How do children construct the color lexicon? : Restructuring the domain as a connected system

Noburo Saji (saji@kamakura-u.ac.jp)
Department of Child Psychology, Kamakura Women's University, 6-1-3 Ofuna,
Kamakura, Kanagawa 2570056 Japan

Michiko Asano (asano@rikkyo.ac.jp)
College of Contemporary Psychology, Rikkyo University, 1-2-26 Kitano,
Niiza, Saitama 3528558 Japan

Midori Oishi (rosecoloreddays@gmail.com) and Mutsumi Imai (imai@sfc.keio.ac.jp)
Faculty of Environment and Information Studies, Keio University, 5322 Endo,
Fujisawa, Kanagawa 252-0882 Japan

Abstract

The present study investigated how children learn the meanings of basic color words and are immersed into the language-specific system of the color lexicon. The study examined how children discover the boundaries of color names by having 3-, 4-, and 5-year-old children produce names for 93 color patches. We found that even 3-year-olds children can map color words to its typical referents. At the same time, they struggle to delineate the boundaries between neighboring color words. The results indicated that, in learning color words, children continuously restructure the entire semantic domain by discovering and adjusting the linguistic boundaries between the neighboring words.

Keywords: lexical acquisition; color word; reorganization process of word meaning

Introduction

The meaning of a word is not determined in isolation from other words: the boundary of the category denoted by the word is determined in relation to other words belonging to the same semantic domain. For example, the meanings of the color words such as "red", "orange", "yellow", "green", "blue" and "purple" are established based on the knowledge of how the meaning of each word differs from the meanings of the others. This means that, to acquire adult-like word meanings, children need to learn a cluster of words in the same semantic domain and delineate the boundaries among them. Previous studies have documented that it requires many years to attain adult-like representation of the entire semantic domain. Ameel, Malt, & Storms (2008) studied how Dutch-speaking children aged 5-14 years and adults named various kinds of containers, and compared their naming patterns. Saji et al. (2011) examined the process in which Chinese children acquire representation of 13 Chinese verbs that all denote actions that are named as "carrying/holding" in English. Both studies found that children's pattern of labeling objects or actions evolved gradually over years, during which the semantic domain is repeatedly restructured.

Thus, to acquire the adult-like representation of the meaning of a word, children need to know how the word differs from other words that surround it. This is particularly critical for color words, as compared to object names, because the continuous visible color spectrum does not have natural partitions. Although there has been debates as to whether universal focal colors exist (Davidoff, 2001; Regier, Kay, & Cook, 2005), boundaries for referents of color words are hardly available in the environment. As a consequence, languages differ widely in the way in which they divide the continuous visible spectrum by color names (e.g., Berlin & Kay, 1969; Cook, Kay, & Regier, 2005). Even among languages that seem to have highly comparable color vocabulary withthe same number of words and the corresponding color categories, the boundaries can still vary considerably. For example, although English and Japanese both have "brown" (*chairo*) and "orange" (*orenjiiro*), the range covered by the English word "orange" is much wider than *orenjiiro* in Japanese. Consequently, the range covered by *chairo* is substantially broader than the range English word "brown" refers to. Given this, children cannot learn a color name from individual word-to-world mappings alone; they have to discover the conventional manner in which the perceptually continuous color spectrum is divided by a set of words in the particular language they are learning. Here, however, if the meaning of a color name cannot be learned without knowing the boundary of the neighboring words, how can children who know few color words ever get off from the merry-go-round of not being able to learn a new color word?

Difficulty in learning color words has long been noted by developmental psychologists. Bornstein (1985a; 1985b) suggested that color words were not reliably acquired until around four to seven years old, which is much slower than object names or names of other semantic properties such as shape. Developmental psychologists have proposed that the color names are difficult because the concept of color is difficult to understand (e.g., Kowalski & Zimiles, 2006). To understand the conventional meaning of a color term, children must understand that the term refers to a property that is abstracted out from diverse kinds of objects that differ in other properties such as shape, size, and functions.

