to the McGurk effect. If, as is predicted by the Bruce and Young (1986) functional model of face processing, facial speech is entirely independent from facial identity processing, and also from gender processing (see Green et al., 1991), there should be no difference between the different groups of subjects in their degree of susceptibility to the McGurk effect.

METHOD

Subjects

A total of 36 subjects were recruited for this experiment. Eighteen subjects were students and staff from within the psychology department of the University of Nottingham. These 18 subjects, comprising 10 females (1 final-year undergraduate, 6 postgraduates, and 3 secretarial staff) and 8 males (4 final-year undergraduates and 4 postgraduates) were all very familiar with the target faces/voices (i.e., those with which they were to be presented). For the purposes of this study, *familiarity* was defined in terms of subjects having had face-to-face interaction with the target faces/voices on several, if not many, occasions, and was not merely based on interaction through a lecturer/student association during lectures alone.

The remaining 18 subjects were not familiar with the faces they were to see. These 18 subjects were undergraduates and postgraduates from other departments within the University of Nottingham, and they comprised 9 females and 9 males. *Nonfamiliarity* was defined in terms of subjects never having seen or heard any of the target faces/voices.

The subjects were within an age range of 19–37 years. All were native English speakers, had normal or corrected-to-normal vision, and had no reported speech or hearing disabilities. They were paid £2 each for their participation in this experiment.

Materials and Apparatus

Recording syllables. A color camera (Sony M7), a microphone (Sony ECM55S), and a video recorder (Panasonic MII AU65) were used to record each of four speakers (two male and two female members of teaching staff in the psychology department at the University of Nottingham) while they produced several instances of each of the four syllables /ba/, /ga/, /bi/, and /gi/. Each speaker, in turn, was seated at a table in front of a plain yellow background, with the microphone positioned on the table, at a distance of 30–40 cm directly in front of him or her. The camera was focused on the speaker's face. The video lighting was set up to British standard broadcast specification: one spotlight was positioned at an angle of 45° to the left of the speaker's nose-line, a second spotlight (the hair-light, used to ensure that the face of the speaker stood out from the background) was placed at the back of the speaker's head, and a fill-light (a softer light used to remove any shadows cast by the spotlights) was positioned at an angle of 40° to the right of the speaker's nose-line. The setup of the equipment ensured an exceptionally clear view of the speaker's face and mouth, with the whole face, including the hair, being entirely visible. The recording session resulted in the production of 20–24 recordings of each of the four syllables /ba/, /ga/, /bi/, and /gi/ from each of the four speakers.

Editing and dubbing of the recorded stimuli were carried out using two video recorders (MII AU63 and MII AU65) and an edit controller (Panasonic AGA 800). For each of the four speakers, a single instance of each of the two syllables /ba/ and /ga/ was selected such that (1) overall duration of lip movement and (2) length of auditory signal were similar (to within 20 msec) within and across all four speakers. Instances were selected such that there were no extraneous lip movements, and such that there was a 1-sec video clip before and after the lip movement, during which there was a neutral facial expression and no lip movement. This was then repeated for the /bi/ and /gi/ syllables. This procedure gave a total of 16 CV utterances (four syllables from each of four speakers), which were subsequently used for the production of the McGurk stimuli.

Creating stimuli. Equipment previously described for editing and dubbing was used for the creation of stimulus materials. Table 1 shows how auditory stimuli from the four speakers were combined with the visual stimuli for one of the four speakers (Speaker 1—a female speaker), using the /ba/ and /ga/ CV syllables.

In this particular case, the selected auditory stimulus (from each of the four speakers), was dubbed onto the selected visual stimulus (from Speaker 1), such that there was coincidence of the release of the consonant for the auditory-visual signals, to within an accuracy of 20 msec. This resulted in four different face/voice pairings for each of the four combinations of the CV syllables. There were therefore three different congruency types: congruent (C), in which Speaker 1 visual stimuli were paired with Speaker 1 auditory stimuli; incongruent same gender (IC/SG), in which Speaker 1 visual stimuli were paired with Speaker 2 auditory stimuli; and incongruent cross gender (IC/XG), in which Speaker 1 visual stimuli were paired with Speaker 3 (and Speaker 4) auditory stimuli. Hence there was a total

Table 1
Combinations of /ba/–/ga/ Syllables for Face of Speaker 1 Paired With Voices of Speakers 1, 2, 3, and 4

Visual Stimuli	Auditory Stimuli			
Speaker 1 (Female)	Speaker 1 (Female)	Speaker 2 (Female)	Speaker 3 (Male)	Speaker 4 (Male)
ba	ba	ba	ba	ba
ga	ga	ga	ga	ga
ba	ga	ga	ga	ga
ga	ba	ba	ba	ba

of 16 novel auditory-visual stimuli made up of the visual stimuli from Speaker 1 paired with auditory stimuli from all four speakers. This procedure was then repeated using the visual stimuli of Speakers 2, 3, and 4, giving a total of 64 novel auditory-visual stimuli (16 from each of the four speakers). The whole procedure was then repeated using the /bi/ and /gi/ syllables.

These newly created auditory-visual combinations were then edited onto a new MII format videotape which became the master tape. This master tape was then copied onto VHS video format before copying, blocking, and presentation of stimuli took place.

