Spatial pattern of BOLD fMRI activation reveals cross-modal information in auditory cortex

P.-J. Hsieh, J. T. Colas and N. Kanwisher

J Neurophysiol 107:3428-3432, 2012. First published 18 April 2012; doi:10.1152/jn.01094.2010

J Neurophysiol 107: 3428–3432, 2012.
First published April 18, 2012; doi:10.1152/jn.01094.2010.

Spatial pattern of BOLD fMRI activation reveals cross-modal information in auditory cortex

P.-J. Hsieh,[1] **J. T. Colas,**[2] **and N. Kanwisher**[2]

[1]*Neuroscience and Behavioral Disorders Program, Duke-NUS Graduate Medical School, Singapore; and* [2]*Department of Brain and Cognitive Sciences, McGovern Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts*

Submitted 15 December 2010; accepted in final form 19 March 2012

Hsieh PJ, Colas JT, Kanwisher N. Spatial pattern of BOLD fMRI activation reveals cross-modal information in auditory cortex. *J Neurophysiol* 107: 3428–3432, 2012. First published April 18, 2012; doi:10.1152/jn.01094.2010.—Recent findings suggest that neural representations in early auditory cortex reflect not only the physical properties of a stimulus, but also high-level, top-down, and even cross-modal information. However, the nature of cross-modal information in auditory cortex remains poorly understood. Here, we used pattern analyses of fMRI data to ask whether early auditory cortex contains information about the visual environment. Our data show that *1*) early auditory cortex contained information about a visual stimulus when there was no bottom-up auditory signal, and that *2*) no influence of visual stimulation was observed in auditory cortex when visual stimuli did not provide a context relevant to audition. Our findings attest to the capacity of auditory cortex to reflect high-level, top-down, and cross-modal information and indicate that the spatial patterns of activation in auditory cortex reflect contextual/implied auditory information but not visual information per se.

hierarchical inference; feedback; top-down modulation

RECENT EVIDENCE HAS SHOWN that early sensory cortex encodes not only low-level sensory properties of a stimulus, but also mid- to high-level perceptual information. For example, early visual cortex has been implicated in figure-ground discrimination (Heinen et al. 2005; Huang and Paradiso 2008; Hupé et al. 1998), shape and size perception (Fang et al. 2008; Murray et al. 2002, 2006), lightness constancy (Boyaci et al. 2007), attentional modulation (Datta and DeYoe 2009; Fischer and Whitney 2009; Ress et al. 2000), tracking stimulus reward history (Serences 2008), conscious perception (Hsieh et al. 2006; Hsieh and Tse 2009, 2010a,b), and even identification of a stimulus (Hsieh et al. 2010). Similarly, responses in early auditory cortex are modulated by attention (Alho et al. 1999; Grady et al. 1997; Hillyard et al. 1973; Jäncke et al. 1999; Lipschutz et al. 2002; Näätänen 1990; O'Leary et al. 1997; Tzourio et al. 1997; Woldorff et al. 1993; Woldorff and Hillyard 1991; Woodruff et al. 1996) and reflect cross-modal processing (Calvert et al. 1997; Calvert and Campbell 2003; Foxe and Schroeder 2005; Ghazanfar and Schroeder 2006; Meyer et al. 2010; Schroeder and Foxe 2005). For instance, responses in auditory cortex can be affected by visual or somatosensory stimuli that accompany auditory stimulation in both humans (Bernstein et al. 2004; Besle et al. 2004, 2008; Calvert et al. 1997; Giard and Peronnet 1999; Foxe et al. 2000, 2002; Gobbelé et al. 2003; Lehmann et al. 2006; Lütkenhöner et al. 2002; Martuzzi et al. 2007; Molholm et al. 2002; Murray et al. 2005; Pekkola et al. 2005; van Atteveldt et al. 2004; van Wassenhove et al. 2005) and monkeys (Brosch et al. 2005; Fu et al. 2003; Ghazanfar et al. 2005; Kayser et al. 2007, 2008, 2009, 2010; Lakatos et al. 2007; Schroeder et al. 2001; Schroeder and Foxe 2002; Schwartz et al. 2004).