Wagner, Dobkins, & Barner (2013) have proposed a different account. They argue that children have abstracted the concept of color by the time they begin using color terms, but they have difficulty in discovering how color boundaries are marked by the specific language they are learning. To establish this argument, these authors asked 3-year-olds to label 11 pieces of colored pasteboard. By and large, children's performance did not conform to adults' convention, but it was not random, either: The children consistently overextended the color words to proximal hues (see also Braisby & Dockrell, 1999). Although Wagner et al.'s work shed new light on the process of color name acquisition, it is not yet known how children learn to carve up the entire visible color spectrum by a set of words according to the convention of the child's ambient language.

The Present Study

In this study, we attempt to uncover how children's initial representation of the color lexicon is like, and how it evolves to the mature representation possessed by adults. Our approach to these questions is different from Wagner et al.'s study in one important respect. Wagner et al. (2013) tested children's knowledge of color words using only typical referents (i.e., focal colors). In the everyday situations, however, children see a wide range of colors, and have to learn how to label them. The real boundaries of color categories should exist around these non-focal colors rather than focal colors; to be able to map a color word to its typical referent does not guarantee that the child understand the correct (adult-like) range of colors the word refers to.

To investigate how children divide the entire spectrums into color name categories, we used a total of 93 color patches that covers the entire spectrum of visible colors, including not only typical referents but also those in the periphery of the 11 basic color names. We asked Japanese-speaking children of three age groups (3-, 4-, and 5-year-olds) and adults to label the 93 color patches. Japanese is considered to have 11 basic color words: *shiro* ('white'), *kuro* ('black'), *haiiro* ('gray'), *aka* ('red'), *kiiro* ('yellow'), *midori* ('green'), *ao* ('blue'), *chairo* ('brown'), *orenjiiro* ('orange'), *pinku* ('pink'), *murasaki* ('purple') (Uchikawa & Boynton, 1987). We attempted to capture the process of color word learning in two steps. We first examined whether Japanese-speaking children can map the basic color words to their typical referents, as was done in earlier studies. We then explored how children establish lexical boundaries among neighboring color words. These two steps allowed us to see the developmental process through which children gradually acquire both the center and boundaries of color word meanings.

Experiment

Method

Participants.

A total of 72 native Japanese speaking children and adults participated in the experiment. The production data was collected from 20 3-year-olds, 18 4-year-olds, 19 5-year-olds, and 15 adults. Children were recruited from several preschools in located in the Tokyo Metropolitan area. Adult subjects were undergraduate and graduate students at Keio University.

Stimuli.

Ninety-three colors were selected from Practical Color Co-ordinate System (PCCS) developed by Japan Color Research Institute, which is widely used for organizing color for industry and education in Japan. PCCS consist of 14 "tone" categories, each of which has 24 hues. Tone is a compound concept of lightness and metric chroma (see http://www.diccolor.com/knowledge/images/pccs06_l.jpg; also see Nayatani, 2003).

We used colors from seven tones ("light", "bright", "soft", "vivid", "dull", "deep", and "dark") out of the 14 tones, which varied in lightness and chroma, so that the stimulus colors covered the entire color spectrum. Only half of the 24 hues (with every other even number) in each tone were used to reduce the number of stimuli. In addition to these 84 chromatic colors (7 tones x 12 hues), we included nine achromatic colors (black, white, and five different grays varied in lightness). Each of the 93 colors was presented as a 2.5cm square patch, which was created by cutting colored pasteboard distributed by Japan Color Research Institute. See Figure 1 for the complete stimulus set.

To identify the 11 typical referents for the basic color words in the current stimuli set, we conducted a pre-study with 20 adults who did not participate in the main experiment. In the pre-study, we presented the 93 colors at once and asked them to choose colors which they thought were the best referent for each of the 11 basic color words. We then determined the most frequently selected color as the typical referent of each basic color word for Japanese speakers (see Table 1 for 11 selected colors selected).

Procedure.

Children received warm-up trials before the main test. Children were presented with six pictures of three different cats, two different dogs and one rabbit, and were asked to name each picture. This process also allowed us to check whether the child understood that they could say the same name more than once across different trials.

