

# Shared Temporal Accuracy of Action Execution and Sensory Perception

**Tomomitsu Herai (herai@brn.dis.titech.ac.jp)**

Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku,  
Yokohama, 226-8502, Japan

**Ken Mogi (kenmogi@csl.sony.co.jp)**

Sony Computer Science Laboratories, Takanawa Muse Bldg., 3-14-13, Higashigotanda,  
Shinagawa-ku, Tokyo, 141-0022, Japan

## Abstract

Integration of an action and its sensory feedback is important in interacting with an uncertain environment and construct a consistent model of the world. In this process, multisensory data need to be processed, in which audition and vision play important roles. Subjective simultaneity of audiovisual stimuli is affected by various factors. To investigate the relation between our subjective simultaneity of audiovisual stimuli and our action, we conducted an experiment in which subjects' action affected the temporal patterns of resulting stimuli. The modes of contingency between action and stimuli were made variable. We found significant correlations between the accuracies of actions and the "window" of subjective simultaneity among subjects, although their task performances were widely varied. In addition, the correlation patterns were found to depend on the contingency between the key pressing and stimuli. These results suggest that the subjective simultaneity of audiovisual stimuli correlates with the accuracy of execution of action, indicating a common mechanism engaging the perception of subjective simultaneity in sensorimotor integration and action execution.

**Keywords:** sensorimotor contingency; audiovisual integration; subjective simultaneity; timing

## Introduction

We interact with our environments through voluntary movements and its sensory feedback. We recognize ourselves and the external world through sensorimotor interactions. However, it is rarely the case that we have a complete knowledge about the sensorimotor contingency in a given context. Understanding the mechanism of integrating motor and sensory information in the presence of uncertainty provides us with important information as to how we construct a consistent model of the world.

The sensorimotor contingency affects both our perception and action. Recent studies have shown that if sensory stimuli are the results of our own action (i.e. pressing a key which generates beep), our perceived timing of those sensory feedback are closer to the timing of action than actually (Haggard, Clark & Kalogeras, 2002; Tsakiris & Haggard, 2003). In addition, this temporal shift was affected by the event probability of sensory stimuli (Engbert & Wohlschläger, 2007; Moore & Haggard, 2008). A higher event probability has been found to lead to a larger temporal shift. Furthermore, the adaptation of this temporal shift recalibrates the perceived timing of the action and the visual

stimulus, leading to a illusionary temporal reversal in which the subject perceive the flash before the action (Stetson et al., 2006). These studies suggest that our cognitive processes related to agency and expectation affect the perceived timing of sensory events.

Temporal shifts in perception are ubiquitous in our daily life. We have to always treat multisensory information, where the knowledge about the external world is mainly from vision and audition. Since the light travels much faster than the sound, the delay of the timings at which each sensory organ receives its appropriate stimuli becomes progressively larger depending on the distance from the event. Despite this temporal disturbance, we tend to perceive the light and sound signals from a single source simultaneously (Kopinska & Harris, 2004; Stone et al., 2001; Sugita & Suzuki, 2003). The temporal window of subjective simultaneity has shown to be affected by various factors. Studies have shown that adjacent adaptation affected the width of simultaneous temporal window and subjective simultaneity (Fujisaki et al., 2004; Navarra et al., 2005; Vroomen et al., 2004). Spatial position (Zampini, Shore & Spence, 2003; Zampini et al., 2005) and attention changed the subjective simultaneity of audiovisual stimuli (Zampini, Shore & Spence, 2005). The subjects in these studies were typically presented with the stimuli passively. The relation between our action and the audiovisual integration remain still unclear.

As noted above, intentional action changes the perceived timing of its sensory feedback. Based on this point, we hypothesized that the subjective judgment of the simultaneity of audiovisual stimuli would be affected by the processes of sensorimotor integration. To examine this hypothesis, we conducted an experiment using two keys which generated either a beep sound or flash. The timings of key pressings by the subjects were reflected in the following generations of flash and beep. In general, variances exist in the accuracy of motor performance and the sensitivity of perception. We were interested if there was a common mechanism involving action execution and simultaneity perception as reflected in the performances of subjects.

