

The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity

Vanessa Harrar · Laurence R. Harris

Received: 26 July 2006 / Accepted: 8 December 2007 / Published online: 9 January 2008
© Springer-Verlag 2007

Abstract Information about an event takes different amounts of time to be processed depending on which sensory system the event activates. However, despite the variations in processing time for lights and sounds, the point of subjective simultaneity (PSS) for briefly presented audio/visual stimuli is usually close to true simultaneity. Here we confirm that the simultaneity constancy mechanism that achieves this for audio/visual stimulus pairs is adaptable, and extend the investigation to other multimodal combinations. We measured the PSS and just noticeable differences (JNDs) for temporal order judgements for three stimulus combinations (sound/light, sound/touch, and light/touch) before and after repeated exposure to each one of these pairs presented with a 100 ms asynchrony (i.e., nine adaptation combinations). Only the perception of simultaneity of the sound/light pair was affected by our exposure regime: the PSS shifted after exposure to either a temporally staggered sound/light or light/touch pair, and the JND decreased following exposure to a sound/touch pair. No changes were found in the PSSs or JNDs of sound/touch or light/touch pairs following exposure to any of the three time-staggered combinations. Participants' reaction times (RT) to the three stimuli were also tested before and after each adaptation exposure. In general, exposure did not affect attention or processing time; the only change in RTs (of the 9 tested) was an increased RT for light following exposure to a sound/light pair with light leading. We suggest that the neural correlates of multisensory sound/light processing are resynchronised by a separate, more flexible

simultaneity constancy mechanism than the light/touch or the sound/touch simultaneity processing systems.

Keywords Adaptation · Simultaneous · Touch · Vision · Audition · Temporal order judgements · Point of subjective simultaneity · Multisensory processing

Introduction

An asynchrony can arise in sensory signals associated with synchronous stimuli for a variety of reasons including the time it takes for the energy to reach the end organs, different transduction latencies (King and Palmer 1985; Poppel et al. 1990), intensity effects (Wilson and Anstis 1969), and attention effects (Spence et al. 2001), variations in latency with visual eccentricity (Nickalls 1996), and different nerve lengths (Von Békésy 1963; Harrar and Harris 2005). Such differences provide a challenge for the brain if it is to perceive the temporal relationship between multiple stimuli correctly. For example, touching a key on a keyboard activates a multisensory event. The click of the key, the sight of the finger touching the surface, and the feel of contact on the skin, all generate information that reaches the brain at different times, yet these stimuli are all associated with a single event at a single moment in time.

The accurate perception of simultaneity in the face of these variations is called “simultaneity constancy” (Kopinska and Harris 2004). Simultaneity constancy is a type of perceptual constancy that maintains a stable percept of simultaneity despite variations in the timing of the internal representation of individual stimuli.

How is the brain to know the relative time of the different components of a multimodal stimulus and achieve simultaneity constancy in the face of the different processing times?

V. Harrar (✉) · L. R. Harris
Department of Psychology, Centre for Vision Research,
York University, Toronto, ON, Canada M3J 1P3
e-mail: vharrar@yorku.ca

The true correspondence between different sensory signals cannot be deduced from the temporal lag of the neural signals alone. Deducing the correspondence requires additional information not present in the stimulus, such as previous experiences, sometimes thought of as forming Bayesian priors (Miyazaki et al. 2006). A strategy that the brain might adopt is to assume that sensory stimuli that repeatedly occur in a fixed temporal relationship in fact originate from a multimodal event and should be regarded as simultaneous. Indeed, the relative timing of a sound and light that is perceived as simultaneous (the point of subjective simultaneity, PSS) can be shifted by even just a few minutes of exposure to a repeated time-staggered sound/light pair (Vroomen et al. 2004; Fujisaki et al. 2004) or to videos of people talking with the sound and video tracks out of synch (Vatakis et al. 2007). This suggests that the some statistical correspondence between the old PSS and the repeatedly presented delay is treated as the new simultaneity.

Alternatively, the perception of simultaneity might not shift to a new value, but might instead (or in addition) accept a broader range of delays as simultaneous. Such an increase in range would make it more likely that the repeatedly presented delay would be perceived as simultaneous. Indeed, repeated exposure to asynchronous audio/visual or audio/tactile stimulus pairs has also been found to increase the “window of acceptance” within which stimuli are perceived as simultaneous (Navarra et al. 2005, 2007). Increasing the acceptance window means that larger variations in the actual timing of stimuli are still regarded as simultaneous. That is, the ability to discriminate pairs of stimuli based on their temporal properties becomes worse: the just noticeable difference (JND) increases.

