

---

幼児の音声獲得に関わる神経機構と学習原理の計算論的研究

---

15500166

平成15年度～平成17年度科学研究費補助金  
(基盤研究(C)) 研究成果報告書

平成18年6月

研究代表者 二見 亮弘  
福島大学共生システム理工学類教授

## はしがき

ヒトが言葉を獲得していく過程を計算論的に理解するために、聴覚短期記憶の時系列情報に対する特性や、言語に関わる各種高次機能に関わる弁別課題におけるヒトの心理物理特性を知った上で、それらを計算論的に表現する神経機構モデルを構成するという研究手法がある。本研究ではそのような立場から、言語課題の評価に新しい計算論的背景を与えるための「ヒトの反応時間に関する心理物理実験およびそのモデリング」と、「ヒトの時系列情報短期記憶特性の解析」を行った。本報告はそれらの結果をまとめたものである。

## 研究組織

研究代表者 二見 亮弘

(東北大学大学院工学研究科助教授 ; 平成17年3月まで)

(福島大学共生システム理工学類教授 ; 平成17年4月以降)

研究分担者 なし

## 交付決定額 (配分額)

(金額単位：円)

	直接経費	間接経費	合計
平成15年度	2,500,000	0	2,500,000
平成16年度	600,000	0	600,000
平成17年度	600,000	0	600,000
総計	3,700,000	0	3,700,000

## 研究発表

### (1)学会誌等

- 加納慎一郎, 二見亮弘, 星宮望 : ヒト聴覚系の感覚記憶における時間統合過程のミスマッチ陰性電位による解析, 生体医工学, Vol.41, No.2, pp.97-104 (2003)
- S.Kanoh, R.Futami and N.Hoshimiya: Sequential grouping of tone sequence as reflected by the mismatch negativity. Biological Cybernetics, Vol.91, No.6, pp.388-395 (2004)
- ◎ Iwaizumi A, Futami R, Kanoh S, Gyoba J: Characteristics of Human Luminance Discrimination and Modeling a Neural Network Based on the Response Properties of the Visual Cortex, Biological Cybernetics, 2006 Mar 18; [E-pub ahead of print] ISSN: 0340-1200(Paper) 1432-0770(Online).
- Iwaizumi A, Futami R, Matsuki H, Suzuki Y: A neural network model of human response in single- or two-choice luminance-discrimination reaction time task, Biological Cybernetics, 発表予定

## (2)口頭発表

- 岩泉, 加納, 二見: ヒトの輝度弁別特性とその神経回路モデル, 電子情報通信学会技術研究報告, NC2004-1 (2004)
- A.Iwaizumi, S.Kanoh and R.Futami: Psychophysical experiments of human luminance discrimination and the model neural network, Proc. of the Fourth International Symposium on Human and Artificial Intelligence Systems, pp.467-472 (2004)
- 岩泉, 加納, 二見: 視覚野応答の時間特性を考慮した輝度弁別神経回路モデル, 電子情報通信学会技術研究報告, NC2005-2 (2005)
- ◎ 岩泉, 二見, 松木, 鈴木: 2刺激に微小な呈示時間差を持った輝度弁別RTタスクの弁別特性を有する神経回路モデル, 電子情報通信学会技術研究報告 NC2006-5 (2006)

## (3)その他

- ◎ 坂井靖文: 短音と長音で構成される時系列の短期記憶に関する研究, 東北大学工学部電子工学科卒業論文 (2005)

## 研究成果による工業所有権の出願・取得状況

該当なし

## 研究成果

脳における言語処理などの高次機能について定量的に調べる際には、被験者に視覚的な言語課題を呈示し、スイッチ押しまでの反応時間の平均や分散を課題の困難さの指標として用いることが多い。本研究では、各種の高次機能に関わる遂行課題の反応時間や誤り率に新しい計算論的背景を与えることを目的として、心理実験における反応時間分布の生成モデルを構成し、単純化された輝度弁別課題の結果との比較検討を行ってきた。実験では、被験者に左右二つの光刺激の内、より明るいと判断した方に対応するボタンをできるだけ早く押してもらい、刺激からボタン押しまでの反応時間を測定する心理物理実験を行うとともに、その反応時間分布と誤り率を再現できる Winner-take-all 型のパルス神経回路網モデルを生理学的知見に基づいて構成して評価した。

さらに、2刺激の輝度弁別タスクにおいて呈示刺激がより多様な時間構造をもつ場合、すなわち2刺激のオンセットに微小な時間差がある場合、および閾上単一刺激を用いた場合の実験を行うとともに、その結果を説明するためのモデルの拡張を行った。2刺激の実験では、強刺激と弱刺激のオンセット時刻差が $-20\text{ms}$ ,  $0\text{ms}$ ,  $+20\text{ms}$ の3条件の刺激対を用いた。これらは、視覚的に時間差を知覚できない程度の小さな時間差であり、競合型の強弱判定モデルの妥当性を評価するために重要な意味をもつ刺激である。実験の結果から、呈示条件によって反応時間と誤り率に統計的に有意な差があることが明らかになった。また、その実験事実を説明するためのモデルの拡張が必要であることが分かった。モデルの自然な拡張として、時間差を強調する機能を導入するなど3種類の新しいモデルを構成し、評価したところ、競合回路へのグローバル抑制とニューロンチェーンによる時間差強調機構を有するモデルが最もよく実験結果を再現できることが明らかになった。

以上の成果に対応する研究発表として、前項の研究発表リストで◎印を付したものを次頁以降に掲載する。

# Characteristics of human luminance discrimination and modeling a neural network based on the response properties of the visual cortex

Akiko Iwaizumi<sup>1</sup>, Ryoko Futami<sup>2</sup>, Shin'ichiro Kanoh<sup>1</sup>, Jiro Gyoba<sup>3</sup>

<sup>1</sup> Department of Electronic Engineering, Tohoku University, Sendai 980-8579, Japan

<sup>2</sup> Department of Human Support System, Fukushima University, Fukushima 960-1296, Japan

<sup>3</sup> Department of Psychology, Tohoku University, Sendai 980-8576, Japan

Received: date / Revised version: date

**Abstract** Reaction time (RT) and error rate that depend on stimulus duration were measured in a luminance-discrimination reaction time task. Two patches of light with different luminance were presented to participants for 'short' (150 ms) or 'long' (1 s) period on each trial. When the stimulus duration was 'short', the participants responded more rapidly with poorer discrimination performance than they did in the longer duration. The results suggested that different sensory responses in the visual cortices were responsible for the dependence of response speed and accuracy on the stimulus duration during the luminance-discrimination reaction time task. It was shown that the simple winner-take-all-type neural network model receiving transient and sustained stimulus information from the primary visual cortex (V1) successfully reproduced RT distributions for correct responses and error rates. Moreover, temporal spike sequences obtained from the model network closely resembled to the neural activity in the monkey prefrontal or parietal area during other visual decision tasks such as motion discrimination and oddball detection tasks.

---

*Correspondence to:* A.Iwaizumi

e-mail: akiko@ecei.tohoku.ac.jp,

Tel.:+81-22-7957059, Fax:+81-22-2639307



## 1 Introduction

Measuring reaction time (RT) in sensory discrimination or detection tasks is one of the most popular methods to investigate the information system of the brain from sensory to higher-order processing, because response speed and trial-to-trial fluctuation of RT are considered to reflect the neural processing for decision formation in the brain. So far, many psychological and computational models have tried to explain responses that depend on various stimulus factor such as luminance, contrast, spatial frequency, etc. The most typical is ‘Evidence accrual models (Sequential-sampling models)’, in which the information of stimulus, ‘evidence’, is collected and integrated, and discrimination is completed when the ‘evidence’ reaches a threshold (cf. Rouder 2000; Smith and Ratcliff 2004; Ratcliff 2002). Among those models, the diffusion models and several leaky competing accumulator models successfully reproduced RT and accuracy, or error rate in various discrimination tasks (cf. Ratcliff 2002; Ratcliff and Smith 2004; Usher and McClelland 2001), however, they were insufficient to account for neural mechanism that enables perceivers to discriminate stimuli.

Meanwhile, neural activity showing similar temporal development to the ‘evidence’ in accumulator models has been found in monkey’s parietal and prefrontal area such as the lateral intraparietal area (LIP) and the frontal eye fields (FEFs) during visual motion discrimination tasks in which a monkey reported its decision by saccadic eye movements (Roitman and Shadlen 2002; Schall and Bichot 1998; Shadlen and Newsome 1996; Smith and Ratcliff 2004). According to the studies, the neural firing rates in those decision-related areas increased gradually; the saccadic eye movements then occurred after a fixed period when the firing rates reached a threshold. Moreover, response speed and monkey’s decision seemed to be predicted by the neural activity of those decision-related areas.

As for the real neural circuits in the brain which accomplish to integrate and discriminate sensory stimuli, several models have given plausible explanations for the relation between neural activity of decision-related areas and perceptual decisions (Mazurek et al. 2003; Machens et al. 2005; Shadlen et al. 1996; Wang 2002). For example, monkey’s psychometric function and the neural activity in LIP during visual motion discrimination tasks were reproduced by the biophysically realistic neural network model that included fast and slow recurrent excitation and feedback inhibition (Wang 2002). But these models did not address the detailed description of RT distributions for both correct and error

responses. The models are mainly for neural computation in fixed-delay tasks where participants or monkeys need to keep their decisions during delay period before response. Sensory stimuli represented in their models are relatively too simplified compared to the real activity in sensory cortices; these studies are, therefore, insufficient to discuss the contribution of various response dynamics in sensory cortices to decision formation.

Kinoshita and Komatsu (2001) recorded strong transient visual responses at the onset and offset of a pulsed surface stimulus and sustained activity accompanying luminance intensity during stimulus presentation in monkey's primary visual cortex, V1. Moreover, it has been suggested that these V1 activities are involved with the dependence of brightness perception on the temporal context between two neighboring stimuli (Eagleman et al. 2004). Therefore, we should investigate more carefully the relation between sensory response dynamics and perception.

In this paper, we conducted a psychophysical experiments of luminance-discrimination reaction time tasks, in which participants discriminated two flashes with a slight difference in luminance. They responded more rapidly and with more errors in shorter stimulus duration (max 150 ms) than they did in longer stimulus duration (max 1 second). Next, assuming that the brain utilizes not only sustained information but also transient information of the stimuli, we proposed a neurally appropriate network model that receives both transient and sustained inputs representing sensory responses in V1 and that comprises fast and slow excitation as introduced in Wang's model. The predicted data from the model such as changes in shape of RT distributions and error rates that occur across conditions are consistent with those of our experiments. The spike sequences produced by the model have common features to those of neural activity in the decision-related areas during visual motion discrimination tasks. According to the model, more rapid responses in shorter stimulus presentation is due to the transient input at the offset of stimuli, and smaller difference of synaptic input frequencies between the stronger and weaker flashes results in an increase of error responses. Furthermore, we evaluated the power ratio of the transient and sustained inputs sufficient for our experimental results by modulating relative synaptic strengths of the two inputs.

## 2 Psychophysical experiments of luminance discrimination

### 2.1 Methods

Five university students (GA, SS, NS, SY, KT, YS; three females and three males, mean age 24.2 years ranging from 22 to 26 years) participated in the experiment. All had normal or corrected-to-normal vision. Each participant was required to compare two flashes appearing simultaneously on the right and the left of a fixation point (FP), keeping his/her gaze to the FP and to press a button that corresponds to the perceived brighter flash as rapidly and accurately as possible using his/her index or middle finger. Since the stronger flash was presented randomly on each side with equal probability, the difference in response speed between those two fingers was counterbalanced, if any. The participants were not forced to respond; they did not press a button if they could not recognize the luminance difference between the two flashes.

Green light-emitting diodes (LEDs) with dim white covers set on a grey board (20.2 degrees high and 14.4 degrees long) were used for stimulus and the FP. Luminance of the board was 3 cd/m<sup>2</sup> when all the LEDs were extinguished. As in Fig. 1, the two LEDs for stimulus (0.57 degrees each) were placed at the left and the right (2.7 degrees) of the FP.

The two flashes of those stimulus LEDs had a slight difference in luminance on each trial. In 'hard' condition, the stronger luminance was 10.3 cd/m<sup>2</sup> and the weaker was 7.7 cd/m<sup>2</sup>, while in 'easy' condition the stronger was 10.8 cd/m<sup>2</sup> and the weaker was 7.2 cd/m<sup>2</sup>; the average luminance of two flashes was fixed to 9 cd/m<sup>2</sup> for both conditions. The luminance of the FP was 20 cd/m<sup>2</sup>. A luminance meter (Minolta, LS-100) was used for measurement of luminance.

On each trial, the two flashes were presented simultaneously to a participant after variable intervals (ranging 600 - 1000 ms, 800 ms on average) of FP presentation (Fig. 1b). The pulsed stimulus flashed for 150 ms in the 'short' conditions and for 1 second in the 'long' conditions. Irrespective of the stimulus duration, all lights were extinguished at the time the participant responded to the stimulus. Four conditions (the two luminance differences and the two stimulus durations) were chosen in random order through the task. It took 4 seconds for one trial. Each participant took an appropriate break after every 100 trials and performed approximately 800 trials, which resulted in approximately 200 responses for each condition.

Stimulus presentation and collection of responses were controlled by a laptop computer (CPU=700MHz, RTLinux) through an analogue multi-function I/O card (Contec AD12-8PM) with a sampling time of 100  $\mu$ s. If the participant mistakenly pressed a button inconsistent with his/her decision, the response was excluded as a motor error by pressing a cancellation button by the end of the trial. The time from stimulus onset

to the button pressing was defined as an RT at each trial. RTs that exceeded 700 ms were removed from data analysis. An RT distribution as a probability density function for each stimulus condition was calculated by fixed-width histogram estimate with smoothing parameter of 2 ms (Zandt TV, 2000). An error rate for each condition represents the ratio of error responses to the total responses excluding the motor errors. An answering rate is defined as the ratio of response occurrences to total trials for each condition.

## 2.2 Results

Mean RT and standard deviation (SD) of RT on correct responses for each participant are presented in Fig. 2. RT distributions of three participants in the hard condition are shown in Fig. 3. As can be seen from Fig. 2 and 3, the participants responded faster in the short condition than they did in the long condition ( $**p < .01$ ,  $F = 126.1$ , Two-factor ANOVA), and the SDs were also bigger in the short than in the long condition ( $**p < .01$ ,  $F = 29.8$ ). Tolhurst (1975) also found smaller RT variability in shorter stimulus presentation through the luminance detection tasks of sinusoidal gratings; he used near threshold contrast. On the other hand, suprathreshold stimulus was used for our experiment, therefore the answering rates of most of participants reached nearly 100 % in both short and long conditions (see Fig. 3).

As is seen from Fig. 4, the shorter stimulus duration brought about more error responses; the difference that depends on stimulus duration was statistically significant ( $**p < .01$ ,  $F = 45.1$ ). As the task gets easier, the difference between the short and the long conditions of error rates would be reduced.

In the following chapter we discuss the dependency of luminance discrimination performance on stimulus durations, providing a simple neural network model based on the response dynamics of visual cortex.

## 3 Modeling of luminance discrimination

### 3.1 Methods

**3.1.1 Visual stimuli** The brain should have different channels that encode specific stimulus attributes such as intensity; it should select among these channels as sensory information which is necessary to

execute a task, and integrate them to decision formation. Tolhurst (1975) predicted the contribution of a transient pathway encoding sudden changes of stimulus intensity under luminance detection tasks. Moreover, several computational models for luminance detection and discrimination tasks have been proposed, in which both transient and sustained information are crucial (e.g. Rouder, 2000).

From electrophysiological recordings of surface-responsive neurons in V1 during surface luminance stimulation, the transient and sustained responses have been observed; those neurons responded sharply at the onset and the offset of an uniform surface stimulus and show sustained activity during stimulus presentation (Kinoshita and Komatsu, 2001). The intensity of both transient and sustained responses varies as a function of luminance when the surface luminance is stronger than that of its surround.

How did these two different response dynamics influence luminance discrimination in our reaction time task? It seems reasonable to suppose that a transient information accompanying the stimulus offset would influence at least the short condition but not the long condition, because all participants completed their responses within 700 ms from the onset on almost all trials whereas the flashes lasted 1 second unless they pressed a decision button in the long condition (see answering rate in Fig. 3). Thus transient, offset-responses would be essential as well as onset-responses for discrimination in the short condition.

Although luminance processing begins in the retina, there is no doubt that luminance discrimination takes place in the higher-order cortices above V1 as suggested in other visual discrimination tasks (Shadlen and Newsome, 1996; Romo and Salinas, 2003; Eagleman et al., 2004 etc). Thus the neural responses to luminance in V1 would indicate more valid or essential sensory information that the higher cortices use for decision formation than the responses in the retina or the lateral geniculate nucleus (LGN).

Although most responses of single surface-responsive neurons in V1 have both transient and sustained components to a pulsed surface stimulus, completely separated transient and sustained responses were prepared for sensory information in our model to study contributions of the respective responses to luminance discrimination. The dependence of response intensity on luminance in the transient and sustained pathways was, on the other hand, based on the data from surface-luminance

stimulation by Kinoshita and Komatsu (2001). Time courses of the transient and sustained responses derived from our model are depicted in Fig. 5 (also see Appendix). Interspike intervals (ISIs) with certain average spike frequency that depends on luminance are noisy on every trial in identical stimulus conditions; the ISIs obey inhomogeneous Poisson process (Dayan and Abbott, 2001; Koch and Segev, 1998).

*3.1.2 Neural network model for luminance discrimination* The model network realizes a typical winner-take-all typed structure representing the cerebral cortices higher than V1, such as LIP. As indicated in Fig. 6, two groups (1 and 2) comprising excitatory and inhibitory neurons compete each other through mutual inhibition, according to the essential neural computation for comparison and discrimination of stimuli (for example, Machens et al., 2005). The excitatory neurons  $E_{ij}$  ( $i = 1$  or  $2$ ,  $j = 1$  to  $20$ ) in each group receive sensory input through AMPA-mediated synapses and excite other E neurons through AMPA- and NMDA- mediated synapses. The interneurons  $I_{ij}$  ( $i = 1$  or  $2$ ,  $j = 1$  to  $5$ ) receiving excitatory inputs from E neurons of the same group  $i$  weakly inhibit E neurons in the group  $i$  and strongly those in the opposite group  $j$  ( $j \neq i$ ) through GABA-mediated synapses. Neurons  $Y_i$  ( $i = 1$  or  $2$ ) are introduced for detection of strong network activity by excitation from all E neurons of the group  $i$  through AMPA- and NMDA- mediated synapses. Patterns of sparse connections between  $E_{ij}$  neurons or  $E_{ij}$  neurons and  $I_{ij}$  neurons are randomized. Since our goal here is to provide an essential and sufficient model that explains the neural computation of the luminance-discrimination reaction time task, detailed description of neuronal or synaptic mechanisms and cortical connectivity were beyond the scope of our study.

It can be assumed that each group (1 or 2) receives sensory inputs corresponding to either stronger or weaker flash respectively because the two flashes were fully separated (5.3 degrees) from each other; each stimulus would be, therefore, encoded in the different places at least in V1. Although in our experiments the stronger flash appeared in the left or the right side of the FP, for simplicity, the side for the stronger flash was fixed - that is, the group 1 always received the stronger flash and group 2 received the weaker flash.

*3.1.3 Decision rule* In the experiments, participants were instructed to press a decision button at the time he/she discriminated the brighter flash by using the specified fingers; the index finger to the left and the middle finger to the right. Thus a restricted area that controls the movement of either finger in the motor cortices would be activated when a participant started the response. In the model, the firing of neuron  $Y_i$  (see Fig. 6) is assumed to correspond to the origin of motor program or motor signal. We defined an RT as the first firing time of neuron  $Y_i$  ( $i = 1$  or  $2$ ) from the stimulus onset. Since excitatory spike input representing the stronger flash is always presented to the group 1, a decision is correct when neuron  $Y_1$  first starts firing, then it becomes an error response when neuron  $Y_2$  first starts firing.

## 3.2 Results

*3.2.1 Reproduction of discrimination properties* Mean and standard deviation (SD) of RT derived from the model and the experimental data are shown in Fig. 7a, respectively; each value is an average of five different patterns of the random interconnections in the model and of six participants in the experiment, respectively. The dependence of mean RT and response variability on stimulus duration derived from the model is consistent with that of the experiments. In the model, the faster and the less variable responses in the short presentation are due to the transient input at the stimulus offset. The predicted mean RTs are, however, approximately 170 ms shorter than those of the experiment. The reason for this would be that the model only expresses encoding and discrimination process of stimuli but not including motor programming and signal transmission to the peripheral nerves and musculo-skeletal system. In view of the fact that the motor preparation time was optimized to be approximately 200 ms through neural activity in monkey's LIP during motion detection task (Cook and Maunsell, 2002), the discrepancy between the predicted and experimental values of mean RT would be compensated.

Regarding error rates, the model also reproduces the similar features to those of the experiment: the number of error responses increases in the shorter presentation (Fig. 7b). This would be due to the frequency difference of the stronger and the weaker synaptic inputs associated with stimulus duration in each sensory pathway; the bigger the frequency difference becomes, the more easily the activity of

group 1 that receives stronger input grows enough to activate  $Y_1$  neuron, which results in less error responses. In the short condition, the frequency difference in a sustained pathway grows first but starts to decrease approximately at 150 ms due to the stimulus offset, while the difference still grows in the long condition at least about 150 ms (see thick lines in Fig. 5 a and b, and continuous lines in Fig. 8). This is one reason for error increase in the short condition. Another reason lies in the frequency difference in transient inputs since sufficient amount of the difference can only be found at limited phase of the onset and offset of stimulus (see thin lines in Fig. 5 a and b, and broken lines in Fig. 8). In simulations, the synaptic strength of the transient pathways  $C_{tra}$  was set to 36 % of that of the sustained pathways  $C_{sus}$  ( $C_{tra}=0.35$ ,  $C_{sus}=0.97$ ). Since a weak connection causes low spike sensitivity, the frequency difference of the stronger and the weaker inputs would have only a little effect on E neurons. In the following chapter we discuss how strengths of synaptic input from the transient and sustained pathways to the network affect response characteristics that the model predicts.

