Interaction of Perceptual Grouping and Crossmodal Temporal Capture in Tactile Apparent-Motion

Lihan Chen[1,2], Zhuanghua Shi[2]*, Hermann J. Müller[2,3]

1 Department of Psychology and Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, P. R. China, **2** Department Psychologie, Ludwig-Maximilians-Universität München, München, Germany, **3** School of Psychological Science, Birkbeck College, University of London, London, United Kingdom

Abstract

Previous studies have shown that in tasks requiring participants to report the direction of apparent motion, task-irrelevant mono-beeps can "capture" visual motion perception when the beeps occur temporally close to the visual stimuli. However, the contributions of the relative timing of multimodal events and the event structure, modulating *uni-* and/or *cross*modal perceptual grouping, remain unclear. To examine this question and extend the investigation to the tactile modality, the current experiments presented tactile two-tap apparent-motion streams, with an SOA of 400 ms between successive, left-/right-hand middle-finger taps, accompanied by task-irrelevant, non-spatial auditory stimuli. The streams were shown for 90 seconds, and participants' task was to continuously report the perceived (left- or rightward) direction of tactile motion. In Experiment 1, each tactile stimulus was paired with an auditory beep, though odd-numbered taps were paired with an asynchronous beep, with audiotactile SOAs ranging from −75 ms to 75 ms. Perceived direction of tactile motion varied systematically with audiotactile SOA, indicative of a temporal-capture effect. In Experiment 2, two audiotactile SOAs—one short (75 ms), one long (325 ms)—were compared. The long-SOA condition preserved the crossmodal event structure (so the temporal-capture dynamics should have been similar to that in Experiment 1), but both beeps now occurred temporally close to the taps on one side (even-numbered taps). The two SOAs were found to produce opposite modulations of apparent motion, indicative of an influence of crossmodal grouping. In Experiment 3, only odd-numbered, but not even-numbered, taps were paired with auditory beeps. This abolished the temporal-capture effect and, instead, a dominant percept of apparent motion from the audiotactile side to the tactile-only side was observed independently of the SOA variation. These findings suggest that asymmetric crossmodal grouping leads to an attentional modulation of apparent motion, which inhibits crossmodal temporal-capture effects.

Citation: Chen L, Shi Z, Müller HJ (2011) Interaction of Perceptual Grouping and Crossmodal Temporal Capture in Tactile Apparent-Motion. PLoS ONE 6(2): e17130. doi:10.1371/journal.pone.0017130

Editor: Justin Harris, University of Sydney, Australia

Received October 20, 2010; **Accepted** January 20, 2011; **Published** February 23, 2011

Funding: The study was support by German DFG SFB453 (http://www.sfb453.de) and China Postdoctoral Science Foundation (20100470113). The funders had no role in study design,data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: shi@psy.lmu.de

Introduction

Apparent motion is a common perceptual phenomenon in our daily life. For example, two brief flashes of light separated in both time and space create an illusion of movement from the location of the first flash to that of the second flash when the spatiotemporal display parameters are within appropriate ranges [1]. Apparent motion has been observed in the visual, auditory, and tactile modalities, given the respective physical stimuli. A number of studies have shown that apparent motion in a particular modality may be influenced by static or dynamic events in another modality [2–4]. For example, the direction of auditory motion in one direction can be *captured* by concurrent visual motion in a conflicting direction; by contrast, the perceived direction of visual motion is not affected by incongruent auditory motion [4]. Recent work on crossmodal temporal integration has also shown that apparent motion in one modality can be modulated solely by the timing of events in another modality [5,6]. For example, using a visual apparent-motion paradigm, Freeman and Driver [5] found that, in a repeated two-flash visual apparent-motion stream with equal inter-flash intervals (for which, when presented alone, the perceived motion direction would be ambiguous), auditory beeps slightly lagging or leading the flashes strongly influenced the perceived direction of visual motion - even though the beeps themselves did not provide any spatial information. Following the modality precision hypothesis [7,8], on which the sensory modality with the highest temporal acuity dominates the perception of events in other modalities, Freeman and Driver attributed their results to the timing of the beeps influencing the perceived timing of the visual stimuli. Similar audiovisual temporal interactions have also been found in temporal-order judgment tasks and replicated in a number of other studies. Such influences have been referred to as 'temporal ventriloquism' effect, that is: when auditory and visual stimuli occur slightly asynchronously, the visual stimulus is *pulled* (being captured) into temporal alignment with the auditory stimulus [9–12].

