

Journal Club

Editor's Note: These short reviews of a recent paper in the *Journal*, written exclusively by graduate students or postdoctoral fellows, are intended to mimic the journal clubs that exist in your own departments or institutions. For more information on the format and purpose of the Journal Club, please see http://www.jneurosci.org/misc/ifa_features.shtml.

Reinforcing the Weight Attributed to Auditory Space across Saccades

Olivier Collignon

Centre de Recherche en Neurosciences, Université catholique de Louvain, B-1200 Brussels, Belgium, and Centre de Recherche en Neuropsychologie et Cognition, Université de Montréal, Montréal, Québec, Canada H3C 3J7

Review of Binda et al. (<http://www.jneurosci.org/cgi/content/full/27/32/8525>)

The world around us is made up of events that generally stimulate more than one sense simultaneously. The multisensory nature of our environment offers many benefits such as enhanced discrimination and accelerated reaction to objects, which arise because the brain is able to construct a unified and more robust percept through the integration of information acquired by the different senses (Stein and Meredith, 1993). However, the presence of conflicting information across different sensory modalities can challenge the creation of a single representation of the object or the event. Because such inconsistent situations give us important insights into how the brain synthesizes information, they have become standard paradigms in the field of multisensory perception (Calvert et al., 2004). In the spatial domain, the best example is certainly the “ventriloquist effect,” where discrepancies are introduced to alter the spatial relationship between auditory and visual stimuli that subjects believe originate from the same location. In this case, be-

cause the visual system typically provides the more accurate and reliable spatial information, the brain affords more weight to visual information in localizing the audiovisual event, thus inducing a “visual capture” of acoustic space (Pick et al., 1969). This explains why although a movie actor’s voice comes from loudspeakers far away from the screen, our brain recalibrates this discrepancy to give us the false impression that sound is actually coming from their mouth. However, such visual dominance does not occur in a rigid, hardwired manner, but follows flexible situation-dependent rules that allow information to be combined with maximal efficacy. For example, though vision provides the most reliable information to localize an object in daylight, at night, auditory information is often more useful. Investigators have already addressed the question of the influence of unimodal reliability on multisensory interaction by “artificial” manipulation of the salience of the stimuli, for example, by adding noise to the most reliable modality (Ernst and Banks, 2002).

In their recent article published in the *Journal of Neuroscience*, Binda et al. (2007) explored this topic by cleverly using saccades, which occur more than three times per second in humans, as a natural cause of visual distortion. The authors took advantage of the fact that saccades have little influence on auditory space perception but distort visual space to explore how vi-

sual reliability affects auditory influence on audiovisual localization. During the task, participants were asked to determine the right/left position of a test stimulus relative to a probe presented shortly before. The two stimuli were both either unimodal (visual or acoustic) or bimodal (audiovisual). These three conditions (visual, auditory, and bimodal) were tested with fixed gaze or when the test stimulus appeared at the time of a saccade. During fixation, visual localization was far more precise than auditory localization, confirming the superior spatial precision of vision than audition in humans. In the bimodal condition, during fixation, the authors requested subjects to compare the apparent position of an incongruent test stimulus, which had conflicting locations of the simultaneous auditory and visual stimuli, with a bimodal congruent probe. In this case, conflicting stimuli were always localized toward the location of the visual input. This result clearly demonstrates that, during fixation, vision dominates and spatially captures the auditory stimulus (the ventriloquist effect). Interestingly, in the saccadic condition, visual localization was less precise and grossly biased in the direction of the saccade whereas auditory localization remained as precise as during fixation. Indeed, even if the bimodal test stimulus was congruent (auditory and visual stimuli came from the same place), vision and audition provided conflicting spatial cues across sac-

Received Sept. 6, 2007; revised Sept. 21, 2007; accepted Sept. 22, 2007.

O.C. is a postdoctoral researcher at the Belgian National Funds for Scientific Research. I thank Dr. F. Lepore for comments on a previous draft of this manuscript.

Correspondence should be addressed to Olivier Collignon, Université de Montréal, Département de psychologie, Centre de Recherche en Neuropsychologie et Cognition, 90 Vincent d’Indy, CP 6128, Succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7. E-mail: olivier.collignon@umontreal.ca.

