

Auditory-Somatosensory Multisensory Processing in Auditory Association Cortex: An fMRI Study

JOHN J. FOXE,^{1–3} GLENN R. WYLIE,¹ ANTIGONA MARTINEZ,¹ CHARLES E. SCHROEDER,^{1,2} DANIEL C. JAVITT,^{1,4} DAVID GUILFOYLE,⁵ WALTER RITTER,¹ AND MICAH M. MURRAY^{1,2}

¹The Cognitive Neurophysiology Laboratory, Cognitive Neuroscience and Schizophrenia Program, and ⁵Center for Advanced Brain Imaging, Nathan S. Kline Institute for Psychiatric Research, Orangeburg 10962; ²Department of Neuroscience and ³Department of Psychiatry and Behavioral Science, Albert Einstein College of Medicine, Bronx 10461; ⁴Department of Psychiatry, New York University School of Medicine, New York City, New York 10016

Received 20 August 2001; accepted in final form 2 April 2002

Foxe, John J., Glenn R. Wylie, Antigona Martinez, Charles E. Schroeder, Daniel C. Javitt, David Guilfoyle, Walter Ritter, and Micah M. Murray. Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J Neurophysiol* 88: 540–543, 2002; 10.1152/jn.00694.2001. Using high-field (3 Tesla) functional magnetic resonance imaging (fMRI), we demonstrate that auditory and somatosensory inputs converge in a subregion of human auditory cortex along the superior temporal gyrus. Further, simultaneous stimulation in both sensory modalities resulted in activity exceeding that predicted by summing the responses to the unisensory inputs, thereby showing multisensory integration in this convergence region. Recently, intracranial recordings in macaque monkeys have shown similar auditory-somatosensory convergence in a subregion of auditory cortex directly caudomedial to primary auditory cortex (area CM). The multisensory region identified in the present investigation may be the human homologue of CM. Our finding of auditory-somatosensory convergence in early auditory cortices contributes to mounting evidence for multisensory integration early in the cortical processing hierarchy, in brain regions that were previously assumed to be unisensory.

INTRODUCTION

Inputs from the different sensory modalities are combined to form a single integrated experience of the world (Stein and Meredith 1993). For this integration to occur, the various sensory inputs must ultimately converge on the same neurons. At the cortical level, it has been widely assumed that information from the different sensory systems is first processed extensively through unisensory systems and then integrated in multisensory regions at a higher level of the hierarchical processing stream (e.g., Massaro 1998).

Recently however, both electrophysiological (Foxe et al. 2000; Giard and Peronnet 1999; Molholm et al. 2002) and functional imaging (e.g., Calvert 1999, 2001; Macaluso et al. 2000) studies in humans have suggested that cortical multisensory integration can occur at very early stages of the cortical processing hierarchy—stages previously thought to be purely unisensory. Moreover, intracranial recordings in macaques have shown direct somatosensory input to the caudomedial

(CM) belt area of auditory association cortex (Schroeder and Foxe 2002; Schroeder et al. 2001). Critically, both the auditory and somatosensory activation profiles in CM showed characteristic feedforward patterns with initial excitation beginning in Lamina 4 at very early latencies (approximately 12 ms for both). As CM is directly adjacent to A1, it was clear from these data that multisensory convergence in the macaque was occurring at early stages of the cortical processing hierarchy in regions that until then were assumed to be unisensory.

The present investigation employed high-field functional magnetic resonance imaging (fMRI) to investigate auditory-somatosensory convergence in auditory association cortex of the human. Our findings confirm the view that in humans, as in monkeys, auditory-somatosensory integration occurs early in the auditory cortical hierarchy. These data contribute to mounting evidence that multisensory processing involves integration at early stages of the cortical processing stream.

METHODS

Subjects

Twelve subjects (4 males) participated [age, 24.8 ± 5.1 (SD) yr]. All were right handed and were neurologically normal. All subjects provided written, informed consent, and were paid for their participation. The Institutional Review Board of the Nathan Kline Institute approved all procedures.

Design

A pseudo-random block (“boxcar”) design was used. There were three main conditions: auditory stimulation alone (Aud), somatosensory stimulation alone (Soma), and simultaneous auditory and somatosensory stimulation (Aud + Soma). In a single session, each of the three main conditions was presented for 30 s separated by 15 s of rest. Each session both began and ended with a block of rest. The order of the three main conditions was counterbalanced. Full counterbalancing resulted in three sequences, but due to software problems, only two of these were used: Aud + Soma..Aud..Soma; Aud..Aud + Soma..Soma. Two sessions per subject were included in the analysis.

