

PERCEPTION OF FACIAL EXPRESSIONS AND VOICES AND OF THEIR COMBINATION IN THE HUMAN BRAIN

Gilles Pourtois^{1,2,CA}, Beatrice de Gelder^{1,2}, Anne Bol³ and Marc Crommelinck²

¹Donders Laboratory for Cognitive and Affective Neuroscience, University of Tilburg, Tilburg, The Netherlands; ²Laboratory of Neurophysiology, University of Louvain, Brussels, Belgium; ³Positron Emission Tomography Laboratory, University of Louvain, Louvain-La-Neuve, Belgium)

ABSTRACT

Using positron emission tomography we explored brain regions activated during the perception of face expressions, emotional voices and combined audio-visual pairs. A convergence region situated in the left lateral temporal cortex was more activated by bimodal stimuli than by either visual only or auditory only stimuli. Separate analyses for the emotions happiness and fear revealed supplementary convergence areas situated mainly anteriorly in the left hemisphere for happy pairings and in the right hemisphere for fear pairings indicating different neuro-anatomical substrates for multisensory integration of positive versus negative emotions. Activation in the right extended amygdala was obtained for fearful faces and fearful audio-visual pairs but not for fearful voices only. These results suggest that during the multisensory perception of emotion, affective information from face and voice converge in heteromodal regions of the human brain.

Key words: face expression, multisensory integration, audio-visual perception, affective process, emotion, convergence region, PET, middle temporal gyrus, amygdala

INTRODUCTION

Perception of emotion in the voice and the face has a central function in communication and normally proceeds effortlessly and accurately. Recent studies using brain imaging have extended our knowledge about the neuro-anatomy of emotional face and voice recognition and have added support to insights obtained with neuropsychological methods (Gainotti, 2000; Adolphs, 2002). But in many natural circumstances we perceive a face and a voice expression at the same time. This situation of multisensory perception of emotion can be compared with other better known multisensory phenomena like audio-visual speech (McGurk and MacDonald, 1976) or ventriloquism (Bertelson, 1999) in which the perceptual system extracts relevant information for a given event and combines it to form one unified percept. The goal of our study was to build upon previous studies of this phenomenon (de Gelder and Vroomen, 2000; Pourtois et al., 2000; Dolan et al., 2001) and to address the question of the integrated perception of emotion in the voice and the face.

Previous studies on how emotion is perceived when provided by either the face or the voice have contributed valuable insights on processes in each modality separately. Specific areas have been found for processing the affective information provided by the face (George et al., 1993; Adolphs et al., 1996; Morris et al., 1998). Depending on whether a facial expression has a positive or a negative valence, laterality effects were observed indicating that visual stimuli with positive valence are processed predominantly in the left hemisphere and

negative ones in the right hemisphere (see Davidson and Irwin, 1999 for a review). Beyond these differences in laterality a more specific picture of separate neuro-anatomical structures for some specific facial expressions is emerging. Amygdala activates to fearful (Morris et al., 1996) and also to angry faces (Morris et al., 1998). The amygdala activation is usually less consistent for positive emotions than negative emotions (see Zald, 2003). Much less is known about the neuro-anatomy of the perception of affective expressions in the voice or affective prosody (Ross, 2000; Buchanan et al., 2000). In comparison with the processing of facial expressions, there have been only a few attempts to identify the specific neural sites for processing emotions in the voice (George et al., 1996; see Ross, 2000 for a review). Recent neuropsychological data (Adolphs et al., 2002) confirm that lower emotional prosody recognition is associated mostly with damage in the right frontal cortex but also indicate a role of the temporal polar cortex in this process.

Other studies have investigated what is usually referred to as modality independent or amodal emotional processing. In this case researchers look for common processing resources and overlapping brain structures for face and voice expressions (Borod et al., 2000; Royet et al., 2000). The issue of overlapping brain structures has mainly been addressed by looking for correlations between emotional processing in different sensory channels. For example, parallel impairments were observed in recognition of fear in the face and in the voice in patients with amygdectomy (Scott et al., 1997; but see Anderson and Phelps, 1998). Right brain

damaged patients were impaired in emotional perception tasks carried out through different communication channels (facial, prosodic and lexical). It is worth noting though that in these experiments the question whether there exists an amodal emotion processor has mainly been addressed by looking for modality independent or supramodal representations (Farah et al., 1989; Borod et al., 2000) which might mediate inter-sensory correspondences across different sensory systems in a top down fashion. How the brain combines multiple sources of affective information is not something we can learn more about by juxtaposing results obtained in studies that have looked at visual and auditory emotion perception separately. The issue of multisensory integration is altogether a different one from that concerning common processing resources or reliance on common, amodal representations.

Evidence from behavioural and electrophysiological studies clearly indicates that information in the visual modality has an impact on the subject's perception of the auditory information and vice versa during perception of emotions (de Gelder and Vroomen, 2000). When presented simultaneously with a facial expression and a sentence spoken with an affective prosody, subjective ratings of the face are biased in the direction of the accompanying voice expression (de Gelder et al., 1995; de Gelder and Vroomen 2000; Massaro and Egan, 1996). The gain in response latencies for bimodal stimuli provides evidence for the notion that perceivers integrate the two sources of information and that this integration is automatic and mandatory (Massaro and Egan, 1996; de Gelder and Vroomen, 2000). These cross-modal bias effects seem to take place at perceptual level (de Gelder et al., 1999; Pourtois et al., 2000; 2002), independent of attentional factors (Vroomen et al., 2001b) and of awareness of the visual face stimulus (de Gelder et al., 2002).

The first study that directly addressed the integration question used fMRI and revealed that a mechanism for such cross-modal binding in the case of fearful face-voice pairs could be found in the amygdala (Dolan et al., 2001). When fearful faces were accompanied by verbal messages spoken in a fearful tone of voice an increase in activation was observed in the amygdala and the fusiform gyrus. This result indicated a combination of face and voice expressions, confirming the role of the amygdala in this process (see Goulet and Murray, 2001 for a critical review). The present PET study extends the research carried out by Dolan and collaborators (2001) in two directions. First, we introduced single modality conditions in order to obtain insight into the difference between each single modality separately and the combination. On this issue we had three *a priori* predictions. The perception of face-voice emotion pairs would yield activation in brain regions known to be heteromodal (like the MTG,

BA21; Damasio, 1989; Mesulam, 1998). This activation would be specific for each emotion pair. Two standardised opposite emotions (fear and happiness) were selected to test this prediction. And finally, activation in heteromodal areas would also be accompanied by increased activation in modality-specific cortices (like the primary auditory cortex or primary visual cortex). The latter phenomenon has been reported previously and can tentatively be viewed as a downstream consequence in modality-specific cortex of multisensory integration (see Calvert et al., 2000; Macaluso et al., 2000; de Gelder, 2000; Driver and Spence, 2000; Dolan et al., 2001). Such feedback or top down modulations could be the correlate of the well-known cross-modal bias effects typically observed in behavioural studies of audio-visual perception (Bertelson, 1999).

Moreover, we opted for an indirect or covert processing task in which the participants are not consciously attending to the emotional meaning of the stimuli, namely a gender decision task. This task should reveal attention independent activation and integration patterns. This issue is important in the light of findings that indicate differential activation patterns as a function of conscious perception of the emotional stimuli (Morris et al., 2001). Modulation by stimulus awareness and/or attentional demands does not imply that attention is not itself at the basis of inter-sensory integration (Bertelson et al., 2000; Vroomen et al., 2001a; McDonald et al., 2001 for a discussion). In fact our results indicate it is not (Vroomen et al., 2001b). The gender decision task used here does not require attention to the emotion expressed in the stimuli and does not require participants to be aware of experiencing an emotion upon stimulus presentation. The gender task has the added advantage of presenting the participants with a situation that is transparent and therefore does not elicit responses that would be under strategic control (see Bertelson and De Gelder, in press, for discussion). With respect to this issue, our *a priori* prediction was that the use of a covert task would still provide us with evidence for audio-visual integration during the perception of emotion.

METHODS

Subjects

Eight right-handed male subjects (age range: 21-23 yrs) participated in the Experiment after written consent was obtained according to institutional guidelines. They were paid for their participation.