Despite this ample evidence for cross-modal influences on responses in auditory cortex, the sources of these influences remain poorly understood. Here, we investigated the nature of one such case of cross-modal modulation. A recent study by Meyer et al. (2010) presented visual stimuli silently and showed that activity in auditory cortex differentiated among various sound-implying animals, musical instruments, and objects. However, this study was limited by its inability to determine whether the pattern information in auditory cortex actually reflects implied sound information or merely visual information. Here, we tested whether the representation of a stimulus in auditory cortex reflects contextual/implied auditory information or visual information per se.

Our experiment was conducted with functional MRI (fMRI) and included three conditions: silence, action control, and sound. In the silence condition, we presented different visual stimuli that implied similar knocking sounds. Our prediction was that the spatial pattern of blood oxygen level-dependent (BOLD) activation in auditory cortex would represent these visual stimuli differently. To distinguish further whether this pattern information reflects contextual/implied auditory information or visual information per se, we presented actions that do not imply sounds in the action control condition. If the pattern information observed in the silence condition reflects implied auditory information, we should expect the pattern information to be absent for the action control condition. However, if the pattern information observed in the silence condition reflects purely visual information, the pattern analysis should reveal information for the action control condition as well. Moreover, to determine whether this top-down signal is robust enough to persist with the addition of a salient bottom-up signal, we paired the visual stimuli from the silence condition with identical sounds in the sound condition.

METHODS

Participants. Ten volunteers between 18 and 30 yr old participated in the study. All of them were healthy and right-handed and had normal or corrected-to-normal visual and auditory acuity. All subjects gave informed written consent within a protocol passed by the Duke-NUS Graduate Medical School or the Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects and were compensated with 60 dollars for their participation.

Experimental procedures. Scanning was performed at the McGovern Institute at the Massachusetts Institute of Technology in

---

Address for reprint requests and other correspondence: P.-J. Hsieh, Program in Neuroscience and Behavioral Disorders, Duke-NUS Graduate Medical School, 8 College Rd., Singapore 169857 (e-mail: hsieh.pj@gmail.com).

Fig. 1. Stimuli. Subjects viewed blocks of 6 2-s color videos that corresponded to the 3 experimental conditions with 2 sequences of visual stimuli each (woodpecker agent and hand agent). Videos from the 1st condition ("sound") visually depicted a sequence of 3 knocking actions that were accompanied by a sequence of generic knocking sounds in lieu of the original audio tracks, such that auditory stimulation was identical across visual stimuli. Videos from the 2nd condition ("silence") included identical visual stimuli without any auditory stimulation. Videos from the 3rd condition ("action control") depicted the same 2 agents in motion without performing any knocking actions and also without any auditory stimulation.

Cambridge, MA, with the Athinoula A. Martinos Imaging Center 3T Siemens Trio scanner. fMRI runs were acquired using a gradient-echo, echo-planar sequence [repetition time (TR) = 3 s, echo time (TE) = 30 ms, $2 \times 2 \times 2$ mm + 20% spacing]. Forty-six slices were collected with a 32-channel head coil. Slices were oriented roughly perpendicular to the calcarine sulcus and covered the whole brain.

Subjects viewed blocks of six 2-s color videos that corresponded to the three experimental conditions with two sequences of visual stimuli each (Fig. 1 and supplementary material available online at the *Journal of Neurophysiology* web site). The agent in the video, which was either a woodpecker or a human hand, distinguished the visual stimuli. Videos from the first condition (sound) visually depicted a sequence of three knocking actions that were accompanied by a sequence of generic knocking sounds in lieu of the original audio tracks, such that auditory stimulation was identical across visual stimuli. The intensity of the sound was subjectively selected by each subject at the beginning of the experiment to be as loud as possible without causing discomfort. Videos from the second condition (silence) included the visual stimuli from the sound condition without any auditory stimulation. Videos from the third condition (action control) depicted the same two agents in motion without performing any knocking actions and also without any auditory stimulation.

Appearing at the beginning of a 3-s TR, the 2-s videos had a resolution of 960 × 540 pixels and subtended $12.25 \times 6.75°$ of visual angle against a black background. A white fixation cross was always present in the center of the display and subtended $0.75 \times 0.75°$. Seven repetitions of a given condition-stimulus pair occurred within each of the 12 21-s blocks that were interleaved among 15-s fixation periods within each run (Fig. 2). Six seconds of fixation at the beginning and end of a run were added to yield a total duration of 429 s for each run. For each video presentation, the overall luminance of the video was randomly selected to be 25% greater or lesser than its original value. Subjects were required to press 1 of 2 buttons on a response box (2-alternative forced choice) to indicate whether each individual 2-s video appeared atypically bright or dim, and every subject's performance was near ceiling. The presentation order of the 6 condition-stimulus pairs was randomized within each run. While being scanned, all subjects completed between 9 and 11 runs.

