

# Visual Enhancement of the Information Representation in Auditory Cortex

Christoph Kayser,<sup>1,\*</sup> Nikos K. Logothetis,<sup>1,2</sup> and Stefano Panzeri<sup>3</sup>

<sup>1</sup>Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, 72076 Tübingen, Germany

<sup>2</sup>Division of Imaging Science and Biomedical Engineering, University of Manchester, Manchester M13 9PT, United Kingdom

<sup>3</sup>Robotics, Brain and Cognitive Sciences Department, Italian Institute of Technology, Via Morego 30, 16163 Genova, Italy

## Summary

Combining information across different sensory modalities can greatly facilitate our ability to detect, discriminate, or recognize sensory stimuli [1, 2]. Although this process of sensory integration has usually been attributed to classical association cortices, recent work has demonstrated that neuronal activity in early sensory cortices can also be influenced by cross-modal inputs [3–5]. Here we demonstrate that such “early” multisensory influences enhance the information carried by neurons about multisensory stimuli. By recording in auditory cortex of alert monkeys watching naturalistic audiovisual stimuli, we quantified the effect of visual influences on the trial-to-trial response variability and on the amount of information carried by neural responses. We found that firing rates and precisely timed spike patterns of individual units became more reliable across trials and time when multisensory stimuli were presented, leading to greater encoded stimulus information. Importantly, this multisensory information enhancement was much reduced when the visual stimulus did not match the sound. These results demonstrate that multisensory influences enhance information processing already at early stages in cortex, suggesting that sensory integration is a distributed process, commencing in lower sensory areas and continuing in higher association cortices.

## Results

The term “early” multisensory influence refers to the impact that stimuli presented to one modality have on neuronal activity at lower (early) stages of another modality [3, 4, 6]. One prominent model system for such early multisensory influences is the auditory cortex [7], and functional imaging (e.g., [8, 9]) and electrophysiological studies [5, 10–13] have revealed response modulations in primary or secondary auditory cortex when a sound is accompanied by visual or somatosensory stimuli. Although these response modulations depend on spatiotemporal stimulus alignment, it remains unknown whether multisensory influences actually enhance the reliability of responses and, as a result, enable individual neurons to carry more information during multimodal conditions.

To address this issue, we recorded neuronal responses to naturalistic audiovisual stimuli from auditory cortex of alert

macaque monkeys. Recordings targeted regions to which previous studies already localized multisensory influences (caudal A1 and caudal belt fields [9, 11]). Stimuli were presented as just movie, as just sound, or as bimodal audiovisual pair (see [Supplemental Experimental Procedures](#) available online). Building on previous work, we quantified multisensory influences on the strength and trial by trial reliability of responses, and on the sensory information they convey [5, 14].

## Visual Influences on Neurons in Auditory Cortex

We first consider the effect of visual stimuli on mean (trial-averaged) firing rates of individual units. Responses of example units are shown in [Figure 1A](#) and [Figure S1](#). Both units reveal a clear temporal modulation of firing, where epochs of relatively vigorous responses alternate with periods of low firing. Such modulated responses were typical and illustrate the fact that these neurons encode naturalistic stimuli by temporal modulations of firing. Importantly, and as predicted by previous studies [5, 10, 12, 13], responses to sounds (black lines) were altered when the sound was accompanied by the corresponding movie (orange): in several instances the audiovisual response exceeded or fell below the auditory response (see also response amplitudes in [Figure 1B](#)).

To understand whether visual stimuli differentially modulate responses to effective and ineffective auditory stimuli, we divided the time axis into three regimes, characterized by “weak,” “intermediate,” and “strong” auditory responses (blue, green, red color code in [Figure 1A](#) and other figures). This partition allowed us to differentiate periods of relatively vigorous responses from periods where responses did not differ from spontaneous baseline activity: prestimulus baseline activity mostly covered the low and intermediate regime, but only rarely fell in the strong regime.

On average across the population ( $n = 78$  units, 3 animals) and the different time windows used to compute the response ( $T = 6.6$  to 33 ms), weak auditory responses were enhanced by the visual stimulus, whereas strong responses were reduced ([Figure 1C](#)). Visual modulation was significant for the weak (ANOVA across time windows,  $F = 6.0$ ,  $p < 0.05$ ) and strong regime ( $F = 5.3$ ,  $p < 0.05$ ), but not for the intermediate ( $F = 0.1$ ,  $p = 0.89$ ). The differential modulation of weak and strong responses emerged not only in the population average, but was also clearly detectable in individual units. The total visual modulation of firing rates varied between 51% (mean value, [Figure 1D](#)) for weak and 22% for strong responses (averaged across time windows) and was of similar order as reported in previous studies [5]. Notably, total multisensory modulation was stronger for weak auditory responses, conforming to the principle of inverse effectiveness [1, 15, 16].