Although the model successfully reproduced the dependence of error rates on stimulus durations, the predicted values were rather inconsistent with experimental ones especially on the hard condition. One possible explanation would be that the model estimates smaller frequency difference of visual responses between the stronger and weaker flashes compared with real ones in human V1. Since the level of error rate level can be controlled by modulating the frequency difference in our model, this difference between the model and the experiments is, therefore, not a critical point to be addressed.

*3.2.2 Neural activities during stimulus discrimination* Fig. 9 shows the time courses of average activity functions among neurons  $E_{i,j}$  in respective groups (for the methods, see Appendix). Both activities of two groups start rising at the onset, then, on correct responses (thick lines in Fig. 9), the activity of group 1 (thick continuous line) sharply increases while the activity of group 2 (thick broken line) is suppressed. On the other hand, an opposite characteristic is seen on error responses (thin lines in Fig. 9). In this case, the rising rate in early period is smaller than that on correct responses, which results in longer mean RT (indicated by perpendicular lines within Fig. 9). This would be due to the weaker sensory inputs to the group 2 compared with those to the group 1. It should be noted that such temporal activities have been observed in decision-related cortices during motion discrimination (Roitman and Shadlen, 2002; Schall and Bichot, 1998; Shadlen and Newsome, 1996; etc) and oddball



(Murthy et al., 2001) tasks, although the activities show a small dip immediately after the onset in those cortices, which is not observed in our model. As a whole, our model can predict not only behavioral response properties but also ramped-like activity of decision-related cortices during visual discrimination.

*3.2.3 The role of slow integration through NMDA-mediated synapses* The decaying time constant of NMDA-mediated synapse is much longer than that of AMPA-mediated synapse (see e.g., Dayan and Abbott, 2001): in our simulations,  $\tau_{\text{NMDA,d}} = 100$  ms and  $\tau_{\text{AMPA}} = 2$  ms. The slow time constant like NMDA-type synapse enables the recurrent network to integrate firing information in longer past period, which is impossible for the network with AMPA-mediated excitation alone. Wang (2002) introduced NMDA-mediated synapses to a biophysically realistic neural network model that reproduced both neural activities in LIP and monkey’s psychometric function during random dot motion discrimination tasks (Shadlen and Newsome, 2001), in which a monkey was required to maintain its decision during a fixed delay period before the response by saccadic eye movement to the target. He suggested that slow integration would be also important for reaction time tasks like our experiments. In our model, if NMDA-mediated synapses are not included, synaptic connections of AMPA-mediated synapses need to be much stronger to ensure signal propagation. In this case, however, mean RT and trial-to-trial variability in RT get much smaller than the experimental values, while precise data are not shown in the present paper. In addition, too strong connections of AMPA- and NMDA-mediated synapses result in too rapid excitation within a group, which also diminishes the difference between mean RTs or error rates of the short and the long conditions.

*3.2.4 Synaptic parameters that modulate absolute RT* We confirmed that synaptic strengths between E neurons and  $Y_i$  neuron,  $C_{Y,\text{fast}}$  and  $C_{Y,\text{slow}}$ , did not affect fundamental features of discrimination if they were within a proper range. For example, 30 % loss of  $C_{Y,\text{fast}}$  did not affect RT variability (SD) and error rate, but mean RTs became longer approximately by 12 ms due to the weak connection to neuron  $Y_i$ . Another example is that one and a half as strong as normal strength ( $C_{Y,\text{fast}} = 1.46$ ) led to shorter mean RT approximately by 13 ms while SD and error rate were unchanged. Further

intensification of the synaptic strength, however, diminished the differences of RT and error rate between the short and the long conditions.

## 4 Discussion

### *4.1 Influence of dynamics in sensory response upon luminance perception*

From the psychophysical experiments and modeling study, as we have seen, dynamics in sensory representation in the brain influences luminance discrimination. Recently Eagleman et al. (2004) reported an interesting phenomenon, ‘Temporal Context Effect, TCE’, in which the perceived brightness of a brief flash with identical luminance intensity was modulated as a function of the temporal context between the flash and the spatially neighboring object with constant luminance. They suggested that the TCE would be produced by adapting and non-adapting sensory responses seen among surface-responsive neurons in V1. Although the phenomenon would support our model in term of participation of detailed dynamics of sensory response, it needs further investigation about how temporal dynamics of sensory responses influence perception. Besides, if other kind of stimuli of visual or other sensory modalities cause the similar sensory responses as surface-responsive neurons of V1, we might get the same discrimination properties dependent on stimulus durations as our experiments.

Our model suggests that the faster responses for the short stimulus duration may be due to transient responses in the sensory cortices - that is, the experimental results are mainly based on the response dynamics in the sensory processing stage among the whole process including sensory encoding, comparison and decision, and motor programming. This assumption would be plausible because the participants could not switch their strategies corresponding to each stimulus duration in advance due to the random presentation of either the short or the long duration. However, we cannot completely exclude the possibility of the involvement of some perceptual biases. For example, the participant might have decided that he/she would make a response as rapidly as possible when he/she noticed a sudden disappearance of stimulus, while the rapid response due to the perceptual bias would be driven fundamentally by sensory information of the sudden disappearance. Since faster responses would be based on rapid increase in neural activity of the decision-related networks, thus, even without the transient inputs from the sensory cortex, the network will receive positive feedback

that has similar feature to the transient sensory responses from other cerebral cortices in the short condition under perceptual biases.

#### *4.2 Relation of synaptic strengths between transient and sustained pathways to task performance*

In simulations, as we noted in 3.2.1, the synaptic strength of transient pathways  $C_{tra}$  was 36 % of that of sustained pathways  $C_{sus}$  ( $C_{tra}=0.35$ ,  $C_{sus}=0.97$ ). Then how does the intensity difference of the two pathways influence decision formation in the luminance-discrimination reaction time task? To examine it quantitatively, we compared SD of RT and answering rate with seven different proportions of  $C_{tra}$  to  $C_{sus}$  from 0.1 to 0.7 in the hard condition, keeping the total synaptic strength to be 1.32 as used in previous simulations. As Fig. 10a indicates, the increase of transient strength results in lower RT variability for the short condition whereas inverse characteristic is observed in the long condition. The low variability in the short condition is due to the stronger transient inputs, especially those caused by stimulus offset, and the more dispersive RT in the long condition results from considerable weak sustained inputs. The low proportion of transient input strength such as 0.1 produces different discrimination features compared with the experiment, namely the larger RT variability for the longer stimulus duration. The answering rate derived from the model is also inconsistent with empirical value for the too low or too high proportions of transient input strength as seen in Fig. 10b. The drops in the answering rate for the short condition would be because the transient input in a part of trials was too weak to make the network activity reach the threshold. Similarly, in the long condition, the answering rate at low proportions of sustained input strength such as 0.6 and 0.7 decrease due to the considerable weak sustained inputs. From these considerations, it would be entirely valid to say that the brain uses not only the sustained information but also the transient information to process the luminance discrimination task, although the sustained inputs are main source. The appropriate ratio of transient input strength to sustained input strength predicted at least from the model will be within the range of 0.3 - 0.5.

#### *4.3 Relation between properties of the P and M pathways and luminance discrimination*

Many studies have attempted to determine the link between RTs and the neural mechanisms in the early stages of visual processing (Murray and Plainis, 2003; Burr and Corsale, 2001; Plainis and Murray, 2000). From physiological and psychological studies, it is well known that there are two main retino-cortical pathways, called the magnocellular (M) and parvocellular (P) pathways. The M and P neurons have distinct contrast gain properties, depending on luminance and spatial frequency of sinusoidal gratings (e.g. Murray and Plainis, 2003; Purpura et al., 1990; Purpura, Kaplan and Shapley, 1988). They are also different in the size of receptive fields (RFs). In this study, we did not design the experiment taking precisely account of neural interactions between neurons each RF of which covers distinct region of a flash, and the model has minimal network structure to demonstrate response characteristics in the experiment. Although it is hard to conduct a direct comparison between the M and P pathways and the model of neural processing for luminance-discrimination, we shall consider it for the moment.

It is well known that M neurons have high luminance contrast sensitivity (high contrast gain), but saturate at fairly low contrasts, on the other hand, P neurons have low contrast gain but specialize in extracting colour and, high spatial frequency information (Murray and Plainis, 2003; Purpura et al., 1990). Since the stimulus used in our experiments was chromatic and far above detection threshold, P neurons may contribute mainly to luminance discrimination. As many studies like those by Plainis and Murray (2000) show, M neurons almost exclusively contribute to stimulus detection under the condition of low luminance or low contrast, it remains uncertain how the processing of suprathreshold contrast influences RTs (Murray and Plainis, 2003).

Sensory responses in the transient and sustained pathways of the model at low luminance level have similar gain properties to those of the P and M pathways. According to equations (1) and (6) in Appendix, the response gain in the transient pathway at low luminance near background is relatively high, but that in the sustained pathway is extremely low. Therefore, for stimulus detection of such low luminance and low contrast, the transient pathway is crucial. As for this point the transient pathway can be regarded as M pathway and the sustained pathway as P pathway. At the low luminance, however, current response frequency and synaptic connection with decision network in the transient

pathway are too weak to excite the network activity above the decision threshold with probability at least above chance level. For example, we preliminarily measured RT to single stimulus of 4 or 5  $\text{cd/m}^2$  and obtained 100 % of answering rate, but the answering rate derived from the model dropped markedly. Since reaction time tasks are biased toward transient activity of the stimulus (Murray and Plainis, 2003), the buildup of synaptic strength from transient pathway would be one way of representing the above biases with almost perfect answering rates.

RT vs. the reciprocal of contrast ( $1/C$ ) functions are bi-linear in various spatial frequencies of gratings (Murray and Plainis, 2003). They account for the RT-contrast relationship by that at low contrasts only M neurons having high gain and fast response are activated, and as the contrast of the stimulus increases, P neurons are recruited. The similar relationship also can be expected at low luminance for the model; the gain of the transient response is constantly high whereas that of the sustained response increases with stimulus intensity. Since the synaptic strength of sensory input from the transient and sustained pathways to the decision network is tuned best to the task at mean luminance of 9  $\text{cd/m}^2$  in the present model, it needs further discussion about how the network adjusts luminance-discrimination tasks at various luminance levels, which is due to be dealt with in our study in the near future.

#### *4.4 Effects of instruction for participants*

In our experiments, participants were required to respond as rapidly and accurately as possible. Empirically, speed-stress instructions result in faster and less accurate responses, whereas accuracy-stress instructions result in the converse, which is called ‘speed-accuracy trade-off’ (Smith and Ratcliff, 2004; Reddi and Carpenter, 2000). The phenomenon can be explained computationally by controlling the level of a decision threshold according to instructions; low threshold for speed instructions and high threshold for accuracy instructions (Reddi and Carpenter, 2000).

Is it possible to incorporate the speed-accuracy trade-off in our model? One possible neural representation would be to modulate the synaptic strength of connections between sensory neurons and decision network, where the stronger connections correspond to speed-stress and the weaker connections correspond to accuracy-stress. The other would be introducing a global inhibition to excitatory

neurons of group 1 and 2; intensity of the inhibition will control a decision-threshold in the network activity. The stronger inhibition causes the network to integrate visual information more slowly and securely, which would correspond to accuracy-stress. This issue deserves to be studied more in detail but is beyond the scope of the present paper.

#### *4.5 Origin of RT variability*

Trial-to-trial variability of RT in our model depends on the variability in spike sequences at which frequency  $F_x(t)$  represents temporal intensity of sensory stimulus (see Appendix). The moment-to-moment and trial-to-trial fluctuation of sensory responses in MT (V5) is thought to be responsible for the variability of RT in visual motion discrimination tasks (Shadlen et al., 1996). Our model, therefore, is compatible with their argument. However, there should be also fluctuation of neural activity in parietal and prefrontal cortices, which are higher than sensory cortices. Thus more empirical and theoretical discussion will be needed to elucidate which fluctuation in the brain is essential or how each fluctuation influences decision and perception.

#### *4.6 Error distributions*

Our model succeeded in reproducing RT distributions on correct responses and error rates for two different stimulus durations and difficulties. Neural network models that explain the behavioral responses of discrimination tasks are, however, desirable to also describe RT distribution in error responses (Smith and Ratcliff, 2004). Our model produced error distributions of both short and long conditions as in Fig. 11; rightward shift of the error distributions in the short condition (Fig. 11a) is due to the small difference of input frequency between transient and sustained pathways. However, we could not compare the model and the experimental results because of too small empirical data: 10 error responses occurred approximately in 200 responses for each condition. Thus we should have needed to get more responses from the participants or to manipulate difficulty of discrimination in order to get enough error responses out of the limited trials.

## 5 Conclusions

Our network model based on the response dynamics in the primary visual cortex reproduced successfully behavioral responses and the dependencies on stimulus duration in the luminance-discrimination reaction time task. Moreover, the spike sequences of the model network closely resembled to decision-related cortical activities during other visual discrimination tasks. These results lead to the conclusion that our model is one of the neurally plausible models that link neurobiology and psychology. Models based on more detailed sensory responses would give us better understanding of decision making at the level of neural mechanism.

## Appendix

### *Spike sequence representing transient and sustained responses of V1*

The dynamical change of firing rate for the transient and sustained responses in surface-responsive neurons of V1 ( $F_{\text{tra},i}(t)$  and  $F_{\text{sus},i}(t)$ ,  $i = 1$  or  $2$ ) were modeled computationally, then these firing rates were converted into spike sequences for synaptic input to the decision network. A transient response obeyed the absolute first-order differential equation with second-order decay and a sustained response obeyed the first-order decay of a pulsed stimulation,

$$l_{\text{tra},i}(t) = 25\{\tanh(\log(L_i(t)/5)) + 1\} \quad (1)$$

$$dz_i(t)/dt = (l_{\text{tra},i}(t) - z_i(t))/\tau_{\text{tra}} \quad (2)$$

$$df_{1,i}(t)/dt = (|l_{\text{tra},i}(t) - z_i(t)| - f_{1,i}(t))/\tau_{\text{tra}} \quad (3)$$

$$df_{2,i}(t)/dt = (f_{1,i}(t) - f_{2,i}(t))/\tau_{\text{tra}} \quad (4)$$

$$F_{\text{tra},i}(t) = G \cdot f_{2,i}(t) \quad (5)$$

$$l_{\text{sus},i}(t) = 25\{\tanh(5\log(L_i(t)/13)) + 1\} + 17 \quad (6)$$

$$dF_{\text{sus},i}(t)/dt = (l_{\text{sus},i}(t) - F_{\text{sus},i}(t))/\tau_{\text{sus}} \quad (7)$$

where  $L_i(t)$  represented the magnitude of luminance ( $L_1 = 10.3\text{cd/m}^2$ ,  $L_2 = 7.7\text{cd/m}^2$  for the hard condition and  $L_1 = 10.8\text{cd/m}^2$ ,  $L_2 = 7.2\text{cd/m}^2$  for the easy condition), decay constants of transient and sustained pathways  $\tau_{\text{tra}}$ ,  $\tau_{\text{sus}}$  were 30, 100 ms, respectively. The gain constant  $G$  was 11. The firing frequencies  $F_{x,i}(t)$  were

transformed into spike sequences according to the inhomogeneous Poisson process by ‘thinning technique’ with absolute refractory period of 5 ms (Dayan and Abbot, 2001).

### *Neural network model*

For sensory input to the decision network, 20 % of  $E_{ij}$  neurons in group  $i$  ( $i=1$  or  $2$ ) received two independent excitatory synaptic inputs respectively; both inputs represented identical transient or sustained information about stronger or weaker flash. Every  $E_{ij}$  neuron connected with other half of  $E_{ij}$  neurons that were randomly-chosen, and every  $I_{ij}$  neuron was excited by 20 % of  $E_{ij}$  neurons in the same group. Each  $E_{ij}$  neuron was weakly inhibited by 20 % of  $I_{ij}$  neurons in the same group and strongly inhibited by 20 % of  $I_{kj}$  neurons ( $k \neq i$ ) in the opponent group. In simulations, each group was constituted by twenty  $E_{ij}$  neurons and five  $I_{ij}$  neurons respectively. Patterns of sparse connections between E-to-E neurons or E-to-I neurons were randomized. Neurons  $Y_1$  and  $Y_2$  received excitatory inputs from all  $E_{ij}$  neurons of group 1 or 2, respectively.

### *Single-neuron model*

All neurons were described by conductance-based, leaky integrate-and-fire, single compartment model that was characterized by a resting potential  $V_r = -70$  mV, a firing threshold  $V_{th} = -50$  mV, a reset potential  $V_{reset} = -70$  mV, a membrane time constant  $\tau_m = 20$  ms and a absolute refractory period  $\tau_{ref} = 2$  ms. The temporal change of membrane potential  $V_i(t)$  of neuron  $i$  was determined by,

$$\tau_m \frac{dV_i(t)}{dt} = -(V_i(t) - V_r) + I_{syn}(t) \quad (8)$$

where  $I_{syn}(t)$  represented the total postsynaptic current flowing into the neuron.

### *Synapses*

Excitatory postsynaptic currents (EPSCs) with regard to E neuron-to-E neuron and E neuron-to-Y neuron connections have two components mediated by AMPA and NMDA receptors, respectively, although the EPSCs for sensory inputs are mediated exclusively by AMPA receptors. The inhibitory synaptic currents are described by GABA receptors. Thus the total postsynaptic currents in the network are given by

$$I_{syn,E}(t) = I_{tra}(t) + I_{sus}(t) + I_{fast}(t) + I_{slow}(t) + I_{EI,r}(t) + I_{EI,l}(t) \quad (9)$$

$$I_{syn,I}(t) = I_{IE}(t) \quad (10)$$

$$I_{syn,Y}(t) = I_{Y,fast}(t) + I_{Y,slow}(t) \quad (11)$$



for an  $E_{ij}$ ,  $I_{ij}$ ,  $Y_i$  neurons respectively, in which

$$I_{tra}(t) = C_{tra} A_{AMPA} (V_i(t) - V_E) \sum_{j=1}^2 s_j^{AMPA}(t) \quad (12)$$

$$I_{sus}(t) = C_{sus} A_{AMPA} (V_i(t) - V_E) \sum_{j=1}^2 s_j^{AMPA}(t) \quad (13)$$

$$I_{fast}(t) = C_{fast} A_{AMPA} (V_i(t) - V_E) \sum_j s_j^{AMPA}(t) \quad (14)$$

$$I_{slow}(t) = \frac{g_{NMDA} A_{NMDA} (V_i(t) - V_E)}{(1 + [Mg^{2+}] \exp(-0.062 V_i(t)) / 3.57)} \cdot C_{slow} \sum_j s_j^{NMDA}(t) \quad (15)$$

$$I_{EI,r}(t) = C_{EI,r} A_{GABA} (V_i(t) - V_i) \sum_{i=j} s_j^{GABA}(t) \quad (16)$$

$$I_{EI,l}(t) = C_{EI,l} A_{GABA} (V_i(t) - V_i) \sum_{i=j} s_j^{GABA}(t) \quad (17)$$

$$I_{IE}(t) = C_{IE} A_{AMPA} (V_i(t) - V_E) \sum_j s_j^{AMPA}(t) \quad (18)$$

$$I_{Y,fast}(t) = C_{Y,fast} A_{AMPA} (V_i(t) - V_E) \sum_j s_j^{AMPA}(t) \quad (19)$$

$$I_{Y,slow}(t) = \frac{g_{NMDA} A_{NMDA} (V_i(t) - V_E)}{(1 + [Mg^{2+}] \exp(-0.062 V_i(t)) / 3.57)} \cdot C_{Y,slow} \sum_j s_j^{NMDA}(t) \quad (20)$$

where  $V_E = 0$  mV,  $V_i = -77$  mV,  $[Mg^{2+}] = 1$  mM (see Dayan and Abbott, 2001; Wang, 2002). The  $C_x$  represents the strength of synaptic connection where  $C_x = 1$  indicates the threshold sufficient for a neuron to fire only by single spike mediated through a certain receptor type  $X$ , and the numerical constant  $A_x$  represents the scaling parameter to standardize the synaptic strength to  $C_x$ . The synaptic strengths for interconnections in the network were modulated to make the gradual buildup of the activity (like Fig.9). Following values were used for computer simulations.  $C_{tra} = 0.35$ ,  $C_{sus} = 0.97$ ,  $C_{fast} = 0.06$ ,  $C_{slow} = 0.043$ ,  $C_{IE} = 0.25$ ,  $C_{EI,r} = -0.05$ ,  $C_{EI,l} = -0.1$ ,  $C_{Y,fast} = 0.065$ ,  $C_{Y,slow} = 0.015$ ,  $A_{AMPA} = 446.2$ ,  $A_{NMDA} = 56.3$ , and  $A_{GABA} = -2230.1$ .

