THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 1998, 51A (1), 179–195

I Recognise You but I Can't Place You: An Investigation of Familiar-only Experiences during Tests of Voice and Face Recognition

J. Richard Hanley, S. Tanya Smith, and Jenny Hadfield
Department of Psychology, University of Liverpool, Liverpool, U.K.

In this paper, we examine in detail the situation in which a subject finds that a face or voice is familiar but is unable to retrieve any biographical information about the person concerned. In two experiments, subjects were asked to identify a set of 40 celebrities either from hearing their voice or from seeing their face. Although many more celebrities were identified and named in response to their face than their voice, the results showed that there was a very large number of occasions when a celebrity's *voice* was felt to be familiar but the subject was unable to retrieve any biographical information about the person. This situation occurred less frequently in response to seeing a celebrity's face; when a face was found familiar, the subject was much more likely to be able to recall the celebrity's occupation. The possibility that these results might have come about because subjects were using different criteria to determine familiarity in the face and voice conditions was investigated and discounted. An additional finding was that when subjects found a face to be familiar-only, they were able to recall significantly more additional information about the person when they were cued by the person's voice than when they simply saw the face again. These results are discussed in relation to the models of person recognition put forward by Bruce and Young (1986) and Burton, Bruce, and Johnston (1990).

There are a number of distinct ways in which an attempt to recall the name of a familiar face can break down (Brennen, Baguley, Bright, & Bruce, 1990: Hanley & Cowell, 1988; Hay, Young, & Ellis, 1991; Young, Hay, & Ellis, 1985). On some occasions, subjects are able to recall detailed biographical information about the person concerned but cannot recall their name. Under these circumstances, subjects are frequently in a tip-of-the-tongue state and can sometimes recall information about the physical structure of the elusive name. On other occasions, subjects are unable to "place" the person—that is, they feel that a face is familiar, but they are unable to recall any specific information about the person concerned. These two types of failure can be seen as reflecting a breakdown at

Requests for reprints should be sent to J. Richard Hanley, Department of Psychology, University of Liverpool, Eleanor Rathbone Building, Bedford Street South, Liverpool L69 7ZA, U.K. E-mail: jw24@liv.ac.uk

The authors would like to thank Andy Young, Mike Burton, Andy Ellis, and Stefan Schweinberger for their helpful comments on earlier versions of this manuscript.

different stages in the models of face processing put forward by Bruce and Young (1986) and Burton, Bruce, and Johnston (1990).

According to Bruce and Young's (1986) sequential stage model, a subject who finds a face familiar but is unable to place the person has successfully activated the Face Recognition Unit (FRU) that represents the face of the person concerned but is unable to activate the Person Identity Nodes (PINs) that contain biographical information about the person. A subject who is in a tip-of-the-tongue state has successfully activated the PIN and the semantic information that it contains but cannot access the name code that is stored within the lexical output system at the next level down in the hierarchy. Hanley and Cowell (1988) provided support for such a conceptualization by demonstrating that different types of retrieval cue enabled subjects to recall the correct name of the person when they were in these two different states (see also Brennen et al., 1990). When subjects were able to recall detailed biographical information about the person concerned, then the initials of the person proved to be the most effective cue in helping them to recall the name. When subjects were in a familiar-only state, then the initials proved to be a less effective cue for name retrieval than biographical information about the person concerned. These results provide strong support for the Bruce and Young (1986) model because they demonstrate that the probability that a retrieval cue can elicit a previously unrecallable name is predictable on the basis of the stage in the model at which the original breakdown occurred. Results consistent with those obtained by Hanley and Cowell have recently been reported by Schweinberger, Herholtz, and Sommer (1997) in a study that investigated the effects of cues on subjects' ability to recall the names of famous voices.

In recent years, a great deal of research has investigated Bruce and Young's (1986) claim that names are stored at separate locations within the face processing system from other pieces of information that we know about a person. Much of this (e.g. Bruce, Burton, & Walker, 1994; Hanley, 1995; Harris & Kay, 1995; Stanhope & Cohen, 1993) has come about following the publication of a paper by Burton and Bruce (1992), arguing that names are difficult to recall because they tend to be unique to the person concerned. Little or no research, however, has investigated different theoretical accounts of what might be occurring when we find a face familiar but are unable to recall any further information about the person concerned.

This is an interesting issue, because Burton et al. (1990) made some changes to the role that is played by the PINs when they attempted to implement the Bruce and Young (1986) framework in the form of an "interactive activation and competition" (IAC) computational network (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). These changes are relevant to what might be occurring within the face processing system when a subject finds a face "familiar-only". The first of these changes concerns the origin of the feeling that a face, voice, or name is familiar. Burton et al. (1990) proposed that it is activation of the appropriate PIN above threshold that signals this familiarity. According to Bruce and Young (1986), it was activation at the level of the FRU that indicated familiarity of a face, activation of a voice recognition unit (VRU) that indicated that a voice was familiar, and activation of a name recognition unit (NRU) that indicated that a name was familiar. In the IAC model, the FRUs, NRUs, and

VRUs are nodes that will pass on activation to PINs as soon as they start to become excited. However, they do not themselves signal familiarity.

The second change made by Burton et al. (1990) was that the PINs do not themselves *contain* any semantic information about people. Instead, biographical information is stored in a separate pool of units, which can only be *accessed via* the PINs. The PINs, therefore, become nodes by which semantic information is retrieved rather than locations where semantic information is stored. There are at least two reasons for believing that this modification offers an advantage over Bruce and Young (1986). First, in contrast to the clear role that they play in the IAC model, the way in which it might be possible for PINs both to access and to store semantic information was not made explicit by Bruce and Young (1986). Second, the view that PINs do not contain any semantic information is consistent with the pattern of performance shown by patient ME (de Haan, Young, & Newcombe, 1991), who performs within the normal range when asked to judge whether faces, voices, or names are familiar but is severely impaired when asked to recall biographical information about these people. The key piece of evidence about ME is that she can match the faces and names of familiar people even though she cannot recall semantic information about them (de Haan et al., p. 61). Such a pattern of performance is not always observed in patients who make a large number of familiar-only responses (c.f. Warrington & McCarthy, 1988). Nevertheless, it is difficult to explain in terms of Bruce and Young's (1986) account. Because the matching task requires processing at a point at which the face and name recognition systems have converged, the name and face recognition units could not in themselves form the basis for her success. If ME was using the PIN to make the matches, then according to Bruce and Young (1986) she should have also been able to recall semantic information about the people.

