
Communicative relevance: Color references in bilingual and trilingual speakers

ANETA PAVLENKO, SCOTT JARVIS, SVITLANA MELNYK and ANASTASIA SOROKINA

Bilingualism: Language and Cognition / *FirstView* Article / May 2016, pp 1 - 14
DOI: 10.1017/S1366728916000535, Published online: 20 May 2016

Bilingualism: Language and Cognition: page 1 of 14 © Cambridge University Press 2016 doi:10.1017/S1366728916000535

Communicative relevance: Color references in bilingual and trilingual speakers*

ANETA PAVLENKO
Temple University
SCOTT JARVIS
Ohio University
SVITLANA MELNYK
Indiana University
ANASTASIA SOROKINA
Temple University

(Received: July 20, 2014; final revision received: April 1, 2016; accepted: April 1, 2016)

The study examined granularity of lexical partitioning of the blue area in speakers of English, which encodes the term blue*; Russian, which encodes two terms*, sinij *[dark/navy blue] and* goluboj *[light/sky blue]; and Ukrainian, which encodes the terms* synij *[dark/navy blue] and* blakytnyj *and* golubyj *[light/sky blue]. Five groups of participants took part in the study: (1) 30 L1 speakers of English, (2) 30 L1 speakers of Russian, (3) 30 Russian–English bilinguals, (4) 30 English–Russian bilinguals, and (5) 25 Ukrainian–Russian–English trilinguals. Quantitative and qualitative analyses revealed that L1 Russian speakers referred to different types of blue significantly more frequently than all other groups, while bilinguals patterned with L1 English speakers. These findings suggest that classroom exposure to L2 Russian does not make the distinction between* sinij *and* goluboj *communicatively relevant for L1 English speakers and that everyday use of L2 English may trigger attrition of the contrast in L1 Russian.*

Keywords: color, communicative relevance, bilingual lexicon, crosslinguistic influence, Russian

Introduction

An omnipresent feature of our physical environment, color is one of its most enigmatic dimensions because at the heart of our color experience is a subjective interpretation of the wavelengths of light manufactured by our neural responses. The nature of this interpretation is all the more interesting in light of cross-linguistic variation: some languages do not encode an abstract category of 'color' and others vary in the number of 'basic' color terms and in the foci and boundaries of the categories referred to by these terms (Berlin & Kay, 1969; Bricker, 1999; Kuschel & Monberg, 1974; Levinson, 2001; Lucy, 1997; Turton, 1980; Wierzbicka, 2005).

This variation raises a number of questions in the study of language and cognition. The first involves the nature and evolution of color lexicons: is our linguistic division of the light spectrum arbitrary or is it systematic and guided by the 'universals' of color experience? The second concerns the relationship between language and color cognition: Are our perception of and memory for colors independent from or affected by the languages we speak? The third considers the acquisition of color terms: Given the ever-changing nature of lighting conditions and physical environments, how do children acquire systematic patterns of color reference? The diversity of color lexicons also raises an interesting question for research on the bilingual lexicon and second language (L2) learning and use: In the context of cross-linguistic variation, how do bi- and multilinguals learn, use, and maintain color terms that refer to distinct categories?

The purpose of our study is to address the latter question. We begin with a brief overview of the key findings in studies of color naming and cognition, followed by a synthesis of the findings in research on color naming by bilingual speakers. Then, we introduce our own study of color reference by monolingual, bilingual and trilingual speakers of English, Russian, and Ukrainian.

1. Language and colors: Do we see them as we call them?

The Berlin-Kay theory (Berlin & Kay, 1969) posits that the physiology of vision constrains linguistic color categories to a small number of those theoretically possible, with focal colors serving as prototypes for *basic color terms* (BCTs), defined as high frequency monolexemic items, psychologically salient to informants, whose meaning is not included in other terms and whose use is not restricted

* The authors are deeply grateful to Drs. Nina Vyatkina and Victoria Driagina-Hasko for their help with data collection and to anonymous peer-reviewers for insightful comments that greatly improved the quality of the final version of the manuscript. All remaining errors and inconsistencies are exclusively ours.

Address for correspondence:
Dr. Aneta Pavlenko, Professor of Applied Linguistics, Department of Teaching and Learning, College of Education, Temple University, Philadelphia, PA 19122
apavlenk@temple.edu

to a narrow class of objects. The BCTs, in this view, evolved around the same foci, in a certain order, from two to eleven 'basic' terms. One of their distinguishing features is high *codability*, that is the efficiency with which a referent can be named in a given language (Brown & Lenneberg, 1954), seen in high inter-speaker agreement on the meaning and/or referential range of the term.

The idea of perceptual universals found support in experiments by (Rosch) Heider (1972; Heider & Olivier, 1972), who compared memory for colors in speakers of English and Dani, a New Guinean language with two BCTs, *mili* [dark/cold] and *mola* [light/warm]. The studies showed that both groups remembered focal colors better than non-focals (even though the Dani performed significantly worse than the speakers of English) and allowed Heider (1972) to argue that memory for colors is influenced by their perceptual salience, rather than linguistic encoding.

These conclusions were challenged by Roberson, Davies and Davidoff (2000), who failed to replicate Heider's (1972) findings with speakers of Berinmo, a New Guinean language with five basic color terms, including *nol* [live, extends to green, yellow, blue, and purple] and *wor* [leaves ready to fall, extends to yellow, orange, brown, and khaki]. A comparison of performance by Berinmo and English speakers revealed no recognition advantage for focal colors and showed that language facilitates both memory and categorical perception: each group performed better on the contrast encoded in their own language (*blue-green* in English and *nol-wor* in Berinmo). These results were replicated with speakers of another language with five BCTs, Himba (Namibia), whose speakers showed a performance advantage in relation to their own categories but not those of English or Berinmo (Roberson, Davidoff, Davies & Shapiro, 2005), and with speakers of English, Korean, Russian, Setswana and Tarahumara, who displayed categorical perception at language-specific boundaries (Davies & Corbett, 1997; Kay & Kempton, 1984; Roberson, Pak & Henley, 2008). In the context of a great amount of variation in foci placement across and within language groups (e.g., Davidoff, Davies & Roberson, 1999; Jameson & Alvarado, 2003; Webster, Webster, Bharadwaj, Verma, Jaikurnar, Madan & Vaithilingham, 2002; Webster & Kay, 2007), these findings have challenged the idea of perceptual salience of 'universal foci' and have led Webster and Kay (2007) to acknowledge that color judgments are inherently subjective and "can be biased by individually experienced linguistic or cultural contexts" (p. 37).

One language-specific boundary investigated in past research involves two 'basic' terms for blue found in several languages, including Greek and Russian. This encoding presents a challenge to the Berlin-Kay theory, which posits the existence of eleven BCTs, which happen to mirror the color lexicon of English. Dismissing this 12th term challenge, Berlin and Kay (1969) and MacLaury (1997) have argued that in Russian only one term, *sinij* [navy blue, commonly glossed as 'dark blue'], is dominant, while *goluboj* [sky blue, commonly glossed as 'light blue'] is recessive. However, studies using large groups of participants and a variety of stimuli, including the Munsell Color array, Color-Aid samples, and the Swedish Natural Color System, have confirmed the BCT status of both terms on the basis of their high frequency of occurrence, high salience in elicited lists of color terms, high naming agreement, short naming times, derivational productivity, and the use of the modifiers *svetlo-* [light] and *temno-* [dark] with each term (Andrews, 1994; Corbett & Morgan, 1988; Frumkina, 1984; Morgan & Corbett, 1989; Moss, Davies, Corbett & Laws, 1990; Paramei, 2005, 2007; Vasilevich, Kuznetsova & Mishchenko, 2008).