The time courses of gating variables  $s^X$  caused by an activation of AMPA, NMDA and GABA receptors respectively were described as follows,

$$\frac{ds_j^{AMPA}(t)}{dt} = -\frac{s_j^{AMPA}(t)}{\tau_{AMPA}} + \sum_k \delta(t - t_j^k) \quad (21)$$

$$\frac{ds_j^{NMDA}(t)}{dt} = -\frac{s_j^{NMDA}(t)}{\tau_{NMDA,d}} + \alpha x_j(t) (1 - s_j^{NMDA}(t)) \quad (22)$$

$$\frac{dx_j(t)}{dt} = -\frac{x_j(t)}{\tau_{NMDA,r}} + \sum_k \delta(t - t_j^k) \quad (23)$$

$$\frac{ds_j^{GABA}(t)}{dt} = -\frac{s_j^{GABA}(t)}{\tau_{GABA}} + \sum_k \delta(t - t_j^k) \quad (24)$$

with  $\tau_{AMPA} = 2$  ms,  $\tau_{NMDA,r} = 2$  ms,  $\tau_{NMDA,d} = 100$  ms,  $\alpha = 0.5$  ms<sup>-1</sup>, and  $\tau_{GABA} = 5$  ms.

*Neural activity function*

An average activity function of group  $i$  was obtained by counting total spikes emitted by all E neurons within the time window of past 25 ms from present time  $t$ , slid with a time step of 5 ms, and being multiplied by 40.

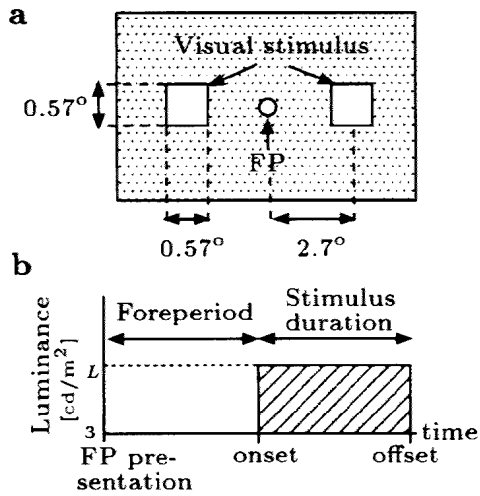
*Simulations*

Computer simulations were run on a Unix workstation using a third order Runge-Kutta method, with a time step  $dt=0.01$  ms. For a given parameter set, discrimination properties such as RT distribution and error rate for each stimulus condition were derived from an average of five sets comprising 300 independent trials respectively; different random interconnections within a group were prepared for each set. An RT and a decision result (correct or error response) of the model for each trial were obtained from the output of the model network within total time steps of 550 ms.

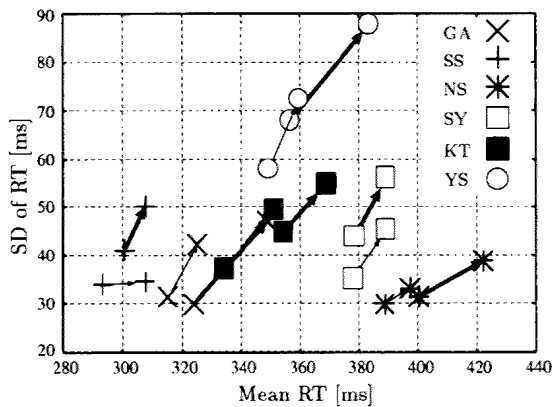
## References

1. Burr DC and Corsale B (2001) Dependency of reaction times to motion onset on luminance and chromatic contrast. *Vision Res* **41**:1039-1048
2. Cook EP and Maunsell JHR (2002) Dynamics of neural responses in macaque MT and VIP during motion detection. *Nat Neurosci* **5**: 985-994
3. Dayan P and Abbott LF (2001) *Theoretical Neuroscience; Computational and Mathematical Modeling of Neural Systems*. The MIT Press, Cambridge, Massachusetts London, England
4. Eagleman DM, Jacobson JE and Sejnowski TJ (2004) Perceived luminance depends on temporal context. *Nature* **428**:854-856
5. Kinoshita M and Komatsu H (2001) Neural representation of the luminance and brightness of a uniform surface in the macaque primary visual cortex. *J Neurophysiol* **86**:2559-2570
6. edited by Koch C and Segev I (1998) *Methods in Neural Modeling: From Ions to Networks*. The MIT Press
7. Mazurek ME, Roitman JD, Ditterich J and Shadlen MN (2003) A role for neural integrators in perceptual decision making. *Cereb Cortex* **13**:1257-1269
8. Machens CK, Romo R, Brody CD (2005) Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* **307**:1121-1124
9. Murray IJ and Plainis S (2003) Contrast coding and magno/parvo segregation revealed in reaction time studies. *Vision Res* **43**:2707-2719
10. Murthy A et al. (2001) Dynamic dissociation of visual selection from saccade programming in frontal eye field. *J Neurophysiol* **86**:2634-2637
11. Plainis S and Murray IJ (2000) Neurophysiological interpretation of human visual reaction times: effect of contrast, spatial frequency and luminance. *Neuropsychologia* **38**:1555-1564
12. Purpura K, Kaplan E and Shapley M (1988) Background light and the contrast gain of primate P and M retinal ganglion cells. *Proc Natl Acad Sci USA* **85**:4534-4537
13. Purpura K et al. (1990) Light adaptation in the primate retina: analysis of changes in gain and dynamics of monkey retinal ganglion cells. *Vis Neurosci* **4**:75-93
14. Ratcliff R (2002) A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychon Bull Rev* **9**:278-291
15. Ratcliff R and Smith PL (2004) A comparison of sequential sampling models for two-choice reaction time. *Psychol Rev* **111**:333-367
16. Reddi BAJ and Carpenter PHS (2000) The influence of urgency on decision time. *Nat Neurosci* **3**:827-830

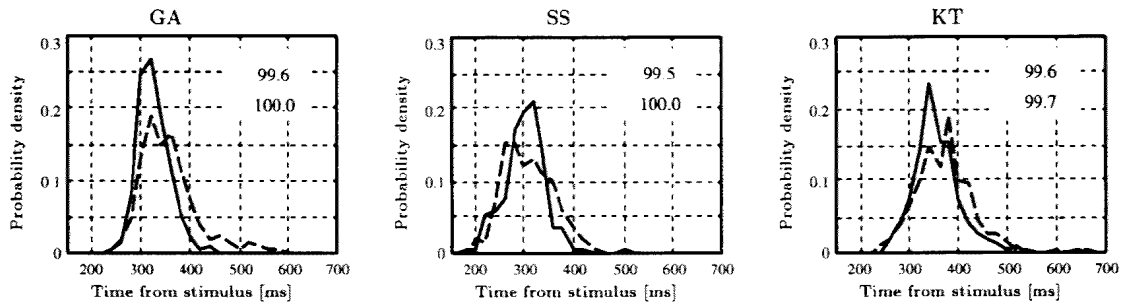
17. Roitman JD and Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci* **22**:9475-9489
18. Romo R and Salinas E (2003) Flutter discrimination: Neural codes, perception, memory and decision making. *Nat Rev Neurosci* **4**: 203-218
19. Rouder JN (2000) Assessing the roles of change discrimination and luminance integration: evidence for a hybrid race model of perceptual decision making in luminance discrimination. *J Exp Psychol Hum Percept Perform* **26**:359-378
20. Shadlen MN et al. (1996) A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci* **16**:1486-1510
21. Shadlen MN and Newsome WT (1996) Motion perception: Seeing and deciding. *Proc Natl Acad Sci USA* **93**:628-633
22. Shadlen MN and Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* **86**:1916-1936
23. Schall JD and Bichot NP (1998) Neural correlates of visual and motor decision processes. *Curr Opin Neurobiol* **8**:211-217
24. Smith PL and Ratcliff (2004) Psychology and neurobiology of simple decisions. *Trends Neurosci* **27**:161-168
25. Tolhurst DJ (1975) Reaction times in the detection of gratings by human observers: a probabilistic mechanism. *Vision Res* **15**:1143-1149
26. Usher M and McClelland JL (2001) The time course of perceptual choice: the leaky, competing accumulator model. *Psychol Rev* **108**:550-592
27. Wang XJ (2002) Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**:955-968
28. Zandt TV (2000) How to fit a response time distribution. *Psychon Bull Rev* **7**:424-465



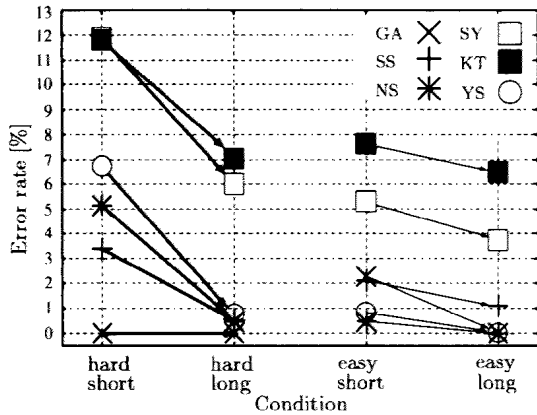
**Fig. 1** Visual stimulus and its time course. (a) Two LEDs for visual stimulus were spatially arranged on either side of the fixation point (FP). Two stimulus LEDs flashed simultaneously after variable foreperiod (ranging 600 - 1000 ms, 800 ms on average) of the FP presentation. (b) Luminance of stimulus increased abruptly to certain magnitude of  $L$ , keeping the intensity level during a fixed period ('short' or 'long'), and dropped to the baseline ( $3\text{cd}/\text{m}^2$ ).



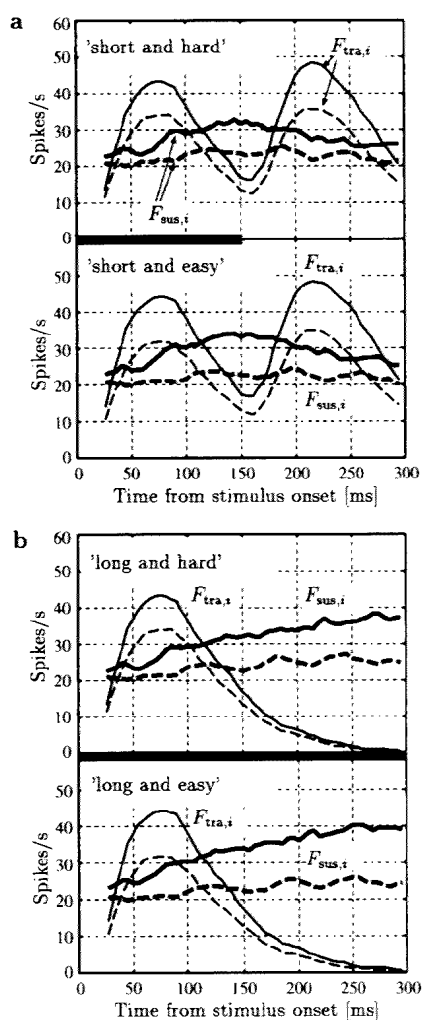
**Fig. 2** Mean and standard deviation (SD) of individual RT distributions. The abscissa represents mean RT and the ordinate represents SD. Data of each participant is represented by identical marks. Thick arrow denotes the hard condition with a mark at its origin representing the short condition and with a mark at its tip representing the long condition. Right upward arrows represent mean and variance of RT are bigger on the long condition.



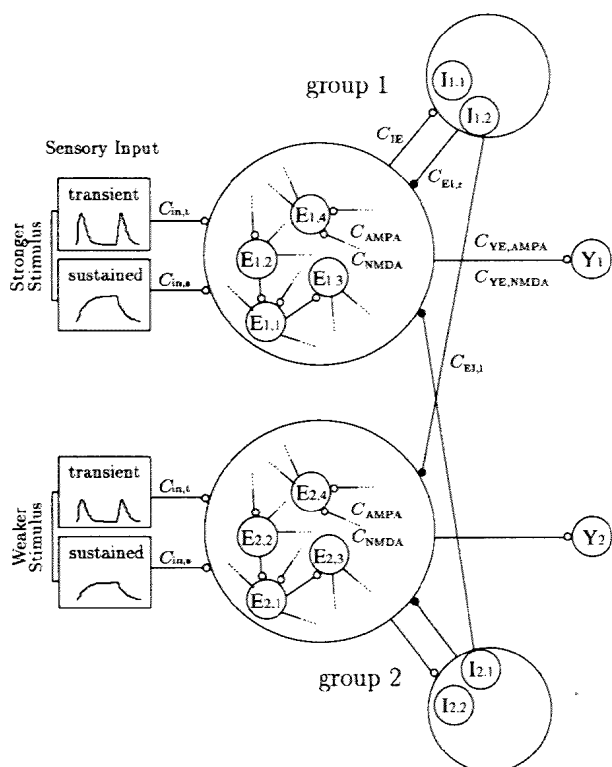
**Fig. 3** Example of RT density distributions (in the hard condition) of three participants (GA, SS and KT). The abscissa represents time from stimulus onset and the ordinate represents probability density of certain RT (continuous lines for the short condition and broken lines for the long condition). Numbers inside windows represent answering rate (upper: the short condition, lower: the long condition).



**Fig. 4** Error rate associated with each condition. Individual data is represented by identical marks. Thick arrow denotes the hard condition with a mark at its origin representing the short condition and with a mark at its tip representing the long condition.

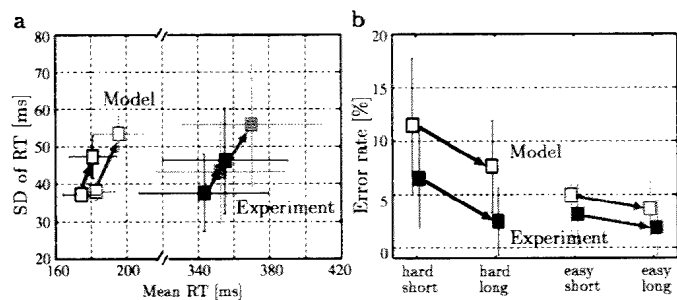


**Fig. 5** Example of time courses of sensory input frequencies,  $F_{tra,i}(t)$  and  $F_{sus,i}(t)$  ( $i = 1$  or  $2$ ), on the short condition **a** and the long condition **b** (thin lines for the transient input and thick lines for the sustained input). Continuous lines are for the stronger flash and broken lines for the weaker flash. The thick black line lying center within each figure indicates the period of stimulus presentation.

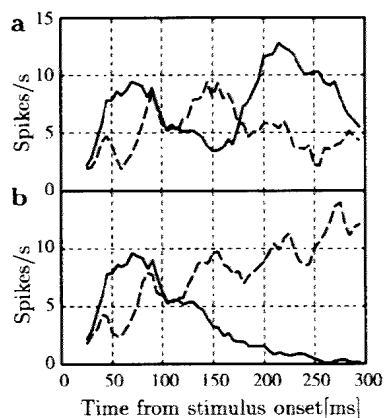


**Fig. 6** Structure of the neural network model. Neurons  $E_{ij}$  in each group  $i$  ( $i = 1$  or  $2$ ) are interconnected with half of other  $E_{ik}$  ( $k \neq j$ ) neurons in the same group through excitatory AMPA- and NMDA-mediated synapses. A part of  $E_{ij}$  neurons receives sensory inputs from transient and sustained pathways. For the sake of simplicity, we assume that neurons in group 1 always receive the inputs representing the stronger flash and neurons in group 2 receive the inputs representing the weaker flash. An  $E_{ij}$  neuron in group  $i$  sends excitatory input to several interneurons  $I_{ij}$  in the same group. Neuron  $Y_i$  is excited by all  $E_{ij}$  neurons of the group  $i$ . The faster firing of neuron  $Y_i$  than neuron  $Y_j$  ( $i \neq j$ ) represents the winning of the group  $i$ . Interneurons  $I_{ij}$  weakly inhibit  $E_{ij}$  neurons of the same group and strongly inhibit  $E_{kj}$  ( $k \neq i$ ) of the opponent group.

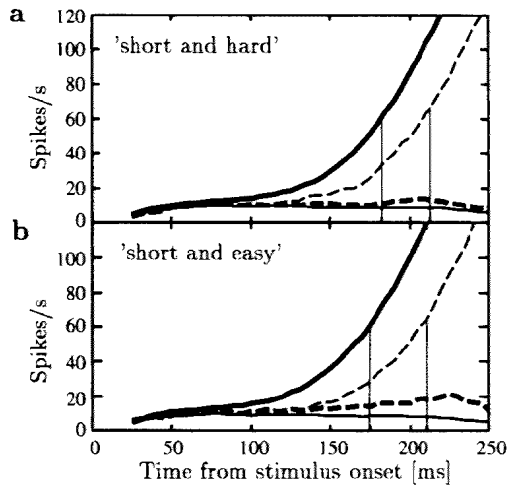




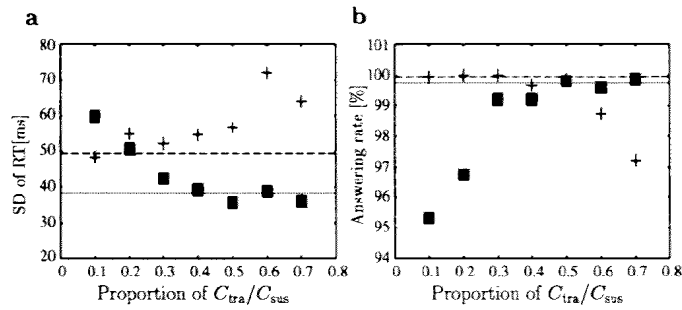
**Fig. 7** (a) Mean RT (the abscissa) and SD of RT (the ordinate) obtained from the experiments and the model. Black marks indicate the hard condition and grey marks signify the easy condition. A mark at the origin of arrow represents the short condition, and a mark at the tip represents the long condition. At both level of difficulty, an average RT predicted from the model is approximately 170 ms shorter than that of the experiments. (b) Error rate of the experiments and the model on each condition.



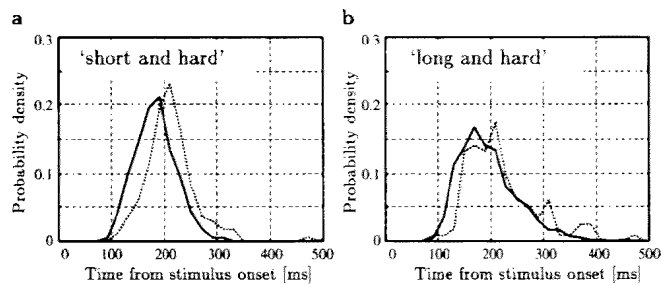
**Fig. 8** Example of time courses of input frequency difference between the stronger and weaker flashes for the transient (continuous line) and the sustained (broken line) pathways. (a) The hard and short condition. (b) The hard and long condition.



**Fig. 9** The average spike density functions of group 1 and 2 on the short and hard condition (a) and short and easy condition (b). Continuous lines indicate average activities of the group 1 and dashed lines indicate activities of the group 2. The thick lines stand for correct responses and thin lines for error responses. Two vertical lines represent average RT on correct and error responses, respectively.



**Fig. 10** SD (a) and answering rate (b) of seven different proportions of the transient input strength to the sustained input strength ( $C_{tra}/C_{sus}$ ) in the hard condition with squares representing data from the model on the short condition and crosses representing those on the long condition. Dotted lines indicate experimental values on the short condition and broken lines indicate that on the long condition.



**Fig. 11** RT distributions on correct (continuous lines) and error (dotted lines) in the short and hard (a) and the long and hard (b) conditions.

## 2刺激に微小な呈示時間差を持った輝度弁別 RT タスクの 弁別特性を有する神経回路モデル

岩泉亜希子<sup>†</sup> 二見 亮弘<sup>††</sup> 松木 英敏<sup>†</sup> 鈴木 陽一<sup>†††</sup>

<sup>†</sup> 東北大学 工学研究科 〒980-8579 仙台市青葉区荒巻字青葉 6-6-05

<sup>††</sup> 福島大学 共生システム理工学類 〒960-1296 福島県福島市金谷川 1

<sup>†††</sup> 東北大学 情報科学研究科 〒980-8577 仙台市青葉区片平 2-1-1

E-mail: †{akiko,matsuki,yoh}@ecei.tohoku.ac.jp, ††futami@sss.fukushima-u.ac.jp

**あらまし** 2刺激の輝度弁別タスクにおいて、より多様な刺激時間構造の輝度弁別における神経機構を探るため、2刺激のオンセットに微小な時間差 (Stimulus Onset Asynchrony, SOA) を持った輝度弁別 RT タスクを行った。そして、これまでの研究から提案された2刺激同時呈示の輝度弁別特性を説明する神経回路モデルに対して、感覚情報の不要な蓄積を禁止する機構と2刺激の非同時性を検出する機構を追加することで、モデルが2刺激同時・非同時呈示の両方の反応を再現するだけでなく、閾上単一刺激検出タスクの実験結果も説明することを示した。

**キーワード** 弁別, SOA, 反応時間, フィードフォワード抑制

## Model neural network for responses in two-choice luminance-discrimination RT task with stimulus onset asynchrony

Akiko IWAIZUMI<sup>†</sup>, Ryoko FUTAMI<sup>††</sup>, Hidetoshi MATSUKI<sup>†</sup>, and Yōiti SUZUKI<sup>†††</sup>

<sup>†</sup> Graduate School of Engineering, Tohoku University Azaaoba 6-6-05, Aramaki, Aoba-ku, Sendai,  
980-8579 Japan

<sup>††</sup> Faculty of Symbiotic Systems Science, Fukushima University Kanetanikawa 1, Fukushima-shi,  
Fukushima 960-1296 Japan

<sup>†††</sup> Graduate School of Information Sciences, Tohoku University Katahira 2-1-1, Aoba-ku, Sendai,  
980-8577 Japan

E-mail: †{akiko,matsuki,yoh}@ecei.tohoku.ac.jp, ††futami@sss.fukushima-u.ac.jp

**Abstract** Reaction time (RT) and error rate were influenced by small stimulus onset asynchrony (SOA) in a two-choice luminance-discrimination RT task. The dependence on the SOA was successfully reproduced by introducing global inhibitors to prevent unnecessary integration of sensory information, and neuron chains with feedforward mutual inhibition as a asynchronism detector of two flashes, into the previous model that Iwaizumi et al. proposed. The revised model not only reproduced responses of these two-choice luminance discrimination tasks both with and without the SOA but also reproduced those of suprathreshold single-stimulus detection tasks in a wide range of stimulus luminance.

