Recalibration of subjective simultaneity between self-generated movement and delayed auditory feedback

Koichi Toida[a,c], Kanako Ueno[a,c] and Sotaro Shimada[b,c]

Temporal contingency between motor commands and corresponding auditory feedback is crucial for perception of self-generated sound as well as external auditory events. The present study examined whether delay detection of self-generated sound was affected by the range of delayed auditory feedback used during the experiment. Participants pressed a button with their right index finger and judged whether auditory feedback was delayed compared with the sensation of finger movement. The range of auditory feedback delay was varied across conditions. To calculate the delay detection threshold (DDT), that is, the point at which the delay detection rate was 50%, we fitted a logistic function to the delay-detection probability curve. The DDT was significantly different across conditions (Tukey–Kramer's honestly significant difference test, $P<0.01$). Specifically, the DDT became longer as the mean delay of the range increased. However, this shift was not observed for the delay range with a minimum delay greater than 250 ms. We propose that the subjective simultaneity of auditory feedback and self-body movement is, to some extent, automatically recalibrated toward the mean delay of the delay range used in the experiment. *NeuroReport* 25:284–288 © 2014 Wolters Kluwer Health | Lippincott Williams & Wilkins.

NeuroReport 2014, **25**:284–288

Keywords: delay detection threshold, delayed auditory feedback, perceptual recalibration, self-body recognition, simultaneity judgment

Departments of [a]Architecture, [b]Electronics and Bioinformatics, School of Science and Technology, Meiji University, Kawasaki and [c]Japan Science and Technology Agency, CREST, Tokyo, Japan

Correspondence to Sotaro Shimada, PhD, Department of Electronics and Bioinformatics, School of Science and Technology, Meiji University, 1-1-1 Higashi-Mita, Tama-ku, Kawasaki 214-8571, Japan
Tel/fax: +81 44 934 7291; e-mail: sshimada@isc.meiji.ac.jp

Received 4 October 2013 **accepted** 7 October 2013

Introduction

Temporal contingency between body movement and the associated sensory feedback is crucial to perceive self-generated effects on the surrounding environment. However, the judgment of synchrony between a body movement and the associated sensory feedback is not a straightforward process. For instance, several factors can introduce temporal delay between an external event and the arrival of auditory feedback to the brain. The velocity of sound is ~340 m/s, meaning that a sound made within 1 m of the body will have a 2–3 ms delay before it arrives to the ear. Similarly, there is a delay (in the order of tens of milliseconds) in the time required for sound information to travel from the auditory receptor in the inner ear to the primary auditory area in the cortex [1]. When playing a musical instrument, such as a horn or electric guitar, there is an intrinsic delay between the self-body movement and the generation of the sound. Further, visual, tactile, and auditory information all have unique latencies in the time required to reach the brain, and these differences must be integrated to judge synchrony among different modalities.

Our previous studies revealed that a delay in visual feedback of less than 200–300 ms is perceived as 'not delayed' with associated tactile or proprioceptive feedback [2,3]. This indicates that there is a specific time window, interpreted as 'the present', in which multisensory information can be integrated into a neural self-body representation. Although the subjective simultaneity between self-body movement and associated auditory feedback has also been examined, the findings are controversial. Whereas some studies found the time frame for subjective simultaneity between self-body movement and the associated auditory feedback to be relatively short, that is, 40 ms [4] and 55 ms [5], others found a longer delay, such as 100 ms [6]. Our preliminary experiment (reported in this manuscript as condition 2) found that this delay could exceed 200 ms in certain experimental settings. This discrepancy among findings may be explained by the presence of an automatic (unconscious) recalibration process that integrates the delayed auditory feedback presented during the experiment [7]. Thus, the delay range used in each experiment might have influenced the judgment of subjective simultaneity, as the participants adjusted their criterion of simultaneity to the exposed delay in the experiment. However, most studies that examined the temporal recalibration of multisensory integration assessed temporal discrepancies of tens of millisecond, and whether longer temporal discrepancies such as hundreds of millisecond, which was the case with our preliminary observation, is also within the scope of the recalibration process has not been fully investigated.

In this study, we used several different delay ranges to investigate the extent to which the delay between self-body movement and corresponding auditory feedback affects perceived simultaneity with five experiments. The delay range for each experiment varied from

0959-4965 © 2014 Wolters Kluwer Health | Lippincott Williams & Wilkins DOI: 10.1097/WNR.0000000000000079

19–253 ms (minimum: experiment 1) to 286–519 ms (maximum: experiment 5). We systematically introduced various delay ranges in the timing of auditory feedback and measured the delay detection threshold (DDT). Precisely, DDT is defined as the delay length at which the probability that the participants detect the temporal delay between self-body movement and the associated auditory feedback is 50%. Recalibration of subjective simultaneity in response to delayed auditory feedback would presumably be indicated by variations in the DDTs according to the delay range used.

