

Neural activity in speech-sensitive auditory cortex during silence

M. D. Hunter^{*†}, S. B. Eickhoff[‡], T. W. R. Miller^{*}, T. F. D. Farrow^{*}, I. D. Wilkinson[§], and P. W. R. Woodruff^{*}

^{*}Sheffield Cognition and Neuroimaging Laboratory (SCANLab), Academic Clinical Psychiatry, Division of Genomic Medicine, and [§]Academic Unit of Radiology, University of Sheffield, The Longley Centre, Norwood Grange Drive, Sheffield S5 7JT, United Kingdom; and [†]Institute for Medicine, Research Center Juelich, 52425 Juelich, Germany

Edited by Dale Purves, Duke University Medical Center, Durham, NC, and approved November 9, 2005 (received for review July 22, 2005)

That auditory hallucinations are voices heard in the absence of external stimuli implies the existence of endogenous neural activity within the auditory cortex responsible for their perception. Further, auditory hallucinations occur across a range of healthy and disease states that include reduced arousal, hypnosis, drug intoxication, delirium, and psychosis. This suggests that, even in health, the auditory cortex has a propensity to spontaneously “activate” during silence. Here we report the findings of a functional MRI study, designed to examine baseline activity in speech-sensitive auditory regions. During silence, we show that functionally defined speech-sensitive auditory cortex is characterized by intermittent episodes of significantly increased activity in a large proportion (in some cases >30%) of its volume. Bilateral increases in activity are associated with foci of spontaneous activation in the left primary and association auditory cortices and anterior cingulate cortex. We suggest that, within auditory regions, endogenous activity is modulated by anterior cingulate cortex, resulting in spontaneous activation during silence. Hence, an aspect of the brain’s “default mode” resembles a (preprepared) substrate for the development of auditory hallucinations. These observations may help explain why such hallucinations are ubiquitous.

auditory system | baseline activity | functional MRI

Tonic baseline activity is thought to be essentially uniform across the human brain volume (1). Specific foci (e.g., medial prefrontal regions) reduce their activity to below baseline upon commencement of structured mental processes, suggesting that an organized default mode of brain function is temporarily suspended during such processes (2). However, less is known regarding how focal baseline activity might naturally vary over time and whether episodic increases in activity might be sufficiently large to resemble statistically defined, externally evoked activation reported in functional MRI (fMRI) experiments. This question is important from both a neuroscientific and a clinical perspective, because spontaneous activation of the auditory cortex during silence is implicated in the emergence of auditory verbal hallucinations (3–6); perceptions in the absence of external stimuli, which occur in numerous states of health and disease, including reduced arousal, hypnosis, drug intoxication, delirium, and psychosis (7). This paper describes a fMRI experiment designed: (i) to examine variation of baseline activity in speech-sensitive auditory regions, in the absence of external stimuli, and (ii) to define the brain-wide functional anatomy associated with the emergence of spontaneous auditory activation during silence.

Previous work has shown that neural activity in the auditory cortex is subject to modulation by factors other than auditory sensory input. Attention to the auditory modality (8) and utilization of auditory imagery (9, 10) increase imaging parameters thought to represent markers of neural activity, including the blood oxygenation level-dependent response measured by fMRI (11). Hence, we hypothesize that auditory baseline activity is sensitive to modulation by background physiological processes that support cognitive biasing effects and, therefore, may vary in

such a way that intermittent episodes of increased activity during silence are observed. The anterior cingulate cortex (ACC), part of the limbic system adjacent to the medial prefrontal cortex, is involved in the recruitment of attention (12, 13) and is a putative source of a signal that might increase baseline activity in sensory cortex (14). With respect to the auditory system, this idea is supported by evidence for coactivation of ACC and auditory regions during silence in an auditory imagery task (15) and by the widely replicated observation of ACC activation during the experience of auditory hallucinations (6, 16–20).

In the current study, our prediction (based upon the hypothesis above) is that spontaneous activation of speech-sensitive auditory cortex occurs during silence, in conjunction with activation of ACC.

Methods

This study was approved by the North and South Sheffield research ethics committees.

Voice Stimuli. To identify speech-sensitive auditory regions, we created 24 unique stimuli for use in the fMRI experiment. Each stimulus consisted of a single emotionally neutral spoken phrase of three to four words and 1- to 2-s duration [e.g., “close the door” (21–23)]. All stimuli were digitally recorded at 44.1 kHz and 16 bits and presented within the scanner (below) at ≈50 dB sound pressure level over Commander XG electrostatic headphones (Resonance Technologies, Dallas).

