Biol. Cybern. 91, 388–395 (2004)
DOI 10.1007/s00422-004-0521-3

Biological Cybernetics

Sequential grouping of tone sequence as reflected by the mismatch negativity

Shin'ichiro Kanoh, Ryoko Futami, Nozomu Hoshimiya

Department of Electronic Engineering, Graduate School of Engineering, Tohoku University, Aoba-yama 6–6–05, Sendai, 980–8579 Japan

Received: 8 April 2004 / Accepted: 13 September 2004 / Published online: 19 November 2004

Abstract. The human sequential grouping that organizes parts of tones into a group was examined by the mismatch negativity (MMN), a component of event-related potentials that reveals the sensory memory process. The sequential grouping is accomplished by the combinations of some factors, e.g., temporal and frequency proximity principles. In this study, auditory oddball stimuli in which each of the stimuli consisted of series of tone bursts, were applied to the subjects, and the MMN elicited by the deviation of the frequency of the last tone in the stimulus was investigated. The relationship between the expected phenomena of sequential grouping of tones and observed magnitudes of MMN was evaluated. It was shown that the magnitudes of MMN changed according to the configuration (number of tones, frequency) of tone sequence to be stored. This result suggested that the sequential grouping of presented tones was achieved on the preattentive auditory sensory memory process. It was also shown that the relative change of MMN magnitudes corresponded to the conditions of sequential grouping, which had been proposed by the auditory psychophysical studies. The investigation of MMN properties could reveal the nature of auditory sequential grouping.

1 Introduction

In the process of human auditory information processing, it is generally thought that much of the presented information is handled not only serially but also in parallel using "buffer" memory.

The auditory sensory memory is a memory system that stores auditory information within a short period of time (a few seconds).

The property of the sensory memory process can be revealed by the mismatch negativity (MMN), a component of event-related potentials (ERPs) (Näätänen 1992). MMN is closely related to the auditory discrimination and storage functions in the sensory memory (Tiitinen et al. 1994; Kanoh et al. 1996). The neuronal generators of the auditory MMN are thought to be located mainly in the superior temporal cortex (Alho 1995).

The MMN studies showed that the successive auditory information was integrated into meaningful entities and was represented as a unitary event in the sensory memory system (temporal integration) (Tervaniemi et al. 1994; Yabe et al. 1997; Kanoh et al. 2001). It was also shown that basic information processing was applied to the stored information during the auditory sensory memory process (Näätänen et al. 2001).

Sequential grouping (Darwin 1997), which binds some successive information into one entity, is one of the functionalities of such a primitive processing. Like the perceptual grouping in the visual system, sequential grouping is accomplished by the combinations of some factors, e.g., temporal and frequency proximity factors (Bregman 1990). A computational model of sequential grouping of auditory stimuli has been proposed in which such auditory psychophysical findings were taken into account (Baumann 2000).

However, no physiological evidence that would validate these factors on sequential grouping has been reported. Observing the neuronal correlates of sequential grouping will help to determine the principle of auditory information processing and to hypothesize about the neuronal representation of temporal tone sequence, e.g., how the order and frequency of the successive tones are encoded.

In this study, the sequential grouping in human auditory cortex was examined by MMN, a component of ERPs that revealed the sensory memory process. By changing the configuration (number of tones, frequency) of tone stimulus that consisted of some tone bursts in MMN measurement, the relationship between the amplitudes of elicited MMN and the expected grouping of these tones was investigated. It was shown that the sequential grouping of successive tones was achieved during the preattentive auditory sensory memory process, and the properties of the sequential grouping were revealed by the MMN activities. These results showed the neurophysiological evidence of preattentive auditory sequential grouping and will help

Correspondence to: S. Kanoh
(e-mail: kanoh@ecei.tohoku.ac.jp)

in discovering the cortical representation of auditory stimuli and will contribute to a computational theory of auditory information processing.

2 Background and working hypothesis

2.1 Mismatch negativity

The MMN, an ERP component, is elicited when some frequently presented ("standard") stimuli are randomly replaced by infrequent ("deviant") stimuli with some different attributes from the standard ones (oddball paradigm) (Näätänen 1992).

The MMN is a measure of preattentive sensory memory properties and is closely related to the auditory discrimination and storage functions. The important property for the present study is that the MMN magnitude increases as a function of the extent of deviancy (difference between the physical attributes in the standard and the deviant stimuli) (Näätänen 1992; Tiitinen et al. 1994; Kanoh et al. 1996).

2.2 Temporal integration

In the auditory sensory memory, the function called temporal integration, which integrates successive information into some entities, has a functional role for the perception of auditory events (Bregman 1990).

The sensory information is integrated and stored in the sensory memory, where primitive processing is applied to the stored information (Näätänen et al. 2001).

The width of the temporal window of integration was evaluated by several studies on human MMN, e.g., $150 \sim 170$ ms (Yabe et al. 1997), ~ 200 ms (Winkler et al. 1998), $200 \sim 250$ ms (Tervaniemi et al. 1994). The authors determined the width of the temporal windows to be $300 \sim 400$ ms (Kanoh et al. 2001), which was longer than that reported by other authors. The authors also investigated the attention effect on the temporal integration process, and it was shown that the attention could prolong the period of temporal window of integration and that the MMN activities with attention had more complex distributions in the frontal area than in an ignored condition (Kanoh et al. 2001).

2.3 Sequential grouping

The sequential grouping, which organizes or binds some successive information into one entity, is one of the functions of the primitive processing of sensory information. In auditory modality, this function serves as a foundation for the segmentation of speech sounds, and it is believed that the auditory sequential grouping is achieved in the process of temporal integration in the auditory sensory memory (Näätänen et al. 2001).

Like the perceptual grouping in the visual system, it is hypothesized that sequential grouping is accomplished by the combination of various factors (Bregman 1990). Some of the factors are as follows.

- *Proximity*: sounds that are near to one another in time (temporal proximity) or in frequency (frequency proximity) will group with one another.
- *Good continuation*: sounds with good continuation (e.g., frequency change) will tend to bind together.

The properties of sequential grouping have also been investigated by ERP studies. The relationship between the sequential grouping due to the temporal proximity factor and the observed MMN was studied, and it was shown that the sequential grouping could occur in sensory memory preperceptually without willful attention (Winkler et al. 2001; Sussman et al. 1998; Schröger et al. 1996; Winkler and Schröger 1995). In addition, it was suggested that the auditory streaming effect due to temporal proximity and frequency similarity in the multiple-tone stream (i.e., frequency distribution in the series of tones) was also achieved in the preattentive sensory memory process (Shinozaki et al. 2000; Sussman et al. 1999).

But the relationship between the temporal proximity factor and the degree of sequential grouping has not been evaluated by MMN experiments. Furthermore, evidence for sequential grouping as a result of the interaction between temporal and frequency factors in the sensory memory process was not demonstrated directly by the previous MMN studies.

2.4 Working hypothesis

In this study, preperceptual sequential grouping was evaluated by observed MMN magnitudes by assuming the following hypothesis on the relationship between MMN amplitudes and sequential grouping phenomena.

Consider an experiment with oddball tone sequence in which each of the stimuli consists of several tones and only the frequency of the last tone (deviated tone) is different between the standard and deviant stimulus.

The MMN is generated by comparing the memory trace of the standard stimuli and the neuronal representation of an incoming deviant stimulus.

Let us assume that the neuronal activations of each grouped auditory stimuli are bound together as a unitary representation on the auditory cortex.

Incorporating this assumption, when the last tone is grouped with other tones, MMN will be elicited by comparing two unitary neuronal representations of the standard and deviant stimuli, both of which encode the same set of grouped tones, including the deviated tone.

In this case, the higher the number of tones in the grouped unitary entity, the smaller the relative change of attributes in the corresponding neuronal representations of the standard and deviant stimuli. As described in Sect. 2.1, MMN magnitude decreases with less deviancy in the standard and deviant stimuli. A relative difference between the neuronal representations of two sensory events in the sensory memory can be evaluated by MMN measurements.

The present study was then informed by the following working hypothesis: "When the number of tones grouped together with the deviated tone increases, the elicited MMN magnitude will decrease."

In the present study, the two experiments were conducted to determine whether the sequential grouping was achieved in the sensory memory process and to find out what "rules" for sequential grouping are applied to various configurations of tone series.

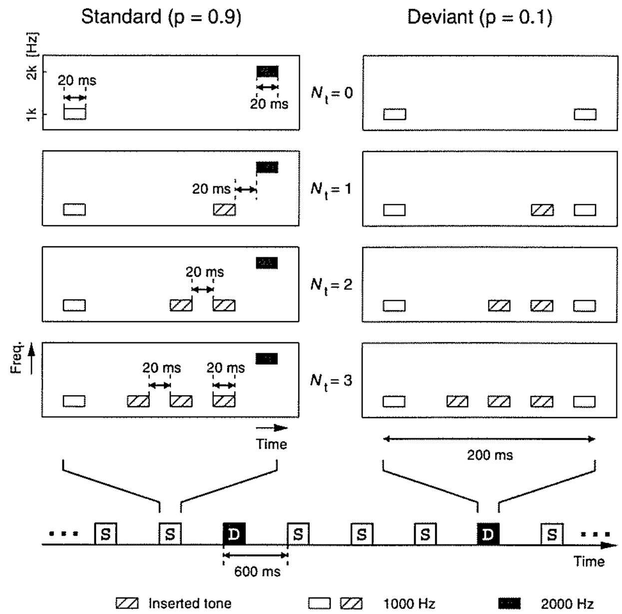


Fig. 1. (Experiment 1) Schematic diagram of stimulus on each condition

3 Experiment 1: Temporal proximity factor on sequential grouping

At first, the relationship between the number of grouped tones due to temporal proximity and the elicited MMN amplitude was examined.

In this experiment, the oddball sequence in which each stimulus consisted of several tones (first tone, inserted tones, and last tone) was used. The relationship between the number of inserted tones and elicited MMN amplitude elicited by the deviation of the frequency of the last tone was investigated, and the tendency of sequential grouping by temporal proximity was discussed.

3.1 Method

The subjects were six male volunteers with normal hearing. Auditory stimuli were applied to both ears by headphones. Each subject, in an electromagnetically shielded room, was instructed to read a self-selected book and to ignore the presented auditory stimuli (reading condition).

The schematic diagram of stimuli is shown in Fig. 1. Each stimulus consisted of several tone bursts (20 ms, 60 dB SL, rise/fall time 5 ms), i.e., the first tone, last tone, and tone(s) inserted between the first and last tones. The total duration of each stimulus was 200 ms. The intervals between the last inserted tone and the last tone, and between inserted tones were set to 20 ms. Interstimulus interval (onset to onset) was set to 600 ms.

In the standard stimulus ($p = 0.9$), frequencies of the first, inserted, and last tones were 1000, 1000, and 2000 Hz, respectively. And in the deviant stimulus ($p = 0.1$), only the frequency of the last tone was deviated to 1000 Hz.