Blocking stimuli. For the /ba/–/ga/ pairings, in any one block, only one congruency type (either C, IC/SG, or IC/XG) and only one face gender (either male faces or female faces) were used. For the female C block (i.e., Speaker 1 visual stimuli combined with Speaker 1 auditory stimuli, and Speaker 2 visual stimuli combined with Speaker 2 auditory stimuli), there were five repetitions of each of the four different /ba/–/ga/ pairings for each of the two female speakers, giving a total of 40 trials—20 from the first speaker and 20 from the second speaker. Each of the 20 trials in these two sub-blocks was randomly placed onto a new tape. Each 40-trial block was immediately preceded by four practice trials. The same blocking procedure was used for the female IC/SG block (Speaker 1 visual stimuli combined with Speaker 2 auditory stimuli, and Speaker 2 visual stimuli combined with Speaker 1 auditory stimuli), and for the female IC/XG block (e.g., Speaker 1 visual stimuli combined with Speaker 3 auditory stimuli, Speaker 2 visual stimuli combined with Speaker 4 auditory stimuli, and so forth). For the female faces, this gave a total of 3 blocks of stimuli (one for each congruency type), with each block containing 40 trials—giving a total of 120 trials plus practice trials. The procedure was repeated for the male faces where, again, 3 blocks of stimuli (one for each of the three congruency types) were created.

Exactly the same blocking procedure was adopted for the /bi/–/gi/ combinations, resulting in a grand total of 12 blocks of trials—6 with female faces (3 = /ba/–/ga/ and 3 = /bi/–/gi/), and 6 with male faces (3 = /ba/–/ga/ and 3 = /bi/–/gi/).

A single trial consisted of a single auditory-visual stimulus, preceded and followed by a 1-sec clip of the speaker's face with a neutral expression and no lip movement. Between trials was a 5-sec gray screen.

Presenting stimuli. Stimuli were presented on the 22-in. screen of a Sony Trinitron color television set (Model KV-202OUB) using a JVC video recorder (Model HR-265OEK). The sizes of the facial images when presented on the TV screen were approximately 29 cm long × 27 cm wide (for both female faces, including hair), and 29 cm long × 24 cm wide (for both male faces, including hair).

Design

This was a 2 × 2 × 3 mixed design, with two between-group factors and one within-group factor.

The first between-group factor was the familiarity of faces presented to subjects (familiar or unfamiliar). The second between-group factor was the sex of the seen faces—the subjects were presented with either the two female faces or the two male faces. This factor was selected as pilot testing had revealed that the experiment took well over an hour for both male and female faces to be shown, and this was considered to be too tedious for subjects to endure in a single session. Half of the subjects in each of the two familiarity groups were randomly allocated to the male faces condition, and half were allocated to the female faces condition.

The within-group factor was that of face/voice congruency. Faces presented to subjects were either congruent with the voice (C), incongruent but of the same gender (IC/SG), or incongruent and of the opposite gender (IC/XG). Order of presentation of the three congruency types was counterbalanced across subjects.

All subjects were presented with syllables in both /a/ and /i/ vowel contexts, and were counterbalanced according to the order of presentation of these two contexts.

Each subject was thus presented with a total of 6 blocks of test stimuli (both vowel contexts, in each of the three congruency types). Each block was made up of 40 trials, giving a grand total of 240 test trials.

There was a break of 30 sec between blocks of stimuli within each of the two vowel contexts, and a break of 2 min between presentation of the two different vowel contexts.

Procedure

The subjects were tested individually, and were seated approximately 1.5 m from the TV screen, which was housed within a quiet room. They were instructed that they were to watch and listen to a video recording of two different speakers, appearing separately and saying various syllables (which began with *d*, *b*, *g*, *bg*, or *th*), and that their task, on each occasion, was to repeat aloud what they heard the speaker say. It was emphasized to subjects that they had to watch the TV screen at all times, and that they had to respond as quickly and as accurately as possible. Responses were recorded by the experimenter, who was sitting in the same room as the subjects.

RESULTS

Tables 2, 3, and 4 show the mean percentages of responses given by the subjects in each of the response categories. Table 2 shows the breakdown of responses recorded for familiar and unfamiliar seen faces when the faces and voices were congruent (C); Table 3 shows the recorded responses for familiar and unfamiliar faces when the faces and voices were incongruent but of the same gender (IC/SG); Table 4 shows the recorded responses for familiar and unfamiliar faces when the faces and voices were incongruent and cross gender (IC/XG). In these tables, *blends* are defined as either *d*- or *th*- responses as reported by subjects in the context of an auditory /b/ combined with a visual /g/, and these responses are referred to in the ensuing text as *correct expected-blend responses*. Similarly, *combinations* are defined as *bg*- responses as reported by subjects in the context of an auditory /g/ combined with a visual /b/, and these responses are referred to in the text as *correct expected-combination responses*.

When the faces and voices were congruent (C), the subjects who were familiar with the faces reported 52% correct expected-blend (*d*- or *th*-) responses and 19% correct expected-combination (*bg*-) responses; the subjects who were unfamiliar with the faces reported 44% correct expected-blend and 38% correct expected-combination responses. For the incongruent, same-gender faces/voices (IC/SG), the subjects who were familiar with the faces reported 27% correct expected-blend and 17% correct expected-combination responses, whereas the subjects who were unfamiliar with the faces reported 43% correct expected-blend and 40% correct expected-combination responses.

