Equal Interval Tapping Synchronized with Metronome Ticking or Light Blinking

Masashi Yamada, Noboru Masuda and Hiromi Hashimoto

Faculty of Arts
Osaka University of Arts, Japan
{m-yamada, rs03040, m8897}@osaka-geidai.ac.jp

Abstract

Equal interval tapping synchronized with a ticking sound or blinking light was performed, and the temporal fluctuation in the tapping was analyzed with Fourier analysis. The results show that the temporal control of tapping with sound is much better than the case of tapping with a blinking light. Equal interval tapping with no external equal-interval cues was also performed and the temporal fluctuation was analyzed. The results show that the temporal control without external equal-interval cues is governed by a 20-tap memory mechanism: The information of the preceding intervals of 20 taps is preserved and used for determining the interval of the present tap. Comparison of the fluctuation between tapping with cues and tapping without cues suggested that, in the case of tapping with external equal-interval cues, the temporal control is governed by a combination of the 20-tap memory mechanism and a cue-synchronized mechanism: The timing of each tap performed by a human can be described as exact equal-interval timing plus a value from a Gaussian noise. Therefore, the inter-onset interval of the tapping is characterized by the first order differential of the Gaussian noise. In the case of tapping with a blinking light, the variance of the Gaussian noise is quite larger than the case of tapping with a ticking sound. This increase of the variance of the noise worsens the temporal control in tapping with a blinking light, compared to the case with a ticking sound.

1. Introduction

In recent years, our daily lives are filled with audio-visual stimuli, *e.g.*, TV programs, animated movies, music videos, *etc*. Moreover, there are also many kinds of systems for which we need to synchronize our behavior to audio-visual stimuli. For example, in using video-game software for Sony Play Station 2, “Dram-mania” or “Dance Dance Revolution” it is required to hit a pad or step on a floor-pad in timing synchronized with the audio-visual stimuli. However, players often feel asynchrony even if the game-software gives high scores and *vice versa*. This is at least partly because the perception of synchronization between audio-visual stimuli and our body movement is not correctly reflected in the game-software. To produce software that fits a player’s feeling of “synchrony,” not only the perception of synchrony between audio and visual stimuli but also synchronization mechanisms of our body movements to the audio-visual stimuli has to be clarified.

In recent years, study of the perception of synchrony between audio and visual stimuli has become a subject of renewed interest [1]. Among these studies, Patel and his colleagues conducted an experiment, in which subjects tapped a key in synchronized timing with a equal-interval sound stimulus or light-blinking stimulus [2]. They showed that it was difficult to synchronize the tapping with a blinking light than with a ticking sound. However, they did not clarify the synchronization mechanisms of the tapping with the stimuli or why the synchronization with a light blinking was poorer than with a ticking sound. During the tapping experiment by Patel *et. al.*, subjects listened to the tapping sound they produced. Therefore, in the tapping condition with sound, the subjects could compare the timing of their tapping and of the sound stimulus in the same auditory modality. However, in the case of tapping with a blinking light, they had to make comparisons between different modalities. We can hypothesize that in the case of sound tapping the timing comparison in the same modality produced excellent synchronization.

In the present study, experimental conditions were prepared such that a pink noise masked the tapping sounds in order to test the hypothesis described above. Moreover, by analyzing the temporal fluctuation in the tapping for various conditions, the temporal control mechanisms of tapping synchronized with equal-interval ticking-sounds or blinking-lights were estimated.

2. Experimental Method

Patel *et. al.* used a small light on MIDI equipment as a blinking-light stimulus [2]. In the present study, equal-interval blinking-light stimulus was constructed as follows: In NTSC video-format (29.97 frames / sec), a pattern of 13 blank frames followed by two consecutive white frames was repeated. Therefore, the inter-onset interval of the white screens was 500.5 ms. In the present study, this equal-interval blinking-light is called ‘Flash’ stimulus. A preliminary experiment showed that people can easily synchronize their taps with the Flash stimulus than with the small light blinking used by *Patel* *et. al*. This Flash stimulus was presented through a Sony GLASSTRON 2 head-mount display-screen. A 6-ms tone-burst of 4 kHz was presented at 74 dB SPL every 500.5 ms as the Sound stimulus. This Sound stimulus was presented through SONY CD 900-ST headphones. The subjects also monitored their tapping sound through the same headphones at about 72 dB SPL under some conditions.

