

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 *intramodal* auditory grouping by using evenly alternating high- (H) and low-pitch (L) beeps (i.e., HLLHH...). 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 gap between two light flashes or a short sound intervening between two flashes enhanced reports of continuous visual motion, while there was no such enhancement when the sound was part of a tone sequence that allowed for intramodal (auditory) grouping *prior to* the multisensory integration of the audiovisual stimuli. Bruns and Getzmann argued that auditory events that intervene between two flashes induce the impression of a single, multimodal moving object. In a more recent study, Shi et al. [6] used visual Ternus apparent motion coupled with auditory events. In Ternus apparent motion, participants are presented with a sequence of visual frames each consisting of two horizontally arranged dots that are shifted forth and back by the inter-dot distance in successive frames. Depending on the inter-frame interval, this stimulus gives rise to two alternative motion percepts: either ‘group motion’, where both dots are seen to be moving (long intervals), or ‘element motion’, where only the ‘outer’ dot is seen to be moving while the ‘inner’ dot appears stationary (short intervals). Using this paradigm, Shi et al. demonstrated that merely presenting a single sound near the first or the second visual frame did not give rise to a crossmodal temporal-ventriloquism effect; more technically, single sounds had little effect on the transition threshold between element and group motion percepts. By contrast, crossmodal temporal integration was evident with fully paired audiovisual stimuli, that is, when a sound event occurred closely in time with each visual frame.

It is important to note that the perceptual groupings implicated in the above studies fall in the categories of either *unimodal* grouping (e.g., auditory grouping based on common pitch or temporal proximity) or *crossmodal* (audiovisual) grouping. Both types of perceptual grouping may influence the effects examined in the above studies. Moreover, to date, the modulatory influence of perceptual uni- and, respectively, crossmodal grouping on crossmodal temporal integration has never been systematically compared within one study. On this background, the present study, employing a directionally ambiguous tactile apparent-motion stream with different embedded auditory events, was designed to explore how perceptual grouping influences crossmodal temporal capture (temporal-ventriloquism effect).

Our motive for using the audiotactile modalities is twofold. First, we aimed to examine the crossmodal temporal interaction between two modalities with similarly high temporal acuity (i.e., the auditory and tactile modalities) [21,22]; thus, the present study was expected to extend upon previous conclusions largely based on the use of paradigms with asymmetric temporal sensitivities, and to augment reliability-based theories of multisensory integration

[23,24]. Second, crossmodal temporal integration has, as yet, not been examined systematically with tactile apparent motion (especially movement over an extended, 90-second period of time); thus, the present study was meant to enhance our understanding of crossmodal temporal integration related to the tactile modality.

In our paradigm, participants placed the tips of their left and right middle fingers on the surface of two tactile actuators (one on the left and one on the right side), while wearing headphones. The two tactile actuators produced alternating taps at a rate of 2.5 Hz for 90 seconds; concurrently, a train of mono-beeps was paired with the stream of tactile taps (for details, see Methods and Figure 1). After an initial presentation for 4 seconds, participants started to hold one foot pedal (the left or the right one) pressed to indicate their perceived direction of tactile motion; they were instructed to switch to the other foot pedal as soon as they perceived the motion direction to be reversed. In this way, it was possible to measure the (phase) durations of apparent motion in one or the other direction.

In order to examine the influence of uni- and crossmodal grouping on crossmodal temporal integration, we varied the auditory-auditory interval and the audiotactile interval separately. In more detail, to modulate unimodal (intra-auditory) grouping (see dashed ellipse in Figure 1B), we presented either interleaved short and long auditory intervals or equal auditory intervals within the stream of audiotactile stimuli. And to modulate crossmodal grouping (see dashed rectangle in Figure 1B), we varied the audiotactile pairing, along with the audiotactile stimulus onset asynchronies (SOAs).