In separate blocks, the number of inserted tones (N_t) was set to 0, 1, 2, and 3. The relationship between N_t and elicited MMN by the infrequent change of the frequency of the last tone was observed. The goal of this experiment was to evaluate the number of tones grouped with the last tone.

The EEG signal was recorded with Ag–AgCl electrodes at two locations (Fz, F4) and referred to linked earlobe (A1, A2) with a forehead ground. The vertical electro-oculogram (EOG) was also obtained using bipolar recording from the outer canthus of the left eye. These signals were amplified (gain $\times 500$) and recorded with a bandpass filter setting of 0.1–100 Hz and analyzed -100 to 500 ms from the onset of the first tone (sampling rate 1,000 Hz with antialiasing filter of 200 Hz).

The signals were referred to the mean amplitudes in the prestimulus baseline ($-50 \sim 0$ ms). EEG epochs in which recorded amplitudes exceeded $\pm 100\,\mu$V were automatically rejected. The accepted EEG epochs were separately averaged for each stimulus type. The responses to the standard stimuli just after the deviant ones were excluded from the analysis.

After the application of low-pass filter (30 Hz) to the averaged waveform, the amplitude of the negative peak in the difference waveform (deviant–standard) whose latency was 100–250 ms was evaluated. The data were statistically tested using one-way analysis of variance (ANOVA) with the factors of stimulus type ($p < 0.05$) to determine whether the ERPs associated with the standard stimuli were significantly different than those associated with the deviant stimuli at the MMN peak amplitude.

3.2 Results

Figure 2 shows the averaged responses to the standard and deviant stimuli and their difference waveforms in Experiment 1 (subject A, electrode Fz). From this subject, MMN was elicited when the number of inserted tones N_t was 0, 1, and 2, but no significant MMN was observed at $N_t = 3$.

The normalized amplitudes of the observed MMN on all subjects are shown in Fig. 3. On all subjects, the elicited MMN amplitude taken from Fz was larger than that from F4. The properties of the observed MMN amplitudes could be divided into the following two groups:

- *Group 1*: The observed MMN amplitude was smaller for larger N_t (subjects A–C).
- *Group 2*: The MMN amplitudes at $N_t = 0$ was the largest, and at $N_t = 1$–3 almost the same amblitudes were observed (subjects D–F).

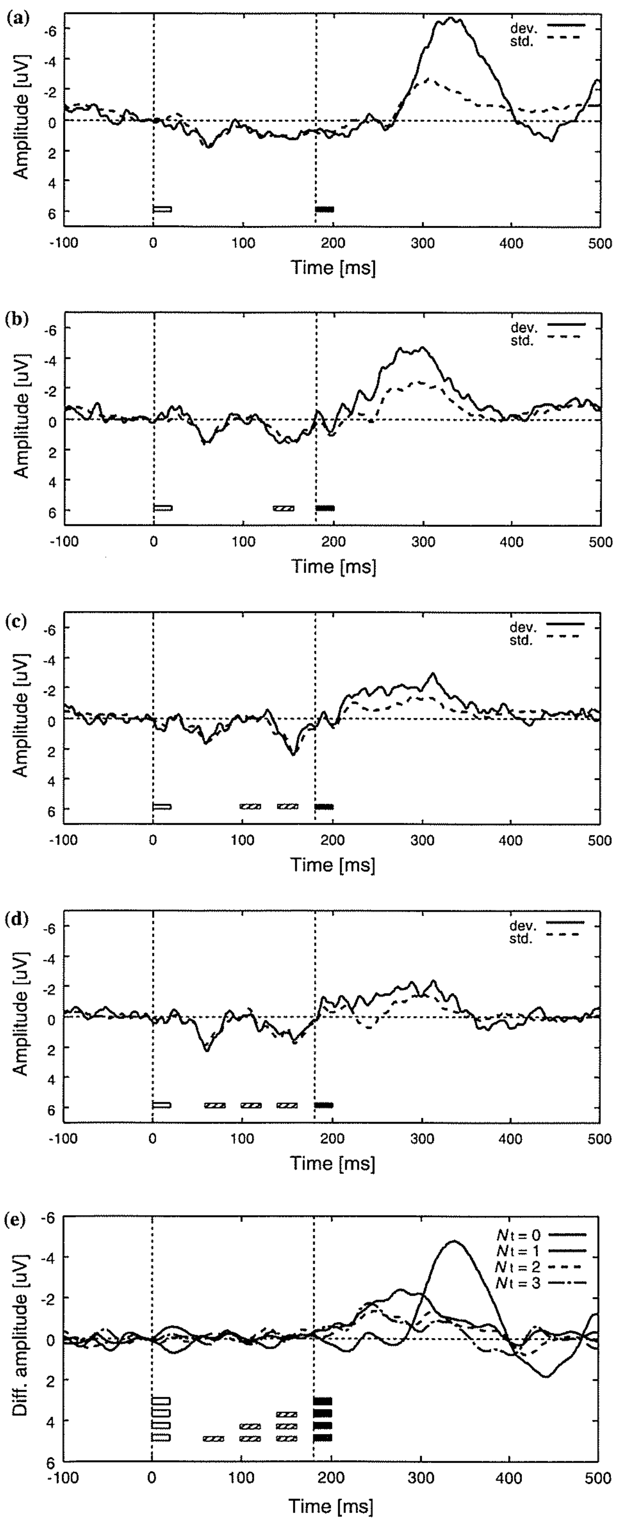


Fig. 2. (Experiment 1) Averaged standard (*dashed*) and deviant (*solid*) waveforms (**a–d**) and their difference waveforms after low-pass filtering of 30 Hz (**e**) (subject A, electrode Fz)

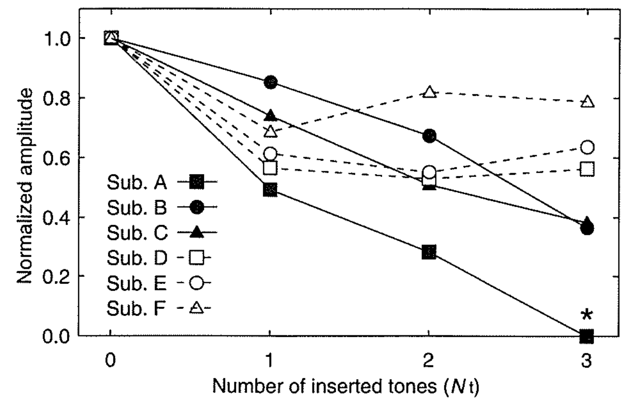


Fig. 3. (Experiment 1) Relationship between N_t and normalized MMN amplitude on all subjects (electrode Fz). From subject A, no significant MMN was observed on $N_t = 3$ (marked *)

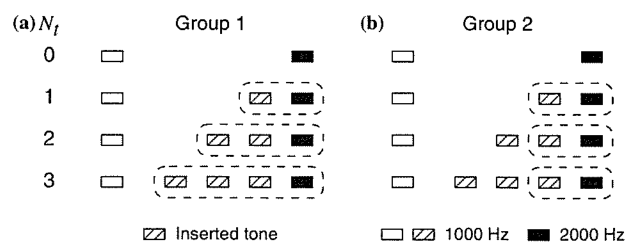


Fig. 4. (Experiment 1) Schematic diagram of possible hypothesis. Temporal configuration of the standard stimulus on each condition is shown

3.3 Discussion

Although the interval between the last inserted tone and the last tone was fixed at 20 ms, the MMN amplitudes were changed depending on the number of inserted tones (N_t). This result could be explained on the assumption of the working hypothesis described in Sect. 2.4 as follows.

On group 1, the MMN amplitude decreased monotonically when N_t increased. This result could be explained by the hypothesis stating that all the inserted tones were grouped together with the last tone by the factor of temporal proximity during each of the experiments (Fig. 4a). And on group 2, the largest MMN amplitude was elicited at $N_t = 0$, but no clear differences of MMN amplitudes were observed at $N_t \neq 0$. If the subjects in group 2 tended to group only the last two tones (the last inserted tone and the last tone), such a result would agree with the present hypothesis (Fig. 4b).

The different phenomena on groups 1 and 2 can then be explained by the different width of the time window for grouping based on temporal proximity: longer than 140 ms (group 1), 60–80 ms (group 2).

From Fig. 2 it was also found that both peak and onset latencies were shorter when N_t was larger. And the duration of MMN was shorter when $N_t = 0$ (120 ms) but longer when $N_t = 1, 2,$ and 3 ($200 \sim 250$ ms). These results could be observed on the other subjects. Such an ERP component with shorter latency might be related to components other than the MMN.

4 Experiment 2: Frequency proximity factor on sequential grouping

Next, the relationship between the tone frequency and sequential grouping was evaluated by the elicited MMN amplitude.

In this experiment, the oddball sequence in which each stimulus consisted of three tones was used, and the relationship between frequency of the second tone and elicited MMN amplitude was examined.

As described in Sect. 2.3, temporal and frequency proximity are known as the factor of proximity for auditory sequential grouping. An attempt was made to control the tendency of sequential grouping, i.e., whether the second tone tended to be grouped with the first tone or the last tone, by changing the frequency of the second tone.

4.1 Method

The subjects were five male volunteers with normal hearing. Auditory stimuli were applied to both ears by headphones. Each subject, in an electromagnetically shielded room, was instructed to read a self-selected book and to ignore the presented auditory stimuli (reading condition).

The schematic stimulus diagram is shown in Fig. 5. Each stimulus consisted of the first tone (duration 40 ms), the second tone (20 ms), and the third tone (40 ms), each of which was 60 dB SL tone burst (rise/fall time 5 ms). Total duration of the stimulus was 200 ms. The interval between the second and third tones was set at 30 ms. Interstimulus interval (onset to onset) was set at 600 ms.

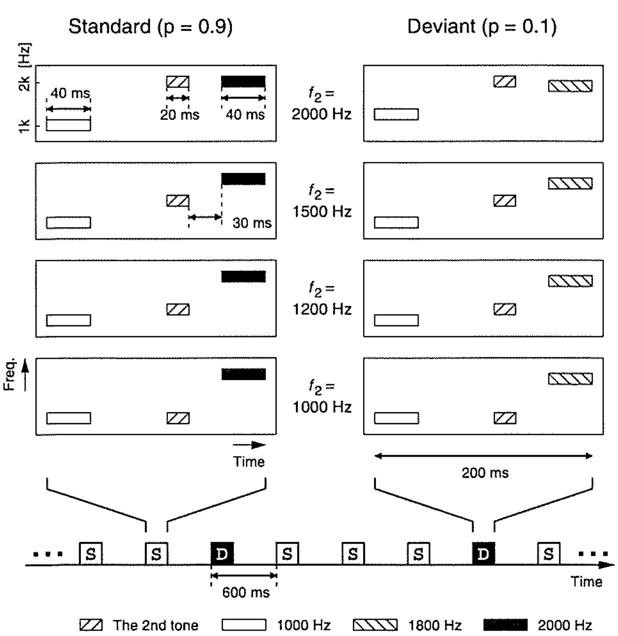


Fig. 5. (Experiment 2) Schematic diagram of stimulus on each condition

In the standard stimulus ($p = 0.9$), the frequencies of the first and third tones (f_1, f_3) were 1000 and 2000 Hz, respectively. In the deviant stimulus ($p = 0.1$), only the frequency of the third tone f_3 was deviated to 1800 Hz.

