Neuropsychologia 47 (2009) 1045–1052

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Selective integration of auditory-visual looming cues by humans

Céline Cappe[a,*], Gregor Thut[b], Vincenzo Romei[b], Micah M. Murray[a,c,**]

[a] The Functional Electrical Neuroimaging Laboratory, Neuropsychology and Neurorehabilitation Service and Radiology Service, Centre Hospitalier Universitaire Vaudois and University of Lausanne, rue du Bugnon 46, 1011 Lausanne, Switzerland
[b] Centre for Cognitive Neuroimaging, Department of Psychology, University of Glasgow, Glasgow G12 8QB, United Kingdom
[c] Electroencephalography Brain Mapping Core, Center for Biomedical Imaging of Lausanne and Geneva, rue du Bugnon 46, 1011 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 25 September 2008
Received in revised form 30 October 2008
Accepted 5 November 2008
Available online 12 November 2008

Keywords:
Multisensory
Crossmodal
Perception
Psychophysics
Movement
Distance

ABSTRACT

An object's motion relative to an observer can confer ethologically meaningful information. Approaching or looming stimuli can signal threats/collisions to be avoided or prey to be confronted, whereas receding stimuli can signal successful escape or failed pursuit. Using movement detection and subjective ratings, we investigated the multisensory integration of looming and receding auditory and visual information by humans. While prior research has demonstrated a perceptual bias for unisensory and more recently multisensory looming stimuli, none has investigated whether there is integration of looming signals between modalities. Our findings reveal selective integration of multisensory looming stimuli. Performance was significantly enhanced for looming stimuli over all other multisensory conditions. Contrasts with static multisensory conditions indicate that only multisensory looming stimuli resulted in facilitation beyond that induced by the sheer presence of auditory-visual stimuli. Controlling for variation in physical energy replicated the advantage for multisensory looming stimuli. Finally, only looming stimuli exhibited a negative linear relationship between enhancement indices for detection speed and for subjective ratings. Maximal detection speed was attained when motion perception was already robust under unisensory conditions. The preferential integration of multisensory looming stimuli highlights that complex ethologically salient stimuli likely require synergistic cooperation between existing principles of multisensory integration. A new conceptualization of the neurophysiologic mechanisms mediating real-world multisensory perceptions and action is therefore supported.

1. Introduction

An organism's evolutionary success partially depends on both the ability to reliably detect and discriminate between predators and prey in the environment and also to appropriately respond to them. When encountering an approaching or looming object, one must determine whether to avoid it (a defensive action) or confront it (an aggressive action). Similarly, when encountering a distancing or receding object, one can on the one hand be more assured of one's own safety or can alternatively use this information to determine whether or not pursuit would be worthwhile. In these (and other) ways, simple spatial cues can confer ethologically meaningful information. Given the potentially mortal cost of missing or misinterpreting looming signals, it is unsurprising that ethologists and neuroscientists consider preferential responsiveness to looming signals to be an evolved capacity (Ghazanfar, Neuhoff, & Logothetis, 2002; Graziano & Cooke, 2006; Maier, Chandrasekaran, & Ghazanfar, 2008; Maier, Neuhoff, Logothetis, & Ghazanfar, 2004; Neuhoff, 1998, 2001; Schiff, 1965; Schiff, Caviness, & Gibson, 1962; Seifritz et al., 2002). Moreover, these situations, like many perceptual events, can likely be facilitated by the integration of multisensory cues to enhance perception and render behavior quicker and/or more accurate (Stein & Meredith, 1993; Welch & Warren, 1980).

Multisensory interactions are a fundamental feature of brain organization (Calvert, Spence, & Stein, 2004; Ghazanfar & Schroeder, 2006; Stein & Meredith, 1993; Stein & Stanford, 2008). Studies are increasingly revealing how the brain achieves such multisensory integration. Anatomic evidence now exists for direct projections between unisensory, even primary, cortices (Cappe & Barone, 2005; Falchier, Clavagnier, Barone, & Kennedy, 2002; Rockland & Ojima, 2003). At a functional level, auditory-visual multisensory interactions occur early in time post-stimulus onset and also within areas typically considered unisensory, again including even primary cortices (e.g. Giard & Peronnet, 1999; Martuzzi

* Corresponding author.
** Corresponding author at: The Functional Electrical Neuroimaging Laboratory, Neuropsychology and Neurorehabilitation Service and Radiology Service, Centre Hospitalier Universitaire Vaudois and University of Lausanne, rue du Bugnon 46, 1011 Lausanne, Switzerland. Tel.: +41 21 314 13 21; fax: +41 21 314 46 45.
E-mail addresses: celine.cappe@chuv.ch (C. Cappe), micah.murray@chuv.ch (M.M. Murray).

