Typicality Knowledge and the Interpretation of Adjectives

Choonkyu Lee

Abstract In this paper, we discuss our experimental results involving color preference and yes/no categorization judgments that provide insight into the interpretation of color adjectives. We selected a set of object categories that show a consistent color typicality bias, and presented them with varying degrees of color manipulation in our experiments. In Experiment 1, Dutch speakers performed a forced-choice picture-phrase matching task. Between a photograph of an object in its typical color (e.g., a light green tomato) and another photograph in a focal color (a darker green tomato), participants showed a significantly higher proportion of preference for the typical, nonfocal color for categories with a color bias (e.g., tomato; 52%) than for categories without a bias (e.g., box; 36%). In Experiment 2, we conducted a categorization task in which participants judged whether an image was an example of the target adjective-noun combination or not, in yes-no format, for 14 adjective-noun combinations including color and other adjectives, such as pattern and material adjectives. When presented with nonfocal images, participants were much more likely to give ‘Yes’ responses for categories with a typicality bias (55%) than for those without a bias (27%), demonstrating an effect of world knowledge in yes-no categorization judgments as well.

1 Introduction

The domain of colors was discussed frequently in the early literature on concept composition (e.g., Smith et al. 1988; Medin and Shoben 1988), and also played a key role in the development of prototype theory (Heider 1972; Berlin and Kay 1969). It continues to be central in the literature on top-down influences on categorization (e.g., Mitterer et al. 2009). In the semantics literature, however, color adjectives have usually been considered intersective adjectives whose composition with nouns simply involves set intersection and whose truth-conditions are inde-

C. Lee (✉)

Department of Linguistics, University of Potsdam, Potsdam, Germany
e-mail: chlee@uni-potsdam.de

© The Author(s) 2017

J.A. Hampton and Y. Winter (eds.), *Compositionality and Concepts*
in *Linguistics and Psychology*, Language, Cognition, and Mind 3,
DOI 10.1007/978-3-319-45977-6_5

pendent of the context they appear in (e.g., Keenan and Faltz 1985; Chierchia and McConnell-Ginet 1996; Drašković et al. 2013).

Previous discussion of adjective-noun combinations involving color adjectives revealed some general issues in meaning composition. Smith et al. (1988) used examples such as *red apple* and *brown apple* to illustrate how their model of meaning composition, ‘selective modification,’ worked. In this model, a noun representation has multiple attributes, including its color (when it is relevant), and adjectival modification of the noun with a color term selectively influences the color attribute. A mental representation of apple thus includes relevant attributes such as its color, shape, texture, and freshness, each with default values set according to our world knowledge. In their example, Smith et al. (1988) assigned the values 25 for red, 5 for green, and 0 for brown in the color attribute of *apple* to model our expectation about different possibilities of the color of an apple. In this model, modifying the noun *apple* with a color adjective *red* or *brown* simply re-assigns the color-attribute values so that the corresponding color gets all the values (30 red, 0 green, 0 brown for *red apple*; and 0 red, 0 green, and 30 brown for *brown apple*), while also increasing the diagnosticity of the color attribute in similarity judgments. When Smith et al. (1988) tested their model’s predictions against participants’ typicality ratings of various items for concepts such as *red vegetable* and *red fruit*, they obtained correlations between 0.70 and 0.97, which were in a similar range to other adjective-noun combinations such as *round vegetable* and *round fruit*.

Medin and Shoben (1988), however, pointed out the limitations of Smith et al.’s (1988) selective modification model with an emphasis on correlated attributes in the world. In the *brown apple* example above, the brown color usually indicates that the apple is not fresh, and neglecting such world knowledge would be an oversimplification for a model of conceptual representation. Participants in Medin and Shoben’s (1988) typicality rating study demonstrated their awareness of correlated attributes, judging white clouds to be more harmless than gray clouds, and black-and-white TVs to be more likely to be small than color TVs. Moreover, their similarity judgments demonstrated that the internal structure of color space changes according to the noun context, with the color pair *white-gray* judged to be more similar than the pair *gray-black* in the context of *hair*, but the other way around in the context of *cloud* or *bear*.