The IAC model is able to explain ME's performance without any difficulty, however. As Burton, Young, Bruce, Johnston, and Ellis (1991, p. 160) point out, ME's ability to do the matching task can be explained by assuming that she can access the PINs; her inability to recall information about people can be explained by assuming that she is suffering from a block between the PINs and the semantic information pool. This account represents a quite different explanation from that proposed by Bruce and Young (1986) of the situation in which a subject feels that a face is familiar but is unable to recall anything about the person concerned. In Burton et al.'s (1991) account of ME, the problem occurs because the PIN reaches threshold but fails to activate biographical detail about the person in the semantic information pool. Bruce and Young, however, argued that familiar-only experiences occur when an FRU reaches threshold without activating the PIN.

The purpose of the present paper is to investigate under laboratory conditions the precise nature of the familiar-only responses that normal subjects experience when attempting to recognize familiar people. Is it the case that they generally reflect a blockage between the PIN and person-specific semantic information, as appears to be the case with ME (Burton et al., 1991)? This would not follow from Bruce and Young (1986), as their model does not make a clear separation between PINs and person-specific semantic information. Alternatively, is there any evidence that familiar-only responses can be caused by a blockage between the FRU and the PIN? Because it is a cascade model, evidence of this kind would not be inconsistent with Burton et al.'s (1990) framework. If the associative links from the FRU to the PIN were relatively weak, this might lead to a

situation in which there was enough activation of the PIN for it to signal familiarity, but insufficient activation at the level of the PIN for nodes in the semantic information pool to become strongly activated themselves. A blockage between the FRU and the PIN would therefore produce a familiar-only experience in the IAC model also.

In the experiments reported below, normal subjects were asked to recognize celebrities either from their face or from their voice. Recent research has suggested that there are important parallels between the ways in which familiar faces and voices are processed (Ellis, Jones, & Mosdell, 1997). On the basis of research that has compared episodic memory for unfamiliar faces and voices (e.g. Yarmey, Yarmey, & Yarmey, 1994), however, we expected that the group of subjects who were given faces would perform better overall than the group given voices. The critical question, though, was as to the precise way that recognition performance might differ in the two groups. One possibility is that subjects in the voice group would simply find fewer of the celebrities to be familiar. This would come about if the VRU for a person is seen as being more difficult to activate, or as requiring more activation to reach threshold, than the FRU.

A more interesting possibility is that there will also be more "familiar-only" responses in the voice condition than in the face condition. Such an outcome could not be explained in terms of weaker links between the PIN and the semantic information pool in the voice processing system than in the face processing system. This is because the link between the PIN and the semantic information pool occurs beyond the point in the system at which face and voice recognition have converged. If, therefore, there are weakened connections between the PIN and the semantic pool, then this should affect the processing of voices and faces equally; there would be no reason to expect that familiar-only responses would be more prevalent in response to hearing voices than in response to seeing faces. In contrast, such a pattern of results could easily be accommodated if familiar-only experiences occur when there are relatively weak associations between the FRU or VRU and the PIN. In the case of voices, it is quite conceivable that there will be a relatively large number of situations in which a VRU becomes activated without there being a strong enough association from the VRU to the PIN for semantic information about the person to be retrieved.

In addition to comparing the overall number of familiar-only responses in the two conditions, we also investigated the effects of cueing the subject with a presentation of the celebrity in the other modality, when the subject was in this state. The critical question was whether subjects who find a voice familiar-only will be more likely to retrieve semantic information about the person if they are subsequently presented with the person's face compared with the control condition of simply hearing the voice saying exactly the same things a second time. Similarly, will subjects who find a face familiar-only be helped by hearing the person's voice relative to seeing the same views of their face for a second time?

If a familiar-only experience reflects reduced activation between the FRU or VRU and the PIN, then this should be possible. If a voice has failed to activate a PIN strongly because of relatively weak links between the VRU and PIN, then a face cue may activate more strongly the PIN via the separate FRU-PIN pathway. Similarly, if a face has failed to activate a PIN sufficiently strongly, then hearing the person's voice could in principle activate the PIN via the pathway from the VRU to the PIN. In contrast, if the problem in a familiar-only situation always occurs because of a block

between the PIN and the semantic information pool (i.e. the PIN has reached the required level of activation, but semantic information cannot be recalled), there is no reason why presenting the person in a new modality should prove beneficial. Presenting the person in a new modality would maintain the activation level of the PIN but would not provide an alternative means of accessing the semantic information pool.

In the experiments reported below, therefore, we presented subjects with video recordings of interviews with celebrities. In Phase 1, half the subjects saw the face with no sound, and half the subjects heard the voice without seeing the face. When subjects failed to identify any celebrities fully, they were given a second opportunity to recall information about them in Phase 2. In Phase 2, half the subjects always saw or heard the celebrity in the same modality as previously, and half of the subjects both saw the celebrity's face and heard their voice. The two critical questions were whether there would be an equivalent number of familiar-only responses in the face and voice conditions in Phase 1 and whether the presentation of the celebrity in a new modality in Phase 2 would help subjects when they had found the celebrity to be familiar-only in Phase 1. When subjects reported that they found a face or a voice to be familiar, we also asked them to indicate the degree of familiarity (low, medium, high) that they felt towards that face or voice in order to ascertain whether any additional familiar-only responses in the voice condition might be associated with relatively low levels of subjective familiarity.

EXPERIMENT 1

Method

Subjects

Sixty subjects who were undergraduate students at Liverpool University were used as subjects. Their ages ranged from 18 to 35 years. Half the subjects were randomly allocated to the voices group and half to the faces group. Half of the subjects in the voices group saw the faces of the celebrities in Phase 2, and half of them heard the voices again in Phase 2. Half of the subjects in the faces group heard the voices of the celebrities in Phase 2, and half of them saw the faces again in Phase 2.