**Key words** discrimination, SOA, reaction time, feedforward inhibition

### 1. はじめに

我々はこれまで、2刺激の輝度弁別反応時間 (RT) タスクにおいて、被験者の反応時間やエラー率が刺激時間長によって有意に変化することを見つけた。また、一次視覚野 (V1) で見られる輝度に対する神経活動 [3] を模擬した transient 応答と sustained 応答を感覚入力として受け取るパルス型 Winner-Take-All 型神経回路モデルにより、実験で得られた刺激時間長

に依存した反応特性を再現した [1]。具体的には、2刺激の呈示時間長が 150ms (‘short’ 条件) の場合、1s (‘long’ 条件) の場合に比べて平均反応時間とそのばらつきが小さくなる一方、エラー率は上昇した。しかし、一般的な輝度弁別処理の神経機構を考察する上で、一実験結果を説明する上記のモデルを提案するだけでは不十分であり、タスクにおける刺激面積や刺激時間構造などさまざまな刺激要因に対する反応特性を神経回路レベ

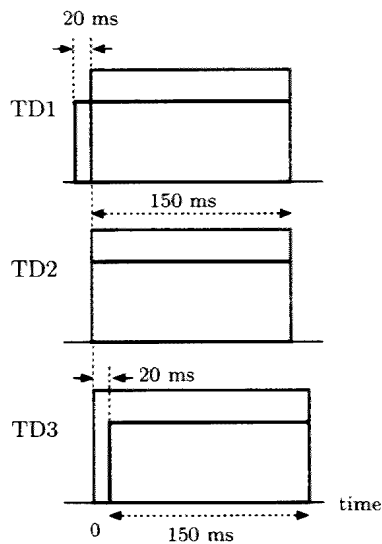


図1 各条件 TD1, TD2, TD3 における強刺激と弱刺激の輝度の時間変化。黒の実線は弱刺激、灰色の実線は強刺激を表す。なお、本タスクでは RT の基準値を強刺激のオンセットと定義する。

ルで自然に説明する必要がある。そのためには、上記モデルにどのような点が必要かを検証し、必要な修正をするべきである。

今回、以上の背景を受け、Iwaizumi et al. [1] が用いた 2 刺激同時呈示の実験環境と同じ条件で、2 刺激のオンセットに微小な時間差 (Stimulus Onset Asynchrony, SOA) を設け、反応特性を調べ、同時呈示・非同時呈示のいずれの輝度弁別特性も説明する神経回路モデルについて検討する。そして最終的に、様々な刺激時間構造を持つ視覚刺激強度弁別を可能にする神経回路構成および処理メカニズムについて、神経回路モデルの立場から今後の生理学的・人間工学的研究の指針となるようなモデルを提案することを本研究の目的とする。

## 2. オンセット時間差を刺激要因とした 2 刺激輝度弁別実験

### 2.1 実験条件と方法

本研究では、輝度の異なる 2 刺激のオンセットの時間差 (SOA) を刺激要因とし、3 条件を設定した。具体的には、弱刺激が強刺激よりも先行する TD1 条件 (SOA=-20ms)、弱刺激と強刺激が同時に呈示される TD2 条件 (SOA=0ms)、強刺激が弱刺激よりも先行する TD3 条件 (SOA=20ms) である。なお、設定した SOA の大きさは検知閾値以下であり、試行錯誤的に求められた。被験者には刺激に SOA が含まれることは知らせずに実験を行い、「2 刺激のうち明るい方をできるだけ早く正確に弁別」するよう教示を与えた。試行中に SOA に気づく被験者はいなかった。オンセットと異なり、オフセットは強刺激と弱刺激で同時刻とした。刺激時間長は SOA を除いて 150ms としたので、SOA が 0 で無い場合、先行する刺激は呈示時間長が 180ms、遅れて呈示される刺激は 150ms の呈示時間長になる。反応時間は強刺激のオンセットを基準とした。

当該実験では、2 刺激の輝度差は 1 条件のみとした。強刺激

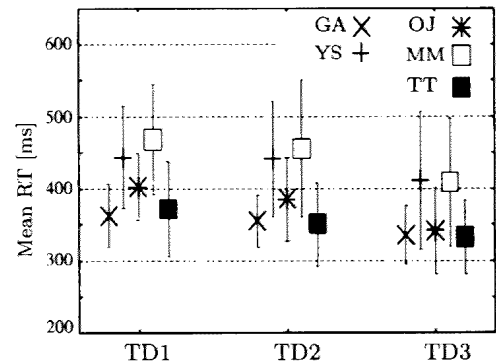


図2 呈示時間差刺激に対する各被験者の正解反応時の RT。縦軸は平均 RT を表す。エラーバーは各時間差条件 (横軸左から TD1, TD2, TD3) の被験者毎の RT の標準偏差 (SD) を表す。

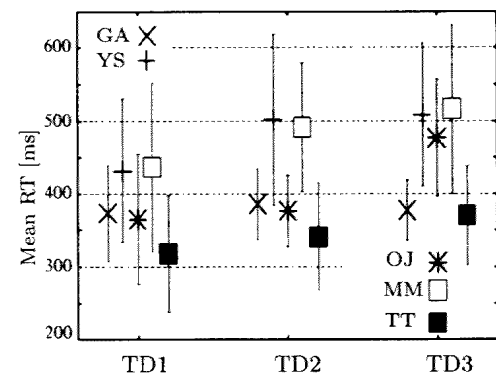


図3 呈示時間差刺激に対する各被験者のエラー反応時の RT。縦軸は平均 RT を表す。エラーバーは各時間差条件 (横軸左から TD1, TD2, TD3) の被験者毎の RT の標準偏差 (SD) を表す。

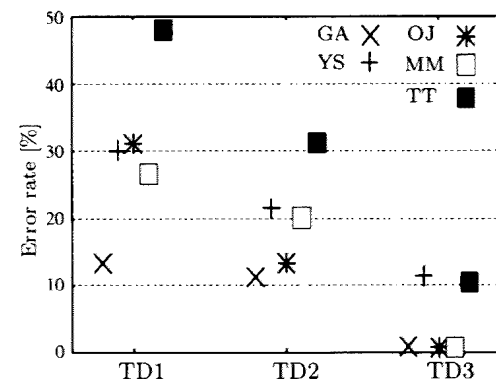


図4 呈示時間差刺激に対する各被験者のエラー率。縦軸はエラー率を、横軸は左から TD1, TD2, TD3 条件を表す。

の輝度は  $9.5 \text{ cd/m}^2$ 、弱刺激の輝度は  $8.5 \text{ cd/m}^2$  とした。2 刺激の輝度差は、10% 程度のエラー反応が生じるよう試行錯誤により決定した。

一回の試行に要する時間は 6 秒である。連続 90 試行で 1 セットとし、1 セット中に 1 呈示時間差条件当たり 30 試行が含まれる。適度に休憩を入れながら被験者ごとに合計 8 セット前後行った。以後、ある 1 条件のデータは各被験者当たり合計約 240 試行分から得られたものである。

### 2.2 実験結果

まず、全被験者の各 SOA 条件における正解反応およびエラー

反応時の RT を図 2 と図 3 に示す。正解反応時の平均 RT に関して、SOA を要因とした 1 要因の分散分析を行なった所、主効果が認められた ( $F = 32.4, **p < .01$ )。平均 RT に関する違いを明確にするため、Tukey 法により多重比較を行なった [4]。その結果、TD1 条件と TD3 条件の間、TD2 条件と TD3 条件の間に水準 5% で有意差が見られた。また、エラー反応時の平均 RT についても同様の分析を行なった所、SOA 要因に関する主効果が認められた ( $F = 7.9, *p < .05$ )。分散分析の結果、TD1 条件と TD3 条件の間に有意水準 5% で有意差が認められた。

エラー率の 2 刺激の SOA に対する変化を図 4 に示す。弱刺激が強刺激より先行することで明らかにエラー率が上昇している。前節と同様の分散分析を行なった所、呈示時間差に対する主効果が認められた ( $F = 28.3, **p < .01$ )。さらに、Tukey 法により多重比較を行なった結果、全ての呈示時間差条件間で有意水準 5% 以下の有意差が見られた。

以上の結果から、脳には 2 刺激の微小な SOA が輝度弁別処理に影響するような機構が存在する事が強く示唆される。ただし、SOA は検知閾値以下であることから、このような SOA 依存性がトップダウン的な修飾により引き起こされたとは考え難い。具体的な神経機構としては、弱刺激が先行する場合 (TD1 条件)、強刺激が先行する場合 (TD3 条件) よりエラー反応時の RT が小さくなり、かつエラー率が増加したのは、弱刺激が先行している時間帯に脳内でレベル付けされた弱刺激の強度が TD3 条件時よりも高まることで、刺激開始後直ちに脳内蓄積機構が閾値を越える割合が増えたためと考えられる。さらに、TD1 条件と TD2 条件を比較した場合、強刺激の呈示時間長は同じであるにも関わらず、TD1 条件の正解反応の平均 RT が TD2 条件のそれより遅れ、エラーが増えたのは、TD1 条件において弱刺激情報を蓄積する脳内機構から強刺激情報を蓄積する脳内機構の方に一時的に強い抑制が存在していると考えられる。

### 3. 神経回路モデル

Iwaizumi et al. は、2 刺激同時呈示の輝度弁別 RT タスクにおける反応時間やエラー率を再現する神経回路モデルを提案した [1]。このモデル (今後モデル A と定義する) は、図 5 に示すように、初期視覚経路の輝度に対する transient 応答と sustained 応答 [3] において強刺激か弱刺激のいずれかに対応する応答を感覚入力として受け取る 2 つの神経集団 (集団 1 と 2) と、集団の活動状態を比較するために 2 集団のうち一方のみから興奮入力を受けるニューロン (Y) から成る。各集団は、興奮性ニューロンと抑制性介在ニューロンから構成され、集団内の興奮性ニューロンによる再帰的興奮性シナプス結合により、各集団は感覚情報を蓄積し、介在ニューロンを介した側抑制結合により他方の集団の活動を抑制する。なお、輝度刺激を表象する感覚入力は、Poisson 過程に従うスパイク時系列に変換し、各集団の興奮ニューロンに対する興奮性シナプス入力として各集団に与えた。モデルでは、Y ニューロンが一方よりも先に発火する場合、その Y ニューロンが結合する集団が受け取る刺激

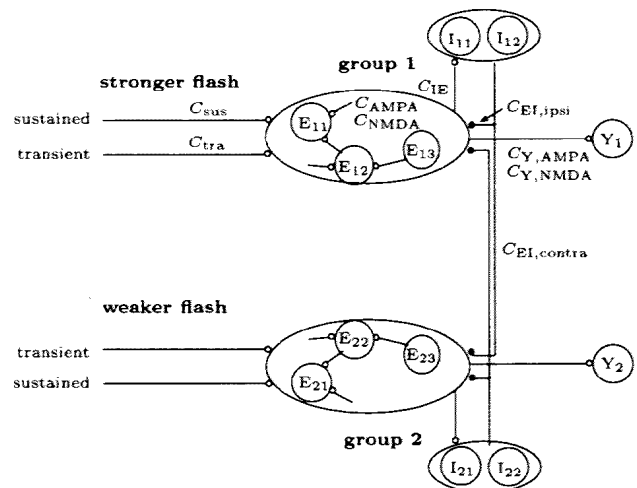


図 5 輝度弁別神経回路モデル (モデル A) の構成

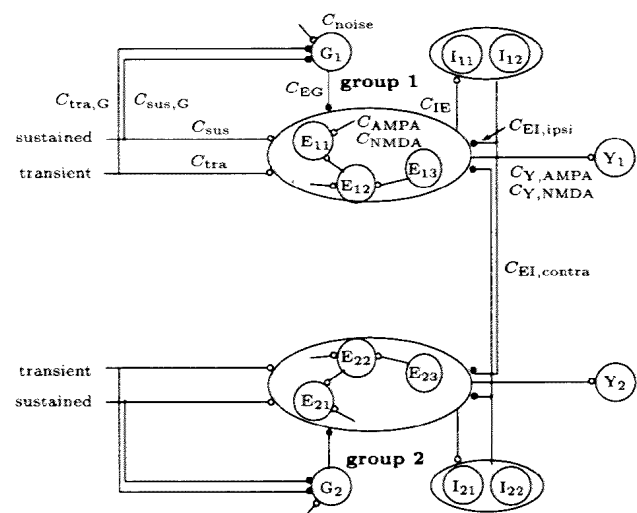


図 6 モデル B の構成

に対応するものを「より明るい」と弁別するよう定義した。

今回、様々な刺激条件下で自然な動作をし、かつ刺激時間構造などの刺激要因に対する反応特性を再現可能なモデルを構築するためにモデル A に必要な修正について検討した。

#### 3.1 グローバル抑制の追加

モデル A の問題点としてまず挙げられるのは、無刺激の状態では背景輝度 ( $3\text{cd/m}^2$ ) に対応する sustained 入力が平均  $17\text{Hz}$  の持続的でランダムな興奮入力として集団に与えられるが、sustained 入力の比較的強い興奮性シナプス結合強度が原因で、無刺激状態でも  $400\text{ms}$  程度の時間で入力を蓄積して Y ニューロンが発火する閾値まで達してしまう割合が高い (約 34%) という問題である。つまり、刺激が呈示されていないのに勝手に 2 刺激の明るさの判断をしてしまうという非現実的な動作をすることになる。このようなモデルの動作の原因として、刺激が無い時には弁別回路の段階において感覚情報の蓄積を行わないという機構が抜け落ちているためと考えられる。

そこで、モデル A に、図 6 で示されるようなグローバル抑制ニューロン (以下 G ニューロンと呼ぶ) を導入する。2 つの G

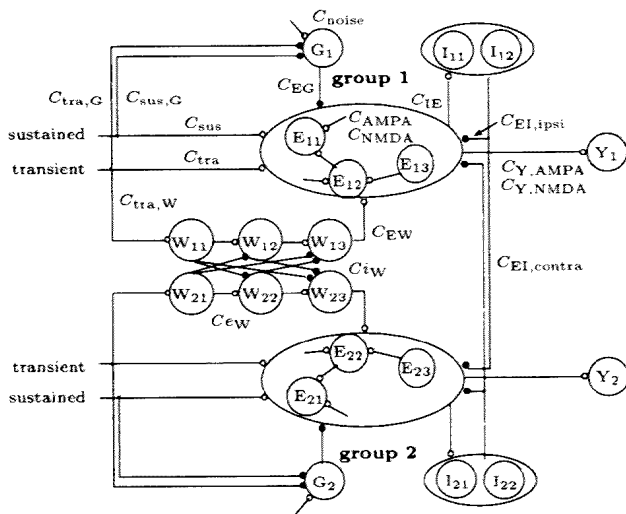


図7 モデルCの構成

ニューロンは常に平均 60 spikes/s のランダムな興奮性入力を受け、それぞれが各集団の全ての E ニューロンを抑制シナプス結合 (強度は  $C_{EG}$ ) を介して抑制する。一方、G ニューロンは transient 経路からの強い抑制性シナプス結合 ( $C_{tra,G}$ ) を持つので、一旦刺激がオンになると、E ニューロンは G ニューロンによる抑制から解放されることにより集団内の刺激情報の蓄積が可能になる。sustained 経路から G ニューロンへの入力強度 ( $C_{sus,G}$ ) は、背景輝度のような、持続的な弱い輝度入力時に神経集団内の発火活動が伝搬するのを抑制するために、transient 入力強度よりも大幅に小さく設定した。今後、このようなグローバル抑制ニューロンを導入したモデルを 'モデル B' を呼ぶ。この修正により、無刺激条件において誤って集団の勝者が決まってしまう割合を 1.5% まで低減した。

### 3.2 フィードフォワード抑制付きニューロンチェーンの追加

次に、第2節の実験から、検知閾値以下の微小な SOA 刺激に対して反応が有意に変化したことから、輝度弁別において集団は SOA 情報を何らかの形で受け取っているはずである。実験で、例えばエラー反応時に TD1 条件では TD3 条件と比べて被験者の反応が早くなったことから、TD1 条件では一時的に弱輝度入力先行に伴う強い興奮入力があったと考えられる。モデルには集団同士の側抑制結合が存在するので、同様の機構が正解反応時における TD1 条件での反応時間の増大の原因として説明できるだろう。以上をまとめると、2 刺激が同時に存在しない時だけ、その情報を集団に興奮入力として強く伝える神経機構が存在すると考えられる。

集団に対する transient 入力はその候補となり得るが、設定した transient 入力シナプス強度は非常に弱く、20ms の SOA により実験で得られたオーダーで集団の活動をより早く閾値まで興奮させる効果は期待できない。また、transient 入力強度を強くするという方法も考えられるが、その場合、Iwaizumi et al. の実験から得られた反応時間やエラー率の刺激呈示時間長に対する依存性が失われてしまう。

そこで、別の方法として、transient 入力から興奮入力を受け、いずれかの集団の興奮ニューロンに興奮入力するフィード

フォワード抑制付きのニューロンチェーンを提案した。興奮性シナプス結合をする 2 本のニューロンチェーンが相互にフィードフォワード抑制性シナプス結合を持つ構造により、2 本のチェーンは 2 刺激の非同時性を検出する機能を持つ。このチェーンをモデル B に導入したものをモデル C と定義し、神経回路構造を図 7 に示す。

第2節の実験について、提案した 3 モデル A, B, C 間の出力を比較する。まず図 8 に各モデルが出力した正解反応における平均 RT・RT の標準偏差 (SD) の特性を示す。ただし、モデルの平均 RT 値は運動プログラムにかかると推定される時間 (200ms) を加算した値となっている。モデル C の出力は、実験と同様に弱刺激が先行すると正解反応時の RT が大きくなる SOA に依存した反応時間変化を示すが、モデル A と B では TD1, TD2, TD3 の 3 条件に大きな違いは見られなかった。よって、モデル C に導入した 2 刺激の非同時性に関する情報を集団に入力する機能を持つチェーンの効果が大きいことが分かる。次にエラー反応における反応時間について比較する。実験では、エラー反応時の平均 RT 特性は正解反応とは全く逆になった。図 9 に示されるように、モデル A, B の出力はその変化率が実験よりも多少小さく、また、モデル C では変化率が実験値よりも大きいものの、全てのモデルが SOA に対する反応の変化傾向を再現しているのが認められる。

エラー率の比較を図 10 に示す。2 刺激の非同時性を検出するチェーンが存在しないモデル A では条件間の変化率は非常に小さいものの強刺激が先行するほどエラー率は低下するが、エラー率のばらつきが小さくなっていく傾向を全く示さなかった。また、モデル B ではエラー率に SOA 依存性は見られなかった。一方、モデル C では (変化率は大きいものの) 実験と同様の SOA 依存性を明瞭に再現できた。

以上のシミュレーションから、第2節で行なった 2 刺激のオンセットに時間差がある輝度弁別実験に関して、提案した 3 モデルの内、SOA を検出するフィードフォワード抑制付きチェーンを導入したモデル C のみが実験で得られた全ての刺激呈示時間差依存性 (正解およびエラー反応時の平均 RT, エラー率) を再現できることが示された。

モデル C は、上記の 2 刺激弁別タスクである 2 実験における反応特性だけでなく、注視点を挟んだ左右の刺激用 LED の内、確率的に一方のみが検知閾値以上の輝度で呈示される、いわゆる単一刺激検出タスクにおける反応時間をも、広い刺激輝度レベルに渡って再現することに成功した。図 11 に実験値と 3 モデルの出力を示す。

しかし、Eagleman et al. (2005) が行った、2 刺激に大きな呈示時間長 (長刺激: 768ms, 短刺激: 54ms) およびオンセットあるいはオフセットの呈示時間差を持った輝度弁別タスク [2] の場合に本研究で提案したモデルは適用できない事が示唆された。彼らの実験では同輝度でも短刺激が長刺激のオフセット時に呈示されると短刺激の方が明るく知覚されるというものが、本研究のモデルは感覚情報を蓄積する神経集団の高い活動状態の持続を抑制する効果が弱いために、先行する長刺激が常に明るく知覚される結果になってしまった。Eagleman et al. が用

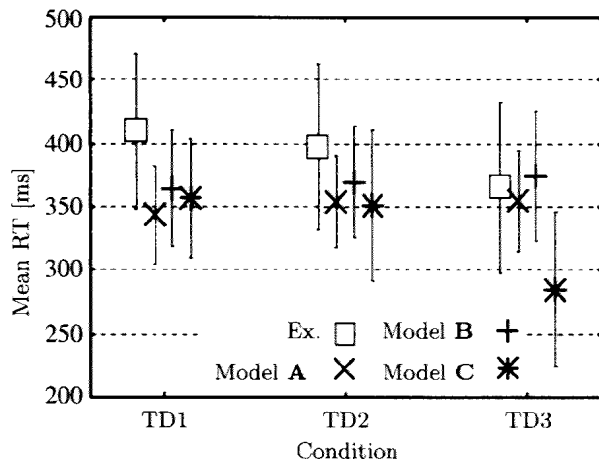


図8 平均 RT(正解反応時)と3モデルの出力: モデルの平均 RTは運動処理推定時間(200ms)を加算して表示. エラーバーは RTのSDを表す.

