日本語の基本色名における色の中心度と短期記憶パフォーマンスとの関係に対する実験的研究
An Experimental Study of the Relationship Between Focality and Short-Term Memory of Colors in the Japanese Language

方 思源†，松居 辰則‡
Siyuan Fang, Tatsunori Matsui
† 早稲田大学大学院人間科学研究科，‡ 早稲田大学人間科学学術院
Graduate School of Human Sciences of Waseda University, Faculty of Human Sciences of Waseda University
siyuanfang@asagi.waseda.jp

Abstract

In this study, we investigated the relationship between color focality and short-term memory (STM) performance for colors in the case of the Japanese language, firstly aiming to evaluate the universality of the focality effect of language-specific basic color terms (LSBCTs). Noticing that the LSBCTs in the Japanese language and the universal basic color terms (UBCTs) are similar in focal color location, we also employed our experiment data to approach the debates on the focality effect of UBCTs. Our experiment found no correlation between color focality and STM performance for colors, but detected a strong correlation between color discriminability and STM performance for colors, suggesting that STM performance for colors was mainly determined by color discriminability. Considered together with the results of some previous studies, our experiment results implied the possibility that the strength of the focality effect of LSBCTs may differ greatly across languages, and that there lacks enough evidence supporting the existence of the focality effect of UBCTs.

Keywords — color, category, Japanese, focality, short-term memory, discriminability, basic color term

1. Research Background and Objective

The research history on the relationship between linguistic color categorization and color perception began with Brent Berlin and Paul Kay[1]'s study. Berlin and Kay investigated 98 languages and found that the basic color terms in any language are restricted in the eleven terms *white*, *black*, *red*, *green*, *yellow*, *blue*, *brown*, *purple*, *pink*, *orange* and *grey*. To explore the cause of this phenomenon, Eleanor Rosch Heider[2] conducted a series of psychological experiments and put forward the hypothesis (called "Rosch's hypothesis" for short in this paper) that the special status of these eleven color terms are derived from the universal perceptual saliency of their focal colors. One of her experiments delved into the relationship between focality of colors and humans' short-term (STM) performance for colors. The focality of a color means the goodness of the color as a representative of the category it belonged to. In this experiment, Rosch used a simplified version of the color array used by Berlin and Kay[1], and selected from it 24 color chips as experimental stimuli of which eight were of the highest focality for each of the eight basic color terms *red*, *green*, *yellow*, *blue*, *brown*, *purple*, *pink*, *orange* (i.e. focal colors), eight of intermediate focality (i.e. internominal colors) and eight of low focality (i.e. boundary colors). Each of the stimuli was tested on two groups of subjects, one speaking English and the other Dani. In each trial, the subject was instructed to watch the color chip for five seconds and search for it in the color array after a 30-second interval. The results showed that for both subject groups, the focal colors were better recognized than the non-focal colors, namely the internominal and the boundary colors. Considering that the Dani language has only two basic color terms *white* and *black*, this result implies, in Rosch's view, that in any language the focal colors of the basic color terms are more perceptually salient, and hence tend to be better recognized through STM than other colors, no matter whether the lexicon of the language actually includes all these color terms or not.

In this paper, we name the correlation between color focality and STM performance for colors as "focality effect". We call this effect "focality effect of universal basic color terms (UBCTs)" if the calculation of color focality is based on the universal basic color terms, which are what we get if we, as Rosch did, average language-specific basic color terms (LSBCTs) across languages. And, we call this effect "focality effect of LSBCTs" if the calculation of color focality is based on the LSBCTs of a specific language. Because Rosch used the UBCTs to determine the focality of her test colors, the focality effect reported in her study was the focality effect of UBCTs.

Debi Roberson et al.(2005)[3]'s study marks a turning point. Although using Rosch's experiment design, they failed to find the same effect on a group of Himba-speaking subjects. When they noticed that some of the internominal colors for UBCTs were within the focal color range for the LSBCTs of the Himba language, they compared the colors that were focal only in English, those focal in both languages, and those focal only in Himba, in terms of STM performance. The results showed that while there was no difference in STM performance between the colors that were focal only in Himba and those focal in both languages, the colors that were focal only in Himba were better recognized than those focal only in English. A re-analysis of a data set collected from a group of Berinmo-speaking subjects using the same method obtained a similar result. In view of the absence of the focality effect of UBCTs in Himba, and the discovery of the focality effect of LSBCTs in Himba and Berinmo, Roberson et al. argue that what is universal is the focality effect of LSBCTs, rather than the focality effect of UBCTs as proposed by Rosch, thus casting doubt on the validity of Rosch's hypothesis.

