The planum temporale as a computational hub

Timothy D. Griffiths and Jason D. Warren

It is increasingly recognized that the human planum temporale is not a dedicated language processor, but is in fact engaged in the analysis of many types of complex sound. We propose a model of the human planum temporale as a computational engine for the segregation and matching of spectrotemporal patterns. The model is based on segregating the components of the acoustic world and matching these components with learned spectrotemporal representations. Spectrotemporal information derived from such a 'computational hub' would be gated to higher-order cortical areas for further processing, leading to object recognition and the perception of auditory space. We review the evidence for the model and specific predictions that follow from it.

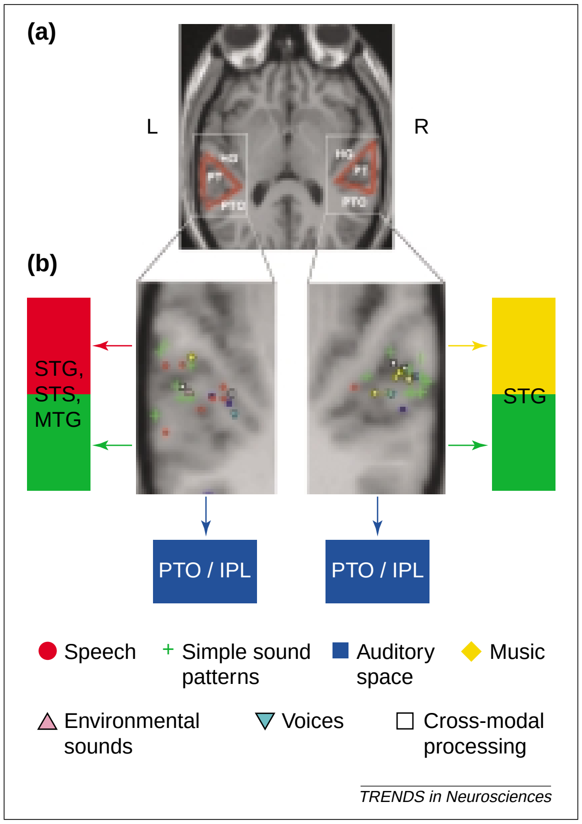

Fig. 1. The planum temporale (PT) as an anatomical and functional hub. (a) Tilted axial section through the superior temporal plane of the human brain. The PT lies posterior to Heschl's gyrus (HG), the site of the primary auditory cortex, and is contiguous posteriorly with the parieto–temporal operculum (PTO). Ninety-five percent probability maps for the boundaries of left and right PT in humans (derived from Ref. [1]) are outlined in red. (b) Insets centred on left and right PT, showing functional activation peaks within PT associated with different types of complex sound processing (see Table 1). Symbols are explained underneath. The functional relationships between the PT and higher cortical areas that are coactivated in processing simple sound patterns (green), music (yellow), speech (red) and auditory space (blue) are indicated schematically. Arrows indicate postulated flow of information from the PT to these higher areas; in many cases, however, exchange of information is likely to be reciprocal. We propose a generic computational mechanism within the PT for the analysis of spectrotemporal pattern. Computation uses information about sound objects derived from higher cortical areas linked to the PT, and the output of the PT is used to update stored information in these same areas. Abbreviations: IPL, inferior parietal lobe; MTG, middle temporal gyrus; PTO, parieto–temporal operculum; STG, lateral superior temporal gyrus; STS, superior temporal sulcus.

What does the human planum temporale (PT) do? This large region, which occupies the superior temporal plane posterior to Heschl's gyrus, is generally agreed to represent auditory association cortex. However, disagreement exists regarding its anatomy [1] and structure–function relationships [2–4]. In the left hemisphere, most definitions of Wernicke's area include part of PT [5] and, indeed, the human PT has traditionally been viewed as a language processor [2]. However, functional imaging indicates that the PT processes diverse types of sound (Fig. 1, Table 1). This article develops a functional model to explain this.

The PT is concerned with analysis of sounds that are spectrally and temporally complex, comprising several component frequencies that change over time (Fig. 2). Such sounds are common in nature. The brain is continuously required to analyse these incoming spectrotemporal patterns and to compare them with those previously experienced, during the process known as auditory scene analysis [6]. Such analysis allows the identification and assignment of position to a mixture of acoustic sources (sound objects) heard simultaneously. This demands both segregation of the spectrotemporal pattern associated with each sound object and separation of each object from the spectrotemporal effects of its location. We argue that the PT solves this daunting computational problem.

