Volumetric vs. surface-based alignment for localization of auditory cortex activation

Rutvik Desai,* Einat Liebenthal, Edward T. Possing, Eric Waldron, and Jeffrey R. Binder

Department of Neurology, 8701 Watertown Plank Road, MEB 4550, Medical College of Wisconsin, Milwaukee, WI 53226, USA

The high degree of intersubject structural variability in the human brain is an obstacle in combining data across subjects in functional neuroimaging experiments. A common method for aligning individual data is normalization into standard 3D stereotaxic space. Since the inherent geometry of the cortex is that of a 2D sheet, higher precision can potentially be achieved if the intersubject alignment is based on landmarks in this 2D space. To examine the potential advantage of surface-based alignment for localization of auditory cortex activation, and to obtain high-resolution maps of areas activated by speech sounds, fMRI data were analyzed from the left hemisphere of subjects tested with phoneme and tone discrimination tasks. We compared Talairach stereotaxic normalization with two surface-based methods: Landmark Based Warping, in which landmarks in the auditory cortex were chosen manually, and Automated Spherical Warping, in which hemispheres were aligned automatically based on spherical representations of individual and average brains. Examination of group maps generated with these alignment methods revealed superiority of the surface-based alignment in providing precise localization of functional foci and in avoiding mis-registration due to intersubject anatomical variability. Human left hemisphere cortical areas engaged in complex auditory perception appear to lie on the superior temporal gyrus, the dorsal bank of the superior temporal sulcus, and the lateral third of Heschl's gyrus.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Auditory cortex; Talairach normalization; Surface-based alignment

Introduction

A common technique in functional magnetic resonance imaging (fMRI) studies is to compare the location of functional activations under varying conditions in a group of subjects, and to display these activations in the form of a group statistical map. The anatomy of the human brain is highly variable, however, and brain structures vary in their size, shape, position, and relative orientation. This variation is even more significant for pathological brains. This makes the problem of pooling data across different anatomies non-trivial (Rademacher et al., 1993; Roland and Zilles, 1994). In order to compare and overlay anatomies and functional activations, a one-to-one mapping needs to be specified so that each location in one brain corresponds to a unique location in another brain.

A common method for specifying this registration between brains is based on the 3D normalization described by Talairach and Tournoux (1988). This method accounts for differences in brain size and changes in head position inside the scanner across subjects. It is simple to use and is applicable to both cortical and subcortical structures. However, it does not afford a high degree of anatomical accuracy. Several studies have shown that the 3D distance between anatomical landmarks in different brains after Talairach normalization can be over 10 mm (Steinmetz et al., 1990; Thompson and Toga, 1996; Van Essen and Drury, 1997). Thus, precise localization of functional areas with respect to gyral and sulcal landmarks is not possible. Furthermore, due to the highly folded nature of the cortical sheet, the distance between two points in the 3D space is often not representative of the true distance between these points on the 2D cortical sheet. A small inaccuracy resulting from the Talairach normalization may translate into a large inaccuracy in terms of the true distances on the 2D sheet, for example, for points located on the opposite banks of a sulcus.

The basic reason for the inadequacy of the Talairach normalization is that it is based only on linear transformations of scaling, translation, and rotation. If nonlinear transformations such as local dilation, contraction, or shearing can be applied to deform the images, registration can be significantly improved. Considerable research has been directed towards development of nonlinear methods in the last decade (Maintz and Viergever, 1998; Thompson and Toga, 2000; Toga, 1999). High-dimensional morphing methods have been suggested that can morph an entire 3D volume to match the intensity values with a canonical anatomy (Christensen et al., 1996, 1997; Evans et al., 1994; Joshi et al., 1997; Woods et al., 1998). However, the sulcal and gyral landmarks, which are often correlated with function, are a property of the 2D cortical sheet. Intensity-driven high-dimensional morphing does not guarantee the alignment of sulcal and gyral

* Corresponding author. Fax: +1 414 456 6562.
E-mail address: rhdesai@mcw.edu (R. Desai).

Available online on ScienceDirect (www.sciencedirect.com).

patterns; an explicit representation of the surface is needed. Furthermore, anatomical landmarks do not necessarily have a fixed relationship with functional areas. In some cases, it may be desirable to identify the functional areas in individual brains and align them explicitly. Talairach normalization and automated 3D morphing are not suitable for this purpose.

Several surface-based methods allow alignment based explicitly on surface landmarks. These methods optimally transform the cortical sheet to a mathematically simpler shape such as a 2D sheet (Drury et al., 1996, 1997, 1999; Van Essen et al., 1998), an ellipsoid (Sereno et al., 1996), or a sphere (Davatzikos, 1996; Fischl et al., 1999; Thompson and Toga, 1996, 1998; Van Essen, 2004).

Van Essen et al. (1998) modeled the flattened surface as a viscoelastic fluid sheet (Joshi and Miller, 2000). In this approach, landmarks are manually identified on the surface, and the surface is deformed so as to align them closely with the corresponding landmarks on the target surface. The viscoelastic properties of the sheet reduce the distortions in the deformed surface and allow a satisfactory compromise between the objectives of bringing the landmarks into register as closely as possible and minimizing the distortions in the surface. Furthermore, working with 2D surfaces makes the morphing process computationally more tractable than 3D deformation of the entire volume. We refer to this approach as the Landmark Based Warping (LBW) method.

Fischl et al. (1999) provided an automated method for registration based on spherical representation of the cortex. Spherical representation provides a mathematically simple surface suitable for deformation. The surfaces of 40 individual brains were transformed to spherical forms while minimizing metric distortions using an energy function based on convexity. These individual spherical maps were combined to construct an average map of the large-scale folding patterns on a unit sphere. An automated method then non-rigidly aligns the surface of any individual brain, converted to the spherical representation, to this average sphere. We refer to this as the Automated Spherical Warping (ASW) approach.

While there is considerable ongoing research on registration methods, the current practice of fMRI group analysis lags behind to some extent. Our goals in this study were to (1) empirically evaluate and compare the registration based on Talairach transformation with the registration based on two commonly available surface-based methods, LBW and ASW, specifically with respect to auditory cortex activation; and (2) obtain high-resolution maps of the areas activated by speech sounds compared to non-speech sounds. Experiments in the monkey suggest that the auditory cortex may contain 12 or more subdivisions in a relatively small region (Kaas et al., 1999). Thus, intersubject alignment is potentially critical in elucidating the precise topographical arrangement of different auditory fields.

We designed an fMRI experiment to generate blood oxygenation (BOLD) signals in temporal areas involved in speech perception (Binder et al., 2000; Hall et al., 2003; Liebenthal et al., 2003). Subjects performed a two-alternative forced-choice discrimination task with tokens consisting of phonetic sounds (/ba/ and /da/) or tones. By comparing the BOLD signal elicited by the phonetic and tone conditions to a baseline of silence, one can identify areas involved in the perception of speech and complex non-speech sounds, respectively. The phonemes > tones contrast can identify areas that are activated more exclusively by speech sounds.

Note that there are three sources of variance when combining data across subjects: (1) the variance in the location of the anatomical landmarks, (2) the variation in the location of cortical fields with respect to the anatomical landmarks, and (3) the variation in the activation of various cortical fields given a task, and in the location of activation within a cortical field. Alignment methods based on anatomical landmarks (either manually or automatically chosen) can only reduce variance due to source (1). If the variance due to sources (2) and (3) is high, it is difficult to assess the anatomical accuracy of alignment based on functional data (Kang et al., 2004). To better assess and compare alignment accuracy with different methods based on anatomical criteria, we also calculated dispersion indices of several anatomical fiducial points after alignment with different methods.

