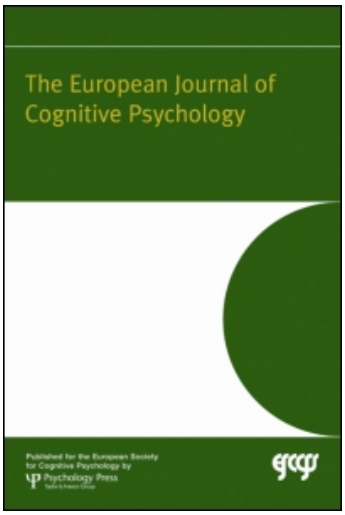

European Journal of Cognitive Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713734596

Recalling semantic information about personally known faces and voices

Serge Brédart [a]; Catherine Barsics [a]; Rick Hanley [b]
[a] University of Liège, Liège, Belgium [b] University of Colchester, Colchester, UK

First Published on: 01 March 2009

To cite this Article Brédart, Serge, Barsics, Catherine and Hanley, Rick(2009)'Recalling semantic information about personally known faces and voices',European Journal of Cognitive Psychology,21:7,1013 — 1021
To link to this Article: DOI: 10.1080/09541440802591821
URL: http://dx.doi.org/10.1080/09541440802591821

EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY
2009, 21 (7), 1013–1021

Psychology Press
Taylor & Francis Group

Recalling semantic information about personally known faces and voices

Serge Brédart and Catherine Barsics
University of Liège, Liège, Belgium

Rick Hanley
University of Colchester, Colchester, UK

Previous research that investigated whether biographical information about familiar people is harder to retrieve from voices than from faces produced contrasting results. However, studies that used a strict control of the content of spoken extracts reported that semantic information about familiar people is easier to retrieve when recognising a face than when recognising a voice. In all previous studies faces and voices of famous people were used as stimuli. In the present study, personally familiar people's voices and faces (standard faces and blurred faces) were used. Presenting such people (i.e., participants' teachers) allowed controlling still more strictly the content of the spoken extracts since it was possible to ask all the target persons to speak the same words. In addition, it was previously stressed that we encounter famous people's faces in the media more frequently than we hear their voice. This methodological difficulty was presumably reduced when teachers' faces were presented. Present results showed a significant decrease in retrieval of biographical information from familiar voices relative to blurred faces even though the level of overall recognition was similar for blurred faces and voices. The role of the relative distinctiveness of voices and faces is discussed and further investigation is proposed.

Keywords: Face; Voice; Person recognition.

Access to semantic and lexical information during person identification from faces and proper names has been extensively investigated during the last two decades (for reviews see Gobbini & Haxby, 2007; Hanley & Cohen, 2008; Valentine, Brennen, & Brédart, 1996; Young & Ellis, 1989). In recent years, several studies have started to investigate the retrieval of semantic information when familiar voices are recognised, and more precisely, to compare the retrieval of biographical information about familiar people from faces and

Correspondence should be addressed to Serge Brédart, Department of Cognitive Science (B-32), University of Liège, B-4000 Liège, Belgium. E-mail: serge.bredart@ulg.ac.be

http://www.psypress.com/ecp DOI: 10.1080/09541440802591821

voices (e.g., Damjanovic & Hanley, 2007; Hanley, Smith, & Hadfield, 1998; Hanley & Turner, 2000). The present study was aimed at examining further the retrieval of semantic information associated with the recognition of familiar individuals from faces and voices by using personally familiar individuals rather than celebrities as target people.