Stimuli

Short extracts from television interviews with 40 celebrities were used as the stimuli in this experiment. The 40 celebrities comprised approximately equal numbers of politicians, sportsmen/women, actors/actresses, TV personalities, comedians/comediennes, and pop stars. Each extract lasted 10–15 sec and included the celebrity saying something neutral that gave no clues as to identity or occupation. In addition, there were 20 similar extracts from television interviews with people who are not famous (20 rather than 40 unfamiliar people were chosen to keep the duration of the experiment as short as possible). The order of the 60 extracts on the videotape was randomized.

Procedure

Subjects in the voices group heard the 60 voices one at a time through the television's speaker but were seated to the side of the TV screen, so that they could not see the person's face. Subjects in the faces group were exposed to the same extracts but were seated in a position such that they could see

the face of the person on the television screen. However, the sound was turned off, so that they could not hear the voice.

Subjects in both groups were asked first of all to indicate whether they found the face or voice unfamiliar or familiar. If they found the face or voice familiar, they were asked to rate the degree of familiarity on a scale of 1–3, where 3 represented high familiarity, 2 represented medium familiarity, and 1 represented low familiarity. They were then asked to recall the occupation and the name of the celebrity if they could.

A subject who failed to find a face or voice familiar or to recall full details about a celebrity was presented with the extract of the celebrity once again. All extracts of celebrities were recorded twice consecutively on the tape, so that they could easily be presented for a second time if they were not fully identified on the first presentation. We will refer to the second presentation as Phase 2. The tape was stopped between Phase 1 and Phase 2 while the subject made a response. The second presentation of a celebrity took place approximately 15 sec after the original presentation had finished. For half of the subjects in each group, Phase 2 was always simply a rerun of Phase 1. The other half of the subjects in both groups saw the face and heard the voice of the celebrity in Phase 2. Only celebrities were presented in Phase 2. In Phase 2, subjects were given a second opportunity to recall the occupation and name of the celebrity. Subjects in the face group who were presented with the face for a second time in Phase 2 were also given a card prior to the second presentation, giving details of what the celebrity was saying. This was to ensure that any advantage from hearing the voices in Phase 2 was due to the voice rather than to the content of what the person was saying.

Results and Discussion

Information Recalled in Phase 1. Figure 1 summarizes the performance of subjects when initially presented with a celebrity's face or voice. It provides details of the number of times that subjects (a) reported that a celebrity's face or voice was unfamiliar, (b) responded familiar-only, (c) recalled the celebrity's occupation but not their name, and (d) recalled the celebrity's name. An analysis of variance (ANOVA) of the failures to name the celebrity revealed that subjects failed to name the celebrity significantly more often in response to a voice than in response to a face, $F(1, 56) = 107.99$, $p < .01$. There was also a significant Type of Failure × Modality of Presentation interaction, $F(2, 112) = 35.86$, $p < .01$.Tests of simple main effects showed that subjects in the voice condition made significantly more unfamiliar decisions than did subjects in the face condition, $F(1, 56) = 85.86$, $p < .01$, and significantly more familiar-only responses than did subjects in the face condition, $F(1, 56) = 54.50$, $p < .01$. Subjects in the face condition recalled the celebrity's occupation but not their name significantly more often than did subjects in the voice condition, $F(1, 56) = 5.59$, $p < .05$. In none of the analyses was there any significant effect on Phase 1 responses of whether or not the subjects saw the celebrity in the same or in both modalities in Phase 2($F < 1$ in all cases).

When subjects responded that a celebrity was familiar without recalling their occupation or name, they were asked to indicate how familiar they felt the celebrity to be, on a scale of 1–3. It may well be, for instance, that the voice condition is associated with a comparatively large number of familiar-only responses that are rated to be of relatively low subjective familiarity. It can be seen from Figure 2 that this is not the case, however, A 3 × 2 ANOVA revealed a significant effect of subject group, $F(1, 58) = 59.38$, $p < .01$, but

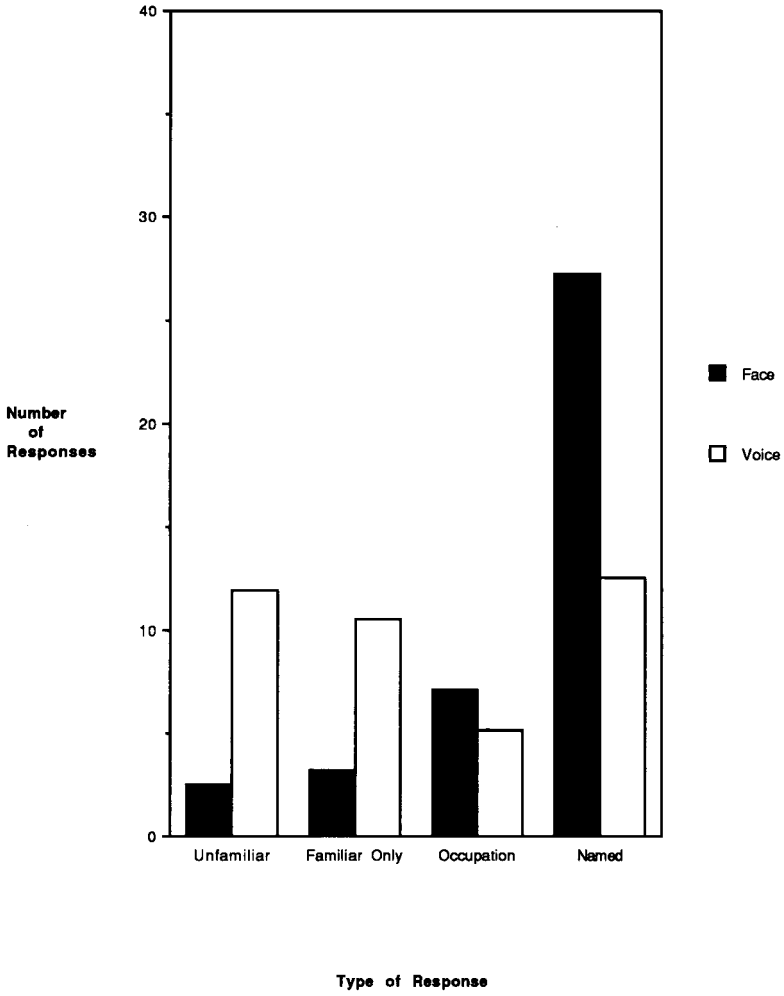

FIG. 1. The number of responses of each type (of a total of 60) that subjects made in Phase 1 while seeing a celebrity's face or hearing the celebrity's voice in Experiment 1.