ROI identification. Functional localization of the region of interest (ROI) was based on three runs of a separate auditory localizer. It consisted of 30-s blocks in which repeated 100-ms pure tones were presented with an interstimulus interval of 400 ms while subjects fixated. Conditions were defined by the three frequency ranges of the blocks, namely low (340–870 Hz), middle (880–2,170 Hz), and high (2,370–5,900 Hz). The frequencies of tones within each block increased linearly with modulo 5.2 s, corresponding to slopes of 102, 248, and 679 Hz/s, respectively. For example, the first tone in a low-frequency block was always 340 Hz, and the frequency increased by 51 Hz for the next tone 500 ms later. If, at the end of a 5.2-s loop, the frequency exceeded 870 Hz by x, it would be reset to $340 + x$ Hz and looped again throughout a 30-s block. Since we did not find a significant difference between the BOLD responses to any of the three conditions, auditory cortex was defined bilaterally as the contiguous regions in the vicinity of the superior and transverse (Heschl's) temporal gyri that responded significantly more robustly to auditory stimulation, combined across the three conditions, than to background noise alone (e.g., from the scanner; $P < 10^{-12}$). ROI sizes ranged from 177 to 432 voxels with a mean of 331.4 voxels and a standard deviation of 70.7.

Data analysis. Data analysis was conducted using the fMRI software package FreeSurfer (http://surfer.nmr.mgh.harvard.edu) and MATLAB (The MathWorks). The preprocessing steps for both the localizer and experimental runs included motion correction and intensity normalization. Preprocessing for the localizer runs also included spatial smoothing using a Gaussian kernel with a full-width at half-maximum of 6 mm. A gamma function with $\delta = 2.25$, $\tau = 1.25$, and $\alpha = 2$ was used to estimate the hemodynamic response for each condition in both the experiment and the localizer.

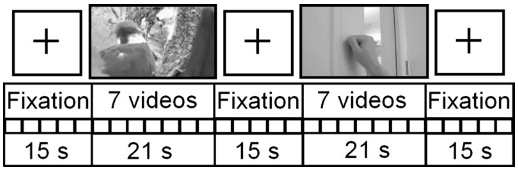

Fig. 2. Experimental design. Appearing at the beginning of a 3-s scanning repetition (TR), the 2-s videos were presented in front of a black background and behind a centered white fixation cross that was always present. Seven repetitions of a given condition-stimulus pair occurred within 21-s blocks that were interleaved among 15-s fixation periods within each run. The presentation order of the 6 condition-stimulus pairs was randomized within each run.

In addition to a standard univariate analysis of the mean BOLD response to each condition, a correlational analysis was performed on the β-weights between visual stimuli for each condition in each voxel with a standard multivariate pattern analysis method (Haxby et al. 2001). Data were split into odd and even runs, and spatial patterns of response were extracted from each subset of data for the six condition-stimulus pairs. The patterns were first normalized, such that the mean response in each voxel across the visual stimuli to be compared was subtracted from the response to each stimulus for each half of the data before the correlation values were calculated. Within each ROI, we then computed the split-half correlations as Pearson correlation coefficients between the normalized activity patterns in response to the two sequences of visual stimuli for each of the three experimental conditions, that is, sound, silence, and action control. These correlations were computed for each subject and then averaged across subjects by condition.

RESULTS

The results of the univariate analysis (Fig. 3) revealed increased mean BOLD activation in auditory cortex for the sound condition relative to baseline ($P < 10^{-4}$) and to the silence and action control conditions ($P < 10^{-3}$). However, no differential mean activation was observed between the silence and action control conditions ($P > 0.05$).