Both faces and voices are means of person identification that we use in everyday life. However, available data suggest that semantic information about familiar people is easier to retrieve when recognising a face than when recognising a voice. Hanley et al. (1998) showed that participants whose task was to recognise famous people from hearing their voice reported more "familiarity-only" experiences than participants who recognised the same celebrities by seeing their face. In other words, participants who heard voices were more frequently unable to recall a target person's occupation while they found the voice familiar than participants who found the corresponding face familiar. However, in Hanley et al.'s study the recognition level was lower (60–70%) for voices than for faces (more than 90%), and the rate of false alarms was higher in the voice condition (about 30%) than in the face condition (about 20%). A face or a voice was considered to be recognised when a participant judged it as being familiar. It was possible that this pattern of results reflected the fact that the participants produced "familiar" responses on the basis of guesswork more often in the voice condition than in the face condition. To avoid this problem, Hanley and Turner (2000) used blurred faces as stimuli in order to bring down face recognition performance to the same level as recognition in the voice condition. They found that the numbers of familiar-only experiences were similar when blurred faces were presented and when voices were presented. Therefore, the recall of occupation was not more difficult for voices than for blurred faces. Such results did not support the view that it is more difficult to associate semantic information with people's voice than with people's face. Nevertheless, more recent studies (Damjanovic & Hanley, 2007; Hanley & Damjanovic, 2008) suggested that the latter results were due to methodological problems. The spoken extracts used in the voice condition of the Hanley and Turner study were likely to provide cues as to the occupation of some of the target celebrities. When these problems were avoided by employing the Schweinberger, Herholz, and Steif (1997) procedure designed to limit the extent to which the speech content of the extracts could give clues to a speaker's identity, the results unambiguously indicated that more semantic details could be recalled from blurred faces than from voices even though overall recognition performance was similar for both types of stimuli. Although, these two studies (Damjanovic & Hanley, 2007; Hanley & Damjanovic, 2008) confirm the view that access to semantic information is easier from faces, even blurred, than voices, further studies are needed.

The aim of the present study was to examine further whether or not semantic information is more difficult to retrieve from faces than from voices since previous research produced inconsistent results. Up to now access to semantic information from familiar faces or voices has been studied with famous people as stimuli. However, famous people do not seem to be the more suitable stimuli in order to compare the ability of faces and voices to allow access to semantic information. We are probably more often exposed to famous people's faces than to their voices. Indeed, we presumably see the faces of actors, sport people or politicians without hearing their voices in magazines or newspapers much more frequently than we hear their voices without seeing their faces. This problem was acknowledged earlier (see Hanley et al., 1998; Hanley & Turner, 2000). An exception is probably the category of pop stars because we often hear their voices on the radio without seeing their faces. But, unfortunately, previous studies did not separately report data for this particular category of people.

Therefore, it is possible that faces are not special in their ability to allow access to biographical information but that the superiority of faces is due to the fact that we are more frequently exposed to famous people's faces than to their voices. Faces and voices of personally familiar people such as teachers seem to be interesting stimuli in order to compare access to semantic information from faces and voices because when we meet such people we usually both see and hear them. It is difficult to quantify this, but intuitively one might even think that students are more massively exposed to their teachers' voices than to their faces because during lessons they need to take notes and look at slides without looking at the teachers. For this reason, in the present study we investigated the retrieval of semantic information when faces and voices of professors or teaching assistants are recognised.

In summary, the main empirical questions addressed in the present paper was to know whether semantic information (e.g., the subject of a professor's course) is more often retrieved from faces than from voices when personally familiar people such as teachers are used as stimuli.

METHOD

Participants

Fifty-four second year undergraduate volunteer students (39 female) of the University of Liège took part in the experiment. They were aged between 19 and 24 (mean age = 20.2 years).

They had known the presented familiar persons for at least 3 months (they used to attend the professors' or teaching assistants' courses 2 hours a week). All the participants had normal or corrected-to-normal vision and

normal hearing. All were native French speakers. They all gave written informed consent.

Stimuli and materials

Sixteen familiar and 16 unfamiliar individuals were filmed while they were reading a prompter that showed the first article of the Universal Declaration of Human Rights: "All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood." These individuals had the opportunity to refamiliarise with the wording of this article before the recording and were instructed to read the article in an emotionally neutral tone and in keeping a neutral facial expression. They spoke a French translation of the article.

Familiar and unfamiliar items were matched for gender (11 men and five women in each type of items) and mean age (men 43.2 years, $SD = 10.7$; women 43.6 years, $SD = 11.1$). All the faces were filmed in front of the same off-white wall. All people wore the same black hairdressing gown in order to avoid any sartorial cue to the person's identity.