Under the other conditions, the tapping sound was not presented to the subjects, and a 79-dB SPL pink-noise was presented through the headphones to mask the external tapping sound. This low-pitch noise perfectly masked the tapping sound but allowed the 4-kHz ticking sound to be heard clearly.

To clarify the temporal control mechanism in the case of the stimulus-synchronization tapping, experimental conditions without external equal-interval cues were also employed. In the present study, these conditions are called ‘Free’ tapping as opposed to ‘Synchronized’ tapping. Under the Synchronized conditions, the Sound or Flash stimulus was presented during tapping, under the Free conditions, the Sound or Flash stimulus was presented for 20 sec before starting the tapping, and no stimulus was presented while tapping.

By combinations of [Sound / Flash], [Synchronized / Free] and [with / without Masking Noise], the following seven experimental conditions were constructed.

1. Sound Synchronized
2. Flash Synchronized
3. Sound Synchronized (Noise)
4. Flash Synchronized (Noise)
5. Sound Free (Noise)
6. Flash Free (Noise)
7. Sound Free

Five (three male and two female) students, aged 21-24 years from the Osaka University of Arts participated as subjects. All subjects had experience in playing piano and other instruments, but only at intermediate levels. The subjects were instructed to tap a 40 cm x 50 cm aluminum board with their right middle fingers, synchronized to the audio or visual stimuli under the Synchronized conditions. For Flash stimuli, they were requested to tap synchronized to the onset of the white screens. In the case of Free tapping, they were requested to tap at an equal interval, identical to the preceding 20 sec of Flash or Sound stimulus.

One trial of tapping consisted of 800 taps. Five successive trials for each of the seven conditions were performed but with the order of the conditions randomized for each subject. A 3-min rest period separated the trials and a 20-min rest period separated the conditions. Small speakers attached to the aluminum board converted the pressure of the fingers to a voltage, and a computer system preserved this voltage data as a function of time. In the conditions of without Noise, this voltage data was also used to allow subjects to monitor their tapping.

3. Results and Discussion

Inter-onset intervals (IOIs) were obtained from the voltage data and plotted as a function of tap order. The IOI was not generally stable during the initial 100 taps, so these 100 taps were removed and the IOI waveform of the following 600 taps was used for the analysis. This waveform was decomposed into Fourier components by DFT with Hanning window. The power spectra of the temporal fluctuation were averaged out for every half-octave range and then averaged out for each condition. The resulting power spectra are shown in Fig. 1. The frequency, f is shown in cycles for 600 taps, therefore the period, p [taps] for the frequency can be calculated by $p = 600 / f$. The difficulty of control for the fluctuation of a given frequency is indicated by the power of that frequency.

In Fig. 1, the power for the Flash Synchronized condition is larger than for the Sound Synchronized condition, especially in the low frequency region. This indicates that, for a ling period, the temporal control of synchronization with the light more difficult than with the sound. This is consistent with Patel *et. al.* [2]. The power for the Flash Synchronized condition is even larger than the Sound Synchronized (Noise) condition, in which subjects could not compare the timing between their own tapping and the stimulus in the auditory system. This means that the hypothesis, “the timing comparison in the same modality produced excellent synchronization in the case of sound tapping,” must be rejected.

The power spectrum for the Sound Free condition indicates a critical phenomenon: the power decreased as the frequency increased in the low frequency region, whereas the power is almost constant for the high frequency region. The critical frequency, the boundary between the two regions, is observed at approximately 30 cycles. This means that the temporal control is excellent for a short period of less than 20 taps, but for a period over 20 taps, the controllability decreases as the period increases. This is consistent with the results of our previous Free tapping experiments [3,4,5]. These previous studies clarified that the temporal fluctuation in Free tapping can be simulated by the following 20th order auto-regression model.

$$IOI(i) = \sum_{j=1}^{20} a_i \cdot IOI(i-j) + \xi_i \qquad (1)$$

where, ξ_i indicates a Gaussian noise. In other words, the temporal control of Free tapping is governed by a 20-tap memory mechanism. In Fig.1, you can see that the power for the Sound Free (Noise) and the Flash Free (Noise) conditions is larger than that for the Sound Free condition, even though the spectral feature is the same. This means that the variance of the Gaussian noise ξ_i increases when, in the same auditory system, the comparison of timing is not available.

