Chapter 8
Using Multisensory Integration to Understand the Human Auditory Cortex

Michael S. Beauchamp

Abstract Accurate and meaningful parcellation of the human cortex is an essential endeavor to facilitate the collective understanding of brain functions across sensory and cognitive domains. Unlike in the visual cortex, the details of anatomical and functional mapping associated with the earliest stages of auditory processing in the cortex are still a topic of active debate. Interestingly, aspects of multisensory processing may provide a unique window to meaningfully subdivide the auditory sensory areas by exploring different functional properties other than the traditional tonotopic approach. In this chapter, a tour of the auditory cortical areas is first provided, starting from its core area, Heschl's gyrus, then moving onto surrounding areas. Evidence from different sources, including postmortem studies of the human auditory cortex, resting-state functional connectivity derived from the Human Connectome Project, and electrocorticographic studies, is presented to better understand how different subdivisions of the human auditory cortex and its surrounding areas are involved in auditory and multisensory processing. The chapter concludes with the remaining challenges to account for individual variability in functional anatomy, particularly pertaining to multisensory processing.

Keywords Auditory cortex · Cross-modal · Electrocorticography · Functional anatomy · Functional connectivity · Heschl's gyrus · Human Connectome Project · Sensory integration · Superior temporal sulcus · Temporal cortex

8.1 Introduction

The human cerebrum is divided into sensory cortices specialized for the processing of a specific sensory modality, with the visual cortex located in the occipital lobe and the auditory cortex centered on Heschl's gyrus on the plane of the superior

---

M. S. Beauchamp (✉)
Department of Neurosurgery and Core for Advanced MRI, Baylor College of Medicine, Houston, TX, USA
e-mail: michael.beauchamp@bcm.edu

A. K. C. Lee et al. (eds.), *Multisensory Processes*, Springer Handbook of Auditory Research 68, https://doi.org/10.1007/978-3-030-10461-0_8

temporal gyrus. The visual cortex and auditory cortex may be further subdivided into multiple cortical areas, each specialized for performing a specific computation on the incoming sensory data. The best-known example is Felleman and Van Essen's (1991) subdivision of the macaque visual cortex into 32 areas in a 10-level hierarchy. More than 25 years later, neuroscientists are still struggling to develop a similarly detailed description of the auditory cortex. Even at the earliest stages of cortical auditory processing, the number of areas, their anatomical layout, and their nomenclature are topics of active research and debate. One reason for this slow progress is the difficulty in finding functional properties that allow the auditory cortex to be subdivided. In this chapter, the possibility is explored that consideration multisensory integration processes, here the integration of both auditory and nonauditory stimuli, may lead to a better understanding of the human auditory cortex. The organization of the human auditory cortex is presented first (Sect. 8.2) and is framed around the general division into core, belt, and parabelt regions. Next, several subdivisions of the human parabelt cortex are examined from different perspectives (Sects. 8.3 and 8.4). Finally, the chapter concludes by using the auditory cortex as an example of the challenges that face functional brain mapping from the perspective of incorporating individual variability into the process of drawing meaningful functional distinctions between brain regions (Sect. 8.5).

8.2 Organization of the Human Auditory Cortex

The auditory cortex is located in the temporal lobe of the human neocortex (Fig. 8.1). Moving from dorsal to ventral, the temporal cortex encompasses the superior temporal gyrus (STG), the superior temporal sulcus (STS), the middle temporal gyrus (MTG), and the inferior temporal sulcus (ITS). Heschl's gyrus (HG) is a short gyrus engraved on the superior surface of the STG that begins on the lateral convexity and runs lateral-anterior to medial-posterior before terminating in the insula.

HG is the location of so-called "core" auditory areas. The use of the term core instead of "primary" denotes the idea that there are multiple areas that coexist at the

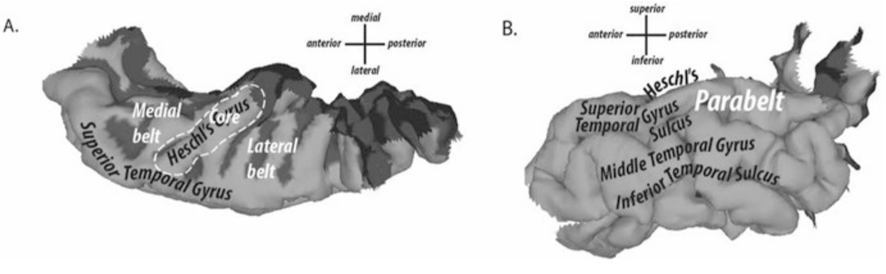

Fig. 8.1 Human auditory cortex. (**A**) Virtual dissection of the human temporal lobe, viewed from above. *Black labels*, anatomically defined structures; *white labels*, functionally defined regions. *Lighter grays,* gyri; *darker grays,* sulci. (**B**) Lateral view of virtual dissection of the pial surface of the temporal lobe

first stage of cortical processing. Each core area contains a tonotopic representation or map. In these maps, there is a gradual change in the preferred auditory frequency of the neurons across the area, with neighboring neurons having similar frequency tuning, and the entire range of perceptible frequencies is represented. At the boundaries between areas, the gradient of change in preferred frequency reverses so that adjacent areas have mirror-symmetrical tonotopic representations. Although the organization of these tonotopic maps in the core auditory cortex has been well established in animal models (Kass et al. 1999), the organization in humans has remained elusive. Data from ultrahigh-field 7-tesla functional magnetic resonance imaging (fMRI) of the blood oxygen level-dependent (BOLD) signal in human subjects led to a proposal by Moerel et al. (2014) that low frequencies are represented along the posterior edge of HG, whereas high frequencies are represented along the anterior edge, and that the core auditory cortex consists of three complete tonotopic maps. These three maps are named based on the conventions applied from studies in nonhuman primate models and consist of area A1 at the medial-posterior edge of HG, the rostrotemporal area at the lateral-anterior edge of HG, and the rostral area in the center of HG. Although this classification scheme is still a subject of debate, it provides a view into the current state of understanding of the functional architecture of the human core auditory cortex.