Although mechanisms for the accurate representation of incoming acoustic spectrotemporal signals exist in the ascending auditory pathways and primary auditory cortex (PAC) [7–9], it would be surprising if *a priori* this system were sufficient for auditory scene analysis. Even the discrimination of a single sound object from the effect of spatial position (Fig. 3) requires learned information about how the external ears filter sound signals arising in different locations, in addition to accurate representations of the sound waveform at the eardrums. This demanding computation might be achieved serially in the PT after initial processing in PAC, using the modular architecture in the PT [10] and inputs from other cortical areas [11–14].

Such computation would transform incoming auditory patterns into information about acoustic objects and position that could be used in other cortical areas. In this model, the PT thus represents a computational 'hub' that directs further processing in other cortical regions, consistent with studies of the

Timothy D. Griffiths*
Jason D. Warren
Newcastle University Medical School, Framlington Place, Newcastle-upon-Tyne, UK NE2 4HH.
Wellcome Dept of Imaging Neuroscience, Institute of Neurology, Queen Square, London, UK WC1N 3BG.
*e-mail: t.d.griffiths@ncl.ac.uk

0166-2236/02/$ – see front matter PII: S0166-2236(02)02191-4

Table 1. Functional imaging data on planum temporale (PT) involvement in different aspects of spectrotemporal pattern processing[a]

Principal contrast	Side	PT peak activation (mm)			Regions activated concurrently	Refs
		x	y	z		
Spatial analysis						
Sound source rotation minus stationary sound object (fMRI)	L	–42	–34	4	Bilat. PTO, L IPL	[16]
	R	64	–24	12		
Sound source rotation minus stationary sound object (PET)	L	–48	–32	8	Bilat. PTO, Bilat. Premotor Area	[16]
	R	60	–36	12		
Simple sound patterns						
Duration sequences minus silence	L	–66	–38	12	Cb, Bilat. HG, Bilat. STG, Bilat. IPL, Bilat. Frontal Lobe	[38]
	R	68	–26	2		
Harmonic complex minus pure tones	L	–54	–22	2	R HG, Bilat. STG	[34]
	R	66	–18	10		
Frequency-modulated minus unmodulated tones	L	–64	–14	2	Bilat. HG, Bilat. STG	[34]
	R	56	–16	4		
Amplitude-modulated minus unmodulated noise	L	–60	–28	8	Bilat. HG, L STS, L STG, L IPL	[35]
	R	62	–22	6		
Spectral motion versus stationary stimuli	L	–66	–16	8	Bilat. STG	[61]
	R	62	–30	12		
Spectrotemporal minus fixed external sound	L	–56	–32	12	Bilat. STG	[16]
	R	66	–24	12		
Pitch sequences						
Pitch sequences minus silence	L	–60	–28	6	Cb, Bilat. HG, Bilat. STG, Bilat. IPL, Bilat. Frontal Lobe	[38]
	R	66	–30	6		
Tone sequences minus words (active task)	L	–55	–31	17		[44]
Tone sequences minus noise	L	–59	–32	14	Bilat. STG, R STS	[36]
	R	48	–23	12		
Environmental sounds						
Passive listening minus rest	L	–56	–30	16	Bilat. HG, R Inf. Frontal Lobe, R Insula, R IPL	[39]
Voices						
Vocal minus non-vocal sounds	L	–40	–37	13	Bilat. STS, R MTG	[40]
	R	56	–30	6		
Music						
Deviant minus standard chords (pre-attentive)	R	58	–24	8	R STG	[62]
Melodies minus noise	R	62	–25	3	R STG, R Fusiform Gyrus	[41]
Listening to familiar songs minus visual baseline	L	–55	–18	5	Bilat. HG, Bilat. STG, Bilat. Frontal Lobe, L IPL, R SMA	[63]
	R	59	–23	6		
Maintenance of pitch while singing minus complex pitch perception	R	51	–30	15	R HG, Bilat. Frontal Lobe, Bilat. Insula, Bilat. IPL, Bilat. Occipital Lobe, Cb	[64]
Musical imagery (imagining continuation of a tune minus listening)	R	56	–30	8	Bilat. Frontal Lobe, L SMA	[42]
Speech and speech-like sounds						
Speech minus noise	L	–58	–21	8	Bilat. STG, L MTG, L Inf. Frontal Lobe	[65]
Speech minus complex non-speech	L	–64	–44	12	Bilat. MTG, Bilat. STG, R Inf. Frontal Lobe	[46]
Speech minus tones	L	–44	–32	8	Bilat. MTG, Bilat. STG, R Insula	[46]
Complex non-speech minus tones	L	–64	–24	4	Bilat. STG, R MTG	[46]
Consonant–vowel syllables minus vowels	L	–48	–32	12	R STG, R STS	[66]
	R	44	–28	12		
Unvoiced minus voiced consonants	L	–44	–28	12	L HG	[66]
Verbal self-monitoring (reading aloud with distorted feedback minus reading aloud)	L	–52	–36	16	L STS, R STG, L Insula	[50]
Dichotic listening						
Listening to dichotic minus diotic speech	L	–52	–30	13	Bilat. STG, Bilat. STS, Bilat. Inf. Frontal Lobe, R Insula	[52]
Active listening						
Active target detection minus passive listening	L	–58	–52	18	L STS, L IPL, L Frontal Lobe, L Thalamus, L Insula	[51]
Cross-modal processing						
Coherent visual motion minus stationary stimulus	L	–41	–30	12	Bilat. V5, Bilat. V3	[54]
Optical flow minus randomized optical motion	R	57	–20	8		
Lip-reading minus watching meaningless facial movements	L	–58	–28	3	L PTL, R IPL	[56]
	R	61	–22	13		
Auditory plasticity						
Sign language minus visual fixation in deaf subjects	L	–52	–38	12	R STG	[57]
Post-training minus pre-training deactivation	L	–48	–48	12	Bilat. STG, Bilat. STS, R HG	[58]