DOI:10.1523/JNEUROSCI.4089-07.2007

Copyright © 2007 Society for Neuroscience 0270-6474/07/2712438-02\$15.00/0

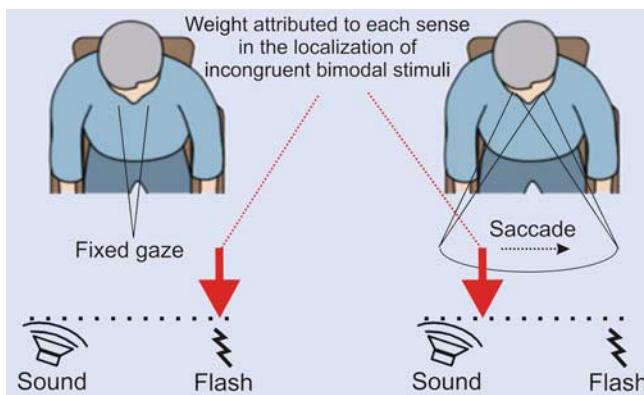


Figure 1. Schematic representation of the main results of Binda et al. (2007). In the fixed gaze condition (left), when spatially conflicting auditory and visual stimuli are presented simultaneously, participants attribute more weight to vision which induce a visual capture of acoustic space. However, when the salience of visual input is challenged, as during saccades (right), the relative weight attributed to sounds increase in the localization of bimodal stimuli.

cades, like the audiovisual conflicting stimuli used during fixation. The authors found that, whereas the position of sounds had no influence on the localization of audiovisual targets during fixation, it significantly affected audiovisual spatial localization during saccades by reducing visual bias (Fig. 1) [Binda et al. (2007), their Fig. 2 (<http://www.jneurosci.org/cgi/content/full/27/32/8525/F2>)]. This clearly demonstrates that when the salience of visual input is challenged, as during a saccade, the relative reliability of the auditory spatial cues increases and more weight is attributed to sounds in the localization of bimodal stimuli. However, one may regret that, because of hardware constraints, the eccentricity of auditory stimuli could only be varied by steps of 6°, compared with the eccentricity of visual stimuli, which could be manipulated by steps of 1°. The use of more precise variations in the location of auditory sources in future studies would allow a more direct comparison between auditory and visual localization threshold, which in turn would certainly refine the conclusions of Binda et al. (2007) regarding auditory–visual interactions.

To better understand the mechanisms underlying such multisensory integration process, the authors investigated whether the Bayesian model may account for the weighting process of each unimodal cue in bimodal integration. This model assumes that when a specific signal has high variance, the system gives it low reliability and, thus, decreases its weight at the time of creation of a unified percept (Ernst and Bülthoff, 2004). Authors report that sensory fusion is extremely well predicted by Bayesian inferences, supporting the notion that auditory and visual spatial cues are combined in a statistically optimal manner, where bimodal localization depends less on visual signals at the time of saccades than during fixation because of increased variance [Binda et al. (2007), their Fig. 5 (<http://www.jneurosci.org/cgi/content/full/27/32/8525/F5>)].

Another frequent situation inducing a decrease in visual saliency is when the same object appears in the peripheral visual field rather than in the fovea. Testing audiovisual integration in the central and the peripheral visual field would thus be interesting to observe whether the results obtained by Binda et al. extend to all nat-

ural situations of visual impoverishment or are specific to precise situations, such as during saccades.

In summary, Binda et al. (2007) provide a clear example of how combining approaches from different disciplines like cognitive psychology, psychophysics, and mathematics can shed new light on how we combine different sources of information. The authors further demonstrate that optimal integration of signals acquired by different sensory modalities has to be made in the context of frequent postural changes involving modifications in the reliability of sensory information. Moreover, it appears that such integration processes are realized in a way that is well predicted by statistical inferences using Bayesian models, allowing clear predictions of audiovisual integration based on the reliability of each modality. However, how this optimal estimation is achieved at the neural level remains far from resolved and should be explored in future studies. The experimental paradigm of Binda et al. (2007) thus provides a new approach for neurophysiological studies in multisensory processing.

References

- Binda P, Bruno A, Burr DC, Morrone MC (2007) Fusion of visual and auditory stimuli during saccades: a Bayesian explanation for perisaccadic distortions. *J Neurosci* 27:8525–8532.
- Calvert GA, Spence C, Stein BE (2004) Handbook of multisensory process. Cambridge, MA: MIT.
- Ernst M, Bülthoff H (2004) Merging the senses into a robust percept. *Trends Cogn Sci* 8:162–169.
- Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–433.
- Pick HL, Warren DH, Hay JC (1969) Sensory conflict in judgements of spatial direction. *Percept Psychophys* 6:203–205.
- Stein BE, Meredith MA (1993) The merging of the senses. Cambridge, MA: MIT.