In separate blocks, the frequency of the second tone (f_2) was set at 1000, 1200, 1500, and 2000 Hz. The relationship between f_2 and the MMN elicited by the infrequent change of the frequency of the third tone was observed. The EEG recording and analysis procedure was the same as in Experiment 1.

4.2 Results

Figure 6 shows the averaged responses for the standard and deviant stimuli and their difference waveforms in Experiment 2 (subject A, electrode Fz). From this subject significant MMN was elicited on each condition.

The normalized amplitudes of the observed MMN for all subjects are shown in Fig. 7. On all subjects the elicited MMN amplitude taken from Fz was larger than that from F4. The elicited MMN amplitude was largest when f_2 was 1000 Hz on all subjects, but the tendency of MMN amplitudes could be divided into the following two groups:

- *Group 1*: The smallest amplitude was observed when f_2 was 1500 Hz, and the amplitudes on $f_2 =$ 1200 and 2000 Hz were almost the same (subjects A, B).
- *Group 2*: The smallest MMN amplitude was observed when $f_2 =$ 1200 Hz, but there was no significant difference between $f_2 =$ 1200, 1500, and 2000 Hz (subjects D–F).

4.3 Discussion

In Experiment 2, the frequency of the second tone f_2 was changed to control the contribution of the frequency proximity factor to sequence grouping. Taking only the factor of frequency proximity into account, it was expected that the second tone would tend to organize together with the first tone when $f_2 = f_1$ (1000 Hz). But when $f_2 = f_3$ (2000 Hz in the standard stimulus), the expected partner to be grouped with would be the third tone.

On the other hand, when only the factor of temporal proximity was taken into account, the second tone tended to be grouped together with the third tone rather than the first tone in this case (intervals between the second and first tone and between the second and third tone were 70 and 30 ms, respectively).

In the former case, the two processes of sequential grouping based on temporal and frequency proximity factors contradicted each other because the temporal position of the second tone was closer to the third tone than to the first tone.

In such a situation, it might be possible for the two factors to be combined (competition or cooperation) to solve the contradiction. By assuming the working hypothesis on this study, MMN amplitude is expected to decrease monotonically when f_2 increases from 1000 Hz to 2000 Hz because the second tone tends to be grouped together with

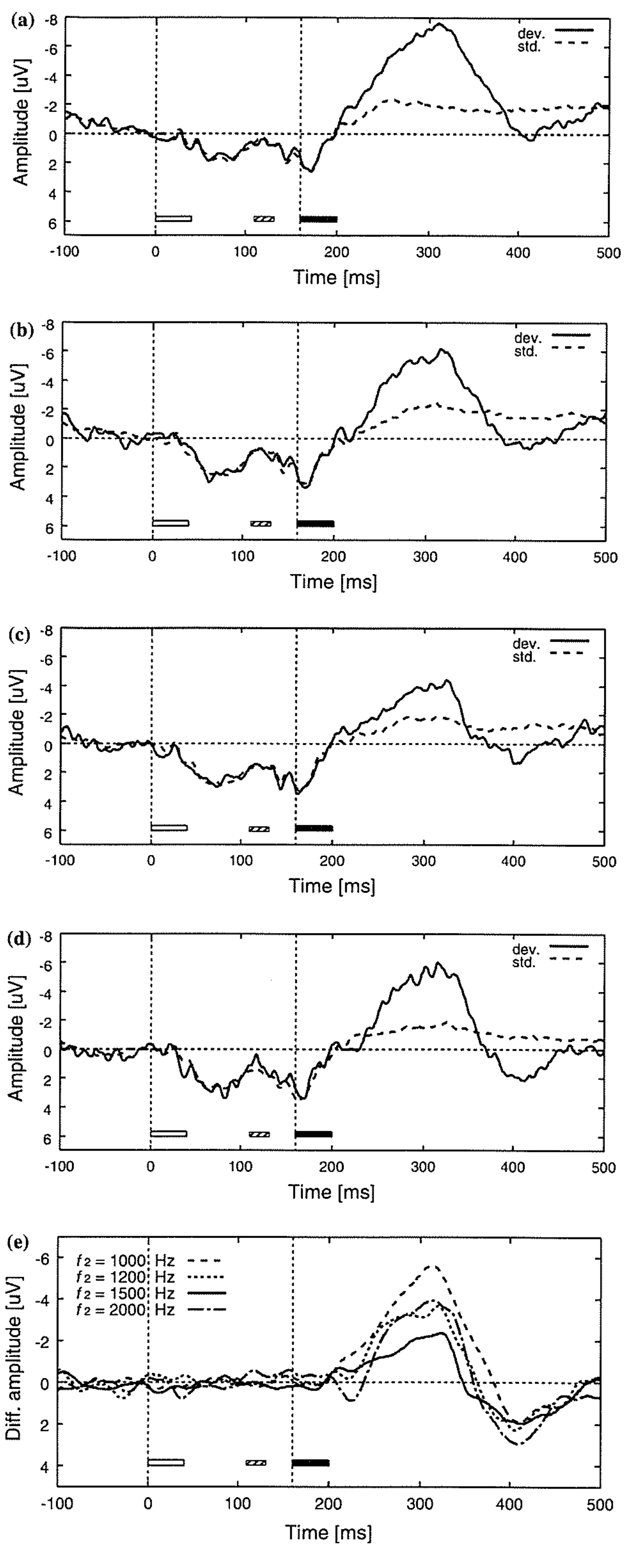


Fig. 6. (Experiment 2) Averaged standard (*dashed*) and deviant (*solid*) waveforms (**a–d**) and their difference waveforms after low-pass filtering of 30 Hz (**e**) (subject A, electrode Fz)

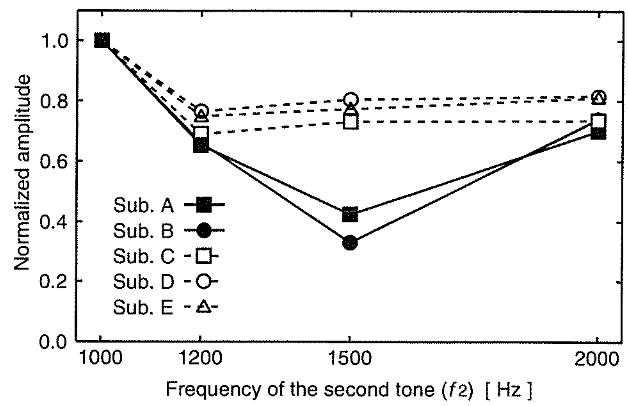


Fig. 7. (Experiment 2) Relationship between f_2 and normalized MMN amplitude on all subjects (electrode Fz)

the third tone easier when f_2 approaches the frequency of the third tone f_3.

But on two subjects (subjects A and B) out of five, the minimum MMN amplitude was observed when $f_2 = 1500$ Hz. This phenomenon could be explained by introducing the "good continuation" as one of the factors in addition to the factors of temporal/frequency proximity for sequential grouping of successive tones.

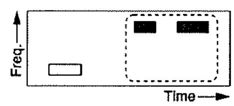
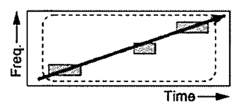
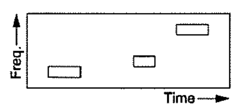
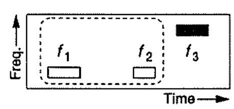
In Experiment 2, it should be noted that the frequencies of tones in the standard stimulus formed specific ratios of musical scale when $f_2 = 1500$ Hz, i.e., $f_2 : f_1 = 3:2$ constituted a musical fifth and $f_3 : f_2 = 4:3$ corresponded to a musical fourth. But as all the subjects in this study had no special musical skills, it is thought that the relative decreases of MMN amplitudes observed from subjects A and B when $f_2 = 1500$ Hz are not due to sensitivity to the musical period.

Table 1 is the schematic diagram of the present hypothesis. When f_2 was 1000 and 2000 Hz, the second tone tended to be grouped together with the first and the third tone, respectively. And when f_2 was 1500 Hz, the frequency change of the first (1000 Hz), second, and third tone (2000 Hz) was roughly regular and continuous. On such a condition, incorporating the factor of good continuation, these three tones were organized together as a tone stream or a unitary entity.

Therefore, when f_2 was 1000, 1500, and 2000 Hz, the number of tones in one group including the third tone (deviated tone) was one, three, and two, respectively. So the results on the group 1 subjects can be explained by assuming the working hypothesis on MMN amplitude. The results on the group 2 subjects are thought to be valid with the lack of the good continuation factor.

The experiments in the present study provide no direct evidence of the neuronal process of the sequential grouping of auditory information. But these results could be explained by the working hypothesis shown in Sect. 2.4, which was based on the electrophysiological characteristics of auditory preattentive sensory memory previously reported by the MMN studies. As the relative change

Table 1. (Experiment 2) Schematic diagram of possible hypothesis

(a) Tone Frequency			(b) Schematic View of Standard Stimulus	(c) Factors for Grouping			(d) Number of Grouped Tones	(e) MMN Amplitude
f_1	f_2	f_3		Temp.	Freq.	Cont.		
1000 Hz	2000 Hz	2000 Hz (Std.) 1800 Hz (Dev.)	 / 1 / 2 3 /	+	++	+	2	++
	1500 Hz		 / 1 2 3 /	+	+	+++	3	+
	1200 Hz		 / 1 /? 2 /? 3 /	+	–	++	1~2	+++
	1000 Hz		 / 1 2 / 3 /	+	––	+	1	++++

The three abstract relative degrees of factors for grouping the second tone together with the third tone: temporal proximity ("Temp."), frequency proximity ("Freq."), good continuation ("Cont."), expected number of tones in the group including the third tone ("number of grouped tones"), and expected MMN amplitude on the assumption of the working hypothesis ("MMN amplitude") were shown for each stimulus condition. f_1, f_2, and f_3 denote the frequencies of the first tone, the second tone, and the third tone, respectively. The expected combination of sequential grouping was also shown under each schematic view of the standard stimulus ("1," "2," and "3" denote the first tone, the second tone, and the third tone, respectively)

of MMN amplitudes between stimulus conditions corresponded to the expected sequential grouping phenomena of tones, these results showed that the auditory sequential grouping was achieved by the preperceptual sensory memory process. The investigation of a neuronal or physiological basis of the present experiments is left for future study.

5 Conclusion

The neuronal process of sequential grouping of auditory information in human preattentive sensory memory was investigated. In two experiments, elicited MMN magnitudes were changed according to the configuration (the number of tones, frequency) of successive tone sequences. From this result it was shown that the sequential grouping of successive tones was one of the processes of the auditory sensory memory. It was also shown that the relationship between the configuration of tones and the elicited MMN magnitudes met the working hypothesis on which the expected MMN magnitudes were assumed for temporal grouping factors (temporal/frequency proximities and good continuation) proposed by the auditory psychophysiology. These results showed that the factors (rules) of the grouping could be revealed by the MMN component. These results will be helpful in investigating cortical representation in the auditory system and the computational theory of auditory information processing.