Subjects and Scanning Protocol. Twelve healthy right-handed [mean ± SD right-hand dominance = 75.00 ± 21.98% (24)] male subjects aged 23 ± 3 years were studied. fMRI was performed by using echo-planar imaging on a 1.5-T Eclipse system (Philips Medical Systems, Cleveland) at the University of Sheffield. To circumvent problems associated with acoustic scanner noise, we used a sparse protocol (25) utilizing haemodynamic response latency to permit stimulus presentation during silent periods (32 × 4-mm contiguous slices at 48 time points, repetition time = 20,000 ms, echo time = 50 ms, field of view = 240 mm, in-plane matrix = 128 × 128). In an alternating stimulus vs. no-stimulus design, functional brain volumes were acquired after (i) the presentation of a single voice stimulus (as described above, every 40 s) and (ii) silence, over a single 16-min run (with the gradient coils turned off, quiescent scanner room sound was effectively cancelled by 30-dB headphone attenuation). We deliberately did not utilize pseudorandom presentation of stimuli because of a concern that this might lead to systematic attentional modulation (8, 14) of the auditory baseline due to vigilance with respect to

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Freely available online through the PNAS open access option.

Abbreviations: ACC, anterior cingulate cortex; BA, Brodmann’s area; fMRI, functional MRI; STG, superior temporal gyrus; TTG, transverse temporal gyrus.

[†]To whom correspondence should be addressed. E-mail: m.d.hunter@sheffield.ac.uk.

© 2005 by The National Academy of Sciences of the USA

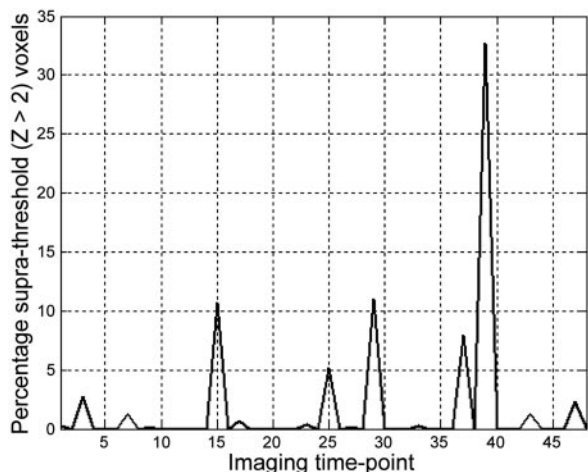


Fig. 2. Fluctuation in baseline activity during silence in functionally defined speech-sensitive auditory cortex. Data are presented from a single subject. The ordinate shows the percentage of functionally defined left-hemisphere voxels that are >2 SD above their own mean value during silence imaging time points (abscissa).

spontaneous activation data (for the example, mean, case) from the binomial distribution.

In the left hemisphere, 11 subjects exhibited two time points of spontaneous activation (8.33% of sampled baseline time), and one subject exhibited three such time points (12.50% of sampled baseline time). The mean \pm SD percentage of speech-sensitive voxels with $Z > 2$ during spontaneous activation was $12.41 \pm 8.26\%$. Example data from a single subject are shown in Fig. 2.

In the right hemisphere, eight subjects demonstrated two time points of spontaneous activation (8.33% of sampled baseline time), and four subjects demonstrated a single time point of spontaneous activation (4.17% of sampled baseline time). The mean \pm SD percentage of speech-sensitive voxels with $Z > 2$ during spontaneous activation was $9.52 \pm 6.16\%$.

Of the 35 time points of spontaneous activation occurring in at least one hemisphere, 42% involved only the left hemisphere, 29% involved only the right hemisphere, and 29% simultaneously involved both the left and right hemispheres.

To check for any temporal grouping of those time points where spontaneous activation was observed, we divided the functional scanning sessions into quartile time bins, each containing six silence time points. The occurrence of spontaneous activation time points did not significantly differ among the temporal quartiles ($\chi^2 = 3.27$; $P = 0.35$).

Because images were smoothed with a Gaussian kernel during preprocessing, which might accentuate spatial correlations in the data, we also performed these analyses using unsmoothed images. This did not result in appreciably different percentages of speech-sensitive voxels with $Z > 2$ during spontaneous activation and did not affect the significance of the difference between that expected and observed (Table 1). By way of a further sensitivity analysis, we reanalyzed our data assuming a Rayleigh distribution, which has a heavy positive tail and is thought to describe the worst-case scenario where reconstructed MRI amplitude images consist only of noise (28). This did not materially affect our results.