Together, the studies conducted in the past decade have led to a reconsideration of the relationship between language and color cognition. Even Kay and Regier, staunch supporters of the universalist approach, acknowledged that "language structures do seem to influence cognition or perception" (Regier, Kay, Gilbert & Ivry, 2010: 167). Under the new consensus, language functions as an attention-directing mechanism, making categories referred to by language-specific BCTs more perceptually salient and facilitating color memory, learning, and categorical perception, with the focus on category boundaries (Kay & Regier, 2006, 2007; Regier, Kay & Khetarpal, 2009; Regier et al., 2010; Webster & Kay, 2007). This consensus has raised an interesting question for the study of L2 learning and the bilingual lexicon: If a color lexicon functions as an attention-directing mechanism, what happens when someone starts learning a language whose color categories differ from those of their native language and any other languages they have previously learned?

2. Color categories in the bilingual lexicon: Now you see them, now you don't

Studies of color naming and categorization reveal that bilinguals' color lexicons display all crosslinguistic influence processes documented in other lexical domains: (1) first language (L1) influence on the L2 or additional languages (L3, LX), (2) internalization of L2 or LX categories, (3) co-existence of language-specific naming patterns, (4) L2 influence on L1 naming and categorization, (5) L1 attrition and (6) incomplete acquisition of the L1 in the L2 context.

L1 influence on L2 naming patterns

In a pioneering study of bilinguals' color categories, Ervin (1961) found that Navajo–English bilinguals favored the term *yellow* in contexts where monolingual Navajos favored its translation equivalent *łitso* [yellow] and

monolingual English speakers were split between *yellow* and *brown*. Bilinguals also displayed a preference for *yellow* in the context of the yellow/green boundary, where Navajo monolinguals used *łitso* [yellow] and English monolinguals favored *green*, and a preference for *gray* in the context where monolingual Navajos used the term *liba*, commonly linked to *gray*, and English monolinguals varied between *blue* and *purple*. L1 influence was particularly visible in speakers with minimal English competence, and five decades later Ervin-Tripp (2011) recalled the unforgettable experience of hearing "a purple chip called *green*" (p. 222), which appeared to have been due to the range of colors associated with the term *dootł'izh* [blue/green/turquoise/purple]. Her findings suggested that in the context of referential indeterminacy (i.e., variation in target language naming patterns), L2 learners may be particularly likely to fall back on L1 patterns.

Internalization of L2 color categories

Ervin (1961) also found that English-dominant Navajos split the category *dootł'izh* into *blue*, *green*, and *purple*, suggesting that adults can learn to make more fine-grained distinctions and form new color categories. In other studies, L1 speakers of Zuni learned to differentiate in their L2 English between *yellow* and *orange*, which are referred to with a single term in Zuni (Lenneberg & Roberts, 1956) and L1 speakers of For acquired the terms 'blue' and 'brown' encoded in Arabic (their L2) but not in L1 For (Jernudd & White, 1983). Caskey-Sirmons and Hickerson (1977) found that speakers of L1 Hindi, Cantonese, and Mandarin formed new categories under the influence of L2 English: in Hindi, for instance, there is no word for *gray* and monolingual Hindi speakers did not map the gray area; in contrast, three of five Hindi–English bilinguals did map such an area, showing sensitivity to the new category acquired in L2 English.

Co-existence of language-specific naming patterns

Bilinguals who use both languages on an everyday basis may be able to maintain distinct categories in the L1 and L2. For example, Jernudd and White (1983) found that on the same naming task For–Arabic bilinguals used 5 to 8 terms in L1 For and 6 to 9 in L2 Arabic, where they singled out discrete categories for 'blue', 'brown', 'purple' and 'orange'. Saunders and van Brakel (1997) observed that bilingual speakers of Kwakw'ala (spoken on Vancouver Island) and English differentiate between *yellow* and *green* when speaking English but in Kwakw'ala stick to the term *lhenxa* [yellow-with-green]. In turn, Andrews (1994) found that Russian–English bilinguals who arrived in the US as adults maintain the obligatory distinction between *sinij* [dark/navy blue] and *goluboj* [light/sky blue] in L1 Russian, while using the single term *blue* in L2 English.

L2 influence on L1 naming and categorization patterns

With prolonged exposure, however, the L2 may exert its influence on L1 naming and categorization patterns, both on the individual level, and, in the case of languages that mark fewer distinctions, on the level of the color lexicon. Several studies have documented contact-driven lexical borrowing in the color lexicon, evident in the adoption of the English terms *orange*, *pink*, and *gray* in Japanese (Iijima, Wenning & Zollinger, 1982; Uchikawa & Boynton, 1987), English *blue* in Hausa and other indigenous language in Nigeria (Ibrahim, 2014), Arabic *samawi* [blue, sky-like] and *asfar* [yellow] in For (Jernudd & White, 1983), Afrikaans *pers* [purple] and English *pink* and *orange* in Damara (Davies, Roling, Corbett, Xoagub & Xoagub, 1997), German *blau* [blue] in Herero and Himba (Roberson & Hanley, 2010), and Spanish *azul* [blue] and *naranja* [orange] in Mesoamerican languages (Bricker, 1999; MacLaury, 1997).

L2 influence on the L1 has also been documented in individual bilingual speakers, in the form of convergence of the L1 categories with the L2 foci and boundaries (Caskey-Sirmons & Hickerson, 1977), in the form of category enlargement (Jameson & Alvarado, 2003), and in the form of increased divergence and variability of foci placement (Athanasopoulos, 2009; Caskey-Sirmons & Hickerson, 1977; MacLaury, 1997). In one study, Vietnamese–English bilinguals referring to 'orange' overused the Vietnamese term *cam* (the name of the fruit 'orange') and underused a modified term for 'yellow' favored by Vietnamese monolinguals who treat 'orange' as a subset of 'yellow' (Alvarado & Jameson, 2002; Jameson & Alvarado, 2003). These changes in the mapping of words to referents indicate ongoing restructuring of color categories in the bilingual lexicon. In a study with trilingual speakers, Iijima and associates (1982) found that under the influence of L2 English and L3 German, Japanese children shifted their L1 category *midori* [green] towards yellow and the category *ao-iro* [blue] in the opposite direction, towards violet, thus displaying greater differentiation between the two categories traditionally referred to by the single Japanese term *ao-iro*.

Another potential L2 effect is decreased frequency of categorical distinctions, seen in the weakening of the obligatory contrast between dark and light blue under the influence of L2 English in L1 Russian (Andrews, 1994), L1 Greek (Athanasopoulos, 2009; Athanasopoulos, Dering, Wiggett, Kuipers & Thierry, 2010), and L1 Japanese (Athanasopoulos, Damjanovic, Krajciova & Sasaki, 2011). Athanasopoulos (2009) examined the maintenance of the L1 Greek distinction between *ble* [dark blue] and *ghalazio* [light blue] in Greek–English bilinguals with different levels of L2 proficiency and found that the majority of advanced bilinguals had shifted their foci of *ble* towards the

prototype for the English *blue*. To maintain the perceptual distance between *ble* and *ghalazio*, they shifted the prototype for *ghalazio* away from the English prototype for *blue*, which resulted in a prototype that was lighter than that for Greek monolinguals.

L1 attrition and incomplete acquisition of the L1

In some contexts, L2 influence may also result in L1 attrition of the word-to-referent mappings and category boundaries. Alvarado and Jameson (2002) found that in Vietnamese–English bilinguals, attrition of L1 Vietnamese color terms began with modifiers and specific object glosses. Caskey-Sirmons and Hickerson (1977) identified category attrition in L1 Hindi: monolingual speakers of Hindi differentiated between *gulabi* [rose] and *lal* [red], while Hindi–English bilinguals incorporated the secondary category of 'rose' into the primary category of 'red'. In turn, Andrews (1994) documented incomplete acquisition of L1 Russian BCTs in the context of L2 English. He found that monolingual speakers of Russian and Russian–English bilinguals who had arrived in the US as adults consistently differentiated between *sinij* [dark/navy blue] and *goluboj* [light/sky blue], while childhood bilinguals did not: their performance revealed an expansion of the category *sinij* into the territory reserved for *goluboj* by adult bilinguals.