Methods

Participants

Sixty-nine healthy students took part in the experiment. They received monetary compensation for their participation. Each participant was assigned to one of five experimental conditions, as follows: 12 (six female and six male, age 20.8 ± 1.4 years, mean ± SD) were assigned to condition 1; 26 (12 female and 14 male, age 21.3 ± 1.3 years, mean ± SD) to condition 2; 13 (seven female and six male, age 20.8 ± 1.7 years, mean ± SD) to condition 3; 12 (six female and six male, age 21.0 ± 1.5 years, mean ± SD) to condition 4; and six (two female and four male, age 21.2 ± 2.0 years, mean ± SD) to condition 5. Two participants under condition 2 and one under condition 3 were excluded from the analysis because they reported that the auditory feedback was delayed in every trial. All participants were right-handed and had normal hearing. The participants were naive as to the purpose of the experiment. The experiments were approved by the ethics committee of the School of Science and Technology, Meiji University, and conducted according to the principles and guidelines of the Declaration of Helsinki. Written informed consent was obtained from all participants.

Apparatus and procedures

We generated auditory feedback as regards the finger movements of the participants using a synthesizer (Micron SE; Alesis, Cumberland, Rhode Island, USA). The synthesizer was placed on a table and was connected to an audio delay-inserting hardware device (SPX2000; Yamaha, Hamamatsu, Japan). The delayed auditory feedback was presented to the participants through headphones (HDA 200; Sennheiser, Lower Saxony, Germany). Only one key of the synthesizer was visible; the other keys were masked with black felt. The allowance height of the key was adjusted to minimum so that the sound was generated when the key was pushed down to the bottom.

The participants wore an eye-mask to prevent any visual input during the experiment. They were asked to press the key on the apparatus and judge whether the auditory feedback (full-range pulsed sound, 2 ms in length) was delayed compared with their finger movements. In each trial, the participant moved their right index finger downward to press the key once, and then returned to the start position. The participants were asked to respond verbally in a forced choice manner ('delayed' or 'not delayed') at the end of the trial. A rest period that lasted for ~10 s followed each trial. The presentation order of the delay lengths was pseudorandomized. To avoid potential top–down effects of anticipation with respect to the nature of the auditory feedback, participants were not informed that the key was that of a synthesizer.

The intrinsic delay in auditory feedback with this setup was about 19 ms. There were eight auditory feedback delay lengths for each condition in the simultaneity judgment task. Each delay length was presented eight times, for a total of 64 trials. The range of delay lengths (including the intrinsic delay) varied across conditions as follows: from 19 to 253 ms under condition 1, from 119 to 353 ms under condition 2, from 186 to 419 ms under condition 3, from 19 to 119 ms under condition 4, and from 286 to 519 ms under condition 5. Under conditions 1–3 and 5, the delay lengths were in 33.3 ms intervals, and under condition 4, the delay lengths were in 14.3 ms intervals.