This study was approved by the Ethics Committee on Clinical Investigation, Graduate School of Engineering, Tohoku University and was carried out in accordance with the policy of the Declaration of Helsinki.

References

Alho K (1995) Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound change. Ear Hear 16:38–51

Baumann U (2000) Identification and segmentation of multiple auditory objects. In: Auditory worlds: sensory analysis and perception in animals and man. Wiley-VCH, Weinheim, Germany, pp 274–278

Bregman AS (1990) Auditory scene analysis. MIT Press, Cambridge, MA
Darwin CJ (1997) Auditory grouping. Trends Cogn Sci 1(9):327–333
Kanoh S, Itoh T, Kazami K, Futami R, Hoshimiya N (1996) A study on mismatch response to a change in tone duration by EEG and MEG. Jpn J Med Electron Biol Eng 34(4):128–136
Kanoh S, Arai T, Futami R, Hoshimiya N (2001) Properties of auditory temporal integration revealed by mismatch negativity. In: Proceedings of the 23rd annual IEEE EMBS international conference, 2.7.3.4
Näätänen R (1992) Attention and brain function. Erlbaum, Hillsdale, NJ
Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) 'Primitive intelligence' in the auditory cortex. Trends Neurosci 24(5):283–288
Schröger E, Paavilainen P, Näätänen R (1994) Mismatch negativity to changes in a continuous tone with regularly varying frequencies. Electroencephalogr Clin Neurophysiol 92:140–147
Schröger E, Tervaniemi M, Wolff C, Näätänen R (1996) Preattentive periodicity detection in auditory patterns as governed by time and intensity information. Cogn Brain Res 4:145–148
Shinozaki N, Yabe H, Sato Y, Sutoh T, Hiruma T, Nashida T, Kaneko S (2000) Mismatch negativity (MMN) reveals sound grouping in the human brain. Neuroreport 11:1597–1601
Sussman E, Ritter W, Vaughan HG Jr (1998) Predictability of stimulus deviance and the mismatch negativity. Neuroreport 9:4167–4170
Sussman E, Ritter W, Vaughan GH Jr (1999) An investigation of the auditory streaming effect using event-related brain potentials. Psychophysiology 36:22–34
Tervaniemi M, Saarinen J, Paavilainen P, Danilova N, Näätänen R (1994) Temporal integration of auditory information in sensory memory as reflected by the mismatch negativity. Biol Psychol 38:157–167
Tiitinen H, May P, Reinikainen K, Näätänen R (1994) Attentive novelty detection in humans is governed by pre-attentive sensory memory. Nature 372:90–92
Winkler I, Schröger E (1995) Neural representation for the temporal structure of sound patterns. Neuroreport 6:690–694
Winkler I, Czigler I, Jaramillo M, Paavilainen P, Näätänen R (1998) Temporal constraints of auditory event synthesis: evidence from ERP. Neuroreport 9:495–499
Winkler I, Schröger E, Cowan N (2001) The role of large-scale memory organization in the mismatch negativity event-related brain potential. J Cogn Neurosci 13(1):59–71
Yabe H, Tervaniemi M, Reinikainen K, Näätänen R (1997) Temporal window of integration revealed by MMN to sound omission. Neuroreport 8:1971–1974

2004 International Symposium on Nonlinear
Theory and its Applications (NOLTA2004)
Fukuoka, Japan, Nov. 29 - Dec. 3, 2004

Neuronal Correlates of Human Auditory Grouping: An ERP Study

Shin'ichiro Kanoh†, Ryoko Futami† and Nozomu Hoshimiya‡

†Graduate School of Engineering, Tohoku University, Aoba-yama 6-6-05, Sendai, 980-8579 Japan
‡Tohoku Gakuin University, 1-3-1 Tsuchitoi, Sendai, 980-8511 Japan
Email: kanoh@ecei.tohoku.ac.jp

Abstract—The human sequential grouping which organizes parts of tones into a group was examined by the mismatch negativity (MMN), a component of event-related potentials that reveals the sensory memory process. The sequential grouping is accomplished by the combinations of some factors, e.g. temporal and frequency proximity principles. Experimental results showed that the sequential grouping of presented tones was achieved on the auditory pre-attentive sensory memory process, and MMN amplitudes correlated to the temporal configurations of tones to be grouped. The investigation of MMN properties could reveal the nature of auditory sequential grouping.

1. Introduction

The auditory sensory memory is a memory system that stores auditory information within a short period (a few seconds). The basic information processing is applied to the stored information during the auditory sensory memory process [1].

Sequential grouping [2], which binds some successive information into one entity, is one of such a primitive processing. This function is one of the bases of higher-order auditory processing, like the segmentation and "cocktail-party effect" of speech sounds.

By the auditory psychophysical studies, it is found that the sequential grouping is accomplished by the combinations of some factors [3], e.g.

- Proximity: sounds that are near to one another in time (temporal proximity) or in frequency (frequency proximity) will group with one another.
- Good continuation: sounds with good continuation (e.g. frequency change) will tend to bind together.

But no physiological evidence to validate these factors on sequential grouping has been reported. Observing the neuronal correlate of sequential grouping will be a help to determine the principle of auditory information processing and to hypothesize the neuronal representation of temporal tone sequence, e.g. how the order and frequency of the successive tones are encoded.

In this study, the sequential grouping in human auditory cortex was examined by the mismatch negativity (MMN) [4], a component of event-related potentials (ERPs) that revealed sensory memory process. It was shown that the sequential grouping of successive tones was achieved during the auditory sensory memory process, and the properties of the sequential grouping were revealed by the MMN activities. These results will help to find out the cortical representation of auditory stimuli and will contribute to a computational theory of auditory information processing.

2. Working hypothesis

The mismatch negativity (MMN), one of the event-related potential (ERP) components, is a measure of pre-attentive sensory memory properties [4]. It is elicited when some of the frequently presented ("standard") stimuli are randomly replaced by the infrequent ("deviant") stimuli with some different attributes from the standard ones (oddball paradigm). MMN is generated by comparing the memory trace of the standard stimuli and the neuronal representation of incoming deviant stimulus.

Consider the experiment with oddball tone sequence in which each of the stimuli consists of several tones, and only the frequency of the last tone (deviated tone) is different between the standard and the deviant stimulus.

Let us assume that the neuronal activations of each grouped auditory stimuli would be bound together as a unitary one on the auditory cortex.

Incorporating this assumption, when the last tone is grouped with other tones, MMN will be elicited by comparing two unitary neuronal representations on the standard and the deviant stimuli, both of which encode the same set of grouped tones including the deviated tone.

In this case, when the number of tones in the grouped unitary entity was more, relative change of attributes in the corresponding neuronal representations on the standard and the deviant stimulus would be smaller.

The important property of MMN is that the MMN magnitude increases as a function of the extent of deviancy (difference between the attributes in the standard and the deviant stimuli) [5, 6, 7]. A relative difference between the neuronal representations of two sensory events in the sensory memory can be evaluated by MMN measurements.

Therefore, the present study was examined under the working hypothesis: "when the number of tones grouped together with the deviated tone increases, the elicited MMN magnitude will decrease."

In the present study, the two experiments were executed to determine whether the sequential grouping was achieved in the sensory memory process, and to find out what "rules" for sequential grouping are applied to various configurations of tone series.

3. Experiment 1: Temporal proximity factor on sequential grouping

At first, the sequential grouping of multiple tones by the temporal proximity factor was evaluated by the elicited MMN amplitude.

3.1. Method

Six male volunteers with normal hearing ability participated in the experiment as subjects. Auditory stimuli were applied to both ears by headphones. Each subject in an electro-magnetically shielded room was instructed to read a self-selected book and to ignore the presented auditory stimuli (reading condition). The EEG signal was recorded with Ag–AgCl electrodes at two locations (Fz, F4), and referred to linked earlobe (A1, A2) with a forehead ground.

Each stimulus consisted of several tone bursts (20 ms, 60 dB SL, rise/fall time 5 ms), i.e. the first tone, the last tone, and the inserted tone(s) which was (were) inserted between the first tone and the last tone. Total duration of the stimulus was 200 ms. The intervals between the last inserted tone and the last tone and between inserted tones were set to 20 ms. Inter-stimulus interval (onset to onset) was set to 600 ms. In the standard stimulus ($p = 0.9$), frequencies of the first, inserted and the last tones were 1000 Hz, 1000 Hz and 2000 Hz, respectively. And in the deviant stimulus ($p = 0.1$), only the frequency of the last tone was deviated to 1000 Hz.

The number of inserted tones (N_t) was set to 0, 1, 2 and 3, and was fixed within a single session. The relationship between N_t and elicited MMN by the infrequent change of the frequency of the last tone was observed. In this experiment, it is intended to evaluate the number of tones that are grouped with the last tone.

After the application of low-pass filter (30 Hz) to the averaged waveform, the amplitude of the negative peak in the difference waveform (deviant − standard) whose latency was 100 – 250 ms was evaluated. The data were statistically tested using one-way analysis of variance (ANOVA) with the factors of stimulus type ($p < 0.05$) to determine whether the ERPs associated with the standard stimuli were significantly different from those associated with the deviant stimuli at the MMN peak amplitude.

3.2. Results and Discussion

Figure 1 shows the difference waveforms of averaged responses to the standard and the deviant stimuli in Experiment 1 (Subject A, electrode Fz). From this subject, MMN was elicited when the number of inserted tones N_t was 0, 1 and 2, but no significant MMN was observed on $N_t = 3$.

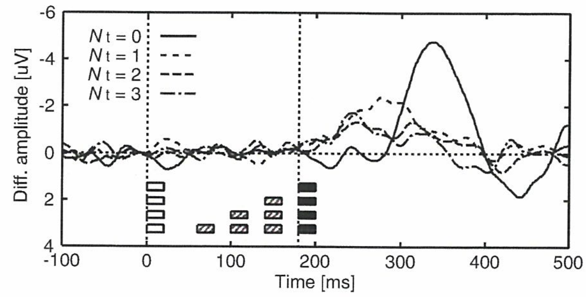


Figure 1: [Experiment 1] Averaged difference waveforms (deviant − standard). (Subject A, electrode Fz)

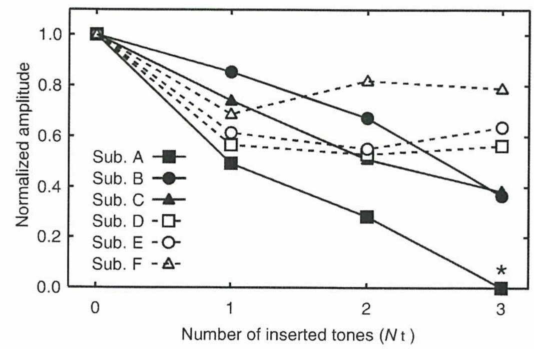


Figure 2: [Experiment 1] Relationship between N_t and normalized MMN amplitude on all subjects (electrode Fz). From Subject A, no significant MMN was observed on $N_t = 3$ (marked *).