Table 2
Mean Percentage Responses for Male and Female Faces With Congruent (C) Face and Voice

Visual Stimuli	Auditory Stimuli	Response Categories					
		b-	*d-*	*th-*	*g-*	*bg-*	Other
		Familiar Faces					
b-	*b-*	100	0	0	0	0	0
g-	*g-*	0	0	0	97.5	2.5	0
g-	*b-*	40	47*	5*	7.5	0.5	0
b-	*g-*	17	1	0	62	19†	1
		Unfamiliar Faces					
b-	*b-*	100	0	0	0	0	0
g-	*g-*	0	0	0	99.5	0.5	0
g-	*b-*	48	31*	13*	8	0	0
b-	*g-*	26	0	0	36	38†	0

*Blend responses. †Combination responses.

Table 3
Mean Percentage Responses for Male and Female Faces With Incongruent Same-Gender (IC/SG) Face and Voice

Visual Stimuli	Auditory Stimuli	Response Categories					
		b-	*d-*	*th-*	*g-*	*bg-*	Other
		Familiar Faces					
b-	*b-*	99.5	0	0	0	0.5	0
g-	*g-*	0	0	0	98.5	1.5	0
g-	*b-*	69	24*	3*	4	0	0
b-	*g-*	14	0.5	0	68.5	17†	0
		Unfamiliar Faces					
b-	*b-*	100	0	0	0	0	0
g-	*g-*	0	0	0	97.5	2.5	0
g-	*b-*	49	26*	17*	7	0	1
b-	*g-*	24	0	0	36	40†	0

*Blend responses. †Combination responses.

Table 4
Mean Percentage Responses for Male and Female Faces With Incongruent Cross-Gender (IC/XG) Face and Voice

Visual Stimuli	Auditory Stimuli	Response Categories					
		b-	*d-*	*th-*	*g-*	*bg-*	Other
		Familiar Faces					
b-	*b-*	99	1	0	0	0	0
g-	*g-*	0	1.5	0.5	95	3	0
g-	*b-*	76	21*	1*	2	0	0
b-	*g-*	16	2	0	68.5	12.5†	1
		Unfamiliar Faces					
b-	*b-*	100	0	0	0	0	0
g-	*g-*	0	0	0	98	2	0
g-	*b-*	48	28*	16*	8	0	0
b-	*g-*	27.5	0.5	0	35	37†	0

*Blend responses. †Combination responses.

For the incongruent, cross-gender faces/voices (IC/XG), the subjects who were familiar with the faces reported 22% correct expected-blend and 12.5% correct expected-combination responses, and the subjects who were unfamiliar with the faces reported 44% correct expected-blend and 37% correct expected-combination responses.

A 2 × 2 × 3 (familiarity of seen face × sex of seen face × congruency type) analysis of variance (ANOVA) was conducted on percentage scores of correct expected-combination (*bg-*) responses. There was a significant main effect of familiarity [$F(1,32) = 21.707$, $p < .001$]. There was no main effect of sex of seen face and no main effect of congruency, and there were no significant interactions. The subjects who were familiar with the seen faces gave significantly fewer correct expected-combination responses than did the subjects who were not familiar with the seen faces.

A 2 × 2 × 3 (familiarity of seen face × sex of seen face × congruency type) ANOVA was conducted on percentage scores of correct expected-blend (*d-* or *th-*) responses. There was a significant main effect of congruency type [$F(2,64) = 19.129$, $p < .001$], but no other main effects reached significance. However, there was a significant two-way interaction between familiarity and congruency [$F(2,64) = 21.249$, $p < .001$]. There were no other significant interactions. Figure 1 shows the interaction between familiarity and congruency.

Simple main effects were conducted on familiarity at levels of congruency. There was no significant simple main effect of familiarity at the C level of congruency. There was a significant simple main effect of familiarity both at the IC/SG level of congruency [$F(1,96) = 4.588$, $p < .05$] and at the IC/XG level of congruency [$F(1,96) = 7.033$, $p < .01$]. When the faces and voices of the seen faces were congruent (i.e., the face and voice were from the same person), there were no differences in the numbers of correct expected blends reported between the subjects who were familiar with the faces and the subjects who were unfamiliar with them. However, when the faces and voices were incongruent (either IC/SG or IC/XG), the subjects who were familiar with the faces reported significantly fewer correct expected blends than did the subjects who were unfamiliar with the faces.

Simple main effects were also conducted on congruency type at the two levels of familiarity. A significant simple main effect of congruency type for familiar faces was found [$F(2,64) = 40.428$, $p < .001$]; but there were no significant simple main effects of congruency type for unfamiliar faces. Tukey tests were conducted on means for levels of congruency of familiar faces. The difference between C and IC/SG levels of congruency was significant [$q(3,64) = 9.94$, $p < .01$], and there was a significant difference between C and IC/XG levels [$q(3,64) = 11.84$, $p < .01$]. There was no significant difference between IC/SG and IC/XG congruency types. When the subjects were unfamiliar with the seen faces, there were no differences in numbers of correct expected blends reported, regardless of whether the faces/voices were congruent (C) or incongruent (IC/SG or IC/XG). However, for subjects familiar with the seen faces, there were significantly fewer blends reported when the face and voice were incongruent (whether same gender or cross gender) than there were when the faces and voices were congruent (C).

