Available online at www.sciencedirect.com
SCIENCE DIRECT®

Neuroscience Letters 377 (2005) 59–64

Neuroscience Letters

www.elsevier.com/locate/neulet

Intramodal perceptual grouping modulates multisensory integration: evidence from the crossmodal dynamic capture task

Daniel Sanabria[a,*], Salvador Soto-Faraco[a,b], Jason Chan[a], Charles Spence[a]

[a] *Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK*
[b] *Pare Cientific—Universitat de Barcelona, Barcelona, Spain*

Received 20 July 2004; received in revised form 18 November 2004; accepted 23 November 2004

Abstract

We investigated the extent to which intramodal visual perceptual grouping influences the multisensory integration (or grouping) of auditory and visual motion information. Participants discriminated the direction of motion of two sequentially presented sounds (moving leftward or rightward), while simultaneously trying to ignore a task-irrelevant visual apparent motion stream. The principles of perceptual grouping were used to vary the direction and extent of apparent motion within the irrelevant modality (vision). The results demonstrate that the multisensory integration of motion information can be modulated by the perceptual grouping taking place unimodally within vision, suggesting that unimodal perceptual grouping processes precede multisensory integration. The present study therefore illustrates how intramodal and crossmodal perceptual grouping processes interact to determine how the information in complex multisensory environments is parsed.

Keywords: Perceptual grouping; Multisensory integration; Motion perception; Dynamic capture

Despite the subjective impression that perception is an effortless process, there is a wealth of literature pointing to the tremendous complexity involved in constructing appropriate representations of our environments. One of the complicating factors is that the incoming information impinges on a range of different sensory systems, so the brain needs to integrate all of this multisensory information in order to derive faithful and adaptive representations [4]. Although most research on multisensory interaction processes has focused on the integration of static events, there has been a recent growth of interest in multisensory interactions specifically influencing the perception of motion [13].

Soto-Faraco et al. [14] recently developed a paradigm to investigate these multisensory interactions in the perception of motion. They demonstrated that the presentation of a visual apparent motion stream can modulate the perceived direction of an otherwise unambiguous auditory apparent motion stream. Observers in their study were less accurate in their judgments of the direction of an auditory apparent motion stream on trials where the visual distractor stream moved in the opposite direction (that is, they frequently perceived the sounds as moving in the *same* direction as the lights), than on congruent trials, where the two streams moved in the same direction. This crossmodal dynamic capture effect was modulated by the synchrony and spatial coincidence between the auditory and visual stimuli, with larger multisensory interactions being reported when the auditory and visual stimuli were presented from the same set of locations at about the same time. This phenomenon suggests that visual motion captures the perception of auditory motion and can be seen as an example of perceptual grouping occurring crossmodally. However, we still have very little understanding regarding how crossmodal perceptual grouping interacts with the perceptual grouping occurring within each sensory modality (i.e., with unimodal perceptual grouping).

In fact, it is only recently that researchers have started to explore the effects of intramodal perceptual grouping on crossmodal dynamic interactions. For instance, Sanabria et al. [12] demonstrated that, in order to modulate the crossmodal dynamic capture effect, the conditions for visual perceptual grouping have to be established before the onset

* Corresponding author. Tel.: +44 1865 271307; fax: +44 1865 310447.
E-mail address: daniel.sanabria@psy.ox.ac.uk (D. Sanabria).

0304-3940/$ – see front matter
doi:10.1016/j.neulet.2004.11.069

of the audiovisual motion event. In the present study, we used an adaptation of Soto-Faraco et al.'s [14] crossmodal dynamic capture task to examine the influence of intramodal visual perceptual grouping on the multisensory integration of auditory and visual motion signals. In particular, we used the principles of perceptual grouping to vary the perception of motion within the visual (irrelevant) modality. In Experiment 1, we investigated whether the presentation of the multisensory audiovisual apparent motion event (composed of two lights and two sounds) within a more spatiotemporally extended visual apparent motion stream would modulate the crossmodal dynamic capture effect. In Experiment 2, we addressed the question of whether global visual apparent motion elicited by intramodal perceptual grouping would modulate the crossmodal dynamic capture effect elicited by local visual apparent motion. If intramodal perceptual grouping precedes and supersedes crossmodal perceptual grouping, then one should expect to see a reduced crossmodal dynamic capture effect under conditions where intramodal visual perceptual grouping effects were stronger. Namely, intramodal visual percetual grouping would result in a better perceptual segregation of the visual and auditory unimodal events [7].

