Cerebral Cortex July 2012;22:1593–1603
doi:10.1093/cercor/bhr205
Advance Access publication September 5, 2011

Is the Link between Anatomical Structure and Function Equally Strong at All Cognitive Levels of Processing?

Amir M. Tahmasebi[1], Matthew H. Davis[2], Conor J. Wild[3], Jennifer M. Rodd[4], Hélène Hakyemez[5], P. Abolmaesumi[6] and Ingrid S. Johnsrude[3,5,7]

[1]Rotman Research Institute, Baycrest Centre for Geriatric Care, Toronto, Ontario, Canada M6A 2E1, [2]Medical Research Council Cognition and Brain Sciences Unit, Cambridge CB2 2EF, UK, [3]Centre for Neuroscience Studies, Queen's University, Kingston, Ontario, Canada K7L 3N6, [4]Department of Cognitive, Perceptual and Brain Sciences, University College London, London WC1H 0AP, UK, [5]Department of Psychology, Queen's University, Kingston, Ontario, Canada K7L 3N6, [6]Department of Electrical and Computer Engineering, the University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4 and [7]Linnaeus Centre for Hearing and Deafness, Linköping University, SE-581 83 Linköping, Sweden

Address correspondence to Amir M. Tahmasebi, Rotman Research Institute, Baycrest Centre for Geriatric Care, 3560 Bathurst Street, Toronto, Ontario, Canada M6A 2E1. Email: atahmasebi@rotman-baycrest.on.ca or ingrid.johnsrude@queensu.ca.

Whereas low-level sensory processes can be linked to macroanatomy with great confidence, the degree to which high-level cognitive processes map onto anatomy is less clear. If function respects anatomy, more accurate intersubject anatomical registration should result in better functional alignment. Here, we use auditory functional magnetic resonance imaging and compare the effectiveness of affine and nonlinear registration methods for aligning anatomy and functional activation across subjects. Anatomical alignment was measured using normalized cross-correlation within functionally defined regions of interest. Functional overlap was assessed using *t*-statistics from the group analyses and the degree to which group statistics predict high and consistent signal change in individual data sets. In regions related to early stages of auditory processing, nonlinear registration resulted in more accurate anatomical registration and stronger functional overlap among subjects compared with affine. In frontal and temporal areas reflecting high-level processing of linguistic meaning, nonlinear registration also improved the accuracy of anatomical registration. However, functional overlap across subjects was not enhanced in these regions. Therefore, functional organization, relative to anatomy, is more variable in the frontal and temporal areas supporting meaning-based processes than in areas devoted to sensory/perceptual auditory processing. This demonstrates for the first time that functional variability increases systematically between regions supporting lower and higher cognitive processes.

Keywords: anatomy, fMRI, functional variability, hearing, image normalization, language, lexical ambiguity, semantic processing, speech perception, structure–function relationships

Introduction

In relating function to anatomical structure in the human neocortex, it is often assumed that, for any given task, different individuals will recruit the same set of neural processes and that these are located in the same (anatomically defined) regions of neocortex. Accordingly, functional imaging data from different participants are typically combined after spatial registration, which aligns anatomical structures, such as sulci and gyri. However, researchers have observed that, despite reasonably accurate alignment of anatomy (as accomplished using conventional whole-brain spatial registration approaches), the localization of activation can vary from individual to individual for a given task (e.g., Fox and Pardo 1991; Watson et al. 1993; Schlaug et al. 1994; Xiong et al. 2000; Hasnain et al. 2001; Fedorenko et al. 2010). The degree of intersubject variability in location may depend on the cognitive "level" of the processes under study, with greater variability in location for higher level compared with primary sensory and perceptual processes.

Converging evidence from intraoperative stimulation mapping (e.g., Penfield and Rasmussen 1950), studies of microanatomical structure (e.g., Geyer et al. 2000; Grefkes et al. 2001; Rademacher et al. 2001; Hinds et al. 2008, 2009), and functional neuroimaging conclusively tie primary cortical regions to morphological landmarks in cortex. For example, the banks of the calcarine sulcus are consistently activated in functional magnetic resonance imaging (fMRI) studies that recruit primary visual cortex (e.g., DeYoe et al. 1996; Dougherty et al. 2003; Wohlschlager et al. 2005; see Iaria and Petrides 2007, for review), and the anterior bank of the central sulcus is consistently activated in a somatotopic fashion when primary motor cortex is recruited (see Chouinard and Paus 2006, for review). In fact, there is a gross morphological landmark for the motor hand area: the "hand knob" (Yousry et al. 1997). Touch stimuli consistently activate the posterior bank of the central sulcus somatotopically (Boling and Olivier 2004), and auditory stimuli elicit activity in the region of Heschl's gyrus, the morphological landmark for primary auditory cortex in humans (Rademacher et al. 1992; Morosan et al. 2001).

It is not apparent whether higher order, cognitive, functions are as clearly linked to morphological landmarks as sensory and motor functions are. In fact, evidence from intraoperative stimulation mapping and neuroimaging suggests substantial variability among individuals in the location of neocortical regions involved in higher levels of cognitive processing, such as high-order motor control (Schlaug et al. 1994) or complex language tasks (Ojemann et al. 1989; Seghier et al. 2008). Ojemann et al. (1989) mapped language function in 117 patients undergoing surgery, by electrically stimulating focal regions of left posterior frontal and temporal cortex while patients named line drawings. Regions critically involved in naming (indicated by naming errors committed during stimulation of these regions) were quite small in individual patients (1–2 cm^2 in area) and varied markedly in location among individuals. However, it is not clear whether the variability noted by Ojemann et al. (1989) is primarily anatomical or functional in nature.

One approach to this question of whether intersubject variability increases for higher order processes compared with

lower sensory/perceptual processing has been to examine between-subject variability in the location of peak activation foci in functional positron emission tomography studies among subjects (Fox and Pardo 1991; Xiong et al. 2000). Results from these studies suggested that foci in higher order processing regions are not more variable among individuals than foci reflecting lower level processes.

These negative findings may be explained by methodological limitations. First, previous studies utilized affine registration, and so there is probably considerable residual anatomical variability between individuals that could mask any between-region differences of interest (Watson et al. 1993). Second, these studies selected and examined only the largest peak within functional clusters while ignoring statistically smaller peaks that may have been located closer to activations in other subjects. Choosing the peak activation focus is expedient and objective but may not result in the selection of the most appropriate focus of activation in each subject. Furthermore, the selection of the strongest peak from an individual's statistical map is misleading in that it reflects both the magnitude of signal change (the numerator of the individual *t*-statistic, which is considered in the group analysis) as well as the variability in this value over observations (the denominator; which is not considered in the group analysis). A high individual *t*-value in a given voxel can result either from high signal change in that voxel or low variability in that voxel, or a combination of both. The former factor is relevant for group overlap; the latter is not. We propose that it would be more accurate to assess functional overlap among individuals in a way that did not take the coordinates of the peak voxel as the proxy for a cluster of activation within a subject.

Our approach is to compare the overlap in functional activation across subjects resulting from nonlinear registration methods, which accurately align anatomical features across subjects, with the overlap resulting from affine registration, which results in substantial residual anatomical variability (see e.g., Watson et al. 1993). We predict that when there is a close relationship between anatomical landmarks and function (e.g., primary sensory cortices), the use of anatomically more accurate methods should improve functional overlap among subjects. In contrast, in brain regions where function is less closely tied to anatomy (e.g., regions supporting higher order "cognitive" processes), we should see no clear advantage associated with improved anatomical alignment. We assess functional overlap in 2 ways. First, we avoid the problems inherent in selecting peaks in individuals by turning the problem on its head. Instead of asking "how does individual data predict the group peak?," we ask, "how well does the group data predict individual data?" If functional overlap is good, then the peak from a group analysis should consistently predict a location of high contrast values in independent individual data sets. We also compare the magnitudes of the overall group *t*-statistics resulting from the different normalization methods—*t*-statistics directly index both the strength of contrast values and the consistency of contrast values across subjects (see also Tahmasebi et al. 2009).

To examine functional patterns of activity in areas supporting lower and higher levels of processing, we look to the human auditory system, which, like the nonhuman primate auditory system (Romanski et al. 1999; Kaas and Hackett 2000; Hackett 2011) is hierarchically organized (Davis and Johnsrude 2003, 2007; Scott and Johnsrude 2003; Rodd et al. 2005; Hickok and Poeppel 2007; Peelle et al. 2010; Price 2010; Lerner et al. 2011). Auditory signals from the periphery are relayed from brainstem to the medial geniculate nucleus of the thalamus. From there, signals are relayed to primary auditory cortex and hence to other temporal regions, and then at a third or later stage, to frontal regions, with cognitively more complex processes dependent on regions further away from primary auditory cortex (Davis and Johnsrude 2003; Rodd et al. 2005; Davis et al. 2007; Okada et al. 2010; Peelle et al. 2010).