The results of the multivariate pattern analysis (Fig. 4) indicated that the spatial patterns of BOLD activation in auditory cortex contained information that distinguished between the visual stimuli (hand and woodpecker) for the silence condition ($P = 0.015$). However, this was not the case for the sound ($P > 0.05$) and action control conditions ($P > 0.05$). Moreover, decoding accuracy for which visual stimulus was presented in the silence condition was significantly greater than decoding for the sound condition ($P = 0.005$) and the action control condition ($P = 0.039$). These findings demonstrated that *1*) top-down information may dominate processing in auditory cortex only when bottom-up auditory input is absent, and that *2*) there may be no such top-down information in auditory cortex when the visual stimuli do not imply sound. As such, these findings suggest that spatial patterns of BOLD activation in auditory cortex do not reflect purely visual information per se, but rather visually induced top-down information.

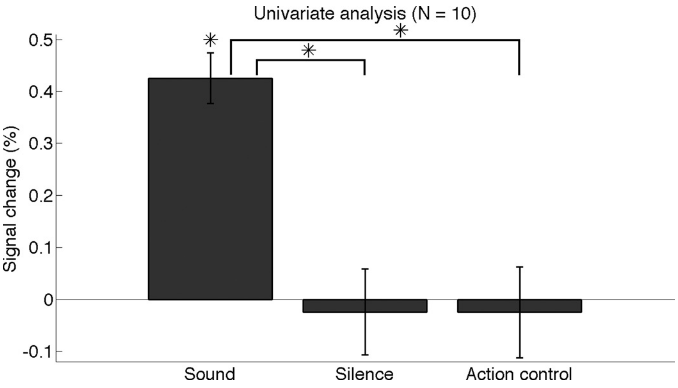

Fig. 3. Results of univariate analysis. Mean blood oxygen level-dependent (BOLD) activation in auditory cortex was significantly greater for the sound condition relative to baseline ($P < 10^{-4}$) and to the silence and action control conditions ($P < 10^{-3}$). However, no differential mean activation was observed between the silence and action control conditions ($P > 0.05$). Error bars indicate standard errors of the means across subjects, and asterisks indicate significance ($P < 0.05$).

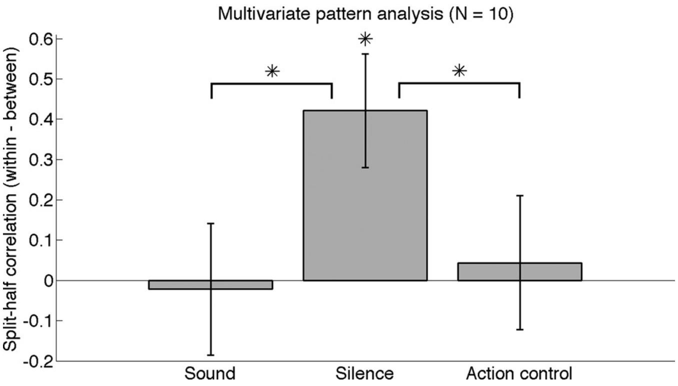

Fig. 4. Results of multivariate pattern analysis. The spatial patterns of BOLD activation in auditory cortex contained information that distinguished between the visual stimuli for the silence condition ($P = 0.015$). However, this was not the case for the sound and action control conditions ($P > 0.05$). Moreover, decoding accuracy for which visual stimulus was presented in the silence condition was significantly greater than decoding for the sound condition ($P = 0.005$) and the action control condition ($P = 0.039$). Error bars indicate standard errors of the means across subjects, and asterisks indicate significance ($P < 0.05$).

DISCUSSION

Our results show that the spatial patterns of BOLD activation in auditory cortex reflected a cross-modal influence of visual information when bottom-up auditory input was absent. In contrast, no top-down information was found in auditory cortex when the visual stimuli did not imply sound (i.e., in the action control condition). These findings suggest that *1*) the activity being captured in the pattern of responses across auditory cortex reflects high-level, top-down, and cross-modal information, and that *2*) stimuli from one modality (e.g., vision) that do not provide a context relevant to another modality (e.g., audition) will not exert a cross-modal top-down influence. Note, however, that we cannot completely rule out the possibility of such top-down effects in the action control condition insofar as top-down activity might simply be weaker, less spatially consistent (i.e., requiring greater spatial resolution), or otherwise structured in a way that cannot be detected with fMRI and multivariate pattern analysis.

Our data did not reveal any visually induced pattern information for the sound condition when the bottom-up auditory signal, which was identical across visual stimuli, was present. One possible explanation is that modulatory top-down information is overshadowed by the more robust bottom-up auditory information. An alternative account is that top-down modulatory signals might be assigned less weight or even disappear altogether when bottom-up signals are strong and unambiguous. Increased mean BOLD activation was observed for the sound condition, however, indicating that the information captured with multivariate analysis, which is encoded in the spatial pattern of activation, is distinct from that captured with univariate analysis, which is encoded in the overall activation across voxels.