With the exception of Harrar and Harris (2005), previous recalibration experiments have only looked at a single multimodal combination at a time (Fujisaki et al. 2004—audio/visual; Vroomen et al. 2004—audio/visual; Navarra et al. 2005—audio/visual; Navarra et al. 2007—audio/tactile). This makes it hard to compare the effectiveness of the brain’s recalibration ability in response to different combinations of modalities because of inevitable differences in the experimental setups between laboratories. Adaptive changes of the audio/visual simultaneity constancy mechanism have previously been shown to generalise to other audio/visual integration conditions (Fujisaki et al. 2004; Navarra et al. 2005) but generalisation to other stimulus combinations is unknown. Further, no one has yet tested the effect of repeated exposure to time-staggered visuo/tactile pairs. Experiment 1 includes an attempt at visuo/tactile recalibration as part of comprehensively trying to recalibrate all combinations of auditory, visual, and tactile simultaneity timings using the same methodology for all combinations.

As a control, experiment 2 tests the effect of repeated exposure on attention. Responses to attended stimuli are

faster than responses to non-attended stimuli. This is known as the prior entry effect (Spence et al. 2001). The exposure sequence may itself cause shifts in attention, which would cause the processing time of one of the stimuli in the pair to be speeded up or slowed down relative to the other. This could provide an alternative explanation to “recalibration” for any PSS shifts found. We therefore measured reaction times to each of the three individual stimuli (lights, sounds, and touches) before and after exposure to the time-staggered pairs. If reaction times to the individual stimuli do not change and the PSSs or JNDs do, then attention shifts are unlikely to provide the explanation.

Spatial proximity affects the perception of simultaneity—stimuli that appear close together in space are more likely to be interpreted as simultaneous (Zampini et al. 2005; Bertelson and Aschersleben 2003; Calvert et al. 2004). Spatial proximity might therefore encourage a staggered pair of multimodal stimuli to be interpreted as simultaneous (Spence et al. 2003; Driver and Spence 2000; Wallace et al. 2004). We therefore constructed a “multisensory cube” (Fig. 1) in which our light, sound, and touch stimuli were presented in close proximity.

General methods

Participants

Participants were volunteers from the graduate and undergraduate pool at York University. The mean age of participants was 23 and they were all right-handed. Some of them were paid. Experiments were approved by the York Ethics Board. Experiment 1 had eight participants (four females), and experiment 2 had nine participants (three females).

Touch stimulation

The touch stimulator was made from a small solenoid mounted in a wooden cup. When the solenoid was powered, a central pin was pushed out.¹ It extended about 1 mm from the edge of the cup and hit the skin surface with a force of a light tap spread over a surface area of about 1 mm². Solenoids were controlled by appropriately amplified 5-volt signals from a CED1401 (Cambridge Electronic Design, UK) interface box controlled by a PC. The duration of the touch was always 50 ms.

¹ A carefully positioned photocell found that it took 5 ms to either extend or retract the pin in the solenoid. This delay was not adjusted for in the reaction time or PSS results that follow since the delay was constant in all conditions; the results below look for differences pre- and post-exposure.

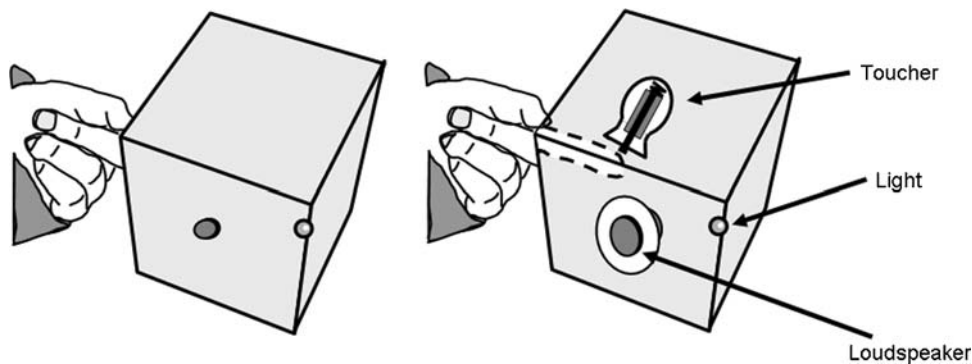


Fig. 1 The Multisensory Stimulus Box. This box contained a touch solenoid, an LED, and a speaker. All stimuli were computer controlled (see general methods). The participants inserted their left index finger into the box through a slit in the foam panel in order to reach the tactile stimulus located inside the box (not visible to the participant). The

foam was stiff enough to hold the finger in place throughout the experimental session. On the *left*, the box is seen as the subject saw it during the course of the experiment. On the *right*, we can see the solenoid and the participant's finger inside the box

Visual stimulation

Visual stimulation was provided by a green LED that was controlled directly by 5-volt signals from the CED1401 interface box controlled by a PC. The luminance was 55 cd/m² measured by Minolta™ luminance meter LS-100, and the duration of illumination was always 50 ms. The room lights were left on throughout all experiments.