The findings so far suggest that when subjects are asked to respond to auditory-visual stimuli in which the auditory and visual channels carry conflicting information, there are some differences between subjects who are familiar with the seen faces and those who are unfamiliar with them.

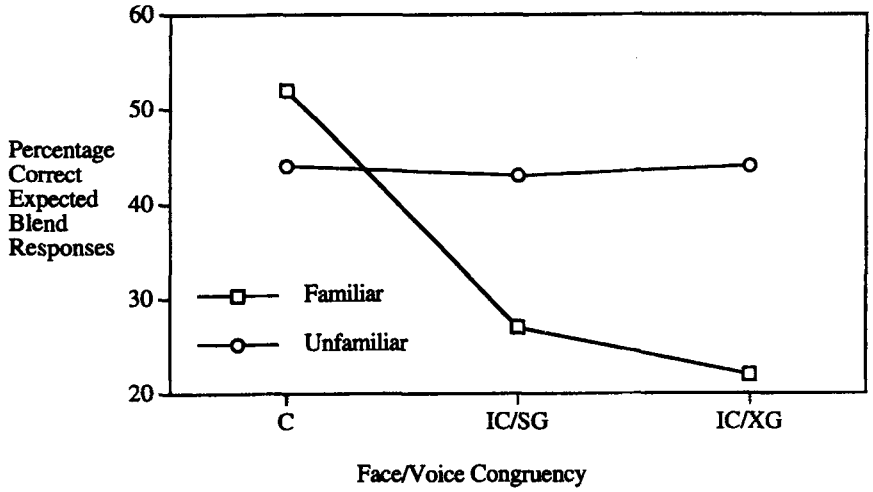

Figure 1. Expected correct-blend responses (familiarity vs. face/voice congruency).

The responses so far considered have been the *correct expected responses* given by subjects when presented with conflicting auditory and visual information, but in cases in which the expected blend or combination response was not given, subjects either responded with the auditory stimulus or with the visual stimulus. For example, subjects reporting a *ba* when presented with a face in which the visual stimulus is /ba/ and the auditory stimulus is /ga/ would be reporting the visual-channel response. The mean percentages of the visual-channel responses given by subjects when the auditory and visual channels conflicted are also given in Tables 2, 3, and 4, for C, IC/SG, and IC/XG faces/voices, respectively. To investigate differences in visual information used by subjects (e.g., how often subjects reported a *ba* when the visual stimulus was /ba/ but the auditory stimulus was /ga/), the visual-channel responses reported when the auditory and visual channels carried conflicting information were analyzed.

A 2 × 2 × 3 (familiarity of seen face × sex of seen face × congruency type) ANOVA was conducted on percentage scores of *b-* responses (i.e., the visual channel) given by subjects when a *bg-* (i.e., a combination) response was expected. There was no main effect of sex of seen face and no main effect of congruency, but there was a significant main effect of familiarity [$F(1,32) = 8.424, p < .01$]. No significant interactions were found. Subjects who were unfamiliar with the seen faces reported the visual-channel response to a greater extent than did subjects who were familiar with the seen faces when presented with a visual-channel stimulus of /b-/ and an auditory-channel stimulus of /g-/.

A 2 × 2 × 3 (familiarity of seen face × sex of seen face × congruency type) ANOVA was also conducted on percentage scores of *g-* responses (i.e., the visual channel) given by subjects when a *d-* or a *th-* response (i.e., a blend response) was expected.

There were no significant main effects of either familiarity or sex of seen face. There was a main effect of congruency type [$F(2,64) = 3.708, p < .05$], but a significant two-way interaction between familiarity and congruency was present [$F(2,64) = 4.724, p < .05$]. There were no other significant interactions. Figure 2 shows the interaction between familiarity and congruency.

Simple main effects were conducted on familiarity at the three levels of congruency type. A significant simple main effect of familiarity at the IC/XG level of congruency was found [$F(1,96) = 9.903, p < .01$]. There were no significant simple main effects of familiarity at C or IC/SG levels of congruency. When the subjects were presented with faces and voices that were incongruent and cross gender (IC/XG), those who were familiar with the seen faces reported the visual-channel response to a lesser extent than the subjects who were unfamiliar with them.

Simple main effects were also conducted on congruency type on the two levels of familiarity. There were no simple main effects of congruency type for unfamiliar faces—for subjects who were unfamiliar with the seen faces, there were no significant differences in numbers of visual-channel responses between C, IC/SG, and IC/XG faces/voices. However, a significant simple main effect of congruency type was found for familiar faces [$F(2,64) = 7.721, p < .01$]. Tukey tests were conducted on the means of congruency types for familiar faces, and the only significant difference found was between C and IC/XG faces/voices [$q(3,64) = 5.52, p < .01$]. Subjects who were familiar with the faces reported significantly fewer visual-channel responses when the face and voice were IC/XG than when they were C. There were also fewer visual-channel responses reported for IC/SG faces/voices than for C faces/voices (C = 7.5%; IC/SG = 4%), although this difference did not reach statistical significance.