## Visual Influences on Response Reliability

In addition to the average response strength, other parameters can affect how neurons encode sensory information, the most important probably being the response variability across repeats of a stimulus [17]. Visual inspection of spike rasters over multiple trials suggests that responses to auditory stimuli are robust across trials and that the variability varies over time windows ([Figure 1B](#)). The Fano factor (ratio between variance

\*Correspondence: [christoph.kayser@tuebingen.mpg.de](mailto:christoph.kayser@tuebingen.mpg.de)

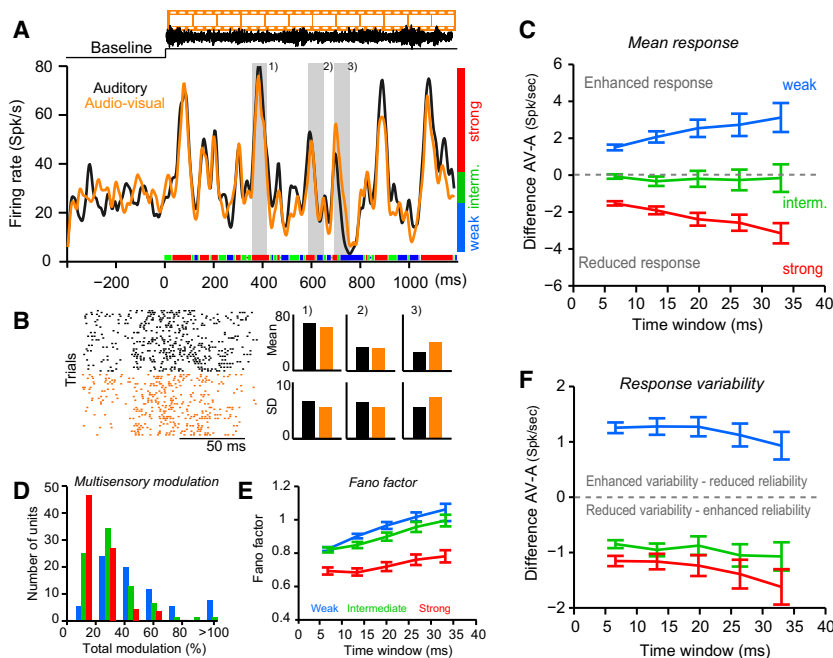


Figure 1. Visual Modulation of Neuronal Firing Rates

(A) Mean response of an example unit for auditory and audiovisual stimulation (trial-averaged firing rate, smoothed with 20 ms Gaussian kernel). The red, green, and blue color code indicates epochs of strong, intermediate, and weak auditory responses, respectively. These are defined according to the neurons' response strength in the auditory condition, as indicated by the colored bar on the right, and were used to group time windows for further analysis (colored dots on x axis). In general, the weak and intermediate regime covered the response range of spontaneous activity, whereas the strong regime comprised the relatively vigorous responses of each unit.

(B) Spike raster plot (left), mean firing rate, and trial-to-trial variability (SD) for selected epochs (right, numbers correspond to gray shading in panel A). Variability was defined as standard deviation of spike count across trials. At different times, the audiovisual response is enhanced or reduced compared to the unimodal response, and changes in mean amplitude are accompanied by changes in variability.

(C) Visual influences on response amplitudes (firing rates) computed in different time windows. Lines denote the mean and standard error of the mean (SEM) across units ( $n = 73$ ) of the difference between auditory and audiovisual conditions ( $AV - A$ ). Positive values indicate response enhancement.

(D) Magnitude of firing rate modulation across units. Modulation was defined as the absolute difference between auditory and audiovisual conditions in percentage, computed regardless of the sign of modulation. Values were averaged across time windows. Mean values are 22%, 31%, and 51% for weak, intermediate and strong responses.

(E) Relative variability of firing rates as a function of time window for unisensory auditory stimulation. The Fano factor is defined as the ratio of variance to mean (both computed across trials), and was averaged over time windows. Lines denote mean and SEM across units.