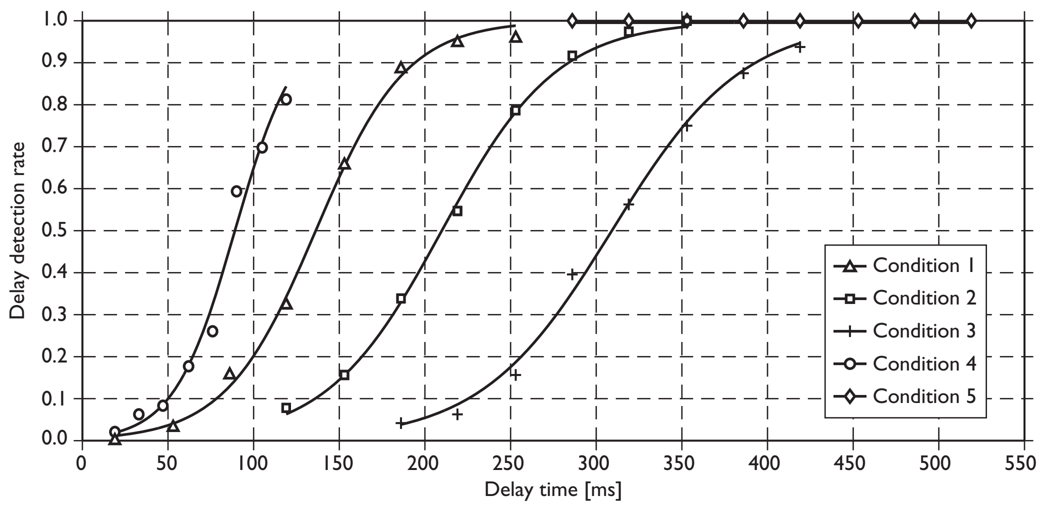
Data analysis

Participants were required to judge whether the auditory feedback was synchronized with their own finger movements or was delayed. We used this information to calculate the asynchrony judgment probability for each delay length. To examine the differences in the shape of the judgment curve between conditions, logistic curves were fitted to the participant responses on the basis of the following formula [3]:

$$P(t) = \frac{1}{1 + \exp[-a(t - t_{\mathrm{DDT}})]}, \tag{1}$$

where t is the auditory feedback delay length and $P(t)$ is the probability of making an asynchrony judgment, a indicates the steepness of the fitted curve, and t_{DDT} indicates the observer's delay detection threshold (DDT), representing the delay length at which synchrony and asynchrony judgment probabilities are equal (50%). In our experiment, t served as an independent variable and $P(t)$ was the observed value. Fitting was performed using a nonlinear least squares method (a trust-region algorithm), provided by the Curve Fitting toolbox in Matlab R2012b (The MathWorks Inc., Natick, Massachusetts, USA), to estimate a and t_{DDT}. In addition, the just noticeable difference (JND) is calculated as half the difference between the lower (25%) and upper (75%) bounds of the threshold, which reflects the subjective sensitivity to the delay near DDT [8]. One-way factorial analysis of variance was carried out using these parameters to investigate differences between conditions. For all statistical tests, the significance level was set at P less than 0.05.

Fig. 1



Asynchrony judgment curves by condition: asynchrony judgment curves averaged across participants for each condition. Logistic curves were fitted to the average asynchrony-judgment probabilities for eight delay lengths in four auditory feedback ranges (conditions).

Results

The average DDT across all participants was 136.3 ± 31.6, 208.9 ± 31.5, 309.1 ± 45.6, and 89.1 ± 13.4 ms (mean ± SD) under conditions 1, 2, 3, and 4, respectively. Under condition 5, all six participants reported that the auditory feedback was delayed in all trials; hence, we were not able to calculate DDT for this condition. The fitted curves for conditions 1–4 are shown in Fig. 1. To assess the differences in DDT and steepness among the conditions, we used one-way (conditions) factorial analyses of variance. We found a significant effect of condition [$F(3,56) = 107.2$, $P < 0.001$]. Subsequent analyses revealed that the DDTs were significantly different between every pair of conditions (Tukey–Kramer's honestly significant difference test, $P < 0.01$).

The average JND across all participants was 18.7 ± 11.8, 28.7 ± 14.4, 30.8 ± 7.8, and 18.2 ± 4.5 ms under conditions 1, 2, 3, and 4, respectively. Although there was a significant effect of the delay in JND [$F(3,56) = 3.26$, $P < 0.05$], the difference between conditions was not significant for each pair of conditions (Tukey–Kramer's honestly significant difference test). In accordance with this result, there was no significant effect of the delay on the steepness (a) of the fitted curve [$F(3,56) = 1.25$, $P = 0.30$].

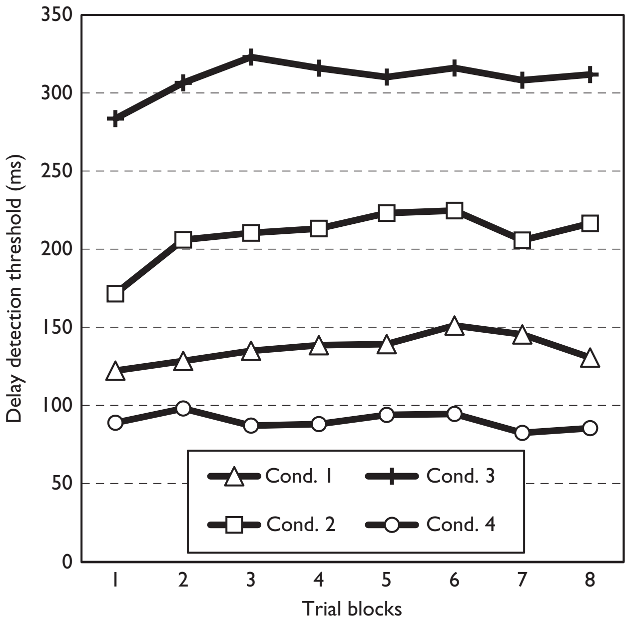
These results indicate that the delay range used under each condition modulated the DDT, but not the JND. To assess the time-series change in DDT during the experiment, we calculated the DDT for each block of eight trials within the total 64 trials. As shown in Fig. 2, the DDT appears to have been recalibrated according to the delay range through the first two to three blocks, as the DDT increases in the later blocks. This effect was most prominent under the condition 2. The proportions of the asynchrony judgment probability curve were significantly different between the first and fourth to eighth (except seventh) blocks [χ^2-test, $\chi^2(7) > 12.9$, $P < 0.05$; Fig. 2]. A similar tendency was also observed under condition 3, although the statistics did not reach a significant level ($P > 0.1$).

