NeuroReport **9**, 1787–1792 (1998)

A three stage model of awareness: formulation and initial experimental support

**J. G. Taylor,[1,2,CA] L. Jäncke,[1,3,CA]
N. J. Shah,[1] T. Nösselt,[4] N. Schmitz,[4]
M. Himmelback,[4] T. Kalenscher
and H.-W. Müller-Gärtner[1]**

[1]Institute of Medicine, Research Centre Juelich, D-52425 Juelich, Germany; [2]Centre for Neural Networks, Department of Mathematics, King's College London, UK; [3]Institute of General Psychology, Otto-von-Guericke-University Magdeburg, Germany; [4]Institute of General Psychology, Heinrich-Heine-University Düsseldorf, Germany

[CA,1] Corresponding Authors and Address

THE various components which together make up the complex state of consciousness require neural support involving a connected network of many brain areas at differing levels. At the lowest level is non-aware processing, of which there is not direct awareness. There are also modules involved in processing with awareness but without focussed attention. Finally there must be a set of modules involved in directing attention in a controlled manner. We expect to be able to dissociate the various components of the three-stage network by using different levels of attention. The results of an auditory experiment performed under three different levels of awareness and attention are analysed to show support for the three-stage model of awareness. The relevant auditory areas are delineated. *NeuroReport* 9: 1797–1792 © 1998 Rapid Science Ltd.

Key words: Attention; Auditory processing; Awareness; Temporal and prefrontal cortices

Introduction

Various proposals have recently been made as to the possible neural sites supporting awareness, such as frontal cortex,[1] hippocampus[2] and neural assemblies that support working memory.[3,4] Not all neural sites in the processing path are of prime importance for consciousness; for example several experiments indicate the secondary character of activity in primary visual area V1 for awareness of visual inputs.[5,6] Thus there is support for a two-stage model of awareness, in which specialized cortical areas are involved in unconscious preprocessing with later areas supporting the direct creation of awareness of inputs. There is also evidence for the relevance of attention to awareness, where changes of response of neurons in various areas of cortex are observed under attended *vs* unattended conditions.[7] It has been suggested that attention is used to bind different codes of objects together to lead to awareness.[8] However while such binding needs to occur in object recognition in which attention is focused it may not do so in a more unfocused attentional state where divided attention occurs. Moreover further cortical regions are activated under focused attentional conditions beyond those solely for awareness.[9] Thus the lowest level of awareness (which we tentatively identify with unfocused or divided attention) must be included. The original two-stage model has thus to be expanded to involve three stages[10]: preprocessing, first emergence of awareness and then attentionally guided awareness. It is this three-stage model (schematically shown in Fig. 1) which is analysed here, and experimental data reported on which support the model at a very preliminary level.

One of the basic predictions made by the three-stage model of awareness of Figure 1 is that activation levels in the three separate stages will, to a first approximation, dissociate under suitable task conditions. Thus under subliminal or unattended conditions it is expected that activity in response to a given input will only be observable in the relevant first-stage modules. Under the same stimulus input conditions but now with passive awareness both first- and second-stage module activity should be observed, whilst under attended conditions to the same input further third-stage modules will be activated by a suitable attentional load. Such dissociation has been observed in the n-back paradigm,[11] in which a subject has to remember the recurrence of an input stimulus

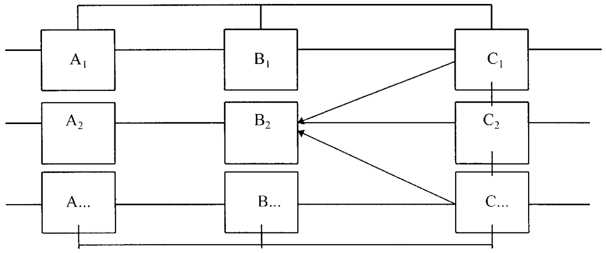

FIG. 1. The three-stage model for the creation and control of awareness. The Modules at the lowest or first stage are denoted A_1, A_2, ... and are involved in low level feature analysis in the appropriate modality. A number of modules are involved in such processing. At the next stage are sites of buffer memory in posterior cortex, denoted B_1, B_2, ... They are fed by preprocessed activity from first stage modules, and are coupled to each other in order to run a competition by lateral inhibition so as to select the most appropriate percept, given earlier context still present in activity traces on these modules, to enter phenomenal awareness. The final third stage is composed of modules which are frontal and possess attentional feedback control both to the second-stage modules, so as to refresh activity there, and to the first-stage modules to reduce or increase stimulus selectivity. Feedback is shown explicitly from the third to lower stages although it will also be present from the second to first stage (not all feedback is shown explicitly).

that occurred n inputs previously. It was found that dorsolateral prefrontal cortex was not activated for short time delays (corresponding to $n < 2$) but was for larger values of n (for delays of $\geqslant 20$ s).

