Effects of Category Labels on P300 in Facial Recognition

Pengfei Yan[1], Yoshiko Yabe[2], and Hiroaki Shigemasu[3]

[1] Graduate School of Engineering, Kochi University of Technology, Tosayamada, Kami-city, Kochi 782-8502 Japan
166001m@gs.kochi-tech.ac.jp

[2] Research Institute, Kochi University of Technology, Tosayamada, Kami-city, Kochi 782-8502 Japan
yy47151@gmail.com

[3] School of Information, Kochi University of Technology, Tosayamada, Kami-city, Kochi 782-8502 Japan
shigemasu.hiroaki@kochi-tech.ac.jp

Abstract. Gordon and Tanaka (2011) suggested that name labels such as "Joe" facilitated face memory and elicited large P300. However, when name labeling was used in Brain Computer Interface (BCI) by which users can choose people with their faces, preserved P300 response to the previous target might be problematic because of the effect of memory. Our study utilized categorical labels of occupation instead of name labels, and investigated the effects of task-relevancy, face exposure and category labels in face selection task. Participants were required to judge whether each stimulus was a target or not. Results showed that although it was consistent with the name-label situation that P300 was enhanced by task-relevant targets, repeated exposure to previous target didn't increase P300 in category-label situation in contrast to the previous study. These results suggest that categorical labeling is more appropriate for BCIs, because task-relevant target face elicits larger P300 than other faces.

Keywords: Face representation, Electroencephalogram (EEG), P300, Categorization.

1 Introduction

Since researches on Brain Computer Interfaces (BCIs) starting in 1970s [1], BCI has received increasing attention due to its capability of assisting, augmenting and repairing human cognitive or sensory-motor functions. It also enables interaction between human and computer without muscular intervention. Recently, a considerable number of studies about human BCI have used Electroencephalography (EEG) because of its high temporal resolution and low set-up cost.

Many EEG-based BCI systems utilized event-related potential (ERP) to construct classifiers, like SVM and LDA, in target selection procedures such as "P300-Speller" [2]. The P300-Speller uses visual stimulation of letter matrix to input words or sentences. A P300 response is elicited when the chosen letter

K. Imamura et al. (Eds.): BHI 2013, LNAI 8211, pp. 33–41, 2013.

is flashed, and it is captured by a subsequent classification procedure. Due to weakness of P300 response, the P300-Speller has to repeatedly present stimuli to select single-letter with high classification accuracy, and further requires many more repetitions to input words or sentences, so the alphabetic BCI system is limited in practice. Here we examined the possibility of BCI system in which desirable human face is selected from candidates under some context, e.g., we want to call a specific person from a group.

In the study of Kaufmann et al. (2011), familiar faces were transparently superimposed on characters of P300-Speller for the purpose of getting clearer P300 response. Under the situation, P300-Speller performance was improved along with significantly enhanced ERP response [3]. To further investigate the feasibility of facial selection in BCIs, it is worthwhile to explore what facial properties are related to ERP components, and how face representations are acquired and maintained for working memory.

In fact, many previous ERP studies have paid much attention on face recognition in terms of P300 component, which is associated with categorization processes. The enhanced P300 were elicited in response to task-relevant targets, and the response was significantly larger especially when target was the face of observer's own [4]. Other self-related information with high social or adaptive value, like self-name, can also produce an increased P300 response [5].

As for the studies of name-face associations using ERP, no less than two different types of tasks were performed. One kind of task required participants to identify named faces as familiar or not [6], while another required participants to use name label to categorize a face as a target [7]. Moreover, a reliable change in the N250 was examined in the latter, possibly because name labels enable observers to be familiar with faces. In more details, participants were asked to monitor for a target face with specific name such as "Joe" presented among a series of non-target distractive faces. At the halfway point, target was switched into another face with the name such as "Bob". The difference between the first and second halves was whether a face was known as "Bob", the target, or not. The result showed that the name-label may anchor acquisition of a face representation, and that label was associated with percept through repeated practice in the process of actively acquiring robust face representations. With the effect of name label, P300 response was still maintained for "Joe" face which was no longer task relevant in name-label situation, so name label is not very suitable to use in real BCI environment, especially for target selection procedure where the P300 response to previous target should contaminate to select present target as minimally as possible. We expect that occupations as category labels provide a possible solution to the problem, considering the flexible anchoring effect of occupational labels. Because occupations are stored separately from names [8] and it takes shorter time for occupations to be retrieved from a face than for names no matter whether the face is familiar or not [9,10], benefitting from its non-arbitrariness, high frequency and high imageability [11]. The study was conduct with two-folded purposes. One was to investigate how face representation is affected by category label, task relevancy and face exposure in terms of P300 component. Besides, we sought to explore effect of semantic information to facilitate face recognition. To this end, P300 responses of target-face were compared between name labeling situation and category labeling situation with distinct capability for expressing semantic information.