In Experiment 1, every trial contained a two-tones auditory apparent motion stream moving from left-to-right or from right-to-left, and a visual apparent motion stream that moved in either the same or opposite direction. The type of visual stream (2 lights versus 6 lights) and the congruency between the direction of auditory and visual motion streams (incongruent versus congruent) were randomized on a trial-by-trial basis. Participants were instructed to indicate leftward versus rightward auditory motion, while attempting to ignore the irrelevant visual stimuli. In the 2 lights condition, the visual apparent motion stream consisted of the sequential activation of the 2 central LEDs in time with the 2 sounds (as in the majority of previous studies of crossmodal dynamic capture). In the 6 lights condition, the flashes from the two central LEDs were embedded within a more spatiotemporally extended apparent motion stream including the sequential illumination of all six LEDs. Note that in both conditions, the onset of the two central flashes coincided temporally with the onset of the two sounds (see Fig. 1a).

Twenty participants (age range: 21–38 years, mean of 27 years) took part in this experiment. All reported normal hearing and normal or corrected-to-normal vision.

Six loudspeakers were positioned in a row (30 cm separation centre-to-centre) at ear level in front of the seated participants. Orange LEDs, one placed just below each loudspeaker, were used to present the visual stimuli. A constantly illuminated red LED, situated between the two central loudspeakers, served as the central fixation point. Participants sat in complete darkness, 90 cm in front of the fixation point, holding a response keypad. The auditory apparent motion stimuli consisted of two 50 ms white noise bursts (70 dB(A) as measured from the participant's head position), one burst presented from each of the two central loudspeakers, and separated by an interstimulus interval (ISI) of 100 ms. The visual apparent motion stimuli consisted of a sequence of two 50 ms light flashes presented from different LEDs, each separated by an ISI of 100 ms. These stimulus parameters have been shown to lead to the perception of apparent motion in both audition and vision [15].

The participants were instructed to prioritize accuracy over response speed, and were also informed about the independence of the direction of the auditory and visual apparent motion streams. The participants were instructed to respond to the direction of the auditory apparent motion stream by making a spatially compatible keypress response (left or right thumb). There was no time constraint on the completion of the trials. The response–stimulus interval was 2000 ms. Participants completed 12 sound-only practice trials (these trials were repeated if participants made more than one error). After the practice session, participants completed two blocks of 96 experimental trials.

The logarithmically transformed accuracy data[1] were submitted to a 2 (Number of Visual Stimuli) by 2 (Congruency) repeated-measures analysis of variance (ANOVA). This analysis revealed a significant main effect of Congruency, $F(1,19) = 61.91$, MSE = 1.41, $p < .01$, with participants performing more accurately on congruent trials than on incongruent trials overall ($M = 94\%$ versus 67%, respectively). The main effect of the Number of Visual Stimuli did not reach significance, $F(1,19) = 3.03$, MSE = 0.27, $p = .09$. Crucially, the interaction between Congruency and Number of Visual Stimuli was significant, $F(1,19) = 7.42$, MSE = 0.61, $p = .013$, with a stronger effect of congruency in the 2 lights condition than in the 6 lights condition (see Fig. 1b). Subsequent post hoc comparisons revealed a significant difference in performance on incongruent trials between the 2 lights and 6 lights conditions, $t(19) = 2.60$, $p = .01$, Bonferroni corrected, $\alpha = 0.025$ [3,6]; but no such difference for congruent trials, $t(19) = 1.91$, $p = .07$, Bonferroni corrected, $\alpha = 0.025$ [3,6]. Note, however, that the numerical trend in the congruent trials toward better performance in the 2 lights condition than in the 6 lights condition (96% versus 93%, respectively) may not have reached statistical significance here because of the near ceiling performance. These two results suggest that intramodal perceptual grouping not only modulates performance on incongruent trials (by weakening the effect of visual motion information on the perception of auditory motion) but may also affect performance on congruent trials.