no significant Subject Group × Degree of Familiarity interaction, $F(2, 116) = 1.50$, $p > .10$—that is, subjects in the voice condition made more familiar-only responses than did subjects in the face condition at all three levels of familiarity. There was no significant difference in the overall number of responses made at the three different levels of familiarity, $F(2, 116) = 1.92$, $p > .10$.

Clearly, then, subjects made significantly more familiar-only decisions in the voice condition than in the face condition. Consistent with this, subjects in the voice condition recalled fewer of the occupations and names of celebrities that they found familiar than did the subjects in the face condition. Subjects in the face condition found an average of 37.4 of the 40 celebrities to be familiar; of this 37.4, 34.3 (92%) were associated with successful recall of the occupation, and 27.2 (73%) were associated with successful recall

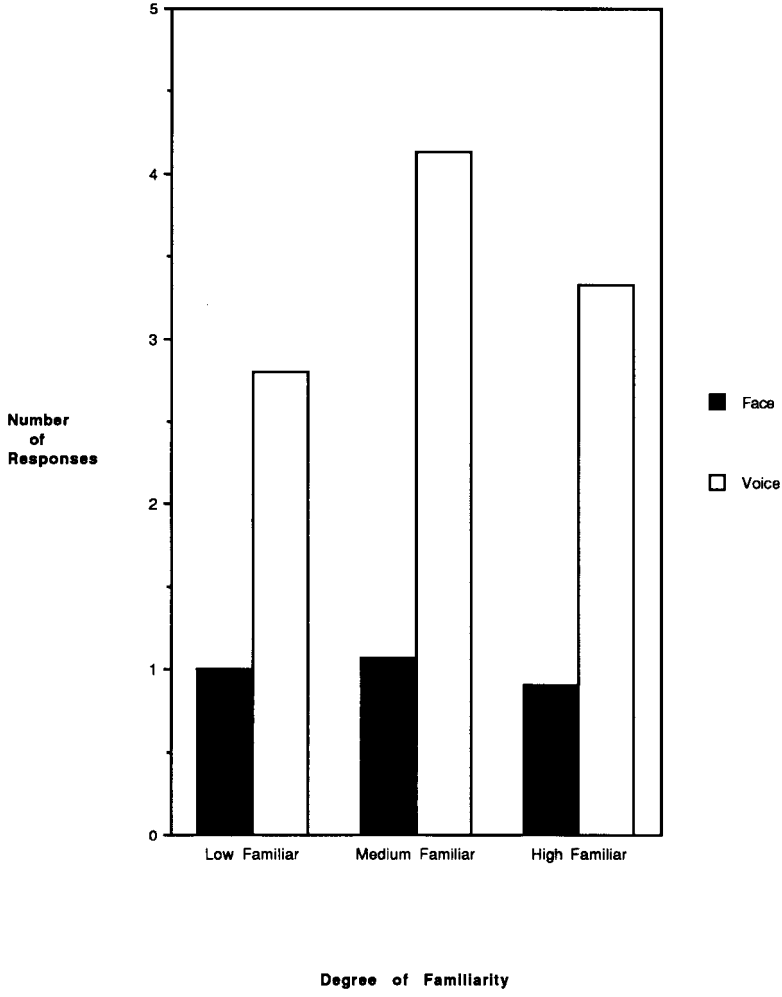

FIG. 2. The number of familiar-only responses that subjects made in the face and voice conditions in Phase 1 as a function of degree of familiarity in Experiment 1.

of the name. The figure of 92% is very similar to that reported by Hay et al. (1991). Their subjects were able to recall appropriate biographical information for 94% of the faces that they found to be familiar. As Hay et al. (p. 779) put it, "the system is extremely efficient in accessing semantic information after an FRU has been activated".

Subjects in the voice condition found an average of 28.1 of the 40 celebrities to be familiar, and of this 28.1, only 17.6 (63%) were associated with successful recall of the occupation, and only 12.5 (44%) were associated with successful recall of the name. Not only do subjects in the voice condition respond familiar to fewer of the celebrities, therefore, they also recall fewer names and fewer occupations as a proportion of the celebrities that they do find to be familiar.

Subjects in the voice condition made more false positive responses on trials where a non-famous person was presented than did subjects in the face condition, $F(1, 56) = 7.20$, $p < .01$. False alarm rates were .32 in the voice condition, and .21 in the face condition. Subjects in the face condition had a mean d' of 2.59, compared to a mean d' of 1.06 in the voice condition, $F(1, 58) = 126.32$, $p < .01$, indicating greater sensitivity in the face condition. In addition, subjects in the voice condition had significantly higher β scores than did subjects in the face condition, $F(1, 58) = 13.08$, $p < .01$. Mean β was 1.25 in the voice condition and 0.57 in the face condition.

As the voice group had significantly higher β scores than did the face group, the large number of familiar-only responses in the voice condition might have come about simply because this pool of items may contain a large number of items to which the subjects responded "familiar" on the basis of guesswork. If this pool of items does contain a large number of items of this kind, it would not be surprising if the subject was unable to recall the appropriate occupation or name. In an attempt to investigate this further, we selected a group of subjects from the two conditions who were as closely matched as possible in terms of number of hits and false alarms.