In the main experiment, participants were presented with

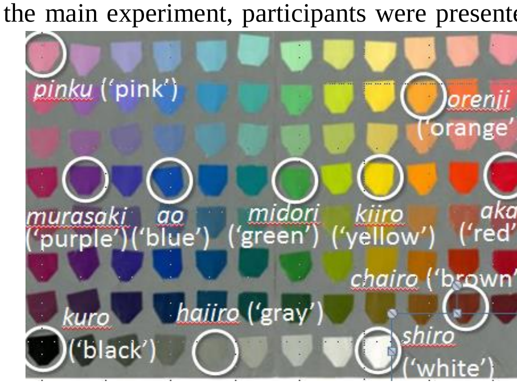


Figure 1: The 93 color patches for stimuli. Those with white circles are the typical referents for the 11 basic color names, which were determined by a pre-test with adults.

the 93 color patches one by one in a random order, and were asked "*kono* (this) *iro* (color) *wa* (particle) *nani* (what) *iro* (color)"? ("what color is this?") by the experimenter. Each color patch was presented only once. Participants observed the color patches under a standardized lighting condition that simulated natural daylight (D55) by using Solax XC-100AF (Seric Ltd.) on a gray background.

Results

Analysis 1: Were children able to produce appropriate basic color terms to their typical referents?

We first counted the number of word types produced by each participant. Compound color names such as *aka-cha* ('red-brown'), or loan words as *howaito* ('white') were counted as different word types from the basic color words. Adults on average produced 19.4 different color words, including ones other than 11 basic color names, such as *giniro* ('silver'), *hadairo* ('skin color') or *shuiro* ('vermilion'). The mean numbers of produced word types by children were 9.5, 12.1, and 13.6 for 3-, 4-, 5-year-olds, respectively. Four-year-olds produced more word types than 3-year-olds (p<.02, Bonferroni corrected), but there was no difference between 4- and 5-year-olds. Although children produced significantly less color terms than adults (ps <.01, Bonferroni corrected), they "knew" most of the 11 basic color names by the age of 3 years.

Table.1: The proportion of children who named colors with appropriate (i.e., adult-like) basic color words.

Basic color word	Sample	3-years	4-years	5-years
shiro ('white')		.80	.89	1.00
kuro ('black')		.80	.78	1.00
aka ('red')		.95	.61	1.00
kiiro ('yellow')		.80	.78	.95
midori ('green')		.65	.56	.84
ao ('blue')		.70	.67	.84
chairo ('brown')		.50	.78	.74
orenjiiro ('orange')		.40	.78	.63
pinku ('pink')		.80	.89	.89
murasaki ('purple')		.55	.78	.79
haiiro ('gray')		.00	.11	.05

Next we examined whether children could apply the 11 basic color words for the corresponding typical referents. Table 1 shows the proportion of children who produced the appropriate basic color words to the correct referents determined by the pre-study. In all age groups, children correctly applied most of the basic color words to the typical referents except *haiiro* ('gray') significantly above chance (binomial test, ps< .01). Consistent with the previous studies, children had a stronger grasp of *kuro*('black'), *shiro*('white'), *pinku*('pink') and *aka*('red') than the other words as *murasaki*('purple') or *orengiiro*('orange'), while they rarely produced *haiiro* in the experiment (See Roberson, Davidoff, Davies & Shapiro, 2004; Pitchford & Mullen, 2002 for similar results).

Analysis 2: How did children's pattern of color naming converge with the adults' pattern?

Analysis 1 showed that even 3-year-old children knew the central meaning (i.e. typical referent) of the 11 basic color words in an adult-like manner. However, how did they delineate the boundaries of the meanings? To examine the question, we adopted the Multi-Dimensional Scaling (MDS) solutions as in Malt & Sloman (1999). MDS provides a geometrical representation of patterns of similarity on dimensions that are extracted to maximize goodness of fit in such a way that inter-point distances on the multidimensional space correspond to dissimilarities between objects. We first created a similarity matrix for each age group. In each matrix, there were 93 rows and 93 columns each representing the 93 stimulus colors. Each cell contained the number of times the given two color patches were named with the same color word (Malt, Ameel, Imai, Gennari, Saji & Majid, 2014).