Experiment 1 was designed to establish crossmodal (audiotactile) temporal integration in tactile apparent motion. Analogously to the paradigm of Freeman and Driver [5], we introduced configurations of full (i.e., one-to-one) pairing audio-tactile stimuli: each tactile tap paired with one beep, where even-numbered beeps were always synchronous with the onsets of the tactile taps on one side and odd-numbered beeps were asynchronous, by a given SOA (−75, −50, −25, 0, 25, 50, 75 ms), with the onsets of the tactile taps on the other side (see Figure 1B). The results revealed a crossmodal (auditory-on-tactile) temporal-capture effect similar to the auditory-on-visual effect reported by Freeman and Driver.

In Experiment 2, we went on to examine the influence of crossmodal grouping on the crossmodal temporal interaction established in Experiment 1, by comparing the influence of an audiotactile SOA of 75 ms (Figure 1B; full-pairing event configuration) with that of 325 ms (Figure 1C; shifted full-pairing configuration). In both conditions, the shorter of the two auditory intervals (between A1 and A2) is pairing the odd-numbered interval between tactile taps (T1-T2, see Figure 1B and 1C). Given this, one would expect the influence of *unimodal* auditory grouping (between A1 and A2) on tactile apparent motion to work in the same direction in both audiotactile SOA conditions (depicted in Figures 1B and 1C, respectively). However, with the audiotactile SOA of 325 ms, *crossmodal* grouping between auditory and tactile events would take place asymmetrically around even-numbered (T2) taps, compared to the more balanced grouping around odd-numbered and even-numbered taps in the 75-ms SOA condition. Thus, if *crossmodal* grouping influenced the temporal capture effect, one would expect differential modulations of tactile apparent motion between the two conditions (as a baseline, a synchronous audiotactile condition, with an SOA of 0 ms, was also included in Experiment 2). The results revealed the direction of the temporal-capture effect to be reversed with the extended audiotactile SOA of 325 ms, compared to the 75-ms SOA, suggestive of an influence of *crossmodal* grouping.

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-pairing’). 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-pairing 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

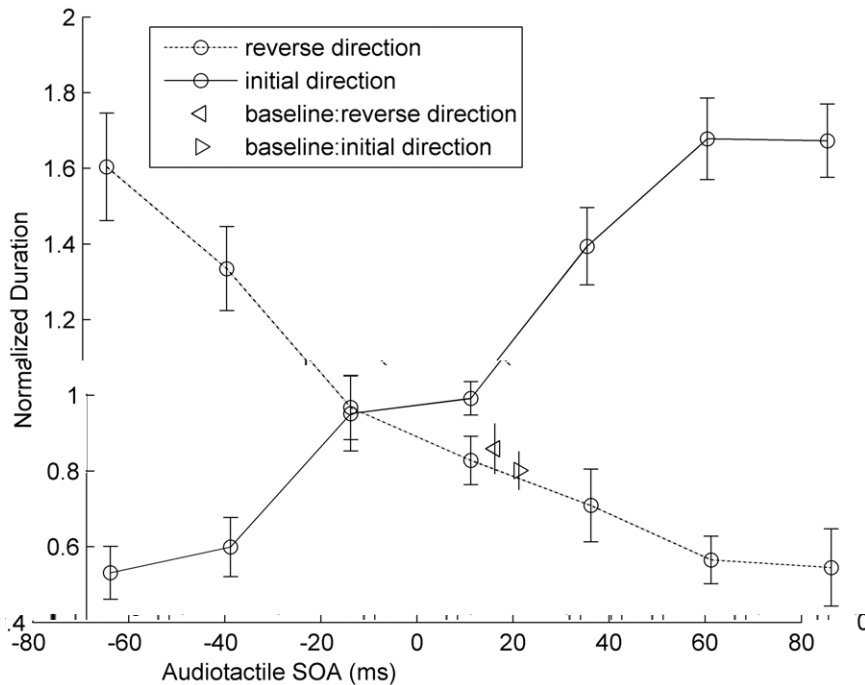


Figure 2. Normalized phase durations of tactile apparent motion in Experiment 1.