In the low frequency region for the Sound Synchronized condition, the power increases proportionally as frequency increases. This illustrates the power spectral feature of a first order differential of a Gaussian noise: In the case of the Synchronized tapping, the subjects can use a rigid timing cue of outer stimuli. The onset time of the *i*-th tick or blink is,

$$t(i) = i \cdot k \tag{2}$$

where *k* indicates the inter-onset interval of the ticks or blinks. Although subjects attempted to tap at the same timing with the cues, their performance included some degree of deviation, *i.e.*, the onset time of the *i*-th tap is illustrated as follows:

$$t(i) = i \cdot k + \zeta_i \tag{3}$$

where ζ_i shows a Gaussian noise. Then the IOI performed by a subject can be illustrated by the first differential of the Gaussian noise.

$$t(i) = k + \left(\zeta_i - \zeta_{i-1}\right) \tag{4}$$

The power spectral feature for the Sound Synchronized condition indicates that the cue-synchronized mechanism is active for a long period of the fluctuation. However, the power spectrum for the Sound Synchronized condition coincides with the power spectrum for the Sound Free condition in the high frequency region. In the case of Synchronized tapping, not only the cue-synchronized mechanism but also simultaneously the 20-tap memory mechanism can be utilized. Consequently, the Sound Synchronized condition shows the power spectral feature of the combination of the two mechanisms. This spectral feature is consistent with our previous tapping experiment synchronized with equal-interval sound ticking in various speeds [6].

Similarly, spectra for the Sound Synchronized (noise), Flash Synchronized and Flash Synchronized (noise) conditions show the combined features of the two mechanisms. This means that both the cue-synchronized mechanism and the 20-tap memory mechanism are both active in the tapping synchronized with equal-interval stimuli.

In the low frequency region, the vertical positions of the Synchronized tapping is as follows:

[Sound Synchronized]
< [Sound Synchronized (noise)]
< [Flash Synchronized]
≈ [Flash Synchronized (noise)] (5)

This correlation suggests that the variance of the Gaussian noise, ζ_i in equation (4) for the case of synchronization with the blinking light is quite larger than with the ticking sound. Moreover, this correlation also suggests that the lack of the timing comparison between the subjects' tapping and outer cues in the same auditory modality slightly increases the variance of the Gaussian noise, ζ_i.

Because of the larger variances in both the Gaussian noises, ξ_i and ζ_i, the temporal control in tapping synchronized with the blinking light is poorer than that with the ticking sound.

In a previous study, it was shown that the temporal control in Free tapping performed by a single-finger does not vary with musical training [4]. Therefore, the control mechanisms shown in the present study will take place not only in musicians but also in people in general.

4. Conclusions

In the present study, it was confirmed that the temporal control in tapping synchronized with an equal-interval blinking light is poorer than with a ticking sound. Moreover, it was shown that in the tapping synchronized with an equal-interval stimulus, both the cue-synchronized mechanism and the 20-tap memory mechanism are utilized.

5. Acknowledgement

The authors wish to thank Reiko Kurai and David Reid for their enthusiastic assistance. The present study was supported in part by an Grant-in-Aid for Research, from Art Institute, Osaka University of Arts.

6. References

[1] Kohlrausch, A., "Perceptual sequences of asynchrony in audio-visual stimuli", *IPO Annual Progress Report* 35: 140-149, 2000.

[2] Patel, A. D., Iversen, J. R., Chen, Y. and Repp, B. H., "Tapping to the beat of auditory and visual rhythmic sequences", *Proc. 7th Int'l Conf. Music Percept. & Cogn.*, Sydney: CD-ROM, 2002.

[3] Yamada, M., "Temporal control mechanism in equaled interval tapping", *Applied Human Science* 15: 105-110, 1996.

[4] Yamada, M. and Tsumura, T., "Do piano lessons improve basic temporal controllability of maintaining a uniform tempo", *J. Acoust. Soc. of Jpn.*, (*E*) 19: 121-131, 1998.

[5] Yamada, M. and Yonera, S., "Temporal control mechanism of repetitive tapping with simple rhythmic patterns", *Acoust. Sci. & Tech.*, 22(3): 245-252, 2001.