Although crossmodal temporal capture has now been demon-

temporally close to each other, or that share common features, are often perceived as forming a coherent "whole". More recently, perceptual grouping has been shown to be an important factor in crossmodal perception [15]. For example, intramodal grouping and segregation of sound pairs can enhance the segregation and discrimination of concurrent visual events [16–18] and bias visual temporal-order judgments [19]. However, the role of perceptual grouping in visual apparent motion is still controversial. For instance, in a control experiment, Freeman and Driver (2008) manipulated *intra*modal auditory grouping by using evenly alternating high- (H) and low-pitch (L) beeps (i.e., HHLLHH…). They found auditory grouping based on pitch alternation to have little influence on visual apparent motion, from which they concluded that audiovisual temporal integration (the temporal-ventriloquism effect) was not due to unimodal (auditory) perceptual grouping. However, evidence from other studies shows that perceptual grouping can influence crossmodal temporal interactions in perceived motion [6,19,20]. For example, Bruns and Getzmann found that either a continuous sound filling in the

Finally, in Experiment 3, we omitted the synchronous beeps, while varying the SOA of the asynchronous audiotactile pairs, in order to further examine the interaction between crossmodal grouping and crossmodal temporal integration (see Figure 1D). With this manipulation, auditory beeps were paired only with one side (either the left or the right) of tactile taps (which is why we refer to this condition as 'half-paring'). If balanced crossmodal grouping is *not* a precondition for the crossmodal temporal interaction, one would expect the results of Experiment 3 (half-pairing condition) to be similar to those of Experiment 1 (full-pairing condition), since the audiotactile SOAs were the same. Alternatively, if asymmetric crossmodal grouping competes with crossmodal temporal capture, one would envisage differential outcomes between the full and the half-paring conditions (realized in Experiments 1 and 3, respectively): the full-pairing audiotactile stream would be subject to a crossmodal temporal-capture effect (as actually observed in Experiment 1); by contrast, the half-pairing condition (realized in Experiment 3) would show little influence of the auditory timing due to the incomplete grouping of the auditory with the tactile events, analogously to the results of audiovisual temporal-ventriloquism study [6,11]. Experiment 3 failed to reveal a significant influence of the audiotactile SOA, consistent with crossmodal temporal capture being prevented under the half-pairing condition; however, apparent motion was subject to a 'global' (i.e., SOA-independent) biasing effect: there was a strong tendency for perceiving motion from the audiotactile side to the tactile-only side. After detailing the