The normalized amplitudes of the observed MMN on all subjects are shown in Figure 2. On all subjects, the elicited MMN amplitude taken from Fz was larger than that from F4. The properties of the observed MMN amplitudes could be divided into the following two groups:

- Group 1: the observed MMN amplitude was smaller for larger N_t (Subjects A, B, C).
- Group 2: the MMN amplitudes on $N_t = 0$ was the largest, and almost the same amplitudes were observed on $N_t = 1$, 2 and 3 (Subjects D, E, F).

This result could be explained on the assumption of the working hypothesis described in Section 2 as follows. On Group 1, the MMN amplitude decreased monotonically when N_t increased. This result could be explained by the hypothesis if all the inserted tones were grouped together with the last tone by the factor of temporal proximity during each of the experiments. And on Group 2, the largest MMN amplitude was elicited on $N_t = 0$, but no clear differences of MMN amplitudes were observed on $N_t \neq 0$. If the subjects in Group 2 tended to group only the last two tones (the last inserted tone and the last tone), such result would agree with the present hypothesis.

Therefore, the different phenomena on Group 1 and Group 2 can be explained by the different width of the time window for grouping based on temporal proximity: longer than 140 ms (Group 1), 60 ms to 80 ms (Group 2).

From Figure 1, it was also found that both peak and onset latencies were shorter when N_t was larger. And the duration of MMN was shorter when $N_t = 0$ (120 ms), but was longer when $N_t = 1, 2$ and 3 (200 ~ 250 ms). These results could be observed on the other subjects. Such an ERP component with shorter latency might be related to components other than the MMN.

4. Experiment 2: Frequency proximity factor on sequential grouping

Next, the sequential grouping of multiple tones by frequency proximity factor, and the effect of interplay with the other grouping factors were evaluated by the elicited MMN amplitude.

4.1. Method

Five male volunteers with normal hearing ability participate in the experiment as subjects. Auditory stimuli were applied to both ears by headphones.

Each stimulus consisted of the first tone (duration 40 ms), the second tone (20 ms) and the third tone (40 ms), each of which was 60 dB SL tone burst (rise/fall time 5 ms). Total duration of the stimulus was 200 ms. The interval between the second tone and the third tone was set to 30 ms. Inter-stimulus interval (onset to onset) was set to 600 ms. In the standard stimulus ($p = 0.9$), the frequencies of the first and the third tone (f_1, f_3) were 1000 Hz and 2000 Hz, respectively. In the deviant stimulus ($p = 0.1$), only the frequency of the third tone f_3 was deviated to 1800 Hz.

The frequency of the second tone (f_2) was set to 1000, 1200, 1500 and 2000 Hz, and was fixed within a single session. The relationship between f_2 and elicited MMN by the infrequent change of the frequency of the third tone was observed. Other experimental procedures were the same as in Experiment 1.

4.2. Results and Discussion

Figure 3 shows the difference waveforms of averaged responses to the standard and the deviant stimuli in Experiment 2 (Subject A, electrode Fz). From this subject, significant MMN was elicited on each condition.

The normalized amplitudes of the observed MMN for all subjects are shown in Figure 4. On all subjects, the elicited MMN amplitude taken from Fz was larger than that from F4. The elicited MMN amplitude was the largest when f_2 was 1000 Hz on all subjects, but the tendency of MMN amplitudes could be divided into the following two groups:

- Group 1: the smallest amplitude was observed when f_2 was 1500 Hz, and the amplitudes on $f_2 = 1200$ and 2000 Hz were almost the same (Subjects A, B).

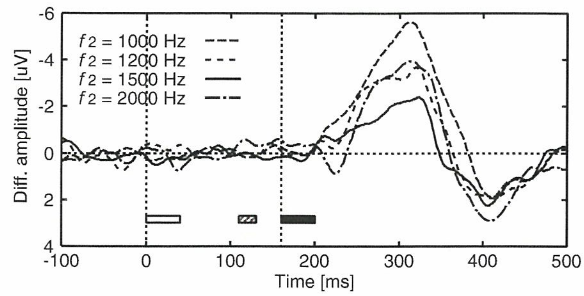


Figure 3: [Experiment 2] Averaged difference waveforms (deviant – standard) (Subject A, electrode Fz)

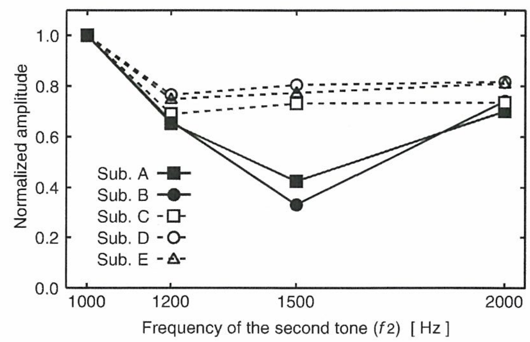


Figure 4: [Experiment 2] Relationship between f_2 and normalized MMN amplitude on all subjects. (electrode Fz)

- Group 2: the smallest MMN amplitude was observed when $f_2 = 1200$ Hz, but there was no significant difference between $f_2 = 1200$, 1500 and 2000 Hz (Subjects D, E, F).

Taking only the factor of frequency proximity into account, it was expected that the second tone tended to organize together with the first tone when $f_2 = f_1$ (1000 Hz). But when $f_2 = f_3$ (2000 Hz in the standard stimulus), the expected partner to be grouped with would be the third tone.

On the other hand, by considering only the factor of temporal proximity, the second tone tended to be grouped together with the third tone rather than the first tone in this case (intervals between the second and the first tone and between the second and the third tone were 70 ms and 30 ms, respectively).

In the former case, the two processes of sequential grouping based on temporal and frequency proximity factors contradicted each other, because the temporal position of the second tone was closer to the third tone than the first tone.

In such a situation, it might be possible that the two factors are combined (competition or cooperation) to solve the contradiction. By assuming the working hypothesis on this study, MMN amplitude is expected to decrease monotonically when f_2 increased from 1000 Hz to 2000 Hz, because the second tone tends to be grouped together with the third

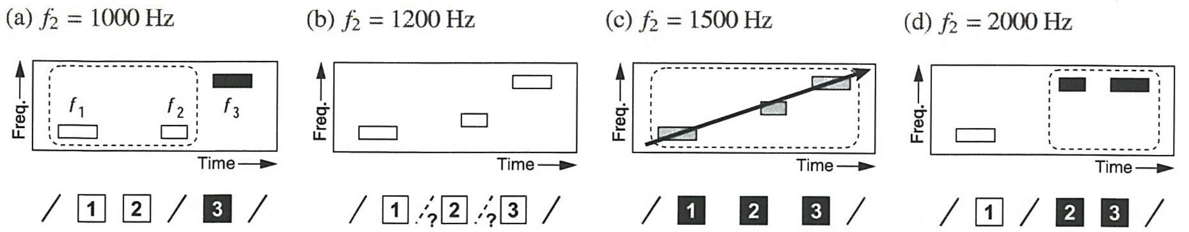


Figure 5: [Experiment 2] The schematic view of the standard tone, and the expected combination of sequential grouping on each condition. See text in details. ("1", "2" and "3" denote the first tone, the second tone and the third tone, respectively).

tone easier when f_2 approaches to the frequency of the third tone f_3.

But on the two subjects (Subjects A, B) out of five, the minimum MMN amplitude was observed when f_2 = 1500 Hz. This phenomenon could be explained by introducing the "good continuation" as one of the factors in addition to the factors of temporal/frequency proximity for sequential grouping of successive tones.

Figure 5 is the schematic diagram on the present hypothesis. When f_2 was 1000 Hz and 2000 Hz, the second tone tended to be grouped together with the first and the third tone, respectively. And when f_2 was 1500 Hz, the frequency change of the first tone (1000 Hz), the second tone and the third tone (2000 Hz) was roughly regular and continuous. On such a condition, incorporating the factor of good continuation, these three tones were organized together as a tone stream, or an unitary entity.

Therefore, when f_2 was 1000 Hz, 1500 Hz and 2000 Hz, the number of tones in one group including the third tone (deviated tone) was one, three and two, respectively. So the results on the subjects in Group 1 can be explained by assuming the working hypothesis on MMN amplitude. The results on the subjects in Group 2 are thought to be valid with the lack of the factor of good continuation.

From the experiments in the present study, no direct evidence on the neuronal process of the sequential grouping of auditory information was given. But these results could be explained by introducing the working hypothesis shown in Section 2. Similar experiments with variable temporal positions of the second tone will be needed for investigating the plausibility of the hypothesis. And by evaluating the frequency of the second tone in various temporal positions on which minimum MMN amplitude is observed, the physiological scale of tone frequency (c.f. the Mel scale) in the auditory cortex might be determined. The investigation of the neuronal or physiological basis of the present experiments is left for further study.

5. Conclusion

The neuronal process of sequential grouping of auditory information in human auditory sensory memory was investigated. In the two experiments, elicited MMN magnitudes were changed according to the configuration (the number of tones, frequency) of successive tone sequences. From this result, it was shown that the sequential grouping of successive tones was one of the processes of the auditory sensory memory. And it was also shown that the relationship between the configuration of tones and the elicited MMN magnitudes met the working hypothesis, on which the expected MMN magnitudes were assumed for temporal grouping factors (temporal/frequency proximities and good continuation) which was proposed by the auditory psychophysiology. This results showed that the factors (rules) of the grouping could be revealed by the mismatch negativity (MMN) component. These results will be helpful to investigate the cortical representation in the auditory system and the computational theory of auditory information processing.

This study was approved by the Ethics Committee on Clinical Investigation, Graduate School of Engineering, Tohoku University, and was performed in accordance with the policy of the Declaration of Helsinki.

References

[1] R. Näätänen, M. Tervaniemi, E. Sussman, P. Paavilainen, I. Winkler, "'Primitive intelligence' in the auditory cortex," *Trends in Neurosciences*, **24**, 5, pp.283–288, 2001.

[2] C. J. Darwin, "Auditory grouping," *Trends in Cognitive Sciences*, **1**, 9, pp.327–333, 1997.

[3] A. S. Bregman, *Auditory scene analysis*, MIT Press, Cambridge, MA. 1990.

[4] R. Näätänen, *Attention and Brain Function*, Erlbaum, Hillsdale, NJ, 1992.

[5] H. Tiitinen, P. May, K. Reinikainen, R. Näätänen, "Attentive novelty detection in humans is governed by pre-attentive sensory memory," *Nature*, **372**, pp.90–92, 1994.

[6] S. Kanoh, T. Itoh, K. Kazami, R. Futami, N. Hoshimiya, "A study on mismatch response to a change in tone duration by EEG and MEG," *Japanese Journal of Medical Electronics and Biological Engineering* (in Japanese), **34**, 4, pp.128–136, 1996.