## Experiment

### Methods

**Subjects** Eight healthy subjects participated in this study (4 females and 4 males, 24 to 45 years old, with a mean of 29 and s.d. of 6.5. The subjects were all right-handed). All had normal or corrected-to-normal vision and audition and motor ability to perform the tasks. Written informed consents were obtained from all subjects. They were naïve about the purpose of the present study.

**Stimuli and Apparatus** The visual and auditory stimuli were produced by a PC (Panasonic CF-W4). The programs were created with DirectX. The visual stimulus (flash) was a white circle ( $2.8^\circ$ ) and was presented for the duration of one frame at the center of a black background ( $17.6^\circ \times 23.0^\circ$ ) on the monitor with a refresh rate of 60Hz. The auditory stimulus was a beep (1,800Hz, 10ms) sound which was presented to both ears through a headphone (Sony MDR-XD100). The manipulations in the experiment were conducted by the keys on the keyboard of the PC. The two stimuli were released by pressing the 'D' and ';' (semicolon) keys. The correspondence between the keys and the stimuli are explained in the next section. Subjects responded to the presentation of stimuli by pressing one of the arrow keys ("<-" (simultaneous) or "->" (not simultaneous) located at the lower right corner on the keyboard with the index finger of their right hand. Participants were seated at a distance of 60cm from the monitor and put their index fingers of both hands with a gap of 12 cm from each other on the keyboard (where there were "D" and ";" keys) (Figure 1).

The experiments were conducted in a dimly lit and sound attenuated room.

**Experimental Design** There were two prominent conditions ("coupled delay" condition and "random delay" condition). 100 trials were conducted for each condition.

In the "coupled delay" condition, the delays between the first and second stimuli (flash to beep or beep to flash) were given as a function of the interval of key presses by the subject (Figure 2). Subjects were instructed to press two keys simultaneously in the experiment, resulting in varied intervals between the two key pressings in physical time, although the subjects might deem them as simultaneous. The actual intervals between the key pressings were reflected in the delays between the flash and beep by twentyfold. Thus, the subject's action was coupled with the resulting sensory feedback in a magnified manner.

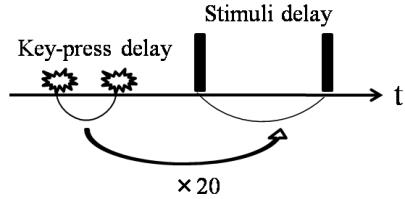


Figure 2: The relation between the key pressing intervals and the stimuli delays.

This experimental condition had three situations categorized by key-stimuli relationships (table 1). In condition 1, pressing the left and the right keys generated the flash and beep, respectively. In condition 2, the key-stimuli relationship was a reversal of that in condition 1. In condition 3, the relation between the keys and stimuli were randomized. The total trial number of the two assignments (left key-flash/right key-beep) and (left key-beep/right key-flash) situations were the same (50 times each) in condition 3. 100 ms after the second key was pressed, either the flash or beep was presented as specified by the conditions 1, 2 and 3. After the designated delay set to be 20 times of the key press interval, the other stimulus (flash or beep) was presented. The subjects judged whether they perceived flash and beep simultaneously or not in a two alternative forced-choice.

Table 1: Relationship between the keys and the stimuli

	Left key	Right key
Condition 1	flash	beep
Condition 2	beep	flash
Condition 3	Key-Stimulus relationship is random.	

As a control experiment, the subjects conducted the random delay condition, in which the delays between the beep and flash were randomly distributed between -270 ms (beep first) to + 270ms (flash first). Delay values of -270ms, -240ms, -210ms, ..., 0ms, ..., +210ms, +240ms, +270ms, (separated by 30ms step), were given. Each value was used 5 times each, except for 0ms which was used 10 times.

No feedback was provided for the accuracy of key pressing and the simultaneity judgment. Subjects conducted the control condition first and then conducted the coupled delay conditions. The orders of experimental conditions were counterbalanced among the subjects.

**Procedures** Before the experiment, to become familiar with the experimental tasks, the subjects practiced twenty trials in the same situation as the control condition except that the delays between the flash and beep were set to be different. They were instructed to press the two keys as simultaneously as possible.