significantly from each other, $t(10) = -1.322$, $p = 0.216$, indicat-

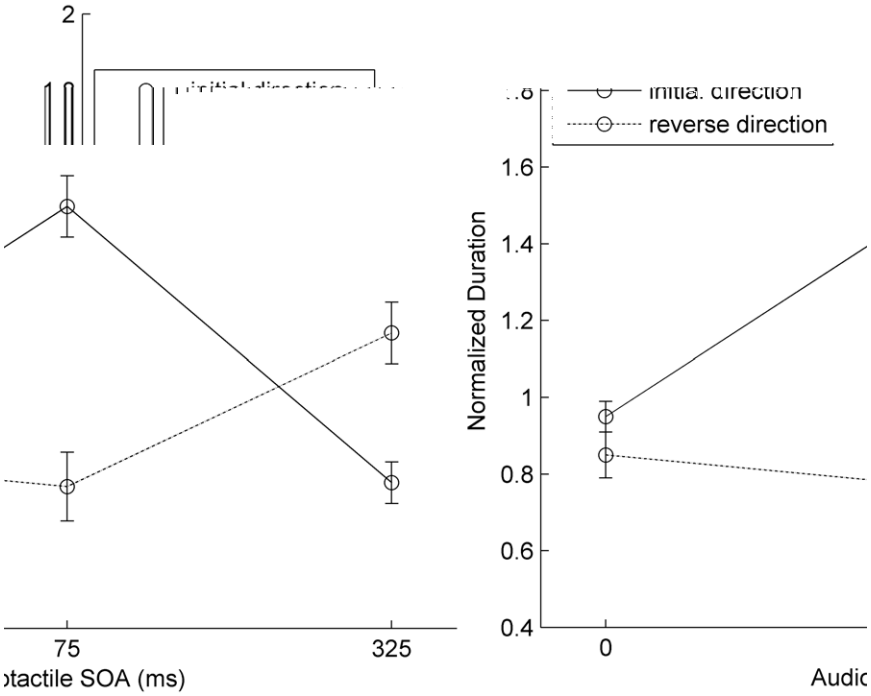

Figure 3. Normalized phase durations of tactile apparent motion in Experiment 2. Normalized phase durations (and associated standard errors) of tactile apparent motion as a function of audiotactile SOA with a shifted full-pairing audiotactile stream.
doi:10.1371/journal.pone.0017130.g003

factor audiotactile SOA, failed to reveal a significant SOA effect, $F(6,60) = 1.069$, $p = 0.391$. Likewise, there were no significant differences among audiotactile SOAs in the phase durations of "reverse-direction" responses, $F(6,60) = 0.451$, $p = 0.841$. Given this, we collapsed the phase durations across all SOAs, separately for "initial-direction" and "reverse-direction" responses, and

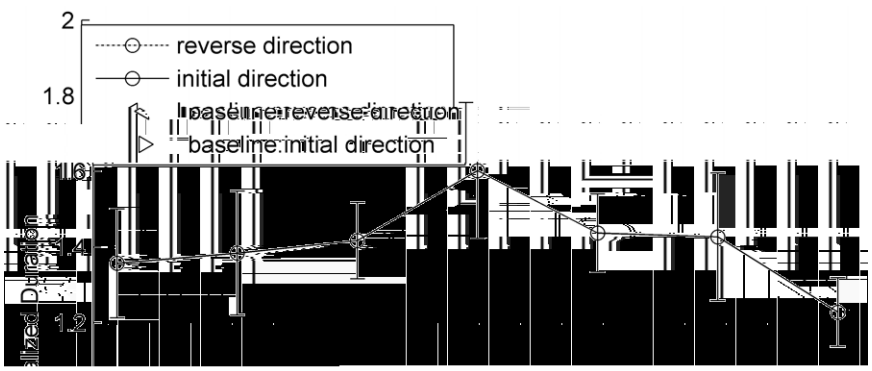

Figure 4. Normalized phase durations of tactile apparent motion in Experiment 3. Normalized phase durations (and associated standard errors) of tactile apparent motion as a function of audiotactile SOA with a half-pairing audiotactile stream. The solid line represents mean phase durations for the "initial direction", the dotted line those for the "reverse direction". Regardless of the audiotactile SOAs, a globally dominant direction of apparent motion, namely, "initial direction", was observed. The rightward-pointing triangle denotes responses of "initial direction", and the leftward-pointing triangle responses of "reverse direction", for the baseline (without-sound) conditions.
doi:10.1371/journal.pone.0017130.g004

compared the resulting values to the corresponding baseline conditions: for the "initial-direction" responses, the phase durations were significantly longer compared to the baseline, $t(10) = 3.140$, $p<0.05$; by contrast, for the "reverse-direction" responses, they were significantly shorter $t(10) = -3.534$, $p<0.01$. Thus, in contrast to Experiment 1, "initial-direction" responses were dominant across all seven audiotactile SOAs, regardless of auditory timing (the audiotactile SOA varied from −75 ms to 75 ms). This indicates that the half-pairing auditory beeps created a "globally" dominant percept of motion direction from the side of the audiotactile stimuli to the side of the tactile-only stimuli.

Discussion

This study examined the influences of perceptual grouping and crossmodal temporal integration of auditory with tactile events in a tactile apparent-motion stream. With a full pairing audiotactile configuration (Experiment 1), we varied the audiotactile asynchronies from −75 ms (beep leading tap) to 75 ms (beep trailing tap) in the odd numbered pairs, while keeping the even numbered pairs synchronous. We observed the (bi-stable) tactile apparent-motion rivalry (i.e., perceived motion going either left- or rightwards) to be systematically resolved by the audiotactile asynchrony. However, contrary to our original expectation, when the audiotactile asynchrony was increased (to 325 ms) such that the (asynchronous) beeps occurred temporally proximal to (i.e., "shifted" towards) the even numbered tactile stimuli, a reversed effect on the direction of apparent motion was found (Experiment 2). In Experiment 3, which used half-pairing audiotactile stimuli, a consistently dominant direction of apparent motion was observed: the dominant direction went from the location (side) with audiotactile stimulus pairings towards the location (side) with a pure tactile stimulus.