Table 1 therefore contains pooled data from the 16 subjects with the highest d' scores from the voice condition and the 12 subjects with the lowest d' scores from the face condition. (Because the face group had much higher d's than did the voice group overall, it was impossible to produce two equal-size groups who had similar numbers of hits and false alarms). It can be seen from Table 1 that the two groups are similar in overall number of familiar responses (439 vs. 435) and in total number of false alarms (68 vs. 72). Table 1 shows that even for this set of items, however, there are almost three times as many familiar-only responses in the voice group (142) than in the face group (48). Furthermore, subjects named only 50.3% (221/439) of the celebrities that they found familiar in the voice group, but they named 69.0% (300/435) of the celebrities that they found familiar in the face group. It therefore seems unlikely that the large number of familiar-only responses in the voice condition relative to the face condition is simply an artifact of the two groups' different overall level of performance, or of the higher number of false alarms made by subjects in the voice condition.

Information Recalled in Phase 2. Table 2 summarizes the findings when subjects who had responded unfamiliar or familiar-only to a voice in Phase 1 were given a second opportunity to recall the occupation of the celebrity in Phase 2. Results showed that

TABLE 1
A Comparison of the Performance of the 12 Subjects with the Lowest d's in the Face Condition with the 16 Subjects with the Highest d's in the Voice Condition

	Type of Response				
	Familiar	*False Alarms*	*Familiar-Only*	*Occupation Recalled but Not Named*	*Named*
Voice group	439	68	142	77	221
Face group	435	72	48	90	300

TABLE 2
The Proportion of Occupations that Subjects Recalled in Phase 2 in Response to Seeing the Person in Either the Same or a New Modality in Experiment 1

		State in Phase 1	
Phase 1	*Phase 2*	*Unfamiliar*	*Familiar-only*
Face	Face	0	.08
Face	Face & Voice	.07	.31
Voice	Voice	.08	.21
Voice	Face & Voice	.73	.81

subjects who were able to see the celebrity's face in Phase 2 recalled significantly more occupations in Phase 2 than did subjects who heard the voice for a second time, $F(1, 28) = 136.22$, $p < .01$. Subjects who had responded familiar-only in Phase 1 were significantly more likely to recall the occupation in Phase 2 than subjects who had found the face unfamiliar in Phase 1, $F(1, 28) = 8.40$, $p < .01$. The interaction between these factors was not significant ($F < 1$); subjects who found the face unfamiliar, as well as subjects who found the face familiar-only, recalled more information in response to the face cue than did subjects in the control condition who simply heard the voice for a second time.

Results from the subjects who saw the face in Phase 1 showed that subjects who heard the celebrity's voice in Phase 2 recalled significantly more occupations in Phase 2 than did subjects who saw the face for a second time, $F(1, 28) = 10.44$, $p < .01$. Subjects who had responded familiar-only in Phase 1 were significantly more likely to recall the occupation in Phase 2 than were subjects who had found the face unfamiliar in Phase 1, $F(1, 28) = 16.91$, $p < .01$. The interaction between these factors just failed to reach significance, $F(1, 28) = 3.51$, $p = .07$. Table 2 shows that subjects who found the face unfamiliar and subjects who found the face familiar-only recalled more information in response to the voice cue than did subjects in the control condition who simply saw the face for a second time.

Finally, there was a significant Modality of Presentation in Phase 1 × Modality of Presentation in Phase 2 interaction on the number of occupations recalled in Phase 2, $F(1, 28) = 40.10$, $p < .01$, when subjects were in a familiar-only state. This shows that a cue in a new modality in Phase 2 was more helpful when it was a face following a voice than when it was a voice following a face.

In this experiment, of course, presenting the person in the same modality meant simply replaying exactly the same stimulus. If a different stimulus had been presented in the same modality condition in Phase 2, then a rather different set of results might have been anticipated, particularly when subjects found the original stimulus unfamiliar (cf. Hanley & Cowell, 1988, Experiment 2).

EXPERIMENT 2

One potentially important difference between the voices and faces condition in Experiment 1 was the fact that the subjects in the voice group had access to the words that the celebrities were saying during Phase 1. Subjects in the face group would only have been

able to determine what the subject was saying during Phase 1 of the experiment if they were able to lip-read. It is conceivable that exposure to the semantic content of the material that the stimulus voice was saying may have influenced the probability that subjects responded "familiar" to the voices. This is a possible reason why subjects may have made more familiar-only responses to voices than to faces in Phase 1 of Experiment 1. In order to investigate this possibility, Experiment 1 was repeated, but with all subjects having access to the information that the subject was saying during phase 1 of the Experiment.

Method

Subjects

Sixty subjects, drawn from the same population as those used in Experiment 1, took part in the Experiment. Half the subjects were randomly allocated to the voice group and half to the face group.

Stimuli and Procedure

These were identical to Experiment 1, with the exception that all subjects were presented with a card before being exposed to each voice or face in Phase 1 of the experiment. This card contained a transcription of what each person (both famous and non-famous) said on the tape. Phase 2 of the experiment took place as in Experiment 1.

Results

Figure 3 summarizes the performance of subjects when initially presented with a celebrity's face or voice. An ANOVA of the failures to name the celebrity revealed that subjects failed to do so significantly more often in response to a voice than in response to a face, $F(1, 56) = 91.80$, $p < .01$. There was also a significant Type of Failure $\times$ Modality of Presentation interaction, $F(2, 112) = 34.15$, $p < .01$. Tests of simple main effects showed that subjects in the voice condition made significantly more unfamiliar decisions than did subjects in the face condition, $F(1, 56) = 52.80$, $p < .01$, and significantly more familiar-only responses than did subjects in the face condition, $F(1, 56) = 45.89$, $p < .01$. Subjects in the face condition recalled the celebrity's occupation but not their name significantly more often than did subjects in the voice condition, $F(1, 56) = 23.80$, $p < .01$. In none of the analyses was there any significant effect on Phase 1 responses of whether or not the subjects saw the celebrity in the same or in both modalities in Phase 2 ($F < 1$ in all cases).