DISCUSSION

The purpose of the present study was to investigate the independence of facial identity and facial speech processing when these two functions are carried out simultaneously, in a dynamic face-to-face situation. To address this issue, a well-reported phenomenon in audio-visual speech

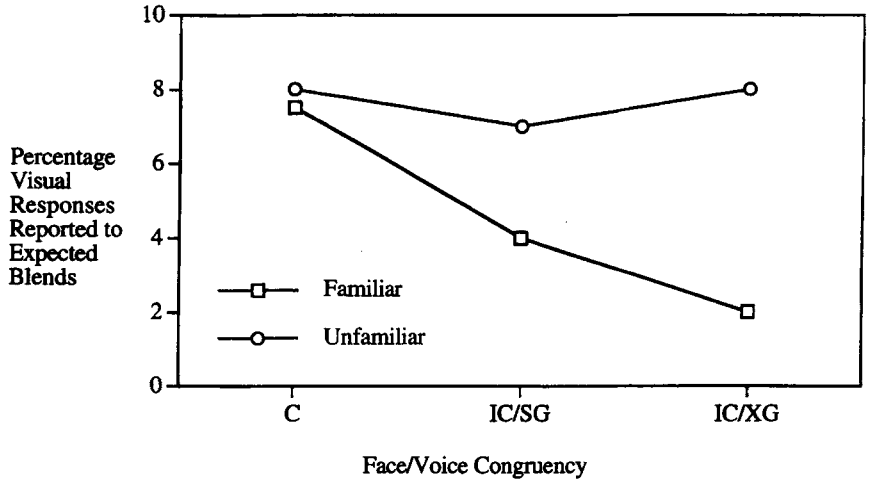

Figure 2. Visual responses reported to expected blends (familiarity vs. face/voice congruency).

perception, the McGurk effect, was used, and a comparison was made between subjects who were familiar with the presented faces and those who were unfamiliar with them, in terms of their susceptibility to the effect. The results of this experiment show that the subjects who were familiar with the faces reported fewer correct expected-combination (*bg*-) responses than did the subjects who were unfamiliar with the seen faces, regardless of whether the faces/voices of the seen faces were congruent or incongruent. The results also show that the subjects who were familiar with the faces reported fewer correct expected-blend (*d*- or *th*-) responses than did the subjects who were unfamiliar with the faces, except when the face and voice of the seen face were congruent (i.e., from the same person), in which case, there was no difference in numbers of correct expected blends reported between familiar and unfamiliar seen faces. In the following discussion, we consider the relationship between these effects and earlier findings, possible explanations of the effects, and the implications for models of person and speech recognition.

Comparison With Previous Findings

Taken overall, these findings suggest that facial identity and facial speech processing are not completely independent of one another, and to this extent, the study does not support the previous research in this area (Campbell et al., 1986; de Gelder et al., 1991) nor does it entirely support the independent-routes theory for facial speech and facial identity, as proposed by Bruce and Young (1986) in their functional model of face processing.

Our data from subjects who were unfamiliar with the seen faces, where there were no differences in responses given between congruent faces/voices and incongruent faces/voices, are in agreement with the findings from Green et al.'s (1991) experiment, and offer support for their claim that a reduction in cognitive congruency (i.e., pairing a male face with a female voice and vice versa) does not lead to a decrease in susceptibility to the McGurk effect. However, for subjects who were familiar with the seen faces, a different pattern of responses emerged, and these differences will now be considered.

The data from expected-combination responses given by subjects who were familiar with the seen faces showed no differences between numbers of correct expected responses, regardless of the face/voice congruency (i.e., same numbers of *bg*- responses are given for congruent faces/voices and incongruent faces/voices), and this pattern is again in agreement with the findings of Green et al. (1991). One question that must be addressed concerns why, when the faces and voices were congruent (i.e., from the same person), subjects who were familiar with the faces reported fewer expected-combination responses than subjects who were unfamiliar with the faces, when they reported just as many correct expected-blend responses as the subjects who were unfamiliar in the same congruent face/voice condition. One explanation for this could be that over previous encounters with the speaker, the listener has been provided with a series of expectations of what speech events are likely, and of how these are realized through facial speech cues. Perhaps a deviation from what is expected, as may be the case with a syllable that begins with *bg*, is therefore more likely to lead to one of the communication channels being ignored in favor of the other. Indeed, previous work by Welch and Warren (1980) does suggest that historical factors, in terms of specific experience that observers have had with the observed event, might have a part to play in the processing of intersensory discrepancies, and this issue will be discussed in more detail later.

Data from expected-blend responses given to familiar faces show a significant decrease in numbers of blends reported between congruent faces/voices and incongruent faces/voices, irrespective of whether the incongruence is due to a same-gender pairing (a familiar female face paired with a different familiar female voice) or to a cross-gender

pairing (a familiar female face paired with a familiar male voice), and it is here that our findings are at odds with those of the Green et al. (1991) study.

The differences between findings reported from the present study and those of Green et al. (1991), are apparent when the issue of face-identity processing is addressed, but before turning our attention back to the face-processing literature, one aspect of these data not yet considered is that of voice familiarity, and it is this issue that we now wish to address.

Voice Familiarity or Face Familiarity?