The cortex anterior and posterior to HG on the superior surface of the STG is the location of “belt” areas of the auditory cortex, so-called because of their anatomical location encircling HG. The areas anterior and medial to HG are referred to as the medial belt areas. Posterior and lateral to HG is a triangular patch of cortex termed the temporal plane (*planum temporale*), the location of the lateral belt areas. Data from tonotopic mapping at 7 tesla was used by Moerel et al. (2014) to divide the belt areas into six subdivisions, with anatomical labels derived from the nomenclature developed in physiological studies of nonhuman primate models. Moving from lateral to medial across the cortex, these areas are rostromedial, mediomedial, and caudomedial within the medial belt and anterolateral, mediolateral, and caudolateral within the lateral belt.

Responses to auditory stimuli extend laterally and posteriorly from the lateral belt areas onto the lateral surface of the STG and into the STS. Collectively, this region is termed the auditory cortex “parabelt.” Although the auditory parabelt is larger than the core and belt areas, it fails to show a robust tonotopic organization, making functional parcellation based on frequency tuning impossible. However, as detailed in Sect. 8.3, substantial effort has been made to better delimit the functional organization of the parabelt areas.

8.3 Subdivisions of the Human Parabelt Cortex

8.3.1 Postmortem and In Vivo Magnetic Resonance Imaging

Classic and more contemporary cytoarchitectonic studies derived from human postmortem tissue may shed some light on the functional organization of the parabelt regions. Although Brodmann in his atlas (1909) classified the entire posterior two-thirds of the STG and STS, extending posteriorly all the way to the inferior parietal lobe, as a single cytoarchitectonic region (i.e., area 22), more recent studies suggest additional subdivisions (Fig. 8.2A; Morosan et al. 2005). Using multimodal architectonic mapping contrasting neuronal cell types, neuronal packing density, columnar organization, and neurotransmitter receptor distributions, the Jülich group identified the existence of a distinct area, labeled Te3, on the lateral bulge of the STG that does not extend onto the dorsal or ventral banks of the STG (Morosan et al. 2005).

A second valuable source of evidence about the functional organization of the human parabelt cortex is the multimodal MRI dataset derived from 210 subjects as part of the Human Connectome Project (HCP; Van Essen et al. 2013). Using this dataset, Glasser et al. (2016) subdivided the cerebral cortex into 180 areas in each hemisphere, including 10 distinct areas located in parabelt cortex (Fig. 8.2B).

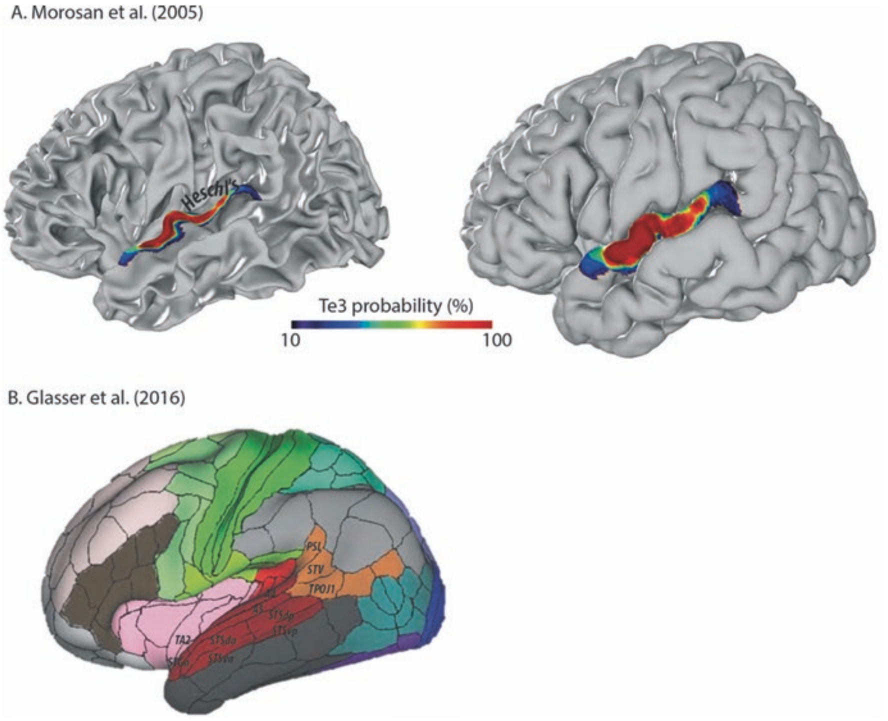

Fig. 8.2 Anatomical-functional subdivisions in lateral temporal cortex. (**A**) Morosan et al. (2005) described a cytoarchitectonic region termed Te3. Ten postmortem human brains were examined. The color scale shows the probability of the brain area containing Te3, visualized in a white matter cortical surface (*left*) and a pial cortical surface (*right*). *Dashed white line*, Heschl's gyrus; *red line,* cutting plane defined by Heschl's gyrus. (**B**) Inflated cortical surface model showing the Human Connectome Project (HCP) 1.0 brain parcellation. *Labeled areas*, parabelt cortex. *A4*, auditory 4 complex; *A5*, auditory 5 complex; *PSL*, perisylvian language area; *STG/S*, superior temporal gyrus/sulcus; *STV*, superior temporal visual area; *TPOJ1*, temporo-parieto-occipital junction 1; *TA2*, area TA2; *d*, dorsal; *v*, ventral; *a*, anterior; *p*, posterior