Spatial domain. Contrasting time points of spontaneous activation in the left hemisphere with the remainder of the silence baseline [l-spont vs. baseline; Fig. 3 (yellow)] revealed foci in the left lateral superior temporal gyrus [(STG); Brodmann's area (BA) 22; coordinates: $-55, -11, 4$; 145 voxels exceeded $P < 0.001$, uncorrected; $t = 9.22$] and more posterior STG (BA 22; coordinates: $-44, -31, 11$; 129 voxels; $t = 7.51$), overlapping with the

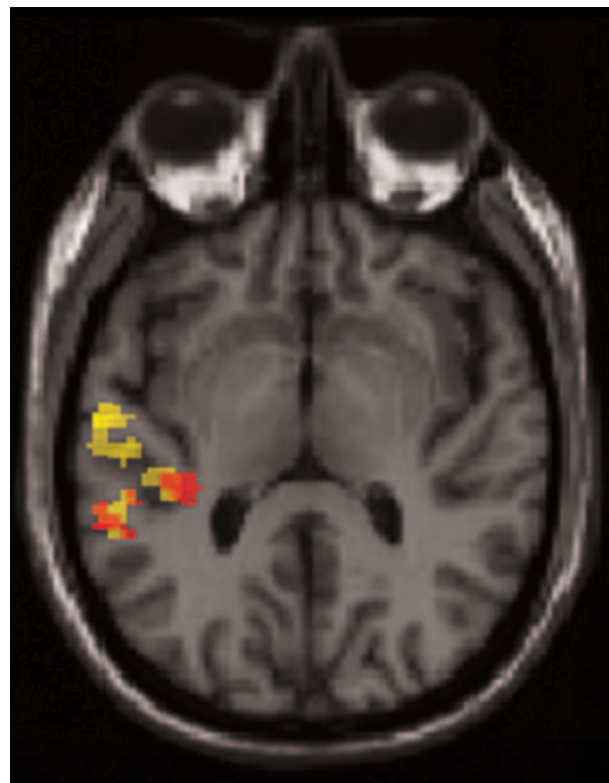


Fig. 3. At the group level ($n = 12$), foci for left- (yellow) and right- (red) hemisphere spontaneous activation phenomena are located in the left primary and association auditory cortices. Functional imaging data in a mixed-effects model are presented. These data show brain regions where statistically defined activation ($P < 0.001$, uncorrected) is associated with high levels of activity in the left- (yellow) and right- (red) hemisphere speech-sensitive brain volume. Note that this is not a simple tautology (i.e., finding sites of activation where activity is known to be high), because it represents the degree of anatomical constraint that can be applied to identifying foci of spontaneous activation in the population from which the subjects were drawn. This is well illustrated by the observation that high levels of activity in right-hemisphere speech-sensitive brain volume are associated with foci in the left auditory cortex. Data are displayed against a tilted axial slice through a canonical T1-weighted image, parallel to the superior temporal plane.

transverse temporal gyrus (TTG; BA 41). Furthermore, the r-spont vs. baseline contrast [Fig. 3 (red)] also revealed foci in the left STG (BA 22; coordinates: $-57, -38, 18$; 63 voxels; $t = 7.72$) and TTG (BA 41; coordinates: $-36, -28, 14$; 52 voxels; $t = 6.63$); significant overlap between l- and r-spont conditions in these regions was confirmed by conjoint statistical analysis [$t = 9.19$ (STG); $t = 7.69$ (TTG)].

In the l-spont vs. baseline contrast, we found a single further focus in the left ACC (BA 32; coordinates: $-2, 53, 12$; 56 voxels; $t = 6.56$), which extended to include part of the medial frontal gyrus (BA 9/10). The glass brain in Fig. 4 demonstrates the specificity of this result.

Further Analyses. We repeated the main analysis, substituting $Z < -2$ for $Z > 2$ to identify time points of spontaneous deactivation during silence. Using a within-subject repeated measures ANOVA, we compared total percentage $Z > 2$ voxels during spontaneous activation (reflecting both number of time points and voxels) with $Z < -2$ voxels during spontaneous deactivation. In the left hemisphere, more voxels achieved $Z > 2$ during spontaneous activation than $Z < -2$ during spontaneous deactivation, but this effect was not observed in the right hemisphere