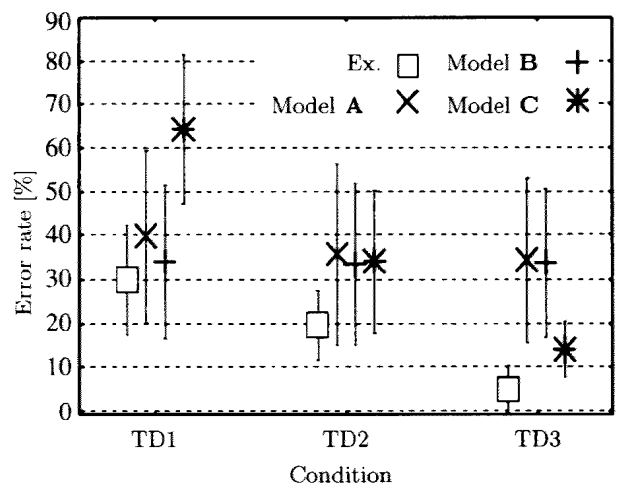


図10 エラー率と3モデルの出力

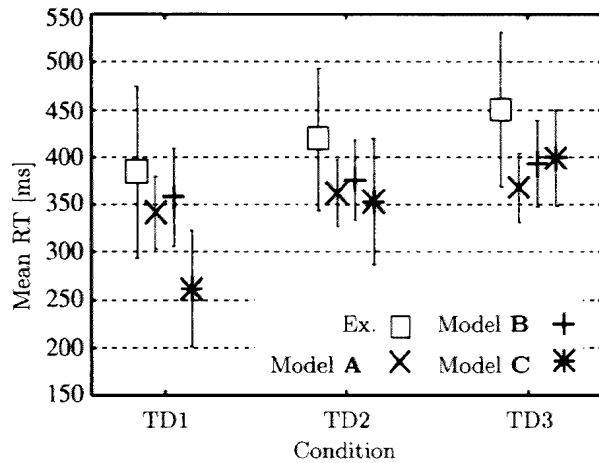


図9 平均 RT(エラー反応時)と3モデルの出力: モデルの平均 RTは運動処理推定時間(200ms)を加算して表示. エラーバーは RTのSDを表す.

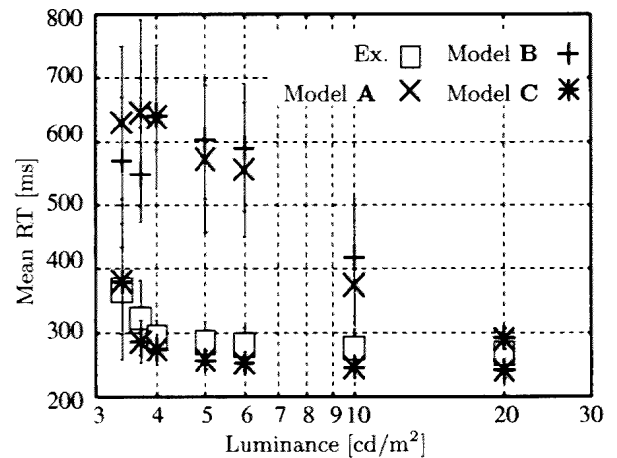


図11 単一閾上刺激検出実験の平均 RTと3モデルの出力. エラーバーは RTの標準偏差を表す.

いたようなより大きな刺激呈示時間差の輝度弁別処理においては短期記憶の関与が大きいと考えられ, その神経機構の表現はより複雑になると思われる.

#### 4. 結論

本研究では, より多様な刺激条件下においても実験結果と対応する自然な動作をするモデルにするためにモデル A に対する必要な修正として輝度情報の不要な蓄積を禁止する機構や2刺激の非同時性検出機構の追加を提案した. グローバル抑制ニューロンとフィードフォワード抑制付ニューロンチェーンを導入したモデル C は, 無刺激状態や同時呈示あるいは微小な SOA を持った2刺激弁別タスクおよび閾上単一刺激検出タスクにおけるの実験結果をも再現することに成功した. モデル C は, 検知閾値以下のある程度小さな呈示時間差の範囲においてのみ有効であるものの, 脳の視覚刺激弁別処理機構を統一的に理解するモデルとして, モデル A や B と比べてより妥当性が高いと思われる. 今後, 本研究で提案したモデルについて生理学的・心理実験的検証が成されることが期待される. また,

刺激時間構造に対する実験・モデルを使ったさらなる検討により, より多様な刺激時間構造の刺激強度弁別に関する脳の情報処理機構の解明が前進することが期待される.

#### 文献

- [1] Iwaizumi A et al. Characteristics of human luminance discrimination and modeling a neural network based on the response properties of the visual cortex. Biol Cybern DOI:10.1007/s00422-006-0058-8 (in printing)
- [2] Eagleman DM, Jacobson JE and Sejnowski TJ (2004) Perceived luminance depends on temporal context. Nature 428:854-856
- [3] Kinoshita M and Komatsu H (2001) Neural representation of the luminance and brightness of a uniform surface in the macaque primary visual cortex. J Neurophysiol 86:2559-2570
- [4] 森 敏昭, 吉田 寿夫 編著 (1990) "心理学のためのデータ解析テクニックブック", 北大路書房

# 卒業論文

## 短音と長音で構成される時系列の 短期記憶に関する研究

東北大学 工学部 電子工学科

坂井靖文

2005年3月24日

指導教員： 二見亮弘（東北大学大学院工学研究科）



# 目次

第1章 序論	1
第2章 知見	3
2.1 記憶システム	3
2.2 二重貯蔵モデルにおける短期記憶	3
2.2.1 短期記憶と長期記憶の区分	4
2.2.2 短期記憶の特性	5
2.2.3 二重貯蔵モデルにおける短期記憶概念の問題点	6
2.3 作動記憶	6
2.4 記憶測定の方法	7
2.4.1 再生	7
2.4.2 再認	7
第3章 短音と長音からなる信号による再生実験	9
3.1 目的	9
3.2 実験方法	9
3.3 実験装置	9
3.4 実験	10
3.4.1 7音から構成される信号による再生実験	11
3.4.2 多数の音から構成される信号による再生実験	13
3.4.3 7音から構成される信号による再生実験(複数被験者)	15
3.5 考察	16
第4章 短音と長音からなる信号による再認実験	17
4.1 はじめに	17
4.2 目的	17
4.3 実験方法	17
4.4 実験	18
4.4.1 熟練被験者による再認実験	18
4.4.2 非熟練被験者による再認実験	22
4.5 系列位置曲線の比較	24
第5章 結論	25
謝辞	27
参考文献	28

# 第1章

## 序論

脳は多くの画期的な発見が行われる可能性を秘めている研究対象であり、21世紀に残された数少ない巨大フロンティアのひとつである。また、脳科学の進歩は、人間たる所以の根元である脳を知ることにつながり、脳を知るとは即ち人間を理解することにつながる。また、脳科学研究の成果は、脳の老化の防止、アルツハイマー病等脳・神経系の困難な病気の克服、脳の原理を生かしたコンピュータやロボットの開発による新技術・新産業の創出につながる。このような意味で脳科学の推進を図り、脳機能の解明を行うことは、正に人類課題となってきた<sup>[1]</sup>。また、脳における時系列パターン処理、特に音声処理の原理や機構を明らかにすることは、言語を介した人間の知的情報処理への一つのアプローチであり、非常に重要である<sup>[2]</sup>。

時系列パターンの処理(認識・発生)に関する神経回路モデルでは、相互結合型モデルに非対称結合を持たせて状態遷移を起こす機構<sup>[3]</sup>がよく用いられる。しかし、状態遷移型モデルだけで脳における種々の“時間”の取り扱いをうまく説明できるわけではない。特に、聴覚系における比較的短い時間オーダーの時系列パターンの知覚や短期記憶の機能を説明できる処理原理・機構モデルに関しては未解決の問題が残されている。また、短期記憶に限らず、脳機能研究において、聴覚の研究もあまり多くないのが現状である<sup>[4]</sup>。

ここで、モールス信号について考えてみる。国際的規約によるモールス信号は図1に示すような符号の列であるが、音声に見られるような時間伸縮の影響を受けにくい特徴は何もない。よって、モールス信号のような短音と長音のみで構成される信号を取り上げることによって聴覚系における情報処理の原理・機構について考えるとき“時間”に関するより本質的な問題に注意を集中しやすくなる可能性があると考えられる。また呈示する刺激の種類は短音と長音の2種類しかないので、刺激の種類は2種類なのだが、信号構成音がスロット数を上回ることに伴って、短期記憶機能を調べるにあたって新たな視点から探れるかもしれない。

そこで、本研究ではモールス信号を元として、最も重要な要素音である短音と長音を用いた実験により、聴覚系の時系列における短期記憶に関する考察を行う。

A	- - -	N	- - -
B	- - - -	O	- - - - -
C	- - - - -	P	- - - - -
D	- - - -	Q	- - - - -
E	-	R	- - - -
F	- - - - -	S	- - - -
G	- - - - -	T	- - -
H	- - - -	U	- - - -
I	- -	V	- - - - -
J	- - - - -	W	- - - - -
K	- - - - -	X	- - - - -
L	- - - - -	Y	- - - - -
M	- - - -	Z	- - - -

図1. 欧文モールス符号

本論文の構成は以下の通りである。

第1章は序論である。

第2章では本研究の背景となる知見を述べる。

第3章では短音と長音より構成される信号による再生実験について述べる。

第4章では短音と長音より構成される信号による再認実験について述べる。

第5章は結論である。

## 第2章

# 知見

### 2. 1 記憶システム

人間の記憶はただ一つの過程からなるものではない。記憶の働きには少なくとも三つの異なる相があることがわかっている<sup>[5]</sup>。第一の相は、主としてパターン認知のメカニズムを含めて、知覚処理が適切に働く上で重要なものである。特定の感覚器官に到達した感覚情報の詳細なイメージを（数分の一秒間）保持しつづける記憶システムで、感覚情報貯蔵（Sensory Information Storage : SIS）と呼ばれている。第二の相は、数秒間、場合によっては数分間情報を保持しつづける短期記憶（Short-Term Memory : STM）システムである。しかし、STM と SIS は同じものでない。STM において情報はすでにパターン認知機構によってコード化され、カテゴリー化されているからである。STM は必要な情報を数分間保持し続けたり、情報を体制化して、永久に貯蔵しようとする段階でもある。記憶の第3の相は、長期記憶（Long-Term Memory）システムである。ここでは、私たちの経験の記録がほぼ永久に保存される。この記憶層の容量には本来限りがないと考えられている。

### 2. 2 二重貯蔵モデルにおける短期記憶<sup>[6]</sup>

過去の出来事や知識のデータベースをなす記憶と異なる、限られた量の情報を短期間保持しておくための記憶システムの存在は、かなり前から想定されていた（James, 1890）。この直観的にうなずける記憶システムの区分を情報処理の観点から本格的に理論化したのが、二重貯蔵モデルである。その集大成である、アトキンソンとシフリン（Atkinson & Shiffrin, 1968）のモデルの機略を図2に示す。

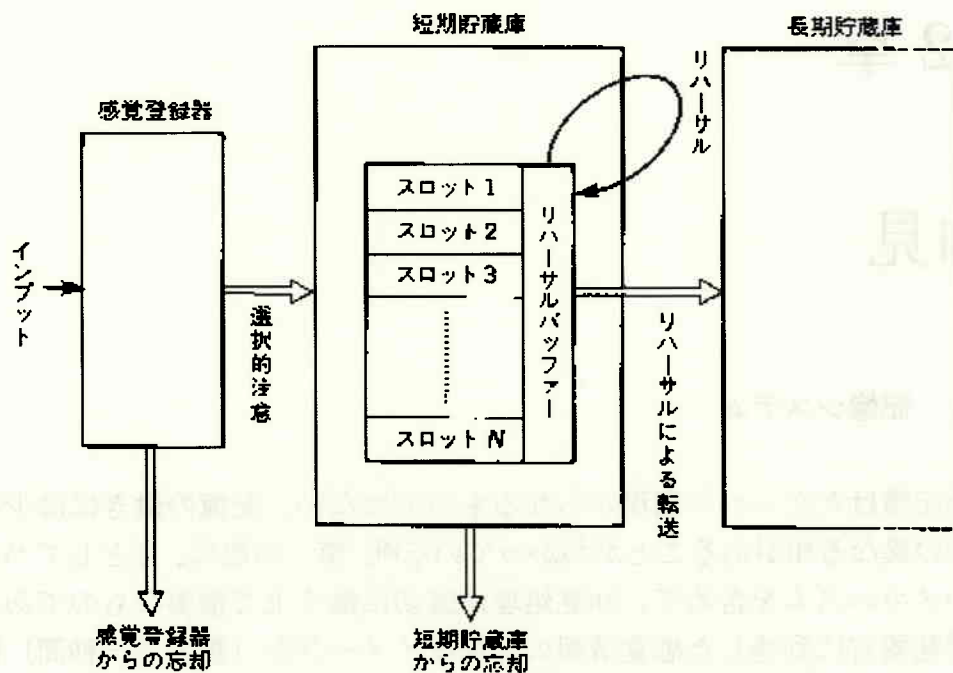


図2. アトキンソンとシフリンの二重貯蔵モデルの概略(Atkinson & Shiffrin, 1968)

このモデルによると、外界からの情報は、まず感覚登録器に入り、選択的注意によって選び出されたものだけが、短期貯蔵庫に送られ、そこで一定期間保持される。しかし、短期貯蔵庫の容量は限られていて、新しい情報がつぎつぎと入ってくると、古い情報は押し出され忘却してしまう。そこで、情報を忘れないように頭の中で繰り返し唱えるように情報をリフレッシュさせるリハーサルと呼ばれる活動が必要になる。短期貯蔵庫には、このために一定数のスロットを持つリハーサルバッファが想定されていて、スロットの数と同じか、それ以下の量の情報を繰り返し循環させることで情報の保持が可能になるという。

この情報保持の機能に加えて、リハーサルには、情報を次の貯蔵構造である長期貯蔵庫に転送する機能もあるとされる。短期貯蔵庫内に滞在している時間が長いほど、その情報が長期貯蔵庫に転送される確率が高いと仮定されているため、リハーサルの回数が多いほど、その情報が長期記憶として定着する可能性も高いことになる。

### 2. 2. 1 短期記憶と長期記憶の区分

被験者にある項目を提示してそれを順序かまわず再生させたとき（自由再生）、リスト内の項目数に関わらず、最初と最後の数項目の再生率が、リストの中間

部の再生率と比べて高くなる。これを系列位置効果(serial position effect)という。

二重貯蔵モデルによると、系列の最初の部分の再生率が高いという初頭効果(primacy effect)が見られるのは、最初の数項目が、まだ空の短期貯蔵庫に入るため、必然的にリハーサルを多く受け長期記憶として定着する可能性が高いからとされる。これに対して、最後の数項目の再生率が高いという新近性効果(recency effect)は、リスト提示直後、それらの項目はまだ短期貯蔵庫の中にあつて、そこから直接読み出せるからとされる。実際、数字の逆唱などの妨害課題を再生前に課し、再生を遅延させると図3に示されているように、初頭効果には変化がないが新近性効果は消失する。このように短期記憶を長期記憶とは別の散在と仮定することで、系列位置効果がうまく説明できる。

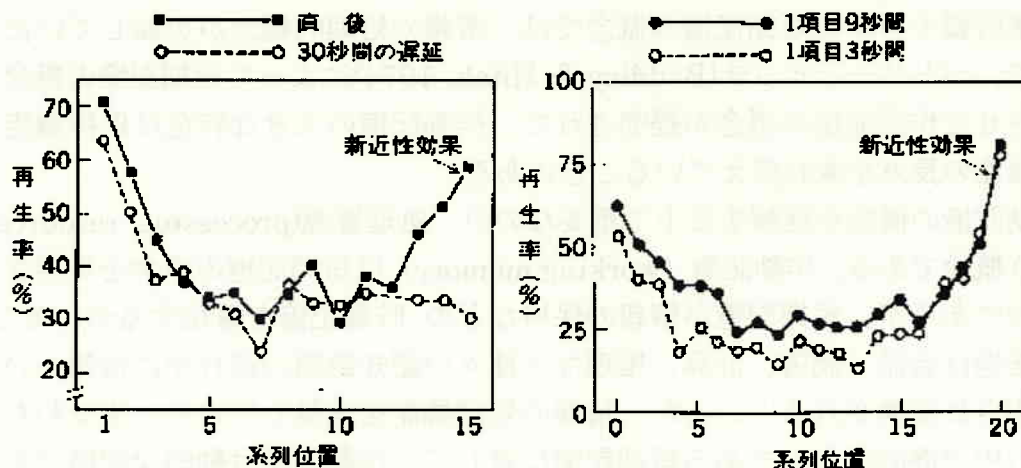


図3. 遅延時間と単語の呈示測度が系列位置曲線に及ぼす効果(Glanzer & Cunitz, 1966)

左の図は遅延時間の効果、右の図は呈示測度の効果を示している。

## 2. 2. 2 短期記憶の特性

短期記憶の容量を見積もるにあたり、情報理論の観点からチャンクという概念を導入したのがミラー(Miller, 1956)である<sup>[7]</sup>。チャンクというのは何らかのまとまりを持つ情報のかたまりのことで、ミラーによれば人間が一度に保持できる情報の量はおよそ $7 \pm 2$ チャンクに限られているという。短期貯蔵庫のリハーサルバッファの観点から言えば、1スロットに1チャンクの情報が入るとして、約 $7 \pm 2$ 個のスロットがあるということになる。

### 2. 2. 3 二重貯蔵モデルにおける短期記憶概念の問題点

以上述べたように、二重貯蔵モデルの理論的枠組みの中では、短期記憶は限られた数のスロットからなるリハーサルバッファの中心的役割に象徴されているように、主に言語的情報を単調な繰り返しによって保持し、長期記憶に転送するための、単純かつ受動的なシステムとして捉えられていたと言える。しかし、二重貯蔵モデルの根本をなす、短期記憶と長期記憶の構造的区分の仮定を疑問視させるデータが数多く示された[8]~[10]。また二重貯蔵モデルの短期記憶概念における大きな問題点として、処理的観念が欠如していることが挙げられる。

### 2. 3 作動記憶

二重貯蔵モデルの短期記憶の概念では、情報の処理的観念が欠如していた。そこで、バドリーとヒッチ(Baddeley & Hitch, 1974)によって短期記憶の概念を発展させた作動記憶の概念が提唱された。作動記憶の大きな特色は保持機能と処理機能の長方を兼ね備えていることにある。

作動記憶の概念を理解する上で重要なのが、処理資源(processing resources)という概念である。作動記憶(working memory)は短期記憶の概念を発展させたものであるが、短期記憶が情報の保持などの貯蔵記憶を重視するのに対し、作動記憶は会話、読書、計算、推理など種々の認知課題の遂行中に情報がいかに操作され変換されるといった、情報の処理機能を重視している。すなわち、記憶の中で静的なものである短期記憶に対して、作動記憶は動的な記憶であると考えられている。

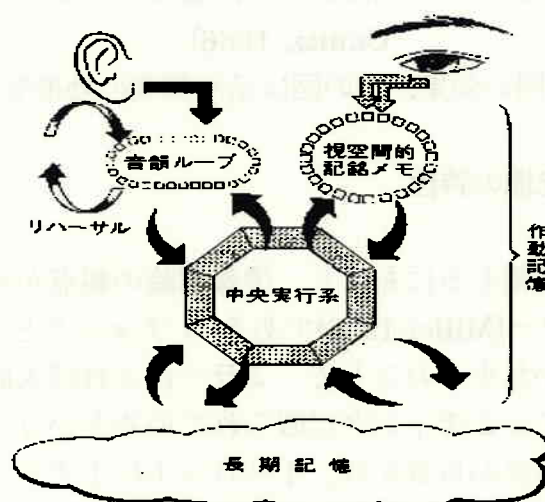


図4. バドリーの作動記憶のモデルの概略図(Baddeley,1986,1992)

Baddeley and Hitch (1990)の作動記憶システムのモデルは主に、下位システムの働きを管理し制御する中央実行系と、下位システムとして音韻情報の保持と処理を行う音韻ループと、視覚情報や空間情報の処理と保持を行う視空間スケッチパッドの3つから構成されているが、それらは入力された情報を操作するためのそれぞれ資源（処理資源）を有し、その資源には一定の容量限界があると考えられている。そして、その資源を利用して下位システムは情報の処理作業および保持作業を行うことができるとされている。しかし、処理および保持される情報量が下位システムの容量限界を超過する場合には中央実行系の資源もその作業に充てられ、他の下位システムの作業に支障が生じることになる。このような干渉が起こる現象をトレードオフと言う。またこの現象は、二重課題法のような2つの課題（例えば、記憶をしながら問題を解いたり、文章を読みながら単語を記憶する）を同時に行わせた時に生じる。

## 2. 4 記憶測定の方法<sup>[11]</sup>

一般に再生は提示された刺激を正確に再現する必要があるために、虚答率は低いが高正答率も低くなるために短期記憶の容量を測る点においては不満足な実験方法である。再認は刺激を呈示後、新たに呈示される項目が刺激と同様か指摘させるために、正答率は高くなるので短期記憶の容量を測るのに適した測度であるが、当て推量で答えても一定の正答率が得られる。このように再生と再認には互いに補完するような長所と短所がある。

### 2. 4. 1 再生

被験者に刺激項目を呈示して、被験者がそれを思い出して正確に再現する方法を再生(recall)と呼ぶ。再生には自由再生(free recall)・系列再生(serial recall)・手がかり再生(cued recall)等の種類がある。自由再生は、被験者が想起した順に自由に再生する方法である。系列再生は刺激を呈示された順序どおりに再生する方法である。手がかり再生は刺激項目に対する手がかりを与え、それをもとに再生する方法である。

### 2. 4. 2 再認

被験者に刺激を提示した後、刺激項目と同じ項目(ターゲット: target)と異なる項目(ディストラクタ: distracter)を呈示し、ターゲットを指摘させる方法を再認という。再認はターゲットとディストラクタの中からターゲットを指摘さ