1.1 Overextension and World Knowledge

Aside from the impact of world knowledge, Hampton (1996) showed non-Boolean processes in concept composition, using letters in different colors, but there is evidence in Smith and Osherson’s (1984) early data that suggests world knowledge also constrains these processes in concept composition. Hampton (1996) conducted categorization experiments on ambiguous-colored letter shapes, whose values on

the dimensions of color and shape similarity to the letters, A or H, varied along 11-point scales. In one experiment, participants gave a forced-choice response between the colors, blue and green, to judge the color of a letter shape, and later gave a yes-no categorization response for the same visual stimuli with regard to adjective-noun descriptions such as *Blue H*. Hampton (1996) observed that in about 10–20% of the trials, participants accepted descriptions such as *Blue H* after choosing the color *Green* over *Blue* for the same image, when the stimulus color was slightly closer to green but the stimulus shape was a very good H. Another experiment using the colors, orange and red, and yes-no questions such as *Is this Orange?* instead of a forced choice between two colors, showed a similar rate of overextension. The main issue in this study was not world knowledge, as there is presumably no clear color preference between blue versus green As and Hs, or between orange versus red As and Hs in our world knowledge. Hampton's (1996) results showed that categorization judgments that are negative for a single dimension (simply color) can be positive when the dimension is combined with another in which the stimulus has high 'goodness-of-membership,' revealing non-Boolean composition.

Smith and Osherson's (1984) typicality rating results (Experiment 2, p. 349) showed a similar 'relaxation' of standards in the interpretation of adjectival modifiers: In response to pictures that were intermediate between two adjective concepts (e.g., an ambiguous red-brown apple drawing), participants' typicality ratings of the drawing as an example of the adjective-noun description (*red apple* and *brown apple*) were significantly higher than those with just the adjective category (*red* and *brown*). Because Smith and Osherson used concepts for which we have stronger knowledge-based biases than colored letter shapes in Hampton's (1996) study, we can observe the impact of world knowledge on judgments involving complex concepts. When the color in the verbal description was an atypical color for the depicted category in the drawing and thus negatively diagnostic (e.g., *brown apple* and *red canary*), there was always reliable overextension in typicality ratings, namely, significantly higher ratings for the adjective-noun descriptions (*brown apple*, *red canary*) compared to just the color adjective (*brown*, *red*), regardless of the actual degree of match between the image and the description. In contrast, when the color of a description was a typical one for the noun category and thus positively diagnostic (*red apple* and *yellow canary*), there was not consistent overextension of a color judgment from simple descriptions (*red*, *yellow*) to complex ones (*red apple*, *yellow canary*). In fact, when the drawing was a poor match (e.g., a drawing of a brown apple) for the target descriptions (*red* and *red apple*), there was even a decrease in typicality ratings from the simple description with the adjective alone (*red*) to the complex one (*red apple*)—in other words, a brown apple was a worse example of *red apple* than of *red*, according to Smith and Osherson's (1984) participants. The fact that overextension for adjective-noun combinations depends on the diagnosticity of the particular adjective reflects our color typicality knowledge about the noun categories (apples and canaries).

1.2 Further Effects of World Knowledge Involving Color Terms

Later developments employing different methodologies such as corpus/dictionary analysis and color labeling experiments shed further light on the role of world knowledge on color-term usage. Steinvall (2002) analyzed color-adjective uses in the Bank of English corpus and color-adjective entries in the Oxford English Dictionary (1993), and made an important distinction between the ‘classifying’ and the ‘characterizing/descriptive’ functions of color adjectives. While the latter use focuses on the color of a specific instance in the referential setting, the former use, which Steinvall (2002) also called *type modification*, picks out a subtype of the noun category in question. For example, natural kinds (e.g., onions, undyed hair) usually exist only in a few colors, rather than in the full range of a color spectrum, and for these, Steinvall (2002) observed that basic color terms are predominantly used to classify the subtypes based on colors without necessarily being descriptively precise with regard to the actual referent object (e.g., *red onion* for a purple hue, since there are no other types of onions closer to a prototypical bright red). These results are consistent with Anishchanka, Speelman and Geeraerts’s (2014) results from an analysis of color-term usage in online marketing, in which the authors found hypernymous usage with a broad referential range of colors for basic color terms, and a much narrower referential range for non-basic color terms. The exact causal history of basic color terms—whether they arose due to the limited color types in frequently mentioned natural categories, or they just happened to be conveniently adequate for type classification—is, however, admittedly unclear (Steinvall 2002).

Aside from adjective-noun combinations, there are experimental studies involving single-word color terms in color labeling and categorization tasks that also demonstrate an effect of typicality biases due to the object category being asked about. Using hand-drawn images of typically orange (e.g., a carrot) and typically yellow objects (e.g., a banana), Mitterer and de Ruiter (2008) demonstrated that participants were more likely to label, for example, a carrot as “orange” and a banana as “yellow” even when these were presented in the same exact hue. A color typicality effect was also evident at the level of perception for the hue midway between orange and yellow: Participants who saw this ambiguous orange-yellow hue on a carrot first categorized the same color sock (i.e., an object with little intrinsic color bias) in a later task as “orange,” and those who saw this hue on a banana first categorized the same color sock in a later task as “yellow”.