[6] Yamada, M., "Temporal control mechanism in equal interval tapping with and without metronome ticking," *The 15th Triennial Congress of the International Ergonomics Association (Seoul) Proceedings 1*: 111-114, 2003.

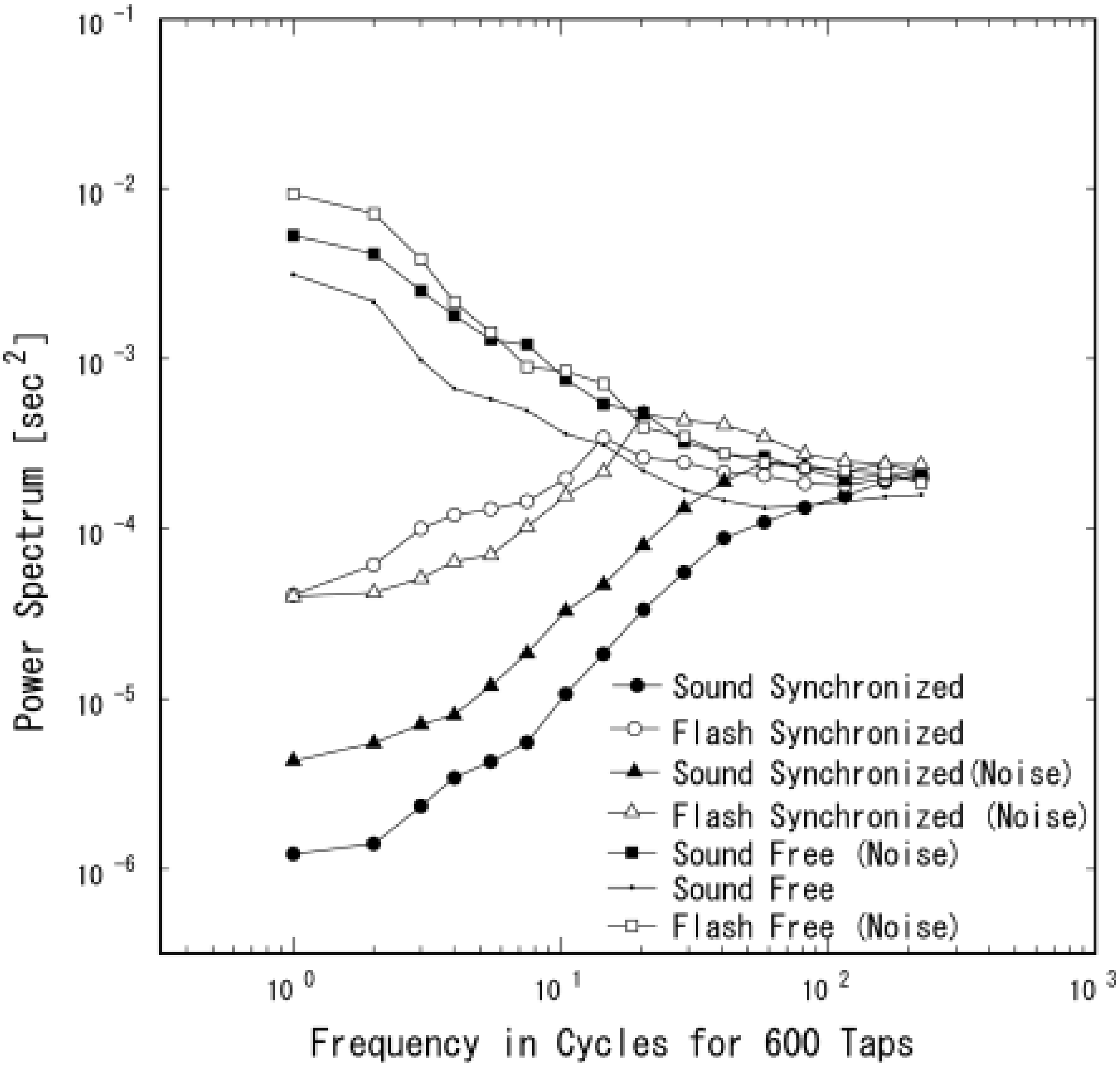

Figure 1: Power spectra of temporal fluctuation in equal interval tapping under various conditions