To differentially activate regions of this network, we adopt an fMRI paradigm that can elicit activity at both "lower" and "higher" levels of the human auditory system. Unintelligible sentence-length sounds ("signal-correlated noise [SCN]") and clear sentences activate auditory thalamus and cortex, relative to silence. Sentences that contain words with more than one meaning (e.g., the *shell* was *fired* toward the *tank*) selectively activate left inferior frontal and left posterior temporal cortex compared with matched sentences without ambiguous words (e.g., her secrets were written in her diary) (Rodd et al. 2005; Davis et al. 2007; Coleman et al. 2009). Interestingly, these meaning-sensitive frontal and posterior temporal regions are much more affected by alterations in level of awareness than are sound- or speech-sensitive regions closer to auditory cortex (Davis et al. 2007; Coleman et al. 2009) consistent with them reflecting cognitively higher processing. We predict that correspondence between function and gross anatomy will be highly consistent among individuals at the level of the thalamus and primary auditory cortex but will be less consistent in these anterior and posterior frontal and temporal regions.

Materials and Methods

Image Acquisition

The fMRI study was conducted at the Centre for Neuroscience Studies at Queen's University, Kingston, ON, Canada. Twenty-eight volunteers (21 females, aged 18–32, right-handed, native English speakers) were scanned. None of the subjects had any history of head injury or neurological illness. All subjects gave written informed consent for their participation. The experimental protocol was approved by the Queen's University Health Sciences Research Ethics Board. Four data sets (all from women) were excluded due to severe head motion during scanning.

MR imaging was performed on a 3.0-T Siemens Trio MRI scanner. T_2^*-weighted Gradient-echo Echo-planar imaging sequences were acquired using a sparse-imaging procedure (Hall et al. 1999) with an 8.04-s silent period between successive scans, permitting auditory stimulus presentation in silence. An interleaved slice acquisition protocol with a typical field of view (FOV) of 211 × 211 mm^2, in-plane resolution of 3.3 × 3.3 mm^2, slice thickness of 4.1 mm (including a 25% gap), acquisition time = 1.96 s, time echo [TE] = 30 ms, and time repetition [TR] = 10 s was used. Additionally, high-resolution T_1-weighted anatomical images were acquired from all volunteers using 3D magnetization prepared rapid gradient echo sequence (TR = 1760 ms, TE = 2.6 ms, flip angle = 9°, and voxel resolution of 1.0 mm^3).

Functional Paradigm

The fMRI study was modeled using methods derived from Rodd et al. (2005). Stimuli included both high- and low-ambiguity sentences, speech-shaped noise matched to the sentences on duration, and silent trials. The ambiguity factor (high vs. low) was crossed with a noise factor (present vs. absent), for reasons that are not relevant to this study; the effect of the noise factor, and interactions with this factor, shall not be discussed here.

The sentences were the same as those used by Rodd et al. (2005). Sentences were 6–18 words long. Fifty-nine items were high-ambiguity sentences that contained at least 2 ambiguous words (e.g., His new *post* was in *China*). The ambiguous words were either homophones (which

have at least 2 meanings with different spelling, e.g., knight/night) or homonyms (2 meanings with the same spelling and pronunciation, e.g., bark). Each high-ambiguity sentence was matched to a low-ambiguity sentence with the same number of words and the same syntactic structure but without homophones (e.g., The old tree was in danger). The 2 sets of sentences were matched on number of syllables, number of words, rated naturalness, rated imageability, and the log-transformed mean frequency of the content words in the CELEX database (Baayen et al. 1995). See Rodd et al. (2005) for further information. The sentences were recorded by a single female talker with a native central/eastern Canadian accent, digitized at a sampling rate of 22.1 kHz, using Audition software (Adobe Systems). The duration of the individual sentences ranged from 1.5 to 4.9 s. Fifty-nine SCN (Schroeder 1968) trials and 59 silent (rest; R) trials were scanned. SCN has the same spectral profile and amplitude envelope as the original speech but consists entirely of noise and is totally unintelligible.

As in Rodd et al. (2005), we included a simple relatedness judgment task to ensure that subjects were paying attention to the materials. Each sentence has a "probe word" that was either semantically related (50%) or unrelated (50%) to the sentence's meaning (Rodd et al. 2005). These words were matched across ambiguity condition for length, frequency, and semantic relatedness to the sentence. Probes were either clearly related or unrelated to the meaning of the sentences and were not related to the unintended meanings of the ambiguous words. The task was only present on half the trials of any condition, counterbalanced across participants. When a probe word appeared on the screen at the end of half of the trials, right before the scan, volunteers were instructed to make a button press to indicate whether the word was related to the meaning of the sentence or not (right index finger or thumb for related, left index finger or thumb for unrelated). In the SCN control condition, the words "right" and "left" were presented as probe words on half the trials (counterbalanced across participants), and volunteers were instructed to press the appropriate button. No effect of this task was observed when probe-absent trials were compared with probe-present trials and so this factor will not be discussed. Half of the 59 high- and 59 low-ambiguity sentences were presented clearly (referred to as HA – N; LA – N), whereas the rest were presented with some background noise at –2 dB SNR (referred to as HA + N; LA + N); these sentences are still highly intelligible (>80% of words reported correctly).

The experiment was divided into 4 sessions of 61 scans (10 min), where the first 2 scans of each session were discarded to allow T_1 signal to equilibrate. A sparse-imaging protocol (Hall et al. 1999) with a TR of 10 and 2 s volume acquisition time was used to image each of 6 different conditions. Stimuli were presented in the 8.04-s silent period between successive scans.

Functional data were motion corrected using SPM8 (Wellcome Department of Imaging Neuroscience, London, UK). Structural MRI data were edited to remove skull and scalp using the Brain Extraction Tool (BET) of the FSL software package (Oxford Centre for Functional MRI, Oxford University, UK) and manual corrections. Stripped structural data were rigidly registered to the motion-corrected functional time series using the Mutual Information coregistration tool of SPM8. Structural (and coregistered functional) data were spatially normalized using 4 different registration techniques, as described in the next section.

Structural Registration

After coregistration with the functional data sets, all 24 structural MR volumes in native space were first affinely (12 parameters) registered to a common reference frame in Montreal Neurological Institute (MNI) space (i.e., Colin27 or CJH27, Holmes et al. 1998) using the SPM8 toolbox. This was done in order to align the volume centers and sizes of all the brains, before further (nonlinear) registration. Note that subsequent (nonlinear) normalization methods used either ICBM452 or the CJH27 templates; these are both in MNI space and have the same brain size, centre, and orientation.

Affinely registered data were further normalized using 3 registration methods. We used multiple methods to ensure that results can be generalized: The different nonlinear normalization methods had different warping power (i.e., deformations with different degrees of freedom) and worked in different ways:

1. DARTEL (Diffeomorphic Anatomical Registration through Exponential Lie Algebra; Ashburner 2007) is an algorithm for diffeomorphic image registration with 6.4×10^6 deformation parameters. DARTEL is a groupwise technique and so subject data are transformed to a group space that is not necessarily identical to the MNI coordinate frame. Therefore, DARTEL-based warped data were normalized to MNI space using a second affine transformation (SPM8).
2. HAMMER (Shen and Davatzikos 2002) is an elastic registration technique that utilizes an attribute vector for every voxel of the image (the number of deformation parameters is equal to the number of voxels in the image). The attribute vector expresses the geometric features, which are calculated from the tissue maps to reflect underlying anatomy at different scales. Our application of the HAMMER algorithm proceeded in 2 steps: first, in order to generate the tissue map, the brain images were segmented into gray matter, white matter, and cerebrospinal fluid using FMRIB's Automated Segmentation Tool (FAST) of the FSL software package. Second, HAMMER registration was applied to warp the brain images to the Colin27 brain (Holmes et al. 1998). HAMMER uses every voxel's information in its hierarchical multiresolution approach, and so the number of parameters for the deformation field is equal to the number of voxels within the volume.
3. Unified segmentation is a nonlinear registration algorithm first implemented in SPM5 (Ashburner and Friston 2005). This method adopts a probabilistic framework that integrates image registration, tissue classification, and intensity bias correction within the same generative model. The number of deformation parameters is on the order of 10^3: The exact number depends on the image FOV (John Ashburner, communication on SPM email discussion list, 24 March 2009). The unified segmentation technique requires tissue probability maps as priors: here, the ICBM452 tissue probability maps were used.

Affine registration with 12 parameters can only afford global alignment of structure (size matching), whereas both DARTEL and HAMMER claim accurate matching of cortical and subcortical folding patterns while preserving topology. Unified segmentation has a lower deformation power (fewer degrees of freedom) compared with DARTEL and HAMMER.

Data were not smoothed after registration, and analysis was conducted using SPM8. A separate design matrix was created for each participant: this included one condition-specific regressor for each of the 5 sound conditions. (Rest was not explicitly modeled). Regressors were created by convolving the stimulus time series (modeling the midpoints of the sentences) with the canonical hemodynamic response. Sentence duration was not modeled. Realignment parameters and a dummy variable coding for each scanning run were included as covariates of no interest. A high-pass filter (cutoff = 128 s) was applied. Fixed-effects analyses were conducted on each participant's data, one for each of the 4 registration methods (affine only and the 3 nonlinear methods).