In order to locate the cause of the color typicality effect precisely by teasing visual and declarative memory from life experience apart, Mitterer et al. (2009) picked traffic lights as their visual stimuli in a later experiment. Due to European Union regulations, EU citizens presumably share a common perceptual experience with regard to traffic lights, but some language groups nevertheless differ in their naming of the color of the middle light. Dutch speakers call the middle light *oranje* (‘orange’), whereas German speakers call it *gelb* (‘yellow’). Mitterer et al. (2009)

found that this difference in color-naming habits led to different color-labeling behavior in the experiment when the same ambiguous orange-yellow hue was presented on a traffic light image: When presented with the same exact hue on a middle traffic light, Dutch speakers were more likely to call it “orange,” while German speakers were more likely to call it “yellow,” reflecting their habit. These two language groups were, however, indistinguishable when the ambiguous hue was presented on other object categories, such as a carrot, a banana, and a sock, for which there is presumably no systematic difference in color-naming habits between the two groups. Mitterer et al. (2009) thus concluded that it is primarily declarative memory (i.e., everyday color-term usage), rather than visual memory from life experience, that gave rise to the world knowledge effect in color labeling and categorization.

In sum, there is reason to believe that our general knowledge of typical properties of and relations between objects plays an important role in using color adjectives for labeling different hues. In our study, we investigated whether our knowledge of typical object properties and relations in the world influences our discrete categorization judgments in the context of adjective-noun combinations as well. Specifically, we were interested in the competing factors of set intersection in concept composition and color typicality knowledge in complex concepts with a bias toward a nonfocal color (cf., Heider 1972; Regier et al. 2005, for the notion of a ‘focal’ color, which refers to the best representative of a color category, as widely recognized across different linguistic communities), such as *red hair* (whose typical red is not the focal, bright red) and *green tomato* (whose typical green is much lighter than the focal green). In a series of pilot studies, we first found a set of object categories whose typical colors are not focal colors in our commonsense knowledge (e.g., *red hair* and *green tomato*). Most previous studies involving color descriptions and visual stimuli presented a single image at a time on a given trial, in which the participant had to make a yes-no response with regard to a color description or a forced choice between two color descriptions for the better match. In our Experiment 1, we had participants make a forced choice between two images, one in a focal color and the other in a nonfocal color, for the better match to an adjective-noun description such as *green tomato*. We predicted that, compared to categories without a color typicality bias (e.g., boxes), categories with a typical nonfocal color bias in the real world such as tomatoes would lead to a higher proportion of responses toward the nonfocal-colored image, against the predictions of stricter accounts of concept composition which might accept only good examples of *green* as good examples of *green tomato*. For example, if participants treat the word meanings of *green* and *tomato* separately first and simply combine them in set intersection for the meaning of *green tomato*, they might prefer a focal-green tomato even though it looks artificial. Forced-choice preference data in Experiment 1 by itself, however, would not establish that people’s discrete categorization judgments with regard to an adjective-noun combination differ depending on the intrinsic color properties of the relevant object category. A simple *preference* for a focal-green box image over a nonfocal-green box image as an example of the description *green box* tells us nothing about whether the participant would

categorize the dispreferred nonfocal image (or even the focal image) as an example of *green box* or not in yes-no format. In order to get at people's truth-judgments directly, in Experiment 2 we conducted another picture-phrase matching experiment in which only one image was presented at a time, and participants judged whether the image was an example of the target adjective-noun description in yes-no format. A world knowledge effect in such discrete judgments would strongly point to flexibility in our truth-evaluations even in non-figurative language, contrary to some traditional assumptions in theoretical accounts of color adjective meanings as intersective (e.g., Keenan and Faltz 1985; Chierchia and McConnell-Ginet 1996; Drašković et al. 2013).

2 Experiments

2.1 *Pretest: Category Confirmation and Color Shift Judgments Along a Spectrum*

We presented photographs of seven categories with an intrinsic color bias: banana, bear, jeans, tomato, egg, grass, and horse. All of these categories had at least two naturally existing typical colors, with one of the colors being more 'canonical' (e.g., yellow bananas and green bananas). In order to investigate the effect of a color typicality bias on color judgments on a fine-grained level, we found digital photographs of objects from these categories, and manipulated the color of each category by starting with the original image and creating a duplicate layer with varying levels of transparency, hue, and/or saturation with a color copied from another object image of the same category in Photoshop.¹ The images varied in color along an 11-level spectrum (similar to Hampton 1996), but we needed to make sure that the color manipulation did not affect the category status—e.g., a banana with an ambiguous yellow-green color is judged to be atypical but nevertheless a banana—as observed in categorization judgments and response times.