The results of Experiment 1 highlight the influence of the direction of visual apparent motion on the perception

[1] When using direct proportion scores, parametric statistical tests may provide unrealistic significance values (i.e., overestimating) certain comparisons because of the artificially reduced variability found near to the ends of the distribution (i.e., when accuracy scores fall close to 0 or to 1). Consequently, the accuracy data were transformed using the following logarithmic equation prior to statistical analysis: $p = \ln[p/(1 - p)]$ to normalize the data [18]. Note, however, that the data in the text and the figures are still described in terms of percentages of correct responses for ease of understanding. The pattern of statistical results was identical when using untransformed accuracy data.

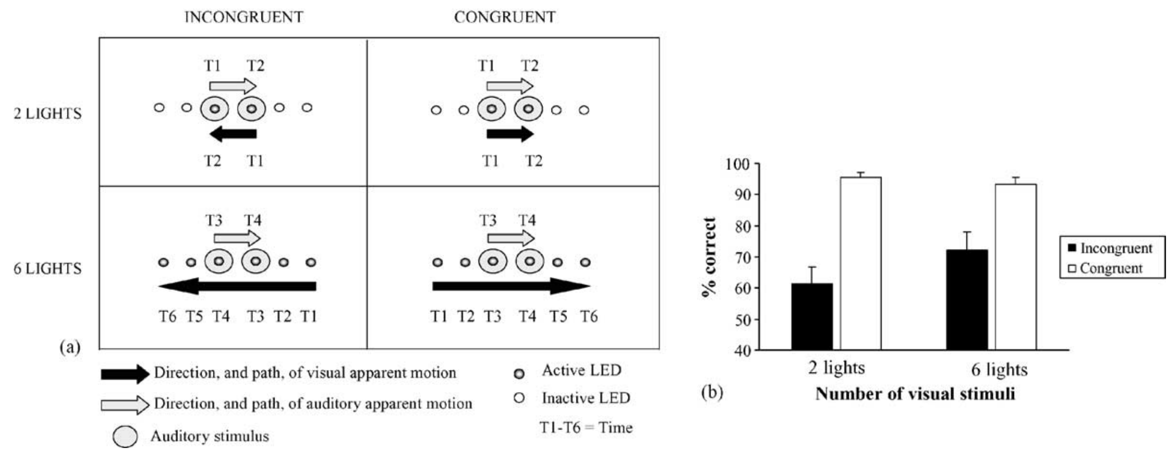

Fig. 1. (a) Four of the trial types presented in Experiment 1, resulting from the crossing of the Number of Visual Stimuli (2 lights vs. 6 lights) and Congruency (incongruent vs. congruent) factors. There were also four more trial types in which the stimuli moved in the opposite direction (i.e., with the direction of auditory apparent motion from right to left, and the direction of visual motion from left to right). T1–T6 represents the sequence of onsets of each of the stimuli (auditory and visual) presented during a particular trial. The upper labels indicate the onset times of the auditory stimuli and the lower labels the onset times of the visual stimuli, (b) mean accuracy (+S.E.) in discriminating the direction of auditory apparent motion as a function of the Number of Visual Stimuli and Congruency.

of the direction of auditory apparent motion, replicating Soto-Faraco et al.'s [14,16] previous findings.[2] Critically, our results also demonstrated that this multisensory integration (or *crossmodal* grouping) effect can be modulated by the perceptual grouping processes taking place *within* vision [12]. When the visual distractor consisted of a 2 lights apparent motion stream, we found significantly more capture than when it consisted of a 6 lights stream. Note that the visual events presented in temporal coincidence with the target auditory stream (the central lights) were identical in both conditions, differing only in that the central lights were embedded within a more spatiotemporally extended stream in the latter condition.