Address for reprint requests: J. J. Foxe, The Cognitive Neurophysiology Laboratory, Cognitive Neuroscience and Schizophrenia Program, Nathan Kline Institute for Psychiatric Research, 140 Old Orangeburg Rd., Orangeburg, NY 10962 (E-mail: foxe@nki.rfmh.org).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Stimuli

AUDITORY. Auditory stimulation consisted of a 962 ms binaural broadband stimulus that was digitally modified to approximate the sound made by sandpaper on a rough surface. The frequency content was that of band-passed white noise such that the log spectral amplitude was relatively constant over frequencies ranging from 0 to 10 kHz. During auditory stimulation blocks, presentation rate was 1 Hz. Intensity was determined on a subject-by-subject basis whereby a mock EPI sequence was run and the sound level of the stimulus was adjusted to a level at which the subject could clearly hear the scratching sound above the magnet noise but at which it was not uncomfortably loud.

SOMATOSENSORY. Somatosensory stimulation was accomplished by manual rotation of a wooden "roller" device against the tips of the index and middle fingers of the right hand of each subject. The rubbing end of the device was covered with 100C-weight sandpaper to provide strong input.

Procedure

Prior to entering the magnet, subjects were informed that they would simply be required to lie in the scanner with eyes closed and passively experience the three conditions. Subjects were fitted with pneumatic ear inserts, and the ears were then covered by custom-built foam earmuffs that served to further attenuate magnet noise while allowing for the delivery of the auditory stimulation. Foam padding was placed around the right arm and hand, allowing them to comfortably extend the hand to the mouth of the bore. One experimenter stood at the mouth of the bore, watching a projection screen. This screen was provided entirely for the experimenter's benefit and was not visible to the subject. When the Aud and Rest conditions were being presented, the word "REST" was presented on the screen, and the experimenter did nothing. When the Soma and the Aud + Soma conditions were being presented, an arrow was presented that alternated direction on the screen at 1-s intervals. This allowed the experimenter to pace the stimulation being delivered, rotating the roller device back and forth in synchrony with the arrows. For the Soma condition, this procedure simply ensured that the stimulation was constant across blocks. For the A + S condition, this procedure ensured that the somatosensory stimulation was roughly synchronous with the sounds. In this condition, each time an arrow was presented on the projection screen, the scratching sound was simultaneously delivered to the subject. The experimenter knew that 30 arrows would

be presented in any block requiring somatosensory stimulation. To further aid the experimenter, the word "REST" was extinguished for 500 ms prior to the beginning of any block requiring somatosensory stimulation to cue them to the beginning of the block.

fMRI data acquisition and analysis

T2*-weighted echo-planar images (EPI; TR/TE/flip angle = 3 s · 40 ms · 90°; voxel size = 3.5 mm³; matrix size = 64 × 64) were acquired on a 3-Tesla SMIS (Marconi) system. In each session, 60 volumes (31 contiguous slices covering the entire brain) were acquired in the axial plane. The first five images were discarded to allow for stabilization of the blood oxygenation level dependent (BOLD) signal. In all subjects, head motion never exceeded 0.75 mm along any axis.

All data processing and analyses were conducted using the AFNI image-analysis software package (Cox 1996). Images were realigned to an image at approximately the mid-point of the time series acquisition (the 25th image) and normalized into Talairach coordinates (Talairach and Tournoux 1988). To estimate the BOLD response associated with each stimulus condition (Aud, Soma, Aud + Soma), regressors modeling the timing of each stimulation epoch were convolved with a canonical hemodynamic response function, and used in a multiple linear regression analysis. The regression coefficients resulting from this analysis were contrasted in three general linear tests: auditory versus rest; somatosensory versus rest; and multisensory versus (auditory + somatosensory). The resulting individual *F* maps were entered into a between-subjects *t*-test that compared the *F* values for each contrast against the null hypothesis. The results ($P < 0.05$) from this random-effects analysis are presented in Fig. 1 and Table 1. Note that extent of activation is reported in units of volume (microliters), where 64 μ l is equivalent to 1 cubic voxel.