significantly from each other, $t(10) = -1.322$, $p = 0.216$, indicating that the initial bias had dissipated after four seconds of stimulus presentation. However, there remained a marginal initial bias after four seconds for tactile apparent motion in the synchronous audiotactile stream (SOA = 0 ms), $t(10) = 2.179$, $p = 0.054$. For the conditions with sounds present, Figure 2 shows a clear audiotactile interaction in the perceived tactile motion across the different audiotactile SOAs. We selected the phase durations of “initial-direction” responses for further analysis of the auditory capture effect (the results would be analogous for the “reverse direction”). A repeated-measures ANOVA revealed a significant main effect of auditory timing, $F(6,60) = 28.534$, $p < 0.001$, and a linear contrast test showed that the phase duration increased linearly with increasing audiotactile SOA, $F(1,6) = 167.289$, $p < 0.001$. This indicates that asynchronous auditory-tactile timing did indeed influence tactile apparent motion, with the influence being systematic and bidirectional. For example, an audiotactile SOA of 50 ms (when the odd numbered beeps lagged the corresponding taps by 50 ms) produced a dominant percept of “initial direction”, while an SOA of -50 ms gave rise to a dominant percept of “reverse direction”. Note that the opposite trends with respect to “initial direction” and “reverse direction” crossed at the audiotactile SOA of -25 ms (rather than the SOA of 0 ms). This slight asymmetry may be attributable to a shift in audiotactile simultaneity resulting from temporal recalibration and adaptation in the extended (and repeated) audiotactile stream [27,28], or the small difference between the auditory and tactile stimulus durations used in the experiment. However, the general trends are consistent with Freeman and Driver’s [5] ‘audiovisual’ study, where auditory timing was found to influence visual apparent motion in a similar way.

Figure 3 presents the mean phase durations for “initial-direction” and “reverse-direction” responses as a function of the (variable) audiotactile SOA. A repeated-measures ANOVA for the “initial-direction” responses revealed the main effect of audiotactile SOA to be significant, $F(2,20) = 11.66$, $p < 0.01$ ($F(2,20) = 7.215$, $p < 0.01$, for the “reverse direction”). Bonferroni-corrected pairwise comparisons showed that for both “initial-direction” and “reverse-direction” responses, the mean phase durations differed significantly between the 75-ms and the 325-ms SOA, $p_s < 0.05$. With an audiotactile SOA of 75 ms, the response pattern was similar to that in Experiment 1, that is, characterized by dominance of “initial direction”. However, the dominant motion direction was changed to “reverse direction” when the audiotactile SOA was increased to 325 ms. The differential dominance patterns of tactile apparent motion between these two conditions is the most interesting finding of Experiment 2, which demonstrates that crossmodal grouping can strongly influence the crossmodal temporal integration.

Experiment 3 was similar to Experiment 1, except that the “asynchronous” beeps were omitted (they were presented in Experiment 1). The mean normalized phase durations are shown in Figure 4.

A pairwise t -test comparing the two perceived directions in the baseline condition (without beeps) revealed no difference, $t(10) = 0.286$, $p = 0.781$. A repeated-measures ANOVA of the phase durations for “initial-direction” responses, with the single

