Brief Communication

Auditory Cortical Neurons Respond to Somatosensory Stimulation

Kai-Ming G. Fu,1,2 Taylor A. Johnston,1,2 Ankoor S. Shah,1,2 Lori Arnold,3 John Smiley,2 Troy A. Hackett,4 Preston E. Garraghty,1 and Charles E. Schroeder1,2

1Albert Einstein College of Medicine, Bronx, New York 10461, 2Nathan Kline Institute, Orangeburg, New York 10962, 3Department of Psychology, Program in Neural Science, Indiana University, Bloomington, Indiana 47405, and 4Vanderbilt University, Nashville, Tennessee 37203

The prevailing hierarchical model of cortical sensory processing holds that early processing is specific to individual modalities and that combination of information from different modalities is deferred until higher-order stages of processing. In this paper, we present physiological evidence of multisensory convergence at an early stage of cortical auditory processing. We used multi-neuron cluster recordings, along with a limited sample of single-unit recordings, to determine whether neurons in the macaque auditory cortex respond to cutaneous stimulation. We found coextensive cutaneous and auditory responses in caudomedial auditory cortex, an area lying adjacent to A1, and at the second stage of the auditory cortical hierarchy. Somatosensory–auditory convergence in auditory cortex may underlie effects observed in human studies. Convergence of inputs from different sensory modalities at very early stages of cortical sensory processing has important implications for both our understanding of multisensory processing and established views of unisensory processing.

Key words: multisensory; auditory; somatosensory; convergence; single unit; cutaneous

Introduction

Because sight, sound, and touch sample unique dimensions of an object, combining information across sensory modalities provides the brain with converging evidence concerning the position, movement, and identity of an object. Multisensory convergence clearly does occur in numerous subcortical structures, including superior colliculus (Stein and Meredith, 1993), inferior colliculus (Groh et al., 2001), dorsal cochlear nucleus (Kanold and Young, 2001), and the reticular formation (Amassian and Devito, 1954; Bell et al., 1964). However, there remains a long-standing view that multisensory convergence in the neocortex is essentially a higher-order process, deferred until each unisensory bit is thoroughly processed through its specific sensory hierarchy (Jones and Powell, 1970). Consistent with this view, physiological studies in monkeys have thus far detected multisensory convergence mainly in higher-order areas of the parietal (Hyvarinen and Shelepkin, 1979; Mazzoni et al., 1996; Duhamel et al., 1998), temporal (Benevento et al., 1977; Leinonen et al., 1980; Bruce et al., 1981; Hikosaka et al., 1988), and frontal lobes (Benevento et al., 1977; Rizzolatti et al., 1981; Graziano et al., 1994). On the other hand, accumulating evidence from noninvasive brain measures in humans (Calvert et al., 1997; Levanen et al., 1998; Giard and Peronet, 1999; Foxe et al., 2002), direct measurements of neural activity in monkeys (Schroeder et al., 2001; Schroeder and Foxe, 2002), and cross connections between low-level cortices, including A1, V1, and S1 (Zhou and Fuster, 1996; Falchier et al., 2002; Rockland and Ojima, 2003), suggests that multisensory convergence may occur at early, putatively unisensory, cortical processing stages.

In this paper, we investigate the nature of somatosensory inputs to the physiologically and anatomically defined caudomedial (CM) region of macaque auditory cortex. CM is adjacent to primary auditory cortex (A1) and receives direct projections from it, as well as several thalamic nuclei (Kosaki et al., 1997; Hackett et al., 1998). CM is considered to be unisensory cortex that participates in early processing of sounds, especially complex noises (Rauschecker et al., 1995). Our previous studies suggest that CM is part of a group of posterior auditory association cortices whose neurons respond to somatosensory and visual, as well as auditory, stimuli (Schroeder et al., 2001; Schroeder and Foxe, 2002) and that these findings pertain to humans as well as monkeys (Foxe et al., 2002). The current study addressed the following questions. (1) Is there cutaneous input to CM, and, if so, which body surface(s) are represented? (2) What other somatosensory submodalities are represented within CM?