The difference in spatial patterns observed in auditory cortex between visual stimuli in the silence condition is likely due to some combination of contextual information and implicit auditory imagery. For example, previous findings have shown that auditory imagery activates secondary auditory cortex (Bunzeck et al. 2005; Halpern et al. 2004; Yoo et al. 2001; Zatorre and Halpern 2005) and that sound-implying visual stimuli can be decoded on the basis of activity in auditory cortex (Meyer et al. 2010). However, it is worth noting that our ROIs were mainly within Heschl's gyrus and primary auditory cortex. This possible discrepancy with work identifying the neural correlates of auditory imagery in secondary auditory cortex is likely due to some combination of the following. First, primary auditory cortex may actually be involved in imagery, but the information contained within this region could only be detectable with pattern analysis. Second, the patterns of responses to the silence condition may not be driven by imagery at all, but rather by visually induced (audition-relevant) contextual information. Little activity in secondary auditory cortex is to be expected in either case, so determining the validity of these possibilities and dissociating their contributions to our main results will require further research.

To conclude, our data show that *1*) auditory cortex contained information about visual stimuli when bottom-up auditory input was absent, and that *2*) no contextual top-down information was observed in auditory cortex when the visual stimuli did not provide a context relevant to audition. Our findings attest to the capacity of early auditory cortex to be affected by high-level, top-down, and cross-modal information and indicate that the spatial patterns of activity in auditory cortex reflect contextual/implied auditory information but not visual information per se.

ACKNOWLEDGMENTS

We thank Mark Tomkins for providing the woodpecker footage.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

P.-J.H., J.T.C., and N.K. conception and design of research; P.-J.H. and J.T.C. performed experiments; P.-J.H., J.T.C., and N.K. interpreted results of experiments; P.-J.H. and J.T.C. prepared figures; P.-J.H., J.T.C., and N.K. drafted manuscript; P.-J.H., J.T.C., and N.K. edited and revised manuscript; P.-J.H., J.T.C., and N.K. approved final version of manuscript; J.T.C. analyzed data.

REFERENCES

Alho K, Medvedev SV, Pakhomov SV, Roudas MS, Tervaniemi M, Reinikainen K, Zeffiro T, Näätänen R. Selective tuning of the left and right auditory cortices during spatially directed attention. *Brain Res Cogn Brain Res* 7: 335–341, 1999.

Bernstein LE, Auer ET, Takayanagi S. Auditory speech detection in noise enhanced by lipreading. *Speech Commun* 44: 5–18, 2004.

Besle J, Fischer C, Bidet-Caulet A, Lecaignard F, Bertrand O, Giard MH. Visual activation and audiovisual interactions in the auditory cortex during speech perception: intracranial recordings in humans. *J Neurosci* 28: 14301–14310, 2008.

Besle J, Fort A, Delpuech C, Giard MH. Bimodal speech: early suppressive visual effects in human auditory cortex. *Eur J Neurosci* 20: 2225–2234, 2004.

Boyaci H, Fang F, Murray SO, Kersen D. Response to lightness variations in early human visual cortex. *Curr Biol* 17: 989–993, 2007.

Brosch M, Selezneva E, Scheich H. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J Neurosci* 25: 6797–6806, 2005.

Bunzeck N, Wuestenberg T, Lutz K, Heinze HJ, Jancke L. Scanning silence: mental imagery of complex sounds. *Neuroimage* 26: 1119–1127, 2005.

Calvert GA, Bullmore ET, Brammer MJ, Campbell R, Williams SC, McGuire PK, Woodruff PW, Iversen SD, David AS. Activation of auditory cortex during silent lipreading. *Science* 276: 593–596, 1997.

Calvert GA, Campbell R. Reading speech from still and moving faces: the neural substrates of visible speech. *J Cogn Neurosci* 15: 57–70, 2003.

Datta R, DeYoe EA. I know where you are secretly attending! The topography of human visual attention revealed with fMRI. *Vision Res* 49: 1037–1044, 2009.

Fang F, Kersten D, Murray SO. Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *J Vis* 8: 2.1–9, 2008.