[a] All local PT maxima fall within the 95% probability anatomical boundaries for human PT proposed by Westbury *et al.* [1]. Studies have been selected to illustrate the variety of types of pattern processing in PT and the different cortical areas coactivated in each case. Abbreviations: Bilat., bilateral; Cb, cerebellum; HG, Heschl's gyrus; Inf., inferior; IPL, inferior parietal lobe; L, left; MTG, middle temporal gyrus; PTL, posterior temporal lobe; PTO, parieto–temporal operculum; R, right; SMA, supplementary motor area; STG, superior temporal gyrus; STS, superior temporal sulcus.

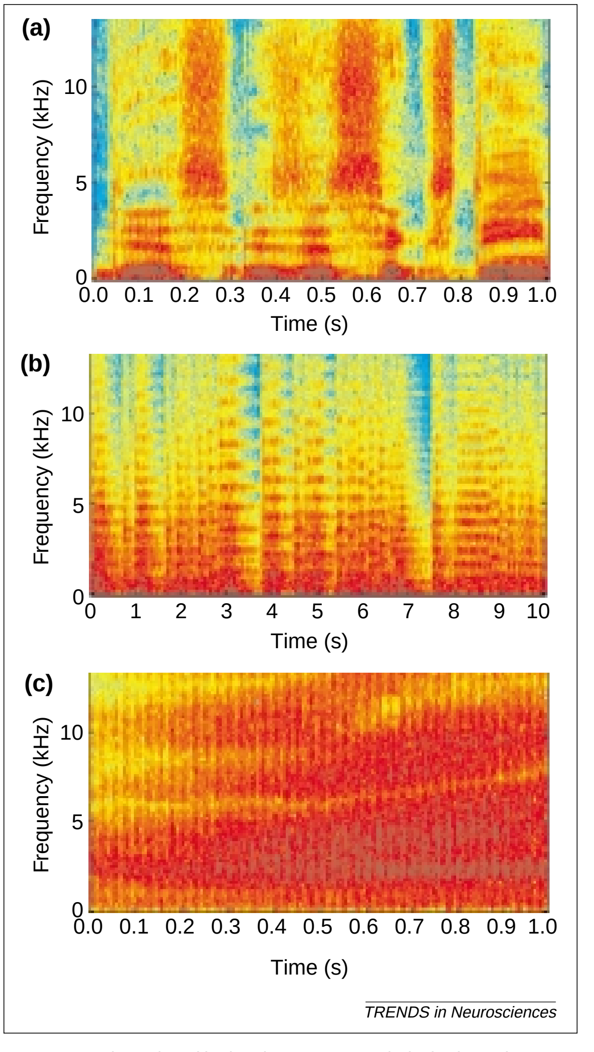

Fig. 2. Sounds analysed in the planum temporale (PT). The PT is involved in the analysis of sounds with complex spectrotemporal structure where there are multiple frequency components that change over time. This is shown in these spectrograms that show the frequency spectrum of the sound at one eardrum as a function of time. (a) 1 s sample of speech. (b) 10 s sample of classical orchestral music. (c) 1 s sample of amplitude modulated noise moving quickly around the head, similar to the sound of an insect.

cortical processing of language [15], auditory space [16] and other types of pattern within complex sound (Table 1, Fig. 1). Such a hub could access distinct cortical mechanisms for sound-object identification and localization [17–19] (Fig. 1).