2 Methods

2.1 Participants

Ten undergraduate students (two females) from Kochi University of Technology, aged 20-25 years ($M = 21.6$ years) participated in this experiment. All were right handed, and all had normal or corrected-to-normal visual acuity. Informed consent was obtained from all participants before the experiment.

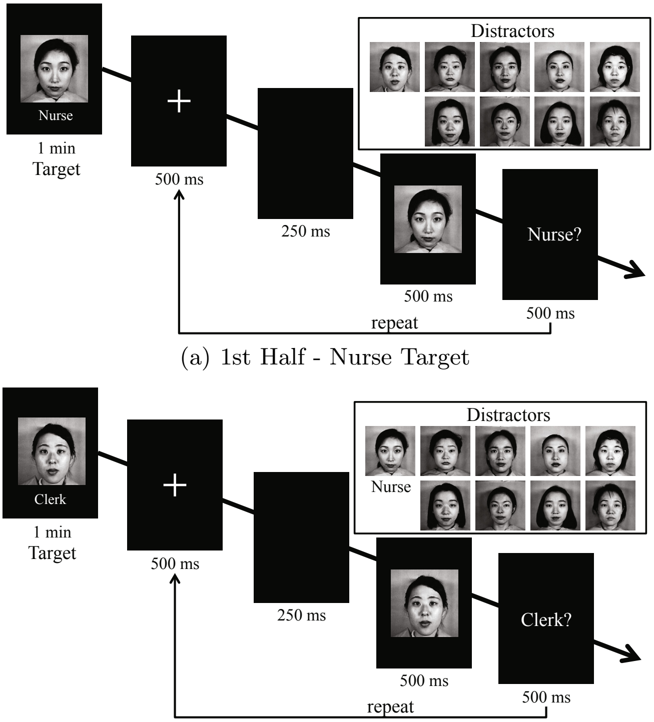
2.2 Materials

The stimuli comprised images of 10 Japanese female frontal faces with neutral expressions from the Japanese Female Facial Expression (JAFFE) database [12]. All of the images were gray-scaled, unfamiliar to all the participants, and cropped to a dimension of 256 pixels around the head area, allowing for a visual angle of 5.9 deg both horizontally and vertically. We used category-labels of occupation such as Nurse and Clerk to categorize one face as target. Assignment of faces to Nurse and Clerk conditions was counterbalanced across participants.

2.3 Procedure

Participants sat in a comfortable armchair at a distance of 650 mm in front of a computer screen (FlexScan L557, 17 in., 1280 $\times$ 1024, EIZO) and were instructed to relax and remain still as possible. Their left hands were asked to place on the amplifier box with metal enclosure for grounding purpose. After the EEG electrodes were applied, participants were introduced to the target (Nurse) face presented at the center of the screen, and asked to remember the face for 1 minute. In the practice phase, participants viewed a series of faces on a computer screen, presented one at a time. They can view each face many times back and forth, and ended up with picking out the target face. After the practice phase, if no further questions, participants proceeded to the experimental phase.

Fig. 1 illustrated the experimental procedure of the first and second halves for one participant. As shown in Fig. 1(a), each trial consisted of a blank screen with a fixation cross at the center for 500 ms, followed by a blank screen lasting 250 ms, then a face stimulus for 500 ms, and ended with a prompt screen reading "Nurse?". Participants were instructed to select the "1" key if the target Nurse face appeared or the "2" key if any other face appeared. At the halfway point in the experiment, participants were shown one of other faces to be the next target, called "Clerk", for 1 minute. For the second half of the experiment, participants were instructed to select the "1" key if they saw "Clerk" face and the "2" key



(a) 1st Half - Nurse Target

(b) 2nd Half - Clerk Target

Fig. 1. Experimental procedure of the first and second halves

if they saw any other face (including "Nurse" face). The prompt screen was altered to reflect the target change, with the prompt reading "Clerk?" as shown in Fig. 1(b). All other aspects of the trial were the same as in the first half the experiment. It is worthwhile noting that each participant, before each half, was naïve to target face and its label of the half. However, "Nurse" face as the target of the first half had been familiar at the beginning of the second half which target was "Clerk" face. Especially, responses elicited by other 8 faces were averaged as "Other" face condition.