Materials

Materials consisted of visual and auditory stimuli and of audio-visual pairs (Figure 1). Visual stimuli consisted of twelve grey-scale full frontal

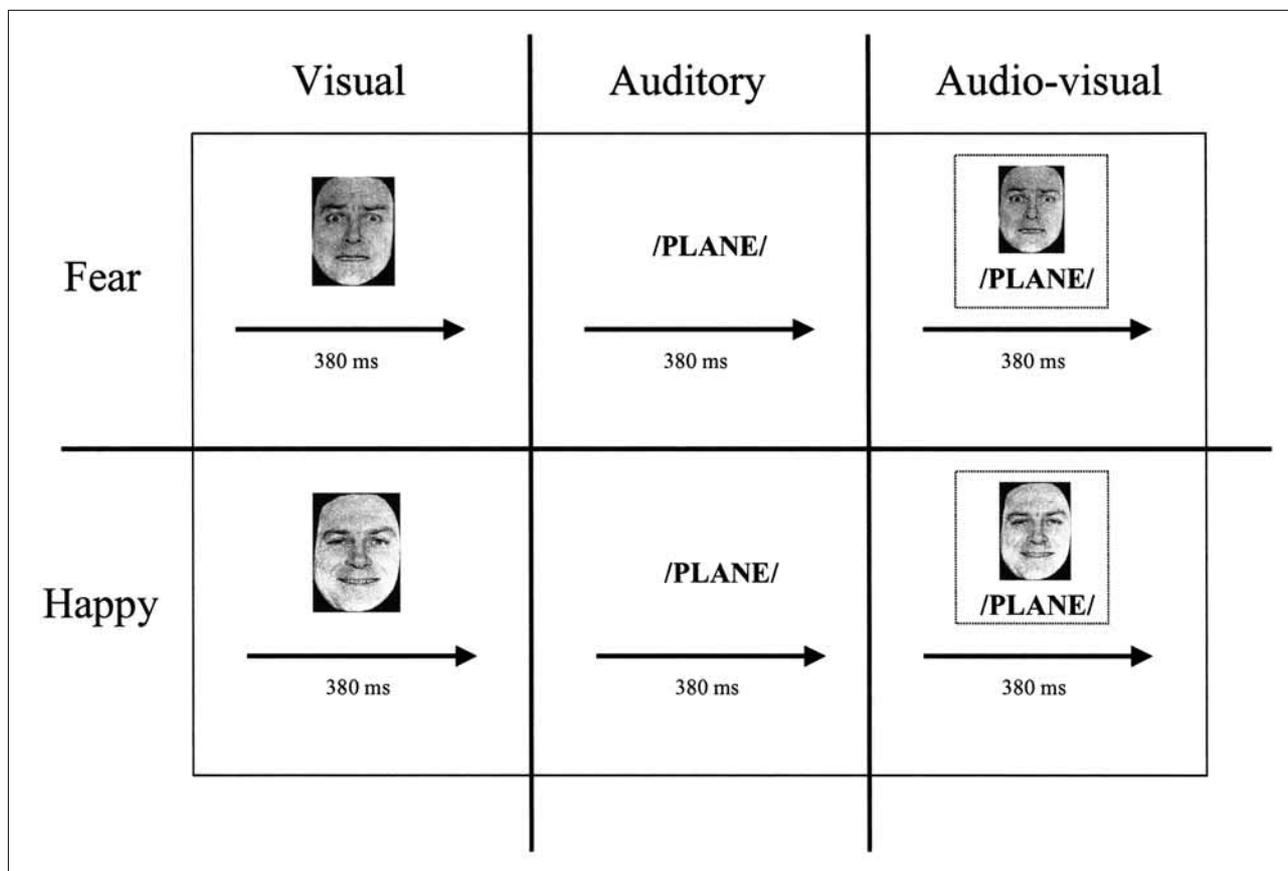


Fig. 1 – Experimental design consisting of two emotion (fear or happy), and three conditions (visual, auditory and audio-visual) with stimulus duration (380 ms) for each trial.

view pictures of 3 males and 3 females presenting a happy or a fearful facial expression (Ekman and Friesen, 1976). Stimulus size was 6-cm width \times 8-cm height. Images were presented on a Macintosh AV17 screen against a black background and viewed at 120-cm distance.

Auditory stimuli were twelve bi-syllabic words spoken by 3 male and 3 female speakers in a happy or a fearful tone of voice. To obtain these materials 12 semi-professional actors (6 males and 6 females) were asked to pronounce a neutral sentence (the French version of /they are travelling by plane/) in a happy or in a fearful tone of voice. They were instructed to pronounce the sentence as when experiencing happy or fearful feelings in real-life situations. Two utterances were recorded for each expression type. Recording was done in an acoustic-isolated room using digital audiotape and sounds were digitised (SoundEdit 16 version 2 running on Macintosh). Only the last bi-syllabic word (/plane/) was selected for each of the 24 productions, and amplified (SoundEdit 16 version 2 running on Macintosh).

In a pilot study, the 24 fragments (12 actors \times 2 tones of voice) were randomly presented to 8 naïve subjects. They were asked to discriminate accurately the tone of voice (either happy or fearful) of each of the 24 productions presented five times. The three best-recognized female and

male speaker fragments were selected for use in the PET Experiment. Mean recognition rate for the three selected female actors was 94% correct, and 95% correct for the male ones. Mean duration was 384 msec (SD: 40 msec) for the 6 happy fragments and 376 msec (SD: 61 msec) for the 6 fear fragments. The mean sound level of the speech as measured at loudspeaker was 75 dB. In the scanner, sounds were directly presented to each ear using earphones. Audio-visual pairs were made by combining each of the 6 voices with one of the face expressions on an arbitrary basis but respecting congruency of gender and of emotion.

Procedure

There were six conditions (Figure 1), each repeated twice: for each emotion (fear and happiness), stimuli consisted of facial expressions (2 scans), auditory fragments (2 scans) or the combination of both (2 scans). Subjects always performed a gender decision task unrelated to the variables manipulated (e.g., Morris et al., 1998). When audio-visual trials were presented (4 scans), subjects were instructed to perform the gender decision task either on the visual channel (2 scans) or on the auditory channel (2 scans). Resting scans were presented at the beginning (first scan) and the end (14th scan) of the Experiment. The block order

of visual (2 emotions \times 2 repetitions), auditory (2 emotions \times 2 repetitions) and audio-visual (2 emotions \times 2 repetitions) conditions was counterbalanced across subjects.

The beginning of a trial was signalled by a small white cross that remained in the centre of the screen for 400 msec. Then, the target stimulus (visual only, auditory only or audio-visual) was presented lasting for about 380 msec. In the audio-visual condition, onset and offset of the face and the sound fragment were synchronized. Interval was set at 1500 ms (black screen) and response was recorded till 1800 msec following onset of the stimulus. Fifty-four trials (6 actors \times 9 repetitions) were presented during a scan. Each scan contained the same proportion (50 % – 27 trials) of male and female stimuli. Before each scan, a block of 20 trials was presented. Subjects responded with the right hand using a 2 buttons-response box.

Data Acquisition

Accuracy and reaction times (RTs) were recorded (Superlab Pro 1.74 running on a PowerPC Macintosh) for each condition and each subject. Each subject was scanned twice in each condition and received intravenous H_2^{15}O (8 mCi, 2.96 e + 02 MBq, 20 s bolus) 10 seconds before starting the task. Measurements of local radioactivity uptake were made using an ECAT EXACT-HR PET tomograph (CTI/Siemens), allowing simultaneous imaging of 47 transaxial slices in three-dimensional (3-D, septa retracted) mode, with an effective resolution of 8 mm full width-at-half-maximum (FWHM) (Wienhard et al., 1994) and a slice thickness of 3.125 mm. Images were reconstructed using filtered back-projection scatter correction, with both transaxial Hanning filter (cutoff frequency of 0.30) and axial Hanning filter (cutoff frequency of 0.50). A two-dimensional transmission scan was acquired before the Experiment for attenuation correction. For each scan, the task started 10 sec after initiation of tracer injection and PET data was acquired simultaneously in a single 100s frame. The integrated counts accumulated during 100 sec were used as an index of rCBF (Mazziotta and Phelps, 1986). The time interval between successive emission scans was 13 minutes, which allowed decay of residual radioactivity. For each subject, 3-D MRI (T1) anatomical data were also obtained on a 1.5 Tesla unit (General Electric Signa).