¹There are two main ways of color matching we used in Photoshop. First, Photoshop has a built-in function 'Match Color' under Image—Adjustments for copying the color of (a selected area in) a source image to (a selected area in) a target image directly. Although this automatic function sometimes affects the luminance patterns in the target image too much, it is possible to control the luminance to preserve the patterns in a precise way. The second way is to create a duplicate layer in the target image file and use an eyedropper to sample a color in a source image to copy onto the target image or simply use the above method and adjust the transparency of the duplicate layer. It is also possible to pre-process the source image by applying an 'Average' filter under Filter—Blur when the internal pattern of the source image is too complex and difficult to transfer naturally to another image using the automatic method. Using the 'Color' mode instead of the default 'Normal' mode in the duplicate layer helps preserve the internal patterns based on contours, etc. Unfortunately there is no single optimal way for all categories, so it is necessary to decide the optimal combination on a case-by-case basis.

In Pretest (a), we looked at three kinds of stimuli: (1) the seven color-biased categories in three different shades (Levels 3, 6, and 9 on the 11-level spectrum, with higher numbers indicating higher color typicality)—a total of 21 Main trials—for detecting any extreme unnaturalness in any direction of the spectrum (3: less typical, 6: midpoint, 9: typical); (2) six Control trials involving categories that require positive extension of the normal noun meaning (stone lion, rubber duck, wooden toy car, model train, Mickey Mouse, and Miffy (an animated rabbit character)), which were included to ensure that our atypical colors in the Main stimuli would not lead to as much surprise as in these noun extension cases; and (3) 24 “No” Filler trials that required a clear “no” response, in order to prevent a set response (see Fig. 1).

Thirteen native speakers of Dutch provided picture-word match decisions in yes-no format under a 5-s time limit, and the response times were measured as well. The judgments were generally consistent with our expectation: For our Main stimuli, we observed around 95% “Yes” responses to all three shades (Levels 3, 6, and 9) of our test categories (258 out of 273 trials, with 1 timed-out trial and 14 “No” responses), confirming that our color manipulation in our stimulus images did not affect their noun category membership—black bears are just as good as brown and black-brown ambiguous ones for the category *bear*, and for Fillers, accuracy was high at 86%. For the six Control categories, for which we expected much greater surprise compared to the atypical colors in the Main trials at least in participants’ reaction times and possibly also in their higher rejection rates, our participants gave 94% “Yes” responses (73 out of 78), accepting the images most of the time for a broader sense of each noun category. In response times, however, these Control categories led to the slowest decisions, as we expected (mean = 1.28 s, see Table 1). Among our Main categories, in contrast, reaction-time differences due to color levels in trials with “yes” responses ($n = 258$: Level 3 average = 1065 ms, Level 6 average = 1009 ms, Level 9 average = 974 ms) were

Fig. 1 Example trials from Pretest (a): (1) Main trial, (2) Control trial, and (3) “No” Filler trial (translated from Dutch). [A color reproduction of this figure appears in the eBook at doi:[10.1007/978-3-319-45977-6_5](https://doi.org/10.1007/978-3-319-45977-6_5)]

Table 1 Response times in noun category confirmation

Stimulus type	<i>N</i>	Mean (ms)	Standard deviation (ms)
Main, Level 3	91	1077 (1065 in “yes” trials)	448
Main, Level 6	91	1073 (1009)	582
Main, Level 9	91	1002 (974)	421
Control	78	1283	589
Fillers	312	1169	518

Fig. 2 An example trial from Pretest 2(b): tomato. [A color reproduction of this figure appears in the eBook at doi:[10.1007/978-3-319-45977-6_5](https://doi.org/10.1007/978-3-319-45977-6_5)]

Table 2 Average level of perceived color shift on a level 1–11 manipulation spectrum

Category	<i>N</i>	Mean	Standard deviation
Banana	13	4.92	1.19
Bear	13	6.08	1.38
Egg	13	5.69	1.65
Grass	13	5.54	1.56
Horse	13	5.54	1.85
Jeans	13	4.85	1.28
Tomato	13	5.08	1.38

small and not statistically reliable ($F(2, 255) = 0.82, n.s.$), confirming our expectation that an atypical color would not make participants hesitate on the category membership of the object shown.