Materials and methods

Subjects

Participants were 18 healthy adults (8 women), 19–50 years of age, with no history of neurological or hearing impairments. Participants were native speakers of English, and right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). The data from 3 other subjects were excluded due to poor behavioral performance. In accordance with a protocol sanctioned by the Medical College of Wisconsin Institutional Review Board, informed consent was obtained from each subject prior to the experiment.

Test items

The phonetic test items were chosen from an 8-token continuum (ba1–ba8) from /ba/ to /da/. The anchor points, /ba/ and /da/, were synthesized (SenSyn Laboratory Speech Synthesizer, Sensimetrics Corp., Cambridge, MA) using pitch, intensity, formant bandwidth, and formant center frequency parameters derived from natural utterances of /ba/ and /da/ syllables produced by a male speaker and sampled at 44.1 kHz. The continuum was created by continuously varying the second formant (F2) between the anchor points. Three sounds from this continuum, divided into two pairs, were used. The ba2–ba4 pair represented a within-category discrimination between two acoustically different tokens of the same syllable /ba/, while the ba4–ba6 pair involved a discrimination across the phonetic category between the syllables /ba/ and /da/ (across-category discrimination).

For the tone condition, three pure tones were similarly divided into two pairs for discrimination: 1000 Hz–1008 Hz, and 1008 Hz–1016 Hz. For the baseline condition, subjects passively listened to silence. A final condition involved passive listening to white noise bursts. Data from the noise condition will not be presented here. Each phonetic, tone, and noise token was 150 ms long. Stimuli were delivered through Koss ESP-959 electrostatic headphones at 85 dB and were attenuated by approximately 20 dB by the earplugs worn as protection from scanner noise.

Experimental paradigm

The subjects were initially familiarized with the task and the test items through a brief practice session outside the scanner. Subjects performed a two-alternative, forced-choice, ABX dis-

crimination task in which they had to decide whether the third token (X) is identical to the first or the second sound in the preceding AB pair. The inter-stimulus interval between A, B, and X was 500 ms, thus a single trial with three stimuli and two inter-stimulus intervals totaled 1450 ms. Subjects indicated their response by pressing one of two keys. During scanning, one trial of the ABX task was presented in each interval between clustered (or 'sparse') image volume acquisitions, beginning 500 ms after the completion of each acquisition. Forty trials were presented for each of the four experimental conditions (phonemes, tones, noise, baseline), distributed across four runs. The four conditions were presented in a random order within each run. An additional image, acquired at the beginning of each run, was discarded.

Image acquisition and analysis

A 1.5-T GE Signa scanner was used to acquire images. One volume of T2*-weighted, gradient echo, echo-planar images (TE = 40 ms, flip angle = 90°, NEX = 1) was acquired every 8 s. Acquisition time was 2200 ms, leaving 5800 ms of silence between images, during which the auditory stimuli were presented. Volumes were composed of 22 axially-oriented, contiguous slices with $3.75 \times 3.75 \times 4$ mm voxel dimensions. Anatomical images of the entire brain were obtained using a 3D spoiled gradient echo sequence (SPGR) with $0.9 \times 0.9 \times 1.2$ mm voxel dimensions.

Within-subject analysis involved spatial co-registration (Cox and Jesmanowicz, 1999) and voxel-wise multiple linear regression with reference functions representing the three stimulus conditions (phonemes, tones, and noise) compared to a silent baseline. A general linear test was conducted to obtain the phonemes > tones contrast. The individual statistical maps and the anatomical scans were projected into standard stereotaxic space (Talairach and Tournoux, 1988) using the AFNI software package (Cox, 1996).

To account for individual differences in anatomy, smoothing or blurring is common in studies using Talairach transformation. We also created smoothed individual maps by applying a Gaussian filter of 6 mm FWHM.

Surface models and flat maps

Three-dimensional surface models of each individual brain, after projecting the anatomy into Talairach space, were created using FreeSurfer (http://surfer.nmr.mgh.harvard.edu) software. To create a 2D flat map, we selected a region of cortex encompassing perisylvian areas in the left hemisphere that contained superior temporal sulcus (STS) and gyrus (STG), middle temporal gyrus (MTG), Heschl's gyrus (HG), planum temporale (PT), and Sylvian fissure (SF) (see Fig. 1). The patch of inflated surface representing this region was cut and flattened. Working with a patch as opposed to the entire hemisphere has two main advantages. First, since the

patch has a limited amount of curvature, it can be flattened without requiring major cuts to overcome intrinsic curvature. This preserves the topology of the region. This does not hold true for an entire hemisphere, which requires cuts to flatten without major distortions. The placement of cuts changes the topology of the surface, since adjacent points on the 3D surface are not necessarily adjacent on the flat map. A second advantage is that the computational time required to flatten and warp the patch is significantly less than the time required for performing the same operations on an entire hemisphere.

For displaying group data, we chose the template anatomy provided by Holmes et al. (1998), which represents an average of 27 scans of the same individual. This "N27" or "Colin" anatomy was processed in an identical manner by preparing a surface model and cutting and flattening the perisylvian patch.

Volume-to-surface mapping of functional data

To display functional data on surfaces, it is necessary to convert the data from a volumetric format to a surface-based form. The precise nature of this volume-to-surface mapping is important, especially if the group statistics are to be computed using the surfaces. Since functional activation along the entire thickness of the gray matter is mapped to a single node of the surface, it can potentially be lost or smoothed in a variety of ways during the transfer, directly affecting the accuracy of the results. One approach for this volume-to-surface mapping is to consider each node of the 3D surface and assign to it the activation of the voxel with which it intersects. A generalization of this approach is to consider symmetrical neighborhoods of nodes or voxels and calculate the activation with a linearly or nonlinearly weighted average of that neighborhood. This method is suitable for surfaces that are near the middle of gray matter (e.g., layer 4 surfaces). We use the FreeSurfer surface representing the gray matter (GM)/white matter (WM) boundary, however. Considering only the voxels that intersect the surface results in capturing the activation near the boundary, but potentially misses the activation that is closer to the outer GM surface. One option is to use a large neighborhood size to capture the entire thickness of GM, but this results in excessive smoothing and loss of resolution, defeating the purpose of high-resolution group averaging techniques.