Studies of the maintenance of the L1 contrast between light and dark blue found that the degree of L2 English influence on L1 was affected by the age of L2 acquisition (Andrews, 1994) and the length of residence in the L2 environment (Athanasopoulos, 2009; Athanasopoulos et al., 2010). Athanasopoulos et al. (2011) also found that the maintenance of the L1 contrast may be affected by the frequency of L2 use: Japanese–English bilinguals who used L2 English more frequently resembled L1 English speakers in their judgments of the perceptual similarity of *ao* [dark blue] and *mizu-iro* [light blue].

Together, these studies suggest that BCTs behave like all other words in the bilingual lexicon and display all types of crosslinguistic influence. Their findings, however, have limited generalizability when it comes to L2 learning and use due to the low ecological validity of their experimental designs. Studies that elicit names for standardized color chips decontextualize colors by dissociating them from their referents, from other colors, and from communicative contexts. These tightly controlled tasks are undoubtedly useful in the study of color perception but tell us little about the learning and use of color words in everyday life, where colors appear as one of many referent properties, in the context of other colors, and in varying lighting conditions, and are invoked spontaneously in communicative contexts (e.g., van Leeuwen, 2011). The present study takes great inspiration from previous research on color perception yet does not aim to contribute to this line of research. As linguists, we are interested in what people do with words in the context of everyday interaction, and, in the present case, in what bi- and trilinguals do with color words. As a consequence, we dispense with variables that cannot be controlled in the study of everyday talk, including minor variations in ambient lighting. Our aim is to take the first step towards a more context-sensitive and ecologically valid design that retains some advantages of experimental procedures, such as control of the referents, yet allows for the investigation of spontaneous color reference in the context of communicative tasks.

3. Research design

3.1. Research aims

The purpose of the study was to compare the granularity of the lexical partitioning of the blue area in monolingual, bilingual and trilingual speakers. Blue was chosen as the focal domain, which allowed us to compare the uses of three languages: English, which encodes the single basic term *blue*; Russian, which encodes two basic terms, *sinij* [dark/navy blue] and *goluboj* [light/sky blue]; and Ukrainian, which encodes the term *synij* [dark/navy blue] and the terms *blakytnyj* and *golubyj*, both of which refer to [light/sky blue] (e.g., Starko, 2013). At the same time, all three languages encode a variety of modifers and non-basic terms that allow their speakers to refer to different shades of blue. The granularity of lexical partitioning, therefore, was treated as an index of the communicative relevance of distinctions between shades of blue (Notably, we do not see the English 'blue' as a more basic or 'natural' category than *sinij* and *goluboj*. The terms 'blue' and 'the blue area' are used here only for the convenience of English-language readers).

3.2. Participants

All participants were asked to fill out a background questionnaire and to rate the four skills (speaking, listening, reading, and writing) in all of their languages on a Likert scale, where 7 was native or native-like and 0 referred to a complete lack of ability. Data from participants who reported advanced levels of knowledge of languages other than those in the study were excluded from data analysis. The remaining 145 participants were distributed across five groups, shown in Table 1 and described below.

(1) L1 speakers of Russian (n = 30; 10 females, 20 males, ages 19–23) were undergraduates, majoring in physics or optics at Tomsk State University, Tomsk, Russia. Russian was their native and dominant language. All reported low levels of competence in English as a foreign language. A few students also reported low

Table 1. *Participant groups.*

Group	*n*	Male	Female	Age range
L1 Russian	30	20	10	19–23
L1 English	30	9	21	18–27
URE trilinguals	25	5	20	16–19
RE bilinguals	30	10	20	18–32
ER bilinguals	30	15	15	19–41
TOTAL	145	59	86	16–41

levels of competence in other foreign (French, German, Japanese) or second languages (Buriat, Kazakh).

(2) L1 speakers of English (n = 30; 21 females, 9 males, ages 18–27) were undergraduates at Temple University, Philadelphia, USA. Their majors included accounting, advertising, criminal justice, education, nursing, psychology, and social work. English was their native and dominant language. All reported low levels of competence in foreign languages, most commonly Spanish or French; a few also studied German, Hebrew, Italian, or Latin.

(3) Russian–English bilinguals (n = 30, 20 females, 10 males, ages 18–32, AoA 0–12) were undergraduate and graduate students at Temple University, Philadelphia, USA. All were childhood bilinguals: three were born in the USA in Russian-speaking families and the rest arrived in the USA between the ages of 1 and 12 as members of Russian-speaking families from Russia (16), Ukraine (5), Moldova (3), Belarus (1), Turkmenistan (1), and Uzbekistan (1). They used Russian at home with family members and, in a few cases, with friends and colleagues at work. All attended secondary school in the USA and used English for interactional and educational purposes. In self-reports, all gave themselves a rating of 7 (native or native-like) in listening, speaking, reading and writing in L2 English. In L1 Russian, their average self-rating scores were 6.3 (listening), 5.5 (speaking), 3.9 (reading), and 2.9 (writing). Some reported low levels of competence in foreign languages (French, Hebrew, Italian, Spanish).

(4) English–Russian bilinguals (n = 30, 15 females, 15 males, ages 19–41, AoA 12–26) were advanced-level L2 Russian students at the Middlebury College summer school, Middlebury, USA. All had extensive exposure to Russian through high school and university study and some also through a semester abroad. Their average self-ratings on the Likert scale for L2 Russian were 4.9 (reading), 4.7 (listening), 4.3 (speaking), and 4.2 (writing).

(5) Ukrainian–Russian–English trilinguals (n = 25, 20 females, 5 males, ages 16–19) were first- and second-year undergraduate students at Kyiv State University, a Ukrainian-medium university in Kyiv, Ukraine. All but one of the participants were ethnic Ukrainians and reported native or native-like proficiency in Ukrainian and Russian and intermediate to advanced competence in English (24 were English majors and one was a Spanish major). Some also reported low levels of competence in other languages (French, Polish, Spanish).

The study did not include monolingual Ukrainian speakers because in the urban student population that matched the other study participants, all L1 Ukrainian speakers had had significant exposure to Russian. The term 'L1 speaker' will refer, in what follows, to participants traditionally termed 'monolinguals' because they are not 'monolingual' in the strict sense of the term, just as some of our 'bilinguals' and 'trilinguals' are, in effect, multilingual. For the purposes of the present study, however, participants will be treated as functional monolinguals, bilinguals, and trilinguals. Given the narrow focus of the study on the lexical partitioning of the blue area, the diverse range of additional foreign/second languages, and participants' low levels of competence in these languages, we have no reason to suppose that exposure to additional languages could have affected the use of the words for blue. Nevertheless, we have made sure that the languages in question did not have more than two BCTs for blue and thus could not have increased the speakers' sensitivity to the distinctions in this domain and interfere with the results (e.g., on Polish, see Stanulewicz, 2010).

As far as gender is concerned, pevious research suggests that females may produce more secondary terms, while males use more BCTs accompanied by a modifier (e.g., Mylonas, Paramei & MacDonald, 2014). Since our study is focused on the granularity of partitioning, rather than on the terms used by participants, such gender imbalance is not likely to affect the final results. Nevertheless, we did consider the gender composition of the groups: the predominance of male participants in the L1 Russian group and of female participants in the L1 English group ensures that the frequency of color distinctions cannot be explained away by gender effects.

3.3. Methods, procedures, and stimuli

Data from L1 English speakers, Russian–English bilinguals and Ukrainian–Russian–English trilinguals were collected by the primary investigator, a Russian-Ukrainian–English trilingual. Data from L1 Russian speakers and English–Russian bilinguals were collected by two other researchers bilingual in Russian and English. The participants were recruited through on-campus advertisements and were not familiar with the interviewers prior to participating in the study. The experiment lasted

about an hour. First, participants completed a background questionnaire and a language-learning interview. Then, they were asked to describe several visual stimuli, including photographs, pictures in children's books (e.g., Mayer, 1969), and reproductions of paintings. Picture description was chosen as a task that allowed us to control the referents and at the same time elicit spontaneous references to the blue area in connected speech.