For each participant, each facial image was presented 72 times of which 36 times for the first half and the other for the second half. A total of 720 trials were divided into six 120-trial blocks. The switch in target faces occurred at 360 trials (the half-way point), such that "Nurse" was the target in the first half of the experiment and "Clerk" was the target in the second half. Unique faces for both "Nurse" face and "Clerk" face were used for each participant. Namely, each facial image was labeled as Nurse or Clerk with the same possibility. Participants had a break for 30 seconds after each block, and remaining time was shown to participants in real time, thereby reminding participants to prepare for next block. To improve data validity of participants, they had to report the number of target faces they recognized during each half. Additionally, an impedance check was performed before each half.

2.4 EEG-ERP Methods

The EEG data were acquired using a g.USBamp (24 Bit biosignal amplification unit, g.tec Medical Engineering, Austria) and sampled digitally at 256 Hz with a bandpass filter of 0.010-60 Hz and a notch filter of 60 Hz. The g.SAHARAclipGND (ground electrode, g.tec Medical Engineering, Austria) was located on the forehead; the g.SAHARAclipREF (reference electrode, g.tec Medical Engineering, Austria) was mounded on the right mastoid. g.SAHARAsys (dry active electrode system, g.tec Medical Engineering, Austria) was used for electrodes. We used dry active electrodes for the convenience of an actual use for BCI and to avoid or reduce artifacts and signal noise resulting from high impedance between electrodes and skin. The EEG was recorded from ten electrodes placed on extended 10/20 system position with g.GAMMAcap[2] (g.tec Medical Engineering, Austria) (shown in Fig. 2). Then the data were converted to double precision, bandpass filtered between 0.5 and 30 Hz on line. Obtained data were then subjected to several further filtering and an ocular correction process off-line using g.BSanalyze (biosignal processing and analysis toolbox, g.tec Medical Engineering, Austria) as follows: first, EEG was filtered using a bandpass filter of 0.5-20 Hz; subsequently, process of removing drift was executed on the obtained data using 50 samples of interval length; thereafter, each ERP segment was divided into 800-msec epochs, beginning 200 ms before stimulus onset and subsiding 600 ms after stimulus onset. Thus, the start of each epoch coincided with the blanking screen stimulus that preceded the presentation of the face stimulus. Trials were rejected if voltage change exceeded 70 μV. All the trials were baseline-corrected 200 ms before stimulus onset.

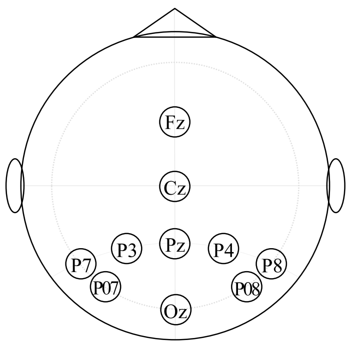


Fig. 2. Electrode configuration

3 Results

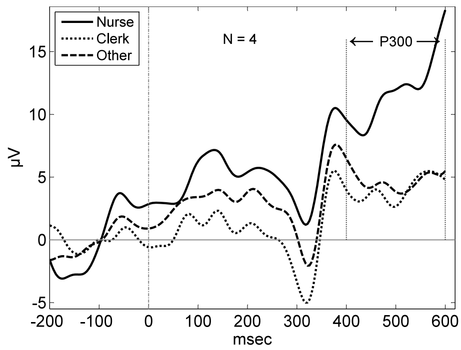
Accuracy of target selection task fluctuated with different participants, although no significant increase or decrease in accuracy was observed as a result of switching the target from Nurse in the first half to Clerk in the second half. Because of imperfect experimental environment and control condition, some artifacts for EEG signals cannot be avoided, resulting in low accuracy for some participants. Another source of the undesirable accuracy may be from unfamiliar faces to be used, and memory decay may occur during the experiment. Under the set-up, it cannot be determined that the low accuracy in second half was from the imperfect experimental environment, the effect of previous target or both. For a more persuasive analysis, we selected data from participants with accuracy over 90% in both halves. Additionally, data from one participant was excluded due to excessive artifacts in the EEG signal. Data of the remained four participants were utilized for data analysis, before all incorrect trials were further discarded.