0028-3932/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2008.11.003

et al., 2007; Molholm et al., 2002; Romei, Murray, Merabet, & Thut, 2007). From such findings, new models of brain organization are being developed that incorporate the occurrence of multisensory interactions and integration both at low and high levels of processes and also at early and late time periods following stimulus presentation (Driver & Noesselt, 2008; Ghazanfar & Schroeder, 2006; Stein & Stanford, 2008; Wallace, Ramachandran, & Stein, 2004).

Given this shift in our conceptualization of brain organization, it is increasingly important to understand the functional significance of multisensory interactions as well as the circumstances governing their occurrence. The seminal works of Stein and Meredith (1993) offer several ‘rules’ of multisensory processing based on receptive field properties of single neurons. More recent data nuance these rules by showing that patterns of interactive effects can be impacted developmentally or through experience (Wallace, Carriere, Perrault, Vaughan, & Stein, 2006; Wallace & Stein, 2007) or even by the spatial heterogeneity within single neurons' receptive fields (Carriere, Royal, & Wallace, 2008). To date, the overwhelming majority of studies have investigated the influences of spatial information on multisensory processing using variation in azimuth or elevation (i.e. 2-dimensional variation in location with respect to the observer). There is comparatively sparse evidence regarding the integration of signals across spatial positions towards versus away from an observer.

Notable exceptions have demonstrated that rhesus monkeys preferentially looked at a looming visual stimulus when presented with a looming, but not receding, sound (Maier et al., 2004). Similarly, 5-month-old infants preferentially looked at matching visual stimuli when presented either with a looming or receding sound (Walker-Andrews & Lennon, 1985). Even though effects were selective for structured sounds instead of noises, the results were only qualitatively suggestive of integrative processes and they did not reveal whether neural response interactions need forcibly be evoked. Likewise, the measurement of looking time cannot differentiate effects occurring at a perceptual level from those driven by biases in attention. Studies of multisensory distance perception by adult humans have predominantly focused on the estimation of time to arrival and remain controversial as whether (and how) auditory and visual distance cues interact and whether or not there is a benefit from multisensory stimulation (Gordon & Rosenblum, 2005; Lewald & Guski, 2004; Sugita & Suzuki, 2003). Moreover, the interpretation of such studies in terms of a neurophysiologic mechanism of either temporal or spatial perception is made complicated by the consistent finding that listeners overestimate the loudness and underestimate the distance of looming sounds (Neuhoff, 1998; Seifritz et al., 2002).

As such, it remains unknown whether multisensory looming/receding signals are integrated to facilitate behavior. Our study addressed this question in humans using a go/no-go motion detection paradigm with unisensory (visual or auditory) and multisensory (simultaneous auditory-visual) stimuli. The perception of visual motion in depth was induced with a central disc that contracted, expanded, or remained constant (i.e. static). The perception of auditory motion in depth was induced with a complex tone that fell or rose in intensity or remained constant (Fig. 1). To ensure that observers used dynamic information in the stimuli, all conditions were initially of the same size/intensity. We assessed multisensory integration of motion perception as measured by reaction times for motion detection (irrespective of its direction or congruence between the senses) and subjective ratings of movement intensity (using a 5-point Likert scale). Performance on multisensory conditions was then compared with that from the constituent unisensory conditions to determine if performance was significantly facilitated to a degree consistent with integrative processes. Finally, the comparison of performance across different multisensory conditions allowed us to determine whether there is selective facilitation for processing multisensory looming signals by humans.