The same pictures and photographs were used with all five groups of participants. For the purposes of the present study, our analysis will ignore the distractor items and focus on four paintings with different shades of blue: Matisse's *The blue window* (1911) (Sternau, 1997, p. 77) and *Portrait of Madame Matisse* (1913) (Sternau, 1997, p. 63), Magritte's *The human condition* (1935) (Meuris, 1994, p. 43), and Van Gogh's *The starry night* (1889) (Cabanne, 2003). The reproductions of these paintings came from high-quality art books and were chosen among a larger array of paintings with shades of blue because all four presented an interpretive challenge for the viewer: *The human condition*, for instance, contains a painting within a painting. As a consequence, the description of the paintings presented a meaningful task where participants tried to figure out the artist's intentions, rather than provide a mechanical description.

Each participant was interviewed individually with the same array of pictures. The pictures were presented in a randomized order that included distractor items. Each reproduction was placed on the table in front of the individual participant in an office with electric lighting, accompanied by the following instructions in the language of the task: "Please describe what you see in this painting". The majority of the participants did mention colors as part of their description. Those who did not were asked a follow-up question: "What do you see as the dominant colors in the painting?" They were not prompted to comment specifically on the blues; therefore, we treat their lexical choices in the blue area as spontaneous reference. Throughout, we have maintained the same stimuli, instructions, and experimental conditions but made no attempt to measure the ambient lighting. Given the fact that data collection took place in three countries (Russia, Ukraine, USA) and four cities (Kyiv, Middlebury, Philadelphia, Tomsk), the maintenance of identical lighting would have been impossible and we will return to this issue in the discussion of our results.

In terms of the language of the task, functional monolinguals described the paintings in their L1, either Russian or English. Bilinguals described all four paintings in the L1/L2 Russian. Trilinguals described the two Matisse paintings in Ukrainian and the paintings by Magritte and Van Gogh in Russian (due to the lack of time, 3 participants did not complete the latter task). Afterwards, bilinguals and trilinguals participated in a debriefing interview, where we asked follow-up questions about their use of the words for blue.

3.4. Data analysis

All picture descriptions were audio-recorded and transcribed in the language of elicitation by bilingual research assistants. Each transcription was then cross-checked for accuracy by two of the authors, Aneta Pavlenko and Svitlana Melnyk. Subsequently, all descriptions were coded for verbal references to the blue area by one of the researchers and cross-checked by the others. In the process, we have developed the coding scheme presented in Table 2, whose purpose is to facilitate comparisons between languages with different numbers of BCTs.

As seen in Table 2, the fact that the three languages had a different number of BCTs for 'blue' with different referential ranges presented a challenge for coding. To ensure consistency of coding across languages, we have focused on the verbal differentiation between different categories of blue, rather than on the terms per se, and coded the uses of modifiers, such as *dark-* or *light-*, and secondary terms, such as *aqua* or *turquoise*. We do not assume, however, that the use of two or more terms always represents the imposition of two or more categories – rather, we have paid careful attention to patterns of use across picture descriptions produced by a single participant. As a consequence, if a participant used a single color term, such as *sinij*, across all four picture descriptions, such uses received a score of 1 (single category) for each use. If, however, they used the word *sinij* in some contexts and the word *goluboj* in others, they received a score of 2 for each choice, based on the fact that to pick each term they had to make a more fine-grained judgment than the participant who used *sinij* or *blue* across the board. A score of 1 was also assigned to picture descriptions where two or more terms applied to the same referent, e.g., *синее-голубое море* [sinee-goluboe sea]. We realize that such usages may reflect differentiation between the two categories, yet without further textual evidence of differentiation we chose to score conservatively, in order not to favor L1 speakers of Russian and Ukrainian.

Below, we will illustrate this usage with excerpts from participants' picture descriptions that appear under the following codes: E (L1 English speakers), R (L1 Russian speakers), RB (Russian–English bilinguals), EB (English–Russian bilinguals), UT (Ukrainian–Russian–English trilinguals), F (females), M (males). Consequently, the code RF3 refers to a female L1 speaker of Russian, numbered 3 in the database, the code UTM2 refers to a male Ukrainian–Russian–English trilingual numbered 2 in the database.

Table 2. *Coding scheme.*

Score	Categories	English-language descriptions	Russian- and Ukrainian-language descriptions
0	Low salience	No references to blue	No references to blue
1	Single category	BCT *blue* without modifiers, e.g., "There's some vases or lamps of some sort against a blue background. Looks like a table up against a window . . . Everything is blue outside the window" (EM6, Matisse (1911) *The blue window*)	(1) One BCT used across all four picture descriptions (2) Two BCTs used jointly to refer to the same referent, e.g., "My море, естественно, оно сине-голубое" [The sea, of course, is *sinee-goluboe*] (RF8, Magritte (1935) *The human condition*)
2	Two categories	(1) BCT *blue* and a modifier, such as *dark-*, *light-* or *navy-*, e.g., "Water is blue and the sky is lighter blue" (EF16, Magritte (1935) *The human condition*) (2) A secondary term, such as *aqua*, indicating differentiation from *blue*	One or two BCTs without modifiers, e.g., "Ма цій картині також переважає синій колір, блакитного тут дуже мало" [This painting is also dominated by *synij*, there is very little *blakytnyj* here] (UTF11, Matisse (1913) *Portrait of Madame Matisse*)
3	Three categories	BCT *blue* with two modifiers or secondary terms, e.g., "It's looking out of an old doorframe, looking out into water, and the light blue sky and the water is darker blue" (EF7, Magritte (1935) *The human condition*)	Three BCTs, or one or two BCTs with modifiers or secondary terms, e.g., "У неё оранжевый шарфик, синее пальто и гол/кофточка где-то цвета морской волны." [She has an orange scarf, *sinee* coat and a *gol*/kind of turquoise blouse] (RF7, Matisse (1913) *Portrait of Madame Matisse*)
4	Four or more categories	BCT *blue* with three or more modifiers or secondary terms, e.g., "lots of different dark blues, navy blues, kind of cornflower blue mixed in, her face is painted a lighter blue, but it's not a sky-blue" (EF24, Matisse (1913) *Portrait of Madame Matisse*)	One, two, or three BCTs with modifiers or secondary terms, e.g., "Море синее, дальше там . . . тёмно-синее, может быть это чуть-чуть светло-голубое . . . А небо бледно бледно голубое" [The sea is *sinee*, further down . . . dark-*sinee*, maybe a little light-*goluboe* . . . And the sky is pale pale *goluboe*] (RM12, Magritte (1935) *The human condition*)

Table 3. *Median number of blues referred to.*

Group	Magritte	Matisse Portrait	Matisse Still Life	Van Gogh	OVERALL
R	3.00	3.00	3.00	3.00	2.75
E	1.50	2.00	1.00	1.00	1.50
RB	2.00	2.00	2.00	1.50	1.50
EB	2.00	2.00	2.00	1.50	1.75
UT	2.00	2.00	2.00	2.00	2.25

4. Results

4.1. Results of the quantitative analysis

The median scores for each group are shown in Table 3. These numeric values show the median number of blues that were referred to by each group in each of the four target paintings. The right-most column in Table 3, labeled "OVERALL", represents the group medians for individual participants' mean number of references to blues in all four paintings. The same information is displayed