Data Analysis

PET images were realigned to the first one (Woods et al., 1998a, 1998b) and then to the MRI, spatially normalized (SPM96-<http://www.fil.ion.ucl.ac.uk/spm/>) to a stereotactic space (Talairach and Tournoux, 1988), smoothed with a Gaussian filter (15 mm FWHM), corrected

for global activity by proportional scaling (Fox et al., 1988) and adjusted to a level of 50ml/g/min. Group statistical maps were made using the general linear model (Friston et al., 1995) in SPM99. Main effects and interactions were assessed with different contrasts using t tests. Three conjunction analyses (Price and Friston, 1997) were performed at the group level using SPM99. A global conjunction analysis for the two emotions together (happy and fear) was first performed [(AV-V) and (AV-A)] and followed by two specific conjunction analyses, one for each emotion separately [(AV_{happy}-V_{happy}) and (AV_{happy}-A_{happy}) in the happy condition and (AV_{fear}-V_{fear}) and (AV_{fear}-A_{fear}) in the fear condition]. Conjunction analyses were used to remove unimodal activations (visual and auditory) and isolate brain activations specific to the processing of audio-visual events. To compute the conjunction analyses, the two simple contrasts were first calculated, and entered in a subsequent stage in the conjunction analysis implemented in SPM99. As required by the conjunction analysis (Price and Friston, 1997), four independent conditions (Visual, Auditory, Audio-visual with gender decision task on the visual channel and Audio-visual with gender decision task on the auditory channel) were used in the statistical analysis to compute the conjunction contrast. Moreover, the [AV – (A + V)] contrast was also calculated to test for the presence of a supra-additivity response enhancement in specific brain regions (i.e., more activation for audio-visual presentations compared to the sum of visual and auditory presentations).

Corrected P values at the voxel level were computed using a spherical Volume of Interest (VOI) with 15 mm radius, centered over the coordinates obtained with SPM 99 (Worsley et al., 1996). These coordinates are similar or close to the centre of activation found in previous studies focussed on face processing (activation in the posterior fusiform gyrus and inferior occipital gyrus, see Hoffman and Haxby, 2000), voice processing (activation in the superior temporal gyrus, see Belin et al., 2000), audio-visual processing (activation in the middle temporal gyrus, see Mesulam, 1998 for a review) and emotional processing (activation in the superior and middle frontal gyrus, see Royet et al., 2000; activation in the amygdala, see Morris et al., 1999a).

RESULTS

Behavioural Data

Correct responses were above 95% for all conditions (98.3%, 96.9% and 96.2% for the audio-visual, visual and auditory condition respectively). A two-way ANOVA (Emotion and Modality) for repeated measures on accuracy rates did not show any significant effect [Emotion: $F(1, 7) = 4.62$,

$p = .07$; Modality: $F(2, 14) = 1.5$, $p = .27$; interaction Emotion \times Modality: $F < 1$). Reaction times were slower for gender in the voices (mean RTs: 553 msec) than in the faces (mean RTs: 493 msec) and in the audio-visual pairs (mean RTs: 501 msec). A two-way ANOVA (Emotion and Modality) showed a significant main effect of Modality [$F(2, 14) = 8.63$; $p < .005$] (all other F s < 1). *Post-hoc* paired t -tests revealed that the difference between visual and auditory conditions was significant [$t_{(7)} = 3.83$, $p = .002$], as well as the difference between auditory and audio-visual conditions [$t_{(7)} = 3.31$, $p = .005$] indicating that subjects were slower in the auditory conditions than in the visual or audio-visual conditions (Ellis et al., 1997). The difference between visual and audio-visual conditions was not significant [$t_{(7)} = 0.53$, $p > 0.5$].

PET Data

Modality-specific brain areas were first identified using the (Faces-Voices) contrast as well as the reverse contrast. Results of these control analyses were first described in order to assess whether the activations related to the perception of audio-visual information and presented subsequently were specific or not.

Modality-specific visual brain areas were obtained by subtracting activation in the Voice only condition from those obtained for the Face only

condition. This contrast revealed bilateral activations in the *posterior lateral fusiform region* (BA19), left *inferior occipital gyrus* (BA18), right *lingual gyrus* (BA18) and right *parahippocampal gyrus* (BA36) (see Table I). This result is consistent with the brain areas known to be associated with face processing (Sergent et al., 1992; the Fusiform Face Area see Kanwisher et al., 1997; inferior occipital and fusiform gyri, see Hoffman and Haxby, 2000). Likewise, when auditory conditions were subtracted from audio-visual conditions (Figure 2), bilateral activation of the *fusiform gyrus* (BA19), the left *lingual gyrus* (BA18) and the left *superior occipital gyrus* (BA19) were manifest (see Table I).

Modality specific auditory activations were obtained by subtracting visual from auditory conditions and indicated bilateral activation in the *superior temporal gyrus* (BA22), an area of the auditory cortex involved in the perception of speech and voice (Liegeois-Chauvel et al., 1999; Belin et al., 2000). Activation was also present in another region of the *superior temporal gyrus* in the left hemisphere, extending towards the *MTG* (BA22). Similarly when visual conditions were subtracted from audio-visual conditions (Figure 3) the comparison revealed bilateral activation in the *superior temporal gyrus* (BA22) and in a more anterior region within the same gyrus in the left hemisphere.

TABLE I

Brain Regions (IOG = inferior occipital gyrus, PFG = posterior fusiform gyrus, LG = lingual gyrus, PHG = para-hippocampal gyrus, FG = fusiform gyrus, SOG = superior occipital gyrus, STG = superior temporal gyrus, MTG = middle temporal gyrus, AFG = anterior fusiform gyrus, SFG = superior frontal gyrus, MFG = middle frontal gyrus, IPL = inferior parietal lobule, M = mesencephalon) with significant rCBF increases when comparing the rCBF images obtained in the 6 conditions (V = Visual, A = Auditory, AV = Audio-Visual, Vh = Visual happy, Vf = Visual fear, Ah = Auditory happy, Af = Auditory fear, AVh = Audio-Visual happy, AVf = Audio-Visual fear)

Contrasts	Cluster size	x, y, z	L/R	BA	Brain regions	T value	Z score	P value uncorrected	P value corrected*	
V - A	1880	- 26, - 80, 0	L	18	IOG	6.45	5.88	< .001	< .001	
		- 32, - 58, - 14	L	19	PFG	5	4.17	< .001	< .001	
	1246	32, - 76, - 6	R	18	LG	5.43	5.06	< .001	< .001	
		32, - 66, - 10	R	19	PFG	5.36	5.01	< .001	< .001	
AV - A	544	28, - 46, - 6	R	36	PHG	4.47	4.26	< .001	.001	
		34, - 64, - 12	R	19	FG	4.96	4.68	< .001	< .001	
	290	- 22, - 80, 0	L	18	LG	4.56	4.33	< .001	.001	
		- 32, - 68, - 8	L	19	FG	3.42	3.32	< .001	.019	
A - V	15	- 32, - 64, 32	L	19	SOG	3.42	3.32	< .001	.019	
		- 58, - 16, 2	L	22	STG	9.91	> 8	< .001	< .001	
	2400	- 50, 12, - 6	L	22	STG-MTG	3.71	3.58	< .001	.009	
		58, - 18, 2	R	22	STG	6.87	6.19	< .001	< .001	
AV - V	1454	- 58, - 18, 2	L	22	STG	7.36	6.55	< .001	< .001	
		- 48, - 34, 10	L	22	STG	4.31	4.12	< .001	.001	
	837	58, - 12, 0	R	22	STG	5.78	5.35	< .001	< .001	
		(AV - A) & (AV - V)	50	- 52, - 30, - 12	L	21	MTG	2.34	3.68	< .001
(AVh - Ah) & (AVh - Vh)	5	- 52, - 38, - 26	L	20	AFG	1.95	3.19	.001	.042	
		- 34, 30, 50	L	8	SFG	2.3	3.64	< .001	.015	
	207	- 54, 30, 22	L	46	MFG	2.17	3.47	< .001	.024	
		- 52, 26, 34	L	9	MFG	2.06	3.33	< .001	.036	
	50	- 36, - 50, 40	L	40	IPL	2.25	3.57	< .001	.018	
		148	- 34, 60, 8	L	10	MFG	2.24	3.56	< .001	.018
	42	- 40, 50, 6	L	10	MFG	2.06	3.32	< .001	.037	
		6, - 32, - 18	R		M	1.96	3.2	.001	.052	
	(AVf - Af) & (AVf - Vf)	41	48, - 42, 6	R	21	MTG	1.89	3.11	.001	.065
	(AVf + Vf) - (AVh + Vh)	94	10, - 8, - 10	R		Amygdala	2.55	2.51	.006	.036

* Corrected P values at the voxel level were computed using a small volume correction approach (see Methods).