Functional Contrasts

For the purposes of this study, we were interested in 3 contrasts: one that would activate auditory thalamus, one that would activate auditory cortex, and one that would activate higher order regions involved in semantic processing. We focused our investigation on 2 foci at each anatomical level using the following contrasts:

Sound versus Rest

"Sound versus rest" consists of the summation of 5 conditions: SCN, (HA + N), (HA – N), (LA + N), and (LA – N) compared with the silent resting baseline condition. This contrast highlights activation along the ascending auditory pathway, including the medial geniculate body (MGB) of the thalamus. This contrast is rather more sensitive to signal change than is the "SCN versus rest" contrast (since it includes more scans) and hence is more suited to detecting signal change in a very small area, such as the MGB. Because it includes all the sentence conditions, it activates both primary auditory regions and surrounding regions of temporal cortex that are speech sensitive, with no clear differentiation. For this reason, this contrast was not used to highlight auditory cortex activity.

SCN versus Rest

SCN (Schroeder 1968) compared with a silent condition, reveals activity in the probable location of primary auditory cortex. Because no speech conditions are included in this contrast, activity ought to be rather focal in the vicinity of Heschl's gyrus.

High Ambiguity versus Low Ambiguity

Rodd et al. (2005) and Davis et al. (2007) document peak activation foci for the high-ambiguity versus low-ambiguity contrast in the left posterior inferior temporal gyrus (ITG) and the left inferior frontal gyrus (IFG).

Assessment of Accuracy of Anatomical Alignment as a Function of Registration Method

To quantify the anatomical alignments achieved by the nonlinear compared with the affine registrations, we first created anatomical regions of interest (ROIs) and then calculated the normalized cross-correlation (NCC) values between all pairs of warped structural data within each of the ROIs.

We used the group *t*-statistics maps (thresholded at $P < 0.001$, uncorrected) for the 3 key contrasts (i.e., sound vs. rest; SCN vs. rest; and high ambiguity vs. low ambiguity) to generate the ROIs. For each contrast, the cluster around the most significant voxel was determined for each of the 4 registration methods. Next, the "union" image of the 4 clusters (corresponding to 4 registration techniques) was created and binarized. This resulted in ROIs in the anatomical locations relevant for our functional questions that could be used to evaluate the accuracy of anatomical registration across the 4 registration methods in an objective fashion (Fig. 1).

NCC is computed by multiplying each voxel within the desired ROI in one warped structural by the spatially corresponding voxel in the other warped structural and summing the products (eq. 1).

$$\mathrm{NCC}(I_m, I_n) = \frac{E[(I_m - \bar{I}_m)(I_n - \bar{I}_n)]}{\sigma_m \sigma_n}, \quad (1)$$

where I_m and I_n refer to the intensity values in the mth and nth subject image volumes, respectively.

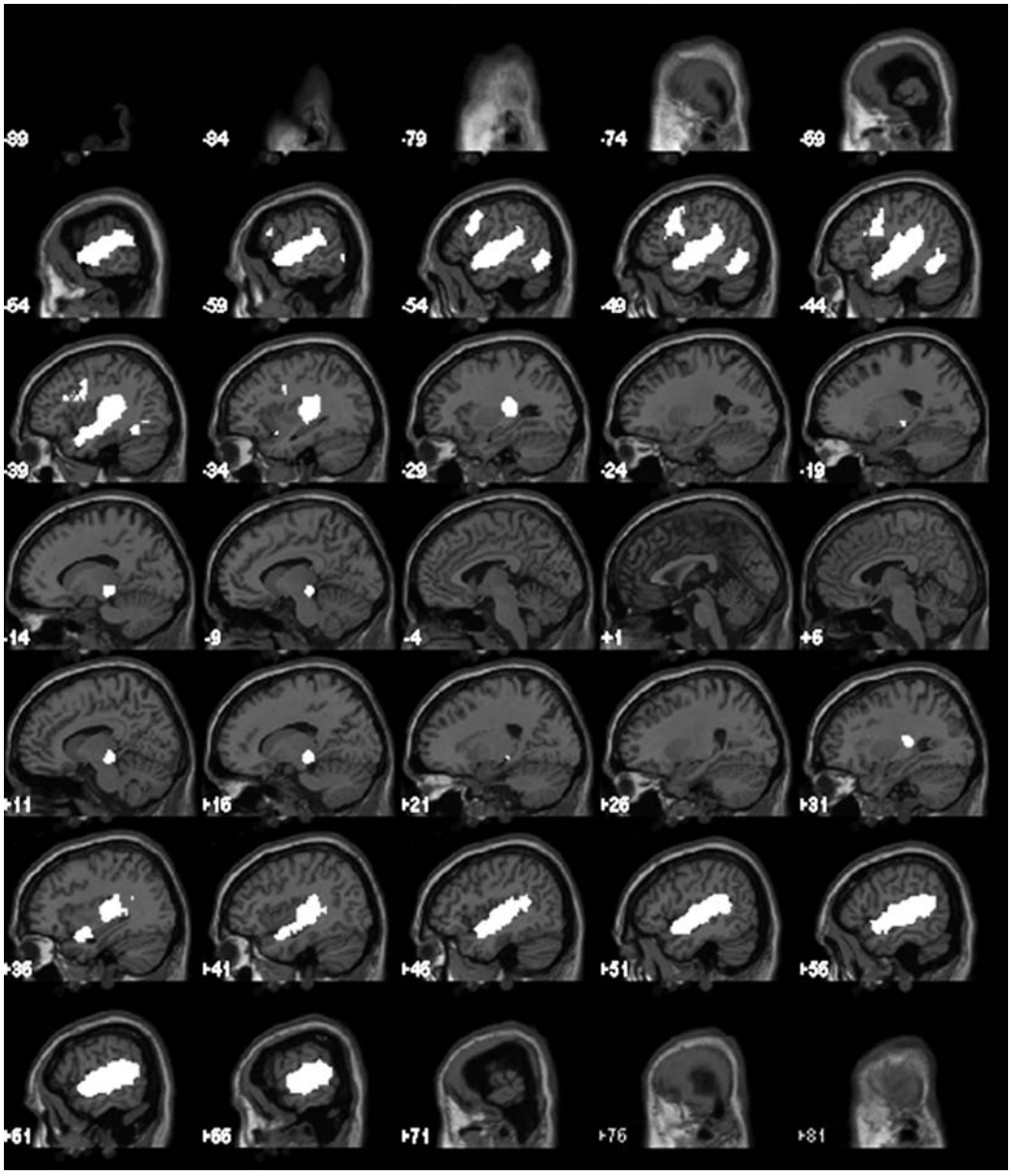

Figure 1. Sagittal views of the functional ROIs used in NCC calculation. ROIs are highlighted in white on the Colin27 brain. The thalamus region highlighted in the sound versus rest contrast is shown between slices x = −19 to x = −9. Auditory cortex, highlighted in the SCN vs rest contrast, is shown between x = −64 and x = −29, and between x = +31 and x = +66. The frontal and temporal association cortex highlighted in the high versus low ambiguity contrast is shown between x = −59 and x = −34.

An analysis if variance (ANOVA) was conducted on NCC scores to assess the effects of registration method (4 levels) and anatomical level (thalamus, auditory cortex, inferior frontal/posterior temporal cortex). Peak (with 2 levels; i.e., left vs. right thalamus/auditory cortex at the 2 lower anatomical levels and inferior temporal region vs. inferior frontal region at the highest processing level) is also included as a dummy factor, although the main effect of peak, and interactions involving peak, will not be considered (since the 2 levels of this factor are not comparable at all 3 anatomical levels, and this factor is not relevant to the hypothesis under investigation).

Assessment of Functional Overlap as a Function of Registration Method

In order to examine whether high-dimensional registration does result in increased functional overlap among subjects, we used a leave-one-out (LOO) method in which we calculated group statistics 24 times for each contrast and each registration method, leaving a different subject out of the analysis each time. We reasoned that, if functional overlap across subjects for a given contrast is improved by nonlinear registration, then the peak location identified in an analysis based on 23 subjects should more accurately predict the signal at the homologous location in the left-out subject. Effect sizes (i.e., contrast values) at these "predicted" voxels in the left-out subjects should be higher and activation in the remaining group of 23 participants should be less variable for nonlinear registration methods than for affine-only registration. Accordingly, for each participant, we extracted the contrast values at the peak voxel coordinates derived from the group analysis of the remaining 23 participants. This was repeated for each of the 2 peaks of interest, for each of the 12 combinations of contrast and registration method. Variability was quite different among the different registration methods, and functional overlap depends both on the magnitude of the contrast values and on the variability in contrast values over subjects. In order to combine single subject contrast magnitude and group variability into a single measure of anatomical consistency, we divided each contrast value in each of the 24 participants by the standard deviation of the remaining set of 23 participants to derive normalized contrast values.