Another approach is to calculate normals to the GM/WM boundary surface and sample the voxel that is closest to the center of the GM on this normal (center-of-normal or CN method). This is a commonly used procedure with FreeSurfer, and takes into account the thickness of the GM at each node. However, this has the potential to miss some of the activation in the GM especially when the data are first transformed to a smaller voxel size using linear interpolation. Here, we use a method available as part of the SUMA (http://afni.nimh.nih.gov/afni/suma) software package, in
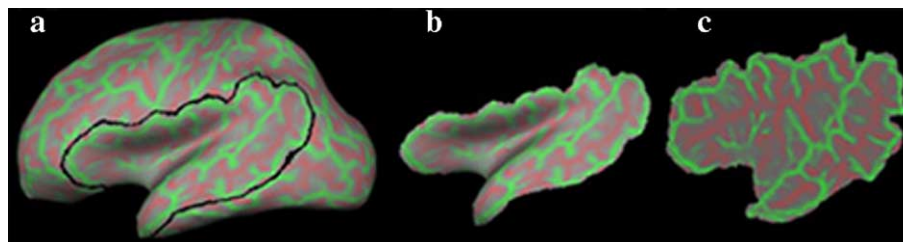

Fig. 1. (a) The inflated surface of a subject with a cut showing the perisylvian patch. (b) The cut 3D patch. (c) The flattened patch.

which activation along the entire GM thickness is mapped. Since the WM and pial surfaces have the same topology, each node on the WM surface has a corresponding node on the pial surface. The maximum activation along this segment is mapped to the surface node (two-surface or TS method). Fig. 2 illustrates an example of a functional volume mapped to the surface using each of the methods in two subjects. Images on the left show activation in the phonemes condition mapped to the subject's own anatomy using the CN method, thresholded at node-wise $P < 0.01$. Images on the right show the same data mapped to the surface using the TS method. More activation in the gray matter is captured by the TS method. Here, we used the TS method, where the absolute maximum value along the segment connecting the two surfaces is mapped to the surface node.

Landmark Based Warping (LBW)

Next, we apply the LBW approach to register the individual brains to the N27 brain. Using Caret (http://brainvis.wustl.edu/caret) (Van Essen et al., 2001), we selected five landmarks on the perisylvian patch to drive the deformation process: (1) HG; (2) Anterior STG, which extends from the anterior tip of STG to the junction of STG and HG; (3) Posterior STG, which extends from the junction of STG and HG to the posterior junction of STG and MTG; (4) MTG; and (5) the gyri bordering the dorsal edge of the patch, whose anterior portion includes the inferior frontal gyrus (IFG) and whose posterior portion includes the supramarginal gyrus. In general, these landmarks are easily identifiable for all subjects, though there is variability especially in the junction points of various landmarks. For example, in some subjects, the posterior

STG and MTG appear to intersect at multiple points. The judgment of the rater was used in such cases to identify unique intersection points. In cases where HG had multiple branches, the branch with higher activation in the tones condition was selected, as that activation is indicative of the location of primary auditory cortex. These landmarks on the template N27 brain and on a subject's brain are shown in Figs. 3a and b, respectively. The individual perisylvian patches, along with the corresponding functional data, were deformed to bring them into register with the landmarks on the N27 brain. The deformed patch for the subject, with the original landmarks, is shown in Fig. 3c.

Automated Spherical Warping (ASW)

We created spherical representations of the individual surfaces and registered them to the canonical spherical brain, using automated procedures implemented in FreeSurfer. To transfer the warped spherical activation to the N27 brain, "standard mesh" surfaces were created for each subject and template brain using utilities in SUMA. The standard mesh surfaces put each node in a given surface in one-to-one correspondence with a node in other standard mesh surfaces.

Group statistics

Using each of the four methods [Talairach (TL), blurred Talairach (TLB), LBW, and ASW], the functional data for each subject were mapped onto the template N27 brain. With the exception of TLB, a small amount of smoothing was applied on the template surface by averaging the activation of each node with that
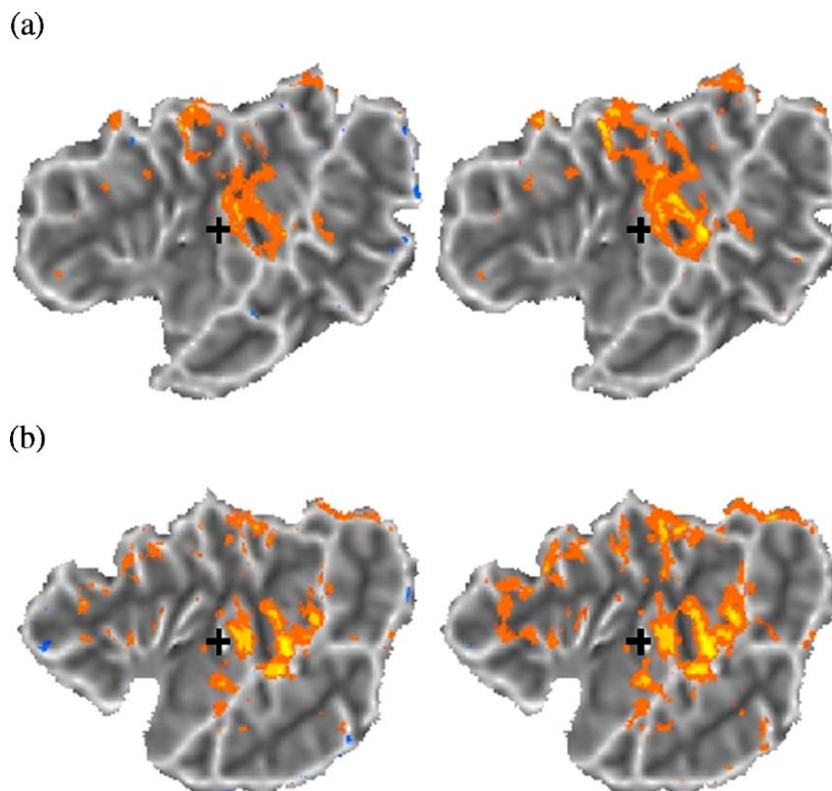

Fig. 2. Comparison of volume-to-surface functional mapping methods. Panels (a) and (b) each show the activation for the phonemes condition thresholded at node-wise $P < 0.01$ for one subject, for center-of-normal (left) and two-surface (right) mapping methods. No cluster size threshold was applied.
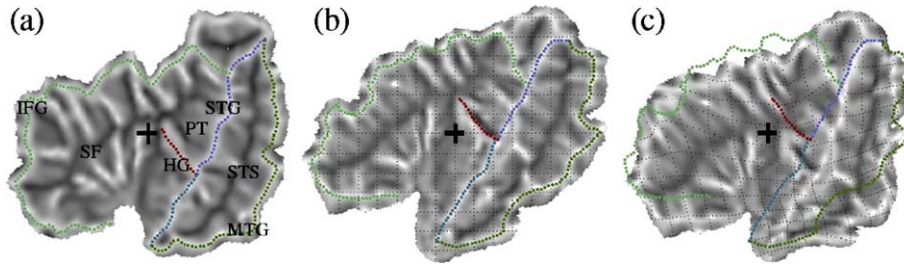
Fig. 3. (a) The flat patch of the template brain with the five anatomical landmarks shown as dotted lines. The areas indicated for reference are Heschl's gyrus (HG), planum temporale (PT), superior temporal gyrus (STG) and sulcus (STS), middle temporal gyrus (MTG), Sylvian fissure (SF), and inferior frontal gyrus (IFG). (b) The patch of a subject with the same landmarks. (c) The subject's patch warped to the template brain. The landmarks are shown in their original position for reference.

of its neighbors. The individual maps were contrasted against a constant value of 0 to create group t maps on a node-by-node basis in a random-effects analysis for all four methods. The resulting t maps were thresholded at node-wise $P < 0.01$, and clusters smaller than 25 mm^2 were removed using Caret.