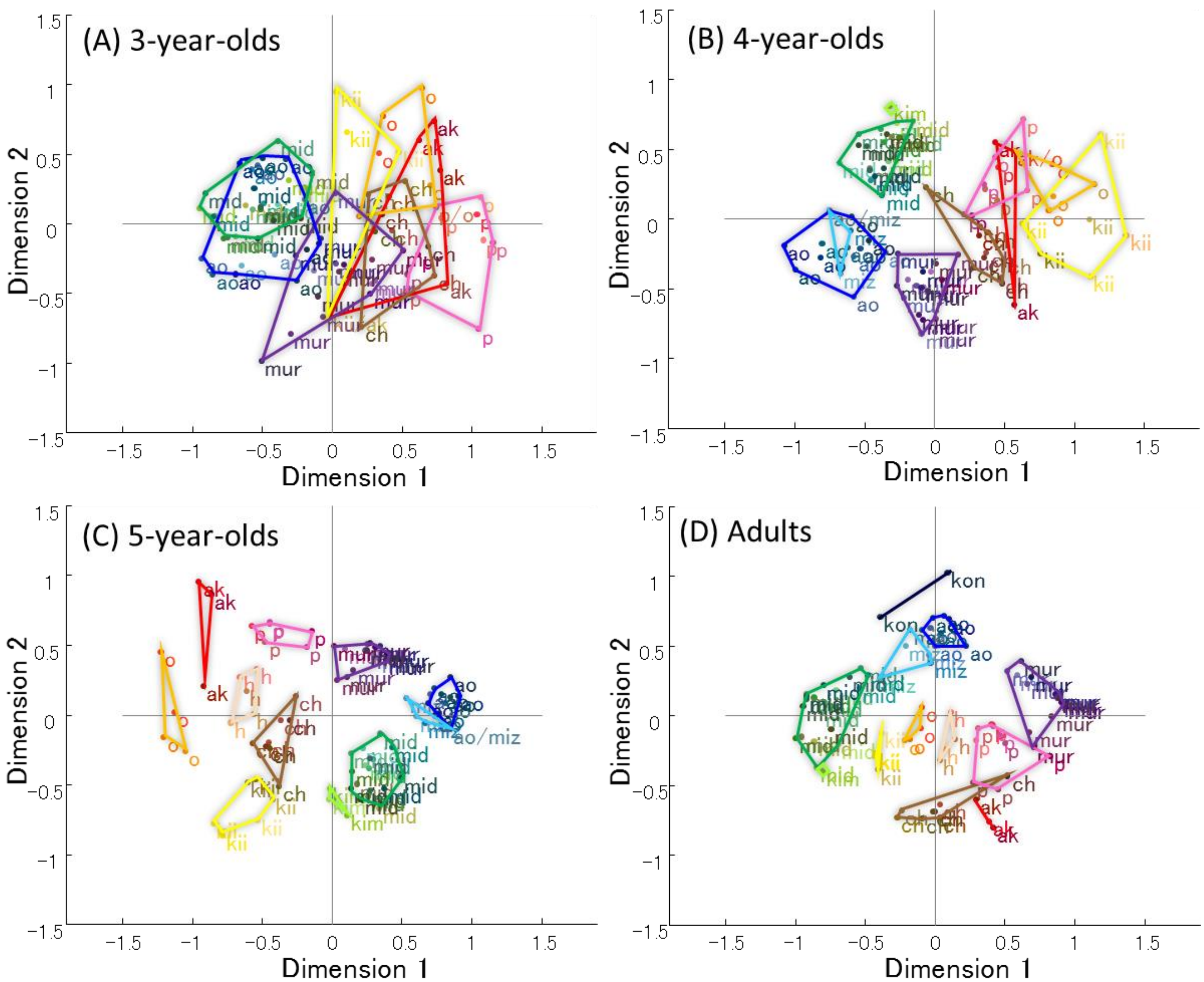
We first conducted the MDS analysis separately for the four age groups using the 93X93 matrices. The detected dimensions mainly divided the chromatic and achromatic color patches only. This means that the participants rarely confused their naming for achromatic color patches with that for chromatic color patches, and this distinction was over-weighted. Because our main interest here is to examine how children and adults categorize continuous color spectrum by names, we carried out a second MDS analysis with the naming data of 84 chromatic patches, again separately for the four age groups (Figure 2). We adopted 2-dimensional solutions for each age group, because the stress values were considered to be acceptable (s = .20, .17, .20, and .07 for 3-, 4-, 5-year-olds and adults, respectively). Each point in Figure 2 represents each color patch, and the distances among every two points reflect the similarity of the two colors in terms of naming. The label for each point shows the most dominantly produced color name for the patch. For ease of viewing, the points and labels were shown in the color of the original stimuli patches. Color patches labeled with the same dominant color names are grouped by solid colored lines. The configurations of the data points were largely different across the four age groups. To capture this trend quantitatively, we calculated the Euclidian distances between all possible pairs of the data points (i.e., color patches) in each MDS space, and calculated the correlation between the adult group and each of the three child groups. The correlation values were .31, .43 and .53 for 3-, 4-, and 5-year-olds, respectively. This indicates that children's color naming behavior steadily develops but only slowly, and it takes a long time for children to become able to name colors in the same way adults do.