Discussion

The current results indicate that the DDT between a self-body movement and the associated auditory feedback is automatically recalibrated toward the mean of the auditory-feedback delay range used in the experiment. The steepness of the fitted curves and JNDs, which express sensitivity to delay were not significantly different among conditions. These results indicate that DDT was shifted as the delay ranges were shifted, with a similar sensitivity to the delay (JND). Importantly, this shift in DDT was no longer observed during exposure to a delay range with a minimal delay greater than 250 ms (condition 5).

It may be possible that the participants intentionally adjusted their threshold for delay detection to balance the number of asynchrony and synchrony judgments. However, during debriefing, no participant reported that he/she had explicitly changed the criteria during the experiment; hence, it is unlikely that the recalibration was accomplished voluntarily. It is more likely for there to be a 'range effect' in which the participants' categorical

Fig. 2



Shift of DDT during the experiment: DDT calculated from the average asynchrony judgment rates per block. One session consisted of eight blocks with eight trials in each block. Cond., condition; DDT, delay detection threshold.

judgment is affected by the minimum and maximum value of the stimuli, and the frequency for each categorical judgment tended to be equal [9]. In our experiment DDT was almost at the center (within ± 30 ms), between the minimum and the maximum delay length (e.g. under condition 1, the minimum delay length, DDT, and the maximum delay length were 19, 136, and 253 ms, respectively). Thus, the participants subconsciously adjusted their criterion for the categorical judgment of 'delay' or 'not delay' during the experiment so that the number of judgments for each category became nearly equal. Nevertheless, our results cannot be explained by this range effect alone, as we did not observe such modulation under the condition 5, in which all the participants responded with a 'delay' response for all trials. The minimum delay under condition 5 exceeded 250 ms, suggesting that there should be a limit for allowable delay length in which the range effect holds.

We suggest that the observed shift in DDT is the result of recalibration of subjective simultaneity through exposure to delayed auditory feedback. Previous studies have demonstrated that exposure to a fixed delay between auditory and visual stimuli for a period of several minutes induces a shift in the audio–visual subjective simultaneity in the direction of the fixed delay [7,10]. Similarly, temporal recalibration has been demonstrated between voluntary self-body movement and the associated visual or auditory stimulus [4,11–13]. Our findings indicate that temporal recalibration occurred even without an explicit adaptation phase (exposure to a fixed delay of some modality for several minutes), which was used in many previous studies. The experimental delay ranges served as adaptation stimuli; thus, the DDT might have been shifted toward the mean delay of the delay range used.

Temporal recalibration of subjective simultaneity might be explained as a relearning of the forward model. It is postulated that when an individual moves his/her body, a motor command is sent from the motor cortex to the muscle and a copy signal of the motor command (the 'efference copy') is simultaneously sent to the parietal lobe. The efference copy makes it possible to internally predict the sensory feedback caused by the self-generated movement (the 'forward model') [14–16]. The present results suggest that the forward model can relearn the temporal contingency between self-body movement and the associated auditory feedback when the prediction about auditory feedback is somewhat deviant from actual feedback timing (less than 300 ms). Exposure to a delay in auditory feedback induces the forward model to adjust to the new temporal relationship between self-generated movement and the associated auditory feedback. Figure 2 shows that this relearning was accomplished relatively quickly, through the first two to three blocks of the experiment.