As in Experiment 1, a 3 $\times$ 2 ANOVA investigated the strength of subjects' familiarity ratings when in a familiar-only state as a function of modality of presentation. This revealed a significant effect of subject group, $F(1, 58) = 42.61$, $p < .01$, but no significant Subject Group $\times$ Degree of Familiarity interaction ($F < 1$). As in Experiment 1, subjects in the voice condition made significantly more familiar-only responses than did subjects in the face condition at all three levels of familiarity. There was no significant difference in the overall number of responses made at the three different levels of familiarity ($F < 1$). Figure 4 contains details of the number of responses of each type made by the subjects.

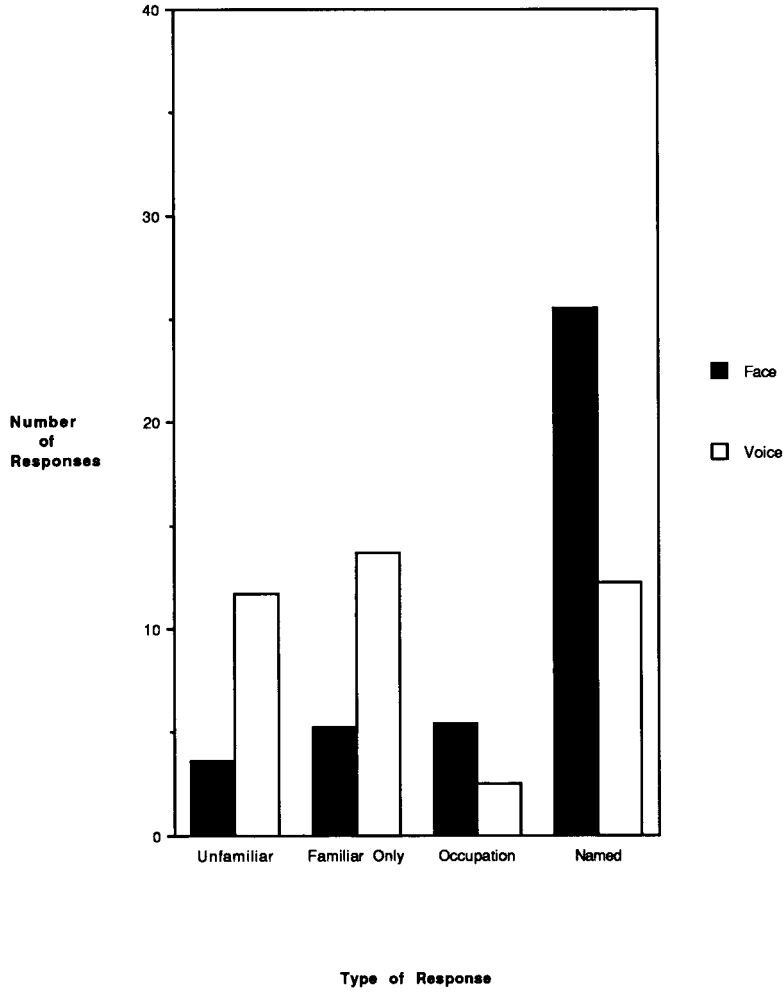

FIG. 3. The number of responses of each type (out of a total of 60) that subjects made in Phase 1 while seeing a celebrity's face or hearing the celebrity's voice in Experiment 2.

Consistent with the results of Experiment 1, subjects in the voice condition recalled less information about celebrities that they found familiar than did the subjects in the face condition. Subjects in the face condition found an average of 36.4 of the 40 celebrities to be familiar, and of this 36.4, 30.9 (85%) were associated with successful recall of the occupation and 25.5 (70%) were associated with successful recall of the name. Subjects in the voice condition found an average of 28.3 of the 40 celebrities to be familiar, and of this 28.3, only 14.7 (52%) were associated with successful recall of the occupation and only 12.2 (43%) were associated with successful recall of the name. As in Experiment 1, therefore, subjects in the voice condition responded "familiar" to fewer of the celebrities and recalled fewer names and fewer occupations as a proportion of the celebrities whom they did find to be familiar.

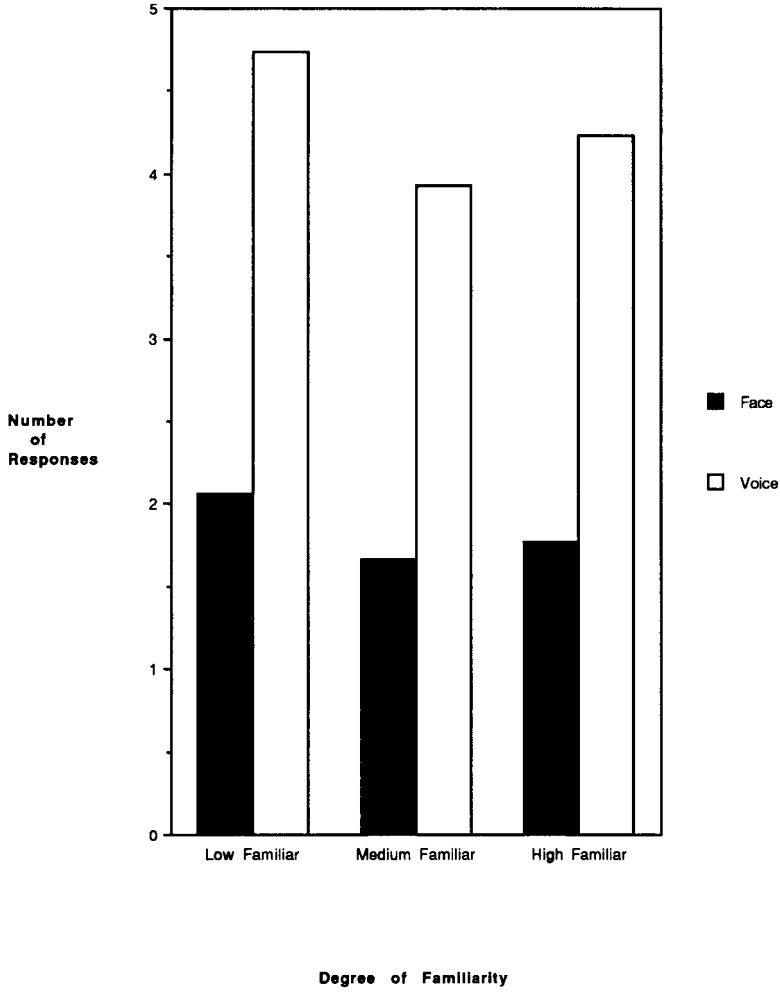

FIG. 4. The number of familiar-only responses that subjects made in the face and voice conditions in Phase 1 as a function of degree of familiarity in Experiment 2.