せる方法なので、当て推量で答えても一定の正答率が得られる(チャンスレベルが存在する)。

表 1. 再生と再認の長所と短所

	長所	短所
再生	虚答率が低い	正答率が低い 運動系の影響が除去できない
再認	記憶の検索過程が不要	チャンスレベルが存在

## 第3章

# 短音と長音からなる信号による再生実験

### 3.1 目的

二重貯蔵モデルにおいて短期記憶の容量は、数字、アルファベットなどの記憶項目に関わらず  $7 \pm 2$  チャンクといわれる。そこで、二重貯蔵モデルにのっとり、直後記憶範囲の実験において、七つ程度の短い音(短音)と長い音(長音)からランダム構成されるパターン信号を被験者に聴取させ、信号を順序どおりに再生(系列再生)してもらふことにより、聴覚系における短期記憶機能を調べる。

### 3.2 実験方法

標準的なモールス信号に合わせて、信号を構成する各音の長さを短音:長音:文字内空白で  $1:3:1$  の比率とする。呈示信号の速さは、短音の長さにして  $100\text{ms}$  とする。短音と長音をいくつか組み合わせた信号を被験者に聴取させ、これを1試行として1セット100試行行う。

信号を聞かせ終わった直後被験者に回答入力装置(電鍵)を使って系列再生させる。この際、誤りの傾向を調べたいので聴取し終わったらなるべく早く入力してもらふよう、指示を前もって与えておく。

### 3.3 実験装置

実験系は二入力の低周波発振器( $1\text{kHz}$ )、スピーカ、直流電圧が出力可能なコンピュータ(動作環境:RT Linux)、電鍵により構成される。発振器は二つの入力に電圧が加わると発振し、入力の一つでも電圧が入力されないと発振しない機構である。スピーカから発生させる音の制御は、発振器の二入力の一つを直流電源( $6\text{V}$ )につなげ、もう一つの入力をコンピュータにつなぎ発振器に入

る電圧を制御することにより行う。電鍵はコンピュータにつながっており、  
が入力した回答をファイルに格納する。

カサ  
被験者

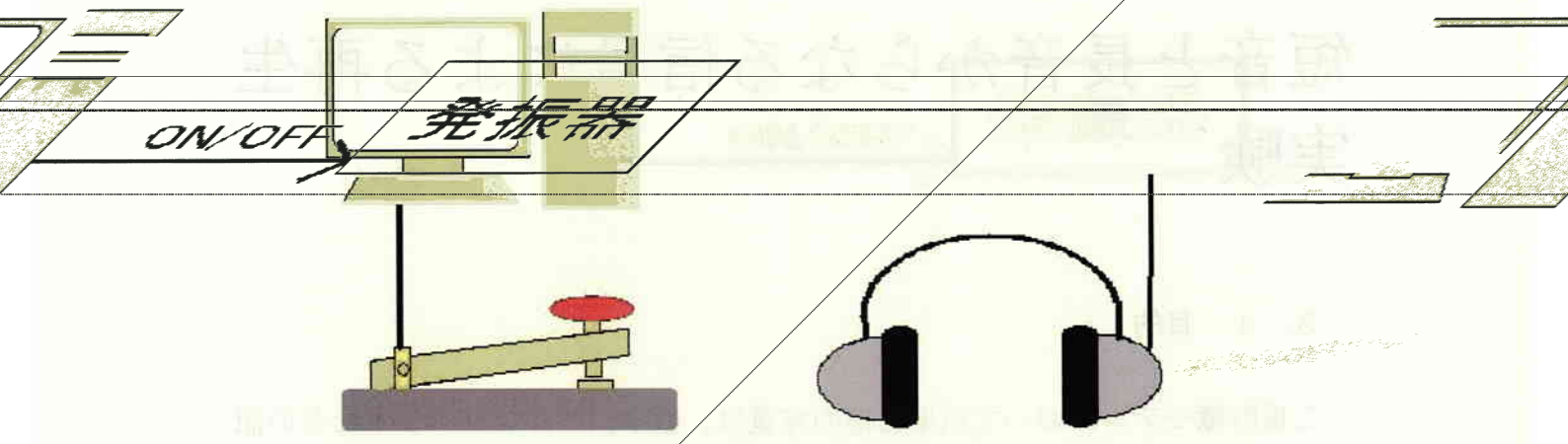


図5. 実験系

### 3.4 実験

長を調べるため  
の長さにして  
鍵を押し

聴取実験を行うにあたり被験者が短音と長音を区別する時間  
に、6人の被験者に短音と長音、合計7音より構成された、短音の  
100msの速さの信号を100の試行聴取・再生させ、再生した際の電  
た時間を計測した。以下がそのヒストグラムである。

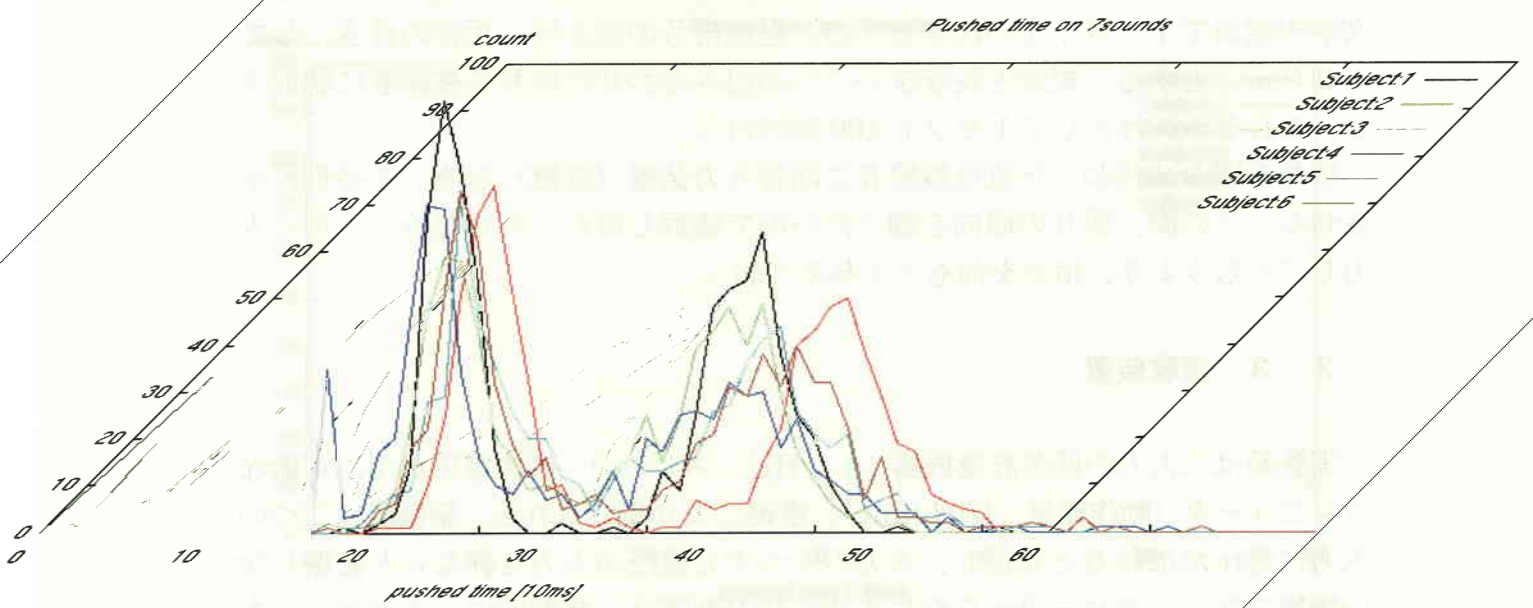


図6. 複数被験者による短音・長音の解答時間分布

図6の分布図において、大きな分布が二つあることがわかる。左の分布が短音を再生したときの分布であり、右の分布が長音を再生したときの分布であると考えられる。この分布図より、被験者は信号を電鍵で再生するとき短音と長音を明確に区別する事が可能であるといえる。

今回の再生実験では単音の長さにして100msの速さの信号のみを扱うので、被験者の回答を短点と長点を区別する際の基準点を二つの分布の谷となっている点(18ms)とした。

### 3. 4. 1 7音から構成される信号による再生実験

はじめに一人の被験者のみで再生実験を行った。

まず7音から構成された信号を再生させ、系列位置ごとの正解率を求めた。

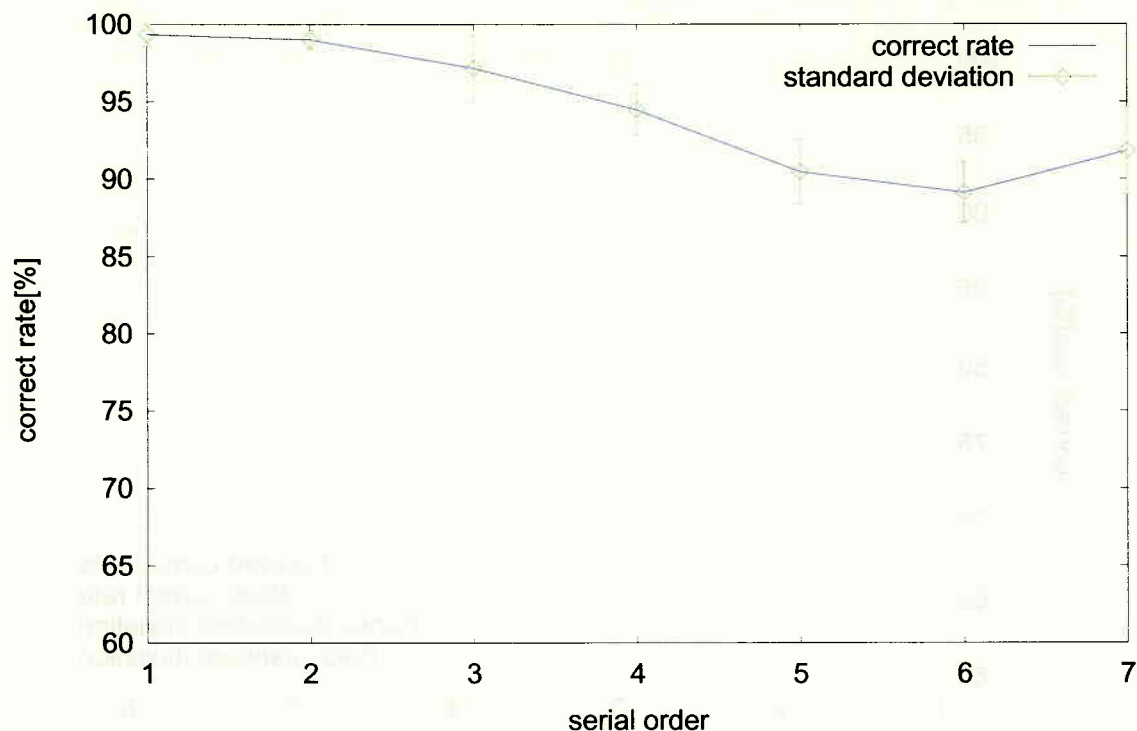


図7. 7音の正解率

正解率は前から順に低くなっていき、最後部で高くなった。

ところで、図7は回答の前から正解と照合させた結果である。次に正解の後ろから解答と照合させた。照合例は図8の通りである。

前から照合					後ろから照合								
1	2	3	4	5	1	2	3	4	5				
解答 :	短	長	短	短	長	解答 :	短	長	短	短	長		
回答 :	短	長	長	短	短	長	回答 :	短	長	長	短	短	長
結果 :	○	○	×	○	×	×	×	○	○	○	○	○	

図8. 照合例

後ろから照合させた理由は、上の照合例のように回答を後ろから照合させると正解と同じになる回答例があり、また聴取させた信号の初めの方は正答率が高いために初めの方は短期記憶ではなく次の段階の記憶に移行していて、後ろの方は短期記憶より再生されているのではないかと考えたためである。後ろから照合させた結果が図9である。

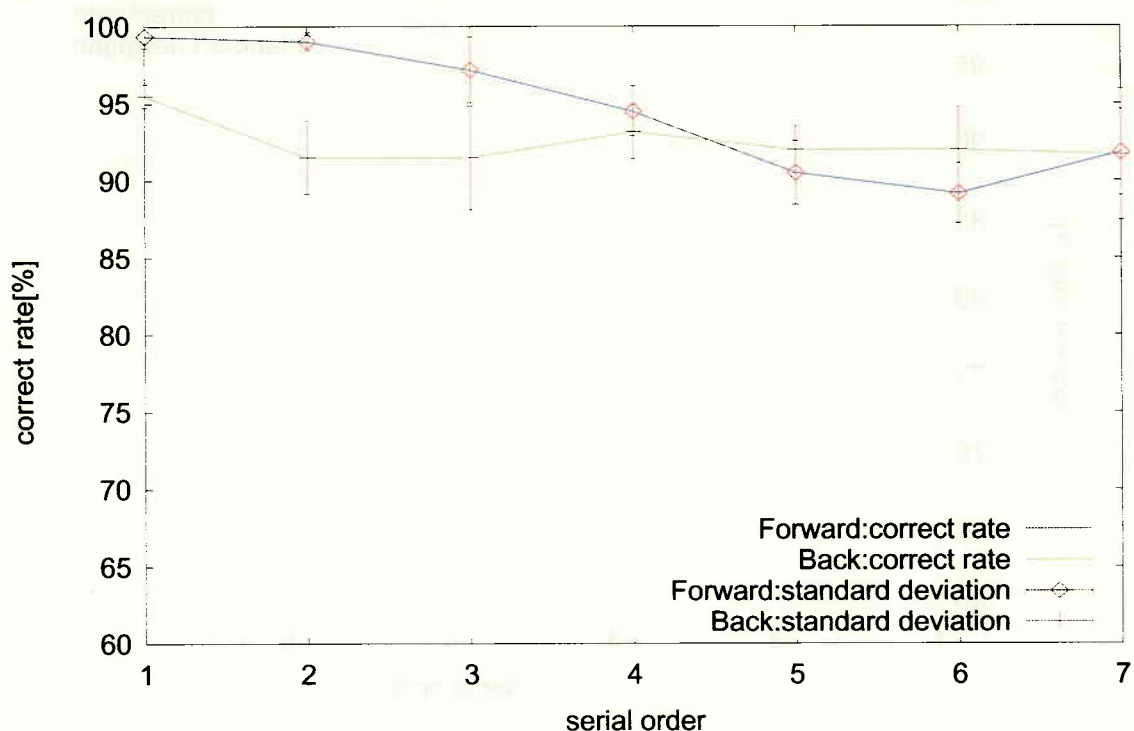


図9. 7音の正解率 (前と後ろから照合)

図9より、後半部の正解率は後ろから照合させたときの方が前から照合したときの正解率よりも高いことがわかる。この結果より、信号の前半部と後半部とでは信号を再生している記憶箇所が異なると推測される。

### 3. 4. 2 多数の音から構成される信号による再生実験

次に、信号の構成音数が変化した場合における正解率の変化を調べるために、6音・8音・9音から構成された信号の再生実験を行った。実験結果は以下のとおりである。

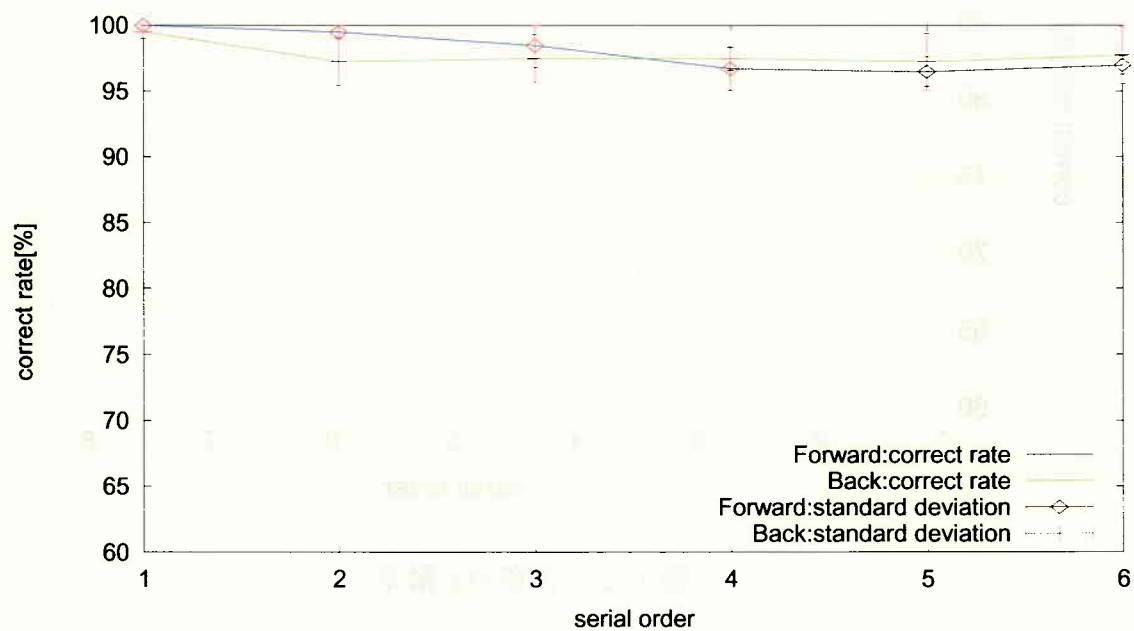


図10. 6音の正解率

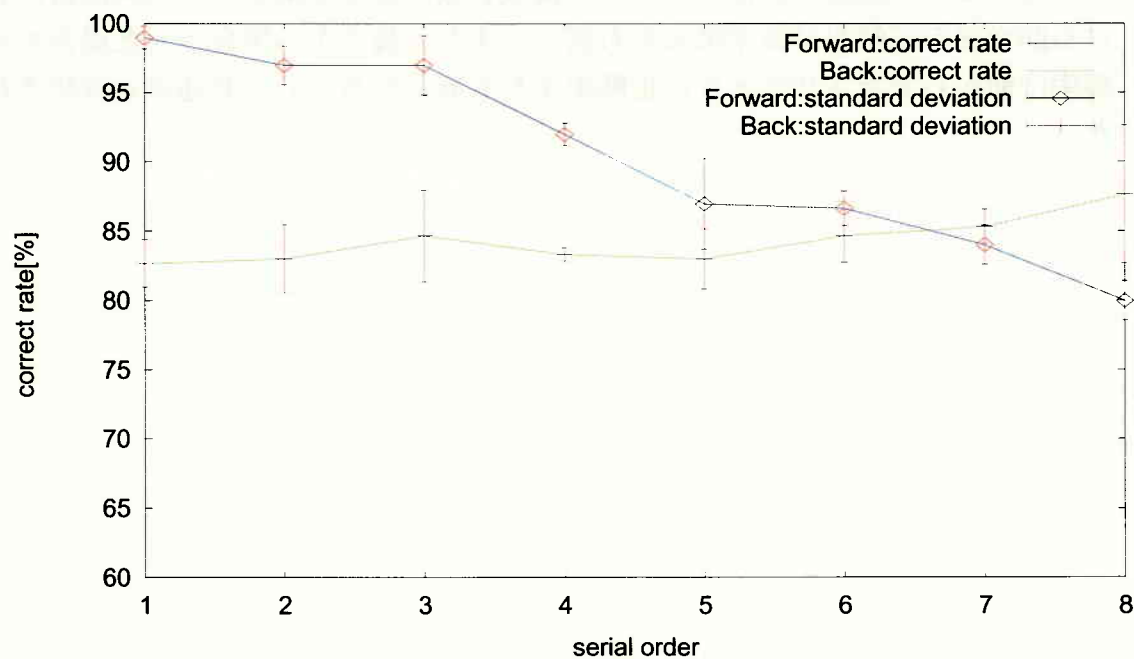


図11. 8音の正解率

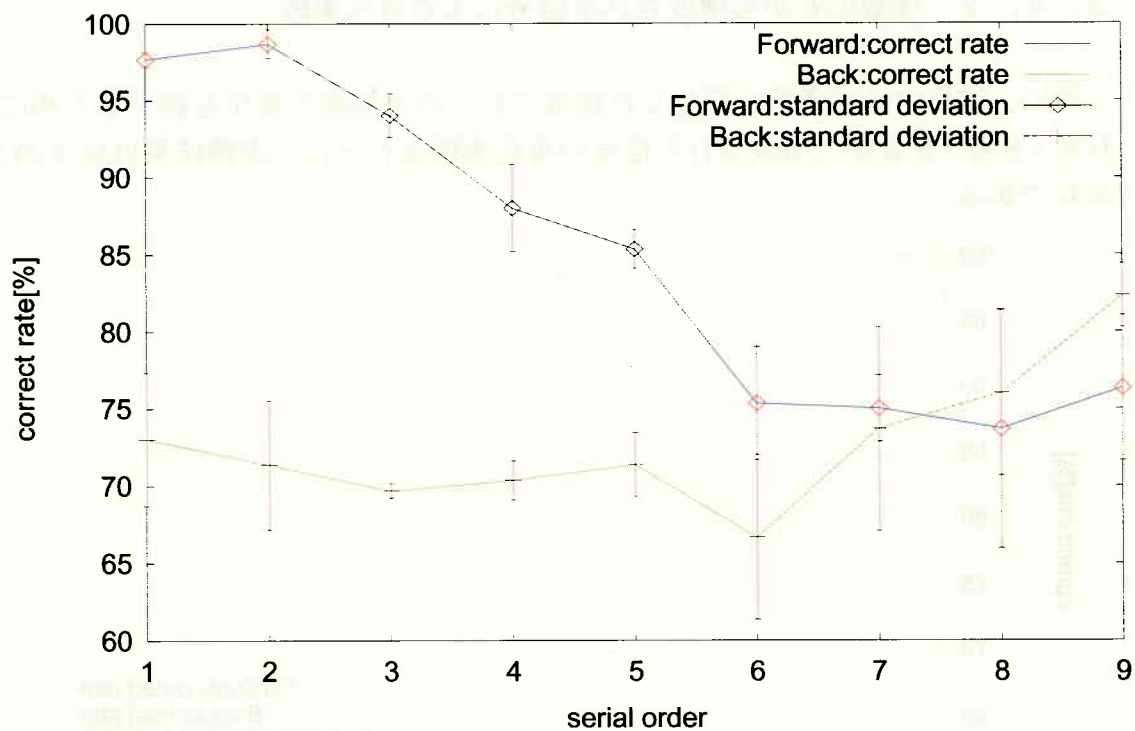


図12. 9音の正解率

6・7・8・9音の再生実験結果より、構成音数が少ない方が全体的な正解率が高いことが確認された。しかし、構成音数が異なる場合でも系列位置における前半部の正解率が後半部よりも高く、また、後ろから照合させた場合の正解率は前から照合させたときの正解率よりも高くなるという共通点も確認された。