In our study, we try to further evaluate the universality of the focality effect of LSBCTs through a psychological experiment to see whether this effect can be detected in the Japanese language. And, owing to the similarity between the LSBCTs of the Japanese language and the UBCTs in coverage distribution, which we discovered during the process of data analysis, we also employ our experiment data to probe into the existence of the focality effect of UBCTs.

2. Experiment Settings

Thirteen subjects (six males and seven females, ages: 32.69 ± 12.78, native Japanese speakers, no art experience) participated in our experiment. All of them passed the Ishihara Color Vision Test (38 plates, International Edition), thus being considered as normal in color vision.

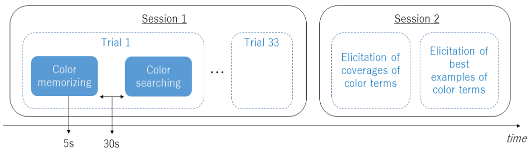

Fig. 1 The experiment procedure.

The whole experiment, which was conducted in Japanese, had two sessions (procedure shown in Figure 1). The first session, which had a similar procedure as Rosch's experiment, aimed to test the subjects' STM performance for colors. It consisted of 33 trials. In each trial, a color chip, mounted on the white surface of a cardboard (5.0 cm*5.0 cm), was presented to the subject for five seconds and then retrieved by the experimenter. After a 30-second interval, the subject was asked to find the color in a color array, which was mounted on the white surface of a cardboard (58.5 cm*28.5 cm), by writing on an answer sheet the index of the color chip that he/she thought was the right answer. There was no conversation between the experimenter and the subject during the interval. The color array was the one developed by Rosch, which had the layout shown in Figure 2. The colors tested in this session (called "test colors" in the following) are those within the bold-line-surrounded area in Figure 2. When an answer sheet was completed, the experimenter retrieved it to prevent the subject from referring to the previous answers in the following trials. Each test color was tested at least once to each subject, and for each subject the order of color-testing was randomly determined. For each subject there were three repeated trials, which were intended to prevent the subject from using the strategy of excluding the already tested colors. Before the formal experiment started, a two-trial training using a different set of test colors was carried out. All the color chips were picked from Munsell Book of Color (Glossy Edition), like those in the aforementioned pre-

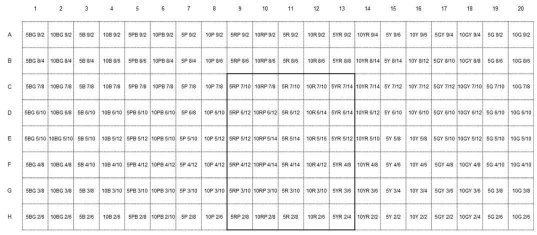

Fig. 2 The layout of the color array.

vious studies.

The second session is targeted to elicit the coverages of the six color categories corresponding to the six basic color terms *Akairo (Red)*, *Pinkuiro (Pink)*, *kiiro (Yellow)*, *Orenjiiro (Orange)*, *Chairo (Brown)* and *Murasakiiro (Purple)*[4], and to obtain the focality of each test color using a modified version of Berlin and Kay[1]'s method. Firstly, the subjects were required to write on six answer sheets, one for each basic color term, all the colors which they thought could be named by each of the color terms. When a subject completed an answer sheet, the experimenter retrieved the answer sheet before giving him/her the next one to prevent the interference of past answers. The six answer sheets were provided to each subject in a random order. Then, the subjects were asked to write on an answer sheet the indexes of the colors that they thought were the best examples of each of the six basic color terms. For each basic color term, reporting more than one colors was allowed, but the subjects were instructed to narrow down their choices as they could. The introduction of the second session was made after the end of the first session, so during the first session the subjects did not know that this experiment was about color categories or color terms.