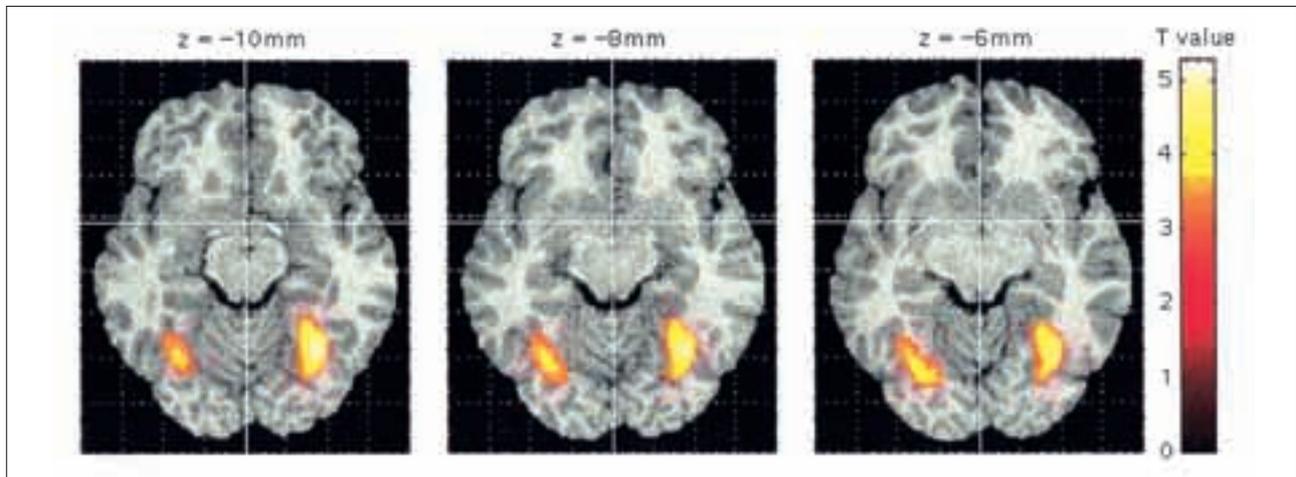


Fig. 2 – Statistical parametric maps ($p < 0.001$) showing significant activation bilaterally in the fusiform gyrus during the perception of facial expressions as revealed by the (AV - A) contrast.

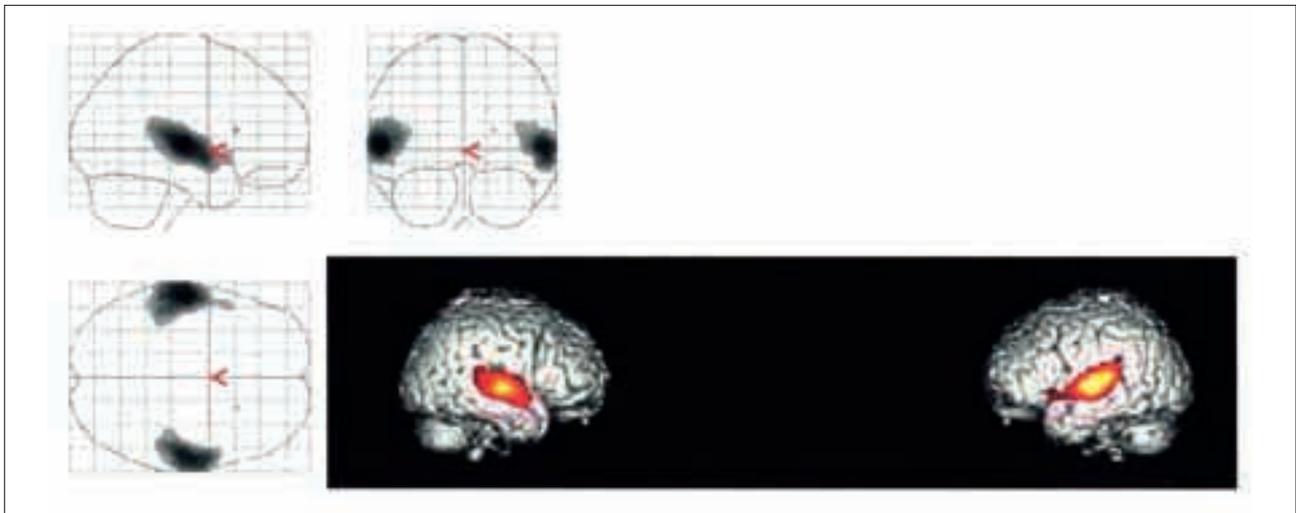


Fig. 3 – Statistical parametric maps ($p < 0.001$) showing significant activation bilaterally in the superior temporal gyrus during the perception of voice expressions as revealed by the (AV - V) contrast.

Next, we carried out a conjunction analysis (see Price and Friston, 1997) in order to determine which brain regions were associated with the perception of audio-visual information. The global conjunction (AV-V) and (AV-A) for the two emotions (Table I and Figure 4) revealed a main region of audio-visual convergence in the *MTG* (BA21) in the left hemisphere as well as a smaller activation, in the *anterior fusiform gyrus* (BA20) in the same hemisphere.

Interestingly, when the [AV - (A + V)] contrast was calculated in order to obtain brain activations specific to the perception of audio-visual events, the analysis confirmed activation in the *MTG* in the left hemisphere ($-52x, -30y, -12z$; $t = 3.23$, z score = 3.14, $p_{\text{uncorrected}} = .001$).

A detailed statistical analysis of the corrected rCBF values in the maxima of activation in the conjunction contrast in the left *MTG* (Figure 5) indicated that this region was significantly more activated by audio-visual stimuli than unimodal

stimuli whether auditory or visual stimuli regardless of the emotion presented. Corrected rCBF values (ml/g/min) were extracted in spherical regions of 3 mm radius centered around the maxima of activation in the left *MTG* ($-52x, -30y, -12z$). A two-way ANOVA (Emotion and Modality) for repeated measurements on these values showed a significant effect of *Modality* [$F(2, 14) = 7.1$, $p = .007$] but no other effects [factor Emotion: $F(1, 7) = 3.58$, $p = .1$; interaction Modality \times Emotion: $F < 1$]. *Post-hoc* t -tests failed to show a difference between visual and auditory levels of activation [$t_{(7)} = 0.95$, $p = .36$], but there was a significant increase in activation for the audio-visual condition as compared to auditory [$t_{(7)} = 2.68$, $p = .02$] and visual [$t_{(7)} = 3.63$, $p = .003$] conditions.