To consider this parcellation of the parabelt cortex in more detail, it is first necessary to briefly review its composition, which consists of four different measurements of brain structure and function. First, functional connectivity or resting-state fMRI measures of BOLD signal fluctuations can be obtained as the subject lays in the scanner without a task. Areas that show synchronous fluctuations in the BOLD signal are presumed to be functionally connected. If adjacent regions of the cortex show very different patterns of connectivity, this is taken as evidence for an areal boundary. Second, task-based fMRI measures the activity in brain areas in response to different stimuli and task conditions. Of particular relevance for parabelt delimitation in the HCP data is the language story condition, in which subjects listened to a story, and the language math condition, in which subjects listened to a math problem. Subtracting the math from the story condition task (story vs. math contrast) reveals areas specific for higher order language processes. Third, MRI pulse sequences and analysis techniques have been developed to measure correlates of the cytoarchitecture, notably myelin content (Glasser and Van Essen 2011; Glasser et al. 2014). Using this method, gradients in myelin content can also be used to distinguish between areas. Fourth, structural MRI can be used to determine cortical thickness and cortical folding patterns. The assumption is that changes in cortical thickness, or differences relative to cortical landmarks such as the fundus of a sulcus, represent areal boundaries.

Use of these four dimensions resulted in the division of the parabelt into ten distinct divisions. The two most anterior parabelt areas are small areas lateral to the medial belt areas and are labeled areas TA2 and STGa. The next four areas tile most of the STS. STSda and STSdp tile the upper bank of the STS in the anterior-to-posterior direction, whereas STSva and STSvp tile the lower bank of the STS in the same direction. Interestingly, the functional connectivity gradient is strongest along the fundus of the STS, providing strong evidence for an important functional distinction between the upper and lower bank areas. Medial to the STS, two areas tile the crown of the STG, area A4 more medially and area A5 more laterally. Finally, the most posterior of the parabelt regions are the superior temporal visual area (STV), made up of the most posterior section of STG before it angles up into the parietal lobe, and the temporo-parieto-occipital junction 1 area, which is the most posterior section of STS (Fig. 8.2B).

8.3.2 *Electrocortigraphic Evidence That Multisensory Integration in the Auditory Cortex Provides Valuable Functional Information*

The aforementioned atlases (see Sect. 8.3.1) derived from postmortem histology and group MRI data clearly suggest the existence of functionally specialized areas within the auditory cortex that are located anterior to posterior along the STG and STS. However, these group atlases are of limited value in understanding the organization of a particular individual brain. For instance, the 180 areas in the HCP atlas were defined by assuming that areas existed with sharp boundaries between them. A natural question is whether the parabelt cortex is best described as a series of areas with sharply defined boundaries or as a broad region of the cortex with gradual transitions between zones with different functional properties.

The ability of a technique such as BOLD fMRI to answer this question is limited because it does not directly measure neural activity. Any property of the cortical map, whether it is a sharp boundary or a gradual transition, could be ascribed to the properties of the cerebral vasculature in the region rather than the functional properties of the underlying neurons. For instance, if the upper and lower banks of the STS are drained by different venous beds, this could create a sharp boundary in BOLD fMRI between the two regions that does not necessarily reflect a functional distinction.

Another method to examine functional specialization in the parabelt cortex is intracranial encephalography (iEEG), also known as electrocorticography (ECoG), a technique to record activity directly from the cortex of awake human subjects (generally patients with conditions such as intractable epilepsy). Relative to BOLD fMRI, iEEG has the advantage that it directly measures neural activity without imposing a blurry hemodynamic filter.

A study used iEEG to probe the functional organization of auditory cortical regions by examining responses to audiovisual speech presented within varying levels of auditory noise (Ozker et al. 2017). The key observation motivating this study is that humans can use the visual mouth movements observed in the face of the talker to better understand the talker's voice and that these improvements grow larger as the auditory speech signal becomes noisier. The hypothesis behind the study was that parabelt areas involved in speech processing would be differentiated into those that process exclusively auditory information and those that integrate visual and auditory speech information.

Posterior portions of the STS/STG are multisensory in that they respond not only to auditory but also to visual and somatosensory stimuli in both humans (Beauchamp et al. 2004a, b) and nonhuman primates (Bruce et al. 1981). Therefore, for the purposes of this study, STG was divided into a posterior section and an anterior section (Fig. 8.3).

The responses to clear versus noisy audiovisual speech were strikingly different between the anterior and posterior STS/STG. Whereas in the anterior half, noisy speech greatly decreased the amplitude of the response when compared with clear speech, in the posterior half there was no decrease in the amplitude of the response. This effect was highly consistent; all anterior electrodes showed larger responses for the stimuli consisting of clear audiovisual speech, whereas all posterior STG electrodes showed similar responses for stimuli consisting of clear or noisy audiovisual speech.