(F) Visual influences on response variability (computed as standard deviation across trials). Lines denote the mean and SEM of the difference between auditory and audiovisual conditions ( $AV - A$ ). Positive values indicate enhanced trial-to-trial variability, hence reduced response reliability.

and mean of spike counts) was, on average, below one and was particularly low for strong stimuli (Figure 1E). In agreement with recent studies [18–20], this demonstrates that responses of sensory neurons are more reliable than expected from a Poisson process (Fano Factor of 1) and that those responses to stimuli eliciting vigorous responses are the most reliable ones (Figure 1E).

Quantifying multisensory influences on the trial-to-trial variability, we found that visual stimuli differentially affected the variability of weak and strong auditory responses (Figure 1F): variability increased for epochs of weak responses (ANOVA,  $F = 4.4$ ,  $p < 0.01$ ), but decreased for epochs of intermediate or strong responses (intermediate:  $F = 3.5$ ,  $p = 0.058$ ; strong:  $F = 14.2$ ,  $p < 10^{-3}$ ). This result was robust across time windows and demonstrates that strong auditory responses, although reduced in response amplitude, become more reliable. Of note, this is not a trivial consequence of strong responses being reduced in multisensory conditions, because the overall response variability was lower than for a Poisson process and the scaling of variance to mean depended on the mean itself (c.f. Figure 1E). Using information theory, we then showed that these increases in response reliability have a direct impact on the encoded stimulus information.

### Visual Influences on Information Coding

Information between stimuli and neural responses is a principled measure of single-trial discriminability (Supplemental Experimental Procedures) [21, 22]. It quantifies the reduction of uncertainty about the stimulus that is gained by observing a single-trial response. Here, we computed information that responses convey about which section of the stimulus was

being presented, by dividing the stimulation time into nonoverlapping windows and considering the neural response collected in each window as elicited by a different “stimulus” [23]. Importantly, this stimulus characterization does not make assumptions about which specific sensory features drive responses, and it takes into account the contribution from any stimulus features presented up to that “stimulus” time window. On the basis of previous findings that auditory neurons likely encode stimuli by temporal patterns of activity [14, 24–27], we characterized responses by using two neural codes: modulations of firing rates (spike count) and precise temporal spike patterns.

The firing rate code quantifies the stimulus discriminability provided by the response strength within individual time windows but discards additional discriminability provided by the precise temporal structure of the response within each time window. Information in firing rates significantly increased with increasing window length (Figure 2A, left column; ANOVA,  $F = 42$ ,  $p \approx 0$ ), because longer windows allow more spikes and thus more “symbols” to carry information [14, 28]. In addition, information conveyed by epochs of weak or intermediate responses was not significantly different from zero (99% confidence interval from randomized data, gray shading Figure 2A), demonstrating that only the relatively more vigorous responses convey information about the stimulus.

Importantly, information provided by firing rates was significantly higher during audiovisual stimulation (ANOVA, epochs of strong responses  $F = 30$ ,  $p < 10^{-7}$ , Figure 2B). This is further illustrated in Figure 2C for firing rates computed in 26 ms windows, where 58 of 78 units (74%) conveyed more information in the bimodal condition (sign-test,  $p < 10^{-4}$ ). The