In Pretest (b), we asked the same group of participants from Pretest (a) for their color-shift judgment along a color spectrum, in order to confirm that they did indeed see a color change in our stimuli, and that the locus of this change was not skewed too much toward one end of our color manipulation spectrum. In this task, the participants saw the entire spectrum of 11 colors of each object category on a single screen and indicated the manipulation level at which they thought there was a color shift by typing in the corresponding number (see Fig. 2).

The direction of the spectrum on the screen (from Level 1 to Level 11, or from Level 11 to Level 1) was randomized for each trial. With Level 6 being the mid-point on the scale of 1–11, participants reported a perceived color shift around an average level of 5–7 for all our Main categories, as we expected (see Table 2).

3 Experiment 1: Forced Choice Between a Focal Color Versus a Nonfocal, Typical Color

In order to test which color people choose between a focal color and a nonfocal but canonical color for the category (e.g., focal green vs. nonfocal, ‘tomato’ green) as the better example of an adjective-noun description (e.g., *green tomato*), we conducted a preference judgment task in which participants had to choose between an image pair. As a control, we also tested categories with no strongly associated color (e.g., box). We predicted that participants’ color typicality knowledge would

influence their color preferences in this task, such that participants will be much more likely to prefer a nonfocal color over a focal one for color-biased categories such as tomatoes than for color-neutral categories such as boxes.

4 Method

We gave 11 adult native Dutch speakers a forced-choice picture-phrase matching task in which they saw two photographs of an object along with an adjective-noun combination and picked the image they preferred as the better match for the phrase. For example, the participant would see on the computer screen a photograph of a green tomato in its typical color (‘nonfocal’ green), another photograph of the same green tomato whose color was manipulated to be a focal green, and the expression *green tomato* (see Fig. 3). ‘Nonfocal’ colors were simply sampled from the web in a search of photographs of our stimulus categories, and for an operational definition of ‘focal’ colors in our digital images, we used the RGB triplets in Table 3.

We picked a category from the pretests which had a nonfocal color as a typically existing color and added more to the list for four color-biased categories (green tomato, green apple, orange sky, red leaf) and, as a control, four color-neutral categories (box, flag, table, T-shirt) whose colors were matched with a color-biased category. There were 16 filler trials with non-color adjective modifiers (such as striped apple, bald man, female scientist, and wooden spoon) or with mismatching noun categories. Four of these filler trials had an image of the canonical color for the category (e.g., red tomato) along with a focal-color image (focal-green tomato) to check that participants actually paid attention to the description (*green tomato*)

Fig. 3 Example trials from Experiment 1: (1) color-biased, and (2) color-neutral categories, in which the figure on the *right* of each pair is a focal *green*, and the figure on the *left* is a non-focal *green* typical of a *green* tomato. [A color reproduction of this figure appears in the eBook at doi:10.1007/978-3-319-45977-6_5]

Table 3 RGB triplets for the colors in Experiments 1 and 2

Color name (based on X11 names)	RGB
Blue	(0, 0, 255)
Dark orange	(255, 140, 0)
Office green	(0, 128, 0)
Red	(255, 0, 0)

Table 4 Mean proportion (%) of focal preferences by category in Experiment 1

Biased		Neutral	
Green apple	45	Green flag	55
Green tomato	45	Green box	73
Orange sky	73	Orange table	100
Red leaf	27	Red T-shirt	27

and pick the focal-color as the better match even when it is an unnatural image, rather than simply choose a more familiar image (red tomato, which is a bad match for the description *green tomato*) on most trials. Participants clicked arrow keys to indicate ‘left image,’ ‘right image,’ or ‘no preference’ (the last option was expected for the noun-mismatching trials).

4.1 Results

We predicted a color bias effect in a specific direction, namely, a lower proportion of focal preferences for color-biased categories. We thus recoded the participants’ responses for a one-tailed logistic regression test: Preference for the focal-color image was ‘1,’ and preference for the nonfocal-color image or no preference was ‘0.’ We analyzed the proportion of focal preferences as a function of color bias (see Table 4). Logistic regression with a random slope and intercept for Participant revealed that participants showed a significantly higher proportion of focal preferences for color-neutral categories (mean = 0.64, *SD* = 0.487, *N* = 44) than for color-biased categories (mean = 0.48, *SD* = 0.505, *N* = 44) ($z = -1.72$, $p = 0.043$, one-tailed).