Because iEEG directly measures neural activity from a small region of the cortex, activity in each electrode can be confidently assigned with anatomical precision. To examine the effect of anatomical location on the response to clear and noisy audiovisual speech with more detail than a simple division of the STS/STG into

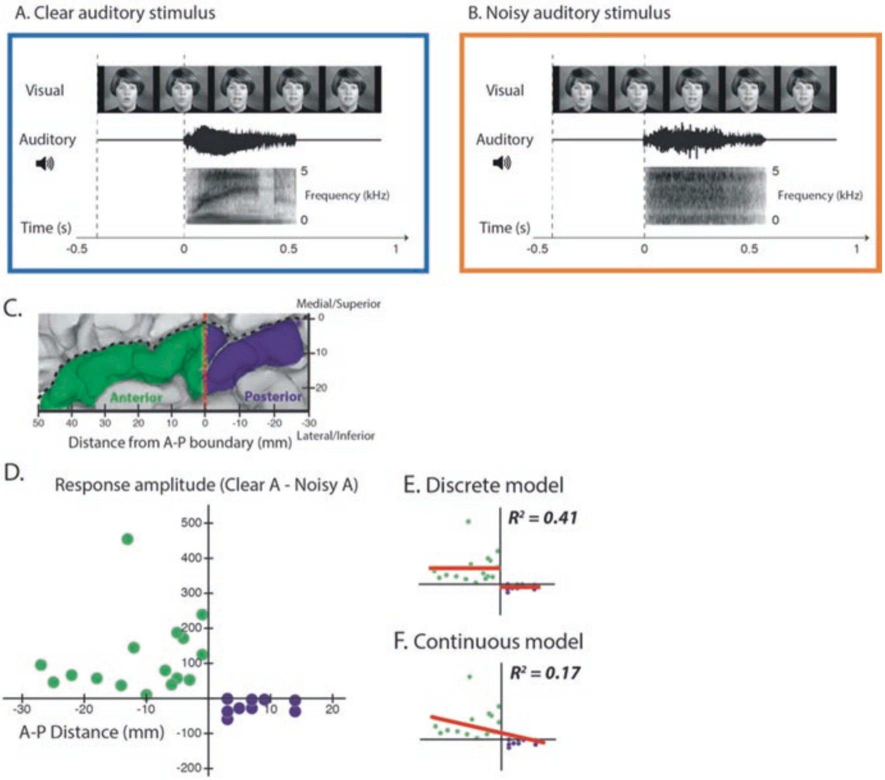

Fig. 8.3 Using electrocorticography to parcellate the parabelt cortex. (**A**) Example stimulus consisting of clear auditory speech (Clear A) and a movie of a talking face. Spectrogram shows clear formant bands in auditory speech. (**B**) Example stimulus consisting of noisy auditory speech (Noisy A) and a movie of a talking face. Spectrogram shows a lack of formant bands. (**C**) Lateral view of a cortical surface model of the temporal lobe showing anterior STG (*green*) and posterior STG (*purple*). Heschl's gyrus (not visible on the superior face of temporal lobe) extends from anterior-lateral to posterior-medial. The posterior most point of Heschl's gyrus is used to define an origin (*red dashed line*). All points anterior to this origin are classified as anterior and given a positive value corresponding to their distance from the origin (values on x-axis). The distance from the origin in the inferior-to-superior direction is shown on the y-axis.*Black dashed line*, distance from medial/superior border of STG. (**D**) Response to Clear A versus Noisy A speech for each individual electrode. *Green circles*, each anterior electrode; *purple circles*, each posterior electrode. The response amplitude is the mean percent change in high-gamma power (70–110 Hz) in the 0- to 500-ms time window relative to the prestimulus baseline (−500 to −100 ms). (**E**) Discrete model: constant values were fit separately to the anterior and posterior electrode data in **B** ($y = a$ for all electrodes with $x > 0$ and $y = b$ for all electrodes with $x < 0$) and the correlation with the data was calculated. (**F**) Continuous model: a linear model with two parameters was fit to both anterior and posterior electrodes ($y = mx + b$, where m is the slope and b is the constant term). Adapted from Ozker et al. (2017)

anterior and posterior segments, the location of each electrode was plotted in a functional reference frame defined by the responses to auditory speech. In creating this reference frame, the location of each electrode was plotted against its preference for clear compared with noisy audiovisual speech. Consistent with the first analysis, anterior electrodes showed greater responses for clear speech, whereas posterior electrodes showed similar or smaller responses for clear compared with noisy speech. Most importantly, when examined on an electrode-by-electrode basis, a sharp transition was found between anterior and posterior electrodes. This observation received quantitative support from a comparison of two Bayesian models, one of which posited a discrete transition (Fig. 8.3E) and one of which posited a gradual transition (Fig. 8.3F). The discrete model was more than 100 times more likely to explain the observed data.

Hence, using iEEG, an electrophysiological method with excellent spatial resolution, it was demonstrated that in the parabelt cortex of individual subjects there is a sharp functional boundary between the anterior and posterior STG, paralleling the findings from group maps created using postmortem anatomy or multimodal MRI. Critically, this functional difference was only evident with the use of multisensory stimuli (i.e., audiovisual speech) because both the anterior and posterior STG respond to unisensory auditory speech (either clear or noisy). It is postulated that the application of such multisensory approaches may allow for the differentiation of additional functionally distinct regions in the parabelt cortex and to other regions beyond the auditory cortex.

8.4 Posterior Boundary of the Parabelt Cortex

Anatomically, the posterior STS/STG is situated between the visual cortex and the auditory cortex, a finding consistent with the general organizational observation that multisensory zones exist at the borders between unisensory cortices (Wallace et al. 2004). If lateral temporal (STG/STS) regions that respond to auditory stimulation are considered as part of the parabelt cortex, the question arises: What is the posterior boundary of the parabelt cortex or where does the auditory cortex end and the visual cortex begin?