Analysis 3: How did children discover the boundaries of different color words?

We next examined how children find boundaries among neighboring color words. In the adult group (Figure 2D), the colors with the same dominant names were grouped in cohesive clusters, and colors with different dominant names were separated from one another. This suggests that the adults named the patches highly consistently across individuals, and boundaries among different color names were clearly delineated. This was not the case with children, 3-year-olds in particular. In the 3-year-olds' MDS configuration (Figure 2A), the areas covered by different dominant names heavily overlapped one another, indicating that 3-year-olds applied each color name to color patches that were outside the range in the adults' convention and that the pattern of naming was highly inconsistent across individuals. It should be noted, however, that they tended to overextend the words to neighboring colors, supporting Wagner et al (2013)'s view. The 4- and 5-year-olds were in intermediate stages between the 3-years-olds and the adults (Figure 2B and 2C). To evaluate this observation quantitatively, we quantified how exclusively each of the dominant chromatic color words was used for color patches against other color words, as given below:

$$Exc(C1, C2) = \frac{D\ (CentC1, CentC2)}{\sum_1^k D(CentC1, C1_k)/k + \sum_1^l D(CentC2, C2_l)/l}.$$

C1 and C*2* represent the two clusters to be compared, and *k* and *l* are the numbers of the points in the clusters. The numerator calculates the distance between the centroid of the two clusters; the denominator takes the summation of the two average distances between the centroid and each point in the two clusters. So the more the two clusters overlap, the smaller the ratio becomes; in contrast, the more



ak = *aka* ('red'), ao = *ao* ('blue'), ch = *chairo* ('brown'), h = *hadairo* ('skin color'), kii = *kiiro* ('yellow'), kim = *kimidori* ('yellowish green'), mid = *midori* ('green'), miz = *mizuiro* ('light blue'), mur = *murasaki* ('purple'), o = *orenjiiro* ('orange'), p = *pinku* ('pink')

Figure 2: Multi-Dimensional Scaling solutions for children and adults

the two clusters were distinguished, the larger the value becomes. We averaged the degree of boundary (non-) overlap among the dominant color names for each age group (Figure 3). The values were 1.02, 1.92, 3.16 and 2.86 for 3-, 4-, 5-year-olds and adults, respectively. The averages are significantly higher in 5-year-olds and adults than in 3- and 4-year-olds (*ps* <.01), suggesting that boundary overlap was more prominent in 3- and 4-year-olds than in 5-year-olds.