The results of Experiment 1 are consistent with Freeman and Driver's [5] finding that auditory beeps leading or lagging visual stimuli can readily bias visual apparent motion. In their study, the target modality (in which to-be-judged apparent-motion stimuli were presented) was vision, which is characterized by low temporal acuity. Our results show that apparent motion in the tactile modality, which has a high temporal resolution, can likewise be influenced by auditory timing. Both findings can be interpreted in terms of a "temporal-ventriloquism" effect [11], that is, the timing of target stimuli (in either the tactile or the visual modality) is systematically influenced by the timing of auditory beeps. In audiotactile streams, lagging odd-numbered beeps *pull* the timing of the corresponding taps closer to the subsequent, even-numbered taps, thus leading to dominant responses of "initial direction". Similarly, leading odd-numbered beeps *push* the timing of the corresponding taps away from the subsequent taps, giving rise to the opposite dominant motion percept of "reverse direction".

However, the temporal ventriloquism account cannot explain the results of the condition with the long audiotactile asynchrony (325-ms SOA, Experiment 2). If the timing of the asynchronous beep captured the timing of either the first or the second tactile tap, the auditory beep at the 325-ms SOA would still enhance the "initial-direction" percept, since the sound would *attract* the two taps (whether by acting on the first or the second tap) closer to each other. Similarly, based on the notion of (intramodal) auditory grouping, with both 75 and 325-ms SOAs, short intervals were paired with odd-numbered tactile intervals – so that one would also expect a dominance of "initial-direction" percepts, rather than the opposite. An alternative explanation, which assumes "bridging" two visual (i.e., by extension to the present scenario: tactile) events by an intervening auditory event [10], would predict similar results to the temporal ventriloquism or auditory-grouping accounts, namely, dominant apparent motion in the "initial direction", for both the 75- and 325-ms SOA conditions. However, (on all these accounts) unexpectedly, the results of Experiment 2 showed exactly the opposite effect: dominant apparent motion in the "reversed direction".

It is known that crossmodal integration takes place within a certain, limited temporal and spatial range [6,15,29–32]. On this background, in the condition with the audiotactile SOA of 325 ms, odd-numbered beeps were shifted close to the even-numbered taps, thus weakening the crossmodal grouping of the odd-numbered audiotactile stimuli (pair) and strengthening the crossmodal grouping of even-numbered stimuli (A1-T2-A2 in Figure 1C). Such asymmetric crossmodal grouping for even- and odd-numbered stimuli may cause an attention shift towards the salient taps (T2) (even though participants were told to disregard the sounds). This, in turn, would prime the following tactile events (T2-T1). This is consistent with previous studies of attentional modulations of apparent motion [33–35]. For example, in the study of the audiovisual or the tactile-visual line motion illusion [36], where a beep sound or an electric pulse (cue) is presented on either the left or the right side and this stimulus is accompanied or followed by a visual line presented in close proximity to the cue, the line is perceived to grow rapidly from the crossmodally stimulated side (this is referred to as the "line motion" effect). The crossmodal line motion effect has been attributed to a spatial-attentional bias induced by the auditory or tactile cue. In our case, strong crossmodal grouping on one side may similarly have served as a "cue" (even though the auditory beeps carried no spatial information), inducing one dominant motion direction.

In Experiment 3, we further examined the interaction between crossmodal grouping and crossmodal temporal interaction by removing the synchronous beeps. Although the audiotactile asynchrony was varied from −75 ms to 75 ms, as in Experiment 1, an overwhelming dominant direction of apparent motion – namely, from the audiotactile side to the tactile-only side – was found across all SOAs. That is, under these conditions, crossmodal temporal timing had no effect on tactile apparent motion. In previous studies of the temporal-ventriloquism effect using temporal-order judgments [11,37], the sensitivity of visual temporal order judgments increased only when two visual stimuli were paired with two auditory stimuli. Analogously to the present results, a single beep failed to produce a temporal-ventriloquism effect. In a more recent study with apparent motion [6], a null effect of single sounds in audiovisual apparent motion has also been reported. Previous accounts of the absence of a temporal ventriloquism effect with single sound configurations have attributed it a violation of the "assumption of unity" [7,8,11]. On this assumption, crossmodal integration makes sense only when the perceptual system has evidence that the two separate multisensory events (e.g., one auditory and one visual) originate from a common source [7]. Although this assumption could explain the null effect of crossmodal temporal modulation in the half-pairing (Experiment 3) and shifted-pairing (Experiment 2) conditions, it does not predict which direction of motion prevails in these conditions. One feasible account may be derived if assuming that a 'biased-competition' mechanism [38,39] is at work. The biased-competition framework assumes that when two (or more) neural assemblies compete with each other for representation, attentional biases in the system operate (over time) to make one assembly win the competition and suppress the competitor(s). Applied to the present paradigm, how an apparent-motion display is perceived depends on the relative balance of crossmodal grouping (the grouping of 'coincident' events in the nontarget and target modality) and crossmodal temporal capture (i.e., modulation of the timing of events in the target