Materials and Methods

Acute microelectrode mapping studies were conducted in two macaque monkeys (3.0 kg Maccaca fuscata and 9.2 kg Maccaca mulatta). All surgical and experimental procedures were approved in advance by the Institutional Animal Care and Use Committee of the Nathan Kline Institute. Under anesthesia (20 mg/kg ketamine and 0.1 mg/kg xylazine) and related support medications (atropine, 0.5 mg/kg Dexamethazone,
**Results**

Figure 1 presents a summary of recording sites in the auditory cortex of one hemisphere, displayed on the anatomical reconstruction of that hemisphere, with the overlying cortex on the right side cut out to the level of the superior temporal plane. The area enclosed in the black outline corresponds to the core auditory areas A1 (posterior) and R (anterior). The large black patch posterior and medial to the core is a fluorescent marker injection made at the posterior margin of our recording field in this hemisphere. This was made to help register the penetration pattern onto the anatomy. The inset presents a composite of the penetrations from all four hemispheres, superimposed on the same anatomic reconstruction. The penetrations displaying only auditory responsiveness are shaded white, whereas those also displaying convergent somatosensory responsiveness are represented by colored circles. Because the majority of the penetrations were vertical ones, and many of these contained more than one recording site, the actual number of recording sites is under-represented here.

**Cutaneous and other submodality representations in CM**

Eighty-three percent of 101 recording sites in the superior temporal plane displayed a response to auditory stimulation. Of the auditory-responsive recording sites posterior to A1 (mainly CM), 72% (33 of 46) were responsive to some form of somatosensory stimulation. No recording site in A1 responded to any form of somatosensory stimulation. Those posteromedial sites displaying multisensory responsiveness had qualitative response properties characteristic of auditory association cortex (Jones et al., 1995; Kaas et al., 1999). That is, the neurons at these sites responded preferentially to complex noise stimuli and displayed broadband tuning relative to neurons in A1. In sites responsive to somatosensory stimulation, through evaluation of responsiveness over the entire body surface revealed a strong bias toward cutaneous input because 26 of 33 sites responded to cutaneous stimulation of the head and hands. The majority of these sites (20 of 26) responded to light stimulation with air puffs or von Frey hairs, whereas six responded to deep pressure stimulation. As discussed above, most of our recording sites were located in lamina 4 or lower lamina 3. Given the small number of observations of deep pressure stimulation, we were not able to resolve a laminar preference for deep versus cutaneous stimulation. Figure 2A illustrates where typical cutaneous receptive fields were encountered during the course of recording. Split receptive fields (e.g., responsiveness on both the forehead and occiput) were not encountered, and the top of the head was not examined because of surgical exposure. The most commonly encountered cutaneous receptive fields were located on the head and neck. Three sites responded to light touch, as well as to air-puff stimuli presented to the dorsal hand surface. Receptive fields located on the head

At the end of recording, monkeys were transcardially perfused with buffered 4% paraformaldehyde, followed by buffered sucrose solutions to cryoprotect the brain. Whole brains were sectioned at 80 μm thickness with a sliding microtome, and sections were stored in serial order in multi-welled plastic boxes. Every 12th section through the brain was Nissl stained with cresyl violet, and adjacent sections of series were processed for acetylcholinesterase histochemistry and parvalbumin immunoreactivity as described previously (Schroeder et al., 2001). During sectioning, a video image of each section was stored onto a Macintosh computer (Apple Computers, Cupertino, CA). The consecutive and aligned video images corresponding to the sections used for Nissl staining were later used for three-dimensional reconstruction of the tissue volume. Video images were processed using NIH Image software, and volume rendering was done with MedX (Ocala, FL) software. The boundary between primary (core) and belt regions of auditory cortex (see Fig. 1, dark line) was determined by examination of the series in which alternate sections were stained for Nissl substance or processed for acetylcholinesterase histochemistry and parvalbumin immunoreactivity.

250 mg/kg Claforan), the subjects underwent resection of the scalp and overlying fascia including peristeum, followed by craniotomy to expose the cortex. The dura was then incised and reflected, and the cortex was covered with silicon to prevent desiccation during recording. At this time, the subject was positioned using either a head post or hollow ear bars that allowed for presentation of auditory stimuli.