5 Experiment 2: Yes-No Categorization Judgment

Although Experiment 1 demonstrated that, given a pair of images with a focal and a nonfocal color, participants’ preferences for an image matching a target phrase showed an effect of the typical color of the target category, it does not establish that people’s truth-judgments or categorization judgments may differ for the same color on two different objects, depending on whether the object categories have an intrinsically typical color or not. In other words, we were interested in finding instances of a nonfocal color that normally falls outside an acceptable range of a certain color term (e.g., a gingery-orange color for the term *red*) to see if this color will be rejected when it is applied to a category without an intrinsic bias in favor of that color (e.g., a gingery-orange car as an example of *red car*), but accepted when it is applied to a category with a color-biased category (e.g., gingery-orange hair as an example of *red hair*). We thus conducted another picture-phrase matching experiment in which only one image was presented at a time, and participants

judged whether the image was an example of the target adjective-noun combination in yes-no format. In addition to color adjectives, we included some other kinds of adjectives, such as pattern and material adjectives, in order to explore the generalizability of a world knowledge effect.

5.1 Method

For our linguistic stimuli, we selected seven adjectives, each combined with two noun categories to be tested for their ‘Biased’ versus ‘Neutral’ status in the relevant adjective dimension (color, pattern, or material) in a pretest: *red hair/car*, *green tomato/chair*, *striped apple/T-shirt*, *straight leg/road*, *cork mug/board*, *wooden bike/frame*, and *woolen shoe/floor-mat*. We conducted a pretest with 17 native speakers of Dutch to establish the Biased versus Neutral distinction in each of the seven category pairs above using three tasks. For color and pattern adjectives (*red*, *green*, *striped*, *straight*), we first conducted a focal preference task similar to Experiment 1 to confirm a higher focal preference for the neutral noun categories (see Table 5).

For material adjectives (*cork*, *wooden*, *woolen*), Biased versus Neutral status of noun categories was confirmed in free-response production and yes-no typicality judgment tasks. In the free response task, participants were shown a category name (e.g., *bike*) and asked to type in the typical material it is made of. Next, in the typicality judgment task, participants were shown an adjective-noun combination (e.g., *wooden bike*), and asked whether it was a typical combination. Our goal in the free-response task was to find no instances of spontaneous production of our target materials (cork, wood, wool) in the Biased categories, but a few instances of the target (or synonymous/hyponymous) materials in the Neutral categories, and our three category pairs confirmed our expected pattern (see Table 6). In the typicality judgments, we also confirmed the expected pattern of higher positive responses to our Neutral adjective-noun combinations than to the Biased counterparts in our category pairs (see Table 7).

For the 14 adjective-noun combinations, we prepared two photographs for each adjective-noun combination, one focal (e.g., red hair with a bright focal red) and one nonfocal (red hair with a more typical orange/copper hue). Within a Biased-Neutral pair, the values on the relevant adjective dimensions were held constant (RGB for color, pixel proportions for source material, and pattern/shape

Table 5 Mean proportion of focal preferences by category pair in Experiment 2 pretest

Biased		Neutral	
Hair (target ADJ: red)	24	Car (target ADJ: red)	94
Tomato (target ADJ: green)	41	Chair (target ADJ: green)	59
Apple (target ADJ: striped)	0	T-shirt (target ADJ: striped)	88
Leg (target ADJ: straight)	41	Road (target ADJ: straight)	71

Table 6 Common responses in Experiment 2 pretest (free response)

Biased		Neutral	
Mug	13 (out of 17) ‘porcelain/stone/clay’	Board	10 ‘cork’
	no ‘cork’		4 ‘cardboard/paper’
Bike	16 ‘metal/steel/iron/aluminum’	Frame	14 ‘wood’
	no ‘wood’		
Shoe	17 ‘leather’	Floor-mat	11 ‘fabric/cotton’
	no ‘wool’		1 ‘wool’

Table 7 Proportion of ‘yes’ responses (%) in Experiment 2 pretest (typicality judgment)

Biased		Neutral	
Cork mug	0	Cork board	94
Wooden bike	6	Wooden frame	100
Woolen shoe	6	Woolen floor-mat	59

for pattern by copying and pasting, see Fig. 4). Twenty-four adult native speakers of Dutch saw a photograph along with an adjective-noun combination, and judged (yes/no) whether the picture matched the expression within a three-second time limit. Each participant saw both photographs for each of the 14 adjective-noun combinations for a total of 28 main trials, along with 28 filler trials.

5.2 Results

In the Focal condition, acceptability judgments were high (>80%) for all adjective-noun combinations except one (*striped T-shirt*, 33%), confirming that participants treated the task as a category- or truth-judgment and not just a typicality/familiarity judgment. In the critical Nonfocal condition, in contrast, Biased categories (*hair*, *bike*, etc.) led to significantly higher ‘Yes’ responses (55%) than Neutral categories (*car*, *frame*, etc., 27%; $p < 0.001$).