The experiment was performed indoors with fluorescent lighting (type: National FHF 32EX-N-H; daylight color; color temperature: 5000K), which was close to the CIE D50 standard illuminant. The experimenter and the subject sat opposite at a table on which the stimuli were presented. The distance between the stimuli and the subject's eyes was fixed at 50 cm. A cardboard separating the two persons was set on the middle of the table, making the subject unable to see the experimenter's face when observing the stimuli, waiting at the time interval, or filling the answer sheets.

3. Data Analysis for Investigating the Focality Effect of LSBCTs

3.1 Coverages of Basic Color Terms and Quantification of Focality

Using the data obtained from the second session of our experiment, we specified the coverages of the color categories corresponding to the six basic color terms, and defined the focality of the test colors in a quantitative way.

First, we introduce Red Index, Pink Index, Yellow Index, Orange Index, Brown Index or Purple Index, which measure how often a color has been named using each of the basic color terms. The Red Index of a color is defined as the percentage of the subjects who named the color as red, and the other five indexes are similarly defined. Then, we define the Overall Index (OI) of a color as the largest of the six indexes of the color. Figure 3 shows the distribution of the non-zero OIs and the partition of the six color categories. We classify a color into the color category Red if its OI is its Red Index, or the color category Pink if its OI is its Pink Index, and so forth.

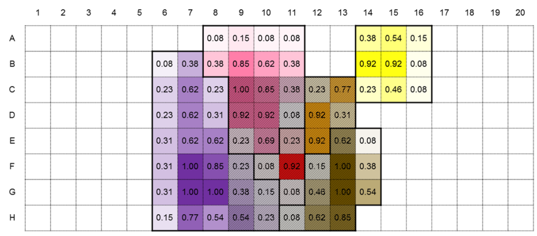

Fig. 3 The distribution of the Overall Indexes of the test colors (colors within the area covered by thin diagonal stripes) and other relevant colors, and the partition of the six basic color categories Red, Pink, Yellow, Orange, Brown and Purple. Color depth corresponds to the magnitudes of Overall Indexes.

Because the colors having the largest OIs and the colors most frequently selected as the best examples (shown in Figure 4) overlapped greatly in location, we think that the OI of a color probably represents the goodness of the color as a typical example of its category. Thus, we quantitatively define the focality score (FS) of a color as its OI value.

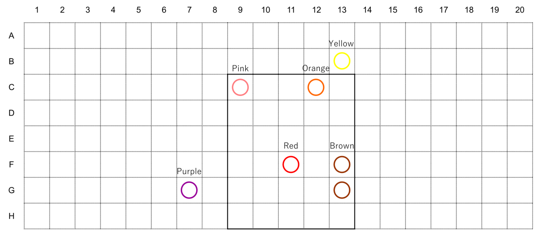

Fig. 4 The colors most frequently selected as the best examples of the color terms *Red*, *Pink*, *Yellow*, *Orange*, *Brown* and *Purple*.

3.2 Relationships Between STM Performance and Focality

We adopt memory accuracy score (MAS) and error distance score (EDS) as two measures of STM performance, like Rosch's and Roberson et al.(2005)'s studies. Because the variable FS is continuous in our study, rather than categorical as in Rosch's and Roberson et al.(2005)'s studies, we think it is best to deem the variables MAS and EDS as continuous. Thus, we define the MAS of a test color as the percentage of the trials where the subjects correctly recognized the color. And we define the EDS of a test color as the mean of the color differences between the test color and the colors mistaken for the test color in the incorrect recognition trials. Color difference is defined as Euclidean distance in the CIE L*a*b* color space throughout this study. So, before calculating color distances, we transformed the Munsell coordinates of the relevant colors into firstly the CIE xyY coordinates (using the O.S.A.-developed conversion tables[5]), then the XYZ coordinates, and finally the L*a*b* coordinates.

To look into whether STM performance for colors correlates with color focality in the Japanese language, we calculated the Person's correlation coefficient between the FSs and the MASs of the test colors, as well as that between the FSs and the EDSs of these colors. No statistically significant correlation was found (FS and MAS: Pearson's correlation coefficient = 0.174, $P = 0.356$, scatter plot shown in Figure 5; FS and EDS: Pearson's correlation coefficient = -0.009, $P = 0.964$, scatter plot shown in Figure 6). It means that, contrary to what Roberson et al.(2005)'s study implies, we failed to find a focality effect of LSBCTs in the Japanese language.