[7] S. Kanoh, T. Arai, R. Futami, N. Hoshimiya, "Properties of auditory temporal integration revealed by mismatch negativity," Proceedings of the 23rd Annual International Conference of IEEE EMBS, 2.7.3.4, 2001.

交流眼電図からのFES制御命令取得に関する基礎的検討

○鈴木信太郎　加納慎一郎　二見亮弘
東北大学大学院 工学研究科

A basic study on FES control command acquisition with AC Electro-Oculogram

○Shintaro Suzuki, Shin'ichiro Kanoh, Ryoko Futami
Graduate School of Engineering, Tohoku University

1 はじめに

機能的電気刺激(FES)による麻痺肢運動制御のためには、刺激装置に患者の随意的な意図を伝達する必要がある。しかし FES の適用対象が脊髄損傷などによる四肢麻痺のように重篤な患者の場合は、刺激装置に与える患者の随意的な意図を反映した生体情報が限られてしまう。そのため、患者の残存機能をできるだけ制限せずに、多くの制御命令を伝達できるような患者—刺激装置間の入力インターフェースの開発が必要となる。そこで本研究では、垂直方向、水平方向の眼電図(EOG)を計測から上下、左右、斜め方向の視点移動方向の検出を行うことでメニュー選択を行うインターフェースシステムの開発のための基礎的検討を行ったので報告する。本報告では、AC 増幅器を用いて計測した EOG から 8 方向の視点移動方向の検出を行うアルゴリズムを提案し、その検出精度について検討を行った結果について述べる。

2 交流増幅による視点移動方向の検出

視点移動によって生じる EOG の計測には、DC 増幅器による計測[1]が一般的であり、また遮断周波数が十分に低い(0.1Hz 程度以下)AC 増幅器を用いることもある[2]。DC 増幅器の場合、視点位置の検出が容易であることが長所としてあげられる。しかし、電極の分極、皮膚の電気的活動や増幅器自体に起因するドリフトによって計測が困難であり、また遮断周波数が低い AC 増幅器を用いた場合は、瞬目などによる大振幅信号やリード線の動揺などによるアーチファクトが混入した場合に波形が安定するまで数秒の時間を要するという欠点があった。そのため通常の EOG 計測に比べて遮断周波数が高い AC 増幅器を用いて視点の移動方向を検出できることが望ましい。しかし、視点移動がステップ状に生じた、過渡的な EOG の変化を計測することは可能であるが、視点を維持している状態の EOG を計測することは不可能である。そのため、計測値からの視点位置の計算は理論的には不可能である。そこで EOG の大振幅を持つデータに注目し、このデータから 8 方向の視点移動方向を検出する方法について検討を行った。

3 実験方法

3.1 計測方法

本研究で用いた眼球運動検出のための実験システムを図1に示す。被験者（健常男性 1 名）は計測・解析用コンピュータのディスプレイの正面（約 70cm）に着座し、ディスプレイ上に呈示される注視点に視点を追従させることを求められた。

実験中、注視点を以下のように移動させた。被験者が任意のタイミングでスタートキーを押下すると, ディスプレイの中央に赤色の注視点が 1.5 秒間表示される。その後注視点は、被験者に予め教示した方向に移動を開始する。注視点は, 初期位置から視野角約 10° の位置に表示される円上まで 0.5 秒で移動し、到達すると停止する。

被験者がスタートキーを押してから5秒間の EOG を計測した。Ag-AgCl 皿電極と電極ペーストを用いて、図 2 に示すように電極を配置し、生体信号増幅器(MEG-6108：日本光電製)を用いて垂直方向(V_v)、水平方向(V_h)の EOG を双極誘導で計測した。増幅率は 5000 倍とした。増幅後 0.5～100Hz に帯域制限をし、サンプリング周波数 250Hz でコンピュータ上に集積した。注視点の移動方向は 45° 刻みの 8 種類（右, 右上, 上, 左上, 左, 左下, 下, 右下）とし、それぞれについて

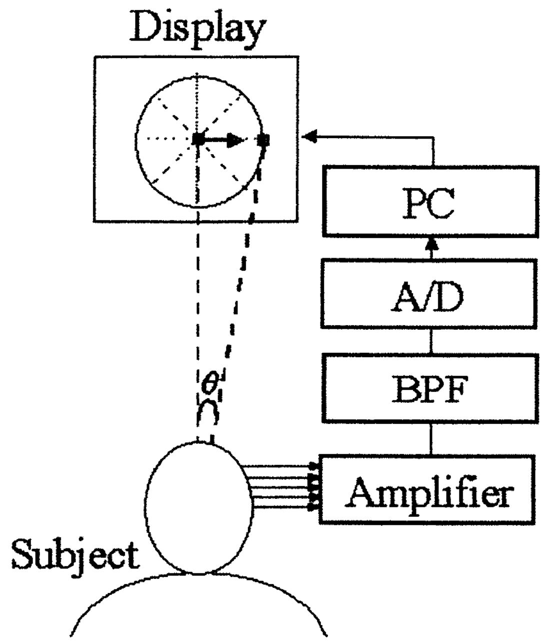


図 1: 実験システムの概略

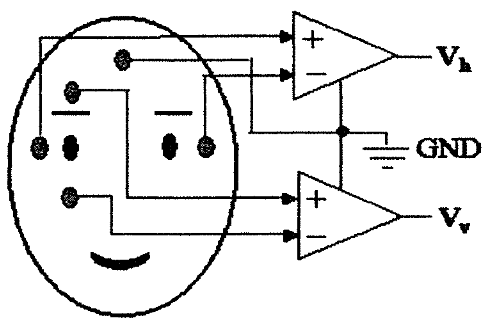


図 2: 電極配置

10 試行の計測を行った。

3.2 解析方法

本研究で用いた8方向の視点移動方向の検出方法を以下に示す。

まず計測した信号 V_h、V_vから式(1)を用いてノルム $R(t)$を計算する。

$$R(t)=\sqrt{(V_v(t))^2+(V_h(t))^2} \qquad (1)$$

2 で述べたように今回の計測では遮断周波数が高い AC 増幅を行うため、計測した信号の大振幅を持つ値に注目する必要がある。そこで本解析手法では$R(t)$の大きさに閾値R_{th}を定め、式(2)を満たすサンプル(V_h,V_v)のみを採用する方式を用いた。ここでいくつかのサンプルを採用する理由は後に述べる。

$$R(t) \geqq R_{th} \qquad (2)$$

次に式(2)によって採用されたサンプルの総数をNとし、式(3)で与えられるようにN個のサンプル(V_h,V_v)の平均値($\overline{V}_h$,$\overline{V}_v$) を決定する。

$$(\overline{V}_h,\overline{V}_v) = (\frac{1}{N}\sum_{i=1}^{N}V_h(i),\frac{1}{N}\sum_{i=1}^{N}V_v(i)) \qquad (3)$$

(0,0)を基準とした($\overline{V}_h$,$\overline{V}_v$) の偏角 θ を式(4)から計算し、θ から実際の視点移動距離を推測できるかどうかを検討した。

$$\theta = \tan^{-1}(\frac{\overline{V}_v}{\overline{V}_h}) \qquad (4)$$

視点移動方向を推定するための一般的な手法として、$R(t)$が最大となるときの(V_h,V_v)を用いて式(4)の計算を行う方法がまず挙げられよう。しかし瞬時値(V_h,V_v)は計測系に起因するノイズ成分が含まれるため、瞬時値から偏角を計算するとその影響が大きくなることが予想される。そこで本手法では振幅のある程度大きなサンプルデータを用い、平均をとることでこれらのノイズを除去し、その平均値から偏角を計算することで視点移動方向を推定する方法を算出した。

4 実験結果

実際の計測波形の例を図 3(a),図 3(b)に示す。図 3(a)は上に視点移動したときの水平、垂直方向の EOG(V_h,V_v)とそのノルムを、図 3(b)は(V_h,V_v)のリサージュ図形と算出した平均値 ($\overline{V}_h$,$\overline{V}_v$) を示す。

上下、左右、斜めの 8 方向に視点を移動させた際の EOG からθを算出した。各方向 10 試行分の θ の平均値とその標準偏差を図 4 に示す。なお、本実験では$R_{th}=50\mu$V とした。図 4(b)より、視点移動方向が左上のときの標準偏差が、他の方向に比べ多少大きくなっていることがわかる。このことから視点移動では、被験者によって安定した視点移動が難しい方向がある可能性が考えられる。

5 考察

図 4(a)より、各移動方向における θ の標準偏差はある程度小さいことがわかる。このことから、求めた θ から移動方向を識別するにあたり、各移動方向に対して θ の範囲を予め設定することで 8 方向の視点移動方向を推定することが容易であることがわかる。なお図 4(a)からは各移動方向における θ の値は完全に一致することはないが、ほぼ

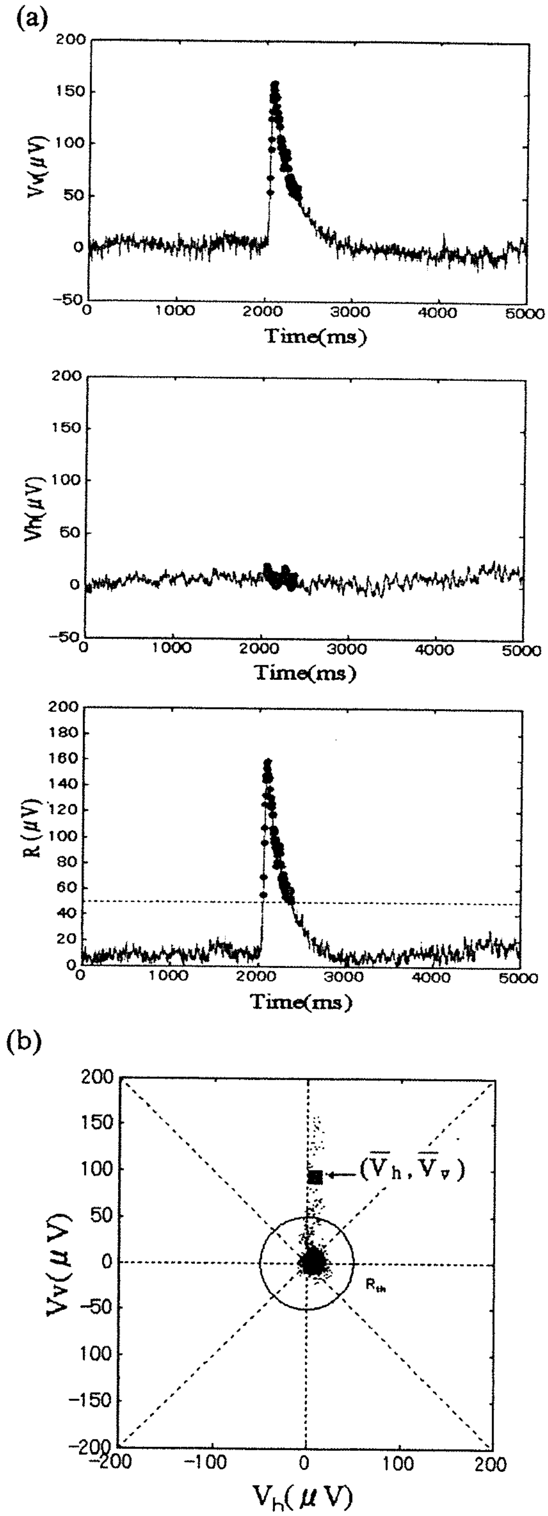


図3 計測波形の例と解析手法の概要

(a)上方向に視点を動かせた際の計測例。V_v(上段)、V_h(中段)および両者から計算されたノルム(下段)を示す。

点で示すものが採用されたデータであり、点線が閾値R_{th}である。

(b)リサージュ図形。閾値R_{th}を半径とする円の外側のサンプルのみを採用し、採用されたサンプルから計算される平均値 $(\overline{V}_h,\overline{V}_v)$を■で示し、(0,0)を基準とした $(\overline{V}_h,\overline{V}_v)$ の偏角を θ とした。