Auditory stimulation

Auditory stimulation was provided by a loudspeaker playing a tone burst of 2,000 Hz with a duration of 50 ms at a volume of 73 dB. The waveform was generated within the CED1401 interface box controlled by a PC.

Equipment

The three stimuli (LED, solenoid, and speaker) were mounted in a 125 cm³ box as shown in Fig. 1. The box was made of and insulated with sound-attenuating materials (cardboard, bubble wrap and foam). Participants inserted their left index finger through a slit in the foam panel and placed it on the tactile stimulator located inside the box. The stimulus sound was loud enough to be heard over white noise—played through a separate set of desk speakers to mask any additional sounds caused by the solenoid.

Experiment 1: The generalizability of temporal recalibration across modalities

In this experiment we measured temporal order judgements (TOJ) of all combinations of touch, light, and sound stimuli after exposure to an asynchronous pair of stimuli staggered by a fixed time delay of 100 ms (which Vroomen et al.

2004) showed is the delay beyond which audio/visual adaptation effects start to level off). Exposure to such a delay has been shown in some situations to shift the PSSs from the initial interpretation of simultaneity towards the newly defined staggered “zero”. In other situations, exposure has been found to increase the JND. The JND is a standardised measure of the precision with which participants could determine the temporal order of the stimuli. If the temporal window for simultaneity changes, as a result of this exposure, then there should be a change in the JND. If, however, the window stays the same width but the perception of simultaneity is recalibrated to a new delay value, then this would appear only as a change in the PSS. We calculated both the PSS and the JND of TOJs in order to find recalibration effects and/or changes in the temporal window.

By testing all three stimulus combinations after exposure to each we were able to further test the prediction that adaptation might generalise to other pairs of stimuli even if only one of the stimuli in the test pair was present in the exposure phase. Finally, we compared the effects between the different pairs to see if the three stimulus combinations were equally sensitive to repeated exposures of delayed stimuli.

Methods for experiment 1

Procedure

Stimuli were presented using the stimulus cube described above. Participants made TOJs to bimodal stimulus pairs (“which stimulus came first”) before and after an exposure phase. TOJs were used in order to have a force choice method as opposed to simultaneity judgements (where subjects determine if the stimuli were simultaneous or not), which are fraught with individuals’ criterion bias. Participants first

had a practice session to familiarise themselves with the response keys so that they would not need to look at them during the experiment. The inter-stimulus delay used for the practice session was 300 ms to make it easy to tell which came first. Participants needed to correctly respond to all stimulus combinations to continue onto the experiment.

Exposure phase

Participants were exposed to a 5 min sequence of one of the following three conditions: sound/light pair with the light leading by 100 ms, sound/touch pair with the touch leading by 100 ms, or a light/touch pair with the touch leading by 100 ms. The pairs of stimuli were presented with a random inter-pair interval varying from 200 to 1,000 ms. Participants were instructed to pay attention to the pair by either counting them or by trying to decide their temporal order. During the TOJ trials that followed the exposure, “top-up” exposure intervals of 10 s duration were inserted after every eight judgements.

Testing phase

There were four testing blocks: one before exposure and three after. In each testing block TOJs were measured for eleven stimulus onset asynchronies (SOAs: -200 , -150 , -100 , -50 , 0 , 50 , 100 , 150 , 200 ms). Each SOA was presented ten times. All stimulus combinations were randomly interleaved. Since there were three pairs, eleven SOAs and ten presentations of each, there were a total of 330 judgements to be made in each testing sequence. The pre-exposure TOJ took about 20 min to run. Post-exposure TOJs took about 30 min because of the inserted top-up exposure periods. All participants completed the four blocks, no exposure, exposure to sound/light, exposure to sound/touch, and exposure to light/touch, in separate sessions run in a partially counterbalanced order.

Responses

Responses were made with the first three fingers of the right hand as follows: participants rested their right hand over three buttons on a keyboard. Each key corresponded to a particular stimulus (light, sound, or touch in alphabetical order) and participants were instructed to press the key that corresponded to the stimulus that appeared to be presented first.

Data analysis

All responses were recorded for further analysis, even when participants chose a stimulus that was not present (for example if a light/sound pair was presented and a participant chose touch as the stimulus that had occurred first).

Response frequencies were calculated as a percentage of the number of trials minus these false positives.

The percentage of trials on which a particular stimulus was chosen first was plotted as a function of the SOA. Using SigmaPlot® 9.0 a two parameter cumulative Gaussian $F(x) = 100/(1 + \exp(-(x - x_0)/b))$ was fitted to the data. The inflection point (x_0) was taken as the PSS. The standard deviation (b) of the cumulative Gaussian was taken as the JND i.e., at 84.13%.