Subjects in the voice condition made more false positive responses on trials where a non-famous person was presented than did subjects in the face condition, $F(1,56) = 14.77$, $p < .01$. False alarm rates were .30 in the voice condition, and .17 in the face condition. As in Experiment 1, d' was higher in the face condition than in the voice condition, and β was higher in the voice condition than in the face condition. Once again, therefore, a group of subjects were selected from the two groups whose d' scores were as closely matched as possible. The closest match was achieved by selecting 8 subjects from the 10 best performers (highest d') in the voice condition, and 8 subjects from the worst 10 performers (lowest d') in the face condition. The mean d' was 1.78 for the voice subset and 1.93 for the face subset. The two groups were similar in overall number of familiar responses (254 vs. 269, respectively) and in total number of false alarms (31 vs. 36). Even

for this set of items, however, there were over two and a half times as many familiar-only responses in the voice group (125) than in the face group (49). As in Experiment 1, therefore, it seems unlikely that the large number of familiar-only responses in the voice condition relative to the face condition is simply an artifact of the two groups' different overall level of performance or of the higher number of false alarms made by subjects in the voice condition.

The results from Phase 2 were similar to those obtained in Experiment 1. Subjects in both the voice and face conditions performed significantly better when they saw the celebrity in a new modality in Phase 2 than in the same modality again. This effect occurred regardless of whether the subject found the item to be familiar or unfamiliar in Phase 1.

Discussion

It is therefore clear that all the major findings from Phase 1 of Experiment 1 have been replicated in Experiment 2. Most important, it is clear that subjects make more familiar-only responses to voices than faces even when the content of what the person is saying is made available to them when they are attempting to recognize the faces.

GENERAL DISCUSSION

The results of both experiments have revealed that people can be successfully identified much more readily from their face than from their voice. This is consistent with research comparing the effectiveness of face and voice recognition in tests of episodic memory (e.g. Yarmey et al., 1994). One problem associated with voices concerns basic recognition: a person's face is more likely to be judged as being familiar than is their voice. In the case of celebrities, this is unsurprising: because we are often exposed to photographs of their faces in magazines and newspapers, we probably perceive celebrities' faces much more frequently than we perceive their voices. This suggests that it may generally be much more difficult to activate a VRU than an FRU for a familiar person.

The most interesting problem associated with voice processing that was observed in this study concerns access to person-specific semantic information about a person from their voice. On a large number of occasions a subject felt that a voice was familiar but was unable to recall any information about the person concerned. This situation occurred significantly more frequently for voices than for faces, suggesting that it is particularly difficult to retrieve semantic information from memory in response to a voice relative to a face. As in previous research (Hay et al., 1991), appropriate semantic information was recalled relatively successfully whenever a face was found to be familiar. The relevant figures were 92% in Experiment 1 and 85% in Experiment 2. When a voice was deemed familiar, however, appropriate semantic information was recalled only 63% of the time in Experiment 1 and only 56% of the time in Experiment 2.

Because both experiments in the present study used a between-subjects design, it is important to consider whether these results might have been caused by subjects using different criteria for familiarity decisions in the voice and face conditions. One possibility is that subjects were simply more cautious about recalling semantic information for people

in response to their voice than in response to their face. This seems unlikely because, if anything, subjects in the voice condition were less cautious, producing significantly more false alarms to non-celebrities than did the subjects in the face condition. This does, however, raise the possibility that in the voice condition there may be more items that were judged familiar without the subject genuinely knowing them than there are in the face condition. That is, there may be an artificially high level of familiar-only responses in the voice condition. Even when we compared a subset of subjects from the face and voice conditions who were matched for number of hits and false alarms (e.g. Table 1), however, there were still very many more familiar-only responses in the voice condition than the face condition. Although it would clearly be interesting to use a within-subjects design in future research, we therefore believe that the finding that voices elicit more familiar-only responses does genuinely reflect the fact that retrieval of semantic information is particularly difficult from voices.

The main theoretical issue addressed by this study is the way in which the familiar-only state should be characterized in contemporary models of person recognition (Bruce & Young, 1986; Burton et al., 1991). One of the possibilities outlined earlier in the paper is that this state may reflect a blockage between the FRU or VRU and the PIN. If there is a stronger activation weight from the FRU to the PIN than from the VRU to the PIN for a large number of known people, then one would expect more familiar-only responses to hearing voices than to seeing faces. This would explain why person-specific semantic information appears to be particularly difficult to retrieve in response to a voice. If, however, familiar-only responses always reflect a block between the PINs and the semantic information pool, then they should have been equally likely to occur for faces as for voices. This is a point in the system at which processing of faces and voices is considered to have merged, and so there is no reason to expect that the number of familiar-only responses would vary as a function of modality. The fact that there was a larger number of familiar-only responses in the present study when subjects were attempting to recognize voices than faces, therefore, indicates that a significant number of familiar-only experiences reflect a block between the VRU and FRUs and the PIN rather than a block between the PIN and the semantic information store.

The finding that seeing the faces in Phase 2 of Experiment 1 helped subjects to retrieve semantic information about people when they were in a familiar-only state from hearing their voice is also consistent with this account. If the block lay between the PIN and the semantic information pool, then presenting the person in a new modality should not prove any more beneficial than presenting the person in the same modality. This is because the new modality would not then provide an alternative means of accessing semantic information about the person. Presenting a person in a new modality *would* be expected to help subjects in a familiar-only state, however, if the block occurs between FRUs or VRUs and the PIN. If a voice has failed to activate the PIN strongly because of a block between the VRU and the PIN, then a face cue may more strongly activate the PIN via the separate FRU–PIN pathway. It is also interesting to note (see Table 2) that when subjects were in a familiarity-only state, they benefited more in Phase 2 when the new modality stimulus was a face rather than a voice. This asymmetric cueing effect is exactly what one would predict if there are generally stronger links between an FRU and a PIN than between a VRU and a PIN. Overall, therefore, the results from Phase 2 also suggest

that a large number of the familiarity-only responses observed in this experiment reflect a block between the VRU and FRUs and the PIN rather than a block between the PIN and the semantic information store.