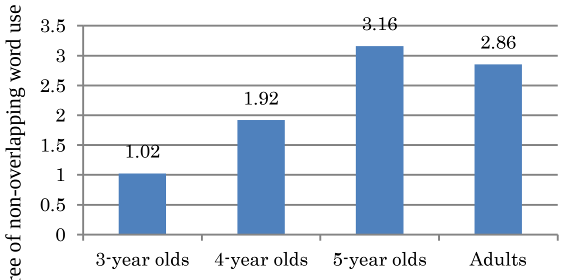


Figure 3: The averages of the degree of non-overlapping

Table 2 and 3 show the degree of boundary (non-)overlap among eight color names for adults and 3-year-olds, calculated by the equation above. The distinction between *aka* ('red') and *ao* ('blue'), and that between *kiiro* ('yellow') and *ao* ('*blue'*) were both evident in adults, as seen in Table 2. However, these distinctions were not so clear in 3-year-olds; they made relatively clear distinctions between *midor*i ('green') vs. *chairo* ('brown') or *ao* ('blue') vs. *pinku* ('pink'), but the distinctions between *midori* ('green') vs. *ao* ('blue') and *aka* ('red') and *chairo* ('brown') were very blurred. The matrices of 3-year-olds and adults are not correlated (*r*= .04), indicating that the boundaries of words that are most clearly delineated in adults were not always clear for children. It is also important to note that, in the 3-year-olds MDS configuration, while the boundary between *ao* ('blue') and *midori* ('green') were largely overlapping (the degree of boundary overlap between *ao* and *midori* was .12), these two words were distinctively separated from other basic colors (the degree of boundary overlap between the two colors and other colors 1.2 on average. See also the blue and green colored lines in Figure 2). This may suggest that children first form separate islands of blue/green vs. others, relying on salient perceptual distinction between cool and warm colors (see Kay & Maffi, 1999).

Table.2: The degree of boundary (non-)overlap among the dominant names in adults.

	orenjiiro (orange)	*pinku* (pink)	*midori* (green)	*kiiro* (yellow)	*murasaki* (purple)	*ao* (blue)	*aka* (red)
pink	2.01						
midori	2.07	3.00					
kiiro	1.82	3.07	1.68				
murasaki	3.53	1.33	4.01	5.06			
ao	3.55	3.01	3.06	5.37	3.15		
aka	4.20	1.67	4.44	5.28	3.53	6.76	
chairo (brown)	2.18	1.66	2.73	2.30	3.24	4.56	1.39

Table.3: The degree of boundary (non-)overlap among the dominant names in 3-years-old.

	orangeiro (orange)	*pinku* (pink)	*midori* (green)	*kiiro* (yellow)	*murasaki* (purple)	*ao* (blue)	*aka* (red)
pink	1.07						
midori	1.81	2.48					
kiiro	0.40	0.96	0.90				
murasaki	1.46	1.32	1.37	0.84			
ao	1.41	1.84	0.12	0.75	0.94		
aka	0.41	0.42	1.34	0.42	0.82	1.08	
chairo (brown)	0.86	0.68	1.87	0.59	0.84	1.35	0.27

Analysis 4: What guides the adjustment of color name boundaries?

Analysis 2 and 3 showed that an adult-like lexical system gradually emerges with development. What guides the delineation of the color name boundaries? It has been shown that children are able to distinguish different colors perceptually long before they initiate word learning (Bornstein, Kessen, & Weiskopf, 1976). However, it is not known how their non-linguistic perception of color words affects acquisition of color names. To address this issue, we correlated the Euclidean distance between every two color patches in the naming MDS space with the Euclidean distance between the corresponding two color patches in the CIE L*a*b* color space, which was designed such that mathematical differences in all color ranges correspond to perceived color differences (Kaatsch, Stadler, & Nietert, 1993). The correlation is high if perceptually similar colors are named similarly.

The correlation coefficients were .43, .56, .75 and .57 for 3-, 4-, 5-year-old children and adults, respectively, which were significantly different between each pair of age groups (*z* =13.6 *ps* <.001). The result showed that the impact of non-linguistic perceptual similarity/distance among color patches on the naming pattern increases from three to five years of age and then declines as they further advances in lexical development. The results may suggest an interesting possibility: Children first map each of the basic words to the color construed as the most typical referent by the adults. They then attempt to find linguistic boundaries among these words by relying on perceptual distance of the given two typical colors: after they tentatively sort out the boundaries by perceptual similarity, they continue to restructure the semantic domain, and are slowly immersed into the adults' way of division of the spectrum, which is to some degree arbitrarily made due to various coincidental forces through the history of the language (cf. Ameel et al., 2008).

Discussion

The present study experimentally examined how children learn the meanings of basic color words and how they are immersed into the language-specific system of the color lexicon. Consistent with previous studies (e.g., Wagner et al., 2013), we found that 3-years old children produced most of the basic color words correctly when the color patch was a typical referent for the name. Furthermore, our findings clearly showed that acquisition of color words continues after 3 years old; even 4- and 5-year-old children still struggled to find adult-like boundaries of color names (see Ameel, et al., 2008 and Ameel, Malt & Storms, 2014 for similar discussion).

