# The voices of wrath: brain responses to angry prosody in meaningless speech

Didier Grandjean[1,5], David Sander[1,5], Gilles Pourtois[2], Sophie Schwartz[2], Mohamed L Seghier[2,3], Klaus R Scherer[1] & Patrik Vuilleumier[2,4]

**We report two functional magnetic resonance imaging experiments showing enhanced responses in human middle superior temporal sulcus for angry relative to neutral prosody. This emotional enhancement was voice specific, unrelated to isolated acoustic amplitude or frequency cues in angry prosody, and distinct from any concomitant task-related attentional modulation. Attention and emotion seem to have separate effects on stimulus processing, reflecting a fundamental principle of human brain organization shared by voice and face perception.**

Detection of potential threats may occur even when they are not initially in the focus of attention, eliciting enhanced perceptual analysis and re-orienting of processing resources[1,2]. Modulation of sensory processing by affective signals has been found in the visual cortex of humans[3,4] and monkeys[5], but it is not known whether such effects exist in modalities other than vision. We used functional magnetic resonance imaging (fMRI) to investigate (i) whether emotional prosody enhances neural activity in the auditory cortex, (ii) whether such enhancement involves voice-selective areas and (iii) whether any emotional effects of prosody depend on selective attention to the voice.

In Experiment 1, 15 healthy right-handed adults who gave informed written consent (7 women, mean age = 24.4 ± 4.6 years) underwent fMRI scanning (see **Supplementary Methods**) while listening to meaningless but word-like utterances pronounced with either angry or neutral prosody. These stimuli were extracted from pseudo-sentences previously validated[6] and matched for mean acoustic energy. To manipulate voluntary attention orthogonally to emotional prosody, we used a dichotic listening paradigm in which stimuli of 750-ms duration were presented simultaneously to either ear (angry and neutral ('AN'), neutral and angry ('NA'), or neutral and neutral ('NN'), respectively, on the right and left sides) during each trial in pseudorandom order (see **Supplementary Figure 1**). In two successive counterbalanced blocks within the same scanning run, participants selectively attended to either the left or right ear and performed a gender decision on the voice heard on the target side (mean accuracy = 88%). Imaging data were analyzed by event-related random-effect statistics across the whole brain volume using the Statistical Parametric Mapping 2 software (see **Supplementary Methods**).

To identify brain regions responding to emotional prosody irrespective of the target side of auditory attention, we compared activity elicited by pairs containing an angry voice (AN plus NA) relative to pairs containing only neutral voices (NN). Results showed selective increases in activity in the middle part of the right superior temporal sulcus (STS) ($x, y, z = 62, -30, 6$; $t = 5.80$; $P < 0.05$ corrected; see **Fig. 1a**) as well as in homologous areas of left STS ($x, y, z = -60, -24, 0$; $t = 4.43$; $P < 0.05$

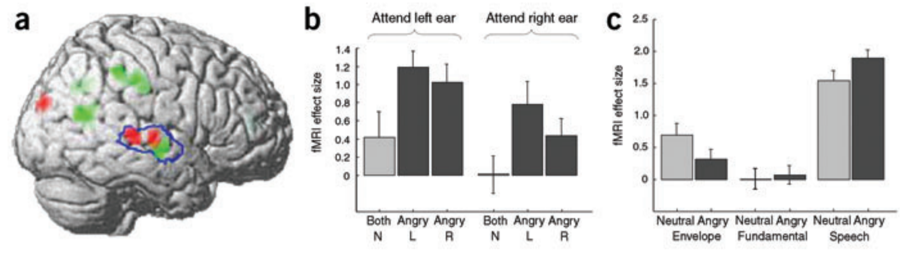


**Figure 1** Cortical activations elicited by spatial attention and emotional prosody. (**a**) Right hemisphere activations: increased responses for angry relative to neutral speech prosody were found in right STS, across all task conditions (red, $P < 0.001$, Experiment 1). An anterior region of right STS was also modulated by attention directed to the left relative to the right ear (green, $P < 0.005$, Experiment 1). These activations by emotion and attention occurred within voice-selective areas responding to speech more than to corresponding fundamental frequency or amplitude envelope cues presented in isolation (blue line, Experiment 2). (**b**) Mean parameter estimates of activity (percentage change relative to the global mean intensity of signal; ± s.e.m.) for right STS in Experiment 1. Blood oxygenation level–dependent (BOLD) responses were higher for angry (dark gray) versus neutral (light gray) speech, irrespective of the side of angry speech and of the side of selective listening. In addition, responses were also increased during attention to the left compared to the right ear, irrespective of prosody (N, neutral; L, left ear; R, right ear). (**c**) The same cluster in right STS in Experiment 2. Activation occurred only in response to vocal stimuli, not to synthetic sounds with matched fundamental frequency level or matched amplitude envelope heard in isolation.