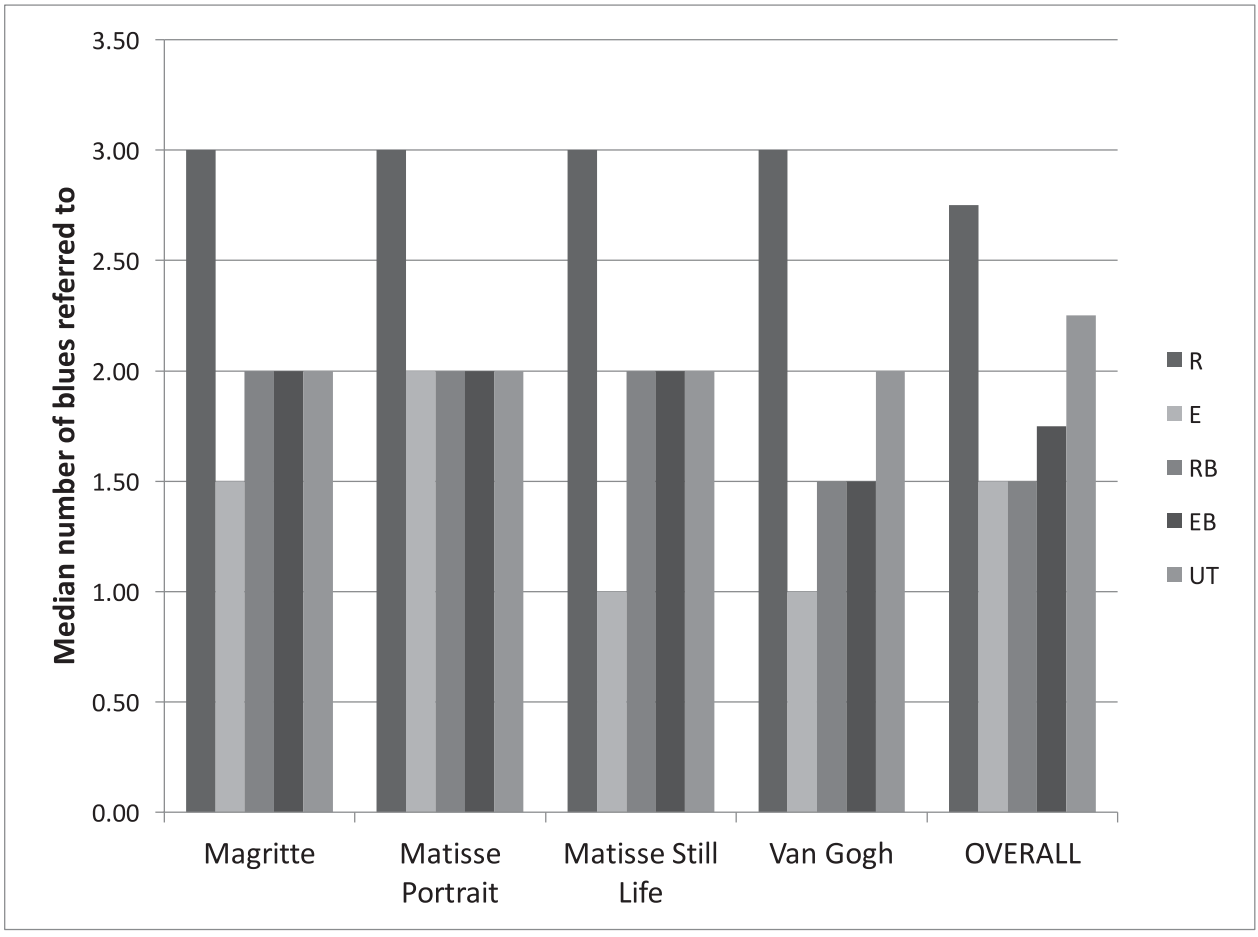

Figure 1. Median number of blues referred to (by group and painting).

graphically in Figure 1, where it is clear that the L1 Russian speakers consistently refer to more blues than any other group. The L1 English speakers, by contrast, have a lower median than all of the other groups for the Magritte, Matisse still life, and Van Gogh paintings. The L1 English speakers also share the lowest median score for the Matisse portrait with the RB, EB, and UT groups, and share the lowest overall median with the RB group. The RB and EB groups exhibit the same median scores for all four individual paintings, but the EB group has a slightly higher median score overall. Finally, the UT group displays the same median score as the RB and EB groups for the Magritte and both Matisse paintings, but has a higher median score than these groups for the Van Gogh painting, and also has a higher overall median score than all groups except the Russian L1 speakers. The descriptive statistics thus show that the L1 Russian speakers refer to more blues than any other group, followed by the UT group, the EB group, the RB group, and finally the L1 English speakers.

A series of nonparametric Kruskal-Wallis tests of the differences in mean ranks across groups showed that these differences were significant at the level of $p < .001$ for all individual paintings and also for the overall values representing the four paintings collectively (see Table 4). (Note: three of the 145 participants failed to produce relevant values for two of the paintings; the overall means for these three participants were thus calculated from their descriptions of just two of the four paintings). The effect sizes shown in Table 4 indicate that the differences across groups were smallest in the participants' descriptions of the Matisse portrait ($eta^2 = .14$) and greatest in the participants' descriptions of the Van Gogh painting ($eta^2 = .36$). Additionally, the effect size was slightly greater for the overall medians ($eta^2 = .39$) than for the medians found in relation to the Van Gogh painting.

Pairwise Mann-Whitney U post-hoc tests with False Discovery Rate (FDR) corrections (see Benjamini & Hochberg, 1995, 2000) showed that the differences lay primarily between the L1 Russian speakers and all other groups. For the Magritte painting (FDR alpha = 0.020), four significant pairwise differences were found, all involving the L1 Russian speakers and each of the other groups: R vs. E ($p < .001$), R vs. RB ($p < .001$), R vs.

Table 4. *Nonparametric (Kruskal-Wallis) tests of differences across groups.*

	Magritte	Matisse Portrait	Matisse Still Life	Van Gogh	OVERALL
H	35.449	22.576	48.027	52.279	56.815
N	142	145	145	142	145
df	4	4	4	4	4
p	<.001	<.001	<.001	<.001	<.001
eta^2	0.24	0.14	0.32	0.36	0.39

EB ($p < .001$), and R vs. UT ($p = .004$). For the Matisse portrait (FDR alpha = 0.015), three significant pairwise differences were found: R vs. EB ($p < .001$), R vs. RB ($p = .001$), and EB vs. UT ($p = .002$). Regarding the Matisse still life painting (FDR alpha = 0.035), seven significant pairwise differences emerged: R vs. E ($p < .001$), R vs. RB ($p < .001$), R vs. EB ($p < .001$), E vs. UT ($p < .001$), EB vs. UT ($p < .001$), RB vs. UT ($p = .004$), and R vs. UT ($p = .011$). Next, for the Van Gogh painting (FDR alpha = 0.040), eight significant pairwise differences were found: R vs. E ($p < .001$), R vs. RB ($p < .001$), R vs. EB ($p < .001$), E vs. UT ($p < .001$), EB vs. UT ($p = .003$), RB vs. UT ($p = .005$), E vs. EB ($p = .018$), and R vs. UT ($p = .037$). Finally, for the overall results (FDR alpha = 0.035), seven significant pairwise differences were found: R vs. E ($p < .001$), R vs. RB ($p < .001$), R vs. EB ($p < .001$), R vs. UT ($p < .001$), E vs. UT ($p < .001$), EB vs. UT ($p < .001$), and RB vs. UT ($p = .008$). The post-hoc tests for the overall results thus place the groups into the following three homogeneous subsets in terms of the mean number of blues they refer to in their descriptions of all four paintings: R > UT > EB, RB, E.

The reader will recall that the UT group described the two Matisse paintings in Ukrainian and the paintings by Magritte and Van Gogh in Russian. In order to determine whether they showed similar patterns in their reference to blues in both languages, and also in order to determine whether the number of blues they referred to differed significantly depending on the language they used for the task, we combined the scores for the two paintings they described in Ukrainian, on the one hand, and the scores for the two paintings they described in Russian, on the other hand, and submitted these scores to a Spearman rank correlation test and a Wilcoxon Signed Ranks test. The Spearman rank correlation test did not find a significant correlation between the participants' scores in the two languages (*rho* = -0.152, $N = 22$, $p = .5$), and the Wilcoxon test also did not reveal a significant difference between their median scores in the two languages (Ukrainian median = 2.50, Russian median = 2.00, $Z = -1.589$, $N = 22$, $p = .112$).