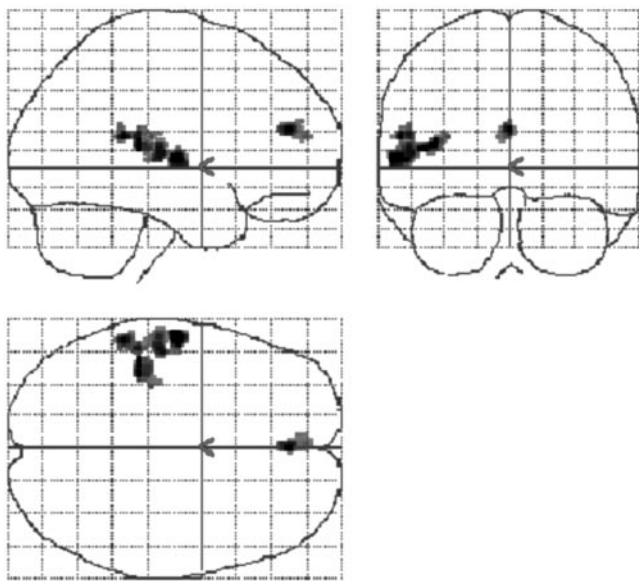


Fig. 4. The ACC demonstrates activation during episodes of high auditory baseline activity. Foci across the whole brain volume where statistically defined activation ($n = 12$; mixed-effects model; $P < 0.001$, uncorrected) is associated with high levels of baseline auditory activity in left-hemisphere speech-sensitive voxels. Such foci are confined to the temporal and anterior cingulate cortices. Data are displayed within a glass brain to demonstrate the specificity of these effects.

(Table 2). Hence, there was a significant direction (positive vs. negative) by hemisphere interaction ($F = 5.79$; $P = 0.03$).

We also examined variation in magnitude of stimulus-evoked activation by repeating the main analysis using data from the speech (not silence) time points and a criterion of $Z > 2$ with respect to the speech-specific mean. This procedure identified time points with more $Z > 2$ voxels than would be expected under Gaussian assumptions, which we refer to as time points of unexpectedly large activation. More left-hemisphere voxels achieved $Z > 2$ during spontaneous activation (in silence) than $Z > 2$ during unexpectedly large stimulus-evoked activation; this was not observed in the right hemisphere (Table 2). A condition (silence vs. speech) by hemisphere interaction was significant ($F = 5.39$; $P = 0.04$).

Table 2. Median time points and mean total percentage speech-sensitive voxels for spontaneous activation and deactivation and unexpectedly large stimulus-evoked activation

Condition	Criterion	Median (range) time points	Mean \pm SD total % voxels
Spontaneous activation			
Left hemisphere	$Z > 2$	2 (2–3)	25.85 \pm 13.27
Right hemisphere		2 (1–2)	15.87 \pm 8.61
Spontaneous deactivation			
Left hemisphere	$Z < -2$	1.5 (1–2)	15.13 \pm 5.40
Right hemisphere		1 (1–2)	16.31 \pm 9.05
Unexpectedly large stimulus-evoked activation			
Left hemisphere	$Z > 2$	1 (1–3)	17.66 \pm 14.24
Right hemisphere		2 (1–2)	17.85 \pm 11.11

The mean total percentage of voxels with $Z > 2$ (or < -2) reflects both numbers of time points and voxels with $Z > 2$ (< -2) at such time points, for each subject ($n = 12$). Note that the mean total percentage of voxels with $Z > 2$ (< -2) is relatively homogeneous across conditions and hemispheres, with the exception of spontaneous activation in the left hemisphere, where more voxels are involved (see text for statistical analysis).

In the entire group of subjects, across all brain voxels and all silence time points, Z score distribution closely approximated to a standard Gaussian distribution (Kolmogorov–Smirnov statistic = 0.005; $P > 0.05$). Repeating our main analysis with data from the whole brain (i.e., without the functional mask of speech-sensitive voxels) revealed some time points of spontaneous activation occurring anywhere in the brain during silence. However, the mean \pm SD percentage of voxels with $Z > 2$ was small ($3.86 \pm 1.66\%$). In particular, this percentage was significantly smaller than in the case of left- ($t = 5.18$; $P < 0.001$) and right- ($t = 4.50$; $P < 0.001$) hemisphere spontaneous activation of speech-sensitive regions. Furthermore, time points of spontaneous activation in auditory regions and anywhere in the brain were not usually identical (80% of left- and 75% of right-hemisphere time points of spontaneous activation in speech-sensitive regions were distinct). Application of the (spatial) mixed-effects analysis revealed that time points of spontaneous activation occurring anywhere in the brain were associated with two foci in the right insula. Hence, there was evidence for perturbation of the overarching Gaussian distribution of Z scores, which we observed across the whole brain volume, at those foci (coordinates: 44, 17, -1 ; 31 voxels exceeded $P < 0.001$, uncorrected; $t = 7.99$ and coordinates: 42, -2 , 2; 30 voxels; $t = 5.21$).