The model: computational analysis of sound patterns

The segregation and matching of spectrotemporal patterns could be achieved in the PT using similar computational mechanisms and neuronal architecture. In this scheme, the PT is a crucial computational interface between incoming sound patterns that are segregated in the PT, and the previously stored patterns with which these are matched. The output after such computation provides information about the acoustic environment that is not immediately available either in the acoustic input or as a result of auditory processing before the PT.

Spectrotemporal analysis in the PT can be considered over three different timespans. The first corresponds to the segregation of simultaneous spectrotemporal patterns, for example, when multiple sound objects are presented. The second corresponds to the segregation of spectrotemporal patterns between successive time points during analysis of the motion of sound objects in space, or analysis of a succession of sounds in time. This timescale corresponds to that of transient acoustic 'memory' suggested by human electrophysiology [20]. Finally, the PT might operate over longer timespans, to effect matching of the incoming spectrotemporal patterns with stored patterns or 'templates'. The timescale of PT computation need not bear a simple relationship to the temporal structure of the acoustic waveform represented in PAC. Depth-electrode studies in humans suggest that rapid acoustic temporal changes, such as voice-onset time, are less clearly represented in the PT than in PAC [21,22], consistent with the processing of stored representations over hundreds of milliseconds rather than the faithful temporal representation of the incoming stimulus.

The analysis that we suggest might be achieved by several different algorithms instantiated in a variety of neural networks. Independent and dynamic component analyses (ICA and DCA, respectively) belong to this family of algorithms [23,24]. Essentially, these achieve separation of the components of a mixture by assuming that these are not correlated [25]. The segregation mechanisms that we envisage could have some of the formal features of ICA, because the auditory system operates under similar constraints to those under which ICA was developed. However, there are no grounds for specifying one member of this family of algorithms, or for suggesting equivalence between a specific artificial neural network used for ICA and the actual neural networks in the PT. Several neural network configurations might achieve the same computational result. We do, however, predict that the required computation is likely to be performed by a neuronal population within the PT, rather than at the level of single neurons.

In terms of information theory, the problem confronting the auditory system is the transfer of maximum information from multiple auditory sources to their neuronal representations. Information about sources forms a 'convolutive mixture' in the acoustic waveform at the eardrum. Here, each sound source is convolved with a filter function that corresponds to the effect of the external ear on sound from a particular region of space [26]; sound sources are further processed by a series of filters in the ascending pathway to PAC. In terms of ICA, the multiple-source problem can be solved when the characteristics of the mixing filters are known (essentially, a form of ICA that is not completely 'blind'). The PT could operate on the information from