Visual stimuli (standard faces and blurred faces) were presented on a 17-inch monitor controlled by a PC computer and were viewed at a distance of 56 cm controlled by means of a chinrest. The computer ran Windows Media Player Software and the video clips were displayed in full screen mode. The same software was also used to present the voices through a pair of speakers. The computer keyboard was used by the experimenter to monitor the presentation of the stimuli.

In the standard and blurred face conditions as well as in the voice condition, the 32 faces were presented in one of two predefined random orders. In the standard and blurred face conditions, each video clip lasted 7 s, and in the voice condition each voice extract also lasted 7 s (therefore the number of words read varied from a target person to another). In the blurred face condition, the same visual stimuli as those used in the standard face condition were edited through the application of a Gaussian blur filter in Adobe Premiere Pro 1.5. All the clips were rendered with a blur value of 37.

As in the Damjanovic and Hanley (2007) study, to ensure that overall familiarity was similar in the voice and the blurred face conditions, 32 other participants drawn from the same population as those described in the Participants section were asked to rate the familiarity of the targets on a scale of 0 = "unfamiliar" to 3 = "very familiar". Sixteen participants rated the blurred faces and gave them a mean familiarity of 1.63 ($SD = 0.57$). The remaining 16 participants rated the voices and gave them a mean familiarity

of 1.75 ($SD = 0.71$). There was no significant difference between these two means, $t(30) = 0.51$.

Procedure

The participants were informed that they would be presented with a sequence of faces, blurred faces or voices, depending on the condition to which they were assigned. They were told that some of these faces/voices were well known to them whereas the other faces/voices belonged to persons that were unknown to them. After the presentation of each item, the clip was paused to enable the participants to respond. For the recognition task itself, the participants had to answer "yes" when they recognised the presented item and "no" when they found it unfamiliar. If they answered "yes", the experimenter also assessed whether the participants were able to recall the target person's name and biographical details such as the subject of a professor's or a teaching assistant's course.

RESULTS

The design included only one between subjects factor: the modality of presentation (standard faces, blurred faces, and voices).

An alpha level of .05 was set for all statistical tests.

Overall recognition

A one-way ANOVA, with the modality of presentation (or condition) as the independent factor, was conducted on the number of hits, and on the number of false alarms respectively. This analysis revealed a significant effect of the modality on the number of hits, $F(2, 51) = 30.38$, $MSE = 0.02$, and also on the number of false alarms, $F(2, 51) = 16.88$, $MSE = 0.01$. HSD Tukey post hoc tests indicated that there was significantly more hits in the standard face condition than in the blurred face and voice conditions, and significantly fewer false alarms in the standard face condition than in the blurred face and voice conditions. The voice and the blurred face conditions did not differ significantly in terms of hits or false alarms. Descriptive data are presented in Table 1.

We used A' as a measure of discrimination, and $B''D$ as a measure of bias (Donaldson, 1996). A one-way ANOVA showed a significant effect of condition on discrimination, $F(2, 51) = 35.68$, $MSE = 0.01$. Post hoc tests indicated that discrimination was significantly higher in the standard face than in the blurred face condition, and in the blurred face than in the voice

TABLE 1
Mean overall proportions (with standard deviations) of hits, false alarms, A' (discrimination) and $B''D$ (bias) measures for the standard face, blurred face, and voice conditions

	Hits		*False alarms*		A'		$B''D$	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Standard faces	0.93	0.11	0.05	0.09	0.96	0.04	0.07	0.54
Blurred faces	0.65	0.15	0.17	0.12	0.82	0.07	0.42	0.41
Voices	0.59	0.16	0.24	0.09	0.76	0.10	0.34	0.38

condition. The effect of condition on bias was not significant although a tendency occurred, $F(2, 51) = 2.91$, $MSE = 0.20$, $p = .06$. However, post -hoc Tukey tests revealed no significant difference between the conditions.