Values of PSSs and JNDs were compared before and after adaptation using t -tests. No adjustments were made for the use of multiple t -tests since each t -test measured a different prediction. T -tests did not overlap: each test was for a particular pre- and post-condition and thus tested specific a-priori predictions.

Results of experiment 1

Effect on PSS

After exposure to a staggered sound/light pair in which the light led by 100 ms, there was a significant shift of the sound/light psychometric curve (paired samples t -test: $t_7 = -3.14$, $P = 0.016$ see top left of Fig. 2). The PSS was shifted by +32 ms (see top left of Table 1A), such that light needed to be on 32 ms earlier than it did before the exposure phase (i.e., the light needed to be on 9 ms before the sound) in order for the stimuli to be perceived as simultaneous. This shift in PSS was in the direction of the exposure: participants adapted to “light first” and the PSS shifted towards “light first”.

After exposure to a staggered light/touch pair in which the touch led by 100 ms, there was a significant difference in the PSS of the sound/light psychometric function (paired samples t -test: $t_7 = -2.91$, $P = 0.023$ see top right of Fig. 2). The shift was in the positive direction (+30 ms, see top right of Table 1A) such that light needed to be presented 30 ms earlier than it did before exposure to the staggered light/touch pair.

The PSS of all the other pairs of stimuli tested were unaffected by exposure to any temporally staggered pairs as measured by paired samples t -tests (see Table 1A for t - and P -values).

Effect on JND

After exposure to a staggered sound/touch pair (touch presented 100 ms before sound) the JND of the sound/light pair was significantly decreased from 76 to 56 ms (paired samples t -test: $t_7 = -3.04$, $P = 0.019$) (see top middle of Fig. 2 and Table 1B). The JND of the other eight adapt-test combinations were unaffected by exposure to time-staggered pairs, as measured by t -tests (see Table 1B for t - and P -values).

