nature neuroscience

Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex

Amir Amedi[1,2], William M Stern[1], Joan A Camprodon[1], Felix Bermpohl[1,3], Lotfi Merabet[1], Stephen Rotman[1], Christopher Hemond[1], Peter Meijer[4] & Alvaro Pascual-Leone[1]

The lateral-occipital tactile-visual area (LOtv) is activated when objects are recognized by vision or touch. We report here that the LOtv is also activated in sighted and blind humans who recognize objects by extracting shape information from visual-to-auditory sensory substitution soundscapes. Recognizing objects by their typical sounds or learning to associate specific soundscapes with specific objects do not activate this region. This suggests that LOtv is driven by the presence of shape information.

Recognition of an object can involve a wide range of cues; for example, a characteristic color, a unique texture or a typical sound. However, shape is a particularly fundamental feature, perhaps because of its evolutionary significance in identifying objects that might serve as tools. Humans obtain most of their information about objects from vision, for which shape is the most salient attribute. In the absence of sight, the world of the blind can appear 'strangely devoid of objects', as John Hull eloquently puts it in his autobiography[1]. Although touch can provide information regarding an object's shape, it may require a slow serial analysis and is limited to the space within the subject's reach. Auditory cues can help to identify objects, but typically provide little shape information, as with the sound of a passing car. Information about objects from the different senses is integrated in our brains to create a coherent and unified perceptual experience, which provides several behavioral advantages, such as speeded responses and improved detection and recognition[2,3]. However, our understanding of the neural substrate for multisensory integration remains limited. Visual and tactile object-related information converges in LOtv, a subregion of the human lateral occipital complex[4]. This region is robustly activated by the exploration of objects in both visual and tactile modalities[4]. It is not activated, however, by various control conditions including visual or tactile textures, visual 'scrambled' objects, the motor aspects of an object recognition task, object naming and the typical sounds made by objects, none of which convey meaningful shape information[2–4]. This suggests that LOtv might be engaged in processing objects' shapes regardless of sensory modality. Visual-to-auditory sensory substitution devices[5–7] (SSD) provide an opportunity to test this hypothesis further. Visual images are captured by a camera and then transformed, according to a predetermined algorithm, into soundscapes that preserve shape information. Although previous studies found that LOtv was not activated by auditory stimuli, we hypothesized that soundscapes synthesized to preserve shape information would activate LOtv.

We studied 12 subjects who gave written informed consent: two blind (one congenital) and five sighted experts of a visual-to-auditory SSD called 'The vOICe', and five corresponding sighted control subjects who learned to associate vOICe soundscapes with object names without extracting the shape information from the sounds. The functional basis of this visual-auditory transformation lies in spectrographic sound synthesis from any input image. Time and stereo panning constitute the horizontal axis in the sound representation of an image, tone frequency makes up the vertical axis and loudness corresponds to pixel brightness[7] (see also **Supplementary Fig. 1** online and http://www.seeingwithsound.com/).

The functional magnetic resonance imaging (fMRI) experiment included six conditions using a block design protocol: (i) tactile object recognition (TacObj), (ii) sensorimotor control (SenMot), (iii) vOICe object recognition using SSD mapping (vOICeObj), (iv) vOICe scrambled images control (vOICeScr), (v) auditory object recognition of 'natural' sounds made by objects (AudObj) and (vi) auditory noise control matched for basic auditory characteristics (AudScr). Stimuli were prepared representing eight objects that were recognizable both by their shape (vision transformed to vOICe, or touch) and by the sounds they make. Sighted experts were trained for 40 h and were on average 69.5% correct on multiple choice testing for recognizing novel objects using The vOICe (range 53–86%, s.d. $= 11.8$; as each question offered four possible choices, this was more than twice the amount of success expected by chance). In contrast, sighted control subjects had 6–8 h of training purely on the eight required associations. To test whether these subjects could interpret novel soundscapes, we asked them to complete the same multiple choice testing as the experts. They averaged 30.8% correct (range 20–36%, s.d. $= 6.2$), which was significantly lower than that achieved by the experts ($P = 0.004$, Wilcoxon rank sum test). All subjects achieved 100% accurate recognition in two complete sets of the eight stimuli to be used in the fMRI experiment on 2 d consecutively before scanning.