[1]Geneva Emotion Research Group, Department of Psychology, University of Geneva, bd. du Pont d'Arve 40, CH-1205 Geneva, Switzerland. [2]Laboratory for Neurology and Imaging of Cognition, Departments of Neurology and Neurosciences, Centre Médical Universitaire, University of Geneva, rue Michel-Servet 1, CH-1211 Geneva 4, Switzerland. [3]Department of Radiology, Geneva University Hospital, Rue Micheli-du-Crest 24, CH-1211 Geneva 14, Switzerland. [4]Department of Psychology, University of Geneva, bd. du Pont d'Arve 40, CH-1205 Geneva, Switzerland. [5]These authors contributed equally to this work. Correspondence should be addressed to D.G. (Didier.Grandjean@pse.unige.ch).

Published online 23 January 2005; doi:10.1038/nn1392

corrected, see **Supplementary Figure 2**). These regions correspond to the location of one of several brain areas previously reported as voice selective[7,8]. Such activations occurred irrespective of which ear was the target of attention (**Fig. 1b**), indicating that the brain could still detect emotional prosody from voices that the participant was supposed to ignore (anger versus neutrality in relevant ear: $t > 4.14$; anger versus neutrality in irrelevant ear: $t > 4.35$; $P < 0.001$, for bilateral STS).

We found no interaction between emotion and attention in the auditory regions (even at a low threshold, $P < 0.05$ uncorrected). Notably, however, attention did modulate activation in the auditory cortex as shown by a general increase in activity in the right STS when participants had to judge voices from the left rather than the right ear, across all stimulus pairs (main effect of attention left side > right side; $x, y, z = 60, -12, -9$; $t = 3.45$, $P = 0.002$; see **Fig. 1a**). The peak of activation occurred in the most anterior part of the right STS region activated by anger. These results further demonstrate that participants followed selective attention instructions and differentially processed voices from the task-relevant ear. No reliable attention effect was found in left STS (right side > left side; $x, y, z = -54, -24, -6$; $t = 2.40$, $P = 0.015$, uncorrected, see **Supplementary Figure 2**).

This pattern of involuntary responses to anger independent of any concomitant modulation by voluntary attention was further supported by a repeated-measure ANOVA on the parameters of activity extracted from bilateral STS using attention (right or left ear), stimulus type (NN, AN, NA), and hemisphere (right or left) as factors. Results confirmed a main effect of stimulus type ($F_{2,13} = 21.8$, $P < 0.0001$) but, critically, no interaction of stimulus type with attention ($F < 1$) or with hemisphere ($F < 1$). Conversely, a 2×2 ANOVA on right STS responses to AN and NA stimuli during left side versus right side attention showed a significant effect of attended side ($F_{1,14} = 4.86$, $P = 0.045$) but no effect of anger side ($F < 1$) and no interaction ($F = 1.16$, n.s.).

Previous work has shown an activation of the middle temporal gyrus when attention is directed toward emotional meaning of vocal prosody as compared with semantic meaning[9], but, to our knowledge, our results provide the first demonstration that the middle right STS is modulated by angry prosody in voices and that such modulation occurs irrespective of current attentional relevance in a selective listening task. This supports recent proposals that this brain region might subserve high-level analysis of complex acoustic information in human voices[8,10].