As Bruce and Young (1986) argued that familiar-only responses come about when an FRU or VRU fires but fails to activate the PIN, the present results can be readily accommodated in terms of their model. The results can also be explained in terms of Burton et al. (1990, 1991), because a weak associative link between the VRU and the PIN could produce a familiar-only response in their model. At first sight, this might seem strange in the sense that familiarity responses are supposed to reflect activity at the level of the PIN rather than activity at the FRU or VRU in the IAC model. In a cascade model, however, reduced activation at the PIN during voice identification could still be strong enough to produce a feeling of familiarity despite being too weak to allow the appropriate occupation to be recalled from the semantic information pool. The consequence of this would be a familiar-only response.

In fact, the Burton et al. model is consistent with a further possible account of why there are more familiar-only responses to voices than to faces. At the start of this section, we pointed out that the relatively high number of failures to recognize voices suggests that it is more difficult to activate a VRU when we hear a person's voice than an FRU when we see a person's face. It therefore appears that all parts of the voice identification system up to the PINs are relatively inefficient compared to face identification. According to a cascade model, weak activation at the level of the VRU might in itself lead to reduced activation at the level of the PIN irrespective of the strength of the activation weight between the VRU and the PIN. Once again, this could lead to a level of activation at the PINs during voice identification that is sufficient to allow voices to be identified as familiar but is too weak to permit recall of semantic information.

An opponent of Burton et al.'s model might argue that the account of the results offered by the IAC model is difficult to reconcile with the finding that there were significantly more familiar-only responses in the voice condition even on trials where the familiarity rating of the voice was reported as being very strong (see Figures 2 and 4). If one assumes that feelings of familiarity reflect the amount of activation at the PIN, then the level of activation on these trials should surely have been strong enough to allow retrieval of semantic information to occur. However, the IAC model makes no attempt to relate level of activation at the PIN to the strength of the feeling of familiarity and is therefore immune to evidence of this kind. An advantage of the Bruce and Young (1986) model is that it can readily explain the data from Figures 2 and 4 because Bruce and Young argue that it is the VRU rather than the PIN that signals familiarity. The VRU could therefore be very strongly activated even if the link from the VRU to the PIN was very weak. As was pointed out earlier, however, the familiar-only responses made by patient ME (de Haan et al., 1991) do appear to reflect a block between the PIN and the semantic information pool (Burton et al., 1991) and are therefore difficult to reconcile with Bruce and Young's account of the locus of the feeling of familiarity. The advantage of Burton et al.'s IAC model in this respect is that it can accommodate both types of familiar-only experience.

REFERENCES

Brennen, T., Baguley, T., Bright, J., & Bruce, V. (1990). Resolving semantically induced tip-of-the-tongue states for proper nouns. *Memory and Cognition, 18*, 339–347.

Bruce, V., Burton, A.M., & Walker, S. (1994). Testing the models? New data and comments on Stanhope & Cohen (1993). *British Journal of Psychology, 85*, 335–349.

Bruce, V., & Young, A.W. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305–327.

Burton, A.M., & Bruce, V. (1992). I recognise your face but I can't remember your name: a simple explanation? *British Journal of Psychology 83*, 45–60.

Burton, A.M., Bruce, V., & Johnston, R.A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology, 81*, 361–380.

Burton, A.M., Young, A.W., Bruce, V., Johnston, R., & Ellis, A.W. (1991). Understanding covert recognition. *Cognition, 39*, 129–166.

Ellis, H.D., Jones, D.M., & Modsell, N. (1997). Intra and inter-modal repetition priming of familiar faces and voices. *British Journal of Psychology, 88*, 143–156.

de Haan, E.H.F., Young, A.W., & Newcombe, F. (1991). A dissociation between the sense of familiarity and access to semantic information concerning familiar people. *European Journal of Cognitive Psychology, 3*, 51–67.

Hanley, J.R. (1995). Are names difficult to recall because they are unique? A case study of a patient with anomia. *Quarterly Journal of Experimental Psychology, 48A*, 487–506.

Hanley, J.R., & Cowell, E.S. (1988). The effects of different types of retrieval cues on the recall of names of famous faces. *Memory and Cognition, 16*, 545–555.

Harris, D.M., & Kay, J. (1995). I recognise your face but I can't remember your name: is it because names are unique? *British Journal of Psychology, 86*, 345–358.

Hay, D.C., Young, A.W., & Ellis, A.W. (1991). Routes through the face recognition system. *Quarterly Journal of Experimental Psychology, 43A*, 761–791.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception. *Psychological Review, 88*, 375–407.

Rumelhart, D.E., & McClelland, J.L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some test and extensions of the model. *Psychological Review, 89*, 60–94.

Schweinberger, S.R., Herholz, A., & Sommer, W. (1997). Recognizing famous voices: Influence of stimulus duration and different types of retrieval cues. *Journal of Speech and Hearing Research, 40*, 453–463.

Stanhope, N., & Cohen, G. (1993). Retrieval of proper names: Testing the models. *British Journal of Psychology, 84*, 51–65.

Warrington, E.K., & McCarthy, R.A. (1988). The fractionation of retrograde amnesia. *Brain & Cognition, 7*, 184–200.

Yarmey, A.D., Yarmey, A.L., & Yarmey, M.J. (1994). Face and voice identifications in showups and lineups. *Applied Cognitive Psychology, 8*, 453–464.

Young, A.W., Hay, D.C., & Ellis, A.W. (1985). The faces that launched a thousand slips: Everyday difficulties and errors in recognising people. *British Journal of Psychology, 76*, 495–523.

Original manuscript received 7 August 1995
Accepted revision received 16 July 1997