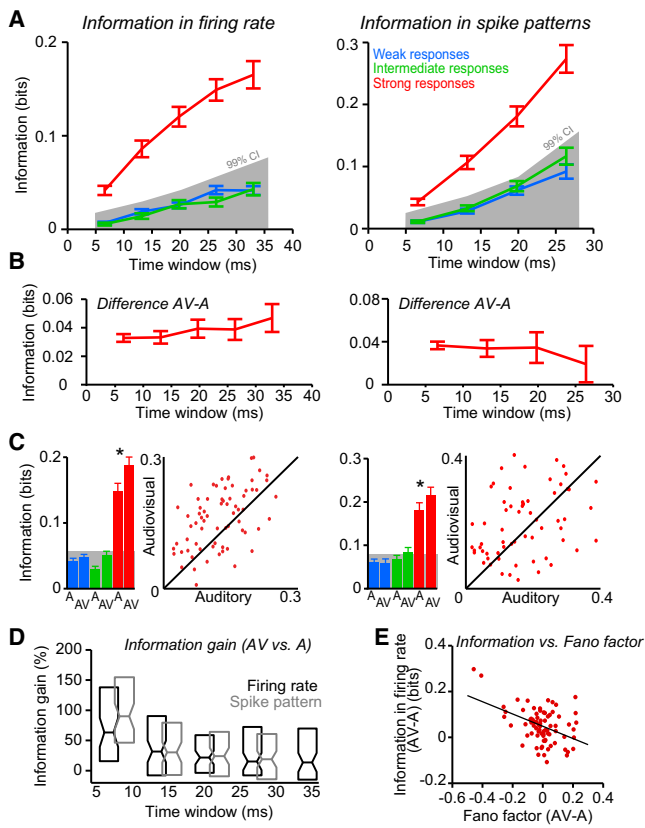


Figure 2. Visual Enhancement of Stimulus Information

(A) Stimulus information (in units of bits) for firing rates (left) and temporal spike patterns (right) computed in different time windows ( $T$ ), shown separately for epochs of weak, intermediate, and strong responses. The red, green, and blue color code indicates epochs of strong, intermediate, and weak auditory responses, respectively. Firing rates were defined as the spike count in the indicated time window  $T$ , spike patterns by subdividing each window into 1–4 bins of length  $\Delta t = 6.6$  ms ( $T = n \cdot \Delta t$ , with  $n = 1 \dots 4$ ). Lines denote mean and SEM across units. Gray shading indicates insignificant information values (99% confidence interval obtained from randomized data), and reveals that significant amounts of stimulus information are only represented during epochs of strong responses (red lines). Note that to avoid biases in information estimates, spike patterns were only computed up to  $T = 26$  ms.

(B) Difference in information between auditory and audiovisual conditions ( $AV - A$ ) as a function of time window for epochs of strong responses (mean and SEM across units). Positive values indicate an information gain in the multisensory condition.

(C) Mean information for firing rates ( $T = 26$  ms window) and spike patterns ( $T = 20$  ms window, subdivided into three time bins, hence characterizing spike patterns comprising three spikes). As in (A), gray shading indicates insignificant information values. Scatter plots display information in auditory and audiovisual conditions for individual units (dots). Asterisks indicate significant differences (paired  $t$  tests,  $p < 0.05$  or smaller).

(D) Information gain in audiovisual relative to auditory condition in units of percent. Box plots display the median and 25th and 75th quartiles across units.

(E) Scatter plot of the information difference in firing rates ( $T = 26$  ms window) between audiovisual and auditory conditions versus the difference in the relative response variability (Fano factor). Each dot denotes one unit. Information gain correlates significantly with a reduction in variability (negative correlation,  $r = -0.44$ ,  $p < 10^{-4}$ ).

relative multisensory information gain ranged up to several hundred percent, with the median being around 30% (Figure 2D). Notably, the multisensory information gain exceeded the linear summation of information available in visual and

auditory conditions, because information for just visual stimuli was insignificant (bootstrap test,  $p < 0.01$ ).

We confirmed that this multisensory gain does not depend on the exact stimulus definition used for information analysis. Although the above results were derived with a feature-independent definition (temporal binning), we found similar multisensory information gains when information was computed with the amplitude of each sound's envelope used as feature (Figure S2). In addition, we verified that the multisensory information enhancement is not dependent on the particular choice of stimulus. Sorting neurons according to whether they were stimulated with behaviorally relevant stimuli (conspecific vocalizations) or with other complex naturalistic stimuli did not reveal an impact of stimulus type on response modulation or information enhancement.

Overall, these results demonstrate that visual stimuli not only modulate the strength of neuronal firing but also enhance the trial-to-trial reliability of responses and the information they carry. In fact, the multisensory information gain (during epochs of strong responses) is directly related to the reduced trial to trial variability: Figure 2E reveals a significant negative correlation between the difference ( $AV - A$ ) in firing rate information and the difference in relative variability (the Fano factor,  $n = 78$ ,  $r = -0.44$ ,  $p < 10^{-4}$ ). Multisensory information gain hence results from a reduction in response variability.

In addition to carrying information by firing rates, auditory neurons may transmit extra information by precise temporal patterns of activity within the encoding time window [14, 24–27]. Addressing multisensory influences on such temporal neural codes, we first considered the information carried by precise temporal spike patterns. As for firing rates, stimulus information in precise spike times also increased with increasing window length (Figure 2A, right column,  $F = 95$ ,  $p \approx 0$ ), and information during epochs of weak or intermediate responses (and for just visual stimulation) was insignificant. Information in spike patterns was also enhanced during multisensory stimulation ( $F = 8.4$ ,  $p < 0.01$ , Figure 2B), as illustrated in Figure 2C for patterns comprising three spikes ( $T = 20$  ms time window divided into 3 time bins): 49 of 78 units (62%) conveyed more information in the audiovisual condition (sign-test  $p < 0.05$ ). The information gain for spike patterns ranged between 10% and 30% (Figure 2D).