### 3. 4. 3 7音から構成される信号の再生実験（複数被験者）

次に、再現性を確認するため5人の被験者に、短音と長音・合計7音より構成された、短音の長さにして100msの速さの信号の再生実験を行った。実験結果は以下のとおりである。

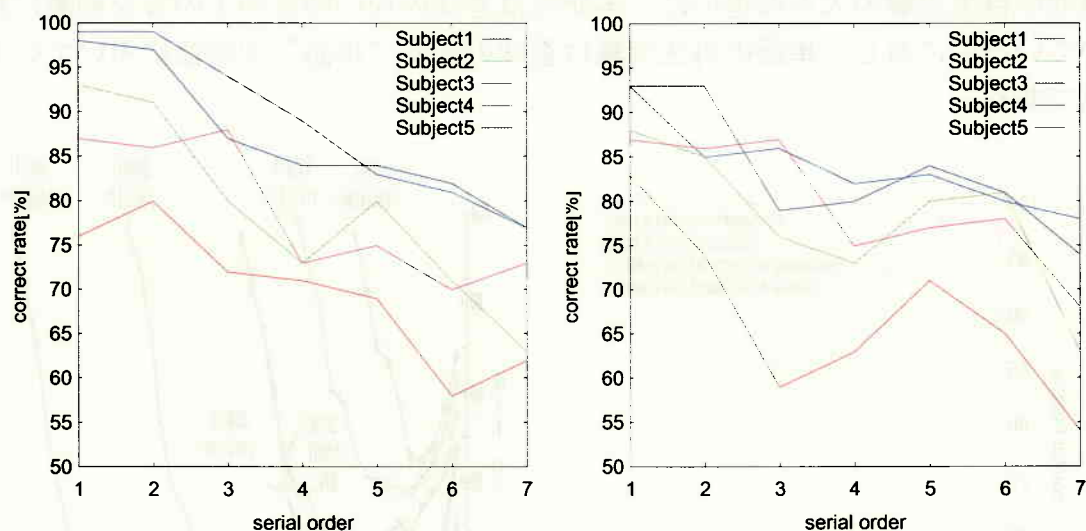


図1.3. 図1.4. 複数被験者による再生実験結果（左：前から照合，右：後ろから照合）

図1.2, 1.3よりどの被験者でも系列位置における前半部の正解率が後半部よりも高くなることが確認された。また、系列位置の後半部について、わずかなではあるが後ろから照合させたときの正解率が前から照合させたときの正解率よりも高いことがわかる。

以上の結果より、聴覚系における時系列上における信号の長短の弁別において、短音の時間長が同じ場合には、構成音数が異なる場合においても、後半に入ってくる信号よりも前半に入ってくる信号の方がより弁別できると考えられる。また、信号の前半部と後半部とでは、信号を再生している記憶箇所が異なると推測される。



### 3. 5 考察

ここで、一般的な短期記憶研究での系列位置曲線と、今回の再生実験の比較を行う。下の図15の右図は被験者にいくつかの単語を1秒または2秒呈示して、それを被験者に自由再生させた実験結果である。単語の自由再生実験と本研究の再生実験の大きな違いは、本研究は意味のない時系列上の音を刺激に用いているのに対し、単語の再生実験は意味のある“単語”を刺激に用いている点である。

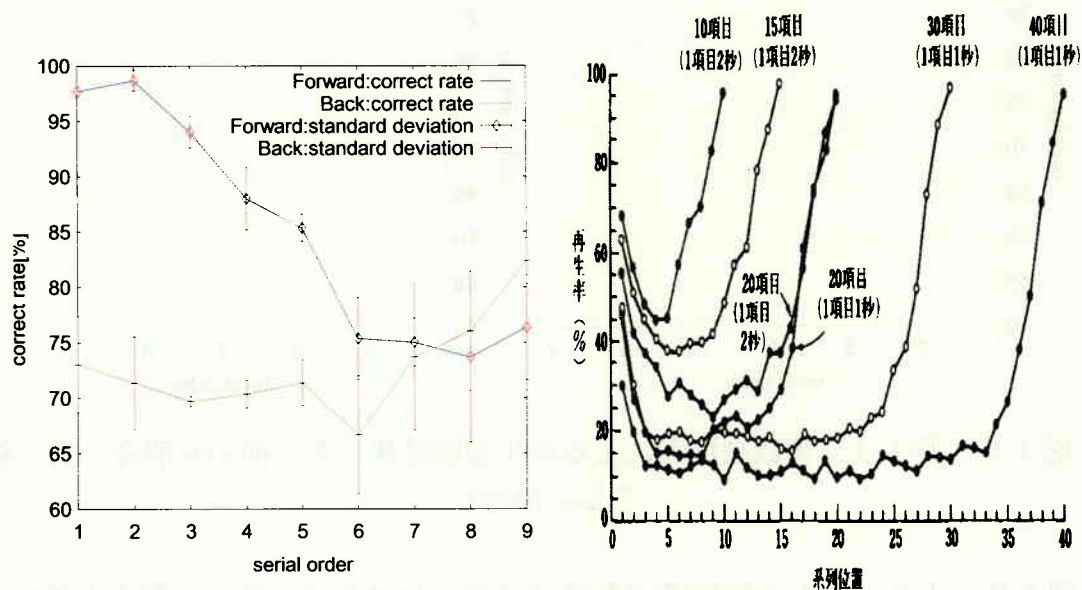


図15. 単語を呈示した場合の系列位置曲線（自由再生）と再生実験結果例

両者の系列位置曲線の共通点としては、呈示刺激項目数が多いほど系列位置における再生率が低い点が挙げられる。これは、呈示刺激項目数の違いによる全体的な再生率の変化については、刺激項目が多くなるにつれて、1つの項目にかけるリハーサルの量が減少するので情報が短期記憶貯蔵庫から忘却されやすくなるため、結果的に全体の再生率が低くなると考えられる。

両者の相違点としては、系列位置曲線の形状が大きく異なることが挙げられる。系列位置効果に注目すると、初頭効果は音系列の再生に強く見られ、新近性効果は単語の再生に強く見られる。二重貯蔵モデルにおける系列位置効果の観点から考えると、音系列の再生時においてはリハーサルを多く受ける初めの方にインプットされる情報は再生しやすいが、言葉の再生では短期記憶貯蔵庫から直接取り出せる最後部の情報が再生しやすいことになる。よって、意味のない時系列上の音と、意味のある言葉とでは短期記憶の運用が異なることを表していると考えられる。

## 第4章

# 短音と長音より構成される信号による再認実験

### 4. 1 はじめに

一般に再生は提示された刺激を正確に再現する必要があるために、虚答率は低いながら正答率も低くなるために短期記憶の容量を測る点においては不満足な実験方法である。再認は刺激を呈示後、新たに呈示される項目が刺激と同様か指摘させるために、正答率は高くなるので短期記憶の容量を測るのに適した測度であるが、当て推量で答えても一定の正答率が得られる。このように再生と再認には互いに補完するような長所と短所がある。

### 4. 2 目的

再生実験では、一概には比較できないが、意味のない音系列と意味のある言葉とでは短期記憶の運用方法が異なることを示唆する結果が得られた。そこで、再生実験のデータを補完する実験として、また詳しい短期記憶の容量を測るために再認実験を行う。

### 4. 3 実験方法

はじめに複数の短音と長音から構成された信号（刺激）を流す。次に刺激と同じ項目（ターゲット）か刺激と1音だけ異なる項目（ディストラクタ）をながし、同じ項目を指摘させる。ターゲットとディストラクタは1：2の比率で呈示される。これを1試行とし、50試行を1セットとして信号の各構成音につき5セット行う。

## 4. 4 実験

### 4. 4. 1 熟練被験者による再認実験

はじめに、記憶できる信号の長さに限界がある可能性があるために、100msと50msの二種類の信号の長さ（短音の長さ）を用いて十分再認実験にこなれている被験者で再認実験を行った。

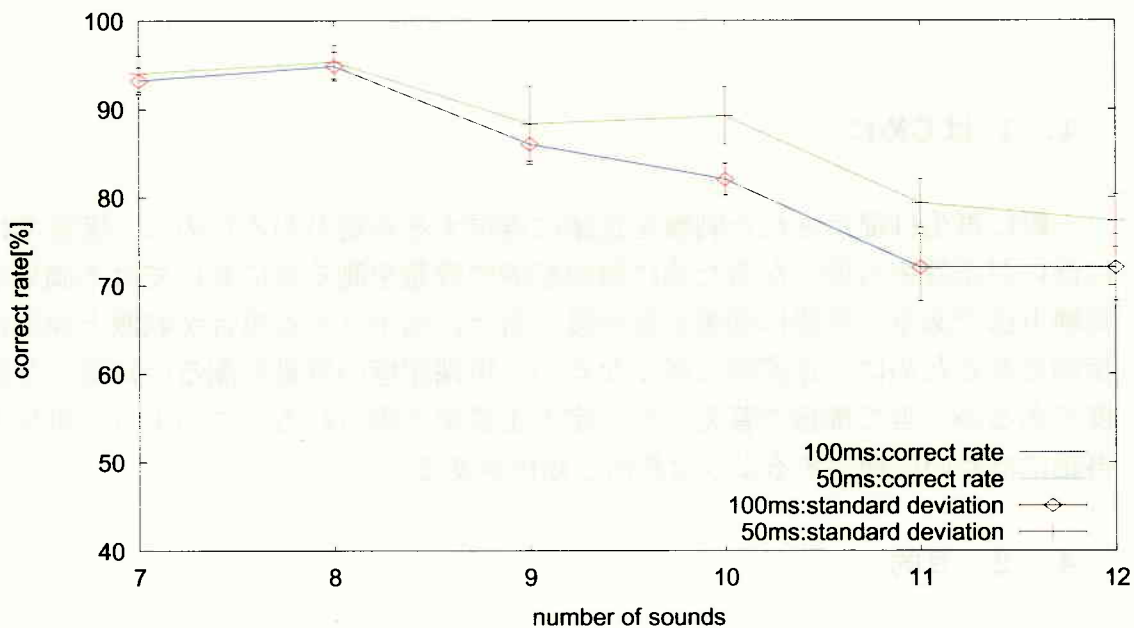


図16. 信号の速さの違いによる正解率

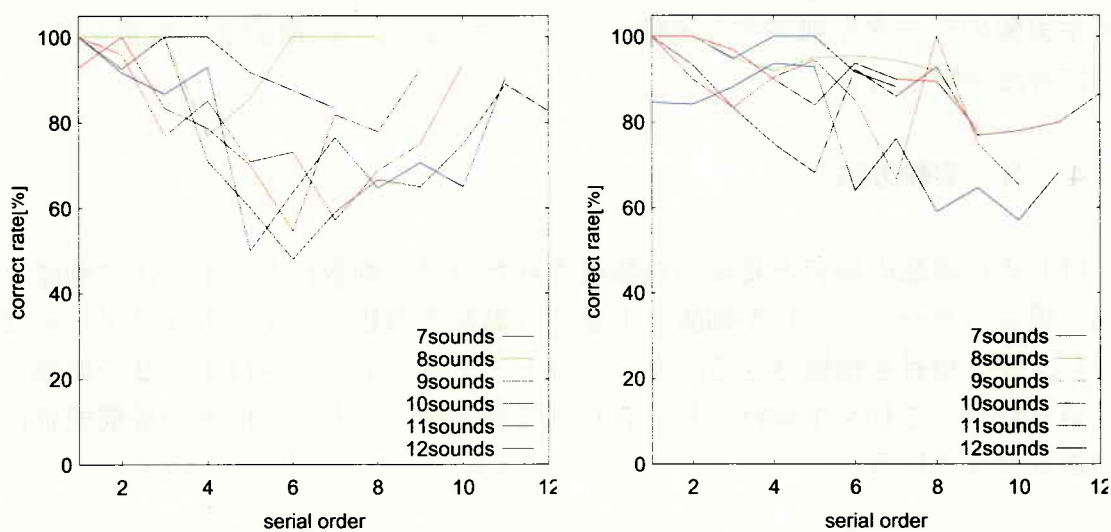


図17. 図18. 系列位置における正解率(右:50ms, 左:100ms)

図16の結果より、信号の速さによって正解率が変化することがわかる。また図17, 18より再生実験結果と同じく正解率が前半部の方が後半部よりも高い事が確認された。

この再認実験より確認された信号の速さの違いによる正解率の変化は、二重貯蔵モデルにおいて短期記憶の容量はスロットの数のみするという絶対容量的な考えを否定するものである。しかし、信号の速さが異なっても、構成音数が増えると正解率が低下していることから、絶対容量的な考えが間違っているわけではない。むしろ、今回の結果は、短期記憶容量は絶対的なものと時間的なものの両者に関係している事を示唆していると考ええる。

次に、信号の速さによって正解率が変化したので、信号の速さを100msから50msまで10msずつ変化させて再び再認実験を行った。

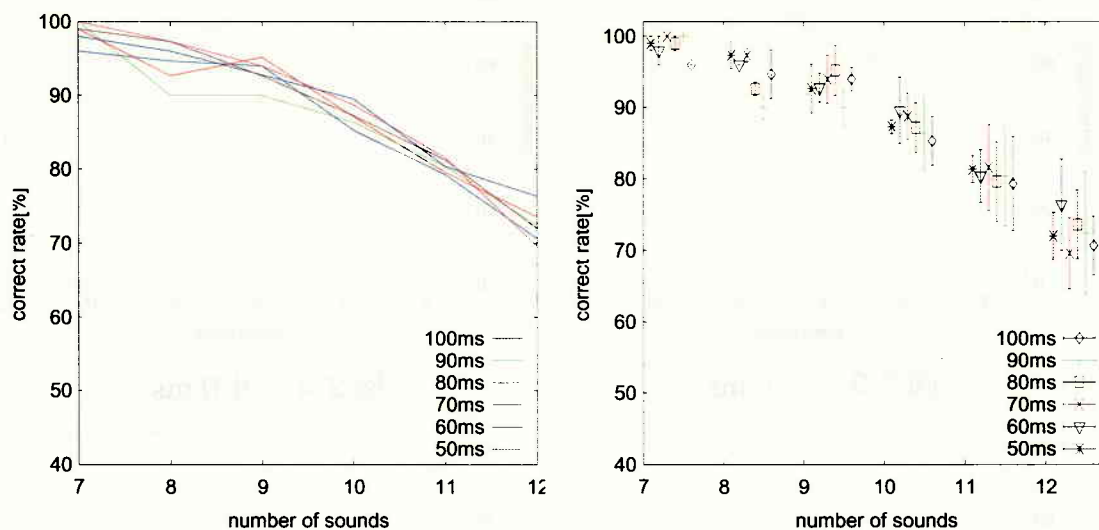


図19. 20. 信号の速さの違いによる正解率 (右図: 標準偏差)

図21から図26は信号の速さ別の系列位置における正解率  
グラフの色の違いは構成音の違いであり、凡例は右の  
にある。

- 7sounds
- 8sounds
- 9sounds
- 10sounds
- 11sounds
- 12sounds

図2  
であり、  
とおりであ

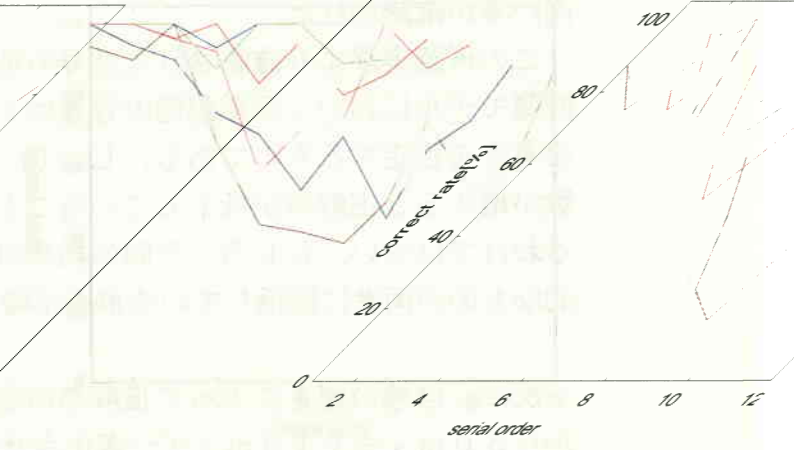
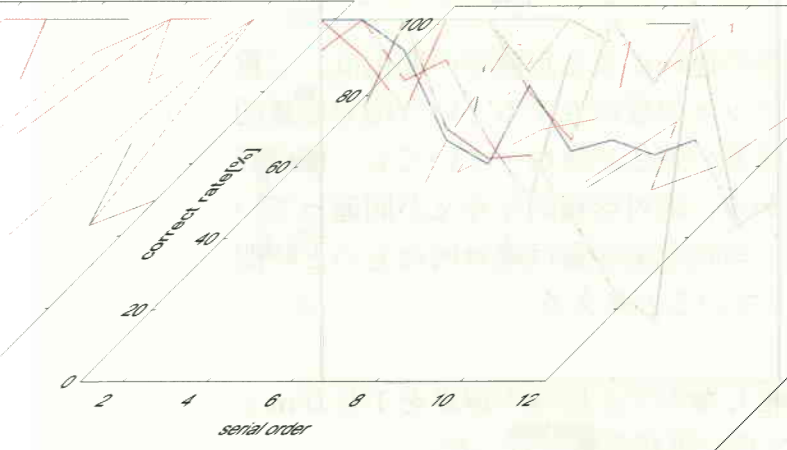


図22. 60ms

図21. 50ms

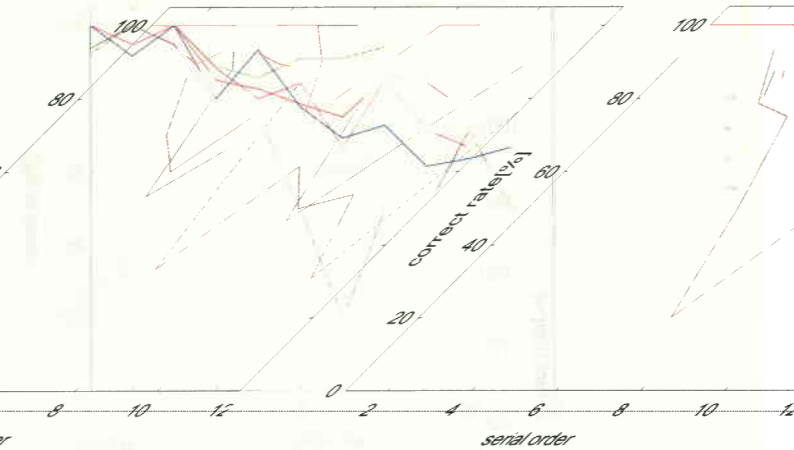
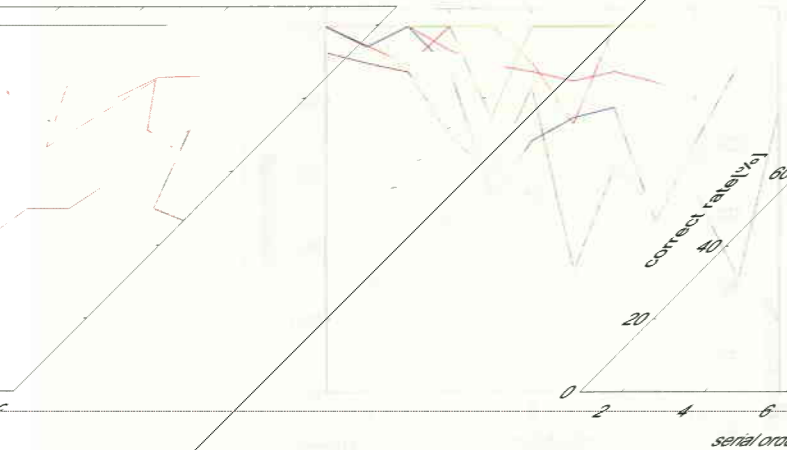