The reduction in the magnitude of the crossmodal dynamic capture effect reported in the 6 lights condition occurred despite the fact that increasing the number of visual stimuli in the visual apparent motion stream resulted in a stronger "impression" of visual apparent motion.[3] These somewhat counterintuitive results (i.e., that better motion produces less capture) make sense if one considers that, as a consequence of the spatiotemporally extended duration of the 6 lights visual apparent motion stream, participants were less likely to perceptually integrate the lights and the sounds, as compared to the 2 lights condition (where the total number and temporal sequence of stimuli in each sensory modality was matched). Hence, the perceptual segregation of the auditory and visual streams would have been better in the 6 lights condition and, as a consequence, the higher level conditions for multisensory integration of visual and auditory motion were weaker [7].

Note that if the influence of the visual apparent motion stream had simply been to elicit a response bias when responding to the direction of auditory motion (i.e., if the direction of visual apparent motion had primed the response consistent with its direction of motion), then *more* (and not less) crossmodal dynamic capture should have been reported in the 6 lights condition than in the 2 lights condition (as the former condition contained stronger directional information). Our results therefore argue against a response bias interpretation of the dynamic capture effect and suggest instead that intramodal visual perceptual grouping modulated multisensory integration (i.e., crossmodal grouping).

One might also argue that the first 2 lights in the 6 lights sequences used in Experiment 1 provided some sort of temporal warning, or alerting, signal [11,17] regarding the imminent onset of the auditory stimuli (which coincided with the third and fourth lights in the stream). According to this view, alerting may have facilitated perceptual segregation of visual and auditory stimuli in the 6 lights condition. However, this possibility has been ruled out elsewhere in studies that have shown that the presentation of a central visual alerting signal (the fixation point flashing with the same temporal profile as the first 2 lights in the 6 lights visual streams used here) has no significant effect on the magnitude of crossmodal dynamic capture [12].

In Experiment 2, we addressed the question of whether the intramodal perceptual grouping elicited by global visual motion would modulate the crossmodal perceptual grouping

2 In order to determine whether the distractor lights facilitated performance on congruent trials or impaired performance on incongruent distractor trials, another group of 10 participants (age range: 19–33 years, mean of 24 years) subsequently completed a block of 32 unimodal auditory motion discrimination trials. Performance was near perfect ($M = 99\%$; just as in the practice trials of the main experiment), suggesting that the visual distractors primarily impaired performance on the incongruent distractor trials, rather than facilitating it on congruent trials.

3 This was confirmed by an additional study in which five new participants rated the quality of apparent motion of the 2 lights and 6 lights unimodal visual apparent motion displays on a 7-point Likert scale (1—poor motion to 7—good motion). A *t*-test comparison revealed that the perceived quality of motion was much higher for the 6 lights ($M = 6.0$) displays than for the 2 lights ($M = 2.6$) displays [$t(4) = -6.97$ $p < .01$], as expected [21].