### Information Gain Depends on Audiovisual Congruency

Benefits of multisensory integration are typically attributed to matching features in both modalities, such as temporal coincidence or semantic congruency [1, 2, 29–31]. Indeed, we found that the information gain was reduced during conditions of audiovisual mismatch. To this end, 33 additional neurons were recorded with an extended stimulus paradigm, containing a mismatch condition that was created by pairing a sound with a movie that had both different semantic content and temporal structure (Figure 3A). As an example, Figure 3B displays the response of one unit, which was considerably more informative during matching than during mismatching bimodal or during unimodal stimulation.

Population analysis confirmed such a reduced information gain in firing rates during the mismatch condition (ANOVA epochs of strong responses,  $F = 8.0$ ,  $p < 0.01$ ), which is shown in Figure 3C for firing rates in  $T = 26$  ms windows and in Figure 3D for all time windows. This demonstrates that information gain by visual stimulation is dependent on the correspondence of movie and sound, and hence reflects the intrinsic correspondence between visual and auditory stimuli,

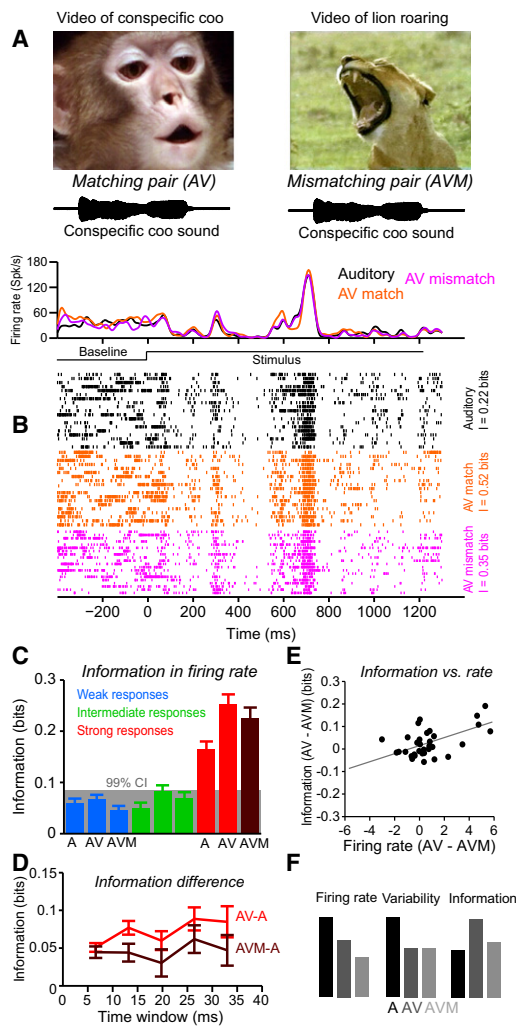


Figure 3. Information Gain Depends on Audiovisual Correspondence

(A) Example stimuli from the mismatch paradigm. Each sound (here a conspecific “coo” vocalization) was presented once accompanied by the corresponding movie (here the face of the vocalizer, audiovisual matching condition, AV), and once by a mismatching movie that shared neither the same image content nor the same temporal structure (here a lion roaring, audiovisual mismatching condition AVM). Note that the mismatch was purposely defined by different dynamics and semantic incongruency.

(B) Mean response (trial averaged firing rate, smoothed with a 20 ms Gaussian kernel), spike rasters, and information values for one example. This unit responded provided 0.22 bits of information during auditory stimulation (firing rate in  $T = 23$  ms windows). In the AV condition, firing became more reliable (e.g., around 700 ms) and exhibited an additional small response peak (around 600 ms), resulting in increased information (0.52 bits). In the AVM condition, reliability and information (0.35 bits) were reduced; hence, this unit exhibited a much reduced information gain during mismatching stimulation.

(C) Information in firing rates ( $T = 26$  ms window, mean and SEM across units,  $n = 33$ ). As in Figure 2, gray shading indicates the 99% confidence interval obtained from randomized data and indicates insignificant information values.

(D) Difference of information in firing rates between audiovisual and auditory conditions for different time windows and epochs of strong responses. Lines denote mean and SEM across units. Across time windows, the difference between matching and mismatching conditions was significant (ANOVA  $F = 8.0$ ,  $p < 0.01$ ).