Two further specific conjunction analyses were carried out separately for happiness and fear in order to investigate whether these two different emotions yield specific activity in the two audio-visual conditions (Table I). The conjunction analysis

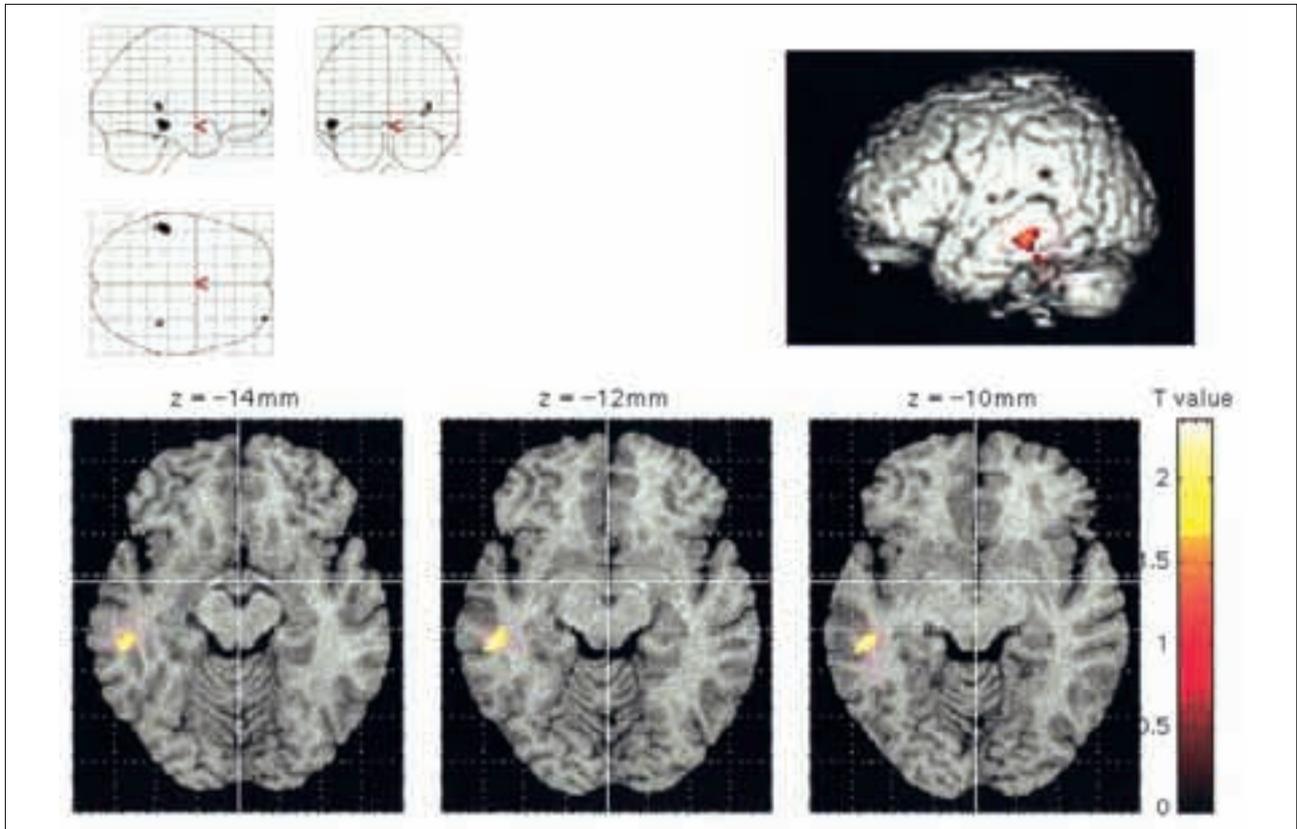


Fig. 4 – Statistical parametric maps ($p < 0.001$) showing significant activation in the left MTG ($-52x, -30y, -12z$) during the perception of audio-visual trials (happy and fearful emotions) when compared to unimodal trials (Visual + Auditory).

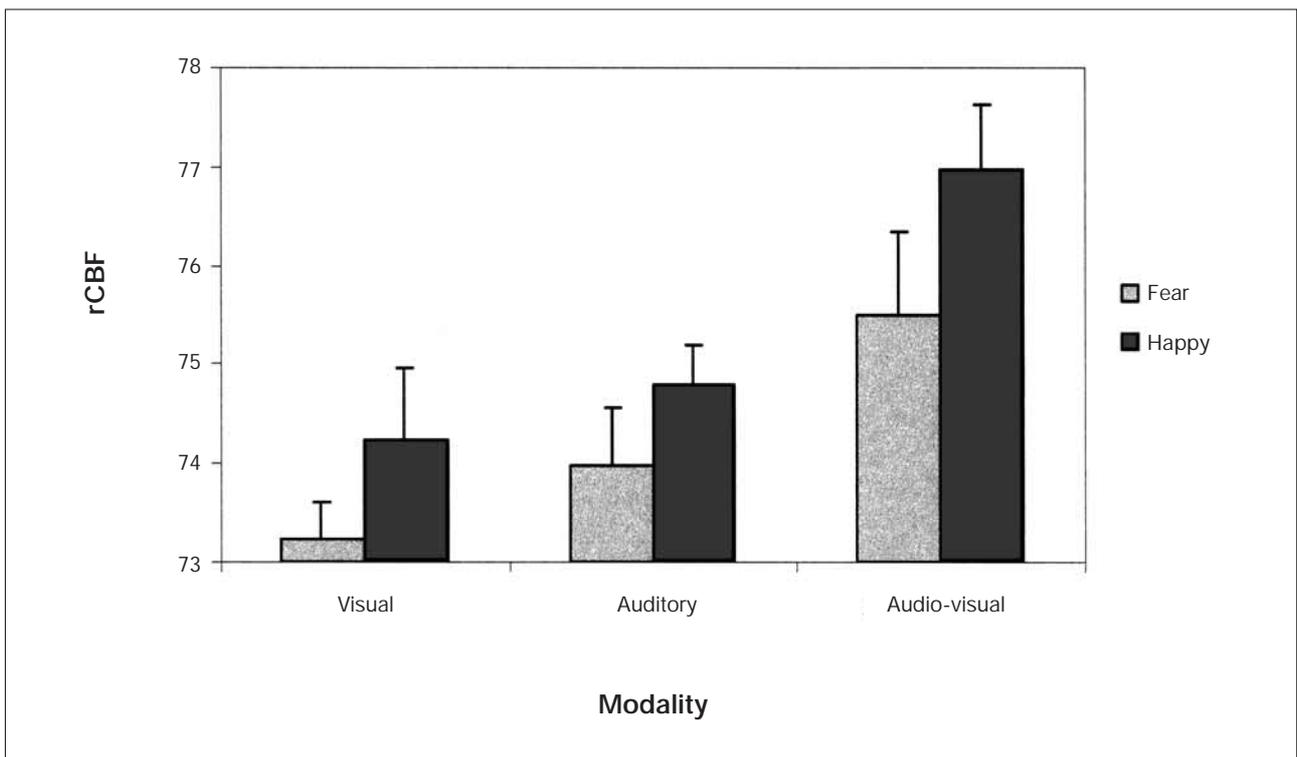


Fig. 5 – Corrected rCBF values (ml/g/min) extracted in spherical regions of 3mm radius centered around the maxima of activation in the left MTG ($-52x, -30y, -12z$) for the 6 conditions. There was no difference between visual and auditory levels of activation but there was a significant increase in activation for the audio-visual condition as compared to auditory and visual conditions. Vertical lines indicate standard error.

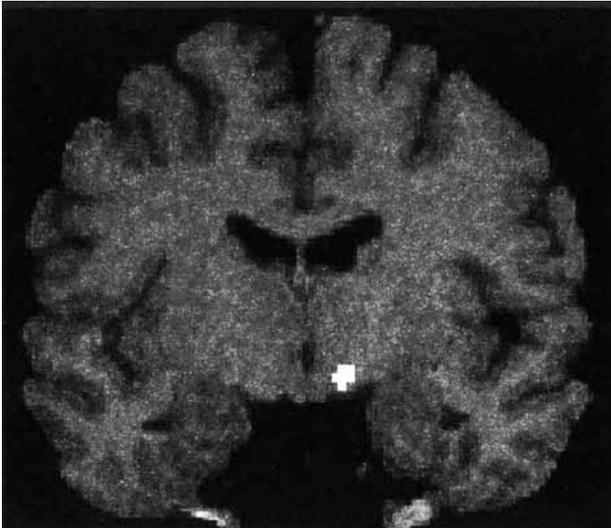


Fig. 6 – Statistical parametric map ($p < 0.001$) showing significant activation in the right extended amygdala ($-10x, -8y, -10z$) during the perception of fearful expressions (Visual + Audio-visual) when compared to happy expressions (Visual + Audio-visual).

for happiness revealed several supplementary frontal brain regions mainly lateralized in the left hemisphere (Table I): in the *superior frontal gyrus* (BA8), the *middle frontal gyrus* (BA46, 10 and 9) and the *inferior parietal lobule*. Moreover, a region of the mesencephalon was activated by audio-visual happiness. The conjunction analysis for fear only revealed activation in a posterior region of the *MTG* (BA21).