During the fMRI subjects had to identify and covertly name each object. In half of the runs, object recognition was assessed using a response box (indicating whether each object was an animal or man-made). All subjects performed this task very well ($>90\%$ correct, see **Supplementary Fig. 2** online). A two-way ANOVA over performance showed no significant group effect (vOICe experts and association

[1]Berenson-Allen Center for Noninvasive Brain Stimulation, Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Medical School, 330 Brookline Ave, Boston, Massachusetts 02115, USA. [2]Department of Physiology and Program of Cognitive Science, Hebrew University, Jerusalem 91220, Israel. [3]Department of Psychiatry, Charité–University Medicine Berlin, 10117 Berlin, Germany. [4]NXP Semiconductors, High Tech Campus 31, 5656 AE Eindhoven, The Netherlands. Correspondence should be addressed to A.A. (aamedi@bidmc.harvard.edu) or A.P.-L. (apleone@bidmc.harvard.edu).

Received 8 February; accepted 27 April; published online 21 May 2007; doi:10.1038/nn1912

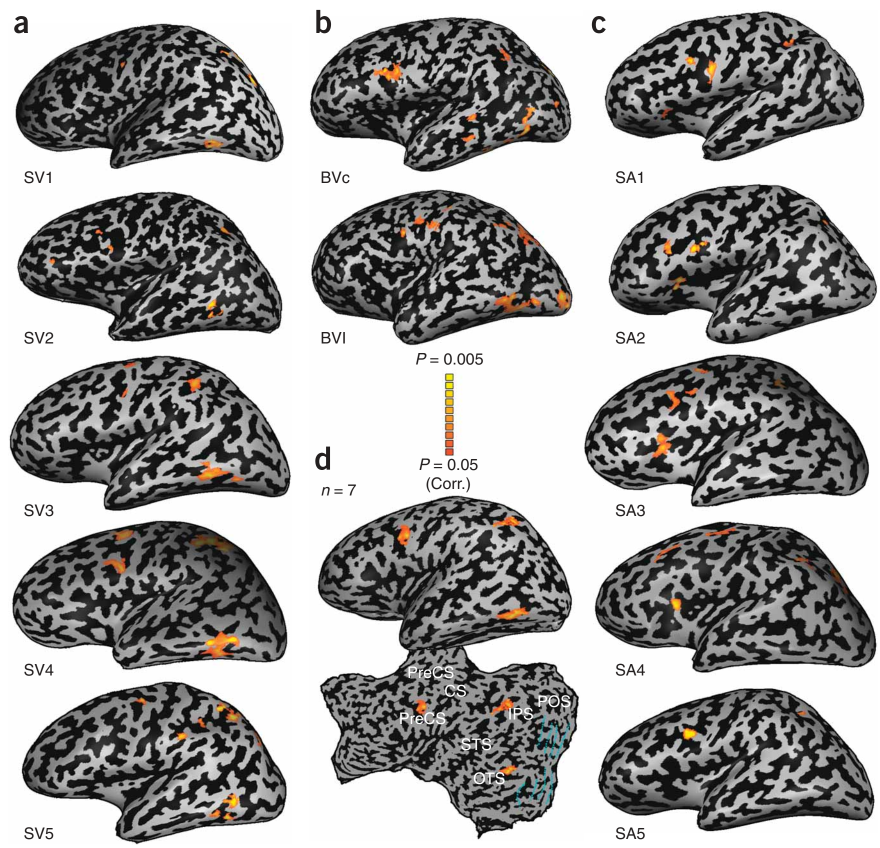

Figure 1 A conjunction analysis for shape across modalities and experimental conditions: vOICeObj > vOICeScr, AudObj, AudScr and SenMot in conjunction with TacObj > vOICeScr, AudObj, AudScr and SenMot. (**a**) Single subject's brain analyses projected on a Talairach normalized inflated cortex reconstruction for five sighted users of The vOICe (SV1–SV5). (**b**) Similar analyses for congenitally blind user (BVc) and late blind user (BVl). All seven expert subjects showed robust activation in left occipito-temporal cortex. Additional common clusters of activation were found in IPS and in prefrontal cortex. (**c**) Similar single subject analysis for the five sighted control subjects (SA1–SA5). Activation was found mainly in IPS and prefrontal cortex, but none of the subjects showed significant activation in occipito-temporal cortex. (**d**) Averages across these seven experts of The vOICe (cyan dotted lines, estimated retinotopic borders).

controls, $F_{1,27} = 0.888$, $P = 0.35$; we did not have behavioral data from one subject) or interaction effect between condition (TacObj, AudObj and vOICeObj) and group ($F_{2,27} = 0.182$, $P = 0.83$).

We used a conjunction analysis to search for brain areas showing selective activation for shape. Specifically, we looked for common areas of activation between object recognition by the soundscapes (vOICeObj) and by touch (TacObj), but not by the typical sounds made by objects (AudObj) or the corresponding sensory controls (vOICeScr, AudScr and SenMot) (**Fig. 1**). Data analysis was performed with BrainVoyager QX1.8 (Brain Innovation) using standard preprocessing procedures and a fixed-effects general linear model. The minimum significance level was set to $P < 0.05$, corrected for multiple comparisons.