The DDT obtained under condition 3 was ∼300 ms, which is far greater than that reported in previous studies [4,5]. Nevertheless, a 200–300-ms delay in sensory feedback is the key to the mechanism of multisensory integration. For example, our previous studies revealed that a delay in visual feedback of less than 200–300 ms was perceived as 'not delayed' with associated tactile or proprioceptive feedback [2,3]. Blakemore *et al.* [17] showed that a delay in tactile feedback of the self-movement delivered by a robotic hand could increase the tickliness of the tactile sensation up to 300 ms. Our result that the limit of DDT shift was about 300 ms (the shift was not observed under condition 5) also supports the notion that the brain can only adapt to the intersensory delay within 200–300 ms.

Another reason for the relatively late DDT, in addition to temporal recalibration, may be that the participant noticed the existence of an electric external apparatus in the experimental setting. A previous study reported that the audio–visual DDT was lengthened when a sound was associated with an object located far away from the participant [18]. This is interpreted as evidence that the brain takes the distance of the sound-making object into account and compensates for the time required for the sound to arrive at the individual. We suggest that a similar effect occurs when individuals engage with an external system like a tool or musical instrument. A previous study reported that the threshold of delay detection varied according to the musical instrument used to generate the sound [19]. Usually the feedback delay of an external system is not clear beforehand, and hence flexibility is understandably advantageous in recalibrating the internal model to match the external system.

Acknowledgements

This work was partially supported by the Japan Science and Technology Agency (JST/CREST) and KAKENHI (25700015 and 22240026).

Conflicts of interest

There are no conflicts of interest.

References

1 Corey DP, Hudspeth AJ. Response latency of vertebrate hair cells. *Biophys J* 1979; **26**:499–506.
2 Shimada S, Fukuda K, Hiraki K. Rubber hand illusion under delayed visual feedback. *PLoS One* 2009; **4**:e6185:1–5.
3 Shimada S, Qi Y, Hiraki K. Detection of visual feedback delay in active and passive self-body movements. *Exp Brain Res* 2010; **201**:359–364.
4 Heron J, Hanson JVM, Whitaker D. Effect before cause: supramodal recalibration of sensorimotor timing. *PLoS One* 2009; **4**:e7681.
5 Sugano Y, Keetels M, Vroomen J. Adaptation to motor–visual and motor–auditory temporal lags transfer across modalities. *Exp Brain Res* 2010; **201**:393–399.
6 Yamamoto K, Kawabata H. Temporal recalibration in vocalization induced by adaptation of delayed auditory feedback. *PLoS One* 2011; **6**:e29414.
7 Fujisaki W, Shimojo S, Kashino M, Nishida S. Recalibration of audio-visual simultaneity. *Nat Neurosci* 2004; **7**:773–778.
8 Spence C. Prior entry: attention and temporal perception. *Attention and time*. Oxford, UK: Oxford University Press; 2010. pp. 89–104.
9 Parducci A. Category judgment: a range-frequency model. *Psychol Rev* 1965; **72**:407–418.
10 Vroomen J, de Gelder B. Temporal ventriloquism: sound modulates the flash-lag effect. *J Exp Psychol* 2004; **30**:513–518.
11 Cunningham DW, Billock VA, Tsou BH. Sensorimotor adaptation to violations of temporal contiguity. *Psychol Sci* 2001; **12**:532–535.
12 Stetson C, Cui X, Montague PR, Eagleman DM. Motor–sensory recalibration leads to an illusory reversal of action and sensation. *Neuron* 2006; **51**: 651–659.
13 Sugano Y, Keetels M, Vroomen J. The build-up and transfer of sensorimotor temporal recalibration measured via a synchronization task. *Front in Psychol* 2012; **3**:246.
14 Wolpert DM, Ghahramani Z, Jordan MI. An internal model for sensorimotor integration. *Science* 1995; **269**:1880–1882.
15 Miall RC, Wolpert DM. Forward models for physiological motor control. *Neural Net* 1996; **9**:1265–1279.
16 Tsakiris M, Haggard P, Franck N, Mainy N, Sirigu A. A specific role for efferent information in self-recognition. *Cognition* 2005; **96**:215–231.
17 Blakemore SJ, Frith CD, Wolpert DM. Spatio-temporal prediction modulates the perception of self-produced stimuli. *J Cogn Neurosci* 1999; **11**:551–559.
18 Sugita Y, Suzuki Y. Audiovisual perception: implicit estimation of sound-arrival time. *Nature* 2003; **421**:911.
19 Lester M, Boley J. The Effects of Latency on Live Sound Monitoring: Audio Engineering Society Convention Paper. Presented at the 123rd Convention 2007, 5–8 October 2007, New York, NY, USA, 7198.