Fischer J, Whitney D. Attention narrows position tuning of population responses in V1. *Curr Biol* 19: 1356–1361, 2009.

Foxe JJ, Morocz IA, Murray MM, Higgins BA, Javitt DC, 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.

Foxe JJ, Schroeder CE. The case for feedforward multisensory convergence during early cortical processing. *Neuroreport* 16: 419–423, 2005.

Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM. Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J Neurophysiol* 88: 540–543, 2002.

Fu KM, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, Garraghty PE, Schroeder CE. Auditory cortical neurons respond to somatosensory stimulation. *J Neurosci* 23: 7510–7515, 2003.

Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci* 25: 5004–5012, 2005.

Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? *Trends Cogn Sci* 10: 278–285, 2006.

Giard MH, Peronnet F. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci* 11: 473–490, 1999.

Gobbelé R, Schürmann M, Forss N, Juottonen K, Buchner H, Hari R. Activation of the human posterior parietal and temporoparietal cortices during audiotactile interaction. *Neuroimage* 20: 503–511, 2003.

Grady CL, Van Meter JW, Maisog JM, Pietrini P, Krasuski J, Rauschecker JP. Attention-related modulation of activity in primary and secondary auditory cortex. *Neuroreport* 8: 2511–2516, 1997.

Halpern AR, Zatorre RJ, Bouffard M, Johnson JA. Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia* 42: 1281–1292, 2004.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293: 2425–2430, 2001.

Heinen K, Jolij J, Lamme VA. Figure-ground segregation requires two distinct periods of activity in V1: a transcranial magnetic stimulation study. *Neuroreport* 16: 1483–1487, 2005.

Hillyard SA, Hink RF, Schwent VL, Picton TW. Electrical signs of selective attention in the human brain. *Science* 182: 177–180, 1973.

Hsieh PJ, Caplovitz GP, Tse PU. Bistable illusory rebound motion: event-related functional magnetic resonance imaging of perceptual states and switches. *Neuroimage* 32: 728–739, 2006.

Hsieh PJ, Tse PU. BOLD signal in both ipsilateral and contralateral retinotopic cortex modulates with perceptual fading. *PLoS One* 5: e9638, 2010a.

Hsieh PJ, Tse PU. "Brain-reading" of perceived colors reveals a feature mixing mechanism underlying perceptual filling-in in cortical area V1. *Hum Brain Mapp* 31: 1395–1407, 2010b.

Hsieh PJ, Tse PU. Microsaccade rate varies with subjective visibility during motion-induced blindness. *PLoS One* 4: e5163, 2009.

Hsieh PJ, Vul E, Kanwisher N. Recognition alters the spatial pattern of fMRI activation in early retinotopic cortex. *J Neurophysiol* 103: 1501–1507, 2010.

Huang X, Paradiso MA. V1 response timing and surface filling-in. *J Neurophysiol* 100: 539–547, 2008.

Hupé JM, James AC, Payne BR, Lomber SG, Girard P, Bullier J. Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* 394: 784–787, 1998.

Jäncke L, Mirzazade S, Shah NJ. Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neurosci Lett* 266: 125–128, 1999.

Kayser C, Logothetis NK, Panzeri S. Visual enhancement of the information representation in auditory cortex. *Curr Biol* 20: 19–24, 2010.

Kayser C, Petkov CI, Augath M, Logothetis NK. Functional imaging reveals visual modulation of specific fields in auditory cortex. *J Neurosci* 27: 1824–1835, 2007.

Kayser C, Petkov CI, Augath M, Logothetis NK. Integration of touch and sound in auditory cortex. *Neuron* 48: 373–384, 2005.

Kayser C, Petkov CI, Logothetis NK. Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. *Hear Res* 258: 80–88, 2009.

Kayser C, Petkov CI, Logothetis NK. Visual modulation of neurons in auditory cortex. *Cereb Cortex* 18: 1560–1574, 2008.

Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53: 279–292, 2007.

Lehmann C, Herdener M, Esposito F, Hubl D, di Salle F, Scheffler K, Bach DR, Federspiel A, Kretz R, Dierks T, Seifritz E. Differential patterns of multisensory interactions in core and belt areas of human auditory cortex. *Neuroimage* 31: 294–300, 2006.