(E) Scatter plot of the difference in information in firing rates ( $T = 26$  ms window) between AVM and AV conditions versus the difference in mean firing rate for individual units (dots). The correlation ( $r = 0.53$ ,  $p < 0.01$ )

rather than just reflecting the presence of a visual stimulus per se. For spike patterns, the information gain also appeared to be reduced in the mismatch condition, but the results did not reach statistical significance (Figure S3). Of note, both for firing rates and spike patterns, stimulus information during audiovisual mismatch was still higher than information during unimodal stimulation. We assume that this residual information gain results from the common onset of sound and mismatching movie relative to the baseline period. This shared timing of both stimuli (a consequence of the specific experimental paradigm used) effectively “binds” any sound with any movie in the temporal domain. Such temporal congruency by synchronous appearance of sound and visual stimulus might be driving a second kind of multisensory modulation that is not specific to more complex audiovisual features, but nevertheless enhances the reliability of auditory responses to some degree. Differentiating such stimulus specific from unspecific multisensory influences will be one challenge for future work.

The reduced information gain during mismatch results from a change in response strength compared to the matching condition and not from a change in response variability. This was revealed by correlating differences in firing rate information (AV – AVM) with the corresponding differences in mean response or variability: information difference did not correlate with variability ( $n = 33$ ,  $r = -0.089$ ,  $p = 0.62$ ), whereas it correlated with firing rate ( $r = 0.53$ ,  $p < 0.01$ ; Figure 3E). Altogether, this promotes the following summary of our findings (Figure 3F): compared with unisensory auditory stimulation (A), mean responses to efficient audiovisual matching stimuli (AV) are reduced, as is their trial-to-trial variability, resulting in a gain of information. During audiovisual mismatch condition, reduced response variability is accompanied by exaggerated response suppression, diminishing the information gain (AVM).

## Discussion

Multisensory integration is a process by which evidence from different modalities is combined in order to facilitate perception and behavioral reactions [1]. Where and how multisensory integration enhances the neural representations underlying such behavioral improvements is only slowly becoming evident.

Recent studies provided evidence for multisensory influences at much earlier processing stages than previously appreciated (e.g., in primary auditory cortex) [3–5, 7, 10–12, 31, 32]. Our results not only confirm the existence of such early multisensory influences but also demonstrate that visual input enhances the trial by trial reliability of responses, which in turn leads to enhanced sensory information provided by firing rates and spike times. This not only suggests that early multisensory influences conform to a strict definition of sensory integration that requires increased reliability of sensory representations or behavior [2], but also enhance auditory representations at

demonstrates that reduced information in during the AVM condition correlates with reduced response strength.

(F) Schematic summary of main findings: Firing rates and response variability for strong auditory responses are reduced during bimodal (AV, AVM), compared with unimodal, stimulation (A). However, this response suppression is further enhanced during mismatching stimulation (AVM), resulting in little information gain.

multiple temporal scales relevant for information processing [14, 24, 25, 33].

Our results demonstrate that the visual influence on information representations in auditory cortex depends on the congruency of visual and auditory inputs. Given that auditory cortex neurons represent the temporal structure of the acoustic environment [34, 35] this congruency effect likely results from the (intentionally) different temporal dynamics between our mismatching auditory and visual stimuli. Although more complex semantic congruency might also play a role in shaping audiovisual interactions, such effects are only now beginning to be investigated at the neural level [30].

Of note, our results reveal a dichotomy between the multisensory impact on response amplitudes and information gain. Several studies emphasize that responses to stimuli eliciting only weak unimodal responses (less effective stimuli) receive proportionally more multisensory enhancement than do responses to more effective stimuli and propose this response enhancement as one important neural correlate of sensory integration [1, 15]. Although we find the same inverse response modulation in our data, our analysis highlights the information gain during strong unimodal responses as another important multisensory effect. To reconcile these findings, it is important to realize that enhancement of little effective stimuli is usually observed in conditions where stimuli are close to psychophysical threshold. Under these conditions, sensory systems are likely in a vigilance mode, and the detection of sudden stimuli from background is an important event that needs to be signaled and hence deserves multisensory amplification [1, 36]. In the present study, in contrast, we used dynamic stimuli at typical everyday intensities, which elicit a train of temporally modulated responses. Under such conditions, individual neurons are likely representing the presence (or absence) of their preferred sensory features, and multisensory influences should rather enhance the reliability or precision of these sensory representations [2]. Together, this suggests that two modes of multisensory integration might exist, one in which the appearance of rare and faint stimuli is enhanced by the enhancement of the respective responses, and one in which the precision and reliability of dynamic representations of complex sensory environments is improved. Whether both modes might cofunction at the same time and how attention or external stimuli cause a switch from one mode to the other remains to be elucidated in future work.