For anatomical localization, high-resolution (.97 × .97 × 1.3 mm) T1-weighted images of the whole brain were acquired using a standard three-dimensional magnetization prepared rapid gradient echo (MPRAGE) pulse sequence. Both anatomical and T2*-weighted images were transformed into the standardized coordinate system of Talairach and Tournoux (1988). Anatomical images from each of the 12 subjects were averaged and used for two-dimensional visualization of the group-averaged functional data (Fig. 1D). Additionally, the structural images from one subject were used to render the cortical surface, according to the method of Sereno, Dale and colleagues (1995). The group-averaged anatomical dataset was registered to the single-subject cortical rendering using blink comparison. The result-

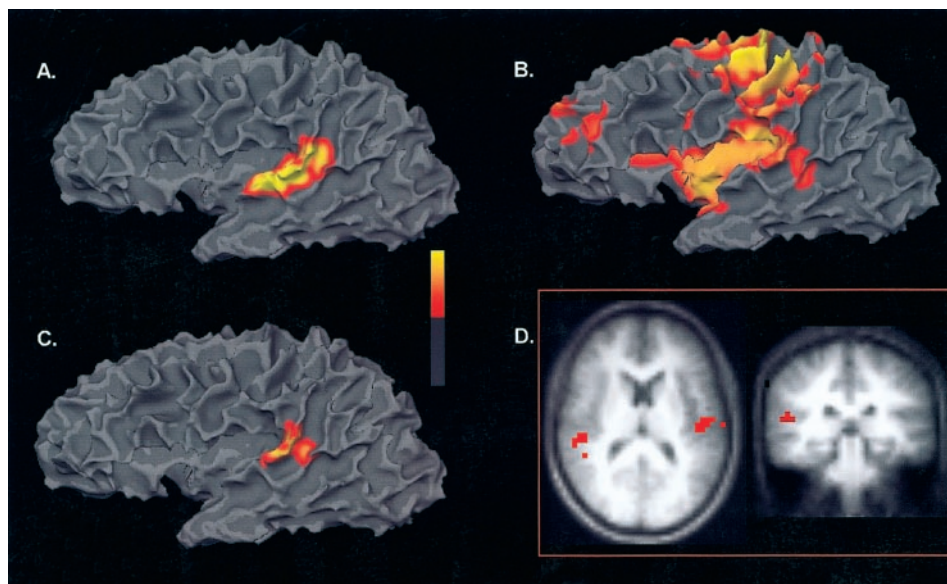


FIG. 1. Group ($n = 12$) activation maps for auditory (A) and somatosensory (B) stimulation, projected onto a 3-dimensional rendering of the left hemisphere, which was derived from one representative subject's anatomic magnetic resonance imaging scan. The color scale indicates significance values from $P < 0.05$ (red) to $P < 0.001$ (yellow). The region of significant overlap between these 2 activation maps (Aud \cap Soma) is shown projected onto the same rendered surface (C). D: the overlap fields in both an axial slice ($z = 12$), which shows overlap fields in both hemispheres, and in a coronal slice ($y = -32$), which shows the overlap field in the posterior portion of the left auditory cortex. These slices were derived from a group-averaged brain, obtained by averaging the anatomic scans of the 12 subjects who participated in the study.

TABLE 1. Results from random-effects analysis

	Brain Region	Brodman Area	Talairach <i>x, y, z</i>	Volume, μ l	$T_{(1,11)}$	<i>P</i> Value (2-Tailed)
AUD v Rest	L. Sup. Temp. Gyrus		-49, -29, 10	3968	2.89	<0.02
			-46, -23, 06		3.67	<0.01
	R. Sup. Temp. Gyrus	BA 22	48, -18, 10	2944	2.57	<0.05
			48, -14, 05		3.27	<0.01
SOM v Rest	R. Sup. Frontal Gyrus	BA 10	26, 59, 22	320	2.96	<0.02
	R. Mid. Temp. Gyrus	BA 39	34, -55, 29	256	2.52	<0.05
	L. Pre/Post Central Gyri	BA 1-4,6	-35, -18, 60	31488	3.52	<0.005
			-53, -21, 55		8.23	<0.001
	L. Insula	BA 13	-38, -10, 11	19136	3.92	<0.005
	R. Insula	BA 13	46, -12, 13	14400	2.52	<0.05
	R. Inf. Parietal Lobe	BA 40	51, -54, 43	2496	2.79	<0.02
	R. Pre/Post Central Gyri	BA 1 & 4	55, -13, 41	2048	2.73	<0.02
	R. Mid. Temp. Gyrus	BA 21	62, -53, 07	704	2.61	<0.05
	R. Sup. Frontal Gyrus	BA 9	24, 57, 30	448	3.16	<0.01
	R. Cerebellum		25, -51, -16	3584	4.91	<0.001
		15, -66, -13	640	3.70	<0.005	
AUD/SOM overlap	L. Sup. Temp. Gyrus	BA 22/39	-50, -32, 14	832	3.07 (Aud)	<0.02
					2.94 (Som)	<0.02
	R. Sup. Temp. Gyrus	BA 22	48, -16, 10	1024	2.57 (Aud)	<0.05
					2.81 (Som)	<0.02