The predictions at the simplest approximation level, according to the model of Figure 1, of activation levels of the modules involved in the processing, are shown in Fig. 2. They are given in terms of hypothetical z-scores representing normalized differences between baseline and current activation, and indicate that under the varying attentional demands of the different task conditions there are expected to be levels of activation which are roughly constant across all conditions for the first-stage modules, the second-stage modules being differentiated only by being activated for the passive or attentional parts of the tasks, whilst the third-stage modules are only active in the attended state. It is these predictions, and more generally the three-stage model itself, which we test in this paper, using the results of the experiment briefly presented in Ref. 12.

Materials and Methods

The study used fMRI scans conducted on 10 normal adults. Functional brain activation was mapped in a 1.5 T Siemens (Magnetom Vision) scanner equipped with echo planar imaging (EPI) and a standard radiofrequency head coil. Informed consent was obtained from all subjects about the nature and possible consequences of the study.

MRI parameters: Sequence parameters were as follows: echo time (TE) 66 ms; repeat time (TR) 5 s; field of view = 200 × 200 mm; spin-flip angle alpha = 90°; matrix size = 64 × 64; in-plane resolution = 3.125 × 3.125 mm; slice thickness = 3 mm; inter-slice gap = 0.3 mm. Using a mid-sagittal scout image, 16 axial slices were oriented along the anterior–posterior commissure (AC-PC) plane, with the lowermost slice positioned to be 20 mm below the AC-PC line. In addition, high resolution, T1-weighted anatomical images of the entire brain were obtained in 3D using a pulse sequence with the following parameters: TR = 40 ms; TE = 5 ms; alpha = 40°; 1 excitation; field of view = 250 mm; in-plane matrix = 256 × 256; 128 sagittal slices with 1.25 mm slice thickness.

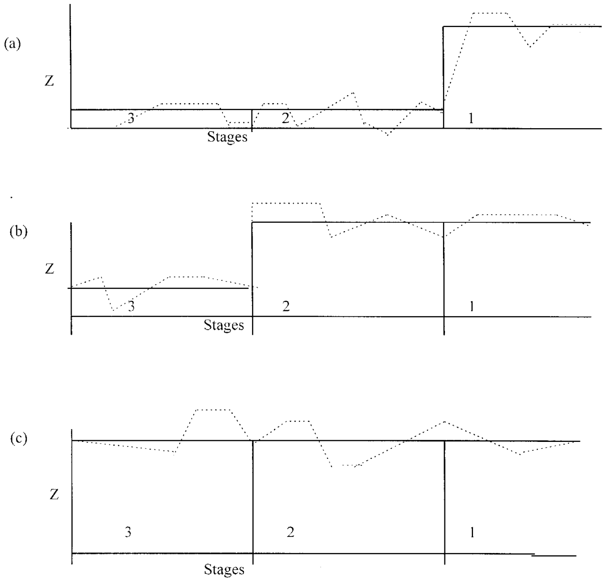

FIG. 2. A schematic of averaged activation levels (given by z-scores) predicted by the three-stage awareness model across the task conditions (**a**) unattended, (**b**) passive sensing and (**c**) attended processing to detect a particular parameter of the incoming stimulus. The same stimulus is assumed used in all the three conditions. The prediction is only the simplest approximation, neglecting modulatory feedback attentional effects from the third stage (and possible additional second) stage modules. The continuous lines represent the averaged activity over the relevant stages; the dashed lines denote a more realistic set of a range of activity levels to be expected form different multi-area stages under the three different paradigms. Moreover there may be difference between the overall activity levels in the three paradigms. Thus the heights of the curves in (a), (b) and (c) are only relative.