ROI analysis

To quantitatively compare the activations of various regions obtained using different alignment methods, we selected eight regions of interest (ROIs) on the flat map in functionally active areas based on anatomical landmarks, using Caret (Van Essen et al., 2001). These ROIs, shown in Fig. 4, contain the anterior supramarginal gyrus (aSMG), the planum temporale (PT), medial Heschl's gyrus (mHG), lateral Heschl's gyrus (lHG), the lateral superior temporal gyrus (STG), the dorsal and ventral bank of superior temporal sulcus (dSTS and vSTS, respectively), and the lateral middle temporal gyrus (MTG). The mean activation in each of these ROIs was calculated for each subject. To calculate the activations for ASW, the ROIs were projected onto the standard mesh surface. The creation of standard meshes changes the distribution of nodes in various areas (the areas that get stretched more during the spherical conversion receive more nodes). To account for this change, the ASW ROI averages were calculated by weighting each node according to the volume of GM represented by it, for each subject.

Since LBW is based on manually identifying and aligning anatomical landmarks, to the extent that landmarks have been chosen correctly, it can be assumed to provide accurate intersubject alignment in an anatomical sense, within small tolerance limits (this can be verified by calculating anatomical dispersion of various points, as described in the next section). We compared TL and ASW activation maps with those obtained by LBW with paired two-tailed t tests for each of the ROIs, to assess if there were reliable differences in activation between these methods.

Anatomical dispersion

To compare the anatomical accuracy of alignment obtained with various methods, we selected five fiducial points on the flat patch of each subject, shown in Fig. 5. The points represent the medial tip of HG (mHG), the intersection of HG and STG (HG/STG), the anterior and posterior intersections of STG and MTG (aSTG and pSTG, respectively), and the projection of the midpoint of STG on the MTG (mMTG). Surface coordinates for each point were collected after the TL, LBW, and ASW normalizations. The group mean was calculated for each point, and the average distance from this mean, the dispersion radius (DR), was computed. For ASW, the transformation of the folded surface to the spherical form causes the DR
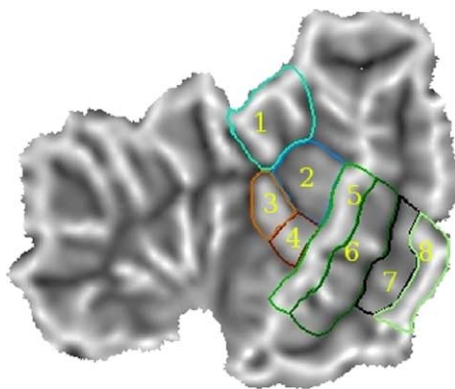

Fig. 4. The ROIs on the flat map. (1) Anterior supramarginal gyrus, (2) planum temporale, (3) medial Heschl's gyrus, (4) lateral Heschl's gyrus, (5) lateral superior temporal gyrus, (6) dorsal bank of superior temporal sulcus, (7) ventral bank of superior temporal sulcus, and (8) lateral middle temporal gyrus.
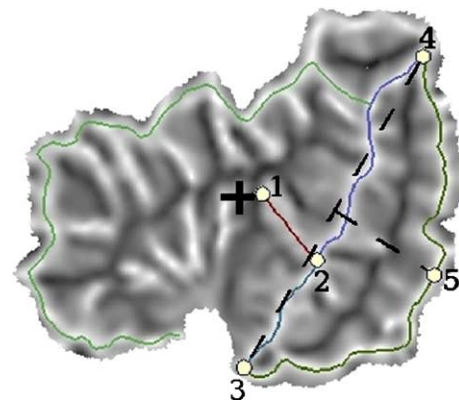

Fig. 5. The five fiducial points chosen for calculating anatomical dispersion. (1) The medial tip of HG, (2) the intersection of HG and STG, (3) the anterior junction of STG and MTG, (4) the posterior junction of STG and MTG, and (5) the projection of midpoint of STG onto the MTG.
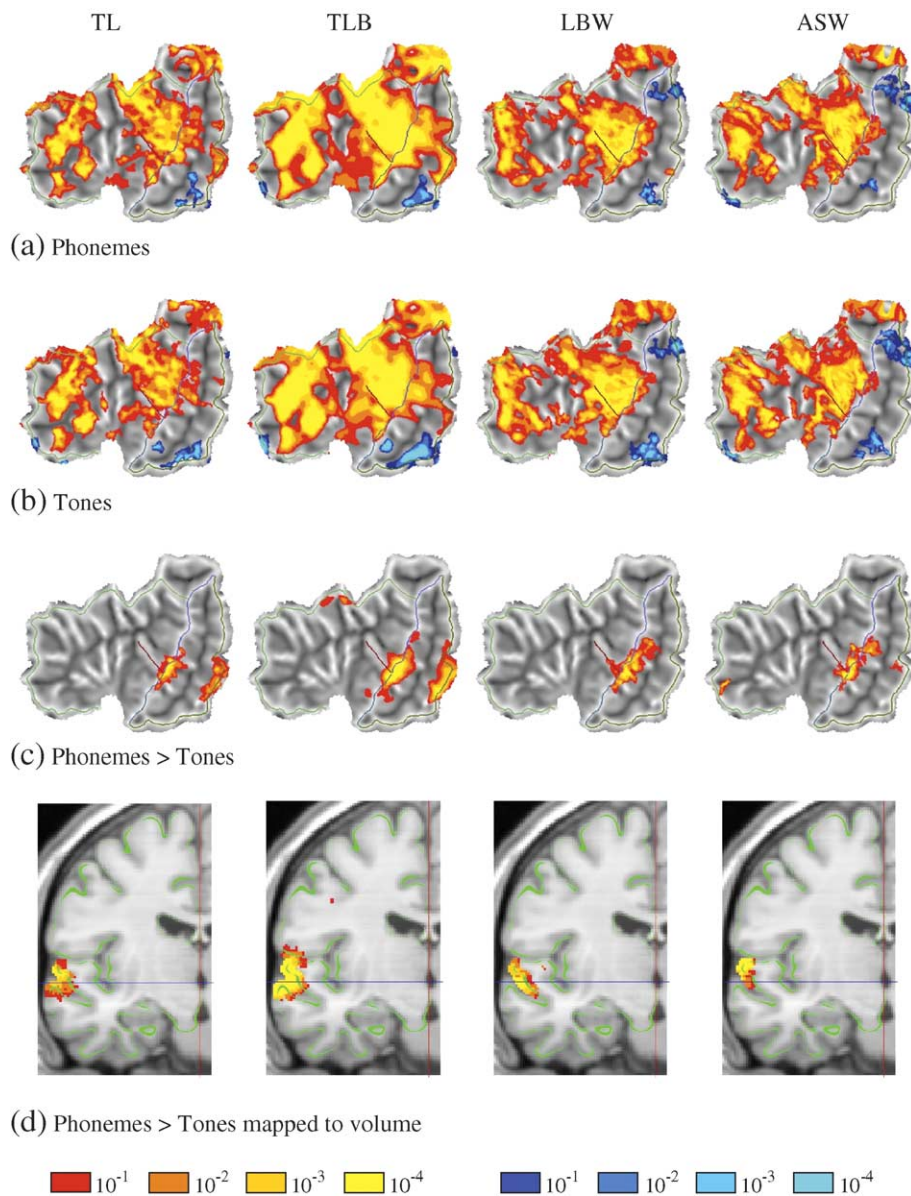
Fig. 6. Group *t* maps for (left to right) Talairach (TL), blurred Talairach (TLB), Landmark Based Warping (LBW), and Automated Spherical Warping (ASW) methods. (a) Phonemes, (b) Tones, and (c) Phonemes > Tones conditions. Panel (d) shows the data in panel (c) mapped to the atlas brain volume. The green line indicates the approximate center of the gray matter. The crosshairs show the stereotaxic *x* (red) and *z* (blue) axis.

to increase due to stretching of portions of the surface. To compensate for this increase, the DR for the warped (registered) spheres in ASW was normalized by multiplying it by the ratio of DR on the original brain with that of the spherical (unregistered) brain for each of the five fiducial points.[1] The ratio of DRs of the original spherical and folded surfaces provides an inflation index that can be used to adjust the morphed spherical DR.