Single neurons and neuron clusters were recorded with tungsten microelectrodes (1.0-1.3 MΩ impedance) during penetrations through auditory cortex via both vertical and lateral penetration angles. For the vertical penetrations, one to three recording sites per penetration were examined. During the lateral penetrations, up to nine recording sites per penetration were examined. On the basis of the baseline neuronal firing rate and our stereotaxic depth measurements, most of our recording sites were located in lamina 4 or lower lamina 3. At each recording site, auditory responses were assessed using a combination of pure tones and complex noises (100 msec, 5 msec on-off ramps), presented using a Tucker-Davis Technologies (Gainesville, FL) System 3 apparatus. This apparatus provided a synchronous logic level output that was recorded concomitant with neuronal activity. Subject to later histological verification, the boundary between core area A1 and belt area CM was determined functionally using characteristic differences in auditory response properties (Rauschecker et al., 1995; Kosaki et al., 1997; Hackett et al., 1998; Recanzone et al., 2000; Schroeder et al., 2001). Also at each recording site, somatosensory responsiveness was assessed through testing with light cutaneous stimulation (supplied by wisps of cotton and von Frey hairs), deep pressure stimulation, joint manipulation, vibration, and air puffs (especially for hairy surfaces). In each case, evaluation included the entire body surface, with the exception of the area precluded from examination by exposure. Quantification of somatosensory responses was effected using an electronically gated air-puff circuit. Logic level output from a pulse generator both activated a solenoid that released a quantity of air (solonoid open time of 0.5 sec) and was recorded in the acquisition computer as a stimulus marker for quantification purposes. The length of the air line between the solenoid and the monkey introduced a delay between trigger acquisition and stimulus delivery. This delay was estimated before the experiment by comparing neuronal responses driven by the air puff with those driven by a time-locked electrical stimulus. This lag was then subtracted from the latency of the somatosensory data in Figure 3 to give an approximation of the real onset latency to air-puff stimulation. Although trial-to-trial temporal alignment is precise, estimation of absolute response onset latency is less so. In a control condition for auditory contamination of somatosensory stimuli (discussed in Results), we tested somatosensory responses after bilateral tympanic membrane destruction. In this condition, auditory stimulation was delivered through bone conduction, using the air-puff device to drive a slave cylinder that delivered light taps to the exposed skull (periorietum removed) over the midline, at a rate of one per second. At the outset, to be sure that we were dealing with a bone-conducted auditory-evoked response, we compared effects of tapping versus scraping, effects of tapping at different locations of exposed bone surfaces, and we tested for the threshold of the effect, by tapping the skull with a series of Semmes-Weinstein monofilament fibers. Although neural responses were extremely sensitive to temporal pattern (e.g., punctate tap versus prolonged scrape), they were remarkably insensitive to the location of the bone surface stimulated. Testing over a range of Semmes-Weinstein fiber sizes (corresponding to 1.08 gm of pressure) down to 1.65 (corresponding to 0.008 gm of pressure), we found responses persisting down to very light skull taps, at forces as little as 0.008 gm of pressure (fiber size, 2.36). In routine testing, it was not possible to calibrate the computer-controlled stimulator for very low force tapping, but we carefully adjusted the apparatus to make the tap as light as possible while still operating reliably. Signals for offline analysis were recorded using Neuroscan (El Paso, TX) Acquire, and subsequent analysis was performed using Neuroscan Edit and Matlab software (Mathworks, Natick, MA).

At the end of recording, monkeys were transcardially perfused with buffered 4% paraformaldehyde, followed by buffered sucrose solutions to cryoprotect the brain. Whole brains were sectioned at 80 μm thickness with a sliding microtome, and sections were stored in serial order in multi-welled plastic boxes. Every 12th section through the brain was

310

to cryoprotect the brain. Whole brains were sectioned at 80 μm thickness with a sliding microtome, and sections were stored in serial order in multi-welled plastic boxes. Every 12th section through the brain was

310

At the end of recording, monkeys were transcardially perfused with buffered 4% paraformaldehyde, followed by buffered sucrose solutions to cryoprotect the brain. Whole brains were sectioned at 80 μm thickness with a sliding microtome, and sections were stored in serial order in multi-welled plastic boxes. Every 12th section through the brain was
were only occasionally broad, that is, larger than that illustrated in Figure 2A (left), with less than a one-quarter (five) demonstrating responsiveness to bilateral stimulation. Approximately one-third (7 of 26) of the receptive fields were smaller than those shown in Figure 2A based on qualitative mapping with fine cotton wisps. Slightly over one-half (14 of 26) of the cutaneous receptive fields were of a moderate size comparable with that illustrated in Figure 2A. A few sites responded to noncutaneous somatosensory stimuli. In eight sites, for example, manipulation of the elbow joint or vibration produced neuronal responses. No site displayed responsiveness to more than one type of somatosensory stimulus. Figure 2B details the submodality preference breakdown for recording sites in posterior auditory cortex.