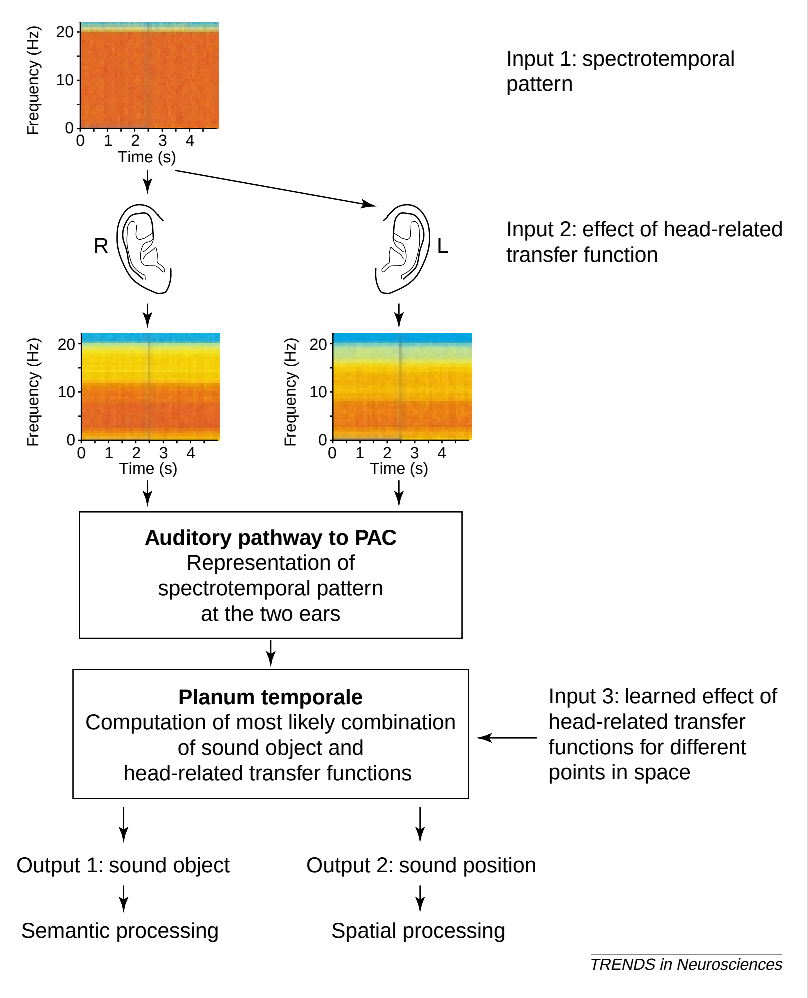

Fig. 3. The computational hub in action: auditory spatial analysis. The spectrotemporal pattern at the two ears results from convolution of the acoustic signal in space (in this example, a square-wave amplitude-modulated noise, similar to the sound of a helicopter) with the head-related transfer functions (HRTF) at the two ears (in this example, corresponding to a location above the subject, to the right). Initial processing of spectrotemporal patterns, including comparison between the ears, occurs in the ascending pathway to the primary auditory cortex (PAC). In the cortex, it is likely that multiple neurons are required to encode a given position in space [30,67]. Our model proposes serial input from PAC to planum temporale (PT), followed by further processing in the PT to compute the most likely combination of sound objects and positions producing the binaural spectrotemporal pattern in PAC. In performing the computation, the PT accesses learned information about the acoustic world (in this example, the HRTF) stored locally or in higher cortical areas. Output from the PT comprises spatial information that passes to the parieto–temporal operculum and inferior parietal lobule, and sound-object information that passes to the temporal convexity for semantic processing.

PAC in such a way by accessing information about the characteristics of these auditory filters acquired as a result of experience [27]. In addition to the mixing filters, we propose that the PT also has access to information about previously experienced sound objects. The filter and object properties might be stored locally in the PT or in other cortical areas. These properties are also key components of other biologically plausible models. For example, in predictive coding [28] they would be described as feedback 'predictions' that allow computation of the source properties of natural scenes.

Output of the PT would consist of segregated spectrotemporal patterns corresponding to auditory objects and their spatial characteristics. This output would feed forward to areas that store information regarding the 'mixing filter' and sound-object characteristics, where it would be used to update this stored information. Accordingly, because reciprocal 'top-down' and 'bottom-up' processing are an essential feature of the segregation process that we envisage as occurring in PT, ours could be classified as a generative model [29]. A core feature common to generative models is plasticity – the capacity for modification of the computational algorithm based on experience.

Evidence for the model

Spatial perception

Auditory spatial analysis is the prototypical application of our computational model (Fig. 3). The PT or its homologues have been implicated in acoustic spatial analysis in both electrophysiological studies in monkeys [19,30,31] and functional imaging in humans [16,17,32,33]. Recent functional imaging experiments using broadband stimuli have demonstrated PT activation when the computation of sound movement requires segregation of the effect of movement from the intrinsic structure of the sound [16,32]. In this situation, a movement trajectory could be computed in the PT by continuous segregation of a spectrotemporal pattern that corresponds to the original sound object from movement-transformed versions of itself.

Elementary acoustic pattern perception

Evidence also exists for specific PT activation in the processing of spectrotemporal patterns that are not spatially determined, including harmonic complexes [34], amplitude modulation [35], frequency modulation [34,36] and sound sequences [37,38]. By contrast, differential activation does not occur in PAC, consistent with the primary cortex acting as a conduit for further processing.

Environmental sound perception

The PT is engaged in processing a range of naturally occurring environmental sounds, including animal cries and inanimate noises [17,39]. Voices are specific examples of ethologically salient environmental sounds for which stored templates might exist. When contrasted with non-vocal sounds of similar frequency distribution, activation of the left PT by voices could reflect processing based on such templates [40].