A 3-factor repeated measures ANOVA was applied to normalized contrast values, to examine the effects of anatomical level (3 levels) and registration method (4 levels). Peak (2 levels) is also included as a dummy factor. Although we recognize that effects of different normalization methods may differ on the left versus the right and in the ITG versus IFG, we are primarily interested here in examining the effects of levels of processing. In consequence, we model level and include "peak" as a factor to capture those much less interesting sources of variability within a level of processing. Accordingly, the main effect of peak and interactions involving peak are not considered.

We conducted a supplementary qualitative analysis to examine directly which of the 3 contrasts yield higher *t*-statistics at the group level, after nonlinear, compared with affine, anatomical registration. Since we only obtain one *t*-value per contrast at the group level, we cannot perform statistical comparisons and so we randomly divided the 24 subjects into 2 sets of 12 (8 females in the first group and 9 females in the second group) so that we could make an observation in one group and then examine whether we could replicate it in the second group. Furthermore, we can examine whether increases in *t*-values are consistent across the 3 nonlinear registrations compared with the affine registration. (Note that the DARTEL registration is used to warp all 24 subjects as a group (group 1 + 2) and also each of 12-person subgroups, separately: This was necessary because DARTEL is a groupwise method.).

Results

Accuracy of Anatomical Alignment as a Function of Registration Method

Figure 2 shows the average of 24 warped images using the 4 registration techniques. DARTEL and HAMMER result in precise alignment of the data: The average images from both techniques have clearly defined gray and white matter boundaries. As expected, the average image resulting from unified segmentation–based registration is not as sharp as DARTEL- and HAMMER-based registration. The affine-based average image has fuzzy boundaries between the 2 tissue types due to poor anatomical alignment.

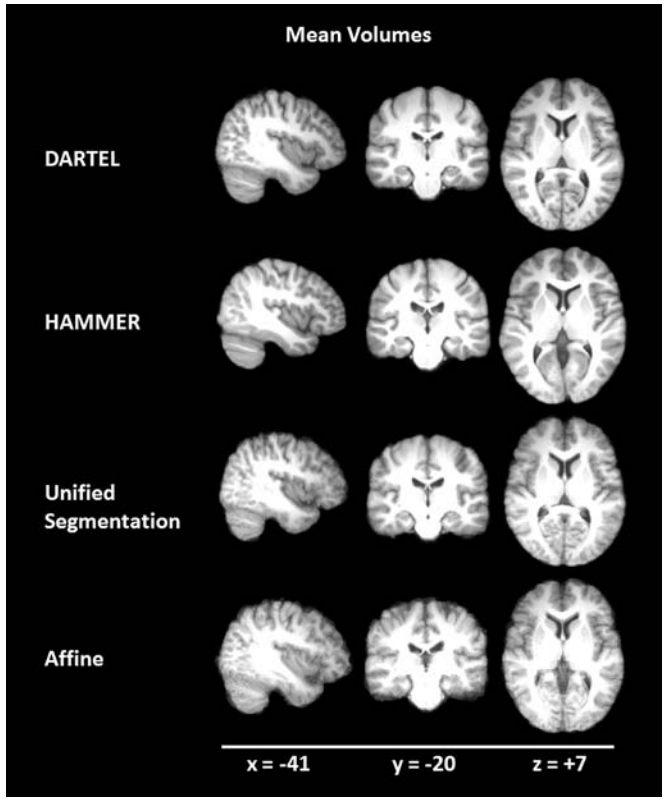

Figure 2. Three cross-sections of the mean warped brain volumes ($N = 24$) generated using 4 different registration techniques shown at ($x = -41$, $y = -20$, $z = +7$) in the MNI coordinate frame.

Figure 3 depicts the NCC scores as a function of registration method (DARTEL, HAMMER, unified segmentation, and affine) and anatomical level (thalamus, auditory cortex, and association cortex). There was a significant effect of registration method ($F_{3,273} = 2389.487$; $P < 0.0001$). Post hoc pairwise comparisons (Sidak-corrected for multiple comparisons to achieve $\alpha = 0.05$) revealed that DARTEL, HAMMER, and unified segmentation outperformed affine registration at all 3 anatomical levels (HAMMER > DARTEL > unified segmentation > affine). Furthermore, there was a significant main effect of the anatomical level ($F_{2,274} = 238.77$; $P < 0.0001$). Post hoc pairwise comparisons revealed the following order of NCC values (higher values index greater alignment among subjects) (mean ± standard error): inferior frontal/posterior temporal cortex (0.726 ± 0.001) > auditory cortex (0.679 ± 0.002) > thalamus (0.661 ± 0.006). Finally, there was a significant interaction between the registration method and anatomical level ($F_{6,270} = 240.39$; $P < 0.0001$). Post hoc pairwise comparisons revealed that although nonlinear registration methods always resulted in greater NCC values than affine, the ordering of the 3 nonlinear methods depended on anatomical level: thalamus: DARTEL > HAMMER > unified segmentation > affine; auditory cortex and inferior frontal/posterior temporal cortex: HAMMER > DARTEL > unified segmentation > affine.

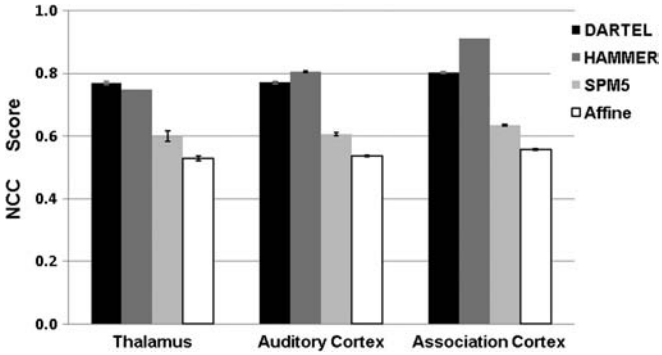

Figure 3. NCC scores, as a function of registration method (DARTEL, HAMMER, unified segmentation, and affine) and anatomical level (thalamus, auditory cortex, and association cortex).

Functional Contrasts

As expected, significant activation foci (corrected at a whole-brain level for multiple comparisons using the false discovery rate correction; Genovese et al. 2002) were observed for each of the 3 contrasts of interest in the anatomical regions predicted. The region of the MGB was observed to be bilaterally active in the sound versus rest contrast; activation was observed bilaterally in the region of Heschl's gyrus for the comparison of SCN and Rest; and activation was observed in left IFG and left posterior inferior temporal cortex for the comparison of high-ambiguity and low-ambiguity sentences (Rodd et al. 2005; Davis et al. 2007). The statistical values for these peaks are reported in Tables 1–3, and they are shown in Figs. 4–6.

Functional Overlap as a Function of Registration Method

We did not apply additional smoothing during preprocessing but it was highly likely that different normalization methods would result in different degrees of effective smoothness. One-way ANOVA confirmed this: contrast images warped using DARTEL and unified segmentation were significantly smoother (average volumetric resolution [mean ± standard deviation]: 12.79 ± 0.03 and 12.86 ± 0.03, respectively) than affine and HAMMER registered images (average volumetric resolution: 7.47 ± 0.03 and 7.57 ± 0.03, respectively); $F_{3,92}$ = 13966.54, P < 0.0001. However, our main prediction is that the influence of registration method on functional overlap will differ at different anatomical levels—overall, differences in smoothness cannot account for such an interaction. Nevertheless, the degree of effective smoothness may influence the activation cluster sizes: the larger effective smoothness, the bigger the cluster sizes one might expect (see Figs. 4–6).

Normalized contrast values (see Materials and Methods) are presented in Figure 7. ANOVA on these values revealed a significant effect of anatomical level ($F_{2,22}$ = 19.2, P < 0.001). Post hoc paired t-tests revealed no difference in normalized contrast values between the 2 lower levels of the auditory system, although these were both higher than those for the high- versus low-ambiguity contrast (Ps < 0.001). A main effect of registration method was also observed ($F_{3,21}$ = 6.1, P < 0.005) such that normalized contrast estimates were higher for DARTEL than any other method (DARTEL vs. other methods; Ps < 0.02) and lowest for Affine (Affine vs. other methods; Ps < 0.02), with no difference between contrast values generated from HAMMER and unified segmentation registrations.

Importantly, the effect of registration method was significantly different at different anatomical levels ($F_{6,18}$ = 10.13, P < 0.001) as is shown in Figure 7. In medial geniculate and auditory cortex, all nonlinear registration methods yielded higher normalized contrast values than did affine registration, according to post hoc pairwise t-tests (Ps < 0.0001). The only exception was that unified segmentation values did not differ significantly from affine values (P = 0.122) in the medial geniculate region. By contrast, in frontal and temporal areas responsive to the ambiguity manipulation, normalized contrast values from affine-registered data were either significantly greater than those from nonlinearly registered data (for DARTEL and HAMMER; Ps < 0.03) or not significantly different (for unified segmentation; P = 0.17).