Experimental protocol: Each series began with three baseline images (15 s interval) which were discarded to allow the signal to reach a steady state, followed by 48 images during which activation alternated with rest every 30 s (60 s/cycle, 12 images/cycle, 4 cycles). Auditory stimuli were presented binuarally using a digital playback system, a magnetically shielded transducer system and air conduction through paired plastic tubes with suppression of ambient scanner noise by about 15 dB. During each experimental condition a series of 51 images was acquired. Each series consisted of multiple periods of baseline, during which subjects heard only the ambient machine noise, alternating with periods of activation, during which prepared auditory stimuli were delivered. The total duration of each image series was about 5 min. The auditory stimuli used throughout, and for which different levels of awareness were required, were 16 bit, digitally sampled consonant–vowel syllables (/ka; /ta, /pa, /ga; /da; /ba/). The onset, duration (500 ms), intensity and fundamental frequency of the stimuli were edited and synchronized. These syllables were randomized. One-third of the trials consisted of the syllable /ta/ serving as target syllable in the experiment. Visual stimuli used as a distractor comprised flickering light delivered in the frequency range of 3–4 Hz. It was not regarded as practical to use a distracting sound input since we had no way of differentiating the regions supporting consciousness for the two inputs. However, the input (the stream of CV syllables) to the subjects was kept constant throughout the experiment. Three conditions were used involving increasing attention to the auditory stimulus: (I) ignore auditorily presented syllables, but attend to a flickering light (inattention); (II) passive listening to the syllables (passive); and (III) discriminate amongst the syllables (attention).[12] Ten subjects, with a range of ages, were scanned with eyes open with digitally recorded consonant (CV)-syllables presented at a rate of 1 Hz. The visual stimulus used in the inattention condition was a flickering light in the frequency range of 3–5 Hz delivered to goggles fitted to the head coil. During rest subjects listened to the scanner noise; these alternated with periods of activation. During the inattention condition subjects were instructed to ignore all of the auditorily presented syllables but to scan the flicker light for changing frequency, while during the attention condition subjects were instructed to detect the syllable /ta/. Occurrence of the targets (/ta/ and changing frequency) was indicated by briefly lifting the index finger of the left hand which was recorded outside the magnet.

FMRI activation tasks: Subjects were scanned with eyes open and room lights dimmed. Auditory stimuli were presented at a rate of 1 Hz. During rest subjects listened to the scanner noise, which alternated with periods of activation. Subjects were exposed to the three experimental conditions in a randomized order, one being the dual task involving both the auditory and light stimuli and the remaining two solely auditory. For the dual task subjects were simultaneously exposed to a flickering light delivered in a frequency range of 3–5 Hz via goggles fitted to the head coil.

Image analysis: Image analysis was performed on a SPARC 20 workstation using MATLAB and SPM96b software.[13–15] All images were realigned to correct for head movement between scans[15] and then were coregistered and transformed into a standard stereotactic space corresponding to the atlas of Talairach and Tournoux.[16] In this space one pixel represents 2 × 2 mm in the x and y dimensions, with an interplanar distance of 2 mm. These sterotaxically transformed functional data sets from each subject were smoothed slightly (Gaussian filter of root mean square radius 4 mm) to compensate for normal variation in sulcal/gyral anatomy across subjects. Voxels that had values > 0.8 of the volume mean in all the images were selected to restrict analysis to intracranial regions. The effects of global whole volume activity and time were removed using linear regression and sine/cosine functions as confounds (corresponding to high-pass filtering the time series to remove low frequency artifacts due to cardiorespiratory and other cyclical components).

Statistical parameter mapping: For statistical analysis the sterotaxically normalized fMRI time-series data of the 10 subjects were pooled. The alternating periods of baseline and activation were modelled using a delayed box-car reference vector accounting for the delayed cerebral blood flow after stimulus presentation. Significantly activated pixels were searched for by using the General Linear Model approach for time-series data suggested by Friston and colleagues.[13–15] For this we defined a design matrix comprising contrasts testing for significant activations during auditory stimulation in each condition separately (tests for simple main effects) and testing for differential activations between conditions (tests for interactions). Only voxels exceeding a z-score of 5 ($p < 0.001$, corrected for multiple comparisons) as height threshold and belonging to a cluster of $\geqslant 20$ voxels (160 μl) were considered as significant. These voxel were superimposed on a rendered standardized brain[16] (Fig. 3).