A region of interest analysis centered on the anterior temporal lobe revealed significant activation in the extended amygdala in the right hemisphere (Figure 6) for fearful faces (visual and audio-visual conditions) when compared to happy faces ($10x, -8y, -10z$; cluster size = 94; $t = 2.55$, $z = 2.51$, $p_{\text{uncorrected}} = .006$; $p_{\text{corrected}} = .036$). The anatomical localisation of this activation corresponded to the extended amygdala, namely the sublenticular region or basal forebrain region. There was no evidence of a rCBF increase in this region for fearful voices when compared to happy voices.

Finally, fear conditions (auditory, visual and audio-visual) and happy conditions (auditory, visual and audio-visual) were contrasted irrespective of modality. In the Fear-Happiness contrast, the analysis showed significant activation in two posterior regions of the brain: in the posterior part of the middle temporal gyrus in the right hemisphere (BA39; $32x, -70y, 28z$; cluster size = 86; $t = 4.18$, $z = 4.0$, $p_{\text{uncorrected}} < .001$) and in a portion of the precuneus in the left hemisphere (BA19; $-14x, -64y, 42z$; cluster size = 26; $t = 3.53$, $z = 3.42$, $p_{\text{uncorrected}} < .001$). The reverse contrast showed activation in the middle frontal gyrus in the right hemisphere (BA10; $28x, 50y, -4z$; cluster size = 70; $t = 3.04$, $z = 2.97$, $p_{\text{uncorrected}} = .001$) and in the middle frontal gyrus

in the left hemisphere (BA10; $-40x, 50y, -8z$; cluster size = 57; $t = 2.76$, $z = 2.7$, $p_{\text{uncorrected}} = .003$). In both cases, the activation was larger in the right hemisphere than in the left hemisphere.

DISCUSSION

Our results indicate that the perception of audio-visual emotions (fear and happiness) activates the left MTG relative to unimodal conditions and to a lesser extent also the left anterior fusiform gyrus. The MTG has already been shown to be involved in multisensory integration (Streicher and Ettliger, 1987) and has been described as a convergence region between multiple modalities (Damasio, 1989; Mesulam, 1998). Brain damage to the temporal polar cortex impairs emotional recognition of both facial and vocal stimuli (Adolphs et al., 2002). Activation of the fusiform gyrus is consistent with results from the fMRI study of Dolan and collaborators (2001) of recognition of facial expression paired with emotional voices. Interestingly, our results show that within the left MTG (BA 21) there was no difference between visual and auditory levels of activation but instead there was a significant increase for the audio-visual condition as compared to the two unimodal conditions (supra-additive response enhancement criterion, Calvert et al., 2000). This clearly indicates that the activation in the left MTG does not correspond to an augmentation of activity in terms of rCBF increase in regions that are modality-specific. This was also demonstrated by a failure to observe the same MTG activation with the two direct contrasts, (A – V) and (V – A) each revealing the corresponding modality specific activations. Recently, such an increased activation in modality specific areas presumably following upon audio-visual integration has been reported. This was found for visuo-tactile pairs in a visual area (the lingual gyrus, Macaluso et al., 2000), in auditory areas for audio-visual pairs (Calvert et al., 1999) and in the fusiform gyrus for audio-visual emotion pairs (Dolan et al., 2001). But it is not clear to what extent this activation pattern might be task related as all these studies required attention to the task relevant modality.

The activation for bimodal pairs relative to unimodal conditions is observed within the MTG in the left hemisphere. This cortical region is not the area most often associated with perception of affective prosody in speech. A recent fMRI study (Buchanan et al., 2000) confirmed a preferential involvement of the right hemisphere for emotional prosody (Van Lancker et al., 1989; George et al., 1996; Adolphs et al., 2002; Vingerhoets et al., 2003). The recognition of emotion in the voice compared with purely linguistic task yielded more activation in the right hemisphere than in the left

hemisphere (in particular in the right inferior frontal lobe, Pihan et al., 1997; Imaizumi et al., 1998). Still right hemispheric dominance for emotional prosody is not fully established yet as affective prosody is also disturbed after left brain damage (Ross et al., 2001). It is also worth noting that all these studies used an explicit emotion recognition task and not a gender classification task as used here. Auditory presentation results in left-lateralized activity in the superior temporal gyrus and the MTG (Price et al., 1996; Binder et al., 1997). Yet our activations are not due to word recognition because activation is stronger for audio-visual pairs than for auditory only stimuli. Such left MTG activation associated with audio-visual processes was obtained in a previous fMRI study of audio-visual speech (see Calvert et al., 2000 for a review). But we argue that this convergence between findings about the role of MTG in the two studies does not mean that audio-visual emotion integration is just another instance of audio-visual speech integration and that by implication the emotional content of the stimuli was either not processed or does not matter. Indeed, here we found different lateralized activations for the two types of emotion pairs. This indicates clearly that subjects did indeed process the emotional content and underscores that the MTG activation is not simply related to linguistic processes. A direct comparison between audio-visual emotion perception and audio-visual speech perception (or at least audio-visual perception of non-emotional stimuli) should help to better understand the exact role of convergence sites within the MTG. Likewise, further research is needed to determine whether the manipulation of other basic emotions (sadness, anger, disgust and surprise) presented in different sensory modalities produce the same pattern of activations.

Our data are compatible with previous PET studies (e.g., Royet et al., 2000) showing that pleasant and unpleasant emotional stimuli from either the visual or the auditory modality recruit the same cortical network (i.e., orbitofrontal cortex, temporal pole and superior frontal gyrus), mainly lateralized in the left hemisphere. Our study adds crucial information about the cerebral network involved in audio-visual perception of specific emotions (happiness and fear) since supplementary activations are also observed separately for the two emotions when visual and auditory affective information are concurrently presented. Activation for the combination of happy face and happy voice is found in different frontal and prefrontal regions (BA 8, 9, 10 and 46) that are lateralized in the left hemisphere while audio-visual fear activates the superior temporal gyrus in the right hemisphere. These data are therefore compatible with current theories of emotional processes (see Davidson and Irwin, 1999 for a review) suggesting hemispheric

asymmetries in the processing of positively (pleasant) versus negatively (unpleasant) valenced stimuli. However, these hemispheric asymmetries are found with an analysis contrasting bimodal vs. unimodal conditions and a simple valence explanation (left hemisphere for happiness vs. right hemisphere for fear) cannot readily account for them. This conclusion is reinforced by the fact that irrespective of modality the happiness-fear contrast shows activation in a portion of the middle frontal gyrus in both cerebral hemispheres. The reverse contrast shows activation in posterior regions (precuneus and MTG) in both cerebral hemispheres as well. In both cases, the activation is larger in the right than in the left hemisphere. However, since we did not use a neutral condition, these data cannot be interpreted as reflecting a right hemispheric dominance for amodal emotional processing (Borod et al., 2000). Our results suggest rather that the left vs. right hemispheric activation within the MTG for happy vs. fearful audio-visual pairs found here might be specific to bimodal conditions.