---

[1] Kang et al. (2004) used a similar method for calculation of DR and adjustment for spherical inflation. They used a constant 42% normalization for all fiducial points. We found that the amount of change in DR due to inflation is not uniform but changes depending on the location; hence, a separate normalization ratio was used for each fiducial point.

Results

Behavioral

Subjects performed the task at 66.4% (SD 17.9) accuracy in the tone condition and 65.8% (SD 20.5) accuracy in the phoneme condition. In a two-tailed paired *t* test, the difference in accuracy between the within- and across-category phonemes discrimination was significant (within-accuracy = 52.2%, across-accuracy = 79.4%; $P < 0.001$), while the difference in accuracy between high- and low-frequency tone discrimination was not significant (high-accuracy = 68.1%, low-accuracy = 64.7%; $P > 0.45$). The results are consistent with categorical perception of the phonemes and continuous perception of the tones.
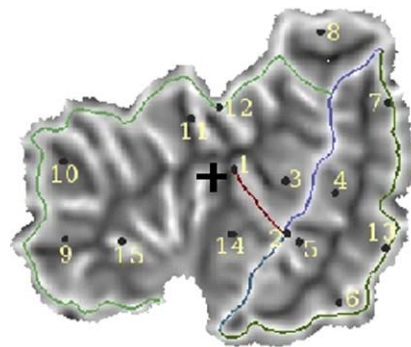
Point	Approx. Brodmann Area(s)	Talairach x, y, z
1	41	-33, -26, +12
2	22	-60, -9, +2
3	41	-50, -25, +8
4	22	-52, -30, +6
5	22	-55, -10, 0
6	21	-55, -12, -17
7	40	-49, -50, +26
8	40	-52, -36, +35
9	45	-39, +27, +6
10	44/45	-50, +16, +19
11	43	-48, -13, +15
12	40	-48, -21, +14
13	21	-63, -24, -2
14	13	-43, -11, -5
15	13	-31, +20, +7

Fig. 7. The Talairach coordinates and approximate Brodmann areas for selected points on the template flat map.

Functional

We present the functional results for all trials in the phonemes (P) and tones (T) conditions compared to the baseline of silence, along with the P > T contrasts, for TL, TLB, LBW, and ASW alignment. For the purpose of comparing the alignment methods, we concentrate on the left hemisphere activations, although some of the areas are activated bilaterally. The results are shown in Fig. 6. Fig. 7 can be used to obtain Talairach coordinates of some of the activation foci. Table 1 shows the results of the ROI analysis. Note that Table 1 shows the ROI analysis on unthresholded data with the mean ROI activations significantly different for from LBW marked, while Fig. 6 shows group *t* maps thresholded at node-wise $P < 0.01$.

With ASW and LBW, both P and T conditions show extensive activation in auditory areas—PT, HG (approximate BA areas 41/ 42), and STG/dorsal bank of STS (BA 22). In addition, areas on the IFG (including the horizontal and vertical ramus and pars

triangularis; BA 44/45), insula (BA 13), postcentral gyrus (BA 43), and posterior supramarginal gyrus (BA 40) are also activated. The dorsal bank of STS is activated more in the P than in the T condition. There is also higher activation for the baseline (or task-induced deactivation) in posterior STS/angular gyrus (BA 39) and in the ventral bank of the anterior STS/MTG.

With TL and TLB methods, HG, PT, and STG/STS are also activated, but these activations are not as strong as with ASW and LBW (see Table 1). Additionally, the ventral bank of STS, central MTG (BA 21), and anterior supramarginal gyrus (BA 40) are also activated. The activation on the IFG in ASW and LBW methods is shifted to somewhat more medial and posterior areas.

In the P > T contrast, ASW and LBW show similar activation patterns. The central STG, the dorsal bank of the STS, and the lateral portion of HG including the HG/STG junction (BA 41/42/ 22) are activated. For ASW, the ventral bank of STS is activated

Table 1

The mean activation (SD) of various ROIs in phonemes, tones, and phonemes > tones conditions with different alignment methods

	aSMG	PT	mHG	lHG	STG	dSTS	vSTS	MTG
Phonemes								
LBW	1.32 (1.31)	2.36 (1.13)	2.17 (0.98)	2.74 (1.12)	2.09 (0.81)	0.53 (0.53)	0.06 (0.33)	0.09 (0.54)
ASW	1.16 (1.06)	2.43 (0.91)	2.83*** (1.11)	3.43* (1.39)	1.85* (1.07)	0.61 (0.65)	0.27** (0.51)	0.51** (0.74)
TL	2.00** (0.83)	2.09 (0.90)	1.09*** (0.77)	1.54*** (1.02)	1.75** (0.74)	0.48 (0.56)	0.24 (0.54)	1.10*** (0.83)
TLB	1.64 (0.69)	1.77*** (0.75)	0.97*** (0.51)	1.29*** (0.76)	1.48*** (0.59)	0.53 (0.45)	0.33* (0.46)	0.88*** (0.65)
Tones								
LBW	1.40 (1.28)	2.47 (1.19)	2.24 (1.02)	2.37 (1.05)	1.53 (0.84)	0.15 (0.50)	−0.09 (0.39)	−0.06 (0.47)
ASW	1.22 (1.10)	2.48 (1.01)	2.86** (1.36)	2.95* (1.53)	1.31 (1.11)	0.21 (0.56)	−0.05 (0.50)	0.31* (0.79)
TL	1.99** (0.92)	2.14* (0.95)	1.08*** (0.74)	1.37*** (0.80)	1.26* (0.73)	0.32 (0.56)	0.03 (0.55)	0.33* (0.71)
TLB	1.64 (0.77)	1.78*** (0.77)	0.99*** (0.47)	1.16*** (0.58)	1.09** (0.60)	0.38 (0.48)	0.16* (0.44)	0.33* (0.55)
P > T								
LBW	−0.04 (0.34)	−0.05 (0.26)	0.00 (0.43)	0.48 (0.57)	0.63 (0.39)	0.40 (0.38)	0.16 (0.36)	0.15 (0.36)
ASW	−0.04 (0.37)	0.01 (0.29)	0.05 (0.47)	0.62 (0.64)	0.62 (0.36)	0.44 (0.41)	0.32* (0.40)	0.20 (0.39)
TL	0.06 (0.34)	−0.01 (0.29)	0.04 (0.37)	0.20* (0.56)	0.57 (0.33)	0.19*** (0.34)	0.22 (0.43)	0.79*** (0.46)
TLB	0.04 (0.27)	0.03 (0.22)	0.00 (0.27)	0.18* (0.36)	0.45** (0.26)	0.18*** (0.25)	0.18 (0.32)	0.57*** (0.35)

aSMG = anterior supramarginal gyrus, PT = planum temporale, mHG = medial Heschl's gyrus, lHG = lateral Heschl's gyrus, STG = lateral superior temporal gyrus, dSTS = dorsal superior temporal sulcus, vSTS = ventral superior temporal sulcus, MTG = lateral middle temporal gyrus. Significant difference in activation from LBW alignment in a paired two-tailed *t* test is marked (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

Table 2
The mean (SD) dispersion radii in millimeters for the three alignment methods

	mHG	HG/STG	aSTG	pSTG	mMTG	Mean
Talairach	3.78 (1.69)	6.91 (2.9)	6.18 (3.89)	10.07 (5.41)	7.85 (3.07)	6.96 (3.39)
LBW	1.03* (0.64)	0.75* (0.44)	1.23* (0.7)	1.12* (0.53)	2.69* (1.68)	1.36* (0.8)
ASW	1.94* (1.18)	4.36* (2.21)	5.78 (4.59)	9.37 (6.66)	6.24 (2.95)	5.54* (3.52)

LBW = Landmark Based Warping, ASW = Automated Spherical Warping. The ASW radii are adjusted for spherical inflation in each of the five locations.
* Indicates a significant reduction in dispersion ($P < 0.05$) over Talairach alignment in a paired one-tailed t test.

more than for LBW. The dorsal bank of STS is activated more the LBW, while the lateral STG is activated more for ASW. No areas are significantly activated for T > P.