As the mean amplitude of P300 component across the midline (Fz, Cz, Pz and Oz) revealed that the largest mean amplitude was at the Cz electrode, we focused on channel Cz for the analysis.

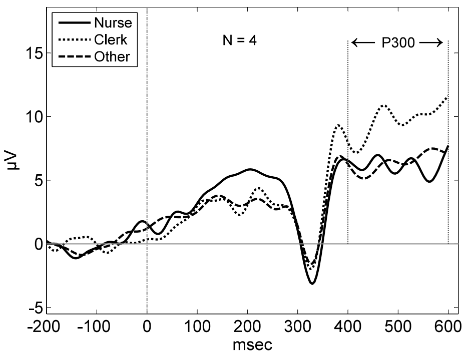
The computation of mean amplitude for component P300 utilized a time window from 400 to 600 ms post-stimulus onset as the previous study [7]. Fig. 3 gave the grand average ERP waveform among valid participants and bar graph of P300 mean amplitude at channel Cz for the Nurse, Clerk and Other condition during the first and second halves.

As can be seen from Fig. 3(a) and 3(b), P300 amplitude was larger in response to the target face than non-target faces not only during the first half but also during the second half, while no significant difference on P300 amplitude between the first and the second halves were observed for Other condition. The P300 amplitude for "Nurse" face as target in the first half reduced to the level of that for "Other" condition in the second half.

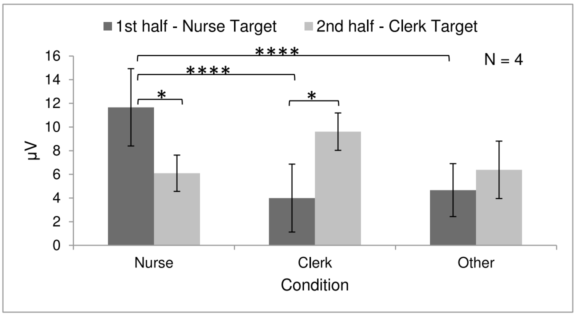
To analyze the difference of peak amplitude of the P300 component, a 2×3 ANOVA was carried out with $Half$ (first, second) and *Condition* (Nurse, Clerk, Other) as within-subjects factors. All reported post hoc comparisons were reliable at the $p = .05$ level. The result of ANOVA showed a significant interaction effect between *Condition* and $Half$ on channel Cz ($F(2, 6) = 14.02, p < .001$). A significant simple main effect of $Half$ was observed for Nurse ($F(1, 9) = 5.80$, $p < .05$) and Clerk ($F(1, 9) = 5.91, p < .05$) but not for Other ($F(1, 9) = 0.55, p = .48$). Additionally, we observed a significant simple main effect of *Condition* for the first half ($F(2, 12) = 13.16$, $p < .001$) but not for the second half ($F(2, 12) = 2.78$, $p = .10$). The further analysis of multiple comparison based on Ryan's method showed a significant difference of Nurse from Clerk ($t(12) = 4.63$, $p < .05$) and Other ($t(12) = 4.23$, $p < .05$). No other main effect and higher-order interactions were found on channels Cz.



(a) 1st Half - Nurse Target



(b) 2nd Half - Clerk Target



(c) P300 mean amplitude at channel Cz

Fig. 3. Grand average ERP waveform and Bar graph of P300 mean amplitude at Channel Cz

4 Discussion

In the current study, the effects of category labels, exposure and task relevancy were tested on the formation of face representation, and were compared with that of name labels of the study of Gordon and Tanaka (2011) [7]. Participants were required to monitor for a target face "Nurse" in the first half and a target face "Clerk" in the second half of the experiment.

The result of this study shed light on the role of the P300 in face-category learning. Congruent with previous studies [7,13], the P300 amplitude was consistently larger to task-relevant target face. However, there was an inconsistency that the P300 response was not enhanced by former target. Although it was difficult to determine how much exposure enhanced P300 response for specific faces, exposure cannot obviously enhance a P300 component, because in the second half, when the target face was switched from "Nurse" to "Clerk", P300 to the target Clerk face increased while that to the previous target Nurse decreased. Moreover, the "Nurse" face produced P300 response with no larger amplitude than "Other" faces, when it was no longer task relevant. So inconsistent with the previous study, only task relevance seemed to be essential to maintain the P300 component and no significant effect of target face of the first half was remained in the second half. Therefore, our findings suggested that categorical labels may be more appropriate for BCI systems related to face recognition compared with name labels, considering the strong task relevance and the weak exposure relevance on P300 response of face changes in categorical labeling situation.