Talairach coordinates are given for the center of mass of a given cluster. *T* values are for these center-of-mass coordinates. In key regions, the coordinates of maximal activation are also reported (gray shading).

ing transformation matrix was used to register the group functional activation maps to the rendered surface.

RESULTS

The first contrast made was between the Aud and Rest conditions. As expected, activations conforming to our inclusion criteria were seen along and surrounding both the left and right superior temporal gyri [primarily Brodmann area (BA) 22; see Table 1]. These two large extended activations included primary auditory cortex bilaterally and surrounding auditory association areas in the belt and parabelt regions. It is worth noting that activation appeared to be stronger for these environmental sounds in the left auditory cortex than in the right. Significant activations were also found in the right superior frontal gyrus and right middle temporal gyrus.

The second contrast was between the Soma and Rest conditions and resulted in an extensive network of activations. These included a very large contiguous activation that spanned the left pre- and post-central gyrus (BA 1-3, BA 4, BA 6) and an extended region of activation in both the left and right insulae (BA 41). The peak activation was seen in BA 1 consistent with activation of the hand area in primary somatosensory cortex (S1). Secondary somatosensory cortex (SII) was also extensively activated (see Fig. 1). Other activated regions included the right inferior parietal lobe, right somatosensory cortex, the superior frontal gyrus near the midline and a number of large clusters in both the anterior and posterior lobes of the right cerebellum (see Table 1).

To assess areas of significant overlap (i.e., multisensory convergence), we determined regions that were significantly activated by both the auditory-alone condition and the somatosensory-alone condition. This analysis resulted in an overlap cluster in the posterior portion of auditory cortex along the left superior temporal gyrus (BA 22/39). A region of overlap was also found in the right auditory cortex (BA22).

The final contrast assessed multisensory interaction effects within the overlap region: (Aud + Soma vs. Rest) - [(Aud vs.

Rest) + (Soma vs. Rest)]. We restricted this analysis to the area of significant overlap in the left hemisphere, based on our hypothesis of auditory-somatosensory interactions in the superior temporal gyrus. In line with our prediction, a cluster of voxels within the overlap region (192 μ l) showed significantly more activation [$T(1,11) = 2.65, P < 0.022$] for simultaneous auditory-somatosensory stimulation than when the respective unimodal responses were summed.

Note that in this final comparison, in addition to the cluster within the overlap region, a system of additional areas was seen to show superadditive effects (not reported here). As we had no specific hypotheses regarding these regions, these findings will be used as a hypothesis generation tool for future studies.

DISCUSSION

The present study demonstrates convergence of somatosensory and auditory inputs within a region of human auditory cortex. Simultaneous stimulation in both modalities resulted in significantly greater activity than would be predicted by summing the responses to the constituent unimodal stimuli, indicating that this area is involved in multisensory integration. These findings are in close agreement with recent electrophysiological findings in awake macaques detailing an area of the auditory association cortex that also showed auditory-somatosensory convergence (Schroeder and Foxe 2002; Schroeder et al. 2001). This region (CM) is directly adjacent to primary auditory cortex and receives early feedforward inputs from both the auditory and somatosensory systems. Our findings provide tentative identification of a human homologue to macaque area CM.

Involvement of auditory cortex in auditory-somatosensory (AS) integration is also in good agreement with a recent high-density electrical mapping study (Foxe et al. 2000). This study showed AS interactions occurring at relatively early latencies in humans, with a topographic distribution over central and postcentral scalp that suggested generators in early

sensory cortices of the somatosensory and auditory systems (see also Lütkenhöner et al. 2002). In light of the close correspondence between the monkey and human data, we propose that there is likely to be a significant feedforward contribution to auditory-somatosensory integration in human auditory cortex.