2. Methods

Sixteen healthy individuals (aged 18–32 years: mean = 25 years; 7 women and 9 men) with normal hearing and normal or corrected-to-normal vision participated. All participants provided written informed consent to the procedures that were approved by the Ethics Committee of the Faculty of Biology and Medicine of the University of Lausanne. The main experiment involved the go/no-go detection of moving versus static stimuli that could be auditory, visual, or multisensory auditory-visual (A, V, and AV, respectively). To induce the perception of movement, visual stimuli changed in size and auditory stimuli changed in volume so as to give the impression of either looming or receding (denoted by L and R, respectively). Static stimuli were of constant size/volume (hereafter denoted by S). We chose these stimulus features as they have been previously shown to be the dominant cue for motion-in-depth. The stimulus conditions are schematized in Fig. 1. Specific multisensory conditions were generated using the full range of combinations of movement type (L, R, and S) and congruence between the senses. For convenience we use shorthand to describe experimental conditions such that, for example, ALVL refers to the multisensory combination of auditory looming and visual looming and ARVL refers to the multisensory combination of auditory receding and visual looming. There were fifteen configurations of stimuli in total (6 unisensory and 9 multisensory). Go trials (i.e. those on which either or both sensory modalities contained moving stimuli) occurred on 80% of the time. Each of the 15 conditions was repeated 252 times across 18 blocks of randomly intermixed trials.

The visual stimulus consisted of a centrally presented disc (either black on a white background or white on a black background, counterbalanced across blocks of trials) that symmetrically expanded (from 7° to 13° diameter with the radius increasing linearly at a constant rate over the 500 ms duration of the stimulus) in the case of looming or contracted (from 7° to 1° diameter) in the case of receding. Auditory stimuli were 1000 Hz tones composed of a triangular waveform and generated with Adobe Audition software (Adobe Systems Inc., San Jose, CA, USA). Auditory stimuli were presented over insert earphones (Etymotic model ER4S). The tones were 500 ms in duration and either rose (looming stimulus) or fell (receding stimulus) 10 dB in intensity approximately linearly over this duration (from 77 dB to 87 dB SPL for the looming sound and from 77 dB to 67 dB SPL for the receding sound). They were sampled at 44.1 kHz, had 10 ms onset and offset ramps (to avoid clicks). Prior research has shown that tonal stimuli produce more reliable perceptions of looming and receding (Neuhoff, 1998) and may also be preferentially involved in multisensory integration (Maier et al., 2004). The particular sounds we used were selected after a pilot study of six participants who used a 5-point Likert scale to rate the strength of movement perceived for each of 12 different pairs of looming and receding sounds that differed in their spectral composition (400 Hz, 1000 Hz), waveform type (triangular, square), and manner of intensity modulation (linear, exponential, and variants of the two). We selected the pair of sounds with the strongest ratings.

The rating experiment involved 14 of the original 16 participants (aged 18–32 years; mean = 25 years; 6 women and 8 men). Their task was to indicate the perceived strength of movement, using a 5-point Likert scale, of each of the 15 conditions from the main experiment. Each condition was presented 48 times and was randomly intermixed within a block of trials.

The follow-up experiment included five naïve participants (aged 26–35 years; mean = 29 years; all women). As in the main experiment, their task was to indicate as fast and as accurately as possible whether or not they perceived movement. There were auditory, visual, and multisensory conditions. Movement types were looming, receding, or static. There were two types of static conditions – low and high – so as to physically match the initial states of the looming and receding conditions, respectively (see Fig. 6A for a schematic including details of stimulus sizes and amplitudes). In contrast to the main experiment, multisensory conditions here always involved the same type of movement (or lack thereof) such that there was never any incongruent multisensory combination. Each of the 12 experimental conditions was repeated 64 times across 4 blocks of randomly intermixed trials.

For all experiments stimulus delivery and response collection were controlled by E-prime software (Psychology Software Tools, Pittsburgh, USA). The stimuli were 500 ms in duration as mentioned above and were presented randomly with an inter-stimuli interval between stimulus presentations varying from 800 ms to 1400 ms.

3. Results

3.1. Multisensory integration of perceived motion in depth

In a first set of analyses, we evaluated if there was evidence for multisensory integration of looming and receding auditory-visual stimulus pairs and if such was affected by the congruence in the direction of perceived motion between the senses. This was done by testing for a redundant signals effect (RSE) (Giard & Peronnet, 1999; Martuzzi et al., 2007; Miller, 1982; Molholm et al., 2002; Raab,