This observation of better functional overlap with nonlinear registration in lower, but not higher, areas is entirely consistent with the relative magnitudes of the t-values generated by the group analyses. DARTEL-, HAMMER-, and unified segmentation–based registration consistently resulted in higher group t-values than did affine registration in the MGB in both hemispheres for the full 24-subject data set as well as for the 2 randomly selected subsets of 12 subjects (see Fig. 4 and Table 1). The same pattern of results is evident in the region of Heschl's gyrus; Figure 5 and Table 2 demonstrate more robust statistics for the 3 nonlinear registration methods, compared with affine registration, across both random subsets and the full group. A different pattern of results is evident for the high-ambiguity versus low-ambiguity contrast. Figure 6 and Table 3 show no clear distinction between DARTEL, HAMMER, unified segmentation, and affine registration in group t-values, in both subgroups. In fact, the t-values are typically highest for the affine-based registration in the full group.

Discussion

Although it is often assumed that the brain regions supporting cognitively higher functions are more variable than those supporting lower sensory or perceptual processes, evidence supporting this assumption has been lacking (Fox and Pardo 1991; Xiong et al. 2000). Here, we address this issue by examining whether improved anatomical alignment of subjects improves functional overlap among subjects to the same degree for lower level (auditory) processing and for higher level (semantic) processing of speech. We did this by comparing functional overlap in the auditory thalamus, auditory cortex, and in frontal/temporal regions, after spatial transformation of data using 3 nonlinear registration methods (DARTEL, HAMMER, and unified segmentation) and a 12-parameter affine registration.

Nonlinear normalization clearly resulted in better anatomical alignment among subjects than did affine normalization at all anatomical levels. DARTEL and HAMMER resulted in greater alignment than did unified segmentation, which in turn resulted in greater alignment than did affine. This is consistent with the number of deformation parameters for each of these methods. DARTEL (Ashburner 2007) and HAMMER (Shen and Davatzikos 2002) are highly nonlinear, whereas unified segmentation (Ashburner and Friston 2005) has a reduced number of deformation parameters (also see Klein et al. 2009).

Functional overlap across subjects was assessed in 2 ways. First, we conducted a leave-out-out analysis in which we measured the contrast value in the excluded subject at the coordinates of the peak revealed in the (n – 1) group. We reasoned that greater overlap among subjects would result in higher and/or more consistent effect sizes in excluded

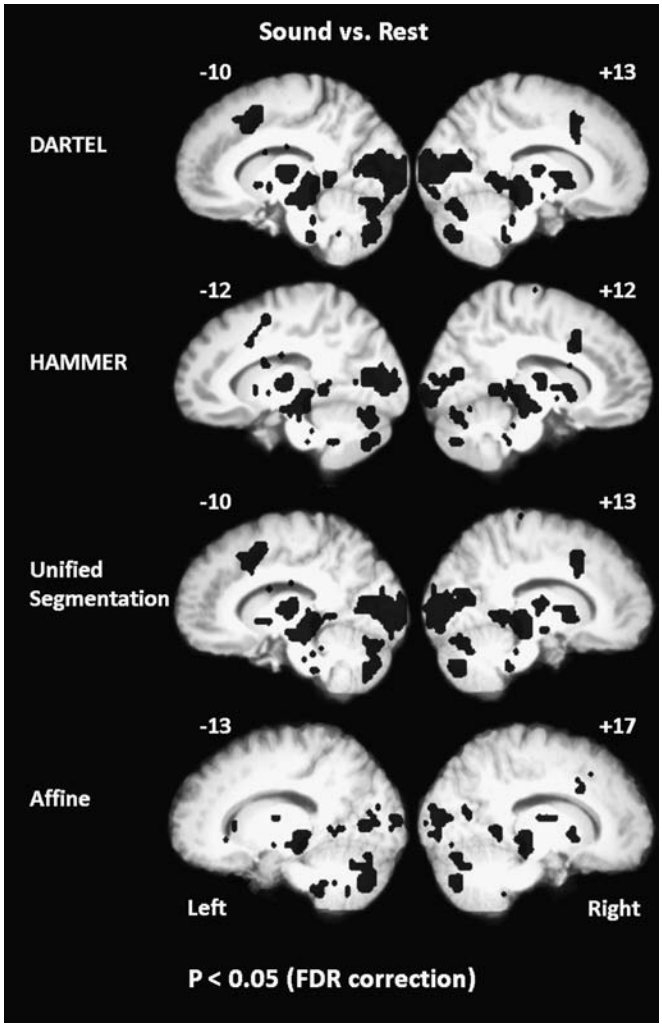

Figure 4. Sagittal views of the activation maps corresponding to the sound versus rest contrast from the group analysis of fMRI data warped using the 4 registration methods; DARTEL, HAMMER, unified segmentation, and affine. The peak activation locations are given in Table 1. All activation maps are extracted from the group analysis with 24 subjects (threshold value for false discovery rate correction, $P < 0.05$: DARTEL: 2.60; HAMMER: 2.68; unified segmentation: 2.61; and affine: 2.76). Coordinates are given in MNI space. Note that cluster sizes are not directly comparable since thresholds are different for the different normalizations.

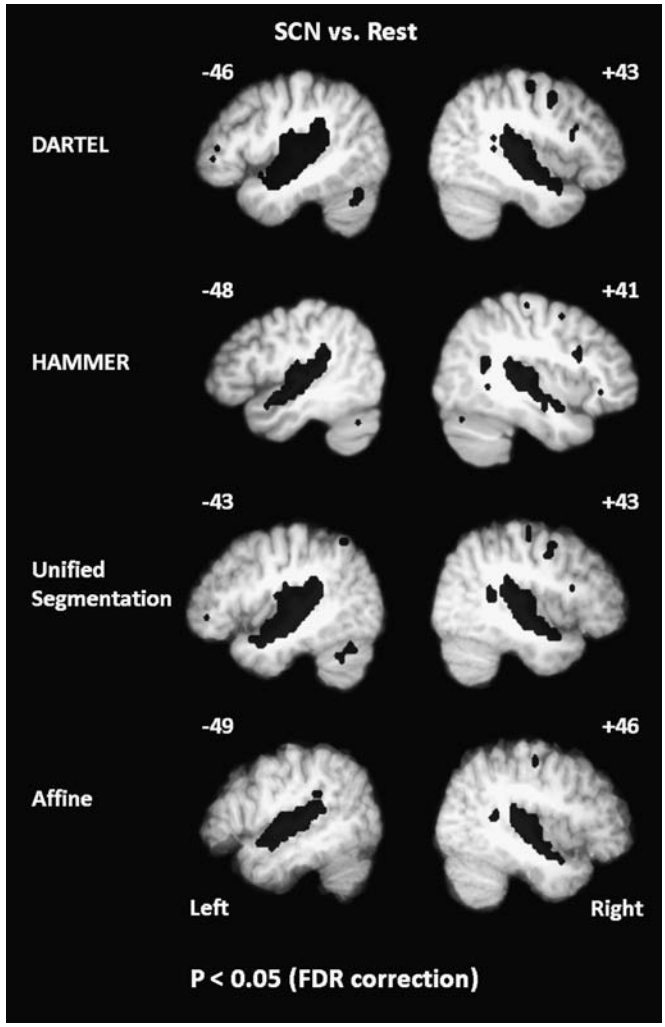

Figure 5. Sagittal views of the activation maps corresponding to the SCN versus rest contrast from the group analysis of fMRI data warped using the 4 registration methods; DARTEL, HAMMER, unified segmentation, and affine. The peak activation locations are given in Table 2. All activation maps are extracted from the group analysis with 24 subjects (threshold value for false discovery rate correction, $P < 0.05$: DARTEL: 2.97; HAMMER: 3.14; unified segmentation: 3.00; and affine: 3.20). Coordinates are given in MNI space. Note that cluster sizes are not directly comparable since thresholds are different for the different normalizations.

Table 1

The significant activation peaks ($P < 0.05$, false discovery rate correction for multiple comparisons) in the region of the MGB for the sound versus rest contrast, in 2 subgroups (12 subjects each), and the total group of 24 subjects

Registration	Group	MGB—L			MGB—R		
		Coordinates	*t*-value	Difference	Coordinates	*t*-value	Difference
DARTEL	Group 1	(−10, −26, −8)	10.21	4.26	(13, −26, −8)	8.15	2.36
	Group 2	(−10, −30, −4)	11.46	4.39	(10, −23, −12)	14.78	8.16
	Group 1 + 2	(−10, −30, −4)	14.85	6.54	(13, −26, −8)	13.73	4.79
HAMMER	Group 1	(−15, −26, −5)	8.55	2.6	(15, −22, −5)	7.83	2.04
	Group 2	(−12, −26, −5)	13.52	6.45	(12, −22, −9)	12.18	5.56
	Group 1 + 2	(−12, −26, −5)	13.93	5.62	(12, −26, −5)	11.88	2.94
Unified segmentation	Group 1	(−10, −26, −8)	9.23	3.28	(13, −30, −8)	8.18	2.39
	Group 2	(−16, −26, −4)	7.54	0.47	(17, −26, −4)	11.16	4.54
	Group 1 + 2	(−10, −26, −8)	10.92	2.61	(13, −26, −8)	12.45	3.51
Affine	Group 1	(−16, −23, −4)	5.95	—	(10, −30, 0)	5.79	—
	Group 2	(−13, −26, −4)	7.07	—	(13, −26, −8)	6.62	—
	Group 1 + 2	(−13, −26, −4)	8.31	—	(17, −23, −4)	8.94	—

Note: The table lists coordinates (in MNI space) and corresponding *t*-values generated by analysis of data warped using 4 different registration methods. MGB—L: left MGB; MGB—R: right MGB. The difference columns give the difference between the *t*-values resulting from data registered using each of the 3 nonlinear methods and affinely registered data (increase in *t*-value).