While both the human and macaque data clearly support the presence of early feedforward multisensory integration, as yet, little is known about the functional role played by such early integration. Perhaps the obvious candidate is the early localization of multisensory objects in space. The early detection and localization of moving or perhaps threatening objects, has clear implications for survival and the presence of coincident sensory inputs is well known to improve detection and localization (e.g., Stein et al. 1996). The anatomic location of CM in posterior auditory cortices is in keeping with this localization role, as intracranial recordings in rhesus monkeys have implicated caudal belt areas in auditory spatial localization (Rauschecker and Tian 2000). Also, to solve the so-called “binding” problem, the system must determine what inputs from one sensory modality belong with what inputs from other modalities. As detailed, this has traditionally been thought to occur late in processing but certainly one alternative is that it occurs in early cortical areas that possess overlapping spatial maps, akin to those found in the superior colliculus (see Stein and Meredith 1993). Preliminary studies in area CM have begun to detail the somatosensory map there and promise to provide answers to this question (Fu et al. 2001).

We express our sincere appreciation to R. Sangoi for help with data acquisition and to B. Higgins and D. Foxe for excellent technical assistance.

This work supported in part by grants from the National Institute of Mental Health (MH-63434 to J. J. Foxe and MH-49334 to D. C. Javitt) and the Burroughs Wellcome Fund.

REFERENCES

- CALVERT GA, BRAMMER MJ, BULLMORE ET, CAMPBELL R, IVERSEN SD, AND DAVID AS. Response amplification in sensory-specific cortices during cross-modal binding. *Neuroreport* 10: 2619–2623, 1999.

- CALVERT GA, HANSEN PC, IVERSEN SD, AND BRAMMER MJ. Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the bold effect. *Neuroimage* 14: 427–438, 2001.
- COX RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29: 162–173, 1996.
- FOX E JJ, MOROCZ IA, MURRAY MM, HIGGINS BA, JAVITT DC, AND SCHROEDER CE. Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain Res Cogn Brain Res* 10: 77–83, 2000.
- FU KG, JOHNSON T, SHAH AS, ARNOLD L, SMILEY J, HACKETT TA, GARRAGHTY PE, AND SCHROEDER CE. Characterization of somatosensory input to auditory association cortex in macaques. *Soc Neurosci Abstr* 390, 2001.
- GIARD MH AND PERONNET F. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cognit Neurosci* 11: 473–490, 1999.
- LÜTKENHÖNER B, LAMMERTMANN C, SIMÕES C, AND HARI R. Magnetoencephalographic correlates of audiotactile interaction. *Neuroimage* 15: 509–522, 2002.
- MACALUSO E, FRITH CD, AND DRIVER J. Modulation of human visual cortex by crossmodal spatial attention. *Science* 289: 1206–1208, 2000.
- MASSARO DW. Speechreading: illusion or window into pattern recognition. *Trends Cognit Sci* 3: 310–317, 1998.
- MOLHOLM S, RITTER W, MURRAY MM, JAVITT DC, SCHROEDER CE, AND FOXE JJ. Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res Cogn Brain Res* 14: 121–134, 2002.
- RAUSCHER JP AND TIAN B. Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proc Natl Acad Sci USA* 97: 11800–11806, 2000.
- SCHROEDER CE AND FOXE JJ. Timing and laminar profile of converging inputs in multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res* 14: 195–207, 2002.
- SCHROEDER CE, LINDSLEY RW, SPECHT C, MARCOVICI A, SMILEY JF, AND JAVITT DC. Somatosensory input to auditory association cortex in the macaque monkey. *J Neurophysiol* 85: 1322–1327, 2001.
- SERENO MI, DALE AM, REPPAS JB, KWONG KK, BELLIVEAU JW, BRADY TJ, ROSEN BR, AND TOOTELL RB. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268: 889–893, 1995.
- STEIN BE, LONDON N, WILKINSON LK, AND PRICE DD. Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. *J Cognit Neurosci* 8: 497–506, 1996.
- STEIN BE AND MEREDITH MA. *The Merging of the Senses*. Cambridge, MA: MIT Press, 1993.
- TALAIRACH J AND TOURNOUX P. *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme, 1988.