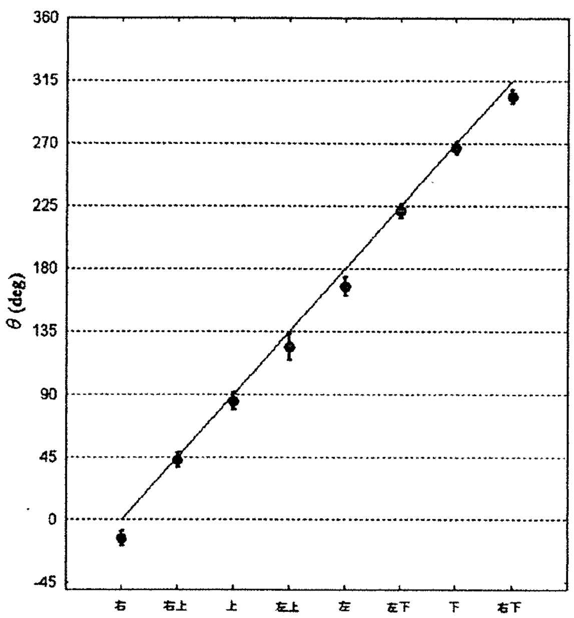
線形な関係となった。この結果は上記の θ の範囲の決定を容易とするものかもしれない。被験者によって容易ではない視点移動方向ある可能性を考慮して範囲を設定するか、あるいは被験者に対してトレーニングを行うことで、推定精度をさらに向上させることも可能であろう。本報告で示した結果は1名の被験者に対して1回の実験のものである。電極の付け替えによる配置のずれ、電極インピーダンスの変化が影響する可能性がある。そのため多くの被験者に対して複数回の実験を行い、結果の再現性を検討する必要がある。また本解析手法を用いて各視点移動方向における θ と標準偏差に着目し、上記で述べた識別のための範囲設定方法や、被験者の眼球運動の特徴について検討する必要がある。

今回は視点の移動方向のみについて検討を行ったが、同一手法を用いて、視点移動距離の長短の情報が検出できる可能性もある。本研究で用いたR_{th}の決定方法が重要となると考えられる。これは、EOG の振幅の大きさが移動距離の長短に依存するため、R_{th}を絶対的な大きさで定義すると条件ごとに採用されたサンプルの振幅分布が重なってしまうために、求める平均値$(\overline{V}_h,\overline{V}_v)$は振幅の大きさを十分反映しないものとなる。これを避けるために$R(t)$の値の分布をもとにR_{th}の値を相対的に設定する必要がある。この方法により、採用されたサンプルを用いて求めた平均値$(\overline{V}_h,\overline{V}_v)$に振幅の大小という情報が大きく反映され、平均値の大小を考慮することで移動距離の長短が検出できると考えられる。今後、本手法の妥当性の検証を行う予定である。

6 まとめ

本稿では AC 増幅器を用いて計測した EOG から、8 方向の眼球動作検出アルゴリズムを提案し、その検出精度について検討を行った。その結果、時定数の小さい AC 増幅器を用いて計測を行った場合でも、8 方向の移動方向の区別が十分に可能であることが示された。本結果から、視点移動方向の検出によってメニュー選択を行う単純な解析アルゴリズムによるインターフェースの開発が可能だと考えられる。今後は同一被験者での結果の再現性を調べることや、被験者を増やすことで解析手法の妥当性を検討する必要がある。

(a)



(b)

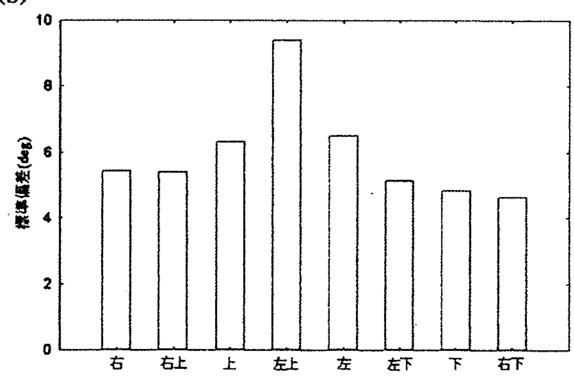


図 4　視点移動方向と θ の関係

(a)視点移動方向と θ の平均値とその標準偏差。直線は移動方向と θ が一致したときである。
(b)視点移動方向と標準偏差。

文献

[1]JOHN R. LaCOUSE and FRANCIS C. HLUDIK,JR,"An Eye Movement Communication-Control System for the Disabled" IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING,VOL.37,NO12,pp1215-1220,DECEMBER 1990

[2]宮田 洋 "新 生理心理学 1巻 生理心理学の基礎",北大路書房(1998)

A Computer Simulation Study of the Feedback Error Learning Controller for FES on the Wrist Joint's 1-DOF Movement

S. Chosa*, T. Watanabe**, M.Yoshizawa** and N.Hoshimiya***

* Graduate School of Engineering, Tohoku University, Sendai, Japan
** Information Synergy Center, Tohoku University, Sendai, Japan
*** Tohoku Gakuin University

seiichi@yoshizawa.ecei.tohoku.ac.jp

Abstract: The Feedback Error Learning (FEL) for Functional Electrical Stimulation (FES) controller was examined through computer simulation using a musculoskeletal model. Dorsi and palmar flexions of the wrist was controlled by stimulating the extensor carpi radialis longus/brevis (ECR) and the flexor carpi ulnaris (FCU). The computer simulation results were similar to those of experiments performed previously with a neurologically intact subject. Furthermore, it was shown that the Artificial Neural Network (ANN) trained with fast movement could control the slower one. The feasibility of the FES controller using the FEL was also shown by the computer simulation. The model simulation was considered to be a good method to advance the FES research.

Introduction

Functional Electrical Stimulation (FES) is an advancing technology for restoring paralyzed motor functions caused by the spinal cord injury or the stroke. One of the main subjects of FES study is to develop a controller that restores movements precisely and stably. Open-loop controllers usually used at clinical sites have problems that they need initial and periodical adjustments in stimulation patterns, which are burdening both patients and medical staffs. The multi-channel PID controller developed by our research group [1] performed good tracking on 2-DOF movements of the wrist joint stimulating four muscles. However, the tracking error increased as the velocity of the desired movement was high.

The Feedback Error Learning (FEL) [2] controller was a hybrid regulator with a feedforward and a feedback controller, which was supposed to make up for their weak points each other. Our research group showed the feasibility of the FES controller using FEL through the experiments with an able bodied subject [3]. Further examinations of the FEL controller under various conditions are desired for clinical application. Computer simulation is considered to be preferable to experiments with subjects for the examination since it is difficult to perform a number of experiments with subjects.

In this paper, the FES controller based on the FEL was examined by using the musculoskeletal model developed by our group [4]. The musculoskeletal model could express different muscle properties easily, which wasn't possible by the forward dynamics model used in the previous study [3]. First, learning and control ability of a trained feedforward controller was tested applying to a different subject model. Then the capability of the FEL FES controller was examined in tracking periodic movements with several different cycle periods.

Materials and Methods

Feedback Error Learning FES Controller: Figure 1 shows the diagram of the control system. The FEL controller consists of a feedforwad controller and a feedback controller. In this study, the PID controller [1] was used as a feedback controller. Three-layered perceptron neural network shown in Figure 2 was used as a feedforward controller. The number of neuron was 18 for the input layer, 18 for the hidden layer and 2 (the number of muscles stimulated) for the output layer. The

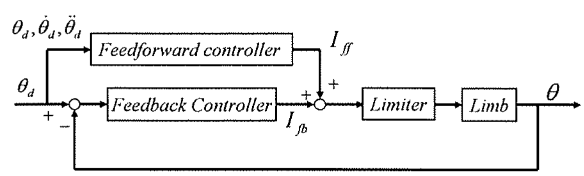


Figure 1. The block diagram of the control system

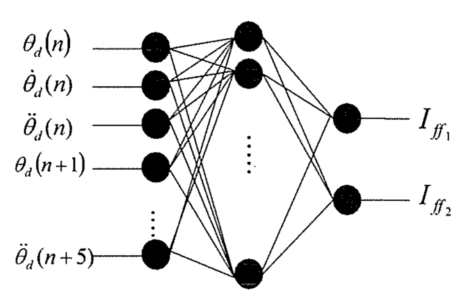


Figure 2. 3-layered perceptron neural network for the feedforward controller

inputs of the neural network were desired angle, angular velocity and angular acceleration at the time t to $t+5$. The output function of the neurons in the second and the third layer was the sigmoid function whose output was between 0 and 1. The neural network was trained by the error back-propagation algorithm using the output of the feedback controller as the error signal.

Computer Simulation: The musculoskeletal model designed based on Hill-type model of muscle [4] was used as the control object. Movements of the elbow, the forearm and the wrist could be controlled by stimulating 15 muscles with the shoulder joint fixed at an arbitrary angle of flexion/extension and rotation. The muscle model had nonlinear length-force and velocity-force relationships, recruitment characteristics and activation dynamics. Nonlinear joint angle dependency of moment arm and passive viscoelastic property were also included into the model.