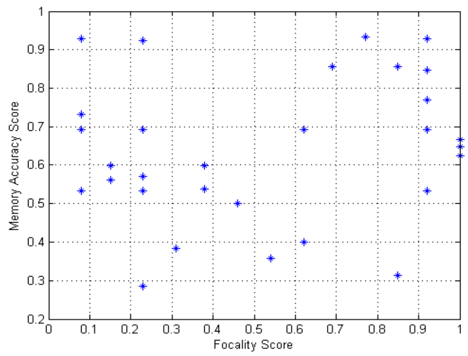

Fig. 5 The relationship between the focality scores and the memory accuracy scores of the test colors.

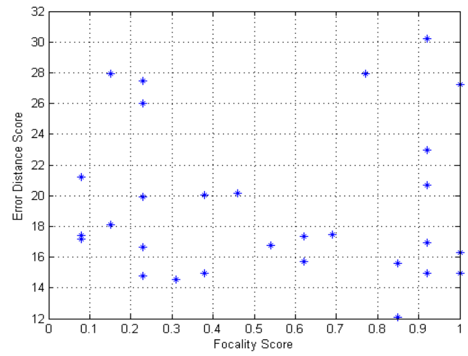

Fig. 6 The relationship between the focality scores and the error distance scores of the test colors.

4. Data Analysis for Investigating the Focality Effect of UBCTs

4.1 Overlapping of the LSBCTs and the UBCTs in the Japanese Language

During data analysis, we noticed that on the color array the locations of the colors that have the highest FSs for the the LSBCTs of the Japanese language are nearly the same as the locations of the focal colors for the UBCTs[1], which suggests a large overlapping of these two sets of color terms in coverage distribution. We then considered that even if the focality effect of LSBCTs does not exist in the Japaese language, if the focality effect of UBCTs exists in this language, we should still have been able to detect the correlation between color focality and the STM performance

[1] To see the locations of the focal colors of the UBCTs in Rosch's color array, please refer to Rosch's paper "Universals in color naming and memory"[2].

for colors. Not founding this correlation in our experiment indicates that the focality effect of UBCTs does not exist in the Japanese language, which runs counter to what Rosch's study suggests.

In order to figure out the cause of the discrepancy between the result of our experiment and that of Rosch's, we searched the literature and found two studies, one by John Lucy and Richard Shweder[6] and another by Debi Roberson et al.(2000)[7], both of which opposed Rosch's proposition. Lucy and Shweder raised the concern that because the discriminability of Rosch's focal colors was higher than that of the non-focal ones in her color array, it was possible that what actually determined the STM performance for her test colors was not their focality, but their discriminability. Here, discriminability of a color means how easily we can discriminate a color from its surrounding colors in the color array. Through discarding some color chips and then randomly shifting the locations of the remaining color chips, Lucy and Shweder changed the color chips which were adjacent to the test colors, and thereby the discriminability of the test colors. Using this new array, they could not detect the focality effect of UBCTs on a group of English-speaking subjects, despite using the same experimental paradigm as Rosch's study. Roberson et al.(2000)'s study, through a successful replication of Lucy and Shweder's experiment results on a group of English-speaking subjects and a group of Berinmo-speaking subjects, supports the idea that it is the discriminability of Rosch's test colors that plays the chief role in determining the STM performance for the colors. In Roberson et al.(2000)'s first experiment, which used Rosch's color array, they obtained a similar result as Rosch's experiment especially in the case of English-speaking subjects. But, when they changed the array configuration to Lucy and Shweder's design and ran the experiment again, they detected the focality effect of UBCTs in neither subject group.

To see whether the discriminability of the test colors, as Lucy and Shweder's and Roberson et al.(2000)'s studies suggest, acts as the determining factor for the STM performance for the colors in our experiment, we firstly defined the discriminability scores of our test colors, and then investigated the relationships between discriminability score and the two indicators of STM performance, namely MAS and EDS.