Finally, activation in the right extended amygdala to fearful faces is compatible with a previous report of right lateralized amygdala activation where the same gender decision task was used as here (Morris et al., 1996), when facial expressions were masked (Morris et al., 1998) or when face expressions were processed entirely outside awareness (Morris et al., 2001). Processing routes in these studies included, besides the amygdala, the superior colliculus and the pulvinar (LeDoux, 1996; Morris et al., 1997, 1999a; Whalen et al., 1998;). There were no indications of similar involvement of the amygdala in discrimination of fearful voice expressions as suggested in a study of amygdectomy patients by one group (Scott et al., 1997) but not by more recent studies (Anderson and Phelps, 1998; Adolphs and Tranel, 1999a). Our results are consistent with the latter negative findings and indicate that the amygdala might not be the locus of amodal fear processing. On the other hand, the amygdala is a complex nucleus but still it may only be involved in processing the expression from the face (Adolphs and Tranel, 1999b). Finally, activation of the amygdala is consistent with neuropsychological and neurophysiological studies (Murray and Mishkin, 1985; Nahm et al., 1993; Murray and Gaffan, 1994) that have indicated a key integrative role for this structure in the association and retention of inter-modal stimulus pairs (i.e., visuo-tactile). This was directly confirmed by the study of Dolan et al. (2001).

CONCLUSIONS

Our study clearly indicates the involvement of the MTG in the audio-visual perception of

emotions relative to unimodal visual and auditory emotion conditions. The analysis of the rCBF increase in the left MTG revealed a significant increase for the audio-visual condition as compared to the two unimodal conditions while there was no difference between visual and auditory levels of activation. Moreover, the observed activations are directly related to emotional valence with happy audio-visual pairs activating structures that are mainly lateralized in the left hemisphere whereas fearful audio-visual pairs activated structures lateralized in the right hemisphere. Our results also confirm and extend earlier studies by indicating that audio-visual perception of emotion proceeds covertly in the sense that it does not depend on explicit recognition of the emotions.

ABBREVIATIONS

PET = positron emission tomography; rCBF = regional cerebral blood flow; SPM = statistical parametric map; MTG = middle temporal gyrus.

REFERENCES

- ADOLPHS R. Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, 12: 169-177, 2002.
- ADOLPHS R, DAMASIO H and TRANEL D. Neural Systems for Recognition of Emotional Prosody: A 3-D Lesion Study. *Emotion*, 2: 23-51, 2002.
- ADOLPHS R, DAMASIO H, TRANEL D and DAMASIO AR. Cortical systems for the recognition of emotion in facial expressions. *Journal of Neuroscience*, 16: 7678-7687, 1996.
- ADOLPHS R and TRANEL D. Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia*, 37: 1285-1292, 1999a.
- ADOLPHS R and TRANEL D. Preferences for visual stimuli following amygdala damage. *Journal of Cognitive Neuroscience*, 11: 610-616, 1999b.
- ANDERSON AK and PHELPS EA. Intact recognition of vocal expressions of fear following bilateral lesions of the human amygdala. *Neuroreport*, 9: 3607-3613, 1998.
- BELIN P, ZATORRE RJ, LAFAILLE P, AHAD P and PIKE B. Voice-selective areas in human auditory cortex. *Nature*, 403: 309-312, 2000.
- BERTELSON P. Ventriloquism: A case of crossmodal perceptual grouping. In G Aschersleben, T Bachman and J Musseler (Eds), *Cognitive Contributions to the Perception of Spatial and Temporal Events*. Amsterdam: Elsevier Science, 1999, pp. 347-369.
- BERTELSON P, VROOMEN J, DE GELDER B and DRIVER J. The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception and Psychophysics*, 62: 321-332, 2000.
- BERTELSON P and DE GELDER B. *The Psychology of Multi-modal Perception*. J Driver and C Spence (Eds). London: Oxford University Press, in press.
- BINDER JR, FROST JA, HAMMEKE TA, COX RW, RAO SM and PRIETO T. Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, 17: 353-362, 1997.
- BOROD JC, PICK LH, HALL S, SLIWINSKI M, MADIGAN N, OBLER LK, WELKOWITZ J, CANINO E, ERHAN HM, GORAL M, MORRISON C and TABERT M. Relationships among facial, prosodic, and lexical channels of emotional perceptual processing. *Cognition and Emotion*, 14: 193-211, 2000.
- BUCHANAN TW, LUTZ K, MIRZAZADE S, SPECHT K, SHAH NJ, ZILLES K and JANCKE L. Recognition of emotional prosody and verbal components of spoken language: An fMRI study. *Cognitive Brain Research*, 9: 227-238, 2000.
- CALVERT GA, BRAMMER MJ, BULLMORE ET, CAMPBELL R, IVERSEN SD and DAVID AS. Response amplification in sensory-specific cortices during crossmodal binding. *Neuroreport*, 10: 2619-2623, 1999.
- CALVERT GA, CAMPBELL R and BRAMMER MJ. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, 10: 649-657, 2000.
- DAMASIO AR. Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33: 25-62, 1989.
- DAVIDSON RJ and IRWIN W. The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences*, 3: 11-21, 1999.
- DE GELDER B. Crossmodal perception: There is more to seeing than meets the eye. *Science*, 289: 1148-1149, 2000.
- DE GELDER B, BÖCKER KBE, TUOMAINEN J, HENSEN M and VROOMEN J. The combined perception of emotion from voice and face: Early interaction revealed by human electric brain responses. *Neuroscience Letters*, 260: 133-136, 1999.
- DE GELDER B, POURTOIS G and WEISKRANTZ L. Fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures. *Proceedings of the National Academy of Sciences of the United States of America*, 99: 4121-4126, 2002.
- DE GELDER B and VROOMEN J. Perceiving Emotions by Ear and by Eye. *Cognition and Emotion*, 14: 289-311, 2000.
- DE GELDER B, VROOMEN J and TEUNISSE JP. Hearing smiles and seeing cries: The bimodal perception of emotions. *Bulletin of the Psychonomic Society*, 30: 15, 1995.
- DOLAN RJ, MORRIS JS and DE GELDER B. Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Sciences of the United States of America*, 98: 10006-10010, 2001.
- DRIVER J and SPENCE C. Multisensory perception: Beyond modularity and convergence. *Current Biology*, 10: 731-735, 2000.
- EKMAN P and FRIESEN WV. *Pictures of facial affect*. Palo-Alto: Consulting Psychologists Press, 1976.
- ELLIS HD, JONES DM and MOSDELL N. Intra- and inter-modal repetition priming of familiar faces and voices. *British Journal of Psychology*, 88: 143-156, 1997.
- FARAH MJ, WONG AB, MONHEIT MA and MORROW LA. Parietal lobe mechanisms of spatial attention: Modality-specific or supramodal? *Neuropsychologia*, 27: 461-470, 1989.
- FOX PT, MINTUN MA, REIMAN EM and RAICHLER ME. Enhanced detection of focal brain response using intersubject averaging and change-distribution analysis of subtracted PET images. *Journal of Cerebral Blood Flow and Metabolism*, 8: 642-653, 1988.
- FRISTON KJ, HOLMES AP, WORSLEY KJ, POLINE JP, FRITH CD and FRACKOWIACK RJS. Statistical Parametric Maps in Functional Imaging: A General Linear Approach. *Human Brain Mapping*, 2: 189-210, 1995.
- GAINOTTI G. Neuropsychological theories of emotion. In JC Borod (Ed), *The Neuropsychology of Emotion*. Oxford: Oxford University Press, 2000, pp. 214-236.
- GEORGE MS, KETTER TA, GILL DS, HAXBY JV, UNGERLEIDER LG, HERSCOVITCH P and POST RM. Brain regions involved in recognizing facial emotion or identity: An oxygen-15 PET study. *Journal of Neuropsychiatry and Clinical Neurosciences*, 5: 384-394, 1993.
- GEORGE MS, PAREKH PI, ROSINSKY N, KETTER TA, KIMBRELL TA, HEILMAN KM, HERSCOVITCH P and POST RM. Understanding emotional prosody activates right hemisphere regions. *Archives of Neurology*, 53: 665-670, 1996.
- GOULET S and MURRAY EA. Neural substrates of crossmodal association memory in monkeys: The amygdala versus the anterior rhinal cortex. *Behavioral Neuroscience*, 115: 271-284, 2001.
- HOFFMAN EA and HAXBY JV. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, 3: 80-84, 2000.
- IMAIZUMI S, MORI K, KIRITANI S, HOSOI H and TONOIKE M. Task-dependent laterality for cue decoding during spoken language processing. *Neuroreport*, 9: 899-903, 1998.
- KANWISHER N, McDERMOTT J and CHUN MM. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17: 4302-4311, 1997.
- LEDoux J. *The Emotional Brain: the Mysterious Underpinnings of Emotional Life*. New-York: Simon and Schuster, 1996.
- LIEGEOIS-CHAUVEL C, DE GRAAF JB, LAGUITTON V and CHAUVEL P. Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cerebral Cortex*, 9: 484-496, 1999.