Controls for auditory contamination of somatosensory stimulation

Given the proclivity of CM neurons to respond to complex sounds (Rauschecker et al., 1995; Recanzone et al., 2000; Schroeder et al., 2001), it is important to control for the possibility that these neurons are responding to a slight noise associated with somatosensory stimulation. In our previous study (Schroeder et al., 2001), this was not a problem, because somatosensory stimulation was supplied by a completely silent electrical shock to the median nerve. We also noted in our previous studies that, when studying auditory–somatosensory convergence sites, loud mask-
recting the stimulus away from the hand, thus isolating any au-

study, a routine control was to test for an auditory input by di-

hand cutaneous receptive fields with air puffs in the present 

stimulation; this control was therefore avoided. When testing 

difficult to drive activity with somatosensory as well as auditory 

ing noise, by “overdriving” auditory neurons, could make it dif-

cult to drive activity with somatosensory as well as auditory 

stimation; this control was therefore avoided. When testing 

hand cutaneous receptive fields with air puffs in the present 

study, a routine control was to test for an auditory input by di-

recting the stimulus away from the hand, thus isolating any au-

ditory response to the noise of the air puff, presented at arm’s 

length from the head. For receptive fields on the head and neck, 

use of air puffs was impractical, because very close to the pinna, 

even the slight noise of the air puff is clearly detectable. In this 

case, only light cutaneous stimuli were applied, and bilateral tym-

panic membrane lesions in one subject provided the means of 

eliminating any (air-conducted) auditory concomitants of these 

stimuli. In this condition, auditory stimulation was effected via 

bone conduction. In the tympanic lesion condition as in the nor-

mal hearing condition, light touch with a cotton wisp was effec-

tive in driving neuronal responses in CM. In neither the tympanic 

membrane lesion nor the normal-hearing condition did light cutaneous 

stimuli drive A1 neurons.

Multisensory convergence at the single neuron level

Although most recordings were of multiunit activity, on four 

occasions we isolated single neurons well enough to assess mul-

tisensory convergence at this level. Two of the isolated neurons 

were multisensory. Figure 3 presents findings from one of these 

cases using raster plots (Fig. 3A, B) and poststimulus time histo-

grams (PSTHs) (Fig. 3C, D), during auditory (left) and somato-

sensory (right) stimulation. Auditory and somatosensory stimula-

tion consisted of complex noise and air puff, respectively, and the 

somatosensory receptive field of this neuron was located on the back 

of the contralateral hand. The neuron had an auditory onset latency 

of ~15 msec, and its estimated somatosensory onset latency was 

~12 msec longer. The auditory- and somatosensory-evoked firing 

patterns appear distinctly different; however, this could stem from a 

variety of factors, including differences in the rise time or duration 

of the stimulus, as well as differences in input modality. These recor-

dings were obtained during the control (bilateral tympanic mem-

brane lesion) condition described above with auditory stimulation 

delivered via bone conduction.

Anatomical confirmation of recording sites

To establish the exact locations of our penetrations with respect 

to the boundaries of A1 and CM, small fluorescent marker injec-

tions were placed in A1 and CM at the end of each experiment. 

Figure 4 shows the histological location of these markers in one 

animal with respect to standard staining techniques for defining 

boundaries between cortical areas (Jones et al., 1995; Hackett et 

al., 1998). The areas determined to be posterior auditory associ-

ation cortex on the basis of their electrophysiological properties 

were indeed caudal to A1 in the superior temporal plane.

Discussion

Source of somatosensory input to CM

At present, there are a number of candidate sources for somato-

sensory inputs to CM, and these fall into two main classes. The 

first is a lateral or feedback projection from other cortical areas, 

such as the multisensory regions of the superior temporal sulcus 

(Hackett et al., 1998) and intraparietal sulcus (Lewis and Van 

Essen, 2000). The second is a feedforward projection from “non-

specific” somatosensory pathway structures, such as the supra-

chiasmatic nucleus (Kaas and Hackett, 2000). The former would 

be consistent with the present findings of specific cutaneous re-

cceptive fields on the head and neck, whereas the latter would 

be consistent with our previous finding that the hand input has a 

short-latency, feedforward profile in CM (Schroeder et al., 2001). 