In the experiment, before starting the each condition, the subjects were instructed of the relationship between the keys and the stimuli by the experimenter. After the experimenter

confirmed that the subject understood this relationship, the experiment started.

Figure 3 shows the procedure for one trial in the experiment. The fixation point was presented for 1.5s at the beginning of the trial. After the fixation disappeared, the subject pressed the two keys simultaneously on their own timings. 100ms after the both keys were pressed the flash and beep were presented depending on the key-stimuli relationship. The subjects judged whether they perceived the flash and beep simultaneously or not by pressing the judging keys with their index finger of the right hand accurately and fast as possible as they could. The inter stimulus interval was set to be 2.5s

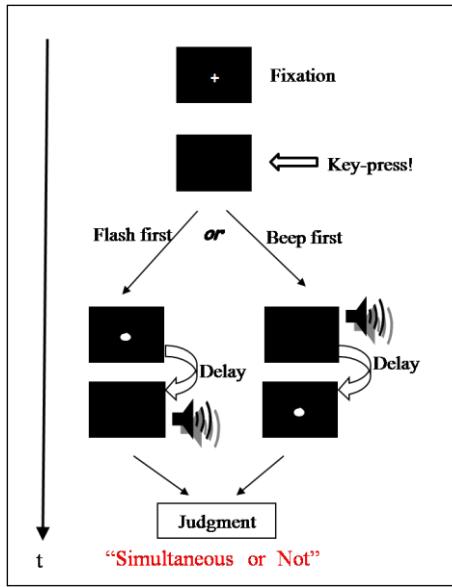


Figure 3: A procedure of one trial.

## Results

Since the delays between the beep and flash depended on the subjects' action, the standard deviations (SDs) of the delays of all stimuli reflected an accuracy of their simultaneous key pressings. On the other hand, the SDs of the delays judged simultaneous by the subject reflected the thresholds of subjective simultaneity. The distributions of the delays between flash and beep were different among the subjects and the conditions except for the control condition (where the interval was given randomly independent of the subject's key pressing). The analysis was done for all stimuli and the subset of stimuli judged as simultaneous by the subject.

Figure 4 shows that the SDs for all stimuli and for the stimuli judged as simultaneous have large positive correlations for conditions 1, 2, and 3, where the correlation of cond.3 was significant [cond.1,  $r=.67$  ( $t_6=2.19$ ,  $p=.072$ ); cond.2,  $r=.68$  ( $t_6=2.24$ ,  $p=.066$ ); cond.3,  $r=.88$  ( $t_6=4.46$ ,  $p=.0035$ )]. One-way within-participants ANOVA showed that the SDs of the delays judged as simultaneous were significantly different among the conditions ( $F(2,14)=4.85$ ,

$p=.025$ ). The SDs of the delays of all stimuli were not different significantly ( $F(2,14)=1.44$ ,  $p=.27$ ). This result indicates that the nature of the subjects' perception changed whereas the action remained constant among the conditions.

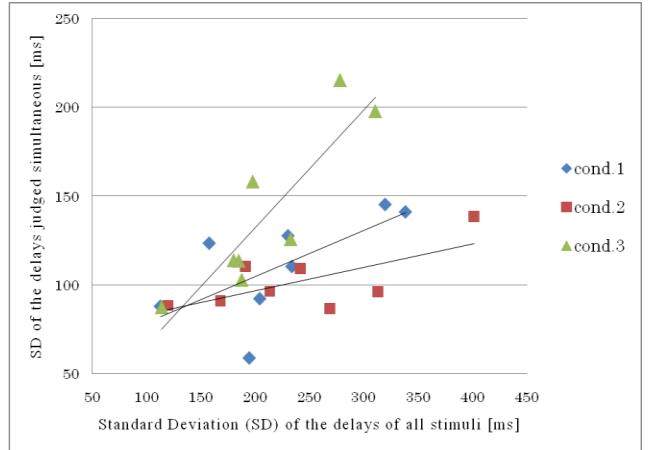


Figure 4: Plot of the standard deviation (SD) of the delays of all stimuli versus the delays judged simultaneous in each condition and each subject. Each point represents a subject.