- MACALUSO E, FRITH CD and DRIVER J. Modulation of human visual cortex by crossmodal spatial attention. *Science*, 289: 1206-1208, 2000.
- MASSARO DW and EGAN PB. Perceiving affect from the voice and the face. *Psychonomic Bulletin and Review*, 3: 215-221, 1996.
- MAZZIOTTA JC and PHELPS ME. Positron emission tomography studies of the brain. In ME Phelps, JC Mazziotta and H Schelbert (Eds), *Positron Emission Tomography and Autoradiography: Principles and Applications for the Brain and Heart*. New York: Raven Press, 1986, pp. 493-579.
- MCDONALD JJ, TEDER-SÄLEJÄRVI WA and WARD LM. Multisensory integration and crossmodal attention effects in the human brain. *Science*, 292: 1791, 2001.
- MCGURK H and MACDONALD J. Hearing lips and seeing voices. *Nature*, 264: 746-748, 1976.
- MESULAM MM. From sensation to cognition. *Brain*, 121: 1013-1052, 1998.
- MORRIS JS, DE GELDER B, WEISKRANTZ L and DOLAN RJ. Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain*, 124: 1241-1252, 2001.
- MORRIS JS, FRISTON KJ, BUCHEL C, FRITH CD, YOUNG AW, CALDER AJ and DOLAN RJ. A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, 121: 47-57, 1998.
- MORRIS JS, FRISTON KJ and DOLAN RJ. Neural responses to salient visual stimuli. *Proceedings of the Royal Society London B Biological Sciences*, 264: 769-775, 1997.
- MORRIS JS, FRITH CD, PERRETT DI, ROWLAND D, YOUNG AW, CALDER AJ and DOLAN RJ. A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*, 383: 812-815, 1996.
- MORRIS JS, OHMAN A and DOLAN RJ. Conscious and unconscious emotional learning in the human amygdala. *Nature*, 393: 467-470, 1998.
- MORRIS JS, OHMAN A and DOLAN RJ. A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Sciences of the United States of America*, 96: 1680-1685, 1999a.
- MORRIS JS, SCOTT SK and DOLAN RJ. Saying it with feeling: Neural responses to emotional vocalizations. *Neuropsychologia*, 37: 1155-1163, 1999b.
- MURRAY EA and GAFFAN D. Removal of the amygdala plus subjacent cortex disrupts the retention of both intramodal and crossmodal associative memories in monkeys. *Behavioral Neuroscience*, 108: 494-500, 1994.
- MURRAY EA and MISHKIN M. Amygdalotomy impairs crossmodal association in monkeys. *Science*, 228: 604-606, 1985.
- NAHM FK, TRANEL D, DAMASIO H and DAMASIO AR. Cross-modal associations and the human amygdala. *Neuropsychologia*, 31: 727-744, 1993.
- PIHAN H, ALTENMULLER E and ACKERMANN H. The cortical processing of perceived emotion: A DC-potential study on affective speech prosody. *Neuroreport*, 8: 623-627, 1997.
- POURTOIS G, DEBATISSE D, DESPLAND PA and DE GELDER B. Facial expressions modulate the time course of long latency auditory brain potentials. *Cognitive Brain Research*, 14: 99-105, 2002.
- POURTOIS G, DE GELDER B, VROOMEN J, ROSSION B and CROMMELINCK M. The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, 11: 1329-1333, 2000.
- PRICE CJ, WISE RJ, WARBURTON EA, MOORE CJ, HOWARD D, PATTERSON K, FRACKOWIAK RS and FRISTON KJ. Hearing and saying. The functional neuro-anatomy of auditory word processing. *Brain*, 119: 919-31, 1996.
- ROSS ED. Affective prosody and the aprosodias. In MM Mesulam (Ed), *Principle of Behavioral and Cognitive Neurology*. Oxford: Oxford University Press, 2000, pp. 316-331.
- PRICE CJ and FRISTON KJ. Cognitive conjunction: A new approach to brain activation experiments. *Neuroimage*, 5: 261-270, 1997.
- ROSS ED, ORBELO DM, CARTWRIGHT J, HANSEL S, BURGARD M, TESTA JA and BUCK R. Affective-prosodic deficits in schizophrenia: Profiles of patients with brain damage and comparison with relation to schizophrenic symptoms. *Journal of Neurology, Neurosurgery and Psychiatry*, 70: 597-604, 2001.
- ROYET JP, ZALD D, VERSACE R, COSTES N, LAVENNE F, KOENIG O and GERVAIS R. Emotional responses to pleasant and unpleasant olfactory, visual, and auditory stimuli: A positron emission tomography study. *Journal of Neuroscience*, 20: 7752-7759, 2000.
- SCOTT SK, YOUNG AW, CALDER AJ, HELLAWELL DJ, AGGLETON JP and JOHNSON M. Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature*, 385: 254-257, 1997.
- SERGEANT J, OTHA S and MACDONALD B. Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain*, 115: 15-36, 1992.
- STREICHER M and ETTLINGER G. Cross-modal recognition of familiar and unfamiliar objects by the monkey: The effects of ablation of polysensory neocortex or of the amygdaloid complex. *Behavioral Brain Research*, 23: 95-107, 1987.
- TALAIRACH J and TOURNOUX P. *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme Medical Publishers Inc, 1988.
- VAN LANCKER D, KREIMAN J and CUMMINGS J. Voice perception deficits: Neuroanatomical correlates of phonagnosia. *Journal of Clinical and Experimental Neuropsychology*, 11: 665-674, 1989.
- VINGERHOETS G, BERCKMOES C and STROOBANT N. Cerebral hemodynamics during discrimination of prosodic and semantic emotion in speech studied by transcranial doppler ultrasonography. *Neuropsychology*, 17: 93-99, 2003.
- VROOMEN J, BERTELSON P and DE GELDER B. The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception and Psychophysics*, 63: 651-659, 2001a.
- VROOMEN J, DRIVER J and DE GELDER B. Is cross-modal integration of emotional expressions independent of attentional resources? *Cognitive and Affective Neurosciences*, 1: 382-387, 2001b.
- WHALEN PJ, RAUCH SL, ETCOFF NL, MCINERNEY SC, LEE MB and JENIKE MA. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, 18: 411-418, 1998.
- WIENHARD K, DAHLBOM M, ERIKSSON L, MICHEL C, BRUCKBAUER T, PIETRZYK U and HEISS WD. The ECAT EXACT HR – Performance of a new high resolution positron scanner. *Journal of computer assisted tomography*, 18: 110-118, 1994.
- WOODS RP, GRAFTON ST, HOLMES CJ, CHERRY SR and MAZZIOTTA JC. Automated image registration: I. General methods and intrasubject, intramodality validation. *Journal of Computer Assisted Tomography*, 22: 139-152, 1998a.
- WOODS RP, GRAFTON ST, WATSON JDG, SICOTTE NL and MAZZIOTTA JC. Automated image registration: II. Intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography*, 22: 153-165, 1998b.
- WORSLEY KJ, MARRETT S, NEELIN P, VANDAL AC, FRISTON KJ and EVANS AC. A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, 4: 58-73, 1996.
- ZALD DH. The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research Reviews*, 41: 88-123, 2003.

Gilles Pourtois, Neurology and Imaging of Cognition, University Medical Center (CMU), Bat. A, Physiology, 7th floor, room 7042, 1 rue Michel-Servet, CH-1211 GENEVA, Switzerland. e-mail: gilles.pourtois@medecine.unige.ch

(Received 24 January 2003; reviewed 21 February 2003; accepted 13 May 2003; Action Editor Edward De Haan)