All seven vOICe experts (**Fig. 1a**,**b**) showed robust activation in the occipito-temporal cortex. Additional common clusters of activation were found in the parietal cortex (intraparietal sulcus, IPS) and prefrontal cortex (mainly in the pre-central sulcus). None of the five control subjects showed activation in the occipito-temporal cortex for the same conjunction analysis, whereas activation was found in the IPS and prefrontal cortex (**Fig. 1c**). We also carried out a random-effect analysis in the experts, which showed robust activation in the same three brain regions with maximal activation in the occipito-temporal cortex (**Fig. 1d**).

To further characterize the interaction between object recognition in the different modalities we carried out an interaction analysis between all four auditory tasks (**Fig. 2a**). LOtv showed a highly significant activation for this analysis. Although some activation is found in the prefrontal cortex, the activation in the parietal cortex did not reach significance.

To test whether the activation for the vOICe objects was confined to LOtv or extended beyond it, we also analyzed the activation pattern of the soundscape objects (vOICeObj > vOICeScr) separately and overlaid it on the activation for tactile objects (TacObj > SenMot) (**Fig. 2b**). In both blind and sighted experts, vOICeObj > vOICeScr robustly activated the occipito-temporal cortex (**Fig. 2b**, red clusters). Overlaying this with the tactile objects versus sensorimotor control contrast (**Fig. 2b**; blue clusters) showed overlap in LOtv (**Fig. 2b**, purple clusters), with some expansion into areas anterior to LOtv.

Other multisensory areas such as posterior parietal (especially IPS) and prefrontal cortex were activated in the conjunction analysis (**Fig. 1**). However, they also showed activation in control subjects (**Fig. 1**), suggesting that these areas participate in associations of cross-modal information about objects. For all analyses shown, similar patterns were found in both hemispheres. In the right hemisphere, however, the findings were weaker (see **Supplementary Figs. 3** and **4** online).

We also examined the activation pattern in LOtv using the TacObj > SenMot localizer. The averaged percent signal change was calculated for each subject. All vOICe experts showed robust activation during vOICeObj. Sighted subjects showed negligible activation for all other conditions, whereas the blind users also had some activation during AudObj. No activation was evident during vOICeObj in the control subjects (**Fig. 2c**). The LOtv percent signal change for vOICeObj in the sighted experts showed significantly higher activation than in control subjects ($P = 0.004$, Wilcoxon rank sum test). Adding the two blind subjects to the expert group yielded an even higher significance level ($P = 0.0013$). For detailed methods and further analyses see **Supplementary Methods** online.