Our results uncovered the developmental trajectory of this process, and what underlies it. By 3 years of age, children "have learned" most basic color names in the sense that they know these words are names of different colors. In their attempt to organize the color semantic domain, children first attempt to locate each color word roughly along the undivided spectrum. Three-year olds are able to map most basic color names to their typical colors, but categories are largely overlapping and their boundaries are overextended. With development, the overextended category boundaries are gradually narrowed down as the boundaries among neighboring words are delineated. Importantly, we observed that the process of boundary delineation occurs simultaneously across different color words, which suggests that development of color words takes place as a consequence of continuous restructuring of the entire color lexicon rather than a local adjustment of a single word.

Acknowledgments

This research was supported by MEXT KAKENHI to Imai (Grant-in-Aid for Scientific Research on Innovative Areas #23120003) and Saji (Grant-in-Aid for Young Scientists (B) #26870573)

References

Ameel, E., Malt, B. C., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle. *Journal of Memory and Language, 58*, 262–285.

Ameel, E., Malt, B. C., & Storms, G. (2014). Steps along a Continuum of Word Knowledge: Later Lexical Development through the Lens of Receptive Judgments. *Language Learning and Development*, *10(3)*, 234–262.

Berlin, B. & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley & Los Angeles: University of California Press.

Bornstein, M. H. (1985a). Colour-name versus shape-name learning in young children. *Journal of Child Language 12*, 387–93.

Bornstein, M. H. (1985b). On the development of colour naming in young children: data and theory. *Brain and Language 26*, 72–93.

Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). Color vision and hue categorization in young human infants. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 115.

Braisby, N., & Dockrell, N. (1999). Why is colour naming difficult? *Journal of Child Language*, *26*, 23-47.

Cook, R. S., Kay, P. & Regier, T. (2005). The world color survey database: History and use. In Cohen, H. & Lefebvre, C. (Eds.), *Handbook of Categorisation in the Cognitive Sciences*. Amsterdam: Elsevier.

Davidoff, J. (2001). Language and perceptual categorisation. *Trends in cognitive sciences, 5*, 382-387.

Kaatsch, H.-J., Stadler, M., & Nietert, M. (1993). Photometric measurement of color changes in livor mortis as a function of pressure and time: Development of a computer-aided system for measuring pressure-induced blanching of livor mortis to estimate time of death. *International Journal of Legal Medicine, 106*, 91–97.

Kay, P., & Maffi, L. (1999). Color appearance and the emergence and evolution of basic color lexicons, *American Anthropologist*, *101*(4), 743-760.

Kowalski, K., & Zimiles, H. (2006). The relation between children's conceptual functioning with color and color term acquisition. *Journal of Experimental Child Psychology*, *94*, 301–32.

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, *40*(2), 230-262.

Nayatani, Y. (2003). Adequateness of a newly modified opponent-colors theory. *Color Research & Application*, *28*, 298-307.

Majid, A., Bositer, J.,S., & Bowerman, M. (2008). The cross-linguistic categorization of everyday events: A stdy of cutting and breaking, *Cognition*, *109*, 235-250.

Malt, B.C., Gennari, S., Imai, M., Ameel, E., Saji, N., & Majid, A.(2014). Human Locomotion in Languages: Constraints on Moving and Meaning.Journal of Memory and Language, *74*, 107-123.

Pitchford, N. J., & Mullen, K. T. (2002). Is the acquisition of basic colour terms in young children constrained? *Perception*, *31*, 1349–1370.

Regier, T., Kay, P., & Cook, R. S. (2005). Focal colors are universal after all, *Proceedings of the National Academy of Sciences, 102*, 8386-8391.

Roberson, D., Davidoff, J., Davies, I.R.L., & Shapiro, L.R. (2004). The development of color categories in two languages: A longitudinal study, *Journal of experimental Psychology: General*, *133*(4), 554-571.

Saji, N., Imai, M., Saalback, H., Zhang, Y., Shu, H., & Okada, H. (2011). Word learning does not end at fast mapping: Evolution of verb meanings through reorganization of an entire semantic domain. *Cognition*, *118*, 45-61.

Uchikawa, K., & Boynton, R. M. (1987). Categorical color perception of Japanese observers: Comparison with that of Americans. *Vision Research*, *27*, 1825-1833.

Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition, 127*, 307–317.