Table 2

The significant activation peaks ($P < 0.05$, false discovery rate correction for multiple comparisons) in the region of auditory cortex for the SCN versus rest contrast, in 2 subgroups (12 subjects each), and the total group of 24 subjects

Registration	Group	HG—L			HG—R		
		Coordinates	*t*-value	Difference	Coordinates	*t*-value	Difference
DARTEL	Group 1	(−46, −20, 4)	15.94	6.48	(66, −23, 8)	13.58	6.73
	Group 2	(−40, −26, 4)	13.32	5.53	(43, −30, 8)	13.26	5.24
	Group 1 + 2	(−46, −20, 0)	14.15	3.7	(43, −30, 8)	14.62	5.28
HAMMER	Group 1	(−48, −16, 3)	14.06	4.6	(45, −29, 7)	10.3	3.45
	Group 2	(−38, −22, 7)	12.38	4.59	(38, −22, 7)	9.65	1.63
	Group 1 + 2	(−48, −16, 3)	13.64	3.19	(41, −22, 3)	11.04	1.7
Unified segmentation	Group 1	(−43, −20, 4)	12.78	3.32	(50, −26, 12)	11.62	4.77
	Group 2	(−46, −16, −4)	12.92	5.13	(43, −26, 8)	13.34	5.32
	Group 1 + 2	(−43, −20, 4)	12.55	2.1	(43, −26, 8)	13.28	3.94
Affine	Group 1	(−46, −20, 12)	9.46	—	(43, −23, 12)	6.85	—
	Group 2	(−46, −23, 8)	7.79	—	(43, −23, 12)	8.02	—
	Group 1 + 2	(−49, −20, 8)	10.45	—	(46, −26, 12)	9.34	—

Note: The table lists coordinates (in MNI space) and corresponding *t*-values generated by analysis of data warped using 4 different registration methods. HG—L: left Heschl's gyrus region; HG—R: right Heschl's gyrus region. The difference columns give the difference between the *t*-values resulting from data registered using each of the 3 nonlinear methods and affinely registered data (increase in *t*-value).

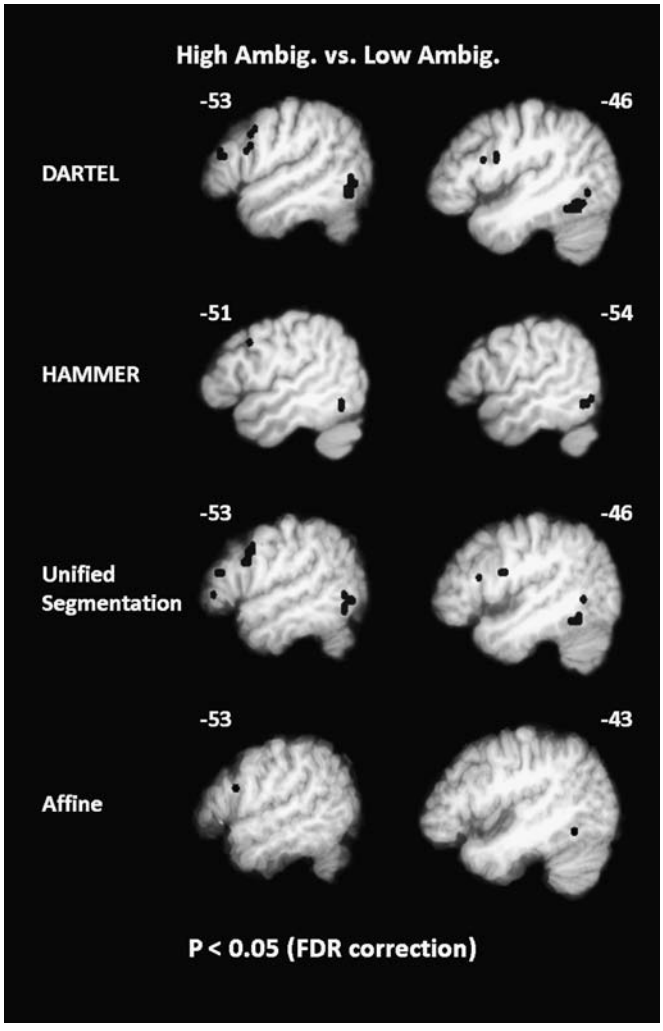

Figure 6. Sagittal views of the activation maps corresponding to the contrast between high-ambiguity and low-ambiguity sentences from the group analysis of fMRI data warped using the 4 registration methods; DARTEL, HAMMER, unified segmentation, and affine. The peak activation locations are given in Table 3. All activation maps are extracted from the group analysis with 24 subjects (threshold value for false discovery rate correction, $P < 0.05$: DARTEL: 4.37; HAMMER: 4.98; unified segmentation: 4.48; and affine: 5.79). Coordinates are given in MNI space. Note that cluster sizes are not directly comparable since thresholds are different for the different normalizations.

subjects. We also qualitatively assessed *t*-statistics at the group level, since group *t*-statistics, which depend both on effect sizes across subjects as well as on the variability in the effect sizes among subjects, index functional signal-to-noise ratios. To the degree that a better anatomical registration among subjects leads to greater functional overlap, we can conclude that "function follows form." Where better anatomical registration does not lead to greater functional overlap across subjects, we conclude that between-subject variation in the patterns of blood oxygen level–dependent (BOLD) activity is not primarily due to variation in anatomical features visible on MRI.

We observed a marked improvement in functional overlap among subjects for high-dimensional registrations compared with affine registration in areas subserving sensory/perceptual analysis (i.e., as the MGB and Heschl's gyrus). This suggests that function is indeed closely tied to the anatomical features matched by nonlinear registration algorithms. In contrast, despite improved anatomical registration in left inferior frontal/posterior temporal regions with nonlinear warping, we did not observe a corresponding improvement in functional overlap in these areas, when we compared activity for high- and low-ambiguity sentences. These higher level meaning-based processes appear therefore not be as closely tied to anatomy compared with lower level sensory/perceptual functions.

Thus, we obtained a dissociation between anatomical alignment across subjects, which is improved at all levels by nonlinear normalization, and functional alignment across subjects, which is not improved at the highest processing level. The fact that nonlinear normalization improves anatomical registration but does not improve the overlap among individuals in semantic processing–related activity may be due to several factors. First, the relative sizes and configurations of microanatomical regions from person to person can vary substantially (Amunts et al. 1999; Scheperjans et al. 2008), even for primary areas (Morosan et al. 2001), and consistent correspondence between macroantomical landmarks and cytoarchitectonic borders is sometimes hard to observe (Caspers et al. 2006). Such variability would result in functional activation foci being observed in somewhat different places relative to morphological features in different subjects. Probabilistic approaches have been proposed to incorporate intersubject variability in the creation of cytoarchitectonic maps (Caspers et al. 2006).

Table 3

The significant activation peaks ($P < 0.05$, false discovery rate correction for multiple comparisons) in the region of left inferior temporal gyrus (ITG—L) and inferior frontal gyrus (IFG—L) for the contrast between high-ambiguity and low-ambiguity sentences in 2 subgroups (12 subjects each) and the total group of 24 subjects

Registration	Group	ITG—L			IFG—L		
		Coordinates	*t*-value	Difference	Coordinates	*t*-value	Difference
DARTEL	Group 1	(−53, −66, −8)	5.87	1.33	(−46, 10, 25)	6.72	1.5
	Group 2	(−46, −69, −8)	6.3	0.23	(−53, 7,3 3)	4.73	0.57
	Group 1 + 2	(−49, −67, −4)	5.84	−0.68	(−53, 10, 25)	4.99	−0.81
HAMMER	Group 1	(−54, −62, −9)	5.9	1.36	(−48, 11, 28)	6.05	0.38
	Group 2	(−48, −59, −17)	5.56	0.51	(−54, 11, 36)	4.2	0.04
	Group 1 + 2	(−54, −59, −9)	5.78	0.74	(−51, 11, 36)	5.24	−0.56
Unified segmentation	Group 1	(−53, −63, 0)	5.31	0.77	(−49, 13, 25)	5.89	0.22
	Group 2	(−46, −59, 0)	6.44	0.37	(−49, 30, 29)	4.16	0
	Group 1 + 2	(−46, −53, −16)	6.64	0.12	(−53, 10, 33)	4.99	−0.81
Affine	Group 1	(−53, −63, −4)	4.54	—	(−53, 17, 21)	5.67	—
	Group 2	(−43, −59, −12)	6.07	—	(−53, 13, 37)	4.16	—
	Group 1 + 2	(−43, −56, −8)	6.52	—	(−53, 17, 25)	5.8	—

Note: The table lists coordinates (in MNI space) and corresponding *t*-values generated by analysis of data warped using 4 different registration methods. The difference columns give the difference between the *t*-values resulting from data registered using each of the 3 nonlinear methods and affinely registered data (increase in *t*-value).