The TL and TLB methods also show activation in the STG and dorsal bank of STS, although there is less central/posterior activation in the STS. The lateral HG is not activated significantly, in contrast to LBW and ASW methods. In addition, a large area on the MTG and ventral bank of STS (BA 21) is activated.

Anatomical

The anatomical alignment accuracy results are shown in Table 2. With TL alignment, the mean DR of all five fiducial points is 6.96 mm (SD 3.39 mm). It reduces to 5.54 mm (SD 3.52) after ASW, and to 1.36 mm (SD 0.8) after LBW alignment. There is significant variation between the dispersion radii of the different fiducial points as well. The TL alignment is based on aligning the anterior and posterior commissures, hence the points closer to these medial structures, such as mHG, tend to show less variability, while lateral areas show higher degree of variation. The posterior STS area shows the highest degree of variation with TL alignment, which is not reduced by a large margin with ASW.

Discussion

We observed clear differences between activation patterns obtained with Talairach and surface-based alignment methods. One striking example is in the P > T contrast, where two distinct foci on the STG and MTG[2] are observed in the TL/TLB maps while only one focus on the STG/dorsal STS is observed using surface-based methods. Visual inspection of individual subjects' data confirms that the MTG is not activated for most subjects in the P > T contrast, which is consistent with the LBW P > T group map. When the same data are mapped to Talairach space and then to the atlas surface, some of the activation in the dorsal bank of STS/STG is mapped to the ventral bank of STS/MTG in the atlas brain due to differences in the precise location of STG and MTG in the atlas and the subject, even after Talairach normalization. This is illustrated in Fig. 8 for four different subjects. The STS is more dorsal in the N27 brain than in these subjects, leading to a systematic bias, as shown in Figs. 8a and b. Using volume slices, Fig. 8e illustrates the misalignment of STS between the subject (red contour) and the N27 brain (green contour).

A similar difference between TL/TLB and surface-based maps, caused by differences in the relative locations of the SF, can be observed in the P and T conditions. When the SF in a subject is more dorsal than in the atlas, some activation in the temporal plane may be erroneously mapped to the dorsal bank of SF in the atlas. As shown by the ROI analysis, anterior SMG is activated to a higher degree with TL while lateral and medial HG and PT are activated less strongly. Due to intersubject differences in the precise location of HG along the ventral–dorsal axis as well as along the anterior–posterior axis, there is less group overlap on the HG. As the results in Table 1 suggest, in addition to causing a loss of spatial precision, smoothing does not compensate for anatomical variability, and the basic problem of misalignment of sulcal and gyral structures is not solved. The precise nature of the misalignment depends on the characteristics of the template brain in relation to those of the individual brains. The TLB maps show higher t values than the LBW and ASW maps in a number of regions, although the mean activation in the ROIs with this method is lower (as indicated in Table 1). Isotropic smoothing causes more overlap in activation among subjects in almost all areas, resulting in reduced variance and higher t values, in spite of the imprecision in alignment.

The overall activation patterns with LBW and ASW are similar, but some differences can be observed. In the P and T conditions, medial and lateral HG is activated more strongly with ASW than with LBW, while the lateral STG is activated less strongly in the P condition. The ventral bank of STS and MTG are also activated more strongly in ASW than in LBW, suggesting some misalignment of STS with the ASW method, though to a much smaller degree than with TL. The difference of activation between these two methods in the lateral STG and the dorsal bank of STS can be seen on the flat maps as well as on the volume slices (Fig. 6d). The standard deviation for most ROIs is higher for ASW than for LBW, suggesting more intersubject variability in activation with ASW. Figs. 8c and d show the P > T activation for four subjects mapped to the template brain after LBW and ASW alignment, respectively, for a visual comparison.

Relatively simple landmarks denoted by lines were used here to drive the deformation process. Note that landmarks of any complexity be used instead, depending on the desired accuracy and the identifiability of the landmark structures. For example, more accurate alignment of the ROIs can be achieved by using ROIs themselves as landmarks. Such options are not available in automated methods such as ASW.

The five fiducial points chosen for the anatomical dispersion comparison lie on the alignment landmarks for LBW, so the relatively accurate anatomical alignment with LBW is not surprising. For ASW, the reduction in DR over TL alignment is not significant for three of the five landmarks.

Kang et al. (2004) present a similar comparison of DR for SPM nonlinear volumetric alignment (Ashburner and Friston, 1999),
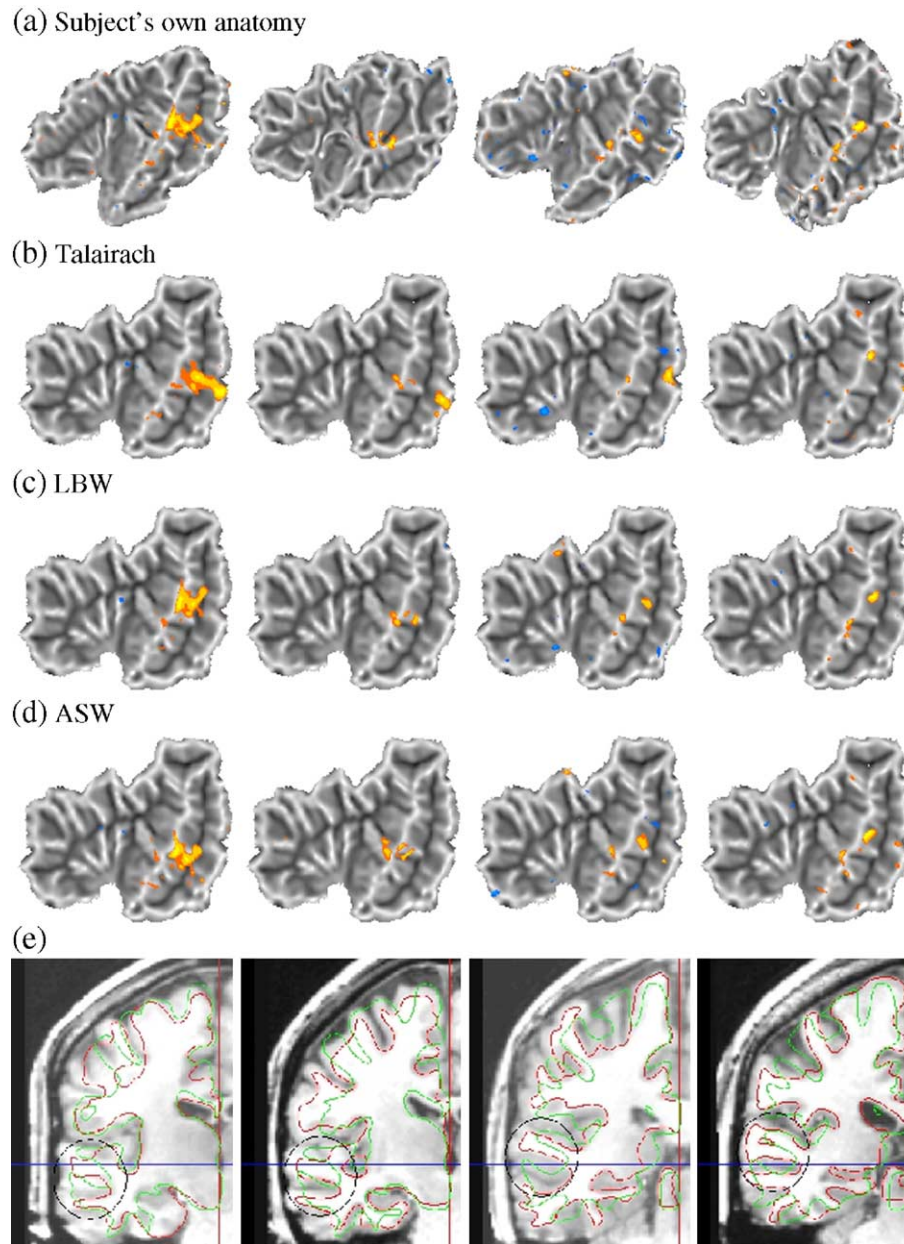
---