5.3 Discussion

Our finding demonstrates that when typical properties of (noun) categories in our commonsense knowledge are biased *against* the ‘focal’ value of an adjective dimension (e.g., focal red in hair, 100% wood throughout a bike, etc.), our standards for categorization are relaxed such that a ‘nonfocal’ value (orange/copper rather than red, or wood only in parts of a bike) is more acceptable for these categories compared to those that have no such typicality bias against a focal value. Experiment 2 suggests that similar effects of typicality knowledge play a role in different

Biased	Focal	Nonfocal	Neutral	Focal	Nonfocal
red hair	87	92	red car	100	17
green tomato	96	58	green chair	96	92
striped apple	91	63	striped T-shirt	33	4
straight leg	96	38	straight road	100	0
cork mug	87	46	cork board	91	14
wooden bike	96	71	wooden frame	100	42
woolen shoe	87	17	woolen floor-mat	83	17

Fig. 4 Stimuli and percentage of “Yes” categorization responses from Experiment 2: non-focal and focal versions of color-biased categories (*left half*) and color-neutral categories (*right half*) (translated from Dutch). [A color reproduction of this figure appears in the eBook at doi:[10.1007/978-3-319-45977-6_5](https://doi.org/10.1007/978-3-319-45977-6_5)]

domains of adjectival meanings, such as colors, pattern, and material, although future research is needed for a much wider range of stimuli. The typicality effect in rapid discrete categorization beyond typicality ratings (e.g., Smith and Osherson 1984) lends support to theoretical accounts that propose a uniform underlying representational space for both typicality and truth judgments, such as Hampton’s (2007) threshold model.

6 Conclusion

Our results point to noun context effects on the interpretation of color adjectives, whose meanings show shifting boundaries for truth-judgments. Color adjectives may seem context-independent and intersective for many categories when the category-specific color spaces converge, but when we consider color judgments for categories with an intrinsic color bias in the real world, we observe context-dependent truth-judgments in uses of color terms. Similar effects of extensional feedback (Hampton 1988) in truth and categorization judgments may arise for many other adjective classes that have traditionally been analyzed as intersective.² Compositional processes that go beyond classical logic and set theory (such as Boolean conjunction and set intersection) are so pervasive in natural language that they cannot simply be set aside as a peripheral issue in semantic theory and pose a serious challenge to accounts of meaning composition as set intersection (e.g., Chierchia and McConnell-Ginet 1996; Heim and Kratzer 1998). It would also be important in future research to pursue further the compositional processes at varying degrees of frequent and conventionalized adjective-noun combinations. Dynamic, context-dependent ‘recalibrations’ of a predicate meaning (Kamp and Partee 1995) or modification of a comparison class to which a predicate applies (Klein 1980) seem to point to general processes of meaning composition in any domain where we have extensional feedback based on our world knowledge, not limited just to a small class of vague predicates.

Experimental studies in categorization and reasoning have made strides in mapping our conceptual space (Gärdenfors 2000) and fine-tuning our ideas about the combination of different conceptual dimensions. Fine-grained quantitative comparisons in the degree or amount of overextension between our study and earlier ones, especially Hampton (1996), would be difficult due to the subtlety of color space and color manipulation (or the lack of detailed descriptions of stimuli in Smith and Osherson 1984). A combination of tasks, such as truth-value judgments, color shift judgments in simultaneous presentation of multiple colors, forced-choice preferences between given colors, and phrase-picture matching tasks, should take us closer to better understanding of concept combination involving color adjectives.

There are important additional insights and challenges from the theoretical literature on color-term interpretation. One is the source or ontological status of the colors—i.e., whether they represent two distinct kinds (e.g., brown vs. black horses) or two stages of the same kind (green vs. red tomatoes that ripen over time). Kennedy and McNally (2010) argue that color adjectives are ambiguous between a gradable reading (denoting a degree scale for the color quality/quantity) and a non-gradable one (denoting a binary presence/absence of an underlying property

²In spontaneous production data (Sedivy 2003), color adjectives do differ, however, from material or scalar adjectives in that they appeared frequently even when they were over-informative and unnecessary, perhaps for a reinforcing effect in referential communication based on the perceptual salience of colors.

correlated with the surface color—e.g., genetic makeup). It remains to be seen whether such taxonomic knowledge is automatically and rapidly accessed, and makes a qualitative difference in our semantic composition of color adjectives with nouns. One may also apply the insights from an account of gradable adjectives such as Toledo and Sassoon's (2011) by analyzing the context-dependent determination of truth-conditions in terms of comparison classes consisting of other members of the same category (in type classification of color-biased categories), or apply the theoretical distinction between stage-level versus individual-level predicates (Carlson 1977) to color adjectives by considering other possible instantiations of an individual (for more gradable usage, e.g., in a maturational sense).