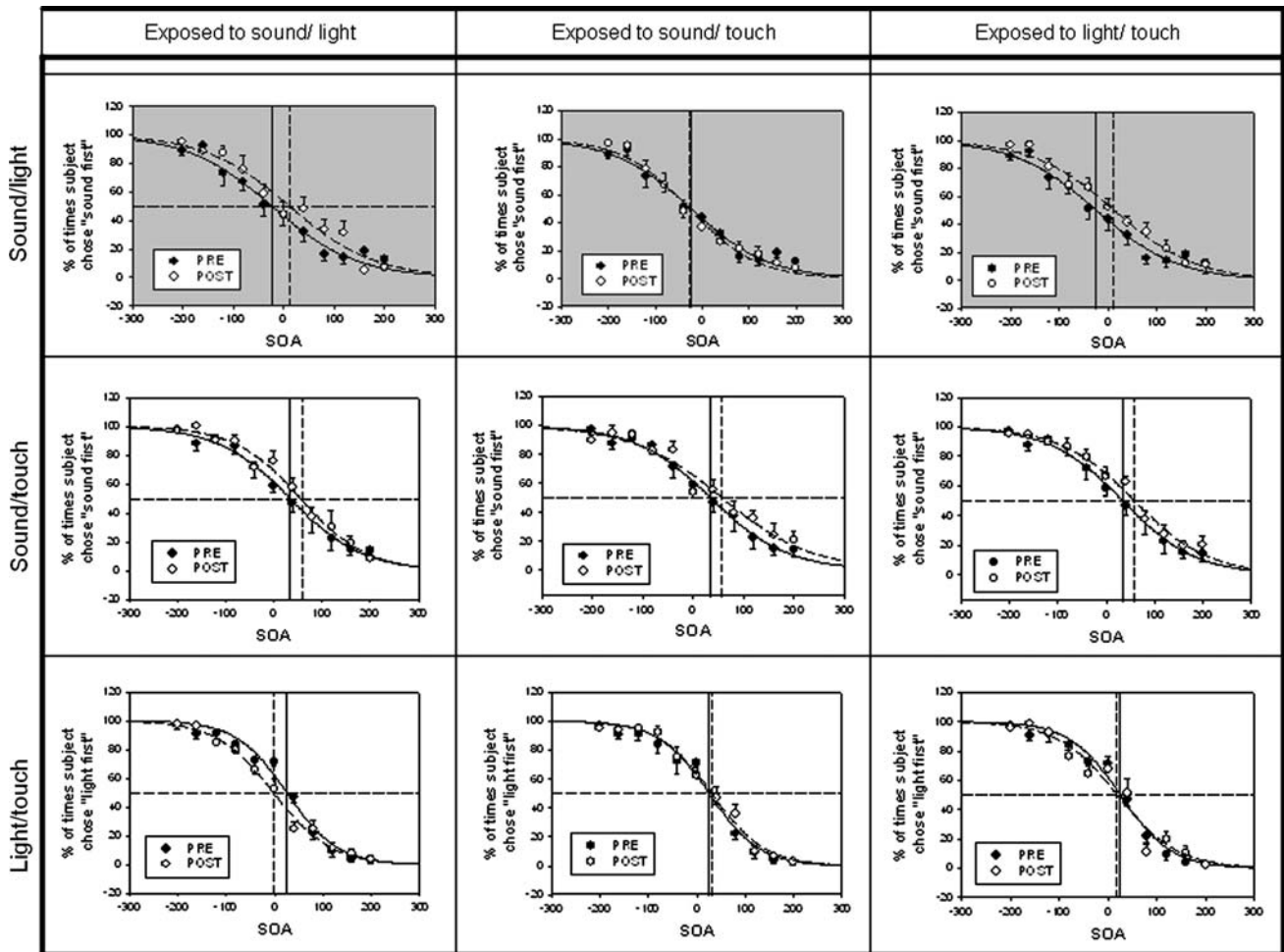


Fig. 2 Average cumulative Gaussian curves before and after exposure to time-staggered bimodal stimulus pairs. The curves are best fits through the means of the percentage of times one stimulus was perceived to be first plotted as a function of the stimulus onset asynchrony (SOA). The *solid curve, solid vertical line and filled black symbols* represent responses pre-exposure. The *dotted curve, dotted vertical line, and open symbols* are post-exposure. The three columns are arranged according to the pair used in the exposure phase: sound/light (light leading by 100 ms), sound/touch (touch leading by 100 ms), and light/

touch (touch leading by 100 ms), respectively. The rows are arranged according to the pairs of stimuli tested: sound/light (positive means “light first”), sound/touch (positive means “touch first”), light/touch (positive means “touch first”). The three *shaded graphs* are the combinations that have significant shifts (exposure to time-staggered sound/light or light/touch pairs shifting the PSS of a sound/light judgements towards “light first”, and exposure to time-staggered sound/touch decreasing the JND of sound/light judgements)

Discussion

This experiment confirmed the adaptability of the audio/visual temporal perception system. However exposure to time-staggered sound/touch and light/touch pairs did not alter the PSS for those stimulus combinations. It seems that the audio/visual system is different from the system that makes audio/tactile and visuo/tactile comparisons. These data were collected after only 5 min of exposure, and using infrequent top-ups (after every eight trials), which underscores the remarkable flexibility of the audio/visual system. Audio/tactile and visuo/tactile PSSs appears to be less flexible. This difference between audio/visual comparisons and those involving touch is also suggested in Navarra et al.

(2007) which shows audio/visual but not the audio/tactile PSS shifts.

Although the PSS of judgements involving touch seem to be relatively fixed, the system is not entirely rigid. Navarra et al. (2007) found a significant increase in the JND of a sound/touch pair after repeated exposure to a 75 ms asynchronous delay between a sound and a touch with *sound* leading. We did not replicate this increased JND of the sound/touch pair. However, our methods were quite different from those of Navarra et al. (2007). Specifically, we exposed subjects to a 100 ms asynchrony with *touch* leading. The audio/tactile simultaneity constancy mechanism may only be adaptable to exposure in one direction. Also, the exposure sequence was very different: Navarra et al.

Table 1 Changes in PSSs, JNDs, and RTs following exposure to temporally staggered pairs

	Exposed to											
	Sound/light				Sound/touch				Light/touch			
	Mean	SE	<i>t</i> (<i>df</i> = 7)	<i>P</i>	Mean	SE	<i>t</i> (<i>df</i> = 7)	<i>P</i>	Mean	SE	<i>t</i> (<i>df</i> = 7)	<i>P</i>
A. PSS difference (post–pre)												
Sound/light	<i>31.82</i>	<i>10.80</i>	<i>–3.14</i>	<i>0.02</i>	15.91	23.10	–0.74	0.49	<i>30.34</i>	<i>11.20</i>	<i>–2.91</i>	<i>0.02</i>
Sound/touch	18.74	22.00	0.92	0.39	–17.16	24.30	0.76	0.47	15.24	9.10	–1.78	0.12
Light/touch	–4.52	17.00	0.28	0.78	6.63	12.20	–0.58	0.58	–4.86	14.70	0.35	0.73
B. JND difference (post–pre)												
Sound/light	–1.31	10.50	–0.13	0.90	<i>–19.41</i>	<i>6.80</i>	<i>–3.04</i>	<i>0.02</i>	7.91	18.00	0.47	0.65
Sound/touch	–7.66	10.90	0.61	0.56	4.61	14.30	0.35	0.74	0.66	10.90	0.06	0.95
Light/touch	1.56	13.80	0.12	0.91	–4.67	9.00	–0.56	0.60	–8.75	8.90	–1.06	0.33
C. RT difference (post–pre)												
Light	<i>27.80</i>	<i>7.20</i>	<i>–4.13</i>	<i>0.00</i>	–0.31	13.40	0.03	0.98	15.66	10.40	–1.61	0.15
Sound	7.63	6.70	–0.26	0.26	22.60	12.90	–2.22	0.06	25.97	12.60	–2.20	0.06
Touch	6.54	9.60	–73.00	0.49	13.18	10.80	–1.30	0.24	9.25	12.50	–0.79	0.46