The lateral occipital-temporal cortex is an important site for multisensory integration of visual, tactile and auditory information[3]. Previous work suggests that LOtv is primarily involved in the integration of visual and tactile information about objects, whereas the superior temporal sulcus integrates visual and auditory information about objects[2–4,8]. Here we report fMRI results showing that in sighted subjects, as well as in a congenitally and a late blind subject, LOtv was activated by object recognition via soundscapes. Notably, soundscapes were only capable of activating LOtv in subjects trained to interpret them and extract their shape information. Subjects taught

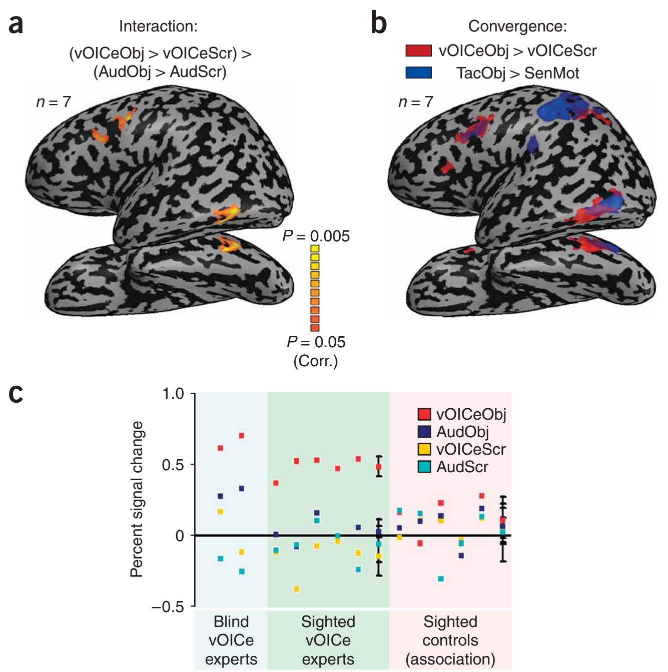

Figure 2 Interaction analysis, convergence analysis and percent signal change analysis of LOtv. (**a**) We conducted an interaction analysis between the tasks labeled as vOICeObj/vOICeScr and AudObj/AudScr based on the first-level single-subject general linear model analysis using random effect analysis. LO shows a highly significant activation. (**b**) The activation pattern of the soundscape objects (vOICeObj > vOICeScr, red clusters) overlaid on the activation for tactile objects (TacObj > SenMot, blue clusters); overlap, purple clusters. (**c**) The magnitude of activation in left LOtv across subjects in the nontactile conditions (using the peak voxel for TacObj > SenMot after convolution with a Gaussian kernel of 4-mm full width at half maximum). Robust activation was found for vOICeObj (red) in all vOICe experts, whereas no such activation was evident in the five sighted controls (percent signal change in each subject and average percent signal change ± s.d. for all sighted vOICe experts and controls).

simple associations between soundscapes and objects did not show this activation. This strongly supports our hypothesis that LOtv is involved specifically in analyzing the shape of objects, regardless of sensory input modality.

Previous neuroimaging studies in the sighted lend support to our hypothesis[2–4,8]. For example, haptic exploration of unfamiliar objects with novel abstract shapes has been shown to activate LOtv. This supports the notion that ecological validity and familiarity (and thus higher chances of invoking visual imagery) with a given tactile object is not required for recruitment of LOtv[2–4]. Furthermore, lesions encompassing the occipito-temporal cortex can result in a severe bimodal visual and tactile agnosia[2,3]. 'Virtual lesions' targeting the right dorsal extrastriate visual cortex have recently been shown to impair object recognition using a different visual-to-auditory SSD in blind, but not in sighted, subjects[8]. Notably, our two blind subjects also showed activation in the posterior occipital cortex corresponding to retinotopic 'visual' areas (**Fig. 1**), as well as hints for additional plasticity in LOtv for auditory objects (**Fig. 2c**), both perhaps due to cross-modal plasticity associated with prolonged sight deprivation[8–10].