Lipschutz B, Kolinsky R, Damhaut P, Wikler D, Goldman S. Attention-dependent changes of activation and connectivity in dichotic listening. *Neuroimage* 17: 643–656, 2002.

Lütkenhöner B, Lammertmann C, Simões C, Hari R. Magnetoencephalographic correlates of audiotactile interaction. *Neuroimage* 15: 509–522, 2002.

Martuzzi R, Murray MM, Michel CM, Thiran JP, Maeder PP, Clarke S, Meuli RA. Multisensory interactions within human primary cortices revealed by B.O.L.D. dynamics. *Cereb Cortex* 17: 1672–1679, 2007.

Meyer K, Kaplan JT, Essex R, Webber C, Damasio H, Damasio A. Predicting visual stimuli on the basis of activity in auditory cortices. *Nat Neurosci* 13: 667–668, 2010.

Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ. Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res Cogn Brain Res* 14: 115–128, 2002.

Murray MM, Molholm S, Michel CM, Heslenfeld DJ, Ritter W, Javitt DC, Schroeder CE, Foxe JJ. Grabbing your ear: rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb Cortex* 15: 963–974, 2005.

Murray SO, Boyaci H, Kersten D. The representation of perceived angular size in human primary visual cortex. *Nat Neurosci* 9: 429–434, 2006.

Murray SO, Kersten D, Olshausen BA, Schrater P, Woods DL. Shape perception reduces activity in human primary visual cortex. *Proc Natl Acad Sci USA* 99: 15164–15169, 2002.

Näätänen R. The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behav Brain Sci* 13: 201–288, 1990.

O'Leary DS, Andreasen NC, Hurtig RR, Torres IJ, Flashman LA, Kesler ML, Arndt SV, Cizadlo TJ, Ponto LL, Watkins GL, Hichwa RD. Auditory and visual attention assessed with PET. *Hum Brain Mapp* 5: 422–436, 1997.

Pekkola J, Ojanen V, Autti T, Jääskeläinen IP, Möttönen R, Tarkiainen A, Sams M. Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *Neuroreport* 16: 125–128, 2005.

Petkov CI, Kayser C, Augath M, Logothetis NK. Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biol* 4: e215, 2006.

Ress D, Backus BT, Heeger DJ. Activity in primary visual cortex predicts performance in a visual detection task. *Nat Neurosci* 3: 940–945, 2000.

Schroeder CE, Foxe J. Multisensory contributions to low-level, 'unisensory' processing. *Curr Opin Neurobiol* 15: 454–458, 2005.

Schroeder CE, Foxe JJ. The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res* 14: 187–198, 2002.

Schroeder CE, Lindsley RW, Specht C, Marcovici A, Smiley JF, Javitt DC. Somatosensory input to auditory association cortex in the macaque monkey. *J Neurophysiol* 85: 1322–1327, 2001.

Schwartz JL, Berthommier F, Savariaux C. Seeing to hear better: evidence for early audio-visual interactions in speech identification. *Cognition* 93: B69–B78, 2004.

Serences JT. Value-based modulations in human visual cortex. *Neuron* 60: 1169–1181, 2008.

Tzourio N, El Massioui F, Crivello F, Joliot M, Renault B, Mazoyer B. Functional anatomy of human auditory attention studied with PET. *Neuroimage* 5: 63–77, 1997.

van Atteveldt N, Formisano E, Goebel R, Blomert L. Integration of letters and speech sounds in the human brain. *Neuron* 43: 271–282, 2004.

van Wassenhove V, Grant KW, Poeppel D. Visual speech speeds up the neural processing of auditory speech. *Proc Natl Acad Sci USA* 102: 1181–1186, 2005.

Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE. Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc Natl Acad Sci USA* 90: 8722–8726, 1993.

Woldorff MG, Hillyard SA. Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalogr Clin Neurophysiol* 79: 170–191, 1991.

Woodruff PW, Benson RR, Bandettini PA, Kwong KK, Howard RJ, Talavage T, Belliveau J, Rosen BR. Modulation of auditory and visual cortex by selective attention is modality-dependent. *Neuroreport* 7: 1909–1913, 1996.

Yoo SS, Lee CU, Choi BG. Human brain mapping of auditory imagery: event-related functional MRI study. *Neuroreport* 12: 3045–3049, 2001.

Zatorre RJ, Halpern AR. Mental concerts: musical imagery and the auditory cortex. *Neuron* 47: 9–12, 2005.