The mean difference, standard errors of the difference, paired samples *t*- and *P*-values are presented for each combination. Columns are arranged according to the pair used in the exposure phase: sound/light (light leading by 100 ms), sound/touch (touch leading by 100 ms), and light/touch (touch leading by 100 ms) respectively. Rows are arranged according to the pairs of stimuli tested: sound/light (positive means “light first”), sound/touch (positive means “touch first”), and light/touch (positive means “touch first”). The first three rows correspond to the PSS difference after exposure (post minus pre) for the three pairs tested. The next three rows correspond to the JND difference after exposure (post minus pre) for the three pairs tested. The last three rows correspond to the reaction time differences after exposure (post minus pre) for the three individual stimuli tested. Values in italic indicate significant changes following exposure.

(2007) had an initial exposure of 240 pairs (~3 min 15 s) and 8 top-ups (~6 s) after every response, while we had an initial exposure of 5 min, with 10 s of top-ups after every 8 responses. Also, the task given to subjects during the exposure phase, to keep them focused, could have an effect: our subjects were told to either count the pairs or attend to the temporal order, while Navarra et al. (2007) had subjects attending to the duration of the individual stimuli—the temporal tasks may have unintentionally pulled the pair farther apart in time. A widening of the acceptance window for simultaneity appears to be the first stage in the recalibration process (Harris et al. 2008; Navarra et al. 2005) and so it seems that audio/tactile is adaptable but requires substantially more exposure than the audio/visual system. It remains to be demonstrated if this increased JND for tactile pairs leads to a subsequent shift of the PSS. Further research should try various asynchronous delays and exposure durations to assess the flexibility of the simultaneity constancy system for pairs involving touch.

The rigidity of crossmodal comparisons involving touch might be related to the fact that the temporal properties of a touch are more predictable and do not generally need to take outside factors into account (Miyazaki et al. 2006). Audio/tactile and visuo/tactile stimulus pairs must occur on the skin (such as watching a bug on your arm or hearing and seeing your hands clap) and therefore, only internal factors need to be taken into account when resynchronising

the stimuli. On the other hand, audio/visual stimuli can arrive at the senses with a wide range of temporal asynchronies (light always first, of course) and so the audio/visual simultaneity constancy system needs the flexibility to cope with this.

One curious result of this experiment was that after exposure to a light/touch pair with touch first, the PSS for the sound/light pair was found to change such that light needed to be presented first, the same shift as was found after exposure to a sound/light pair with light presented first. This very interesting and unexpected finding (unexpected because the light/touch pair was the one being repeatedly exposed and yet this combination did not change its temporal properties) suggests that some temporal recalibration of the audio/visual simultaneity constancy mechanism occurred even after exposure to non-audio/visual temporally staggered pairs. This further supports the conclusion that audio/visual simultaneity is more flexible than other stimulus pairs.

Similarly, the finding that the JND for sound/light TOJs became smaller after exposure to sound/touch suggests increased flexibility of audio/visual simultaneity; only audio/visual simultaneity properties changed while the audio/tactile and visuo/tactile remained rigid throughout. However, it remains a puzzle how exposure to a light/touch pair does not effect a change in the perception of simultaneity for light/touch pairs, but does for sound/light pairs.

Changes in temporal properties might be due to attention changes instead of adaptation of a simultaneity constancy mechanism. We therefore measured reaction times as a control for effects of attention.

Attention control experiment 2: Reaction times after exposure

PSS shifts following repeated exposure may be the result of increases or decreases in attention and consequent changes in processing time for individual stimuli (Spence et al. 2001). During the exposure to our adapting stimuli, the participants may have attended more to the stimulus that came first (the “onset” of the pair), or they may have attended more to the stimulus that came second since it signalled the “offset” of the multisensory stimulus, or they may have attended more to the third stimulus (the one not present in the exposure) since it became more novel. Reaction times to individual stimuli were therefore tested before and after exposure.