#### Experimental Procedures

Neuronal activity comprising local field potentials and spiking responses was recorded from auditory cortex of three rhesus monkeys (*Macaca mulatta*). Recording sites covered regions of primary auditory cortex and the caudal belt (fields A1, CM, and CL). All procedures were approved by the local authorities (Regierungspräsidium) and were in full compliance with the guidelines of the European Community (EUVD 86/609/EEC). Stimuli were presented while the animal was performing a visual fixation task in a dark and anechoic booth. Trials started once the animal engaged in a central fixation dot and consisted of a 500 ms baseline period, a 1200–1400 ms stimulation period (depending on recording session), and a 300 ms poststimulus period. Stimuli consisted of a range of matching pairs of naturalistic sounds and movies and, for each neuron response, were recorded to one of these audiovisual pairs. Auditory, visual, and audiovisual stimuli were presented in a pseudo-random order. In the mismatch paradigm, we presented the original audiovisual pair ( $A_1, V_1$ , and  $A_1V_1$ ), a different (nonmatching) movie by itself ( $V_2$ ) and paired with the original sound ( $A_1V_2$ ). For statistical analysis, we used a parametric analysis of variance (ANOVA) across neurons, including time windows as additional factors. Signal extraction, sensory stimuli, and information estimates are described in full detail in the Supplemental Information and are similar to those of previous

studies [5, 14, 37]. For information calculations, we relied on a recently developed information theoretic toolbox (<http://www.ibtb.org>) [38].

#### Supplemental Information

Supplemental Information includes three figures and Supplemental Experimental Procedures and can be found with this article online at doi:10.1016/j.cub.2009.10.068.

#### Acknowledgments

This work was supported by the Max Planck Society and by the BMI project at the Italian Institute of Technology. We are grateful to Cesare Magri for programming efficient routines for data analysis.

Received: August 23, 2009

Revised: October 25, 2009

Accepted: October 26, 2009

Published online: December 24, 2009

#### References

- Stein, B.E., and Meredith, M.A. (1993). *Merging of the Senses* (Cambridge: MIT Press).
- Ernst, M.O., and Bulthoff, H.H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 8, 162–169.
- Ghazanfar, A.A., and Schroeder, C.E. (2006). Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285.
- Foxe, J.J., and Schroeder, C.E. (2005). The case for feedforward multisensory convergence during early cortical processing. *Neuroreport* 16, 419–423.
- Kayser, C., Petkov, C.I., and Logothetis, N.K. (2008). Visual modulation of neurons in auditory cortex. *Cereb. Cortex* 18, 1560–1574.
- Schroeder, C.E., and Foxe, J. (2005). Multisensory contributions to low-level, “unisensory” processing. *Curr. Opin. Neurobiol.* 15, 454–458.
- Kayser, C., Petkov, C.I., and Logothetis, N.K. (2009). Multisensory interactions in primate auditory cortex: fMRI and electrophysiology. *Hear. Res.* 258, 80–88.
- Murray, M.M., Molholm, S., Michel, C.M., Heslenfeld, D.J., Ritter, W., Javitt, D.C., Schroeder, C.E., and Foxe, J.J. (2004). Grabbing your ear: Rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cereb. Cortex* 15, 963–974.
- Kayser, C., Petkov, C.I., Augath, M., and Logothetis, N.K. (2007). Functional imaging reveals visual modulation of specific fields in auditory cortex. *J. Neurosci.* 27, 1824–1835.
- Bizley, J.K., Nodal, F.R., Bajo, V.M., Nelken, I., and King, A.J. (2006). Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cereb. Cortex* 17, 2172–2189.
- Lakatos, P., Chen, C.M., O’Connell, M.N., Mills, A., and Schroeder, C.E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53, 279–292.
- Ghazanfar, A.A., Maier, J.X., Hoffman, K.L., and Logothetis, N.K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J. Neurosci.* 25, 5004–5012.
- Besle, J., Fischer, C., Bidet-Caulet, A., Lecaigard, F., Bertrand, O., and Giard, M.H. (2008). Visual activation and audiovisual interactions in the auditory cortex during speech perception: Intracranial recordings in humans. *J. Neurosci.* 28, 14301–14310.
- Kayser, C., Montemurro, M.A., Logothetis, N., and Panzeri, S. (2009). Spike-phase coding boosts and stabilizes the information carried by spatial and temporal spike patterns. *Neuron* 61, 597–608.
- Stein, B.E., and Stanford, T.R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* 9, 255–266.
- Wallace, M.T., and Stein, B.E. (1996). Sensory organization of the superior colliculus in cat and monkey. *Prog. Brain Res.* 112, 301–311.
- de Ruyter van Steveninck, R.R., Lewen, G.D., Strong, S.P., Koberle, R., and Bialek, W. (1997). Reproducibility and variability in neural spike trains. *Science* 275, 1805–1808.
- Kara, P., Reinagel, P., and Reid, R.C. (2000). Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron* 27, 635–646.