Finally, to exclude the possibility that our results were influenced by any random walks of the MRI scanner field, we also acquired and analyzed data from a phantom head with internal structure. We did not detect evidence of spontaneous activation.

Discussion

We have shown that temporal variation in the auditory baseline is manifest as intermittent episodes of strikingly increased activity within speech-sensitive regions. These episodes arise in the absence of external auditory stimulation and exceed the statistical definition of activation used in functional image analysis. Hence, we have referred to such phenomena as spontaneous activations. Anatomical foci of spontaneous activation are located in the left STG and medial TTG [site of the primary auditory cortex in humans (29)].

These data suggest that spontaneous auditory activation in both left and right hemispheres can be explained by foci in the left temporal cortex (Fig. 3). Essentially, this means that bilateral spontaneous activation (in the time domain) implicates a specific population of left-hemisphere voxels (in the space domain) that is consistent between subjects. On the other hand, right-hemisphere spontaneous activation (in the time domain) is associated with greater intersubject variability with respect to the right-hemisphere voxels that are involved, but simultaneous intersubject consistency with respect to a population of left-hemisphere voxels (associated with right-hemispheric activation). Overall, we argue these results suggest that spontaneous activation of speech-sensitive regions is driven by a left-hemispheric effect, which impacts upon the right hemisphere via homotopic connections (30, 31). This notion can be further specified by considering similarities and differences in topography between left-hemisphere foci associated with left- and right-hemisphere spontaneous activation. Voxels in the left-lateral STG are associated only with left-hemisphere effects, but voxels in the left medial TTG and posterior STG are associated with left and right spontaneous activation phenomena (Fig. 3). Hence, these latter foci are more likely to be the sources of effects that can be detected in the right hemisphere, perhaps with the posterior STG as predominant due to its involvement in processing complex characteristics of speech (21).

Our analysis of spontaneous deactivation also suggests that spontaneous auditory activation is a left-lateralized phenomenon. Although the data show that Z score distribution in the auditory baseline is somewhat more polarized than throughout the entire

brain (Table 2), with more voxels achieving $Z > 2$ and $Z < -2$ than would be expected under Gaussian assumptions, we can nonetheless detect evidence of spontaneous activation above and beyond any such polarization. Specifically, more voxels achieve $Z > 2$ than $Z < -2$ during silence, but only in the left hemisphere. Similarly, more voxels achieve $Z > 2$ during silence than $Z > 2$ during speech perception (i.e., by comparison with their speech-specific mean) but only in the left hemisphere (Table 2). The latter observation is compatible with speech serving to constrain spontaneous activation in left-sided speech-sensitive regions; an interpretation consistent with the idea that externally evoked activation and endogenous activation may compete for finite neural resources up to a point of saturation (4).

The observed lateralization of spontaneous activation in speech-sensitive auditory cortex is in accordance with many demonstrations of left-hemispheric dominance for speech and language functions (32). It is possible that such dominance derives from interhemispheric differences in neuronal properties such as connectivity patterns and strengths and synaptic gain control, caused by asymmetries in the regional distribution of neurotransmitters (or their receptors), or electrophysiological differences related to the distribution of specific ion channels and transporters (33–37). These mechanisms might also account for the left auditory cortex's propensity to undergo spontaneous (and anatomically focal) activation during silence. Furthermore, although global vascular effects on the blood oxygen level-dependent signal have been shown to demonstrate some regional variation (38), the distinct lateralization of spontaneous activation that we observed (together with our correction for mean voxel value) provides evidence against any such vascular confound in the current data.

As predicted, the ACC was revealed as an extratemporal focus where activity was associated with spontaneous activation of left hemisphere speech-sensitive auditory cortex. This finding is consistent with ACC exerting influence over regionally specific brain activity (12), a process that might amount to the physiological basis of attention (13, 14). Given the overlap between our results and those of an auditory imagery experiment (15), it could be tempting to suggest that we have imaged the correlates of attention to thoughts involving the auditory domain. However, we deliberately sought to examine baseline auditory processes without reference to any specific cognitive task or mental state. Ours is a study of auditory physiology, not phenomenology. Hence, we avoid any dualist interpretation of our results, instead emphasizing that the physiological processes that we have described might represent a generic mechanism by which the brain supports modulation of focal auditory activity [evident in cognition experiments (8–10)].