modality by the timing of events in the nontarget modality) – two mechanisms that may be assumed to be in competition with each other, where spatial attention may exert a biasing influence on how the competition is resolved. In the half-pairing condition realized in Experiment 3, asymmetric audio-tactile grouping on the two sides of stimulus presentation (beep plus tap on one side vs. tap only on the other side) may generate a spatial-attentional bias towards the side of the crossmodal grouping. This would make the tactile stimulus on this side more salient and afford it "prior entry", thus giving rise to apparent tactile motion from the side of the audiotactile grouping to the other side. This is consistent with previous studies [33–36] that have shown attentional modulation of apparent motion to be of considerable strength, such as in the line motion illusion. By contrast, crossmodal temporal capture has been found to be a relatively weak effect [6,19,20]. Consequently, the latter temporal effect may be inhibited (or swamped) by the former spatial modulation.

In summary, examining tactile rivalry apparent motion dependent on different audiotactile configurations, we found a systematic influence of auditory timing on the motion percept in a full-pairing crossmodal condition. However, this temporal ventriloquism effect was abolished under conditions with half-pairing (unbalanced) and temporally shifted full-pairing configurations. Unimodal grouping based on auditory time interval or crossmodal temporal capture cannot readily explain the reversed pattern of audiotactile interaction with an audiotactile SOA of 325 ms. We propose an alternative account, namely, that unequal odd- and even-numbered audiotactile stimulus pairs leads to an attentional modulation of crossmodal grouping, which in turn prevents (or inhibits) crossmodal temporal integration. To test the hypothesis of a general attentional-saliency modulation of crossmodal temporal capture in the apparent-motion paradigm, it would be interesting to compare the present findings (tactile target modality) with conditions in which the target modality is reversed (auditory modality), that is, to examine the influence of touch modulations on auditory apparent motion rivalry.

Materials and Methods

Participants

Eleven paid participants participated in Experiment 1 (6 females, average age 26.6), Experiment 2 (7 females, average age 26.7), and Experiment 3 (7 females, average age 25.5). None of the participants reported any history of somatosensory disorders. They were all naïve as to the purpose of the study and were paid after the experiment. The study was approved by the Ethics Committee, Faculty of Psychology and Education, Ludwig-Maximilian University. All experiments were conducted in accordance with the guidelines of Ethical Principles of Psychologists. Written informed consent was obtained from each participant before experiments.

Apparatus and stimuli

A customized tactile stimulus generator (Heijo Research Electronics, UK) was connected to a HP PC (AMD Athlon 64 Dual-Core processor) via the LPT port. The two solenoid actuators, which were embedded in a sponge with a fix.6(a)-32(ofui)1-dw-.1(5)16.dwthe taspondr-