Results

The qualitative predictions shown in Fig. 2 were tested using the data taken under the three conditions of inattention, passive and attention.[12] In general, the three conditions activated perisylvian brain areas comprising BA 41, 42 ,21 and 22. The conditions with the highest level of awareness additionally activated frontal regions. The smallest spatial extent of activation was found in the inattention condition (Fig. 3), while there was an increase in the total area activated caused by new regions arising in going successively from the inattention to passive conditions and then to the attention condition. The data are analysed and presented more fully in Table 1, which displays stereotaxic centre of

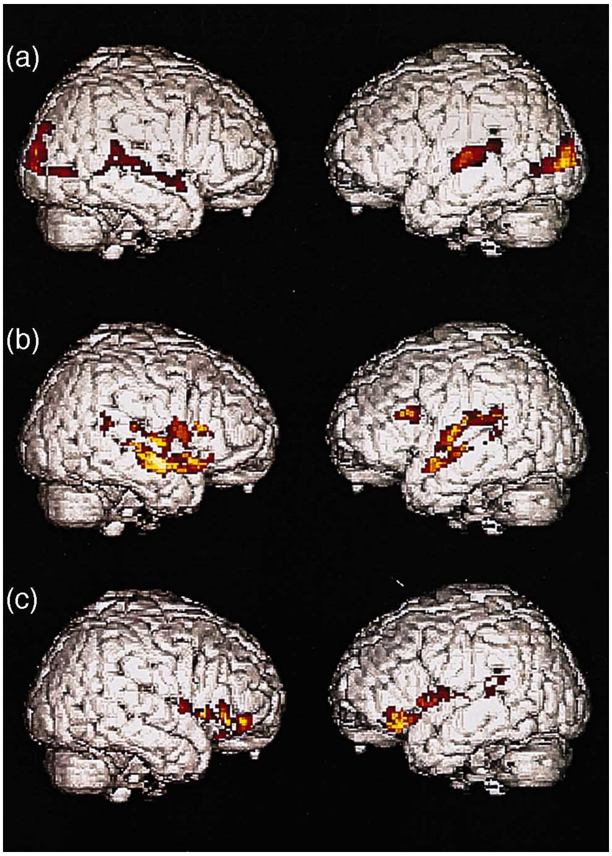

FIG. 3. Surface-projected schematic representation of mean activated areas on a stereotactically normalized rendered brain: (**a**) inattention (ignore auditory stimuli but attend to visual stimuli), (**b**) passive minus inattention, (**c**) discriminate minus passive listening. From the text and Fig. 1 we identify the areas shown under (a) as stage 1 modules, those under (b) as stage 2 modules and those for (c) as the modules of stage 3.

Table 1. Stereotaxic center of mass coordinates and size of activation clusters. We report only positive activation peaks and activation clusters located within the parenchyma that exceed a statistical criterion of $p < 0.001$ (corrected for multiple comparisons), corresponding to a z score $\geqslant 5$

	x	y	z	Z_{max}	No. voxels
Inattention					
R cuneus	10	–94	4	9.3	3846
R superior temporal gyrus	52	–26	6	8	464
L superior temporal gyrus	–52	–22	8	8.7	397
Passive–inattention					
R superior temporal gyrus	50	–24	4	8.41	1236
L superior temporal gyrus	–56	–22	2	8.23	774
L inferior frontal gyrus	–38	16	22	7.22	126
R inferior frontal gyrus	48	20	14	6.89	65
Discriminate–passive					
R inferior frontal gyrus	40	32	–2	7.7	402
L gyri orbitali	–38	28	–10	7.8	298
L inferior frontal gyrus	–54	10	–2	6.7	101
R inferior frontal gyrus	46	6	0	6.1	64
L superior temporal gyrus (posterior)	–54	–34	10	8.8	29
L superior temporal gyrus	–56	–20	6	8.9	26
R gyri orbitali	40	30	–16	6.7	24
R cerebellum	12	–68	–12	5.7	22

Stereotaxic coordinates refer to medial-lateral position (x) relative to midline (positive = right), anterior–posterior position (y) relative to the anterior commissure (positive = anterior), and superior-inferior position (z) relative to the commissural line (positive = superior). L: left, R: right.

mass coordinates and size of activation clusters for the inattention condition as well as the passive–inattention, and discriminate–passive differences.

From both Fig. 3 and Table 1 it can be seen that auditory stimulation under the first three conditions mainly causes activity in perisylvian brain regions. These regions are therefore supportive of the predicted activation levels across the three conditions of the non-aware modules 1 in Fig. 2. Most importantly, under the passive condition there are additional areas activated on the right and left middle temporal gyrus, the right and left superior temporal gyrus as well as the left middle frontal gyrus. Finally the difference between discriminate and passive shows up strong right and left-sided frontal activations, corresponding to the activation of modules 3 of Fig. 2. The modules of classes 1, 2 and 3 of Fig. 2 identified above are summarized in Table 1. With our use of the statistical parameter $z = 5$ as a threshold in Fig. 3 (which is significant at the $p = 0.001$ level, corrected for multiple comparisons), a strong cluster of activity shows up in BA46/45 on the left hemisphere, which is interesting not only because of being Broca's area but also because of the large extent it presents under the discriminate condition.