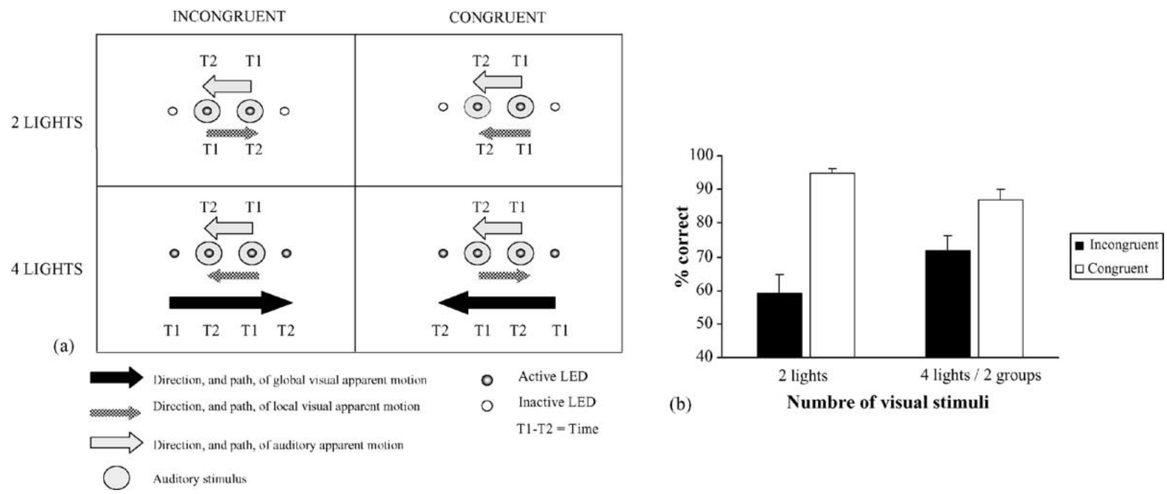

Fig. 2. (a) Four of the trial types presented in Experiment 2, resulting from the crossing of the Number of Visual Stimuli (2 vs. 4 lights) and Congruency (incongruent vs. congruent) factors. The direction of apparent motion refers to the direction of the 2 groups of 2 lights with respect to the two sounds (as shown by the filled black arrows in the figure). In the 4 lights/2 groups condition, two LEDs (one central and the other peripheral on the other side from the centre) were activated simultaneously with each sound. In this situation, the apparent motion of the group of four lights could be congruent with the sounds (while the apparent motion of the two central lights was incongruent with them), or vice versa. There were four more trial types, in which the stimuli moved in the opposite direction (i.e., with the direction of auditory apparent motion from right to left, and the direction of visual motion from left to right). T1 and T2 represents the sequence of onsets of each of the stimuli (auditory and visual) presented during a particular trial. The upper labels indicate the onset times of the auditory stimuli and the lower labels the onset times of the visual stimuli, (b) mean accuracy (+S.E.) in discriminating the direction of auditory apparent motion as a function of the Number of Visual Stimuli and Congruency.

elicited by local visual motion. The 2 lights condition was the same as in Experiment 1, while the 6 lights condition was replaced by a 4 lights/2 groups condition (see Fig. 2a). In this 4 lights/2 groups condition, the two central lights were presented at the same time as the sounds, just as in the 2 lights condition. However, one of the peripheral lights (on the opposite side) was also flashed in synchrony with each of the sounds. Taken together, the 4 lights/2 groups display induced the perception of global motion of 2 groups of 2 lights in the direction *opposite* to the local motion of the two central lights when considered in isolation. If global motion prevails over local motion [10], one would expect that the crossmodal dynamic capture effect should be reversed, and now be determined by the direction of the global motion of the 4 lights/2 groups rather than by the local apparent motion of the two central lights (when considered in isolation). Crucially, we also expected to find *less* capture in the 4 lights/2 groups condition than in the 2 lights condition because the spatiotemporally extended duration of the 4 lights/2 groups stream would give rise to a better segregation of the visual apparent motion stream from the target auditory apparent motion stream.

Thirty participants took part in Experiment 2, 18 of whom had taken part in Experiment 1. The order in which the experiments were completed was counterbalanced across participants.

These were the same as in Experiment 1, with the following exceptions. In the 4 lights/2 groups condition, each visual onset was composed of the simultaneous activation of two LEDs (50 ms duration), one central and the other peripheral (on the opposite side; see Fig. 2a). The outer right LED was always flashed in time with the inner left LED, and the outer left LED was always flashed with the inner right LED. When the central LEDs flashed from left-to-right, it appeared as if 2 groups of 2 lights moved from right-to-left (and vice-versa when the central LEDs flashed from right-to-left). Congruency in the 4 lights/2 groups condition was defined with respect to the global direction of motion of the visual display rather than with respect to the local motion of the two central lights.