For the difference between name labeling and category labeling, one possible explanation is that occupations are conceptual representations different in both spelling and semantics, while names are lexical symbols only different in spelling. Especially names and occupations are stored separately in speech production system and semantic system respectively of human brain [14]. When occupations were rendered meaningless by using nonsense words, recall of these occupations was as poor as recall of names [15]. Accordingly, semantic information took a critical role in label recalling, considering that context and visual appearance can give clues to a person's occupation. In our study, each trial was characterized by semantic information as well as facial features, label spelling and exposure. Benefiting from contribution of semantic information, there was a significant difference between target face and non-target faces. Because of a large difference between these two faces, it was considered that "Nurse" face as previous target did not elicit a significant P300 response as target "Clerk" face did in second half, even when it was known as previous target face at the beginning of the second half. But in name labeling situation, the same factors except semantic information were used to characterize each trial. The lack of semantic information induced the result that the difference between target faces in two halves was not significant contrary to category labeling situation, so the previous target elicited a P300 response of nearly equivalent magnitude with the target in the second half.

In future, we will develop a BCI system related to face selection in terms of the findings of this study showing that categorical labeling is more appropriate for face selection task. Another possible work is to investigate the role of category label in acquisition and maintaining of face representation for further examining relationship between semantic label and face representation.

References

1. Vidal, J.J.: Toward Direct Brain-Computer Communication. Annu. Rev. Biophys. Bio. 2(1), 157–180 (1973)
2. Farwell, L.A., Donchin, E.: Talking off the Top of Your Head: Toward a Mental Prosthesis Utilizing Event-related Brain Potentials. Electroen. Clin. Neuro. 70(6), 510–523 (1988)
3. Kaufmann, T.C., Schulz, S.M., Grünzinger, C., Kübler, A.: Flashing Characters with Famous Faces Improves ERP-based Brain-Computer Interface Performance. J. Neural. Eng. 8(5), 056016 (2011)
4. Ninomiya, H., Onitsuka, T., Chen, C.H., Sato, E., Tashro, N.: P300 in Response to the Subject's Own Face. Psychiat. Clin. Neuros. 52(5), 519–522 (1998)
5. Tacikowski, P., Nowicka, A.: Allocation of Attention to Self-name and Self-face: An ERP Study. Biol. Psychol. 84(2), 318–324 (2010)

6. Kaufmann, J.M., Schweinberger, S.R., Burton, A.M.: N250 ERP Correlates of the Acquisition of Face Representations across Different Images. J. Cognitive. Neurosci. 21(4), 625–641 (2009)
7. Gordon, I., Tanaka, J.W.: Putting a Name to a Face: The Role of Name Labels in the Formation of Face Memories. J. Cognitive. Neurosci. 23(11), 3280–3293 (2011)
8. Bruce, V., Young, A.: Understanding Face Recognition. Brit. J. Psychol. 77, 305–327 (1986)
9. Young, A.W., Ellis, A.W., Flude, B.M.: Accessing Stored Information about Familiar People. Psychol. Res. 50(2), 111–115 (1988)
10. Cohen, G., Faulkner, D.: Memory for Proper names: Age Differences in Retrieval. Brit. J. Dev. Psychol. 4(2), 187–197 (1986)
11. McWeeny, K.H., Young, A.W., Hay, D.C., Ellis, A.W.: Putting Names to Faces. Brit. J. Psychol. 78(2), 143–149 (1987)
12. Lyons, M.J., Kamachi, M., Gyoba, J.: Japanese Female Facial Expressions (JAFFE). Database of Digital Images (1997)
13. Tanaka, J.W., Curran, T., Porterfield, A.L., Collins, D.: Activation of Preexisting and Acquired Face Representations: The N250 Event-related Potential as an Index of Face Familiarity. J. Cognitive. Neurosci. 18(9), 1488–1497 (2006)
14. Hanley, J.R.: Are Names Difficult to Recall Because They Are Unique? A Case Study of a Patient with Anomia. Q. J. Exp. Psychol-A 48(2), 487–506 (1995)
15. Cohen, G.: Why Is It Difficult to Put Names to Faces? Brit. J. Psychol. 81(3), 287–297 (1990)