[2] On a folded surface, or in a volume, the two foci appear as a single focus covering STG and MTG (BA 21/22). The flattened surface pulls the two gyri apart, making the difference between surface-based and TL methods more clear.

Fig. 8. The P > T contrast for four subjects, thresholded at node-wise $P < 0.05$. Each column represents the data from one subject. (a) The activation mapped to the subject's own anatomy. (b) The same activation mapped to the template brain after Talairach normalization. Some of the activation in the dorsal bank of STS shifts to the ventral bank when mapped to the atlas, due to differences in the relative location of the STS in the atlas and these subjects. (c) The activation mapped to the template brain after the Landmark Based Warping (LBW) alignment. (d) The activation mapped to the template brain after Automated Spherical Warping (ASW) alignment (e) Coronal slices of the subjects' own anatomy. The red contour shows the gray/white matter segmentation boundary for the subject. The same boundary for the template brain is superimposed in green. The difference in the location of the STS between the subject and the template is highlighted with a circle, where the dorsal bank of the STS in the subject is close to the ventral bank of the STS in the template. The crosshairs show the stereotaxic x (red) and z (blue) axis.

ASW, and local landmark-based methods (LLMs). In LLMs, they use single-point landmarks rather than contours, and use thin-plate spline and other low-dimensional surface transformation techniques. The pattern of results presented here is similar to the one in that study, though the radii presented here are somewhat smaller (a DR of 8.8 mm and 7.9 mm for SPM and ASW alignment, respectively, in the Kang et al. study, compared to 6.96 mm for Talairach and 5.54 mm for ASW here). This could be due to the larger sample size (10 in the Kang et al. study vs. 18 in the present

study) or to inter-rater differences in the selection of fiducial points. With a larger sample size, the effect of outliers on the group mean is less, yielding a more robust estimate. The LBW dispersion here (1.36 mm) is also lower than the affine (4.1 mm) and Procrustean LLM dispersion (4.6 mm), which could additionally be due to the characteristics of these warping methods.

It is interesting to relate these anatomical alignment results with the functional group maps. In terms of DR, ASW is more similar to TL than to LBW, but group functional maps suggest that it is more

similar to LBW than to TL. This is possibly due to a high degree of variation in functional activation in relation to the anatomical landmarks, in addition to the variability of the folding patterns. As Kang et al. (2004) point out, if this "activation variance" is high, more accurate anatomical alignment does not have a major effect on functional group maps. By aligning major sulci and gyri, such as STG and SF, ASW prevents coarse-grain mis-registration of functional data (e.g., the MTG focus in the P > T contrast obtained with TL), though at a finer grain the anatomical alignment may not be as accurate as LBW.

Auditory processing

We briefly discuss the activation patterns from a neuro-biological perspective; a more detailed discussion will be presented elsewhere. The P and T conditions activate HG and PT, as expected. The medial one half to two thirds of HG is the site of primary auditory cortex and is activated, as expected, by both phonemes and tones. The activation of PT by tones, in addition to phonemes, adds to the growing evidence that PT is involved in the perception of complex sounds as opposed to being dedicated for speech perception (Binder et al., 1996; Griffiths and Warren, 2002). The P > T contrast suggests that left hemisphere areas potentially involved in speech perception lie on the STG, the dorsal bank of STS, and the lateral third of HG, though some of these areas may simply be involved in the perception of sounds with complex spectro-temporal variations. Nevertheless, these results support the view that areas in the left temporal plane are involved in the analysis of acoustic properties of sounds and more ventral areas in the upper STS are involved in phonetic or categorical perception (Binder et al., 2000; Liebenthal et al., 2003; Scott et al., 2000). The involvement of the dorsal bank of left STS in speech perception is also supported by physiological studies in non-human primates, suggesting auditory sensitivity in this region but not in the ventral bank of STS (Baylis et al., 1987; Poremba et al., 2003; Seltzer and Pandya, 1994).

A number of studies have demonstrated activation of inferior frontal regions (BA 44/45) during auditory perception (Binder et al., 2000; Fiez et al., 1995), but the precise role of these areas remains unclear. Here, these areas were activated for both P and T conditions when performing the ABX task. The baseline of silence only involves passive listening; hence, one hypothesis is that this activation is task-related and results from short-term maintenance of stimuli, attention, or response selection processes.

Conclusions

Surface-based intersubject alignment methods, based on non-linear transformations of surfaces to align (manually or automatically chosen) anatomical landmarks, provide a more accurate method of obtaining group data than traditional volumetric alignment based on affine transformations, at least for perisylvian regions. Smoothing of individual functional data can help in increasing overlap between activation in misaligned structures, but it also lowers spatial precision and potentially results in false positive activation.

The ASW method provides greater accuracy than Talairach alignment, preventing coarse misalignment of structures. With ASW, it is not necessary to place cuts in the surface that introduce distortion, and user interaction and training in identifying land-marks are not necessary. The disadvantage of ASW is in the potential lack of significant improvement over Talairach alignment, especially in regions with high anatomical variability such as the left angular gyrus. ASW also aligns the entire hemisphere, which can prevent accurate alignment of a smaller ROI due to variability in other regions. Higher accuracy, with approximately 2 mm or less dispersion, can be achieved with LBW, at the cost of user interaction in the form of identification of landmarks. An additional advantage of LBW is that the alignment need not be restricted to anatomical information only; functional information can be readily incorporated as an alignment metric if necessary. An intermediate approach between LBW and ASW is also clearly possible, in which landmarks are manually identified on an inflated sphere, eliminating the need for cuts. The main disadvantage of surface-based methods is the human and computational effort required in the generation of surface models.

Left hemisphere areas involved in the perception of complex sounds, including speech, lie on the lateral STG, the dorsal bank of STS, and the lateral third of HG.