The logarithmically transformed accuracy data were submitted to a 2 (Number of Visual Stimuli) by 2 (Congruency) repeated-measures ANOVA. Participants responded more accurately on congruent trials (90%) than on incongruent trials (65%), resulting in a significant main effect of Congruency, $F(1,29) = 38.70$, MSE = 2.60, $p < .01$. There was no main effect of the Number of Visual Stimuli, $F(1,29) < 1$. Once again, the interaction between Congruency and Number of Visual Stimuli was significant, $F(1,29) = 14.84$, MSE 1.25, $p < .01$, indicating that the magnitude of the crossmodal dynamic capture effect in the 4 lights/2 groups condition was smaller than that seen in the 2 lights condition ($M = 15\%$ versus 35%, respectively). Subsequent post hoc comparisons revealed that participants responded more accurately on incongruent trials in the 4 lights/2 groups condition than on the incongruent trials of the 2 lights condition, $t(29) = 2.72$, $p = .01$, Bonferroni corrected, $\alpha = 0.025$ [3,6]. The difference between congruent trials in both distractor conditions was also significant, $t(29) = 4.17$, $p < .01$, Bonferroni corrected, $\alpha = 0.025$ [3,6]. Note also that performance on congruent trials in the 4 lights/2 groups condition was worse than in the unimodal auditory baseline task [$t(38) = 3.27$, $p < .05$; see footnote 2].

This latter result seems to suggest that, although global motion prevailed over local motion (both objectively and subjectively; see the following), the incongruent local motion still had a distracting effect on the perception of the direction of motion of the auditory stream.

The results of the post hoc tests in Experiment 2 indicate that the modulation of the visual capture effect of auditory apparent motion as a consequence of the perceptual grouping taking place within vision was not restricted to incongruent situations but also to performance on motion congruent trials. It appears then that the lack of a significant difference between congruent trials in Experiment 1 was probably caused by participants' performing at ceiling levels on those trials, instead of a lack of modulation of the crossmodal dynamic capture effect by visual perceptual grouping.

The results of Experiment 2 show that there was a reduction in the magnitude of the crossmodal dynamic capture effect in the 4 lights/2 groups condition relative to the 2 lights condition (see Fig. 2b). One important consideration before interpreting this result is to ensure that the impression of apparent motion produced by the 4 lights/2 groups condition was no weaker than that in the 2 lights condition. Otherwise, it could be argued that the reduction in the magnitude (though not its sign or direction; see next paragraph) of the crossmodal dynamic capture effect in the 4 lights/2 groups condition was not, in fact, the result of a perceptual modulation of multisensory integration but was instead a consequence of participants experiencing a weaker impression of visual apparent motion in this condition than in the 2 lights condition. To examine this possibility, we conducted an additional study in which 10 participants were asked to rate the quality of apparent motion of the visual stream (2 lights versus 4 lights/2 groups; sixteen trials with the auditory stimuli present and sixteen trials presented in silence), using a 7-point Likert scale (l = poor apparent motion, 7 = good apparent motion). The results showed that participants rated the quality of visual apparent motion to be just as strong in both conditions (4.7 in both cases). Furthermore, participants reported perceiving only one type of apparent motion in the 4 lights/2 groups condition; the one involving the global motion of the 2 groups of 2 lights.[4]

As noted earlier, there were two possible visual apparent motion percepts (i.e., local versus globally defined) in the 4 lights/2 groups condition. These two potential apparent motion streams had *opposite* directions (i.e., while the two central lights considered in isolation flashed from *right-to-left*, taken together the group of 4 lights appeared as 2 groups of 2 lights moving from *left-to-right*). The direction of global motion of the 4 lights/2 groups display prevailed over the local motion (i.e., the movement of the two central lights) in Experiment 2, and only one direction of motion was unambiguously perceived by the majority (all but one) of the participants, as shown schematically by the arrows in Fig. 2a. Note that the overall direction of the congruency effect, if congruency is assessed with respect to the local motion (i.e., the movement of the two central lights), actually *reversed* in the 4 lights/2 groups condition.

Hence, the crossmodal dynamic capture results converge with the subjective ratings indicating that the two central lights were grouped with the peripheral lights and perceived as 2 groups of 2 lights; It was this emergent global motion that modulated participants' perception of the direction of auditory motion, thus reinforcing the case for the modulation of multisensory integration (or crossmodal grouping) by visual perceptual grouping (independently of the strength of apparent motion in either condition).