Journal of Experimental Psychology: Human Perception and Performance 30: 330–345.
5. Freeman E, Driver J (2008) Direction of Visual apparent motion driven solely by timing of a static sound. Current biology 18: 1262–1266.
6. Shi Z, Chen L, Müller HJ (2010) Auditory temporal modulation of the visual Ternus effect: the influence of time interval. Experimental Brain Research 203: 723–735.
7. Welch RB (1999) Meaning, attention, and the "unity assumption" in the intersensory bias of spatial and temporal perceptions. In: Aschersleben G, Bachmann T, Müsseler J, eds. Cognitive contributions to the perception of spatial and temporal events. Elsevier: Amsterdam. pp 371–387.
8. Welch RB, Warren DH (1980) Immediate Perceptual Response to Intersensory Discrepancy. Psychological Bulletin 88: 638–667.
9. Bertelson P, Aschersleben G (2003) Temporal ventriloquism: crossmodal interaction on the time dimension: 1. Evidence from auditory visual temporal order judgment. International Journal of Psychophysiology 50: 147–155.
10. Getzmann S (2007) The effect of brief auditory stimuli on visual apparent motion. Perception 36: 1089–1103.
11. Morein-Zamir S, Soto-Faraco S, Kingstone A (2003) Auditory capture of vision: examining temporal ventriloquism. Cognitive Brain Research 17: 154–163.
12. Vroomen J, de Gelder B (2004) Perceptual Effects of Cross-modal Stimulation: Ventriloquism and the Freezing Phenomenon. In Handbook of mutlisensory processes, MIT Press. pp 141–150.
13. Koffka K (1935) Principles of Gestalt psychology. New York: Harcourt Brace.
14. Köhler W (1930) Gestalt psychology. London: Bell.
15. Spence C, Sanabria D, Soto-Faraco S (2007) Intersensory Gestalten and crossmodal scene perception. In The Psychology of Beauty and Kansei: New Horizons of Gestalt Perception. Ed. K Noguchi (Tokyo: Nihon University College of Humanities and Sciences).
16. O'Leary A, Rhodes G (1984) Cross-modal effects on visual and auditory object perception. Perception & Psychophysics 35: 565–569.
17. Recanzone GH (2009) Interactions of auditory and visual stimuli in space and time. Hearing Research 258: 89–99.
18. Vroomen J, de Gelder B (2000) Sound enhances visual perception: cross-modal effects of auditory organization on vision. Journal of Experimental Psychology: Human Perception and Performance 26: 1583–1590.
19. Keetels M, Stekelenburg J, Vroomen J (2007) Auditory grouping occurs prior to intersensory pairing: evidence from temporal ventriloquism. Experimental Brain Research 180: 449–456.
20. Bruns P, Getzmann S (2008) Audiovisual influences on the perception of visual apparent motion: exploring the effect of a single sound. Acta Psychologica 129: 273–283.
21. Von Bekesy G (1959) Similarities between hearing and skin sensations. Psychological Review 66: 1–22.
22. Hirsh IJ, Sherrick CEJ (1961) Perceived order in different sense modalities. Journal of Experimental Psychology 62: 423–432.
23. Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415: 429–433.
24. Bresciani JP, Ernst MO (2007) Signal reliability modulates auditory-tactile integration for event counting. NeuroReport 18: 1157–1161.
25. Mamassian P, Goutcher R. Temporal dynamics in bistable perception. Journal of Vision 5: 361–375.
26. Pastukhov A, Braun J (2007) Perceptual reversals need no prompting by attention. Journal of Vision 7: 1–17.
27. Zampini M, Brown T, Shore DI, Maravita A, Röder B (2005) Audiotactile temporal order judgments. Acta Psychologica 118: 277–291.
28. Vroomen J, Keetels M (2010) Perception of intersensory synchrony: A tutorial review, Attention, Perception, & Psychophysics. 72: 871–884.
29. Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. Current Biology 14: 257–262.
30. Bresciani JP, Dammeier F, Ernst MO (2006) Vision and touch are automatically integrated for the perception of sequences of events. Journal of Vision 6: 554–564.
31. Gepshtein S, Burge J, Ernst MO, Banks MS (2005) The combination of vision and touch depends on spatial proximity. Journal of Vision 5: 1013–1023.
32. Slutsky DA, Recanzone GH (2001) Temporal and spatial dependency of the ventriloquism effect. NeuroReport 12: 7–10.
33. Berman RA, Colby CL (2002) Auditory and visual attention modulate motion processing in area MT+. Cognitive Brain Research 14: 64–74.
34. Cavanagh P (1992) Attention-based motion perception. Science 257: 1563–1565.
35. Lu ZL, Sperling G (1995) Attention-generated apparent motion. Nature 377: 237–239.
36. Shimojo S, Miyauchi S, Hikosaka O (1997) Visual motion sensation yielded by non-visually driven attention. Vision Research 37: 1575–1580.
37. Scheier C, Nijhawan R, Shimojo S (1999) Sound alters visual temporal resolution. Investigative Ophthalmology &Visual Science 40: S792.