Musical perception and imagery

The perception of music demands the capacity to build and retain long-lasting abstractions of spectrotemporal structures. Right-sided PT activation typically accompanies the perception of melodies (Table 1), and active tasks that involve musical pitch recruit a frontal–temporal network, which includes the PT (in normal subjects) [41]. A frontal–temporal network involving the PT is also activated during musical imagery [42] and during musical hallucinations [43], when subjects perceive music in the absence of any musical stimulus. In the case of music, therefore, the proposed initial spectrotemporal analysis in the PT affords access to widely distributed and lateralized brain processing mechanisms that operate over different temporal scales.

Speech perception

The representation and updating of auditory speech traces are necessary for phonological working memory and speech production [5]. The PT has been implicated in these processes in studies of both normal and brain-damaged subjects [44,45]. It is also activated by natural speech contrasted with acoustically similar non-speech sounds [46], by deviant or unpredictable verbal and nonverbal events [47–49], and in verbal self-monitoring [50]. These observations are consistent with the suggestion [5] that Wernicke's area constructs a transient representation of the spectrotemporal structures embodied in spoken words, regardless of whether these are heard or retrieved from lexical memory (i.e. a phonological 'template'). Such a role would be crucial for distinguishing phonemes as closely related spectrotemporal structures [36]. However, we do not argue that PT is necessarily the primary storage site for such templates.

Other considerations

Attention

Attentional ('top-down') influences might modulate PT computation via its connections to other association areas (Table 1). Activation of the PT in studies that specifically assess the effect of attention [51] and in dichotic listening [52] might be interpreted in this way. However, functional imaging studies of auditory spatial and object processing have demonstrated that PT activation does not depend on whether a task is employed [17]. Moreover, one generator for the pre-attentive mismatch negativity response in electroencephalography and magnetoencephalography [20] (an electrophysiological correlate of 'oddball' or novel stimuli) arises in the vicinity of the PT. We therefore suggest that, although PT contributes to an auditory attentional network, its computational role does not depend on attention.

Cross-modal processing

Area Tpt of the macaque, a potential homologue of the human PT, contains neurons that are responsive to visual and somaesthetic, as well as auditory, stimuli [53]. Cross-modal processing of visual motion has also been demonstrated in the human PT [54]. Activation of the PT during reading [55] and lip-reading [56] can be interpreted as examples of cross-modal processing that involve access to phonological spectrotemporal templates.

Auditory learning

Left PT activation in response to sign language in prelingually deaf individuals [57] is consistent with recruitment of the computational hub by an entirely different sensory modality: a striking example of plasticity. Our model predicts the bilateral deactivation of the PT specifically associated with short-term auditory learning that is demonstrated in normal subjects [58], as such training could lead to the establishment of stored templates for acoustic targets, improvements in computational efficiency and reductions in metabolic demands.

Lateralization

Lateralized PT activation during processing of language and musical stimuli (Table 1, Fig. 1) is not a specific feature of our computational hub model. However, the model could accommodate lateralization of processing determined by stimulus features [59,60] or lateralized downstream cognitive processing.

Predictions

Figure 3 illustrates a scheme for the identification and localization of single sound objects in space. One of the most challenging tasks for the auditory system is to execute this task for multiple objects, in the 'cocktail party effect' (where we perceive and attend to the voice of one speaker when many speakers are present). We predict a crucial involvement of the PT in this task.

Our model is based on the properties of local networks within the PT, and could be tested experimentally using single unit recording in animal homologues of PT, or depth-electrode studies in humans. One core feature of the model that might be examined directly in this way is the plasticity of unit responses to spectrotemporal patterns.

Acknowledgements
Our work is supported by the Wellcome Trust. We thank G. Green, K. Friston and A. Rees for helpful discussion.

References

1 Westbury, C.F. *et al.* (1999) Quantifying variability in the planum temporale: a probability map. *Cereb. Cortex* 9, 392–405

2 Marshall, J.C. (2000) Planum of the apes: a case study. *Brain Lang.* 71, 145–148

3 Keenan, J.P. *et al.* (2001) Absolute pitch and planum temporale. *NeuroImage* 14, 1402–1408

4 Zatorre, R.J. *et al.* (1998) Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. *Proc. Natl. Acad. Sci. U. S. A.* 95, 3172–3177