The experiments reported here demonstrate that multisensory grouping (assessed in terms of the magnitude of the crossmodal dynamic capture effect) depends on how visual stimuli are perceptually grouped. In Experiment 1, the presentation of the 6 lights stream resulted in significantly *less* dynamic capture of the perceived direction of auditory motion than the presentation of just 2 lights, even though the 6 lights condition induced a stronger impression of visual apparent motion. Experiment 2 replicated these findings, again showing how changes in intramodal visual perceptual grouping can modulate the strength of multisensory integration (or crossmodal grouping; see [12,19,20], for related findings). These results support the view that visual perceptual grouping precedes multisensory integration, thus adding weight to the idea that the multisensory integration of motion information depends upon the configuration of unimodal perceptual information. This claim is consistent with Sanabria et al.'s [12] recent demonstration that, in order to modulate the crossmodal dynamic capture effect, conditions for unimodal visual perceptual grouping must concur *before* the presentation of the multisensory audiovisual event.

When investigating crossmodal contributions to the perception of apparent motion an important issue that needs to be addressed regards the role played by static information. For instance, Allen and Kolers [1] observed that under certain conditions, a single visual flash elicited the same effect as a visual apparent motion stream on the perception of auditory apparent motion. In fact, one might wonder whether the crossmodal dynamic capture reported in the present study could be explained in terms of a 'ventriloquist-like effect' [2]; That is, on incongruent trials, the location of the first static auditory stimulus may simply have been mislocalized toward the location of the first static visual distractor stimulus, and the second static auditory stimulus toward the location of the second static visual distractor, thus impairing the correct perception of the direction of the auditory motion stream. However, other recent work from this laboratory [14–16] has shown that the crossmodal dynamic capture effect cannot be accounted for

[4] Only one participant reported perceiving a zig-zagging pattern of lights in the 4 lights/2 groups displays. Interestingly, this participant showed the inverse pattern of performance in the 4 lights/2 groups condition of Experiment 2, exhibiting less accurate performance on congruent trials than on incongruent trials (65% versus 87%, respectively). Note that, in this case, the local motion of the two central lights in the 4 lights/2 groups condition presumably dominated over the global motion of the 2 groups of 2 lights; and so, contrary to the other participants, the congruency effect was still related to the movement of the 2 central lights.

solely by the influence of the static visual stimuli on the perception of simultaneously presented static auditory stimuli. Instead, the emergent property of apparent motion within the display plays a crucial role in eliciting the crossmodal dynamic capture effect. For instance, a split-brain patient who no longer experiences visual apparent motion across the midline (but who is still able to detect and localize static visual stimuli) experience a much reduced crossmodal dynamic capture effect when compared to normal participants (using a design very similar to the 2 lights conditions reported here) [14]. This result, together with other converging psychophysical data from normal participants [15], confirms that it is the dynamic properties of the distracting visual stimulation that mediates the crossmodal dynamic capture effect, over-and-above any effects attributable to static ventriloquist effects.

Beyond the particular case of the multisensory integration of motion information, our results support the idea that multisensory integration, in general, depends on the way our perceptual system configures incoming sensory information. In the present study, the visual and auditory stimuli were perceptually grouped under particular conditions, giving rise to two perceptual objects (one visual and the other auditory; see [7]). Interestingly, when both perceptual objects are spread over the same medium (i.e., simultaneous in time; using Kuvoby and Valkenburg's [7], terminology), and are of the same number (i.e., 2 lights and 2 sounds), more pronounced multisensory integration occurs (especially if the stimuli are presented from the same set of spatial locations; see [8,9,12]). Consequently, people are more likely to perceive a single multisensory object moving coherently through space and time (and so crossmodal dynamic capture takes place). When additional visual stimuli are presented, a different visual object emerges. Thus, because of the different number of components of the more extended visual stream relative to the auditory stream, multisensory integration was weaker, presumably because participants were more likely to group (and hence to perceive) the auditory and visual stimuli as separate unimodal perceptual objects moving independently through space and time. The interplay between intramodal and crossmodal perceptual grouping identified here presumably influences multisensory integration in both humans and animals.