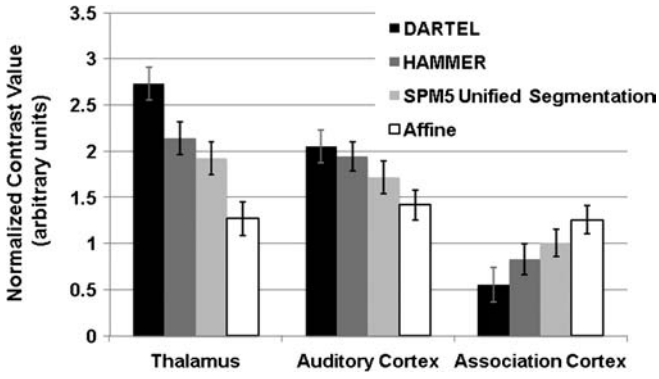

Figure 7. Mean and standard error of normalized contrast values from the LOO analyses (see text for details) for 3 different contrasts revealing activity in 3 different brain regions (sound vs. rest: thalamus; SCN vs. rest: auditory cortex; high ambiguity vs. low ambiguity: association cortex) and 4 different registration methods. Data are collapsed across left and right peaks (for medial geniculate and auditory cortex regions) and collapsed across the inferior temporal and frontal peaks (for association cortex). A significant interaction is observed between brain region and spatial registration method, such that more nonlinear methods increase normalized contrast estimates at the 2 lower levels of processing but not at the highest level ($F_{6,18} = 10.13$, $P < 0.001$).

The observed intersubject variability in this study could be due to microanatomical borders being less bound to macroanatomical landmarks in frontal and posterior temporal regions compared with auditory regions. Aspects of microanatomical structure are visible from in vivo imaging (Klein et al. 2007; la Fougère et al. 2010). Use of such methods may help to assess whether functional anatomy is tied to microanatomical borders.

The observed intersubject functional variability in this study could also be due to the fact that different people rely on different cognitive processes when they are processing ambiguous words. Indeed, several studies have indicated that functional variability among individuals depends on individual differences in cognitive strategies. For example, Grafton et al. (1994) conducted a memory study in which slow learners were shown to rely more on areas associated with visuomotor processing, whereas fast learners relied more on frontal cortex. Many other studies reveal activation differences reflecting individual differences in activation foci during recognition (Kirchhoff and Buckner 2006) and working memory tasks (Feredoes et al. 2007). Schlaug et al. (1994) found significant interindividual variation in activation when individuals were trained and tested on a complicated finger movement sequence. They interpreted such variability as an effect of different strategies used by the individuals to perform the desired task.

Finally, as is well known, the BOLD fMRI signal is a hemodynamic response that only indirectly reflects neuronal activity. In fact, a substantial component of the signal arises in small- and moderate-gauge draining blood vessels (Poser and Norris 2007; Thompson et al. 2010). The vascular component is particularly large both at lower field strengths (such as the 3T field used here) and with gradient EPI acquisition sequences (as used here). Venous vasculature does vary substantially from person to person (e.g., see Stoquart-Elsankari et al. 2009). Given a large vascular component to the observed BOLD response, such intersubject differences in vasculature could manifest as different patterns of BOLD signal, even if the underlying neurally active tissue was the same.

Given functional variability among subjects, between-subject alignment based on functional response has been proposed for group studies (Van Essen et al. 2001; Lashkari and Golland 2009; Sabuncu et al. 2009). CARET software (Van Essen et al. 2001) provides a functional matching method based on landmarks that are drawn on the surface of the brain by the operator. Sabuncu et al. (2009) proposed a surface-based groupwise method for aligning the functional neuroanatomy of individual brains based on the patterns of neural activity evoked by cognitive tasks, by maximizing the correlation of functional time series among subjects while preserving the cortical topology of each subject. Lashkari and Golland (2009) proposed an unsupervised method for functional parcellation of multisubject fMRI data in a visual study. Their proposed method incorporates both intra- and intersubject variability of the fMRI response. These studies demonstrate significant increase in group statistics for specific targeted functional studies (e.g., vision experiments). However, whether such an approach would consistently result in higher statistical power in group studies of higher level cortical function requires investigation.

Another approach to group analysis is to classify individuals based on a factor related to functional organization. Van Horn et al. (2008) review factors, other than anatomical morphology, that play a significant role in defining unique patterns of activity

in individuals. With a sufficient number of data sets, one could categorize subjects with respect to one or combination of such independent factors and conduct analysis on each group of subjects, separately. Such a classification-based approach in group studies provides an opportunity to explore how brain function relates to intersubject variability in anatomical organization by looking at the correlation (if any) between such factors and anatomy. Such studies may result in a finer mapping of function onto structure and may also result in greater understanding of what constitutes "normal" variability in the anatomical and functional organization of the human brain.

Conclusions

Is the link between anatomical macrostructure and function equally strong at all cognitive levels of processing? In this work, we demonstrate that more accurate anatomical registration among participants does not always mean greater functional alignment. Although the 3 nonlinear algorithms we used all resulted in more accurate anatomical alignment at all anatomical levels of interest compared with affine registration, functional overlap was not improved at all anatomical levels. Better anatomical registration led to greater functional registration across subjects at early stages of auditory processing including auditory thalamus and in the vicinity of Heschl's gyrus. In contrast, at anatomically and cognitively higher levels of processing, greater anatomical alignment did not lead to greater functional overlap across subjects. This indicates that the intersubject variability observed in such regions is not solely due to anatomical variability but may instead be due to other factors such as the relationship between microanatomical features (such as cytoarchitecture, chemoarchitecture, and connectivity) and macroanatomical landmarks (such as gyri and sulci), to individual differences in processes recruited, to differences in vascular configuration, or a combination of all 3. We can, however, say with confidence that the degree to which function follows anatomical form does vary systematically as a function of anatomical/cognitive level of processing, and this has important implications for the analysis and interpretation of functional imaging data.

Funding

Operating grants from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada and infrastructure awards from the Canadian Foundation for Innovation and the Ontario Innovation Trust. A.M.T., C.W., and H.H. were supported by a training grant from the Ontario Ministry of Research and Innovation, and I.S.J. is supported by the Canada Research Chairs Program.

Notes

Conflict of Interest : None declared.

References

Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings H, Zilles K. 1999. Broca's region revisited: cytoarchitecture and intersubject variability. J Comp Neurol. 412(2):319–341.

Ashburner KJ. 2007. A fast diffeomorphic image registration algorithm. Neuroimage. 38:95–113.

Ashburner KJ, Friston K. 2005. Unified segmentation. Neuroimage. 26(3):839–851.

Baayen RH, Piepenbrock R, Gulikers L. 1995. The CELEX Lexical Database (CDROM). Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Boling WW, Olivier A. 2004. Localization of hand sensory function to the pli de passage moyen of broca. J Neurosurg. 101(2):278–283.

Caspers S, Geyer S, Schleicher A, Mohlberg H, Amunts K, Zilles K. 2006. The human inferior parietal cortex: cytoarchitectonic parcellation and interindividual variability. Neuroimage. 33(2):430–448.

Chouinard PA, Paus T. 2006. The primary motor and premotor areas of the human cerebral cortex. Neuroscientist. 12(2):143–152.

Coleman MR, Davis MH, Rodd JM, Robson T, Ali A, Owen AM, Pickard JD. 2009. Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. Brain. 132:2541–2552.

Davis MH, Coleman MR, Absalom AR, Rodd JM, Johnsrude IS, Matta BF, Owen AM, Menon DK. 2007. Dissociating speech perception and comprehension at reduced levels of awareness. Proc Natl Acad Sci U S A. 104(41):16032–16037.

Davis MH, Johnsrude IS. 2003. Hierarchical processing in spoken language comprehension. J Neurosci. 23(8):3423–3431.

Davis MH, Johnsrude IS. 2007. Hearing speech sounds: Top-down influences on the interface between audition and speech perception. Hear Res. 229:132–147.

DeYoe E, Carman G, Bandettini P, Glickman S, Wieser J, Cox R, Miller D, Neitz J. 1996. Mapping striate and extrastriate visual areas in human cerebral cortex. Proc Natl Acad Sci U S A. 93(6):2382–2386.