Throughout this paper so far, we have referred to familiarity with respect to *faces*. However, during the course of the study itself, face familiarity covaried with voice familiarity, and so it is possible that our familiarity effects were due in fact not to *face* familiarity but instead to *voice* familiarity. Before addressing the issue of voice familiarity effects per se, one further aspect of our data is pertinent to this issue and therefore merits some discussion at this point—namely, why did our familiar subjects, when reporting on incongruent face/voice trials, tend to report the auditory channel rather than the visual channel when blend or combination responses were expected? (In other words, why was there a shift toward reporting the voice?) To account for these data, three points are worth mentioning. First, in more general terms, our data fits in with those of Green et al. (1991) in that when subjects do not report blends or combinations, they are more likely to report what they *hear* rather than what they *see*. Second, the study reported by de Gelder et al. (1991) showed that although autistic subjects performed at levels equivalent to those of normal controls when judging facial speech on isolated faces, they were not at all susceptible to McGurk effects—instead, they reported what they heard. This finding is consistent with Frith's (1989) account of an autistic cognitive style of *weak central coherence*, whereby autistics are better at part-based tasks than at integrated ones. If, as is suggested by this account, a lack of integration leads autistic subjects to use evidence from a voice rather than a face, then perhaps due to a similar lack of integration (from different face and voice cues), normal subjects adopt the same strategy. This issue of nonintegration for our familiar subjects will be discussed in more detail later. The final point is that, according to research reported by Nygaard, Sommers, and Pisoni (1994), familiarity with a voice aids the perception of speech in noise. If this is the case, perhaps there is less need to supplement familiar (hence clearer) speech with information from the face. With these points in mind, it would appear that before we can go on to discuss our familiarity effect in terms of *face* familiarity, we must first establish whether *voice* familiarity alone is able to account for our findings.

Some evidence from our study, albeit informal, that a voice-familiarity effect did not account for our findings came to light during subject debriefing, when several subjects who were familiar with the target faces declared that although they had always recognized the *faces* during the course of the experiment, when presented with incongruent faces/voices, they had not been able to identify the *voice*, even though they were aware that "the voice did not belong to the face." Thus, at an informal level, it appears that it was not familiar voices that led to our familiarity effect. To more formally assess subjects' abilities to categorize voices as familiar or unfamiliar when presented with short auditory tokens alone, 32 new auditory tokens (made up of equal numbers of each of the syllables /ba/, /ga/, /bi/, and /gi/) were recorded from eight different speakers, four familiar and four unfamiliar.[1] These auditory tokens were presented, in random order, to 10 new subjects, who were asked (1) to categorize each voice presented as familiar or unfamiliar, and (2) for any voice categorized as familiar, to try to identify the voice. For categorization of voices as familiar or unfamiliar, subjects' performance was not statistically above chance level (mean percentage of correct categorizations was 56%). For identification of familiar voices, subjects' performance was very poor, with a mean percentage of correct voice identification of 11%.

These findings are certainly in keeping with the comments made by subjects during our main study—that is, they suggest that identification of a familiar voice from short auditory tokens is extremely difficult. Indeed, from the data reported here, it appears that even categorization of voices as familiar or unfamiliar is at levels that are no better than chance when reports are based on short verbal utterances. At an anecdotal level, many of us may be able to recall experiences of answering a telephone and not being able to immediately recognize the voice at the other end of the line—even though that person may be very familiar to us. Indeed, it has previously been reported that learning to identify voices when hearing only isolated words is an exceptionally difficult task (Williams, 1964). The findings from our post hoc study suggest that the familiarity effect reported in our McGurk-effect experiment is not due to voice familiarity—subjects were quite poor at voice-familiarity categorizations when making decisions based on short utterances.

Overall, our findings lead us to suggest that (1) familiarity with a person has an effect on how facial speech cues are analyzed, and (2) this familiarity effect is not due to subjects' differential processing of familiar and unfamiliar voices. We therefore propose that the familiarity effects obtained are due to the differential processing of familiar and unfamiliar *faces*, and it is with this in mind that we now turn our attention back to the face-processing literature.

Implications for Models of Person Identification and Audio-Visual Speech Processing

The Bruce and Young (1986) functional model of face processing predicts that knowing the identity of a face will have no effect on the processing of any facial speech cues given by that face, and this is clearly at odds with the findings of the current study. Although our data show that subjects who were familiar with the faces reported as many blends as subjects who were unfamiliar with the faces when the face and voice were from the same person, significantly fewer blends were reported for the familiar faces when the face and voice did not match; also, irrespective

of the congruency of face and voice, fewer combination responses were reported by subjects who were familiar with the seen faces. It would appear from our results that when subjects were processing facial speech cues from familiar faces, they were able to use their knowledge concerning those particular faces. One question that must now be addressed concerns whether the Bruce and Young model, as it stands, can incorporate these data.