We also observed some occurrences of confusions between two professors, although this kind of errors remained exceptional, respectively 2.1% in the voice condition, 2.8% in the blurred face condition, and 0% in the standard face condition. These cases were removed in the next analyses.

Recall of semantic information and names

The retrieval of semantic information in response to faces and voices was examined. In any case, when semantic information was recalled, participants specified whether the person was a professor or a teaching assistant, and the subject associated with the person (e.g., "professor of social psychology", or "teaching assistant in statistics"). Table 2 shows the mean proportion of trials for which such semantic information was recalled. A one-way ANOVA revealed a significant effect of condition, $F(2, 51) = 64.19$, $MSE = 0.03$. Post hoc HSD tests indicated that the proportion of trials that yielded the recall of semantic information was higher in the standard face condition than in the blurred face condition, and higher in the blurred face condition than in the voice condition.

The same pattern of results was observed when recall of semantic information was conditionalised on hits, $F(2, 51) = 45.27$, $MSE = 0.02$. Tukey post hoc tests revealed the same "standard faces > blurred faces > voices" significant inequality pattern (see Table 2 for descriptive data).

The recall of names was also analysed. A one-way ANOVA revealed a significant effect of condition, $F(2, 51) = 43.08$, $MSE = 0.03$. Post hoc HSD tests indicated that the proportion of trials for which the correct name was recalled was higher in the standard face condition than in the blurred face condition, and higher in the blurred face condition than in the voice condition. Descriptive data are presented in Table 2. Again the same pattern

TABLE 2
Overall and conditionalised on the hit rate mean proportions of semantic and names recalled in the standard face, blurred face, and voice condition

	Semantic information		*Name*	
Condition	*M*	*SD*	*M*	*SD*
Overall				
Standard faces	0.89	0.11	0.78	0.14
Blurred faces	0.51	0.18	0.46	0.18
Voices	0.30	0.17	0.29	0.16
Conditionalised on hits				
Standard faces	0.96	0.06	0.83	0.23
Blurred faces	0.77	0.18	0.69	0.21
Voices	0.47	0.19	0.46	0.18

of results was observed when recall of names was conditionalised on hits, $F(2, 51) = 21.34$, $MSE = 0.03$. Tukey post hoc tests revealed the same "standard faces > blurred faces > voices" significant inequality pattern (see Table 2).

DISCUSSION

Is semantic information about people more often retrieved from a recognised familiar face than from a recognised familiar voice? As mentioned earlier, previous research that investigated this question provided contrasting results. In the Damjanovic and Hanley (2007) study, semantic information was more frequently reported from standard than from blurred faces and, from blurred faces than from voices whereas, in the Hanley and Turner (2000) study, semantic information was recalled more often from standard than from blurred faces, but blurred faces and voices produced similar rates of semantic information retrieval. For this reason, we decided to examine this question further by using faces and voices of personally familiar persons as stimuli rather than those of famous people. The use of personally familiar persons had two advantages. First, it allows controlling more strictly the content of the spoken extracts presented as stimuli. Indeed, it was possible to ask all the target persons to speak the same words (here the first article of the Universal Declaration of Human Rights). Such a procedure would obviously not have been possible to apply if famous people had been chosen as target persons. Second, as noted in previous papers (e.g., Hanley et al., 1998; Hanley & Turner, 2000), we probably encounter a famous person's face in the media more frequently than we hear their voice. By choosing professors or teaching assistants as target individuals, we think this problem

was reduced, and that the comparison between faces and voices was more equitable.

Present results are straightforward. They are totally consistent with those from the Damjanovic and Hanley (2007) and the Hanley and Damjanovic (2008) studies. Indeed, present findings showed a significant and numerically substantial decrease in retrieval of occupational details and names from familiar voices relative to blurred faces even though the level of overall recognition (e.g., the rates of hits and false alarms) was similar for blurred faces and voices.

In addition to the level of overall recognition, it seems that controlling the recognition speed by recording the participants' reaction times to blurred faces and voices would be particularly useful. Indeed, it is possible that the superiority of blurred faces compared with voices is, at least partly, due to a faster recognition of blurred faces, leaving more time for semantic activation within the 7 s during which the stimuli were presented in the blurred face condition than in the voice condition. This point should be addressed in future research.