There is substantial evidence from fMRI that the inferotemporal sulcus (ITS) is a reasonable boundary for the transition from the visual cortex to the multisensory cortex. Two visual areas are situated along the ITS, area MT, which is typically located on the posterior bank of the ascending limb of the ITS, and area MST, which is typically located on the anterior bank of the ITS. These two areas, although both highly responsive to visual motion, have markedly different response properties. In macaque monkeys, single neurons in area MST are multisensory, responding to both visual and vestibular stimuli, potentially reflecting the role of this area in spatial navigation and postural control (Takahashi et al. 2007). In contrast, similar recordings from area MT in macaques reveal this area to be almost exclusively responsive to visual motion stimuli. A second difference between these areas is that whereas area MT appears to only respond to visual stimuli in the contralateral visual field, area MST responds to both contralateral and ipsilateral visual stimuli (Huk et al. 2002). When using fMRI to measure activation patterns to simple vibrotactile stimuli, area MT was found to respond only to visual stimuli, whereas area MST was found to respond to both visual and somatosensory stimuli (Fig. 8.4; Beauchamp et al. 2007). Therefore, the fundus of the inferotemporal sulcus (which typically corresponds to the border between areas MT and MST) appears to represent the functional boundary between the multisensory cortex (area MST) and visual cortex (area MT; Jiang et al. 2015).

One potential objection to this schema is the claim in the literature that MT in humans is, in fact, multisensory. A number of prior studies have claimed that MT responds to tactile motion (as well as to visual motion), such as an (unseen) brush stroking the arm (Hagen et al. 2002; van Kemenade et al. 2014). These results have been interpreted to mean that in humans, area MT is multisensory and, more generally, serves as a cross-modal motion-processing module (Pascual-Leone and Hamilton 2001; Ricciardi et al. 2007). However, a recent attempt to replicate these results (Hagen et al. 2002) found that in any individual subject, there was no overlap between visual and tactile motion activations in and around area MT (Fig. 8.4A).

So how can these seemingly contradictory results be reconciled? First, some of these studies did not actually locate area MT in individual subjects, instead relying on stereotactic coordinate values (Matteau et al. 2010). This is problematic because atlas values are blind to anatomical or functional landmarks, and it is known that the location of area MT in any individual can vary by a centimeter or more. Thus, multisensory activity on one side of the ITS can easily be confused with unisensory visual activity on the other bank of the ITS. Other studies rely on group-average activation maps to compare the location of tactile and visual motion activations (Ricciardi et al. 2007; Summers et al. 2009). The problem with this approach is illustrated in Fig. 8.4B, in which the absence of overlap between tactile and visual motion in any individual subject can result in an overlapping group activation map. Once again, this misleading activation pattern in the group map can be attributed to variability in the spatial location of activity across individual subjects. Averaging across individuals to create a group map acts as a blurring filter, taking two distinct tactile and visual motion peaks and merging them together. A simple illustration of this effect is shown in Fig. 8.4C. Although in any individual automobile, the front seats and back seats are in completely separate locations along the anterior-to-posterior axis of the auto, a map of the average seat locations across vehicles shows substantial overlap. However, the inference that the front and back seats are in the same location is obviously false. By extension, the use of group activation maps and meta-analysis techniques such as activation-likelihood estimation (Eickhoff et al. 2012) that creates what is, in effect, group maps by combining data from multiple studies must be used with extreme caution when making inferences about the anatomical organization multisensory responses.

A final set of observations relevant to the multisensory character of area MT is the fact that work in the blind and those with some degree of sight restoration have suggested that this area can support the processing of auditory motion in the absence