Methods

The same three stimuli and multisensory cube² as in experiment 1 were used (Fig. 1). One of the three stimuli was presented, in random order, and the participant pressed down a response key as soon as they detected it. Then, following a variable period between trials of from 200 to 1,000 ms, the next stimulus was presented and reacted to. Reaction times faster than 100 ms or slower than 500 ms were automatically discarded and the trial repeated.

There were a total of three blocks corresponding to the same three exposures conditions described in experiment 1 that were run on different days (see “General methods” and “Experiment 1” for further details on the exposure procedures). There were 30 reaction times for each of three stimuli for a total of 90 reaction times before exposure and another 90 reaction times after exposure. In each of the three blocks participants performed the reaction times pre-exposure. Then had the 5 min of exposure. Then performed the reaction time task post-exposure with the top-ups (10 s after every 12 responses). Each block took roughly 30 min to complete.

Results

Overall, reaction times were stable following repeated exposures. Repeated exposure affected the reaction time to

only one stimulus (the light) and only after repeated exposure to a sound/light pair (paired-samples *t*-test: $t_7 = -4.13$, $P = 0.004$). The reaction time for the light increased from 249 to 277 ms after exposure to a staggered sound/light pair, where the light led the sound by 100 ms. No other significant changes in reaction times were observed (see Table 1C for mean RT differences, and *t*- and *P*-values).

Discussion

Exposure did not generally change reaction times to the individual stimuli suggesting no change in the attention of the participants to any of the particular stimuli. While there may have been a trend towards a slightly increased reaction time following the exposure, this increase was, by and large, not significant. The significantly increased reaction time to the light, however, suggests that after exposure to a sound/light pair, the light alone (not both members of the pair) had a longer processing time, suggesting that it may have been relatively less attended. Such a slowing down of the processing of the visual stimulus could have created the shift in the PSS of the sound/light pair in which the light needed to be on earlier to be perceived as synchronous with the sound because it took longer to be processed than usual. However, this interpretation cannot explain the other PSS shift or the JND decrease. Further, reaction times have been shown to use different temporal information than TOJs (Jaskowski and Verleger 2000). Accordingly, processing speed may be differentially affected by attention in the two tasks; while exposure did not affect processing speed as measured by reaction times—suggesting that it did not affect attention—it may still have affected processing speed for TOJs.

General discussion

Our experiments have confirmed that the simultaneity constancy mechanism, which accounts for the different delays that each modality takes to reach the brain, is adaptable. The perception of simultaneity can shift, regarding a new delay between stimuli in different modalities as corresponding to simultaneity, in response to repeated exposure of a particular temporal relationship. Of the three stimulus pairs tested however, only the PSS and the JND of audio/visual comparisons seemed adaptable. Further, changes in attention (or processing speed, as measured by reaction times in a control experiment) certainly do not account for all of the changes found in the perception of simultaneity. It seems that the audio/visual simultaneity constancy mechanism, which arguably encounters the most temporal variability and thus requires the most temporal flexibility, is more easily adapted.

² The experiment was initially performed with the tactile stimulus attached to the finger, a speaker on the desk and an LED attached to the top of the tactile stimulus—not in a stimulus box. Using this setup there were no reaction differences for any of the three stimuli following any of the three adaptation sequences. At the request of a reviewer, the experiment was repeated using the stimulus box used in experiment 1.

Takahashi et al. (2007) also showed how robust the tactile system is, as compared to audio/visual simultaneity recalibration. They found that visual/haptic recalibration indeed occurred when either the visual or the haptic information regarding the deformation of an object is delayed; but this adaptation did not generalise to the other hand. For audio/visual recalibration, however, generalisation to other audio/visual stimuli is common (Fujisaki et al. 2004; Navarra et al. 2005; Vatakis and Spence 2006).

Navarra et al. (2007) suggested that the temporal binding window for tactile stimuli might be more rigid than the temporal integration window for audio/visual stimuli, even though they did find an adaptation for audio/tactile stimuli following exposure. Navarra et al. (2007) and the above experiment used simple tactile stimuli (passive touches). However, complex stimuli may require a different mechanism for resynchronising, which might be more adaptable than the mechanism for simple stimuli. If complex stimuli are more adaptable, then complex tactile stimuli might be as adaptable as audio/visual pairs—where a complex tactile stimulus is one that includes a motor component.

Indeed, visuo/tactile temporal recalibration has been found for complex touches that are more “cause and effect” such as driving a car (Cunningham et al. 2001a, b) or pressing keys on a keyboard and seeing the letters appear on the screen (Stetson et al. 2006). Most readers have experienced these last two examples already: the movement of the car and the appearance of the keys on a computer screen appear to be simultaneous with the action that caused them but of course there is some delay that has been adapted to as a result of experience.