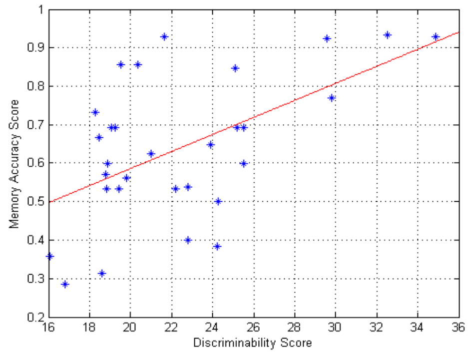

Fig. 7 The correlation between the discriminability scores and the memory accuracy scores of the test colors (Pearson's correlation coefficient = 0.545, P = 0.0018).

4.2 Relationship Between STM Performance and Discriminability

We defined the discriminability score (DS) of a color as the average of the color differences between this color and its eight adjacent colors.

The DSs of the test colors have a standard deviation of 4.61, which is a relatively large quantity as compared to the mean (22.437), and relatively large statistically significant correlations are found between the DSs and MASs of the test colors (Pearson's correlation coefficient = 0.545, P = 0.0018, scatter plot shown in Figure 7) and between their DSs and EDSs (Pearson's correlation coefficient = 0.626, P = 0.0002, scatter plot shown in Figure 8). These results confirm the heterogeneity in discriminability over the color array and suggest the great impact of the discriminability of the test colors on the STM performance for these colors.

5. Discussion

5.1 Implication to the Focality Effect of LSBCTs

Our experiment did not find the focality effect of LSBCTs in the Japanese language, thus casting doubt on the universality of this effect which was proposed

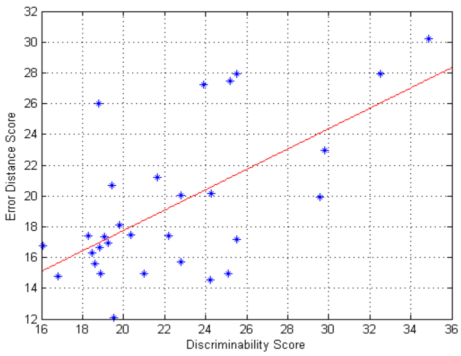

Fig. 8 The correlation between the discriminability scores and the error distance scores of the test colors (Pearson's correlation coefficient = 0.626, P = 0.0002).

by Roberson et al.(2005). In their study, they detected this effect in both the Himba-speaking subject group and the Berinmo-speaking subject group. Because both of the two languages differ from each other and from English in focality distribution[2], it is reasonable to rule out the discriminability of the test colors as the cause of their superiority in STM performance.

On the other hand, like the Japanese language, the LSBCTs of the English language are also similar to the UBCTs in focal color location over Rosch's color array[3][7], which suggests that the coverages of these two sets of color terms overlap greatly. Hence, Lucy and Shweder's and Roberson et al.(2000)'s experiment results, which were originally used to disconfirm the existence of the focality effect of UBCTs, can also be taken as counter-evidence to the existence of the focality effect of LSBCTs in the English language.

Considering these experiment results as a whole, we speculate that the focality effect of LSBCTs exists in some languages (e.g. Himba and Berinmo), but not other languages (e.g. Japanese and English), perhaps due to the differences in living environment between the people speaking the former languages and the people speaking the latter languages.

[2]Both the Himba language and the Berinmo language have five LSBCTs[3][7].

5.2 Implication to the Focality Effect of UBCTs

In Rosch's study, the same test colors and color array are presented to both the English-speaking subjects and the Dani-speaking subjects. As mentioned in Section 4.1, Lucy and Shweder's and Roberson et al.(2000)'s studies found that in Rosch's set of test colors the focal colors are more discriminable than the non-focal ones in average over the color array, and when this discriminability advantage of the focal colors is eliminated, there is no correlation between the focality of the colors and the STM performance for the colors. These results, as the researchers of the two studies argue, imply that the superiority of the focal colors in STM performance observed in Rosch's experiment actually resulted from the higher discriminability of the focal colors, rather than their higher degrees of focality as contended by Rosch. In other words, what seems to be the "focality effect" in Rosch's experiment is probably a "discriminability effect".