19. DeWeese, M.R., Wehr, M., and Zador, A.M. (2003). Binary spiking in auditory cortex. *J. Neurosci.* *23*, 7940–7949.
20. Maimon, G., and Assad, J.A. (2009). Beyond Poisson: Increased spike-time regularity across primate parietal cortex. *Neuron* *62*, 426–440.
21. Borst, A., and Theunissen, F.E. (1999). Information theory and neural coding. *Nat. Neurosci.* *2*, 947–957.
22. Quian Quiroga, R., and Panzeri, S. (2009). Extracting information from neuronal populations: Information theory and decoding approaches. *Nat. Rev. Neurosci.* *10*, 173–185.
23. Strong, S.P., Koberle, R., de Ruyter van Steveninck, R.R., and Bialek, W. (1998). Entropy and information in neural spike trains. *Phys. Rev. Lett.* *80*, 197–200.
24. Schnupp, J.W., Hall, T.M., Kokelaar, R.F., and Ahmed, B. (2006). Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J. Neurosci.* *26*, 4785–4795.
25. Narayan, R., Grana, G., and Sen, K. (2006). Distinct time scales in cortical discrimination of natural sounds in songbirds. *J. Neurophysiol.* *96*, 252–258.
26. Wang, L., Narayan, R., Grana, G., Shamir, M., and Sen, K. (2007). Cortical discrimination of complex natural stimuli: Can single neurons match behavior? *J. Neurosci.* *27*, 582–589.
27. Lu, T., and Wang, X. (2004). Information content of auditory cortical responses to time-varying acoustic stimuli. *J. Neurophysiol.* *91*, 301–313.
28. Panzeri, S., Petersen, R.S., Schultz, S.R., Lebedev, M., and Diamond, M.E. (2001). The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron* *29*, 769–777.
29. van Atteveldt, N., Formisano, E., Goebel, R., and Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron* *43*, 271–282.
30. Ghazanfar, A.A., Chandrasekaran, C., and Logothetis, N.K. (2008). Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *J. Neurosci.* *28*, 4457–4469.
31. Bizley, J.K., and King, A.J. (2008). Visual-auditory spatial processing in auditory cortical neurons. *Brain Res.* *1242*, 24–36.
32. Brosch, M., Selezneva, E., and Scheich, H. (2005). Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J. Neurosci.* *25*, 6797–6806.
33. Engineer, C.T., Perez, C.A., Chen, Y.H., Carraway, R.S., Reed, A.C., Shetake, J.A., Jakkamsetti, V., Chang, K.Q., and Kilgard, M.P. (2008). Cortical activity patterns predict speech discrimination ability. *Nat. Neurosci.* *11*, 603–608.
34. Moore, D.R., Schnupp, J.W., and King, A.J. (2001). Coding the temporal structure of sounds in auditory cortex. *Nat. Neurosci.* *4*, 1055–1056.
35. Nelken, I. (2004). Processing of complex stimuli and natural scenes in the auditory cortex. *Curr. Opin. Neurobiol.* *14*, 474–480.
36. Schroeder, C.E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* *32*, 9–18.
37. Panzeri, S., Senatore, R., Montemurro, M.A., and Petersen, R.S. (2007). Correcting for the sampling bias problem in spike train information measures. *J. Neurophysiol.* *98*, 1064–1072.
38. Magri, C., Whittingstall, K., Singh, V., Logothetis, N.K., and Panzeri, S. (2009). A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. *BMC Neurosci.* *10*, 81.