For a further analysis, we calculated the regression lines for each condition. We analyzed the homogeneity of regression slopes statistically and found that they were significantly different among the conditions ( $F(2,18)=5.93$ ,  $p=.01$ ). We also analyzed the homogeneity of regression slopes in each pair with the Bonferroni method. It was revealed that the regression slopes of cond. 2 and cond. 3 were significantly different ( $F(1,12)=13.91$ ,  $p=.0029 < .05/3 = .017$ ).

Figure 5 shows the mean reaction times in the four (1, 2, 3, and control) conditions. One-way within-participants ANOVA showed that the reaction times were significantly different among the conditions ( $F(3,21)=5.45$ ,  $p=.0063$ ). A post-hoc multiple comparison analysis (Ryan's method) revealed that the reaction times between the control condition and condition 3 were significantly different ( $t_{28}=2.86$ ,  $p=.0078$ ).

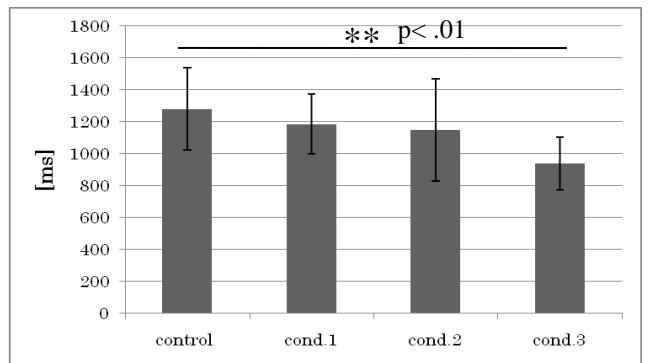


Figure 5: Mean reaction times in each condition.

## Discussion

The present study investigated 1) how the subjective simultaneity of audiovisual stimuli was affected when the delays between beep and flash depended on the subjects' action and 2) how the sensorimotor contingency was engaged in the integration of the action and its sensory feedback.

The standard deviations of the delays of all stimuli and that of the delays judged simultaneous represented the accuracy of simultaneous key pressing and the threshold of subjective simultaneity, respectively. There are large individual differences in subjective simultaneity in audiovisual stimuli (Stone et al., 2001). The values of both axes in Figure 4 depend on each subjects' performances of key pressings and simultaneity judgments, reflecting the individual characteristics. Nevertheless, the values were significantly correlated among the subjects in the condition 3. The correlations in conditions 1 and 2 are also noticeable. These positive correlations are not trivial. The subject's accuracy of key pressing and the accuracy of simultaneity judgment can be in principle independent. Our result suggests the existence of a shared temporal accuracy between action execution (key pressing) and sensory perception (simultaneity judgment). Thus, we suggest that the threshold or accuracy of motor output is associated with the threshold or criterion of the simultaneity judgments of beep and flash. In addition, we suggest that the threshold of subjective simultaneity might depend on the accuracy of simultaneous movement of two hands. There might be a shared threshold of simultaneity within the brain mechanism, which is embodied both in our action and perception.

The standard deviations of the delays judged simultaneous were significantly different among the conditions while that of the delays of all stimuli were not. The results indicated that subjects' threshold of subjective simultaneity in cond.1 and 2 became severer than that in cond.3. This change cannot be explained by adaptation (Fujisaki et al., 2004; Navarra et al., 2005) and spatial factor (Zampini et al., 2003; Zampini et al., 2005). We thought that subjects' knowledge about the contingency of the key and the stimuli induced perceptual and behavioral changes in the present study. Subjects might attend more their key pressing in cond.1 and 2 than in cond.3 because their actions were linked to the stimuli more directly in these conditions. Since a temporal resolution of tactile is higher than audiovisual, they tended to notice an asynchronous of their key pressing and it led severe judgments.

Since the variables of both axes in Figure 4 depended on the subject's action and perception, the regression lines indicate the nature of sensorimotor integration in the each condition. The relation between each subjects' action and perception would affect the slopes of regression lines significantly. The cognitive mechanism engaging this change might be common among the subjects as there were large positive correlations in each condition. These large correlations and the change of regression slope indicated that the sensorimotor contingency affected not only the

perceived timing of sensory stimulus (Haggard et al., 2002; Moore & Haggard, 2008) but also the subjective simultaneity of audiovisual stimuli.