Acknowledgments

References

Ashburner, J., Friston, K.J., 1999. Nonlinear spatial normalization using basis functions. Hum. Brain Mapp. 7, 254–266.

Baylis, G.C., Rolls, E.T., Leonard, C.M., 1987. Functional subdivisions of the temporal lobe neocortex. J. Neurosci. 7, 330–342.

Binder, J.R., Frost, J.A., Hammeke, T.A., Rao, S.M., Cox, R.W., 1996. Function of the left planum temporale in auditory and linguistic processing. Brain 119, 1239–1247.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and nonspeech sounds. Cereb. Cortex 10, 512–528.

Christensen, G.E., Rabbitt, R.D., Miller, M.I., 1996. Deformable templates using large deformation kinematics. IEEE Trans. Image Processing 5, 1435–1447.

Christensen, G.E., Joshi, S.C., Miller, M.I., 1997. Volumetric trans-formation of brain anatomy. IEEE Trans. Med. Imaging 16, 864–877.

Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.

Cox, R.W., Jesmanowicz, A., 1999. Real-time 3D image registration of functional MRI. Magn. Reson. Med. 42, 1014–1018.

Davatzikos, C., 1996. Spatial normalization of 3D brain images using deformable models. J. Comput. Assist. Tomogr. 20, 656–665.

Drury, H.A., Van Essen, D.C., Anderson, C.H., Lee, C.W., Coogan, T.A., Lewis, J.W., 1996. Computerized mappings of the cerebral cortex: a multiresolution flattening method and a surface-based coordinate system. J. Cogn. Neurosci. 8, 1–28.

Drury, H.A., Van Essen, D.C., Snyder, A.Z., Shulman, G.L., Akbudak, E., Ollinger, J.M., Conturo, T.E., Raichle, M., Corbetta, M., 1997. Warping

fMRI activation patterns onto the visible man atlas using fluid deformations of cortical flat maps. NeuroImage 5, S421.

Drury, H.A., Van Essen, D.C., Corbetta, M., Snyder, A.Z., 1999. Surface-based analyses of the human cerebral cortex. In: Toga, A.W. (Ed.), Brain Warping. Academic Press, pp. 337–363.

Evans, A., Collins, D.L., Neelin, P., MacDonald, D., Kamber, M., Marrett, T.S., 1994. Three-dimensional correlative imaging: applications in human brain mapping. In: Thatcher, R., Hallett, M., Zeffiro, T., John, E., Huerta, M. (Eds.), Functional Neuroimaging: Technical Foundations. Academic Press, pp. 145–162.

Fiez, J., Raichle, M.E., Miezin, F.M., Petersen, S.E., Tallal, P., Katz, W.F., 1995. Studies of the auditory and phonological processing: effects of stimulus characteristics and task demands. J. Cogn. Neurosci. 7, 357–375.

Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp. 8, 272–284.

Griffiths, T.D., Warren, J.D., 2002. The planum temporale as a computational hub. Trends Neurosci. 25, 348–353.

Hall, D.A., Hart, H.C., Johnsrude, I.S., 2003. Relationships between human auditory cortical structure and function. Audiol. Neuro-Otol. 8, 1–18.

Holmes, C.J., Hoge, R., Collins, L., Woods, R., Toga, A.W., Evans, A.C., 1998. Enhancement of MR images using registration for signal averaging. J. Comput. Assist. Tomogr. 22, 324–333.

Joshi, S.C., Miller, M.I., 2000. Landmark matching via large deformation diffeomorphisms. IEEE Trans. Image Processing 9, 1357–1370.

Joshi, S., Miller, M.I., Grenander, U., 1997. On the geometry and shape of brain sub-manifolds. Int. J. Pattern Recogn. Artif. Intell. 11, 1317–1343.

Kaas, J.H., Hackett, T.A., Tramo, M.J., 1999. Auditory processing in primate cerebral cortex. Curr. Opin. Neurobiol. 9, 164–170.

Kang, X., Bertrand, O., Kimmo, A., Yund, E.W., Herron, T., Woods, D.L., 2004. Local landmark-based mapping of human auditory cortex. NeuroImage 22, 1657–1670.

Liebenthal, E., Binder, J.R., Spitzer, S.M., Possing, E.T., Medler, D.A., 2003. Neural correlates of categorical perception of speech. J. Cogn. Neurosci., 103–104 (Suppl.)

Maintz, J.B., Viergever, M.A., 1998. A survey of medical image registration. Med. Image Anal. 2, 1–36.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.

Poremba, A., Saunders, R.C., Crane, A.M., Cook, M., Sokoloff, L., Mishkin, M., 2003. Functional mapping of the primate auditory system. Science 299, 568–572.

Rademacher, J., Caviness Jr., V.S., Steinmetz, H., Galaburda, A.M., 1993. Topographical variation of the human primary cortices: implications for neuroimaging, brain mapping, and neurobiology. Cereb. Cortex 3, 313–329.

Roland, P.E., Zilles, K., 1994. Brain atlases—a new research tool. Trends Neurosci. 17, 458–467.

Scott, S.K., Blank, C.C., Rosen, S., Wise, R.J.S., 2000. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123, 2400–2406.

Seltzer, B., Pandya, D.N., 1994. Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. J. Comp. Neurol. 343, 445–463.

Sereno, M.I., Dale, A.M., Liu, A., Tootell, R.B., 1996. A surface-based coordinate system for a canonical cortex. NeuroImage 3, S252.

Steinmetz, H., Furst, G., Freund, H., 1990. Variation of perisylvian and calcarine anatomic landmarks within stereotaxic proportional coordinates. Am. J. Neuroradiol. 11, 1123–1130.

Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain. Thieme Medical, New York.

Thompson, P.M., Toga, A.W., 1996. A surface-based technique for warping 3-dimensional images of the brain. IEEE Trans. Med. Imaging 15, 1–16.

Thompson, P.M., Toga, A.W., 1998. Anatomically-driven strategies for high-dimensional brain image registration and pathology detection. In: Toga, A.W. (Ed.), Brain Warping. Academic Press, pp. 311–336.

Thompson, P.M., Toga, A.W., 2000. Elastic image registration and pathology detection. In: Bankman, I., Rangayyan, R., Evans, A., Woods, R., Fishman, E., Huang, H. (Eds.), Handbook of Medical Image Processing. Academic Press.

Toga, A.W., 1999. Brain Warping. Academic Press.

Van Essen, D.C., 2004. Surface-based approaches to spatial localization and registration in primate cerebral cortex. NeuroImage 23 (Suppl. 1), S97–S107.

Van Essen, D.C., Drury, H.A., 1997. Structural and functional analyses of human cerebral cortex using a surface-based atlas. J. Neurosci. 17, 7079–7102.

Van Essen, D.C., Drury, H.A., Joshi, S., Miller, M.I., 1998. Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. Proc. Natl. Acad. Sci. U. S. A. 95, 788–795.

Van Essen, D.C., Drury, H.A., Dickson, J., Harwell, J., Hanlon, D., Anderson, C.H., 2001. An integrated software suite for surface-based analyses of cerebral cortex. J. Am. Med. Inform. Assoc. 8, 443–459.

Woods, R.P., Grafton, S.T., Holmes, C.J., Cherry, S.R., Mazziotta, J.C., 1998. Automated image registration: I. General methods and intrasubject, intramodality validation. J. Comput. Assist. Tomogr. 22, 139–152.