The findings reported in the present study are also consistent with the existent data from neuroimaging and neurophysiological studies, showing that different areas are involved in the 'early' intramodal perceptual grouping of visual information and the 'late' multisensory integration. For instance, recent neuroimaging studies have suggested primary visual areas as the locus of perceptual grouping by apparent motion [5]. Meanwhile, other studies have shown that a number of 'high level' association areas (e.g., ventral premotor cortex, ventral intraparietal area, lateral parietal cortex, lateral frontal cortex, hMT) appear to play a critical role in the multisensory integration of motion information (see [13], for a review). Added to the behavioural evidence reported here, the neurophysiological studies described above seem to support the idea that multisensory integration of motion information depends upon the configuration of unimodal perceptual information.

Acknowledgement

This study was supported by a Network Grant from the McDonnell-Pew Centre for Cognitive Neuroscience in Oxford to Salvador Soto-Faraco and Charles Spence.

References

[1] P. Allen, P. Kolers, Sensory specificity of apparent motion, J. Exp. Psychol. 7 (1981) 1318–1326.
[2] P. Bertelson, G. Aschersleben, Automatic visual bias of perceived auditory location, Psychon. Bull. Rev. 5 (1998) 482–489.
[3] C.E. Bonferroni, Teoria statistica delle classi e calcolo delle probabilità [Classic theory of the class and probability calculation], Pubblicazioni del Instituto Superiore di Scienze Economiche e Commerciali di Firenze 8 (1936) 3–62.
[4] G. Calvert, C. Spence, B. Stein (Eds.), The Handbook of Multisensory Processes, MIT Press, Cambridge, MA, 2004.
[5] G. Francis, S. Grossberg, Cortical dynamics of form and motion integration: persistence, apparent motion, and illusory contours, Vision Res. 36 (1996) 149–173.
[6] J.C. Hsu, Multiple Comparisons Theory and Methods, Chapman & Hall, London, 1996.
[7] M. Kubovy, D. Van Valkenburg, Auditory and visual objects, Cognition 80 (2001) 97–126.
[8] C.T. Lovelace, B.E. Stein, M.T. Wallace, An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection, Cogn. Brain Res. 17 (2003) 447–453.
[9] G.F. Meyer, S.M. Wuerger, F. Röhrbein, C. Zerzsche, Low-level integration of auditory and visual motion signals requires spatial colocalisation, Exp. Brain Res., in press.
[10] D. Navon, Forest before trees: the precedence of global features in visual perception, Cogn. Psychol. 9 (1977) 353–383.
[11] M.I. Posner, Chronometric Explorations of Mind, Lawrence Erlbaum, Oxford, 1978.
[12] D. Sanabria, S. Soto-Faraco, J. Chan, C. Spence, When does visual perceptual grouping affect multisensory integration? Cogn. Affect. Behav. Neurosci. 4 (2004) 218–229.
[13] S. Soto-Faraco, A. Kingstone, C. Spence, Multisensory contributions to the perception of motion, Neuropsychologia 41 (2003) 1847–1862.
[14] S. Soto-Faraco, J. Lyons, M. Gazzaniga, C. Spence, A. Kingstone, The ventriloquist in motion: illusory capture of dynamic information across sensory modalities, Cogn. Brain Res. 14 (2002) 139–146.
[15] S. Soto-Faraco, C. Spence, A. Kingstone, Assessing automaticity in the audio-visual integration of motion, Acta Psychol., in press.
[16] S. Soto-Faraco, C. Spence, A. Kingstone, Crossmodal dynamic capture: congruency effects of motion perception across sensory modalities, J. Exp. Psychol.: Hum. Perform. Percept. 30 (2004) 330–345.
[17] C. Spence, J. Driver, Audiovisual links in exogenous covert spatial orienting, Percept. Psychophys. 59 (1997) 1–22.
[18] J.W. Tukey, Exploratory Data Analysis, Addison-Wesley, London, 1977.
[19] J. Vroomen, B. de Gelder, Sound enhances visual perception: crossmodal effects of auditory organization on vision, J. Exp. Psychol.: Hum. Percept. Perform. 26 (2000) 1583–1590.
[20] K. Watanabe, S. Shimojo, When sound affects vision: effects of auditory grouping on visual perception, Psychol. Sci. 12 (2001) 109–116.
[21] S. Yantis, T. Nakama, Visual interactions in the path of apparent motion, Nat. Neurosci. 1 (1998) 508–512.