Contrary to a domain such as colors, there are domains that do not have a reasonable context-independent focal point or 'most typical value' (such as size and height—*big*, *tall*). We would expect similar world knowledge effects for these predicates as well (e.g., a man who is 190-cm tall may be considered tall in normal business attire but not tall in basketball gear, showing the typicality bias in height for basketball players as opposed to height-neutral businessmen), but these predicates need to be studied in future experimental research. Another interesting issue for future research is whether the relative order of modifier and head has any impact on the composition of meanings in real time. It would be interesting to see if preference for nonfocal color typicality is facilitated in languages with post-nominal adjectives, such as French and Hebrew, in which one processes the relevant noun category before a color adjective, in ways that are observable through nonfocal preference speed and proportion measures, compared to Dutch or English, with pre-nominal adjectives.

Acknowledgements This work was supported by the NWO VICI project *Between Logic and Common Sense* while the author was a postdoctoral researcher at the Utrecht Institute of Linguistics OTS. I thank Carien Westenbroek, Nina Wiedenhoff, Eva Poortman, and Yoad Winter for discussion and for help with data collection and stimulus and task selection. I also thank the other project team members and the audience at my Leiden-Utrecht Semantic Happenings talk for valuable comments.

References

- Anishchanka, A., Speelman, D., & Geeraerts, D. (2014). Referential meaning in basic and non-basic color terms. In W. Anderson, C. P. Biggam, C. Hough & C. Kay (Eds.), *Colour studies: A broad spectrum* (pp. 323–338). John Benjamins Publishing Company.
- Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Los Angeles: University of California Press.
- Carlson, G. N. (1977). A unified analysis of the English bare plural. *Linguistics and Philosophy*, 1 (3), 413–457.
- Chierchia, G., & McConnell-Ginet, S. (1996). *Meaning and grammar: An introduction to semantics*. Cambridge, MA: MIT Press.
- Dražković, I., Pustejovsky, J., & Schreuder, R. (2013). Adjective-noun combinations and the generative lexicon. In J. Pustejovsky, P. Bouillon, H. Isahara, K. Kanzaki, & C. Lee (Eds.), *Advances in Generative Lexicon Theory* (pp. 181–202). Dordrecht: Springer.

- Gärdenfors, P. (2000). *Conceptual spaces*. Cambridge: MIT Press.
- Hampton, J. A. (1988). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 14, 12–32.
- Hampton, J. A. (1996). Conjunctions of visually-based categories: Overextension and compensation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 22, 378–396.
- Hampton, J. A. (2007). Typicality, graded membership, and vagueness. *Cognitive Science*, 31, 355–383.
- Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology*, 93(1), 10–20.
- Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford, UK: Blackwell Publishers.
- Kamp, H., & Partee, B. (1995). Prototype theory and compositionality. *Cognition*, 57, 129–191.
- Keenan, E. L., & Faltz, L. M. (1985). *Boolean semantics for natural language*. Dordrecht: Kluwer.
- Kennedy, C., & McNally, L. (2010). Color, context, and compositionality. *Synthese*, 174, 79–98.
- Klein, E. (1980). A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, 4(1), 1–45.
- Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, 20, 158–190.
- Mitterer, H., Horschig, J. M., Müsseler, J., & Majid, A. (2009). The influence of memory on perception: It's not what things look like, it's what you call them. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(6), 1557–1562.
- Mitterer, H., & de Ruiter, J. P. (2008). Recalibrating color categories using world knowledge. *Psychological Science*, 19, 629–634.
- Regier, T., Kay, P., & Cook, R. S. (2005). Focal colors are universal after all. *Proceedings of the National Academy of Sciences of the United States of America*, 102(23), 8386–8391.
- Sedivy, J. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research*, 32(1), 3–23.
- Smith, E. E., & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, 8, 337–361.
- Smith, E. E., Osherson, D. N., Rips, L. J., & Keane, M. (1988). Combining prototypes: A selective modification model. *Cognitive Science*, 12, 485–527.
- Steinvall, A. (2002). *English colour terms in context*. Doctoral dissertation, Umeå University.
- The Oxford English Dictionary*. (1993). CD-ROM version. Oxford: Oxford University Press.
- Toledo, A., & Sassoon, G. (2011). Absolute vs. relative adjectives—Variance within vs. between individuals. In *Proceedings of Semantics and Linguistic Theory (SALT 21)*, (pp. 135–154).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.