To verify that activation by anger in STS was driven by vocal prosody rather than by related low-level acoustic features, we conducted a second fMRI experiment with the same participants. As fundamental frequency and distribution of energy through time have a critical role in conveying emotional information in voices[6], three categories of binaural stimuli of 750-ms duration were used in Experiment 2: angry or neutral speech, similar to the stimuli used in Experiment 1 but not heard before (AA-sp/NN-sp); synthesized sinusoid sounds matched for the mean fundamental frequency (F0) of each of the respective vocal stimuli used in Experiment 1 (AA-fo/NN-fo); and sounds consisting of white noise matched for the amplitude envelope of each stimulus used in Experiment 1 (AA-en/NN-en). Participants now had to judge whether two successive stimuli from the same category separated by a 100 ms silence were identical or different.

Results from Experiment 2 showed a selective activation of bilateral STS (right: $x, y, z = 68, -20, -4$; 258 voxels, $t = 12.87$; left: $x, y, z = -59, -12, 1$; 199 voxels, $t = 10.76$; both $P < 0.05$ corrected) in response to speech sounds as compared with both F0- and envelope-matched sounds (**Fig. 1a**). The STS voxels activated by angry prosody in Experiment 1 clearly overlapped with these voice-selective regions. Moreover, in Experiment 2, they also showed greater responses to angry versus neutral speech sounds (AA-sp > NN-sp, $t > 3.31$, $P < 0.002$; see also **Supplementary Table 1**), but no differential increases for the angry versus neutral equivalents of F0-matched sounds (AA-fo > NN-fo, $t = 0.13$) or envelope-matched sounds (AA-en > NN-en, $t = 0.51$) (**Fig. 1c**). Therefore, STS responses to angry speech, as observed in Experiment 1 irrespective of attended ear, were not simply driven by a particular range of frequency or a specific amplitude contour, but rather reflected more specific emotional voice-related processes.

Our data demonstrate that emotional signals from angry prosody may increase activity in the associative auditory cortex, an effect that occurs even when voice prosody and location are irrelevant to the listener's task, being independent or additive to any concomitant modulation by the spatial distribution of auditory attention in right STS. A similar enhancement by emotion additive to spatial attention was found in the face-sensitive fusiform area for fearful relative to neutral faces[11]. This suggests that the right STS might have a function in the auditory domain comparable to that of the right fusiform in the visual domain[11], finely tuned to extract socially and affectively salient signals from conspecifics[12]. Future research, requiring new acoustics and synthesis technology, should address more systematically the range of critical acoustic parameters involved in processing different voice qualities and expression of different emotions in STS.

Enhanced sensory responses to emotional events may constitute a fundamental neural mechanism shared by voice and face recognition systems, enabling emotion and attention interactions that prioritize orienting towards significant stimuli even when these are not in the focus of attention.

*Note: Supplementary information is available on the Nature Neuroscience website.*

ACKNOWLEDGMENTS
We thank T. Bänziger for helpful discussions. Supported by grants from the Swiss National Foundations (K.S., P.V. and S.S.).

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

Received 5 November; accepted 28 December 2004
Published online at http://www.nature.com/natureneuroscience/

1. Taylor, S.F. *et al. Neuroimage* **8**, 188–197 (1998).
2. Armony, J.L. & Dolan, R.J. *Neuropsychologia* **40**, 817–826 (2002).
3. Pourtois, G., Grandjean, D., Sander, D. & Vuilleumier, P. *Cereb. Cortex* **14**, 619–633 (2004).
4. Hadjikhani, N. & de Gelder, B. *Curr. Biol.* **13**, 2201–2205 (2003).
5. Sugase, Y., Yamane, S., Ueno, S. & Kawano, K. *Nature* **400**, 869–873 (1999).
6. Banse, R. & Scherer, K.R. *J. Pers. Soc. Psychol.* **70**, 614–636 (1996).
7. Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P. & Pike, B. *Nature* **403**, 309–312 (2000).
8. Belin, P., Fecteau, S. & Bédard, C. *Trends Cogn. Sci.* **8**, 129–135 (2004).
9. Mitchell, R.L.C., Elliott, R., Barry, M., Cruttenden, A. & Woodruff, P.W.R. *Neuropsychologia* **41**, 1410–1421 (2003).
10. Kriegstein, K.V. & Giraud, A. *Neuroimage* **22**, 948–955 (2004).
11. Vuilleumier, P., Armony, J.L., Driver, J. & Dolan, R.J. *Neuron* **30**, 829–841 (2001).
12. Gervais, H. *et al. Nat. Neurosci.* **7**, 801–802 (2004).