Could activation caused by soundscapes be triggered by visual imagery? This argument would require that: (i) the activation of LOtv in the congenitally blind expert user and in the sighted trained subjects, though identical, be caused by different mechanisms, (ii) that none of the control subjects invoked visual imagery, whereas all of the experts did, and (iii) that all subjects used imagery for soundscapes, but none for the auditory objects. Although visual imagery is a difficult issue to rule out completely, we find it an unlikely explanation, particularly because the pattern of activation differs from that seen in prior visual imagery studies[2–4].

What can explain the recruitment of LOtv to sounds (and tactile input) containing shape information? Notably, similar cross-modal effects have been found in sighted and blind individuals in Brodmann's area BA37 in response to verbal word processing[11], in visual association areas in response to depth perception using visual-to-auditory SSD in the sighted[12] and in the functional relevance of visual cortex to tactile grating orientation[13] and tactile motion[14]. Such instances support a metamodal organization of the brain where specific brain regions are conceptualized as operators that process certain types of data regardless of their input modality[15]. Inputs may shift dynamically depending on availability and the integrative capacity of a region in a functional distributed neural network to optimize behavioral outcome[10]. Our results provide new information suggesting that LOtv acts as a metamodal operator for shape.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS
We thank M. Thivierge for invaluable administrative help and S. Gautam for statistics support. The work on this article was supported by grants from the US National Institutes of Health (K24-RR018875, R21-EY0116168 and RO1-EY12091 to A.P.-L.) and The International Human Frontier Science Program Organization (to A.A).

AUTHOR CONTRIBUTIONS
All authors were involved in the design and implementation of the study and the writing of the manuscript. A.P.-L. and A.A. devised the concept, A.A. carried out the analysis, W.M.S. designed and implemented the vOICe training program and P.M. provided expert technical support regarding The vOICe.

COMPETING INTERESTS STATEMENT
The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/natureneuroscience/.

Published online at http://www.nature.com/natureneuroscience
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Hull, J.M. *Touching the Rock: an Experience of Blindness* (Vintage Books, New York, 1992).
2. Amedi, A., von Kriegstein, K., van Atteveldt, N.M., Beauchamp, M.S. & Naumer, M.J. *Exp. Brain Res.* **166**, 559–571 (2005).
3. Beauchamp, M.S. *Curr. Opin. Neurobiol.* **15**, 145–153 (2005).
4. Amedi, A., Malach, R., Hendler, T., Peled, S. & Zohary, E. *Nat. Neurosci.* **4**, 324–330 (2001).
5. Bach-y-Rita, P. & Kersel, S.W. *Trends Cogn. Sci.* **7**, 541–546 (2003).
6. Rauschecker, J.P. *Trends Neurosci.* **18**, 36–43 (1995).
7. Meijer, P.B. *IEEE Trans. Biomed. Eng.* **39**, 112–121 (1992).
8. Sathian, K. *Dev. Psychobiol.* **46**, 279–286 (2005).
9. Collignon, O., Lassonde, M., Lepore, F., Bastien, D. & Veraart, C. *Cereb. Cortex* **17**, 457–465 (2007).
10. Pascual-Leone, A., Amedi, A., Fregni, F. & Merabet, L.B. *Annu. Rev. Neurosci.* **28**, 377–401 (2005).
11. Buchel, C., Price, C. & Friston, K. *Nature* **394**, 274–2777 (1998).
12. Renier, L. *et al. Neuroimage* **26**, 573–580 (2005).
13. Zangaladze, A., Epstein, C.M., Grafton, S.T. & Sathian, K. *Nature* **401**, 587–590 (1999).
14. Ricciardi, E. *et al. Cereb. Cortex* published online 19 March 2007 (doi:10.1093/cercor/bhm018).
15. Pascual-Leone, A. & Hamilton, R. in *Vision: From Neurons to Cognition, Progr. Brain Res* (eds. Casanova, C. & Ptito, M.) Ch. 27 (Elsevier, New York, 2001).