Dougherty RF, Koch VM, Brewer AA, Fischer B, Modersitzki J, Wandell BA. 2003. Visual field representations and locations of visual areas V1/2/3 in human visual cortex. J Vis. 3:586–598.

Fedorenko E, Hsieh PJ, Nieto-Castañón A, Whitfield-Gabrieli S, Kanwisher N. 2010. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. J Neurophysiol. 104(2):1177–1194.

Feredoes E, Tononi G, Postle BR. 2007. The neural bases of the short-term storage of verbal information are anatomically variable across individuals. J Neurosci. 27(41):11003–11008.

Fox PT, Pardo JV. 1991. Does inter-subject variability in cortical functional organization increase with neural 'distance' from the periphery? Ciba Found Symp. 163:125–140.

Genovese CR, Lazar NA, Nichols T. 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage. 15(4):870–878.

Geyer S, Matelli M, Luppino G, Zilles K. 2000. Functional neuroanatomy of the primate isocortical motor system. Anat Embryol (Berl). 202(6):443–474.

Grafton ST, Woods RP, Tyszka M. 1994. Functional imaging of procedural motor learning: relating cerebral blood flow with individual subject performance. Hum Brain Mapp. 1:221–234.

Grefkes C, Geyer S, Schormann T, Roland P, Zilles K. 2001. Human somatosensory area 2: observer-independent cytoarchitectonic mapping, interindividual variability, and population map. Neuroimage. 14(3):617–631.

Hackett TA. 2011. Information flow in the auditory cortical network. Hear Res. 271(1–2):133–146.

Hall D, Haggard M, Akeroyd M, Palmer A, Summerfield A, Elliott M, Gurney E, Bowtell R. 1999. Sparse temporal sampling in auditory fMRI. Hum Brain Mapp. 7:213–223.

Hasnain MK, Fox PT, Woldorff MG. 2001. Structure-function spatial covariance in the human visual cortex. Cereb Cortex. 11(8):702–716.

Hickok G, Poeppel D. 2007. The cortical organization of speech processing. Nat Rev Neurosci. 8(5):393–402.

Hinds O, Polimeni JR, Rajendran N, Balasubramanian M, Amunts K, Zilles K, Schwartz EL, Fischl B, Triantafyllou C. 2009. Locating the functional and anatomical boundaries of human primary visual cortex. Neuroimage. 46(4):915–922.

Hinds OP, Rajendran N, Polimeni JR, Augustinack JC, Wiggins G, Wald LL, Diana Rosas H, Potthast A, Schwartz EL, Fischl B. 2008. Accurate prediction of V1 location from cortical folds in a surface coordinate system. Neuroimage. 39(4):1585–1599.

Holmes C, Hoge R, Collins L, Woods R, Toga A, Evans A. 1998. Enhancement of MR images using registration for signal averaging. J Comput Assist Tomogr. 22(2):324–333.

Iaria G, Petrides M. 2007. Occipital sulci of the human brain: variability and probability maps. J Comp Neurol. 501(2):243–259.
Kaas J, Hackett T. 2000. Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci U S A. 97(22):11793–11799.
Kirchhoff B, Buckner RL. 2006. Functional-anatomic correlates of individual differences in memory. Neuron. 51(2):263–274.
Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, Christensen GE, Collins DL, Gee J, Hellier P, et al. 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. Neuroimage. 46(3):786–802.
Klein JC, Behrens TE, Robson MD, Mackay CE, Higham DJ, Johansen-Berg H. 2007. Connectivity-based parcellation of human cortex using diffusion MRI: establishing reproducibility, validity and observer independence in BA 44/45 and SMA/pre-SMA. Neuroimage. 34(1):204–211.
la Fougère C, Grant S, Kostikov A, Schirrmacher R, Gravel P, Schipper HM, Reader A, Evans A, Thiel A. 2011. Where in-vivo imaging meets cytoarchitectonics: the relationship between cortical thickness and neuronal density measured with high-resolution [(18)F]flumazenil-PET. Neuroimage. 56(3):951–960.
Lashkari D, Golland P. 2009. Exploratory fMRI analysis without spatial normalization. Inf Process Med Imaging. 5636:398–410.
Lerner Y, Honey C, Silbert L, Hasson U. 2011. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. J Neurosci. 31(8):2906–2915.
Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K. 2001. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage. 13(4):684–701.
Ojemann G, Ojemann J, Lettich E, Berger M. 1989. Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. J Neurosurg. 71(3):316–326.
Okada K, Rong F, Venezia J, Matchin W, Hsieh IH, Saberi K, Serences J, Hickok G. 2010. Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. Cereb Cortex. 20(10):2486–2495.
Peelle JE, Johnsrude IS, Davis MH. 2010. Hierarchical processing for speech in human auditory cortex and beyond. Front Hum Neurosci. 4:51.
Penfield W, Rasmussen T. 1950. The cerebral cortex of man: a clinical study of localization of function. New York: Macmillan.
Poser BA, Norris DG. 2007. Fast spin echo sequences for BOLD functional MRI. MAGMA. 20(1):11–17.
Price CJ. 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann N Y Acad Sci. 1191:62–88.
Rademacher J, Galaburda AM, Kennedy DN, Filipek PA, Caviness VS. 1992. Human cerebral cortex: localization, parcellation, and morphometry with magnetic resonance imaging. J Cogn Neurosci. 4(4):352–374.
Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund HJ, Zilles K. 2001. Probabilistic mapping and volume measurement of human primary auditory cortex. Neuroimage. 13(4):669–683.
Rodd JM, Davis MH, Johnsrude IS. 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. Cereb Cortex. 15(8):1261–1269.
Romanski L, Bates J, Goldman-Rakic P. 1999. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. J Comp Neurol. 403:141–157.
Sabuncu MR, Singer BD, Conroy B, Bryan RE, Ramadge PJ, Haxby JV. 2009. Function-based intersubject alignment of human cortical anatomy. Cereb Cortex. 20(1):130–140.
Scheperjans F, Hermann K, Eickhoff SB, Amunts K, Schleicher A, Zilles K. 2008. Observer-independent cytoarchitectonic mapping of the human superior parietal cortex. Cereb Cortex. 18(4):846–867.
Schlaug G, Knorr U, Seitz RJ. 1994. Inter-subject variability of cerebral activations in acquiring a motor skill: a study with positron emission tomography. Exp Brain Res. 98:523–534.
Schroeder M. 1968. Reference signal for signal quality studies. J Acoust Soc Am. 44:1735–1736.
Scott S, Johnsrude IS. 2003. The neuroanatomical and functional organisation of speech perception. Trends Neurosci. 26:100–107.
Seghier ML, Lazeyras F, Pegna AJ, Annoni JM, Khateb A. 2008. Group analysis and the subject factor in functional magnetic resonance imaging: analysis of fifty right-handed healthy subjects in a semantic language task. Hum Brain Mapp. 29(4):461–477.
Shen D, Davatzikos C. 2002. HAMMER: hierarchical attribute matching mechanism for elastic registration. IEEE Trans Med Imaging. 22(11):1421–1439.
Stoquart-Elsankari S, Lehmann P, Villette A, Czosnyka M, Meyer ME, Deramond H, Balédent O. 2009. A phase-contrast MRI study of physiologic cerebral venous flow. J Cereb Blood Flow Metab. 29(6):1208–1215.
Tahmasebi AM, Abolmaesumi P, Zheng Z, Munhall K, Johnsrude IS. 2009. Reducing inter-subject anatomical variation: effect of normalization in analysis of functional activity in auditory cortex and the superior temporal region. Neuroimage. 47(4):1522–1531.
Thompson R, Correia M, Cusack R. 2011. Vascular contributions to pattern analysis: comparing gradient and spin echo fMRI at 3T. Neuroimage. 56(2):643–650.
Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH. 2001. An integrated software suite for surface-based analyses of cerebral cortex. J Am Med Inform Assoc. 8(5):443–459.
Van Horn JD, Grafton ST, Miller MB. 2008. Individual variability in brain activity: a nuisance or an opportunity? Brain Imaging Behav. 2(4):327–334.
Watson JD, Myers R, Frackowiak RS, Hajnal JV, Woods RP, Mazziotta JC, Shipp S, Zeki S. 1993. Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. Cereb Cortex. 3:79–94.
Wohlschlager AM, Specht K, Lie C, Mohlberg H, Wohlschlager A, Bente K, Pietrzyk U, Stcker T, Zilles K, Amunts K, et al. 2005. Linking retinotopic fMRI mapping and anatomical probability maps of human occipital areas V1 and V2. Neuroimage. 26(1):73–82.
Xiong J, Rao S, Jerabek P, Zamarripa F, Woldorff M, Lancaster J, Fox PT. 2000. Intersubject variability in cortical activations during a complex language task. Neuroimage. 12:326–339.
Yousry TA, Schmid UD, Alkadhi H, Schmidt D, Peraud A, Buettner A, Winkler P. 1997. Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. Brain. 120(Pt 1):141–157.