By calculating the DSs of Rosch's test colors using our definition and calculation method described in Section 4.2, which were based on strict modern colorimetry, we affirmed the higher discriminability of the focal colors as compared to the non-focal colors (the DSs of the focal colors: mean = 26.17, SD = 9.52; the DSs of the internominal colors: mean = 22.31, SD = 8.43; the DSs of the boundary colors: mean = 21.30, SD = 7.03). And, because the focality of the test colors used in our experiment does not correlate with their discriminability (FS and DS: Pearson's correlation coefficient = 0.125, P = 0.509), we are able to investigate respectively the two relationships, one between the focality of the colors and the STM performance for them, and another between their discriminability and the STM performance for them. Our experiment results indicate that the focality effect of UBCTs does not exist in the Japanese language, thus objecting to the universality or even the existence of this effect, and support the decisive role of color discriminability in determining the STM performance for colors. Obviously, these results tally well with the conclusions of Lucy and Shweder's and Roberson et al.(2000)'s studies.

6. Summary and Future Works

In this study, we conducted a psychological experiment to see whether the focality effect of LSBCTs exists in the Japanese language. Although this experiment is originally intended to evaluate the universality of the focality effect of LSBCTs, due to the great possibility that the LSBCTs of the Japanese language coincide with the UBCTs in focality distribution, our experiment data can also be employed to approach the focality effect of UBCTs. Our experiment found no correlation between the focality of the test colors and the STM performance for them, but detected a strong correlation between the discriminability of the colors and the STM performance for them.

Considered together with the results of Roberson et al.(2005)'s study as well as Lucy and Shweder's and Roberson et al.(2000)'s studies, our experiment results suggest that the strength of the focality effect of LSBCTs may vary greatly across languages. And, regarding the focality effect of UBCTs, like the studies by Lucy and Shweder and by Roberson et al.(2000), our experiment data argue against the universality or even the existence of this effect, and support the hypothesis that what appeared to be the "focality effect" in Rosch's study was in fact a "discriminability effect".

To see how robust our experiment results are, we plan to recruit more subjects and use test colors different from this time in our future studies. In addition, a closer scrutiny of Figure 5 makes us feel that there seems to be a weak tendency that STM performance improves as focality gets closer to its two extremes. Although this U-shaped relationship has not reached statistical significance, we think that in future studies it is important to consider the possibility of nonlinear relationships between color focality and STM performance for colors.

Of course, to eventually resolve the long-term debates on the existence, robustness or universality of either the focality effect of LSBCTs or the focality effect of UBCTs, much more studies investigating different languages and using different experimental methods are needed[3]. One thing to note is that, to prevent the factor of color discriminability from disturbing the analysis of the relationship between color focality and STM performance for colors, researchers should avoid using test colors whose focality correlate with their discriminability.

References

[1] Berlin, B., & Kay, P., (1991) "Basic color terms: Their universality and evolution", University of California Press.

[2] Heider, E. R., (1972) "Universals in color naming and memory", Journal of experimental psychology, Vol. 93, No. 1, pp. 10-20.

[3] Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R., (2005) "Color categories: Evidence for the cultural relativity hypothesis", Cognitive psychology, Vol. 50, No. 4, pp. 378-411.

[4] Uchikawa, K., & Boynton, R. M., (1987) "Categorical color perception of Japanese observers: Comparison with that of Americans", Vision Research, Vol. 27, No. 10, pp. 1825-1833.

[5] Newhall, S. M., Nickerson, D., & Judd, D. B., (1943) "Final report of the OSA subcommittee on the spacing of the munsell colors*", Journal of the Optical Society of America, Vol. 33, No. 7, pp. 385-418.

[6] Lucy, J. A., & Shweder, R. A., (1979) "Whorf and his critics: Linguistic and nonlinguistic influences on color memory", American Anthropologist, Vol. 81, No. 3, pp. 581-615.

[7] Roberson, D., Davies, I., & Davidoff, J., (2000) "Color categories are not universal: replications and new evidence from a stone-age culture", Journal of Experimental Psychology: General, Vol. 129, No. 3, pp. 369-398.

[8] Garro, L. C., (1986) "Language, memory, and focality: A reexamination", American Anthropologist, Vol. 88, No. 1, pp. 128-136.

[3] There is a study by Linda Garro[8] whose results conflict with Lucy and Shweder's and Roberson et al.(2000)'s for the reasons still unclear to us, and it is difficult for us to explain why the discriminability advantage of the focal colors did not lead to their superiority in STM performance in the case of Himba-speaking subjects as reported by Roberson et al.(2005)[3]. We hope that more empirical data obtained in future studies will help to answer these questions.