As noticed earlier by Damjanovic and Hanley (2007), such results potentially raise problem for models of person recognition in which familiarity decisions are made on a modality-free person identity node, i.e., after the face and voice recognition systems processed the current stimulus (Brédart, Valentine, Calder, & Gassi, 1995; Burton, Bruce, & Johnston, 1990; for a review see Young & Burton, 1999). Indeed for such models, since the familiarity decision reflects the activation of a person identity node, it should be as difficult to retrieve semantic information and names from a face found familiar as from a voice found familiar especially when target voices and faces were matched for familiarity. The present results seem to be more consistent with the original Bruce and Young model (1986).

Damjanovic and Hanley (2007) considered that results such as the present ones are consistent with the possibility that there are closer connections between the face recognition system and biographical information stored in semantic memory than between the voice recognition system and biographical information (see also Gainotti, Barbier, & Marra, 2003). There is, however, another possible interpretation of such results. It is possible that we distinguish more easily between faces than between voices. Such a hypothesis may be empirically tested. If this hypothesis is correct, in other words if distinctiveness is a key factor, one would expect distinctive voices to yield a better recall of semantic information and name than nondistinctive voices. The same prediction could also be stated for faces. In addition, from such a hypothesis it is possible (although this is not mandatory prediction) to imagine that as much, or even more, semantic information could be retrieved from distinctive voices than from nondistinctive faces even when they are

matched for familiarity. Further research is needed to evaluate such a hypothesis.

In conclusion, present results confirm that semantic information and names are more likely to be retrieved from familiar faces, even when blurred, than from familiar voices.

Original manuscript received August 2008
Revised Manuscript received October 2008
First published online March 2009

REFERENCES

Brédart, S., Valentine, T., Calder, A., & Gassi, L. (1995). An interactive activation model of face recognition. *Quarterly Journal of Experimental Psychology*, *48A*, 466–486.

Bruce, V., Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *76*, 373–383.

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

Damjanovic, L., & Hanley, J. R. (2007). Recalling episodic and semantic information about famous faces and voices. *Memory and Cognition*, *35*, 1205–1210.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory and Cognition*, *24*, 523–533.

Gainotti, G., Barbier, A., & Marra, C. (2003). Slowly progressive defect in recognition of familiar people in a patient with right anterior temporal atrophy. *Brain*, *126*, 792–803.

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*, 32–41.

Hanley, J. R., & Cohen, G. (2008). Memory for faces. In G. Cohen & M. Conway (Eds.), *Memory in the real world* (3rd ed., pp. 107–140). Hove, UK: Psychology Press.

Hanley, J. R., & Damjanovic, L. (2008). It is more difficult to retrieve a familiar person's name and occupation from their voice than from their blurred face. *Manuscript submitted for publication*.

Hanley, J. R., Smith, S. T., & Hadfield, J. (1998). I recognise you but I can't place you: An investigation of familiar-only experiences during tests of voice and face recognition. *Quarterly Journal of Experimental Psychology*, *51A*, 179–195.

Hanley, J. R., & Turner, J. M. (2000). Why are familiar-only experiences more frequent for voices than for faces? *Quarterly Journal of Experimental Psychology*, *53A*, 1105–1116.

Schweinberger, S. R., Herholz, A., & Steif, V. (1997). Auditory long term memory: Repetition priming of voice recognition. *Quarterly Journal of Experimental Psychology*, *50A*, 498–517.

Valentine, T., Brennen, T., & Brédart, S. (1996). *The cognitive psychology of proper names*. London: Routledge.

Young, A. W., & Burton, A. M. (1999). Simulating face recognition: Implications for modelling cognition. *Cognitive Neuropsychology*, *16*, 1–48.

Young, A. W., & Ellis, H. D. (1989). Semantic processing. In A. W. Young & H. D. Ellis (Eds.), *Handbook of research on face processing* (pp. 235–262). Amsterdam: North Holland.