Further support for the predictions of Fig. 2 came from the fact that signal intensity in those regions activated in all conditions only differed marginally (about 1% signal intensity difference from inattention to passive, and passive to discriminate) while there were dramatic signal intensity increases in the additionally activated areas rising from 0 to a maximum of 14%.

It is necessary to ensure that there was little awareness of the stimuli in the inattention condition for the data to be relevant to the model. We have evaluated that question by requiring the subjects to indicate the level of success (on 10-point scales ranging from little success (1) to high success (10)) in performing the first three tasks immediately after the fMRI scanning and found that the subjects succeeded well in inattending (mean ± s.d. 9 ± 0.5), passive listening to (8 ± 2), and discriminating (9 ± 2) the stimuli. This implies that the corresponding levels of awareness and attention of the auditory stimuli under the various conditions were as desired (very low awareness, high awareness, and high awareness and attention for the respective conditions I, II and III) for our interpretation of the data.

Discussion

The simple three-stage model shown in Figs 1 and 2 for the emergence of awareness and its control by attentional processors has been supported by the data reported in Ref. 12, with the results shown in Fig. 3 and Table 1. It is important to try to validate the model shown in Fig. 1 in other modalities, as well as to probe the second approximation to it, taking better account of the differences between activation levels across the cortex, for example by the use of structural modelling. The results presented here in any case indicate that the three-stage dissociation of cortical areas for auditory processing is a valid first approximation to processing of auditory signals. There are, however, other experimental results which seem to be at variance with those presented here, and especially about the change of levels of activation of earlier cortical areas under the highest attentional load as compared to the lowest one.[9,17–20] We will consider each of these results in turn.

Woodruff *et al.*[9] observed important differences in auditory areas (BA41/42/21/22) when attention was paid to simultaneously visually and auditorily presented digits: a decrease was found in the activity in auditory areas when the visual form of the stimuli was attended to which was recovered when the auditory form of the stimuli were noticed. However there may be competition between the two forms of presentation of a given concept; the experimental paradigm is different in our case since quite different stimuli are being presented in the different modalities. Such a reduced effect is also noticed in the lack of an attentional blink between two modalities,[21] which is in support of our results. Woldorff *et al.*[17] reported that the M100 peak, measured by MEG in dichoptic listening, was doubled in height when attention was paid to a tone in that ear while a different tone in the other ear was neglected. However the data were analysed only using a single current dipole, resulting in averaging the neural activity over a considerable area. An increase of the M100 could arise from an increase of the total area of neural activation and not from increase in activation in a given area. Pitch and phonetic discrimination were required of subjects studied by Zattore *et al.*,[18] but no control was made for awareness; the areas observed are consistent with our results. The use of a single dipole (in this case for EEG) also limits the spatial sensitivity of the results reported by Heinz *et al.*,[19] which involved attention to left or right of a stimulus array investigated by both PET for localization and by EEG for temporal resolution; a similar problem arises in the EEG analysis of Clark and Hillyard,[20] in which subjects were presented with a patterned stimulus and an increase of the P1 and N1 amplitudes was detected under increased attentional load.

There are very few previous studies of the site for the emergence of awareness in audition. One result from MEG measurement, of the M100 variation of position with pitch and frequency,[22] claimed that the pitch perception of a tone has a neural substrate place prior to cortex. However, this result was based only on a single dipole fit to the MEG signal, so that it will be necessary to wait for more general distributed source analysis of such a MEG experiment, say along the lines of magnetic field tomography.[23] There is still the difficult problem of reconciling our result that awareness initially arises in cortex with the claim of Pantev *et al.*[22] that it is subcortical. In order to discuss that more fully we turn to a comparison of awareness in vision and hearing. In the former awareness is known to occur after primary processing in V1 in cortex. The feature integration theory of Ref. 8 indicates in what manner the initially fragmented representation of a visual input is brought together by later attention (which could be diffuse or focused). By analogy we would expect that prior primary sensory processing in audition would then lead in later modules (there are now known to be 6 or 7 separate auditory areas in total) to awareness of the sound.

The mismatch-negativity signal (MMN)[24] is one which occurs in primary auditory cortex but is preattentive (it is still present under such a strong distractor load that awareness of the auditory signal is completely lost). This result alone indicates that the interpretation of the results reported by Pantev *et al.*[22] may be premature, and that only in some later processing stage are the signals for the pure and mixed sounds separated. That would agree with the nature of the MMN signal as well as relate suitably to the visual processing stages.