Moreover, if an aspect of the brain's default mode (1, 2) is well suited to support physiological modulation of auditory activity, then it is also possible that the brain might be naturally predisposed to sustain aberrant auditory activity in pathological states, manifest as auditory hallucinations. We believe this is an important issue, because attempts to define a structural abnormality in the auditory cortices of people with auditory hallucinations have been subject to problems of nonreplication (39, 40). Our results suggest that the concept of a focal anatomical lesion need not necessarily be invoked to explain auditory hallucinations. Rather, these states might arise from a functional disturbance of the auditory system, which shows an unstable baseline even in healthy subjects. Fluctuation of activity within the temporocingulate network that we have described is associated with non-perception in health, but it is possible that dysfunction within such a network could lead to false perception in disease (3–6).

Such transformation into perception could arise under conditions of heightened ACC influence on auditory cortical function, a notion consistent with the widely replicated finding of ACC activation during auditory hallucinations (6, 16–20).

Outside the mask derived from functional definition of speech-sensitive voxels, the Z score distribution across the dimensions of whole brain space and scan time was close to Gaussian. Nonetheless, we were able to identify foci in the right insula that, like the left temporal cortex, were associated with spontaneous activation (but at different time points). Although these findings are of their own interest [the functional associations of the insula are numerous (41)], they also enable further specification of our main results. First, it is apparent that time points of defined spontaneous activation detectable within the speech-sensitive and whole-brain volumes are not (generally) the same. Second, even if they were the same, the percentage of voxels with $Z > 2$ anywhere in the brain would be insufficiently large to explain the speech-sensitive voxel data. Hence, the functional mask of speech-sensitive voxels is not a window on more widespread phenomena, and spontaneous auditory activation is not simply an expression of fluctuating Z scores across the whole brain. Moreover, it is possible (indeed, a further testable hypothesis) that the overarching Gaussian distribution of Z scores reflects baseline activity in numerous functionally defined subvolumes of the brain (e.g., auditory, visual, and somatosensory regions), each with the capacity to undergo spontaneous activation, but at different time points.

Although we have used the term spontaneous activation, this phenomenon differs from the concept of activation (within a deterministic system) that is commonly used in models of neuroimaging data (42, 43). In such models, the brain can be activated only by driving external inputs, i.e., experimental stimuli. On the other hand, in this study, we have described a situation where structured activation is generated by the brain itself. This demonstrates one of the major discrepancies between neurocomputational models and physiological observations: the presence of nonrandom internally generated information derived from physiological substrata such as oscillating ion channels, which are found at high concentrations within distinct generator regions (44–46). We thus suggest that the neuroimaging concept of activation could be expanded to include the effects of spontaneously discharging internal generators. In this context, it is important to emphasize that spontaneous is not equivalent to random; the shifts in baseline activity are distinctly clustered in both the temporal (Fig. 2) and spatial (Figs. 3 and 4) domains (i.e., relatively proximal voxels exhibit unusually high signal levels at similar times).

Conclusion

We have provided evidence for a phenomenon of spontaneous activation in the baseline of speech-sensitive auditory regions. Such activation arises in the left primary auditory cortex (TTG) and association auditory cortex (STG), in conjunction with activation of the ACC. If, as we hypothesize, this temporocingulate network represents a generic mechanism by which the brain supports modulation of focal auditory activity, then dysfunction within such a network could be manifest as auditory hallucinations. Hence, we propose a biologically informed model for auditory hallucinations based upon perturbation of normal baseline processes, a model that might explain why auditory hallucinations are ubiquitous.

The Wellcome Trust supported this work.

1. Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 676–682.
2. Gusnard, D. A. & Raichle, M. E. (2001) *Nat. Rev. Neurosci.* **2**, 685–694.
3. Woodruff, P., Brammer, M., Mellers, J., Wright, I., Bullmore, E. & Williams, S. (1995) *Lancet* **346**, 1035.

4. Woodruff, P. W. R., Wright, I. C., Bullmore, E. T., Brammer, M., Howard, R. J., Williams, S. C. R., Shapleske, J., Rossell, S., David, A. S., McGuire, P. K., et al. (1997) *Am. J. Psychiatry* **154**, 1676–1682.
5. Dierks, T., Linden, D. E. J., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H. & Singer, W. (1999) *Neuron* **22**, 615–621.