Of course, one of the problems with functional models of face perception, including that of Bruce and Young (1986), is that they are static models, whereas face perception, and indeed faces themselves, are not. What the current models of face perception actually offer is an account of the early processing stages (i.e., the initial categorization and recognition of faces). According to Bruce and Young, this initial identification of an individual from his or her face takes place via distinct sequential stages. First, each familiar face has its own *face recognition unit* (FRU), and this becomes active when any recognizable view of a familiar face is seen. FRUs respond only to faces; they are at the stage at which perceptual classification is realized, and once activated, an FRU activates the appropriate *person identity unit* (PIN). There is one PIN for each known person, and activation of a particular PIN allows access to semantic information concerning a particular individual (e.g., occupation, address, etc.). Whereas the FRUs can only be activated by a person's face, PINs may be accessed via the face (from the FRU) or via the person's voice, from a written or heard name, or even from an individual's gait or clothing. During the reported study, a familiar face would lead to activation of the appropriate FRU and hence the appropriate PIN, but information from the auditory signal would only lead to activation of the same PIN if the face and voice were congruent. In cases in which the face and voice were incongruent, there would be conflict between the PIN activated via the face of the known individual and the PIN activated via the voice of a different individual—hence there would be two different PINs activated from the same single event, resulting in a *disunity* signal.

The problem with this account of face/voice incongruency is that although it provides an explanation of the data from our McGurk-effect study within the framework of the Bruce and Young (1986) model, it does not entirely fit it with our findings that subjects' performance was no better than chance on voice-familiarity categorizations—if a voice is not recognized as familiar, how could a PIN be activated by it? To answer this question, we must perhaps consider the framework from a broader perspective. One explanation might be that although an incongruent familiar voice activates a second, rival PIN, this particular activation is at subthreshold level (i.e., the rival PIN is active, but the level of activation is smaller than would be required for overt recognition). Under this account, subjects not *overtly* recognizing the voice as familiar may be responding *covertly* to what is still in effect a signal of disunity (i.e., there are still two different PINs active from the same event). In fact, a mechanism of this kind has previously been described as an account of covert face recognition in a prosopagnosic patient (Burton, Young, Bruce, Johnston, & Ellis, 1991). A second explanation might be that during presentation of an incongruent face/voice, there will be *reduced* activation at the appropriate PIN (i.e., there will be activation of the PIN from the face but not from the voice) in comparison with presentation of a congruent face/voice, when there will be activation of the appropriate PIN from both face and voice. In order to tease apart the various aspects of these alternative explanations, further investigation is warranted, and this is currently under way.

For now, it may be useful to consider the observed data with recourse to normal discourse when, for example, there may be several people speaking at the same time (i.e., the "cocktail party" scenario). In this situation, it may well be ambiguous as to which voice goes with which face. Although temporal cues (acoustic signals emanating from lips moving in synchrony) and spatial (location) cues would obviously be helpful in this situation, if the people speaking are unfamiliar, perceived gender of voice may not be a good cue—some males do have high-pitched voices and some females have low-pitched voices. If, however, the people speaking are familiar, *knowing* that a particular voice belongs to a particular identifiable individual might be quite useful for sorting out the various speakers.

The importance of a *unity assumption* during the processing of discrepant signals has been considered by Welch and Warren (1980) in terms of a more general model for the processing of intersensory discrepancies. According to Welch and Warren, it is the unity of the two discrepant signals that is fundamental to how they may or may not be integrated. They claim that when a perceiver receives information from more than one modality (in the case reported here, there are two modalities involved—auditory and visual channels), an assumption is made by the perceiver as to whether there is a single or a multiple event taking place. Welch and Warren refer to this as "the assumption of unity," and suggest that if there is a "strong unity assumption," the perceptual outcome is of a single physical event, whereas if there is a "weak unity assumption," the outcome will be treated as more than one event. Green et al. (1991) suggest that their findings indicate that this need not be the case, as in their experiment, the information from two sources (auditory and visual) was integrated, even though subjects reported low unity of the auditory and visual stimuli. The current findings, however, offer some support for the unity theory—at least for subjects who were familiar with a seen face, and who, when presented with an incongruency between face and voice, were less susceptible to effects of one modality on the other.

Welch and Warren (1980) also suggest that historical factors, in the form of specific experience observers have had with the event being perceived, may play a part in any intersensory bias. In terms of the data from the current study, subjects who are familiar with the seen face have had more experience of the audio-visual stimuli when face and voice are congruent (i.e., when the face seen matches the voice heard) than when there is an incongruency between these two modalities, and therefore they are more likely to observe the congruent faces/voices as a single event and respond accordingly, integrating information

from the two different sources. However, when the familiar faces and voices are incongruent (i.e., when the face seen does not match the voice heard), subjects do not consider the two sources to be the same event, and therefore they are less likely to integrate the two modalities and so report less blend and combination responses. For subjects who are unfamiliar with the seen face, there are no historical factors to influence their responses. These subjects are more likely, therefore, to treat the two sources as one single event and respond accordingly, integrating the information from both modalities.

In conclusion, the results of the present study, which used a more dynamic and simultaneous facial identity and facial speech processing task than had previously been used, do not entirely support the notion of independence of the processing of facial speech and the processing of facial identity, as is claimed by Bruce and Young (1986) in their functional model of face processing; neither are the current findings in agreement with previous research using the more traditional separate static stimuli to assess face-identity processing. We suggest that more research to address the issues surrounding early and late processing using similar dynamic processing tasks is needed, and with current advances in interactive video techniques that permit these types of tasks to be more easily constructed, this will no doubt be forthcoming.