In total the present experimental data do not conflict with our results: attention does not appear to cause an appreciable increase in activity in the preprocessing stage 1 areas of Figs 1 and 2, and awareness (or unfocused attention) occurs in a specific site in cortex separate form the cortical region dedicated to earlier nonconscious processing. Thus the most important result of our paper, consistent with the prediction of Fig. 2 of the three-stage model, is that there are distinct brain areas (stages 2 and 3 of Figs 1 and 2) which are activated to support, first, awareness (or unfocused attention) and then, directed attention. This result is important since we can now extend it to probe in what manner the separate cortical areas for stages 2 and 3 are contributing to allow awareness and directed attention to arise. This is a question to be analysed more carefully by use of higher temporal resolution, say by EEG and MEG.

However from our comments on experiments using such tools single dipole analysis may be too coarse to be effective.

Conclusion

The data analysed here give support to the three-stage model of awareness shown in Fig. 1 for auditory processing. They delineate the relevant areas as superior temporal gyrus, additional areas in the superior temporal gyrus and a small region in the inferior frontal gyrus and mainly inferior frontal and orbital gyrus respectively. The level of activity over these areas was roughly constant across conditions. We also need to explore the validity of the model in vision and somatosensation, as well as determine a more precise and complete map in the cases we have considered.

References

1. Crick FHC. *Nature* **379**, 485–486 (1996).
2. Gray JG. *Behav Brain Sci* **18**, 659–675 (1995).
3. Baddeley A. In: Gathercole SE *et al.* eds, *Theories of Memory*, Hillsdale, NJ: Erlbaum 1993: 11–28.
4. Taylor JG. *Verh Deutsch Zool Ges* **86**, 159–163 (1993).
5. Kolb FC and Braun J. *Nature* **377**, 336–338 (1995).
6. He S, Cavanagh P and Intriligator J. *Nature* **383**, 334–337 (1996).
7. Treue S and Maunsell JHR. *Nature* **382**, 539–541 (1996).
8. Treisman A. *Q J Exp Psychol* **40A**, 201–203 (1988).
9. Woodruff PWR, Benson RR, Bandettini PA *et al. NeuroReport* **7**, 1909–1913 (1996).
10. Taylor JG. *The Race for Consciousness.* Boston: MIT Press (1988).
11. Cohen JD, Perlstein WM, Braver JS *et al. Nature* **386**, 604–608 (1997).
12. Jäncke L, Posse S, Shah NJ *et al. Hum Brain Mapp* **5**, S191 (1997).
13. Friston KJ. *Proc R Soc Lond B Biol Sci* **261**, 401–405 (1995).
14. Friston KJ and Frith CD. *Clin Neurosci* **3**, 89–97 (1995).
15. Friston KJ, Williams S, Howard R *et al Magn Reson Med* **35**, 346–355 (1996).
16. Talairach J and Tournoux P. *Co-Planar Stereotaxic Atlas of the Human Brain. 3-Dimensional Proportional System: An Approach to Cerebral Imaging.* New York: Thieme, 1988.
17. Woldorff MG, Gallen CC, Hampson SA *et al. Proc Natl Acad Sci USA* **90**, 8732–8736 (1993).
18. Zatorre RJ, Evans AC, Meyer E and Gjedde A. *Science* **256**, 846–849 (1992).
19. Heinze H, Mangun GR, Burchert W *et al. Nature* **372**, 543–546 (1994).
20. Clark VP and Hillyard SA. *J Cogn Neuroscience* **8**, 387–402 (1996).
21. Duncan J, Martens S and Ward R. *Nature* **387**, 808–810 (1997).
22. Pantev C, Hoke M, Lutkenhoner B and Lehnertz K. *Science* **246**, 486–488 (1989).
23. Ioannides AA, Bolton JPR and Clarke CJS. *Inverse Problems* **6**, 591–594 (1990).
24. Naatanen R. *Attention and Brain Function.* Hillsdale NJ: Erlbaum (1992).

ACKNOWLEDGEMENTS: This study was supported by a grant of the Deutsche Forschungsgemeinschaft to L.J. The guest scientists (M.H., T.K., L.J., T.N., N.S., J.G.T.) would like to thank the Research Center Jülich for providing the facilities that made this work possible and Dr Stefan Posse for help and advice in data analysis.

Received 3 March 1998;
accepted 8 March 1998