4.2. Results of the qualitative analysis

To elucidate the patterns of color reference identified in the quantitative analysis, we also performed a qualitative analysis of the data. To ensure that our results are not an artifact of differences in ambient lighting, we compared references to the same objects, namely Madame Matisse's clothing, across groups interviewed in different conditions. We found that speakers of L1 Russian and L1 English and Ukrainian–Russian–English trilinguals all described Madame Matisse's jacket as *dark (navy) blue* or *sinij*, while her shirt – when mentioned at all – was described as *aqua* or *turquoise*. This consistency suggests that the description of the same jacket as *goluboj* by Russian–English bilinguals – interviewed in the same room and in the same lighting as the L1 English speakers – reveals incomplete mastery of Russian color terms, rather than a change in lighting. Next, we have considered patterns of color reference produced by the L1 speakers of English and Russian, then the patterns by the Ukrainian–Russian–English trilinguals, and finally the patterns produced by the two groups of bilinguals.

In the case of L1 Russian speakers, our qualitative analysis showed that the majority of the L1 Russian picture descriptions differentiated between three or more types of blue by using the terms *sinij* and *goluboj* as independent BCTs, each with its own modifiers, *svetlo-* [light] and *temno-* [dark], e.g.: "Море синее, дальше там . . . тёмно- синее, может быть это чуть-чуть светло- голубое . . . А небо бледно бледно голубое" [The sea is *sinee*, further down . . . dark-*sinee*, maybe a little light-*goluboe* . . . And the sky is pale pale *goluboe*] (RM12, Magritte (1935) *The human condition*). This pattern of usage resulted in a more complex and fine-grained partitioning of the blue space than is conventional in English. Relatively fine-grained partitioning of the blue space is, of course, possible in English through the use of modifiers (e.g., *light-blue*, *navy blue*) and secondary color terms (e.g., *aqua*, *turquoise*), but the patterns produced by the L1 Russian speakers reflect distinctions whose equivalents in English would normally require a double modifier (e.g., comparable but not fully equivalent to *light-light-blue*).

Our qualitative analysis of the L1 English speakers' patterns showed that some of them used a single modifier with 'blue' or a secondary color term, but the majority of the L1 English picture descriptions referred to a single category of 'blue', e.g., "Looks like there's a house on the beach and the ocean in the background is blue" (EF15, Magritte (1935) *The human condition*),

"A little chalkboard or a table, a piece of jewelry, all done basically in blue" (EF14, Matisse (1911) *The blue window*), "There's some vases or lamps of some sort against a blue background. Looks like a table up against a window . . . Everything is blue outside the window" (EM6, Matisse (1911) *The blue window*). Thus, whereas the majority of the L1 English descriptions made use of a single, unmodified BCT (i.e., 'blue'), the majority of the L1 Russian descriptions were two levels more fine-grained than this, revealing different approaches towards the communicative relevance of encoding the distinction.

Our qualitative analysis of the trilinguals' picture descriptions showed that, similar to the L1 Russian speakers, they systematically differentiated between *synij* [dark/navy blue] and *blakytnyj* or *golubyj* [light/sky blue], with the latter terms used in complementary distribution. Speakers of Eastern varieties of Ukrainian, which developed in contact with Russian, favored the term *golubyj*, borrowed from Russian, e.g., "ну, я бачу картину переважно голубих і синіх кольорів" [so, I see a painting in predominantly *golubi* and *syni* colors] (UTF5, Matisse (1911) *The blue window*). Speakers of Western varieties, which developed in contact with Polish, favored the term *blakytnyj*, borrowed from Polish, in the same context, e.g., "Ма цій картині також переважає синій колір, блакитного тут дуже мало" [This painting is dominated by *synij* color, there is very little *blakytnyj* here] (UTF13, Matisse (1911) *The blue window*). At the same time, both groups were aware of both terms. In the debriefing interviews, the participants reported that they treat the two as synonyms: "As far as I know, in Ukrainian these words are synonyms, and in Russian there is only one word for this color" (UTF18), "*blakytnyj* is Ukrainian and *golubyj*/ although sometimes they say *golubyj* in Ukrainian too" (UTF10), "*blakytnyj* and *golubyj* are identical for me . . . " (UTF8), "*blakytnyj* . . . it is the same as *golubyj* . . . ." (UTF15), "I use *blakytnyj* and *golubyj* as synonyms . . . " (UTF6). These responses show that availability of three BCTs does not necessarily constitute evidence of three distinct color categories – instead, two BCTs may be used as synonyms to refer to the same category.

The fact that Ukrainian makes exactly the same color distinctions in the blue space as Russian explains why the picture descriptions produced by the trilingual speakers do not show a higher level of granularity of the blue continuum than the picture descriptions produced by the L1 Russian speakers. It is less clear, however, why the trilinguals' picture descriptions display significantly lower granularity than those of the L1 Russian speakers. One possibility is that the lower granularity of trilinguals' references might reflect lower numbers of participants and picture descriptions, as well as interviewer effects. The L1 Russian and trilingual data were collected by different researchers, and despite the fact that they followed the same protocol, we cannot rule out the possibility that one researcher had the effect of soliciting more terms for blue than the other. At the same time, it is also possible that the data display genuine effects of learning L3 English on the salience of the light/dark blue contrast in Russian and Ukrainian.

The possibility of L2/L3 influence on the L1 is supported by our qualitative analysis of the Russian–English bilinguals' picture descriptions. Several participants used only one BCT, either *goluboj* or *sinij*, across all four paintings: e.g., "океан такой голубого цвета" [the ocean is of *goluboj* color] (RBF2, AoA 1.5 yrs; Magritte (1935) *The human condition*), "Те дома, это много голубой свет [Those houses, there is a lot of *goluboj* light (meaning: color)] (RBF4, AoA 4 yrs; Van Gogh (1889) *The starry night*). What is particularly striking about such usages is that they extend the uses of *goluboj* to navy blue referents described as *sinij* by L1 Russian speakers, e.g., "Я вижу, что женщина сидит в зелёном стуле или кресле и она одета в голубом пиджаке и в голубых штанах" [I see a woman sitting in (on) a green chair or armchair and she is wearing a *goluboj* jacket and *golubye* pants] (RBF2, AoA 1.5 yrs; Matisse (1913) *Portrait of Madame Matisse*), "На картине номер четыре голубая . . . комната а-а с окном, тоже выходит на голубое . . . с голубыми, с голубыми деревьями" [In the painting number four there is a *golubaja* . . . room with a . . . window, also looking out on a *golubaja* . . . with *golubye*, with *golubye* trees] (RBF13, AoA 12 yrs; Matisse (1911) *The blue window*). These patterns of usage suggest that early childhood bilinguals, like RBF2 and RBF4, may have never fully acquired the distinction between *sinij* and *goluboj*, while late childhood bilinguals, like RBF13, lost the distinction. In either case, they have then extended a single term to cover the entire domain.

These observations raise another important question: If the Russian distinction is difficult to maintain in the context of Russian–English bilingualism, how do L1 English speakers acquire it when learning Russian as an L2? We can address this question to some degree by looking at the patterns produced by the English–Russian bilinguals. As discussed earlier, this group consists of American learners of L2 Russian. Our qualitative analysis of their picture descriptions showed that they tended to use a single term, either *sinij* or *goluboj*, in all four picture descriptions. A few participants also used the terms interchangeably, as if they were synonyms, e.g., "светы синие, голубые" [colors [mispronounced] *sinij, goluboj* (pl)] (EBF9, Van Gogh (1889) *The starry night*). This pattern and the fact that the second term usually appeared in descriptions of the second or third painting may reflect a shift from automatic to conscious color reference at moments when participants might have remembered that Russian has not just one but two terms for 'blue' even if the participants were not completely sure what the

difference was. Interestingly, in our debriefing interviews, we found that a majority of the EB participants – including those who favored a single term for 'blue' in their picture descriptions – were able to define *sinij* as dark blue and *goluboj* as light blue. However, their performance in the picture description task demonstrated that their explicit knowledge of the meanings of these color terms has not yet led to automatic usage. That is, their spontaneous language performance in L2 Russian still betrays a single underlying category for the blue continuum and the lack of communicative relevance of the distinction between navy and sky blue.