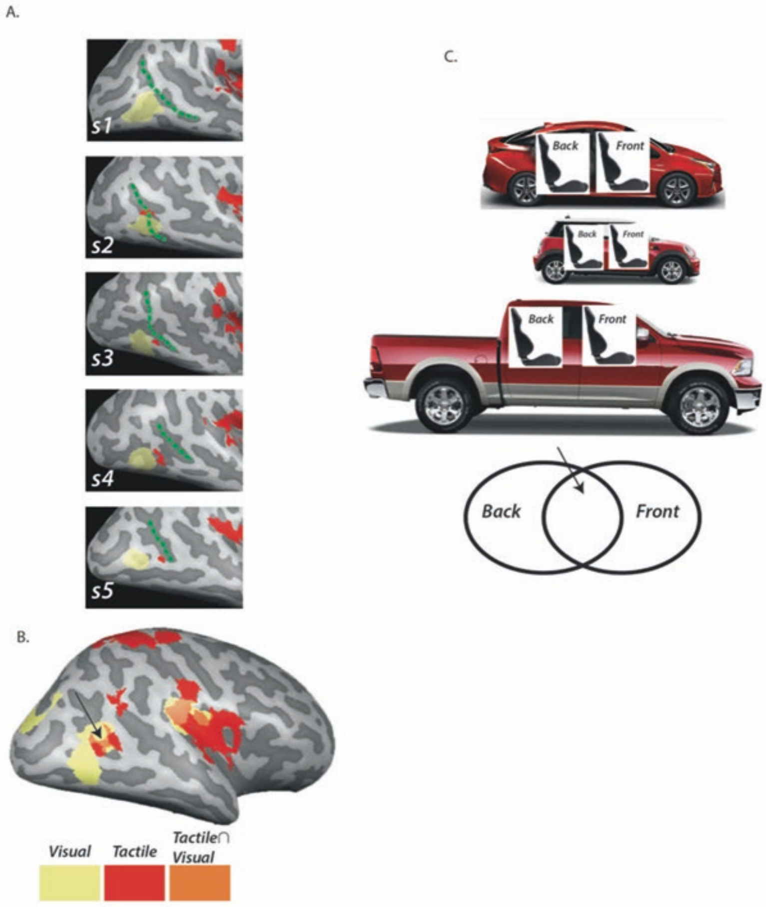

Fig. 8.4 Posterior boundary of parabelt cortex. (**A**) Functional magnetic resonance imaging (fMRI) activation maps from 5 subjects (s1 to s5) show the lateral views of inflated cortical surface models of the gray-white matter boundary for the right hemisphere. *Yellow*, areas with significantly greater blood oxygen level-dependent (BOLD) signal ($t > 2$ uncorrected [t statistic]) for a visual stimulus of moving dots compared with the fixation baseline in the vicinity of human area MT; *red*, areas with significantly greater BOLD signal ($t > 2$ uncorrected) for auditory sounds compared with the fixation baseline; *orange*, areas with significant activity for both visual and tactile stimuli. Adapted from Jiang et al. (2015). (**B**) Group map constructed from the individual subjects shown in **A**. Note that the group map shows overlapping tactile and visual activity in the vicinity of area MT (*black arrow*) even though this overlap is not present in any individual subject. Adapted from Jiang et al. (2015). (**C**) Example illustrating how average position maps can lead to incorrect inferences. In different vehicles, front and back seats are always in different spatial locations. A group map showing the average location of front and backs seats shows overlap between their positions (*black arrow*) even though this overlap is not present in any individual vehicle

of normal visual input (Saenz et al. 2008; Jiang et al. 2016). Although these examples are evidence for cross-modal plasticity, they highlight that area MT has some degree of multisensory or "supramodal" character that may contribute to the confusion as to whether it is truly a visual or a multisensory area.

Although this example of MT and MST illustrates some of the difficulties in drawing distinctions between multisensory and unisensory brain regions, the same challenges and issues are likely to apply to a host of brain regions that have been characterized as multisensory using methods such as fMRI and in which the spatial resolution is sufficiently coarse to result in the blurring of true functional distinctions. Thus, similar arguments can be applied to parietal cortical regions that lie along the intraparietal sulcus (IPS), and that is interposed between more posterior occipital (visual) cortical regions and more ventral temporal (auditory) cortical regions. Although invasive physiological studies in animal models have established the multisensory character of a number of the divisions of the IPS and the important role these areas play in saccadic eye and reach movements and spatial attention (Snyder et al. 1997; Grefkes and Fink 2005), a number of human imaging studies focused on the IPS are subject to the same concerns as articulated in Fig. 8.4C in regard to spatial blurring and the potential overestimation of true regions of multisensory convergence and integration.

8.5 Difficulties with the Use of Task-Based Functional Magnetic Resonance Imaging to Demarcate Area Boundaries

In addition to the concerns about spatial blurring in cross-subject analyses, there are other are fundamental difficulties with using task-based fMRI on its own to define area boundaries. Defining a given cortical area using task-based fMRI requires the use of statistical criteria, with different criteria giving different results (Beauchamp 2005). Even if the statistical criterion is fixed across subjects, a ubiquitous observation is that there is remarkable interindividual variability in fMRI activation maps. For instance, Fig. 8.5 shows the activation patterns in six healthy subjects during the presentation of unisensory auditory and visual stimuli, with multisensory activations generated using a simple conjunction analysis criterion ($t > 2$ for auditory stimulation and $t > 2$ for visual stimulation [t statistic]). Not surprisingly, visual activations are concentrated in the occipital lobe, whereas auditory activations are concentrated in the superior temporal gyrus. Regions responsive to both visual and auditory stimuli are found in the posterior superior temporal sulcus and gyrus (pSTS)/STG, with most activations located between the boundaries of HG and the ITS. However, in general, the activation in these areas is patchy, making it difficult to delineate sharp boundaries between unisensory and multisensory cortices.