When the level of complexity is held constant as in the present experiment where the stimuli were limited to beeps, flashes and prods, it reveals differences between the systems responsible for basic TOJs. The perception of simultaneity is likely achieved by a number of separate mechanisms for the different stimulus combinations, each affected differently by various stimulus contingencies.

Acknowledgments These experiments were sponsored by the Natural Sciences and Engineering Research Council (NSERC) of Canada. Vanessa Harrar holds an NSERC Scholarship. Special thanks to Shin'ya Nishida for fruitful discussions on the topics in this paper.

References

- Bertelson P, Aschersleben G (2003) Temporal ventriloquism: cross-modal interaction on the time dimension. 1. Evidence from auditory–visual temporal order judgment. *Int J Psychophysiol* 50:147–155
- Calvert GA, Spence C, Stein BE (2004) The handbook of multisensory processes. MIT Press, Camb, Mass
- Cunningham DW, Billock VA, Tsou BH (2001a) Sensorimotor adaptation to violations of temporal contiguity. *Psychol Sci* 12:532–535
- Cunningham DW, Chatziastros A, von der Heyde M, Bulthoff HH (2001b) Driving in the future: temporal visuomotor adaptation and generalization. *J Vis* 1:88–98
- Driver J, Spence C (2000) Multisensory perception: beyond modularity and convergence. *Curr Biol* 10:R731–R735
- Fujisaki W, Shimojo S, Kashino M, Nishida S (2004) Recalibration of audiovisual simultaneity. *Nat Neurosci* 7:773–778
- Harrar V, Harris LR (2005) Simultaneity constancy: detecting events with touch and vision. *Exp Brain Res* 166:465–473
- Harris LR, Harrar V, Jaekl P, Kopinska A (2008) Mechanisms of simultaneity constancy. In: Nijhawan R (ed) *Space and time in perception and action*. Cambridge University Press, Cambridge (in press)
- Jaskowski P, Verleger R (2000) Attentional bias toward low-intensity stimuli: an explanation for the intensity dissociation between reaction time and temporal order judgment? *Conscious Cogn* 9:435–456
- King AJ, Palmer AR (1985) Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp Brain Res* 60:492–500
- Kopinska A, Harris LR (2004) Simultaneity constancy. *Perception* 33:1049–1060
- Miyazaki M, Yamamoto S, Uchida S, Kitazawa S (2006) Bayesian calibration of simultaneity in tactile temporal order judgment. *Nat Neurosci* 9:875–877
- Navarra J, Soto-Faraco S, Spence C (2007) Adaptation to audiotactile asynchrony. *Neurosci Lett* 413:72–76
- Navarra J, Vatakis A, Zampini M, Soto-Faraco S, Humphreys W, Spence C (2005) Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Brain Res Cogn Brain Res* 25:499–507
- Nickalls RWD (1996) The influence of target angular velocity on visual latency difference determined using the rotating Pulfrich effect. *Vision Res* 36:2865–2872
- Poppel E, Schill K, von Steinbüchel N (1990) Sensory integration within temporally neutral systems states: a hypothesis. *Naturwissenschaften* 77:89–91
- Spence C, Shore DI, Klein RM (2001) Multisensory prior entry. *J Exp Psychol Gen* 130:799–832
- Spence C, Baddeley R, Zampini M, James R, Shore DI (2003) Multisensory temporal order judgments: when two locations are better than one. *Percept Psychophys* 65:318–328
- Stetson C, Cui X, Montague PR, Eagleman DM (2006) Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron* 51:651–659
- Takahashi K, Saiki J, Watanabe K (2007) Recalibration of vision-haptics temporal simultaneity. Poster presented at the International Multisensory Research Forum in Sydney, Australia
- Vatakis A, Spence C (2006) Audiovisual synchrony perception for music, speech, and object actions. *Brain Res* 1111:134–142
- Vatakis A, Navarra J, Soto-Faraco S, Spence C (2007) Temporal recalibration during asynchronous audiovisual speech perception. *Exp Brain Res* 181:173–181
- Von Békésy G (1963) Interaction of paired sensory stimuli and conduction in peripheral nerves. *J Appl Physiol* 18:1276–1284
- Vroomen J, Keetels M, de Gelder B, Bertelson P (2004) Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Brain Res Cogn Brain Res* 22:32–35
- Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA (2004) Unifying multisensory signals across time and space. *Exp Brain Res* 158:252–258
- Wilson JA, Anstis SM (1969) Visual delay as a function of luminance. *Am J Psychol* 82:350–358
- Zampini M, Guest S, Shore DI, Spence C (2005) Audio-visual simultaneity judgments. *Percept Psychophys* 67:531–544