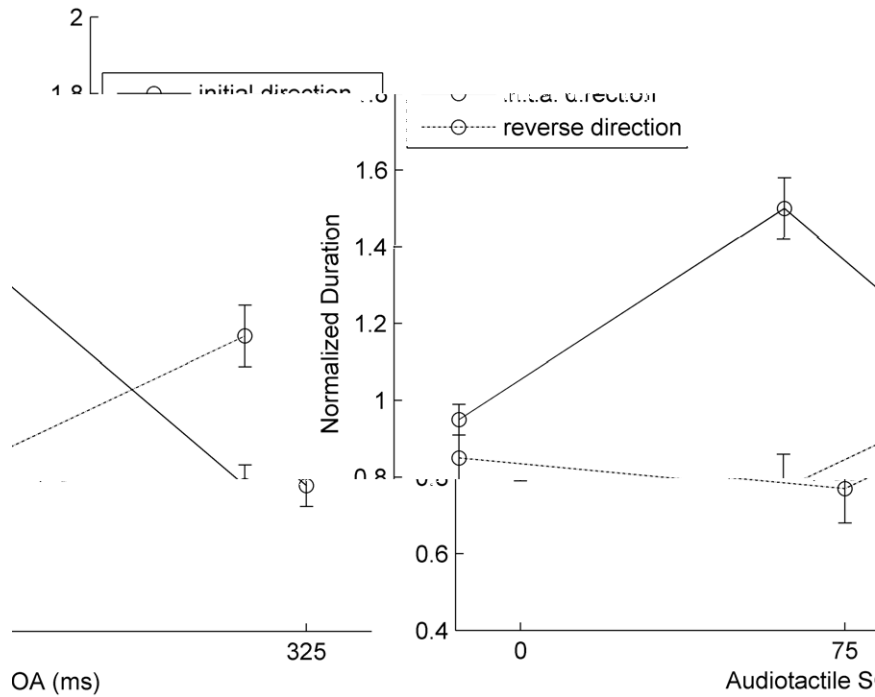


Figure 3. Normalized phase durations of tactile apparent motion in Experiment 2.

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

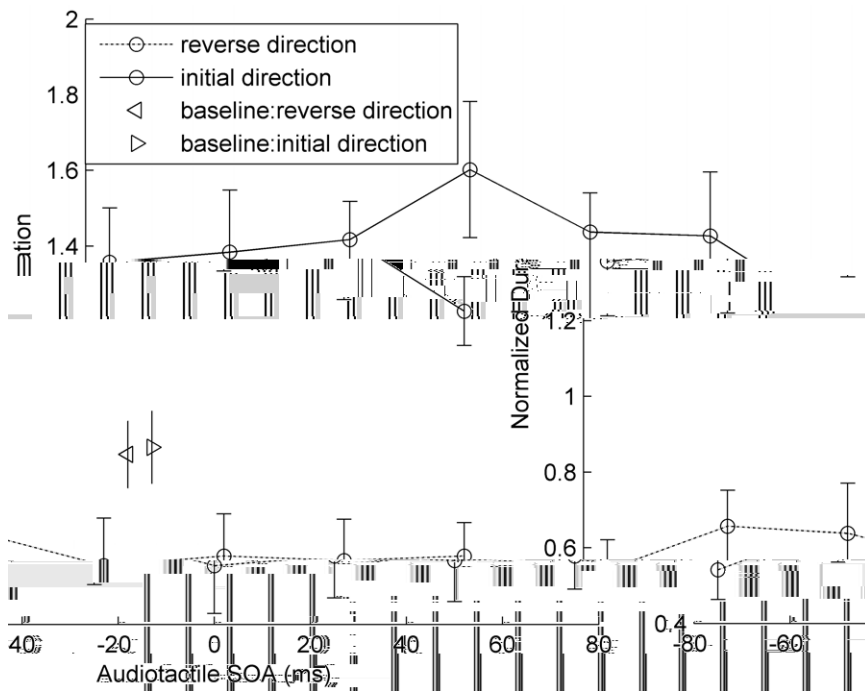


Figure 4. Normalized phase durations of tactile apparent motion in Experiment 3.

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

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

t t

Eleven paid participants participated in Experiment 1]TJ0

- 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 multisensory 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.
 38. Desimone R (1998) Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society B* 353: 1245–1255.
 39. Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG (2010) The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences* 14: 400–410.
 40. Brainard DH (1997) The Psychophysics Toolbox. *Spatial Vision* 10: 433–436.