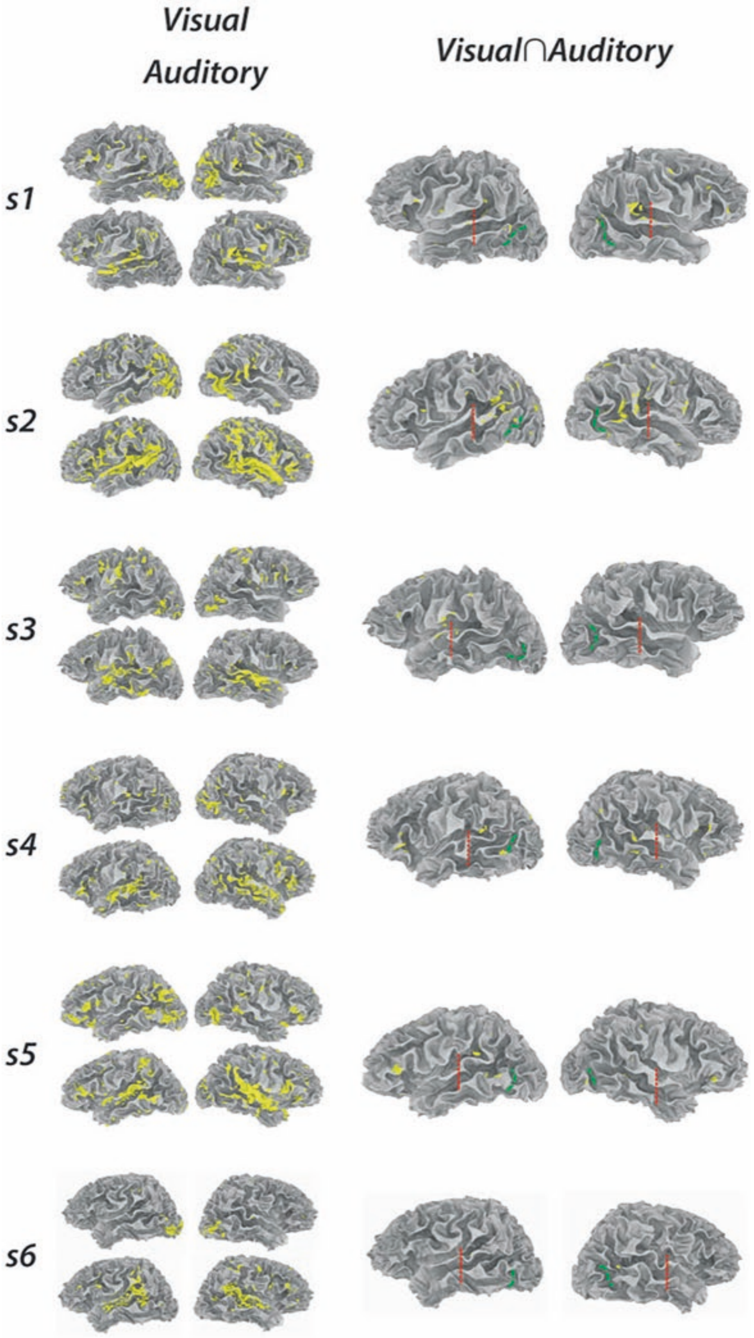

Fig. 8.5 Auditory and visual single subject activation maps. fMRI activation maps from 6 subjects (1 subject per row, s1 to s6) show the lateral views of surface models of the gray-white matter

A related problem is that the definition of a "significant" response in fMRI is strongly dependent on the amount of data collected. Gonzalez-Castillo et al. (2012) measured brain activity with fMRI while subjects viewed a simple visual stimulus consisting of a flashing checkerboard. As opposed to traditional designs for such studies, which would entail 4–6 "runs" or blocks of data collection, the authors carried out 100 runs per subject. Their remarkable observation was that the number of active voxels continued to increase as more data were collected, with no plateau. Using all 100 runs, about 96% of the entire brain was active in response to the simple visual stimulus. Although this is an impractical amount of data to collect under most circumstances, increases in the signal-to-noise ratio with improved scanner hardware, pulse sequences, and cardiac and respiratory noise removal means that similar results could soon be achieved with much less data. Similar results arise from the use of multivoxel pattern analysis methods that attempt to effectively "decode" the presence (or identity) of a given stimulus within a given brain region (cluster of voxels). Increasingly, such studies are illustrating the capacity of unisensory cortices to decode stimuli presented in other modalities; for instance, the visual cortex can decode the identity of auditory stimuli above chance. As a natural extension of these univariate and multivariate analyses, with sufficient data, it is very clear that the entire brain can ultimately be labeled "multisensory." Indeed, an influential review (Ghazanfar and Schroeder 2006) had the provocative title "Is Neocortex Essentially Multisensory" and highlighted the growing number of studies that were illustrating some degree of multisensory influence even in low-level (i.e., primary) sensory cortices. Although this work has changed the traditional views of the sensory cortex, it also serves to unintentionally blur the true functional distinctions between primary sensory cortices (which can be influenced by other senses but which are still largely unisensory) and multisensory brain regions (those with active convergence and integration of information from multiple sensory modalities). Indeed, as highlighted in earlier chapters (see Willet, Groh, and Maddox, Chap. 5; King, Hammond-Kenny, and Nodal, Chap. 6), the evidence for multisensory convergence (as well as some forms of integration) now extends to a number of subcortical loci beyond the classic multisensory structure, the superior colliculus, in which many of the initial formative observations about the behavior of multisensory neurons were first made (Stein and Meredith 1993).

Fig. 8.5 (continued) boundary for the left (*left columns*) and right (*right columns*) hemispheres for three conditions. In the visual condition (*top rows in 2 left columns*), *yellow* shows areas with significantly greater BOLD signal ($t > 2$ uncorrected) for a visual stimulus of moving dots compared with the fixation baseline. In the auditory condition (*bottom rows in 2 left columns*), *yellow* shows areas with significantly greater BOLD signal ($t > 2$ uncorrected) for auditory sounds compared with the fixation baseline. In the conjunction condition (*2 right columns*), *yellow* shows areas with significant activity for both visual and auditory stimuli. *Red dashed line*, Heschl's gyrus landmark; *green dashed line*, inferior temporal sulcus boundary

8.6 Summary and Future Directions

Although a great deal of focus has been placed on understanding the structural and functional organization of the auditory cortex, this work has (not surprisingly) had a strong emphasis on the processing of auditory signals. This chapter posits that a greater understanding of the functional role played by the auditory cortex can also be gained through bringing a multisensory lens to studies of these brain regions. This approach becomes increasingly important as one moves outward from the predominantly auditory core regions into the increasingly complex and multisensory belt and parabelt regions, where influences from other sensory modalities become increasingly prevalent. One of the best illustrations of this comes in the context of naturalistic speech, which is generally encountered in an audiovisual manner in which the spoken signal is accompanied by correlated visual signals largely associated with mouth movements (see Grant and Bernstein, Chap. 3). Indeed, using such naturalistic speech in both quiet and noisy settings reveals a functional distinction in parabelt regions not evident in auditory signals alone.

Given the spatial limitations of fMRI (even at high field) and the enormous amount of temporal information available in other electrophysiological approaches that can be applied to human subjects (i.e., EEG, iEEG, and magnetoencephalography [MEG]), future work that employs a conjunctive set of approaches toward questions of the auditory and multisensory cortical processes are likely to reveal additional insights into the functional organization of this complex set of cortical regions.

Compliance with Ethics Requirement Michael S. Beauchamp declares that he has no conflict of interest.

References

Beauchamp, M. S. (2005). Statistical criteria in FMRI studies of multisensory integration. *Neuroinformatics, 3*(2), 93–113.

Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004a). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron, 41*(5), 809–823.

Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., & Martin, A. (2004b). Unraveling multisensory integration: Patchy organization within human STS multisensory cortex. *Nature Neuroscience, 7*(11), 1190–1192.

Beauchamp, M. S., Yasar, N. E., Kishan, N., & Ro, T. (2007). Human MST but not MT responds to tactile stimulation. *The Journal of Neuroscience, 27*(31), 8261–8267.

Brodmann, K. (1994). Vergleichende Lokalisationslehre der Grosshirnrinde [Brodmann's Localization in the Cerebral Cortex] (L. J. Garey, Trans.) Leipzig: Barth. London: Smith Gordon (Original work published in 1909).

Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology, 46*(2), 369–384.

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage, 59*(3), 2349–2361.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*(1), 1–47.

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences, 10*(6), 278–285.

Glasser, M. F., & Van Essen, D. C. (2011). Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *The Journal of Neuroscience, 31*(32), 11597–11616.

Glasser, M. F., Goyal, M. S., Preuss, T. M., Raichle, M. E., & Van Essen, D. C. (2014). Trends and properties of human cerebral cortex: Correlations with cortical myelin content. *NeuroImage, 93*, 165–175.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2016). A multi-modal parcellation of human cerebral cortex. *Nature, 536*(7), 171–178.

Gonzalez-Castillo, J., Saad, Z. S., Handwerker, D. A., Inati, S. J., Brenowitz, N., & Bandettini, P. A. (2012). Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis. *Proceedings of the National Academy of Sciences of the United States of America, 109*(14), 5487–5492.

Grefkes, C., & Fink, G. R. (2005). The functional organization of the intraparietal sulcus in humans and monkeys. *Journal of Anatomy, 207*(1), 3–17.

Hagen, M. C., Franzen, O., McGlone, F., Essick, G., Dancer, C., & Pardo, J. V. (2002). Tactile motion activates the human middle temporal/V5 (MT/V5) complex. *European Journal of Neuroscience, 16*(5), 957–964.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *The Journal of Neuroscience, 22*(16), 7195–7205.

Jiang, F., Beauchamp, M. S., & Fine, I. (2015). Re-examining overlap between tactile and visual motion responses within hMT+ and STS. *NeuroImage, 119*, 187–196.

Jiang, F., Stecker, G. C., Boynton, G. M., & Fine, I. (2016). Early blindness results in developmental plasticity for auditory motion processing within auditory and occipital cortex. *Frontiers in Human Neuroscience, 10*, 324.

Kaas, J. H., Hackett, T. A., & Tramo, M. J. (1999). Auditory processing in primate cerebral cortex. *Current Opinion in Neurobiology, 9*(2), 164–170.

Matteau, I., Kupers, R., Ricciardi, E., Pietrini, P., & Ptito, M. (2010). Beyond visual, aural and haptic movement perception: hMT+ is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals. *Brain Research Bulletin, 82*(5–6), 264–270.

Moerel, M., De Martino, F., & Formisano, E. (2014). An anatomical and functional topography of human auditory cortical areas. *Frontiers in Neuroscience, 8*, 225.

Morosan, P., Schleicher, A., Amunts, K., & Zilles, K. (2005). Multimodal architectonic mapping of human superior temporal gyrus. *Anatomy and Embryology, 210*(5–6), 401–406.

Ozker, M., Schepers, I. M., Magnotti, J. F., Yoshor, D., & Beauchamp, M. S. (2017). A double dissociation between anterior and posterior superior temporal gyrus for processing audiovisual speech demonstrated by electrocorticography. *Journal of Cognitive Neuroscience, 29*(6), 1044–1060.

Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. *Progress in Brain Research, 134*, 427–445.

Ricciardi, E., Vanello, N., Sani, L., Gentili, C., Scilingo, E. P., Landini, L., Guazzelli, M., Bicchi, A., Haxby, J. V., & Pietrini, P. (2007). The effect of visual experience on the development of functional architecture in hMT+. *Cerebral Cortex, 17*(12), 2933–2939.

Saenz, M., Lewis, L. B., Huth, A. G., Fine, I., & Koch, C. (2008). Visual motion area MT+/V5 responds to auditory motion in human sight-recovery subjects. *The Journal of Neuroscience, 28*(20), 5141–5148.

Snyder, L. H., Batista, A. P., & Andersen, R. A. (1997). Coding of intention in the posterior parietal cortex. *Nature, 386*(6621), 167–170.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: The MIT Press.

Summers, I. R., Francis, S. T., Bowtell, R. W., McGlone, F. P., & Clemence, M. (2009). A functional-magnetic-resonance-imaging investigation of cortical activation from moving vibrotactile stimuli on the fingertip. *The Journal of the Acoustical Society of America, 125*(2), 1033–1039.

Takahashi, K., Gu, Y., May, P. J., Newlands, S. D., Deangelis, G. C., & Angelaki, D. E. (2007). Multimodal coding of three-dimensional rotation and translation in area MSTd: Comparison of visual and vestibular selectivity. *The Journal of Neuroscience, 27*(36), 9742–9756.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., Ugurbil, K., & Wu-Minn HCP Consortium. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage, 80*, 62–79.

van Kemenade, B. M., Seymour, K., Wacker, E., Spitzer, B., Blankenburg, F., & Sterzer, P. (2014). Tactile and visual motion direction processing in hMT+/V5. *NeuroImage, 84*, 420–427.

Wallace, M. T., Ramachandran, R., & Stein, B. E. (2004). A revised view of sensory cortical parcellation. *Proceedings of the National Academy of Sciences of the United States of America, 101*(7), 2167–2172.