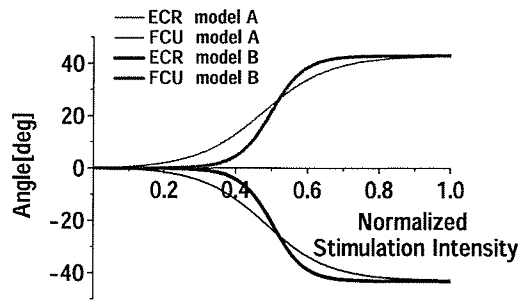
The wrist movement was controlled by the electrical stimulation of the extensor carpi radialis longus/brevis (ECR) and the flexor carpi ulnaris (FCU). The desired trajectory was given by a sinusoidal angle trajectory with 30 degrees in peak-to-peak amplitude (25 degrees in the palmar-flexion, 5 degrees in dorsi-flexion). 6 cycles of the movement were included in one control trial. Computer simulations under the two conditions were performed in this study as follows:

1. Muscle properties for control are different from those for training

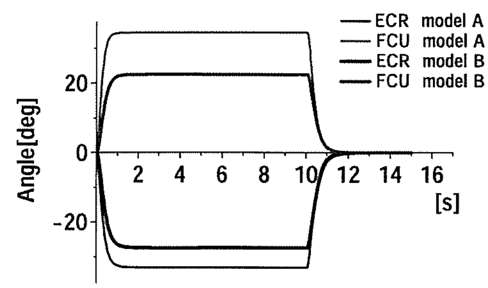
The cycle period of the desired trajectory was 2 seconds. Firstly the neural network was learned on the musculoskeletal model with certatin muscle property (model A). Then the trained ANN was applied to the control of a different musculoskeletal model which had different properties in input-output characteristic and step response (model B). Figure 3 shows the input-output characteristics and step response of the muscles of the model A and the model B.

2. Cycle period of the desired trajectory for control is different from it for training

Firstly the ANN was trained through the control of tracking the sinusoidal angle trajectory with 2, 5 or 10 seconds in cycle period. Then the trained ANN was applied to the control of tracking the trajectory with 2, 5 and 10 seconds in cycle periods.



(a) input-output characteristic



(b) step response

Figure 3. Properties of muscles in the input-output characteristic and the step response used for the ANN training (model A) and for examination with a different subject (model B)

Results

The control results of the first and the 24th trials are shown in Figure 4. It is obvious that the control error and delay seen in the first trial were reduced in the 24th trial. The stimulation outputs of the controller were mainly determined by the PID controller in the first trial [Figure 4 (A) (b)], while they were determined by the ANN in the 24th trial [Figure 4 (B) (b)]. The feedforward controller constructed by learning of the ANN enabled the controlling of fast movements with less delay and error.

Figure 5 shows the changes in mean error and power ratio when the neural network was trained with random weights or the prerained weights as initial value. The error was small and the power ratio was high at the first trial if the neural network was pretrained. Furthermore, the mean error was reduced enough and better control was achieved after a few iterations of training of the pretrained ANN. On the other hand, about 40 iterations were required for the training when the ANN training was started with random weights.

Figure 6 shows the mean error obtained from the control of tracking movements with different cycle periods from those for training. Computer simulation for the training was conducted 5 times changing the initial values of the ANN's weights. The graph shows the mean value of the 5 results. The mean error was smallest when the cycle period for control was same as that for training. The ANN trained with fast movements could also control the slower movements than the movement for training with small mean error.

Discussion

The FES controller using the FEL was examined by the musculoskeletal model. The computer simulation could predict similar results to the experiment with a neurologically intact subject [3]. That is, the mean error and the delay were reduced after the training of the ANN and the number of iterations required for training of the ANN was reduced by the pretraining (Figure 5).

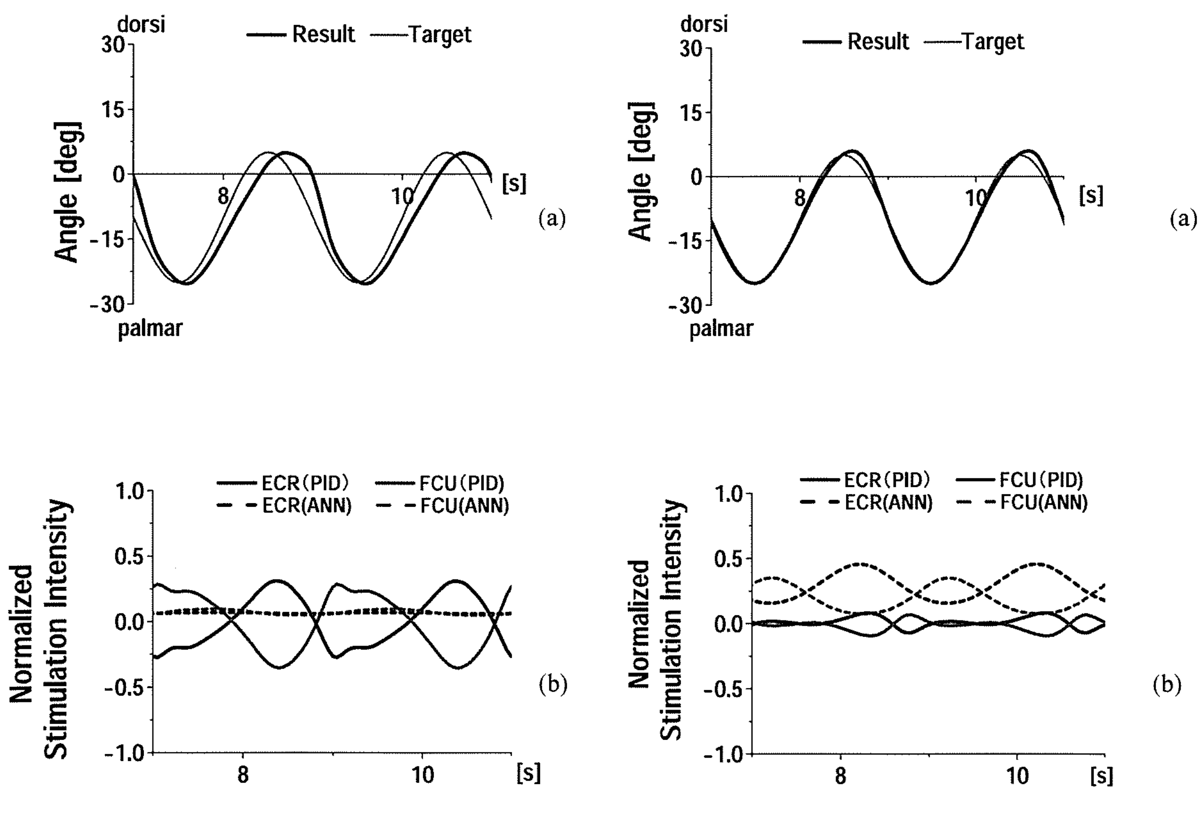
Similar tendency was also observed through the same simulations conducted changing the ANN's initial value of weights. These results suggest that the computer simulation by the musculoskeletal model is useful.

The forward dynamics model used in the previous study had difficulties in modification to express different muscle properties and in expansion to multi-degrees of freedom of movement control. Since the musculoskeletal model used in this study can be modified easily, it is considered to be suitable for this research work compared to the forward dynamics model. It is expected that control capabilities of the FEL controller will be revealed in more detail by performing computer simulations under various conditions modifying properties of this musculoskeletal model.

The ANN trained with fast movements could control the slow movements with good performance. Therefore, training with fast movement may only be required for restoring movements at the clinical site. This will lead to reduce the burden of patients. However, the ANN sometimes failed to learn the inverse dynamics of controlled object. When there was large difference in the output level between the PID controllers for the ECR and the FCU, only one output of the ANN became larger and the training didn't succeed. Modification of the PID controller or the structure of the ANN will be required to solve this problem.

Conclusions

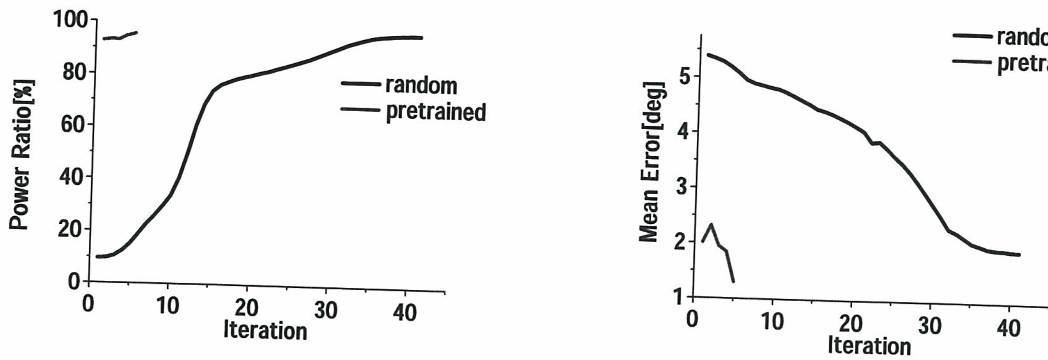
The FES controller using the FEL was found to be a promising controller from the results of computer simulations. However, the PID controller or the structure of the ANN has to be discussed in more detail for further improvement, because the ANN wasn't trained appropriately in some cases. Then it is expected to extend the FEL controller to controlling multiple joint movements. Since the musculoskeletal model was found out to be a useful tool for these studies, further study of the FEL controller can be performed by computer simulation.



(A) iteration 1

(B) iteration 24

Figure 4. An example of simulation results before and after the FEL
(a) angle trajectory (b) outputs of each controller



(a) mean error

(b) power ratio

Figure 5. An example of changes in mean error and power ratio (PR)

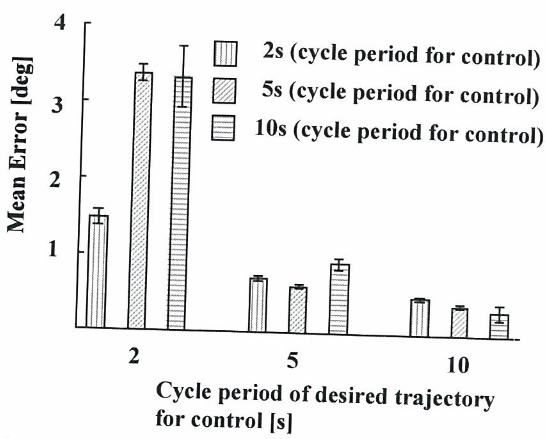


Figure 6. Mean error when the cycle period of the desired trajectory during control is different from it during training

Acknowledgment

This study was partly supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan under a Grant-in-Aid for Scientific Research, and the Ministry of Health, Labour and Welfare under the Health and Labour Sciences Research Grants (Comprehensive Research on Disabillity Health and Welfare)

References

[1] WATANABE T., IIBUCHI K., KUROSAWA K., HOSHIMIYA N., (2002): 'A method of multichannel PID control of 2-degree o freedom of wrist joint movements by functional electrical stimulation', *Trans. of IEICE,* Vol.J85-D-2, pp.319-328, (in Japanese)

[2] MIYAMOTO H., KAWATO M., SETOYAMA T., SUZUKI R.(2003): 'Feedback-error-learning neural network for trajectory control of a robotic manipulator', *Neural Networks,* Vol.1, pp.365-372

[3] KUROSAWA K., FUTAMI R., WATANEBE T., HOSHIMIYA N. (2003): 'Feedback Error Learning for Controlling 1-DOF Joint Angle: Model Simulation and Experiment', *Proc. 8th Ann. Conf. Int. FES Soc.,* pp.65-68

[4] WATANABE T., OTSUKA M., YOSHIZAWA M., HOSHIMIYA N. (2004): 'Computer Simulation for Multichannel Closed-loop FES Control of the wrist Joint', *Proc. Vienna FES Workshop,* pp.138-14