5 Wise, R.J.S. *et al.* (2001) Separate neural subsystems within 'Wernicke's area'. *Brain* 124, 83–95

6 Bregman, A.S. (1990) *Auditory Scene Analysis*, MIT Press

7 de Charms, R.C. *et al.* (1998) Optimizing sound features for cortical neurons. *Science* 280, 1439–1443

8 Nelken, I. *et al.* (1999) Responses of auditory-cortex neurons to structural features of natural sounds. *Nature* 397, 154–157

9 Schnupp, J.W.H. *et al.* (2001) Linear processing of spatial cues in primary auditory cortex. *Nature* 414, 200–204

10 Galuske, R.A.W. *et al.* (2000) Interhemispheric asymmetries of the modular structure in human temporal cortex. *Science* 289, 1946–1949

11 Galaburda, A. and Sanides, F. (1980) Cytoarchitectonic organisation of the human auditory cortex. *J. Comp. Neurol.* 190, 597–610

12 Pandya, D.N. (1995) Anatomy of the auditory cortex. *Rev. Neurol.* 151, 486–494

13 Howard, M.A. *et al.* (2000) Auditory cortex on the human posterior superior temporal gyrus. *J. Comp. Neurol.* 416, 79–92

14 Tardif, E. and Clarke, S. (2001) Intrinsic connectivity in human auditory areas: a tracing study with DiI. *Eur. J. Neurosci.* 13, 1045–1050

15 Karbe, H. *et al.* (1998) Cerebral networks and functional brain asymmetry: evidence from

regional metabolic changes during word repetition. *Brain Lang.* 63, 108–121

16 Warren, J.D. *et al.* (2002) Perception of sound source motion by the human brain. *Neuron* 34, 139–148

17 Maeder, P.P. *et al.* (2001) Distinct pathways involved in sound recognition and localisation: a human fMRI study. *NeuroImage* 14, 802–816

18 Anourova, I. *et al.* (2001) Evidence for dissociation of spatial and nonspatial auditory information processing. *NeuroImage* 14, 1268–1277

19 Rauschecker, J.P. and Tian, B. (2000) Mechanisms and streams for processing of 'what' and 'where' in auditory cortex. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11800–11806

20 Näätänen, R. *et al.* (2001) 'Primitive intelligence' in the auditory cortex. *Trends Neurosci.* 24, 283–288

21 Steinschneider, M. *et al.* (1999) Temporal encoding of the voice onset time phonetic parameter by field potentials recorded directly from human auditory cortex. *J. Neurophysiol.* 82, 2346–2357

22 Liégeois-Chauvel, C. *et al.* (1999) Specialisation of left auditory cortex for speech perception in man depends on temporal coding. *Cereb. Cortex* 9, 484–496

23 Bell, A.J. and Sejnowski, T.J. (1995) An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7, 1129–1159

24 Attias, H. and Schreiner, C.E. (1998) Blind source separation and deconvolution: the dynamic component analysis algorithm. *Neural Comput.* 10, 1373–1424

25 Stone, J.V. (2002) Independent component analysis: an introduction. *Trends Cogn. Sci.* 6, 59–64

26 Wightman, F.L. and Kistler, D.J. (1989) Headphone simulation of free-field listening. II: Psychophysical validation. *J. Acoust. Soc. Am.* 85, 868–878

27 Hofman, P.M. *et al.* (1998) Relearning sound localisation with new ears. *Nat. Neurosci.* 1, 417–421

28 Rao, R.P. and Ballard, D.H. (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87

29 Friston, K.J. and Price, C.J. (2001) Dynamic representations and generative models of brain function. *Brain Res. Bull.* 54, 275–285

30 Recanzone, G.H. (2000) Spatial processing in the auditory cortex of the macaque monkey. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11829–11835

31 Tian, B. *et al.* (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292, 290–293

32 Baumgart, F. *et al.* (1999) A movement-sensitive area in auditory cortex. *Nature* 400, 724–726

33 Lewis, J.W. *et al.* (2000) A comparison of visual and auditory motion processing in human cerebral cortex. *Cereb. Cortex* 10, 873–888

34 Hall, D.A. *et al.* (2002) Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 12, 140–149

35 Giraud, A-L. *et al.* (2000) Representation of the temporal envelope of sounds in the human brain. *J. Neurophysiol.* 84, 1588–1598

36 Binder, J.R. *et al.* (2000) Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528

37 Penhune, V.B. *et al.* (1998) Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm reproduction. *J. Cogn. Neurosci.* 10, 752–765

38 Griffiths, T.D. *et al.* (1999) A common neural substrate for the analysis of pitch and duration pattern in segmented sound? *NeuroReport* 18, 3825–3830