5. Communicative relevance in the bilingual color lexicon: Discussion and conclusions

Following Brown and Lenneberg (1954), studies of color perception and naming often invoke the Sapir-Whorf hypothesis as their *raison d'être*, despite the fact that Boas and Whorf saw color perception – and visual perception in general – as an area NOT influenced by language (Boas, 1911 [1965]: 190; Whorf, [1940a] 2012: 267–268; [1940b] 2012: 209). We fully concur with Boas and Whorf and have no doubt that all healthy adults with normal color vision can differentiate between many more shades of color than they can name. Our interest is in another concept central to Whorf's ([1940a] 2012) thinking about language, ascription of significance:

> We cut nature up, organize it into concepts, and **ascribe significances** as we do, largely because we are parties to an agreement to organize it in this way – an agreement that holds throughout our speech community and is codified in patterns of our language. (p. 272; emphasis ours)

In the present paper, we reinterpret the ascription of significance as COMMUNICATIVE RELEVANCE and operationalize it as the granularity of lexical partitioning of the blue area in spontaneous reference. Perceptual salience is undoubtedly a prerequisite for lexical partitioning, since a distinction needs to be noticed in order to be referred to. At the same time, we do not ascribe the same significance to everything we notice and the fact that something is not named cannot be unambiguously interpreted as evidence that it was not noticed or perceived. Research shows that color naming is affected by both macro factors, such as the sociohistoric context and the speaker's age, linguistic background, education, occupation, and expertise, and by micro factors, including the time of day and lighting, the state we are in, the sociolinguistic context and communicative intent (e.g., Goldstone, 1998; Martinez, 2013; van Leeuwen, 2011). In contexts where other factors are held more or less constant, the absence or the low granularity of lexical partitioning may be interpreted as the lack of ascription of significance or communicative relevance to the distinction in question.

On the face of it, there is no reason for which L1 speakers of English, Russian, and Ukrainian should differ in descriptions of the same visual stimuli in the context of the same task. While the ambient lighting may have varied slightly across contexts, it was always sufficient to see all details of the paintings. The differences, if any, would have affected the perception of individual hues but not the ability to notice and describe the contrasts. All participants had the means to do so, as English encodes a wide range of secondary color terms, such as *teal* or *aqua*, and modifiers, such as *light-* and *dark-*, and the performance of a few outliers shows that speakers of L1 English DO perceive the differences between the shades of blue and CAN talk about them. What they can do, however, differs from what the majority of L1 English speakers did do habitually in the context of the picture description task, namely rely on the single BCT *blue*. L1 Russian speakers, on the other hand, systematically used the terms *sinij* and *goluboj* with modifiers, thus imposing significantly more fine-grained distinctions on the blue continuum. The description of Van Gogh's (1889) *The starry night* by a Russian-speaking participant illustrates the communicative treatment of the two color categories as separated by an imaginary border:

> Значит, с правого края, такой как бы чёрно-сиреневый холм, ниже идут тоже красивого цвета голубого, насыщенного такого, насыщенно голубого, уже на грани с синим, немножко видно так как бы . . . (RF2)

[So, on the right, there is such a black-lilac hill, lower there are of a nice *goluboj* color, such saturated, saturated *goluboj*, on the border of *sinij*, we can see a little bit . . .]

The pattern of L1 Russian color reference identified in the study supports previous claims about the status of both *sinij* and *goluboj* as BCTs (Andrews, 1994; Corbett & Morgan, 1988; Paramei, 2005, 2007; Vasilevich et al., 2008), yet it cannot be explained through this status, first, because such explanation would be circular, and secondly, because both L1 Russian terms were used with modifiers in constructions similar to the English *light sky-blue*. At the heart of the three interrelated phenomena – the BCT status of the two terms, their high codability, and the high granularity of lexical partitioning of the blue area by L1 Russian speakers – are the current norms in the Russian speech community, where the distinction between navy and sky blue is treated as communicatively relevant (Vasilevich et al., 2008).

Communicative relevance, in our view, offers a more interesting and useful framework than codability for understanding the findings of this study: its advantages include its openness to semantic and syntactic categories and its dynamic nature, sensitive to language change, which enables its integration into evolutionary models of the color lexicons (e.g., Narens, Jameson, Komarova & Tauber, 2012). The notion of codability fails to explain

why American L2 learners of Russian do not retrieve the words *sinij* and *goluboj* with the same efficiency as L1 Russian speakers. We have no doubt that L1 English speakers do perceive the difference between dark/navy and light/sky blue. They can also memorize the L2 Russian words corresponding to each color – in fact, in debriefing interviews several study participants assigned distinct meanings to *sinij* and *goluboj*. In picture descriptions, however, American L2 learners differentiated between these colors significantly less frequently than speakers of L1 Russian and Ukrainian. This performance suggests that they have not yet acquired the sense of communicative relevance of the distinction between navy and sky blue and have not formed two distinct multimodal categories that could be automatically triggered in spontaneous speech. Instead, they patterned with L1 English speakers, referring to a single undifferentiated category of 'blue'.

The Russian–English bilinguals also displayed decreased communicative salience of the *sinij/goluboj* contrast. Debriefing interviews, conducted after the experiement was over, confirmed that some participants did not differentiate between *sinij* and *goluboj*:

Segment 1, interview with RBF1 (AoA 1 yr):

Interviewer (I): Here is a question for you, [name]: Is there a difference for you between *goluboj* and *sinij*?

- **Participant (P):** Ah, yes, there should be.
- **I:** There should be but is there one?
- **P:** Uhm, it seems *goluboj* is probably darker and *sinij* is lighter.
- **I:** And in your opinion, which word do you use more frequently? The word *goluboj*?
- **P:** Yes.

Segment 2, interview with RBF4 (AoA 3 yrs):

- **I:** OK, [name], so you say *goluboj*, *goluboj*. Do you see a difference between *sinij* and *goluboj*?
- **S:** Yes, but I always say *goluboj*, I don't know.
- **I:** And can you explain the difference between *sinij* and *goluboj*?
- **S:** I don't know the difference. . . . It is simply *blue*.
- **I**: So for you this is all *blue* [pointing to the painting]?
- **S**: Yes.

These findings are consistent with those of Andrews (1994), who documented semantic extension of a single Russian term to the blue area and with those of Pavlenko and Malt (2011), who documented L2 English influence on the L1 Russian categories *chashka* [cup] and *stakan* [glass] in Russian–English bilinguals. They show that the norms governing communicative relevance of particular distinctions can be transferred from the more dominant language, in the present case L2 English, into the chronologically preceding L1. In the absence of longitudinal data, however, it is impossible to determine whether collapsing the two categories into one and extending the term *goluboj* to settings described as *sinij* by L1 Russian speakers represents L2 influence on L1 or incomplete acquisition of the L1.

Like all studies, the present study also has its limitations, which constitute opportunities for future research. The first involves the use of the picture task, which is insufficient from the point of view of ecological validity and generalizability. To understand the conventions that govern color reference in individual languages, future studies need to use a wider array of approaches, ranging from ethnographic studies to corpus analysis. A particularly interesting issue for future research is the influence of referents on the naming task. Russian speakers, for instance, often rely on collocations of the sea with *sinee* [dark/navy blue] and sky with *goluboe* [light/sky blue], regardless of the actual color of the referent. In experimental tasks it would be interesting to present participants with images of a light-blue sea and navy-blue sky to examine the influence of conversational conventions on the naming of atypical referents.

Another limitation of our study involves comparability: our decision to dispense with controls for chromaticity and changes in ambient lighting means that our findings are not directly comparable to findings in research on color perception. In future studies, it would be interesting to compare the naming of standardized stimuli, such as Munsell color chips, with references to 'the same' colors in picture descriptions and other communicative contexts, in order to determine the degree of correspondence between everyday color reference and labels elicited in the lab.