図23. 70ms

図24. 80ms

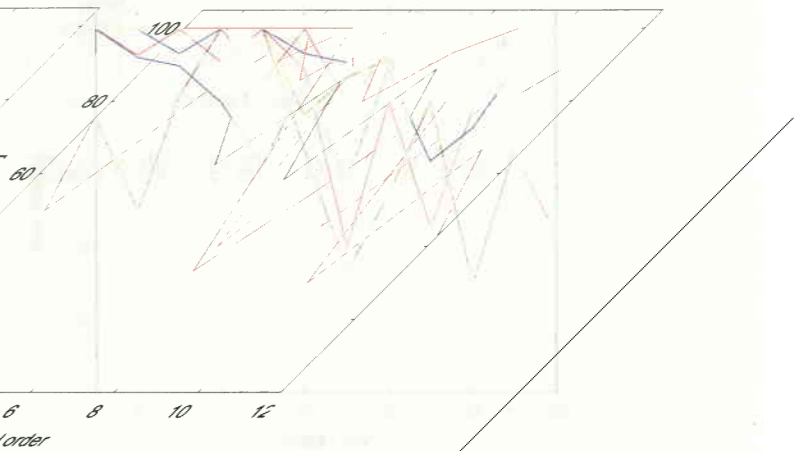
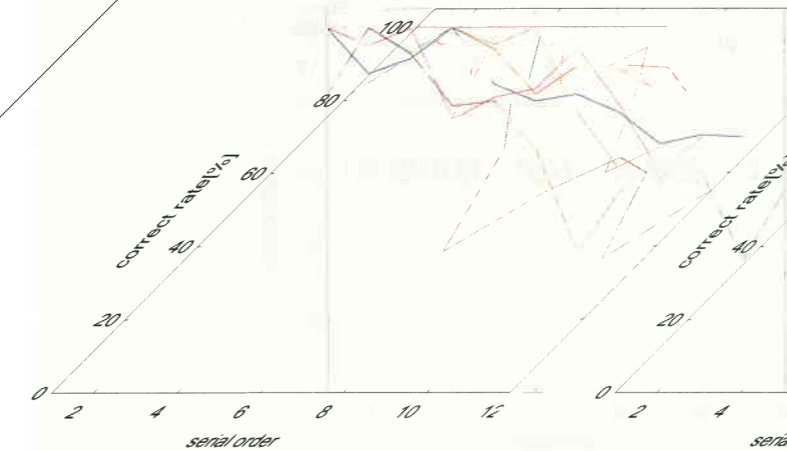


図25. 90ms

図26. 100ms

100ms から50ms まで10ms ごとに再認実験を行ったのだが、先程とは異なり信号の速さの違いによる正解率の変化は見られなかった。むしろ図20の標準偏差が異なる速さの信号でも重なる結果となったことにより、全ての正解率がある曲線に収束するのではないか。これは再認実験を長期にわたって行ったために、被験者が記憶可能な範囲が増大したために、遅い信号(100ms 等)の正解率が全体的に上昇し、その結果速い信号の正解率に近づいたためと考えられる<sup>[12]</sup>。

また、系列位置曲線における正解率については、系列位置曲線が信号の速さの違いと関係なく、再生実験と同様の形状となったことより、再生実験のデータが補完されるものとする。

#### 4. 4. 2 非熟練被験者者による再認実験

次に、再現性の確認と長期訓練による記憶範囲の増大、再認実験になれていない非熟練者によって短音の長さにして100msと50msの信号の再認実験を行った。

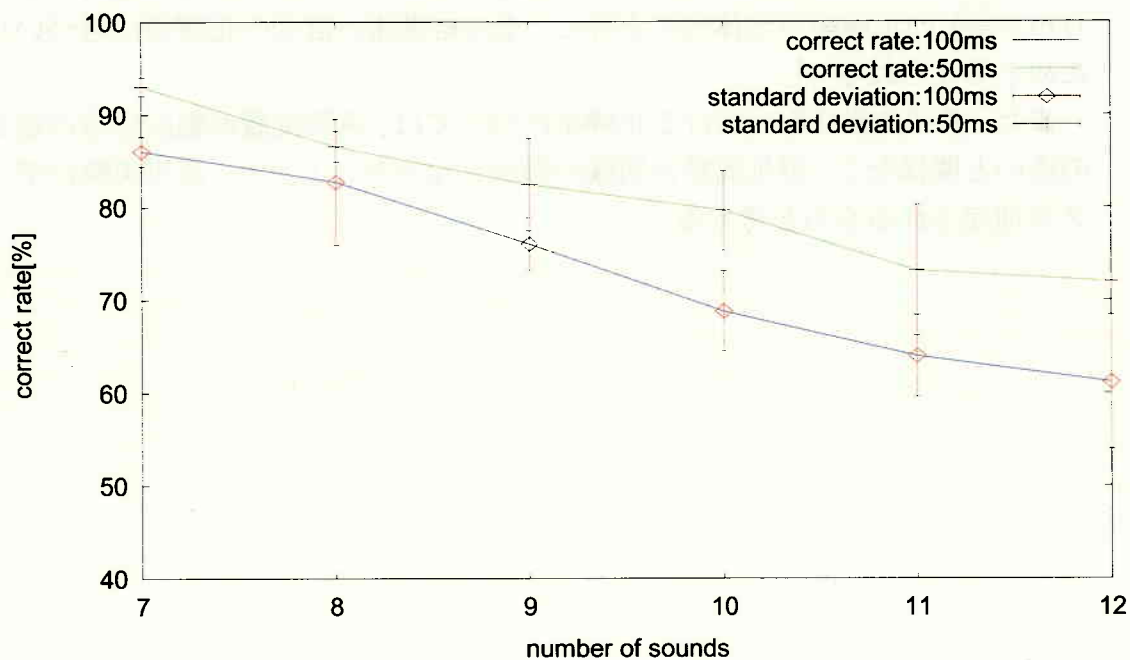


図27. 信号の速さの違いによる正解率(被験者B)

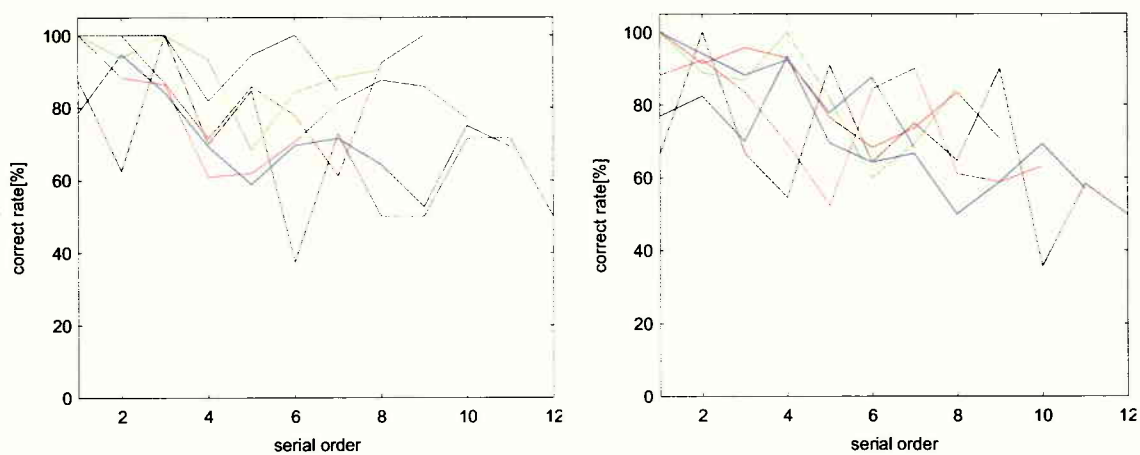


図28. 図29. 系列位置における正解率(左: 50ms, 右: 100ms)



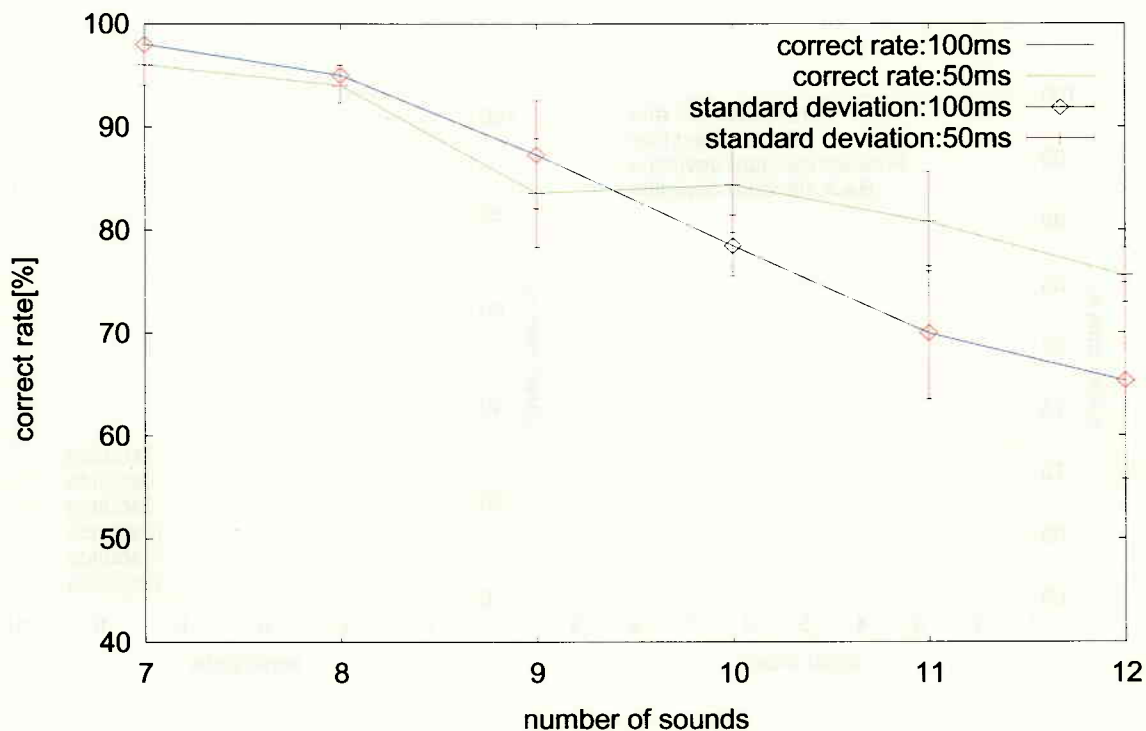


図30. 信号の速さの違いによる正解率(被験者 C)

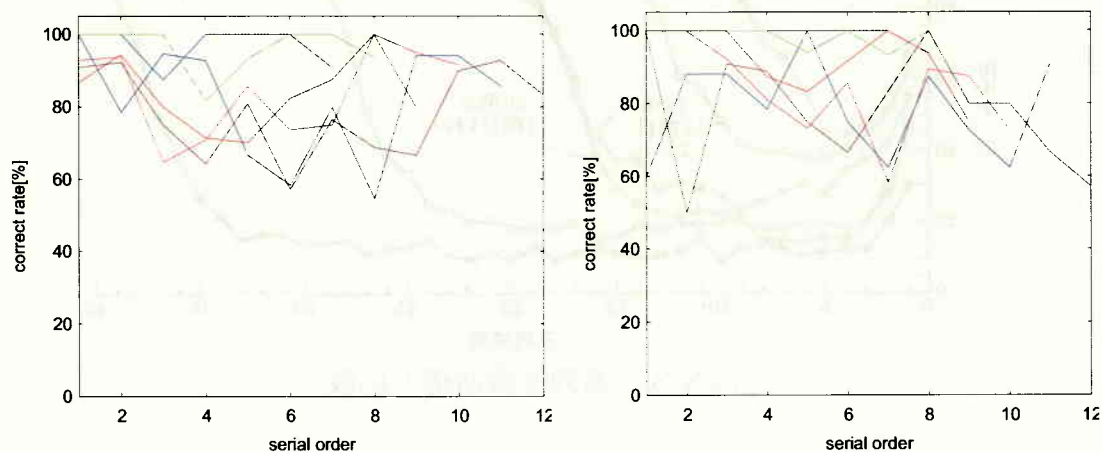


図31. 図32. 系列位置における正解率(左: 50ms, 右: 100ms)

2人の被験者について、信号の速さの違いによる正解率の変化が見られた。系列位置曲線についても、正解率は前半部のほうが後半部よりも高かった。よって再現性が確認されたと考える。

また、系列位置曲線においても熟練被験者と同様の形状となったことにより、再生実験のデータが補完されるものとする。



## 4. 5 系列位置曲線の比較

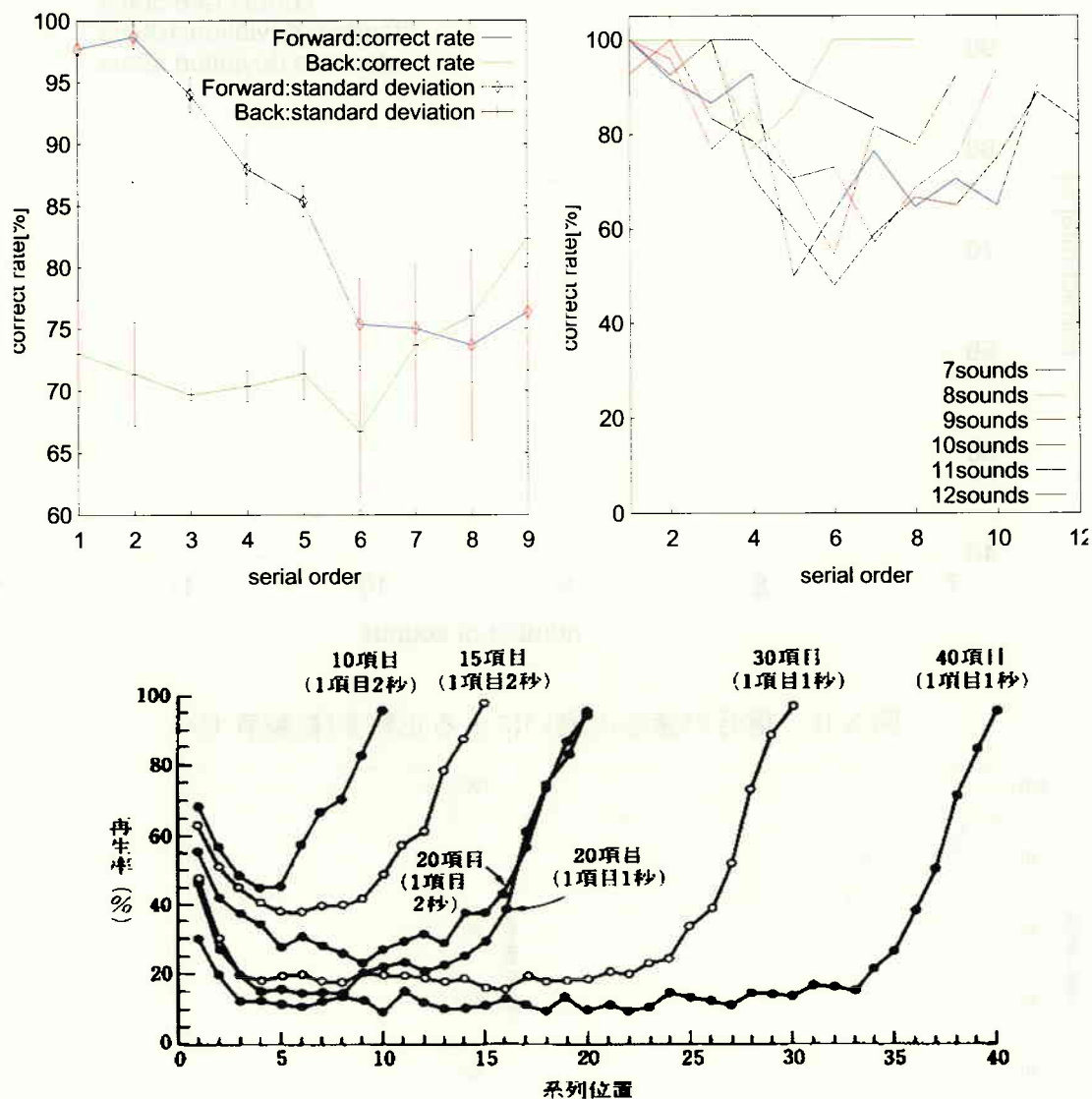


図3.3. 系列位置曲線の比較

再生実験結果と再認実験結果を比較してみると、共に系列位置の前半部の正解率が高く後半部が低いという同じ形状の系列位置曲線である。よって、再生と再認の実験手法の違いによる実験結果の違いは見られない。ゆえに、今回の再生・再認実験から、時系列上の音系列は意味のある言葉とは異なり、初頭効果が新近性効果よりも強く見られることが確認された。

ここで、改めて単語を自由再生させた場合の系列位置曲線と今回の再生・再認実験を比較してみると、系列位置曲線の形状の違いは明確である。よって、言葉と時系列とでの記憶の運用方法の違いが明確に言えるのではないか。

## 第5章

### 結論

再生・再認実験による系列位置による正解率は、前半部が高く後半になるにつれ低くなっていき、最後部でまた高くなった。両実験により得られた系列位置曲線と一般的な系列位置効果を表す系列位置曲線との比較により、時系列上の音系列における短期記憶は言葉や絵などの意味を持つ短期記憶とはその運用方法が大きく異なる事が裏付けられたと考える。

熟練被験者の初期データと非熟練被験者の再認実験結果より、信号が速さで正解率が変化したことは、時系列上の音系列における短期記憶容量が絶対量だけでなく時間も関係していることを示唆するものと考えられる。また、熟練被験者の実験後期データだと、信号時間の違いによる正解率の差は多少あるものの、実験初期のような大きな差はなかった。これは長期訓練により数字記憶範囲の増大した可能性が考えられる。

本研究結果より、私は、少なくとも時系列上の音系列における短期記憶では、二重貯蔵モデルで想定されているスロットはなく、作動記憶における処理資源の容量が限られていて、インプットされた情報の処理レベルが高いほど多くの処理資源(メモリ)を必要とするものと考えられる。(例えば本研究の再認実験だと、1つの刺激項目(音)に対し50msの場合は少しのメモリ、100msの場合は多くのメモリを要すると想定する。)そして、短期記憶内で情報を処理しているとき、使用していないメモリが多いと、その分を処理にまわせるので処理速度が速くなると考えられる。また、熟練被験者の実験後期データより、信号の速さの違いによる正解率の差が見られなくなったことより、訓練を行うことにより、CPUのクロック周波数のようなものが速くなり(神経回路の信号伝達速度が上がる?)、多くのメモリを要するタスクでも時間的に速く処理できると考える。この考えだと、本研究の構成音数が多いほど正解率が低くなることと、信号の速さの違いによる正解率の変化が説明できるのではないかと。

今後の課題として、再認実験の試行回数をさらに重ねることにより、系列位置における信号の速さの違いによる特徴が見られると思う。被験者数が少ないので、多くの被験者で試行を重ね、再現性を確認する必要がある。また、熟練被験者の再認実験において、異なる速さの信号でも正解率がある曲線に収束し

ているように見えることから、多くのデータを積み重ねることにより時系列上における短期記憶容量のロット数と時間の関係をモデル化できるのではないか。また、もしモデル化できたら、現在のコンピュータの音声認識における処理機構よりも、さらに効率のよい音声処理機構が構築可能ではないか。

## 謝辞

卒業研究を進めるにあたり、本研究の機会を与えてくださった東北大学大学院工学研究科電子工学専攻 内田 龍雄 教授に心より御礼申し上げます。

本研究の全般に渡り、適切な御意見・御助言を頂きました東北大学大学院工学研究科電子工学専攻 二見 亮弘 助教授に心より感謝いたします。

本研究を進めるにあたり、貴重な御助言を頂きました東北大学大学院工学研究科電子工学専攻 加納 慎一郎 助手に厚く御礼申し上げます。

本研究のプログラム作成にあたり、快く協力していただいた本研究室の岩泉 亜希子氏に心より感謝いたします。岩泉氏のご協力がなければプログラムを組むのに多大な時間を労したでしょう。本当にありがとうございます。

本研究において、精神的負担の大きい実験の被験者を快く受けてもらいました、本研究室の鈴木 信太郎氏，杉山 貴也氏，伊藤 選氏、東北大学大学院技術社会システム専攻の須田 洋行氏，牧野研究室の梶浦 泰智氏に心より御礼申し上げます。本研究は被験者の方のご協力がなければ成り立ちませんでした。被験者の方々には感謝の気持ちでいっぱいです。特に伊藤 選氏につきましては多くの実験に協力していただき本当にありがとうございます。

最後に内田・二見研の皆様のおかげで、1年間楽しく有意義な研究室生活を送れました。皆様、本当にありがとうございました。

参考文献

参考

事業団：“平成13年度研究年報—戦略的基礎研究推進事業—研究実施  
 計画・戦略目標”，  
<http://www.kisoken/crest/report/heisei13/html/mokuhyou/mokuhyou03.html>

[1] 科学技術振興  
 報告 脳機能の解  
<http://www.jst.go.jp/ku>  
 ml (2002)

列パターン識別機能について—可変スピードのモ  
 音響学会 聴覚研究会資料 H-90-23 (1989)

[2] 二見亮弘, 星宮望：“聴覚系の時系  
 ルス符号列を用いた聴取実験—”, 日本

電気通信学会

[3] 二見亮弘, 星宮望：“時間情報処理のモデル”

システム”, 科学技術振興事業団  
 課題・脳を創る 中間評価結  
 (1998)

[4] 河原英紀：“聴覚の情景分析に基づく音声・音響処理  
 戦略的創造研究推進事業 CREST 平成9年度採択研究  
 果 <http://www.jst.go.jp/kisoken/crest/intro/jigyoku.html> (1

Processing An

[5] Peter H. Lindsay & Donald A. Norman : “Human Information  
 Introduction to Psychology 2nd Edition”, Academic Press (1977)

tion”

[6] Atkinson, R.C. & Shiffrin, R.M. : “The psychology of learning and motiva  
 Advances in research and theory Vol.2 Academic Press (1968)

[7] Miller, G.A. : “The magical number seven, plus or minus two : Some limits on our  
 capacity for processing information”, Psychological Review 63, 81-97. (1956)

[8] Watkins, M.J. & Peymircioglu, Z.F. : “Tree recency effects at the same time”, Journal  
 of Verbal Learning and Verbal Behavior 22, 375-384 (1983)

[9] Bjork, R.A. & Whitten, W.B. : “Recency-sensitive retrieval processes in long-term  
 free recall”, Cognitive Psychology 6, 173-189 (1974)

Crowder, R.G. : “Short-term memory: Where do we stand?”, Memory & Cognition  
 1, 142-145 (1993)

[10]

21,

- [11] 高野陽太郎 編：“認知心理学 2：記憶”，東京大学出版会（2003）
- [12] Chase,W.G. & Ericsson,K.A.：“The psychology of learning and motivation, Vol.16”,  
Academic Press (1982)