REFERENCES

BOWERS, D., BAUER, R. M., COSLETT, H. B., & HEILMAN, K. M. (1985). Processing faces by patients with unilateral hemisphere lesions: 1. Dissociation between judgements of facial affect and facial identity. *Brain & Cognition*, **4**, 258-272.

BRUCE, V. (1979). Searching for politicians: An information-processing approach to face recognition. *Quarterly Journal of Experimental Psychology*, **31**, 373-395.

BRUCE, V. (1983). Recognising faces. *Philosophical Transactions of the Royal Society of London: Series B*, **302**, 423-436.

BRUCE, V. (1986). Influences of familiarity on the processing of faces. *Perception*, **15**, 387-397.

BRUCE, V., & YOUNG, A. (1986). Understanding face recognition. *British Journal of Psychology*, **77**, 305-327.

BURTON, A. M., YOUNG, A. W., BRUCE, V., JOHNSTON, R. A., & ELLIS, A. W. (1991). Understanding covert recognition. *Cognition*, **39**, 129-166.

CAMPBELL, R., LANDIS, T., & REGARD, M. (1986). Face recognition and lipreading. *Brain*, **109**, 509-521.

DE GELDER, B., VROOMEN, J., & VAN DER HEIDE, L. (1991). Face recognition and lip reading in autism. *European Journal of Cognitive Psychology*, **3**, 69-86.

DODD, B. (1977). The role of vision in the perception of speech. *Perception*, **6**, 31-40.

ELLIS, H. D. (1981). Theoretical aspects of face recognition. In G. M. Davies, H. D. Ellis, & J. W. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 171-197). London: Academic Press.

ELLIS, H. D. (1983). The role of the right hemisphere in face perception. In A. W. Young (Ed.), *Functions of the right cerebral hemisphere* (pp. 33-64). London: Academic Press.

ELLIS, H. D. (1986). Processes underlying face recognition. In R. Bruyer (Ed.), *The neuropsychology of face perception and facial expression* (pp. 1-27). Hillsdale, NJ: Erlbaum.

ETCOFF, N. L. (1984). Selective attention to facial identity and facial emotion. *Neuropsychologia*, **22**, 281-295.

FOWLER, C. A., & DEKLE, D. J. (1991). Listening with eye and hand: Cross-modal contributions to speech perception. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 816-828.

FRITH, U. (1989). *Autism: Explaining the enigma*. Oxford, U.K.: Basil Blackwell.

GREEN, K. P., KUHL, P. K., & MELTZOFF, A. N. (1988). Factors affecting the integration of auditory and visual information in speech: The effect of vowel environment. *Journal of the Acoustical Society of America*, **84**, S155.

GREEN, K. P., KUHL, P. K., MELTZOFF, A. N., & STEVENS, E. B. (1991). Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics*, **50**, 524-536.

HASSELMO, M. E., ROLLS, E. T., & BAYLISS, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research*, **32**, 203-218.

HAY, D. C., & YOUNG, A. W. (1982). The human face. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 173-202). London: Academic Press.

MACDONALD, J., & MCGURK, H. (1978). Visual influences on speech perception processes. *Perception & Psychophysics*, **24**, 253-257.

MASSARO, D. W. (1987). Speech perception by ear and eye. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 53-83). London: Erlbaum.

MASSARO, D. W., & COHEN, M. M. (1983). Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 753-771.

MCGURK, H., & MACDONALD, J. (1976). Hearing lips and seeing voices. *Nature*, **264**, 746-748.

NYGAARD, L. C., SOMMERS, M. S., & PISONI, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, **5**, 42-45.

REISBERG, D., MCLEAN, J., & GOLDFIELD, A. (1987). Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 97-113). London: Erlbaum.

SUMBY, W. H., & POLLACK, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, **26**, 212-215.

SUMMERFIELD, Q. (1979). Use of visual information for phonetic perception. *Phonetica*, **36**, 314-331.

SUMMERFIELD, Q. (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. In B. Dodd & R. Campbell (Eds.), *Hearing by eye: The psychology of lip-reading* (pp. 3-51). London: Erlbaum.

SUMMERFIELD, Q., & MCGRATH, M. (1984). Detection and resolution of audio-visual incompatibility in the perception of vowels. *Quarterly Journal of Experimental Psychology*, **36A**, 51-74.

WELCH, R. B., & WARREN, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, **88**, 638-667.

WILLIAMS, C. E. (1964). *The effects of selected factors on the aural identification of speakers* (Report No. EDS-TDR-65-153, Section III). Hanscom Field, MA: Air Force Systems Command, Electronic Systems Division.

YOUNG, A. W., MCWEENY, K. H., HAY, D. C., & ELLIS, A. W. (1986). Matching familiar and unfamiliar faces on identity and expression. *Psychological Research*, **48**, 63-68.

YOUNG, A. W., NEWCOMBE, F., DE HAAN, E. H. F., SMALL, M., & HAY, D. C. (1993). Face perception after brain injury. Selective impairments affecting identity and expression. *Brain*, **116**, 941-959.

NOTE

1. Although auditory tokens alone were presented to subjects, in order to maintain quality and consistency with those used during the main study, all recordings were made using video equipment similiar to that previously described.

(Manuscript received May 26, 1994;
revision accepted for publication March 8, 1995.)