39 Engelien, A. *et al.* (1995) The functional anatomy of recovery from auditory agnosia. A PET study of sound categorisation in a neurological patient and normal controls. *Brain* 118, 1395–1409

40 Belin, P. *et al.* (2000) Voice-selective areas in human auditory cortex. *Nature* 403, 309–312

41 Zatorre, R.J. *et al.* (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–1919

42 Halpern, A.R. and Zatorre, R.J. (1999) When that tune runs through your head: a PET investigation of auditory imagery for familiar meolodies. *Cereb. Cortex* 9, 697–704

43 Griffiths, T.D. (2000) Musical hallucinosis in acquired deafness. Phenomenology and brain substrate. *Brain* 123, 2065–2076

44 Binder, J.R. *et al.* (1996) Function of the left planum temporale in auditory and linguistic processing. *Brain* 119, 1239–1247

45 Giraud, A. *et al.* (2001) Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron* 30, 657–663

46 Vouloumanos, A. *et al.* (2001) Detection of sounds in the auditory stream: event-related fMRI evidence for differential activation to speech and nonspeech. *J. Cogn. Neurosci.* 13, 994–1005

47 Celsis, P. *et al.* (1999) Differential fMRI responses in left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *NeuroImage* 9, 135–144

48 Bischoff-Grethe, A. *et al.* (2000) Conscious and unconscious processing of nonverbal predictability in Wernicke's area. *J. Neurosci.* 20, 1975–1981

49 Shtyrov, Y. *et al.* (2000) Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. *NeuroImage* 12, 657–663

50 McGuire, P.K. *et al.* (1996) Functional neuroanatomy of verbal self-monitoring. *Brain* 119, 907–917

51 Hall, D.A. *et al.* (2000) Modulation and task effects in auditory processing measured using fMRI. *Hum. Brain Mapp.* 10, 107–119

52 Hashimoto, R. *et al.* (2000) Functional differentiation in the human auditory and language areas revealed by a dichotic listening task. *NeuroImage* 12, 147–158

53 Leinonen, L. *et al.* (1980) Functional properties of neurons in the temporo-parietal association cortex of awake monkey. *Exp. Brain Res.* 39, 203–215

54 Howard, R.J. *et al.* (1996) A direct demonstration of functional specialisation within motion-related visual and auditory cortex of the human brain. *Curr. Biol.* 6, 1015–1019

55 Nakada, T. *et al.* (2001) Planum temporale: where spoken and written language meet. *Eur. Neurol.* 46, 121–125

56 Calvert, G.A. *et al.* (1997) Activation of auditory cortex during silent lipreading. *Science* 276, 593–596

57 Petitto, L.A. *et al.* (2000) Speech-like cerebral activity in profoundly deaf people processing signed languages: implications for the neural basis of human language. *Proc. Natl. Acad. Sci. U. S. A.* 97, 13961–13966

58 Jäncke, L. *et al.* (2001) Short term functional plasticity in the human auditory cortex: an fMRI study. *Cognit. Brain Res.* 12, 479–485

59 Schwartz, J. and Tallal, P. (1980) Rate of acoustic change may underlie hemispheric specialisation for speech perception. *Science* 207, 1380–1381

60 Belin, P. and Zatorre, R.J. (2000) 'What', 'where' and 'how' in auditory cortex. *Nat. Neurosci.* 3, 965–966

61 Thivard, L. *et al.* (2000) A cortical region sensitive to auditory spectral motion. *NeuroReport* 11, 2969–2972

62 Tervaniemi, M. *et al.* (2000) Lateralized automatic auditory processing of phonetic versus musical information: a PET study. *Hum. Brain Mapp.* 10, 74–79

63 Zatorre, R.J. *et al.* (1996) Hearing in the mind's ear – a PET investigation of musical imagery and perception. *J. Cogn. Neurosci.* 8, 29–46

64 Perry, D.W. *et al.* (1999) Localization of cerebral activity during simple singing. *NeuroReport* 10, 3979–3984

65 Zatorre, R.J. *et al.* (1992) Lateralisation of phonetic and pitch discrimination in speech processing. *Science* 256, 846–849

66 Jäncke, L. *et al.* (2002) Phonetic perception and the temporal cortex. *NeuroImage* 15, 733–746

67 Middlebrooks, J.C. *et al.* (1994) A panoramic code for sound location by cortical neurons. *Science* 264, 842–844