In cats, there is proprioceptive somatosensory input from the 

pinna into the subcortical auditory pathway at the level of the 

dorsal cochlear nucleus (Kanold and Young, 2001), which, if 

present in primates, could fall into the feedforward class of input 

sources. Both feedforward and feedback–lateral input sources are 

considered active possibilities and are under investigation.

Interrelationship of multisensory processing in cortical and 
tectal regions

The apparent bias of the cutaneous representation in CM toward 

the skin surfaces of the head and neck (receptor surfaces not well 
suited for object identification) is consistent with the hypothesis 

that posterior auditory cortex represents the spatial–movement 

analysis or “where” pathway in auditory processing (Rauschecker 

et al., 1997; Rauschecker, 1998; Kaas et al., 1999; Romanski et 

al., 1999), analogous to the parietal pathway in the visual system 

(Ungerleider and Mishkin, 1982). Our previous findings 

(Schroeder et al., 2001) predicted a wider representation of hand 
inputs than we found in the present study. This is likely attri-

butable to methodological differences, such as use of awake versus 
anesthetized subjects, electrical versus cutaneous somatosensory 

stimulation, or current source density versus action potential 

analyses, but resolving this question will require additional study. 

The emerging conceptual model of an auditory spatial system 

parallels that proposed for the visual system, with components 

devoted to both spatial representation (the interconnected 

parieto-temporal–prefrontal cortical regions) and to motoric “ori-

enting” functions (corticovestibular projections). In this model, 

the multisensory functions of the superior colliculus (Stein and
Meredith, 1993) and inferior colliculus (Groh et al., 2001) are proximal to motor output and, thus, are efferent rather than afferent to cortical multisensory processing, as is often assumed. Such a hierarchical arrangement of cortical and tectal functions is supported by the findings that, in cats, multisensory integration in superior colliculus clearly enhances motoric orienting (Stein and Meredith, 1993) and depends on active cortical inputs from rostral lateral suprasylvian and anterior ectosylvian sulcal regions (Jiang et al., 2001). The proposition that cortical and tectal multisensory functions are specifically related in this way makes several clear and testable predictions. The first is that the auditory spatial receptive fields of CM neurons are in register with their cutaneous receptive fields, as is the case for the neurons displaying somato-auditory convergence in the superior colliculus (Meredith and Stein, 1986; Jiang et al., 2001). Another prediction is that the projections from auditory cortices into the tectal system in monkeys (Casseday et al., 1979), like the cortico-tectal projections in cats (Jiang et al., 2001), are necessary for any multisensory integration involving auditory inputs in superior colliculus. In the larger context, it will be important to determine the relationship among the various signal types present in the auditory pathways, including cutaneous inputs from the head–neck and hand (present results), vibratory and arm position signals (present results), eye position signals (Groh et al., 2001; Werner-Reiss et al., 2001; Fu et al., 2002), and pinna position signals (Kanold and Young, 2001). Ongoing studies are directed at these issues, along with more detailed mapping of the cutaneous representation(s) in posterior auditory association cortices.

Multisensory processing and cortical sensory hierarchy

Demonstration of somatosensory responsiveness in auditory association area CM joins with recent reports of low-level cross connections between “modality-specific” cortices (Falchier et al., 2002; Rockland and Ojima, 2003) to directly challenge the view that neocortical multisensory convergence occurs only in higher-order processing regions. CM is positioned at the second level of the cortical auditory hierarchy, corresponding to visual area V2 (Felleman and Van Essen, 1991) and to somatosensory area 1 (Garraghty et al., 1990) and has not been classified previously as a multisensory region. On the basis of its pattern of interconnectivity with both higher- and lower-order regions (Hackett et al., 1998), CM does seem to be a relatively low-level cortical processing area. The properties of the neuronal responses in CM are typical of low-level sensory areas in that they are robust and show little evidence of habituation under both awake (Schroeder et al., 2001) and anesthetized (present results) recording conditions. Our findings thus underscore a significant, general observation about neocortical mechanisms of multisensory integration; that is, inputs from different modalities converge at very early stages of cortical sensory processing. Both our developing understanding of multisensory processing and established views of unisensory processing must incorporate this observation.

References

Giard M, Peron F (1999) Auditory-visual integration during multimodal