The significant effect on the reaction times might reflect an interaction of the sensorimotor contingency and subjects' knowledge of it. Since the delay between beep and flash were linked to the subject's key pressing in conditions 1, 2, and 3, the subjects were in principle able to use the information related to their key pressing in judging the simultaneity of the beep and flash, starting the process of preparation to make judgments just after they have pressed the two keys. They would only confirm their insight of their own action by referring to the stimuli that follow. On the other hand, in the control condition, the subjects had to wait for the occurrence of the beep and flash, basing their responses solely on the actual properties of the flash and beep stimuli. Therefore, the timing of preparation for the simultaneity judgment in the control condition would be later than that in the conditions 1, 2, and 3, leading to larger values of reaction time.

However, the above assumption cannot explain why the reaction time in the condition 1 and 2 were not significantly different from that in the control condition.

Since the subjects knew the relations between the key pressings and the stimuli and attended to their key pressing in the conditions 1 and 2, they might come to care for the order of key pressing in judging the simultaneity of stimuli. In contrast, in condition 3, they would come to ignore the orders and attended to the simultaneity of action and the stimuli only. These factors would mean that the subjects would take more time to judge in condition 1 and 2 than in condition 3. If the subjects were asked about the temporal order of beep and flash instead of simultaneity, their reaction time would be possibly faster in conditions 1 and 2 than in condition 3.

In conclusion, our results suggest a shared temporal accuracy between the action execution and sensory simultaneity judgment, when the action and sensory feedback are coupled. The subjective simultaneity of visual and auditory stimuli is affected not only by adaptation (Fujisaki et al., 2004; Navarra et al., 2005), spatial factor (Zampini et al., 2003; Zampini et al., 2005) and attention (Zampini, Shore & Spence, 2005) but also by the accuracy of the key pressing that induce the stimuli, possibly affected by the contingency between action and its sensory feedback. When the subjects were able to access to the information as to the relation between the key pressing and the stimuli, their simultaneous judgments became severer. These results indicate that there is a close correlation between action and perception affected by the sensori-motor contingency.

## Acknowledgements

We would like to thank Fumihiko Taya, Qi Zhang, Toru Yanagawa, Kei Omata, Takayasu Sekine, Ayako Onzo, Tamami Sudo, Fumiko Tanabe, Eiichi Hoshino, Shinichi

Nozawa, Tetsuo Ishikawa, Mayumi Toshima, Haruna Takagawa and Miki Kato for their useful comments.

## References

Engbert, K., & Wohlschläger, A. (2007). Intentions and expectations in temporal binding. *Conscious Cogn*, 16(2), 255-264.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat Neurosci*, 7(7), 773-778.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nat Neurosci*, 5(4), 382-385.

Kopinska, A., & Harris, L. (2004). Simultaneity constancy. *Perception*, 33(9), 1049-1060.

Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Conscious Cogn*, 17(1), 136-144.

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(2), 499-507.

Stetson, C., Cui, X., Montague, P., & Eagleman, D. (2006). Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron*, 51(5), 651-659.

Stone, J., Hunkin, N., Porrill, J., Wood, R., Keeler, V., Beanland, M., et al. (2001). When is now? Perception of simultaneity. *Proc Biol Sci*, 268(1462), 31-38.

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*, 421(6926), 911.

Tsakiris, M., & Haggard, P. (2003). Awareness of somatic events associated with a voluntary action. *Exp Brain Res*, 149(4), 439-446.

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(1), 32-35.

Zampini, M., Guest, S., Shore, D., & Spence, C. (2005). Audio-visual simultaneity judgments. *Percept Psychophys*, 67(3), 531-544.

Zampini, M., Shore, D., & Spence, C. (2003). Audiovisual temporal order judgments. *Exp Brain Res*, 152(2), 198-210.

Zampini, M., Shore, D., & Spence, C. (2005). Audiovisual prior entry. *Neurosci Lett*, 381(3), 217-222.