The limited number of trilingual participants also constrains the generalizability of the findings of L3 influence. In future studies, it would be interesting to examine whether post-puberty exposure to a foreign or second language with a single BCT for blue does indeed reduce the communicative relevance of the navy/sky blue contrast in L1 Russian or Ukrainian. In the study of L2 acquisition, it would be interesting to follow up on the dissociation displayed by American L2 learners of Russian between explicit identification of *sinij* and *goluboj* as the terms for dark and light blue and automatic references to a single blue. In our view, part of the problem lies in the fact that Russian dictionaries and textbooks link the two terms to a single English category of blue with modifiers, thus treating the distinction as non-obligatory. If future studies replicate the same dissociation between explicit knowledge and automatic reference, the next step would be to see if the replacement of modifier-based translations with the terms *navy* and *sky blue*, in combination with activities requiring the labeling of different visual referents (e.g., Malt, Jobe, Li, Pavlenko & Ameel, in press), may facilitate the formation of new prototypes. Another promising direction for future research involves the study of the longitudinal

development of color reference in the L2. It is our great hope that future research on color reference will step into the world outside the experimental laboratory to examine how we manage to learn and use basic color terms, with such impressive consistency, in the ever-changing physical environments.

References

Alvarado, N., & Jameson, K. (2002). The use of modifying terms in the naming and categorization of color appearances in Vietnamese and English. *Journal of Cognition and Culture, 2,* 53–80.

Andrews, D. (1994). The Russian color categories *sinij* and *goluboj*: An experimental analysis of their interpretation in the standard and émigré languages. *Journal of Slavic Linguistics, 2,* 9–28.

Athanasopoulos, P. (2009). Cognitive representation of color in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition, 12,* 83–95.

Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, J., & Thierry, G. (2010). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive color perception. *Cognition, 116,* 437–443.

Athanasopoulos, P., Damjanovic, L., Krajciova, A., & Sasaki, M. (2011). Representation of color concepts in bilingual cognition: The case of Japanese blues. *Bilingualism: Language and Cognition, 14,* 9–17.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B, 57,* 289–300.

Benjamini, Y., & Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational & Behavioral Statistics, 25,* 60–83.

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley, CA: University of California Press. (Reprinted in 1991)

Boas, F. ([1911]1965). *The mind of primitive man*. Revised edition. New York: The Free Press.

Bricker, V. (1999). Color and texture in the Maya language of Yucatan. *Anthropological Linguistics, 41,* 283–307.

Brown, R., & Lenneberg, E. (1954). A study of language and cognition. *Journal of Abnormal and Social Psychology, 49,* 454–462.

Cabanne, P. (2003). *Van Gogh*. Paris: Telleri.

Caskey-Sirmons, L., & Hickerson, N. (1977). Semantic shift and bilingualism: Variation in the color terms of five languages. *Anthropological Linguistics, 19,* 358–367.

Corbett, G., & Morgan, G. (1988). Color terms in Russian: Reflections of typological constraints in a single language. *Journal of Linguistics, 24,* 31–64.

Davidoff, J., Davies, I., & Roberson, D. (1999). Color categories in a stone-age tribe. *Nature, 398,* 203–204.

Davies, I., & Corbett, G. (1997). A cross-cultural study of color grouping: Evidence for weak linguistic relativity. *British Journal of Psychology, 88,* 493–517.

Davies, I., Roling, P., Corbett, G., Xoagub, F., & Xoagub, X. (1997). Color terms and color term acquisition in Damara. *Journal of Linguistic Anthropology, 7,* 181–207.

Ervin, S. (1961). Semantic shift in bilingualism. *American Journal of Psychology, 74,* 233–241.

Ervin-Tripp, S. (2011). Advances in the study of bilingualism: A personal view. In Cook, V. & B. Bassetti (eds.), *Language and bilingual cognition,* pp. 219–228. New York: Psychology Press.

Frumkina, R. (1984). Цвет, смысл, сходство: Аспекты психолингвистического анализа. [Color, meaning, similarity: Aspects of psycholinguistic analysis]. Moscow: Nauka.

Goldstone, R. (1998). Perceptual learning. *Annual Review of Psychology, 49,* 585–612.

Heider, E. (1972). Universals in color naming and memory. *Journal of Experimental Psychology, 93,* 10–20.

Heider, E., & Olivier, D. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology, 3,* 337–354.

Ibrahim, R. (2014). Expanding the domains of English influence: The issue of color naming in some indigenous Nigerian languages. *European Scientific Journal, 14,* 312–322.

Iijima, T., Wenning, W., & Zollinger, H. (1982). Cultural factors of color naming in Japanese: Naming tests with Japanese children in Japan and Europe. *Anthropological Linguistics, 24,* 245–262.

Jameson, K., & Alvarado, N. (2003). Differences in color naming and color salience in Vietnamese and English. *Color Research and Application, 28,* 113–138.

Jernudd, B., & White, G. (1983). The concept of basic color terms: Variability in For and Arabic. *Anthropological Linguistics, 25,* 61–81.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist, 86,* 65–79.

Kay, P., & Regier, T. (2006). Language, thought, and color: Recent developments. *Trends in Cognitive Sciences, 10,* 51–54.

Kay, P., & Regier, T. (2007). Color naming universals: The case of Berinmo. *Cognition, 102,* 289–298.

Kuschel, R., & Monberg, T. (1974). "We don't talk much about color here": A study of color semantics on Bellona Island. *Man, 9,* 2, 213–242.

Lenneberg, E., & Roberts, J. (1956). The language of experience: A study in methodology. Memoir 13 of *International Journal of American Linguistics*, supplement to volume 22, 2, April 1956. Baltimore, MD: Waverly Press.

Levinson, S. (2001). Yélî Dnye and the theory of basic color terms. *Journal of Linguistic Anthropology, 10,* 3–55.

Lucy, J. (1997). The linguistics of 'color'. In C. Hardin & L. Maffi (eds.), *Color categories in thought and language,* pp. 320–346. Cambridge: Cambridge University Press.

MacLaury, R. (1997). *Color and cognition in Mesoamerica: Constructing categories as vantages*. Austin, TX: University of Texas Press.

Malt, B., Jobe, R., Li, P., Pavlenko, A., & Ameel, E. (in press). What constrains simultaneous mastery of first and second language word use? *International Journal of Bilingualism*.

Martinez, E. (2013). Listen to that painting! An intercultural experiment based on cross-sensory coupling. *The International Journal of New Media, Technology, and the Arts, 7,* 33–47.

Mayer, M. (1969). *Frog, where are you?* New York: Dial.

Meuris, J. (1994). *Magritte*. Cologne, Germany: Taschen.

Morgan, G., & Corbett, G. (1989). Russian color term salience. *Russian Linguistics, 13,* 125–141.

Moss, A., Davies, I., Corbett, G., & Laws, G. (1990). Mapping Russian basic color terms using behavioral measures. *Lingua, 82,* 313–332.

Mylonas, D., Paramei, G., & MacDonald, L. (2014). Gender differences in colour naming. In W. Anderson, C. Biggam, C. Hough, C. Kay (eds.), *Colour studies: A broad spectrum* (pp. 225–239). Amsterdam / Philadelphia: John Benjamins.

Narens, L., Jameson, K., Komarova, N., & Tauber, S. (2012). Language, categorization and convention. *Advances in Complex Systems, 15,* 1–21.

Paramei, G. (2005). Singing the Russian blues: An argument for culturally basic color terms. *Cross-Cultural Research, 39,* 10–34.

Paramei, G. (2007). Russian 'blues': Controversies of basicness. In MacLaury, R., Paramei, G., & D. Dedrick (eds.), *Anthropology of color: Interdisciplinary multilevel modeling,* pp. 75–106. Amsterdam: John Benjamins.

Pavlenko, A., & Malt, B. (2011). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian–English bilinguals. *Bilingualism: Language and Cognition, 14,* 19–45.

Regier, T., Kay, P., & Khetarpal, N. (2009). Color naming and the shape of color space. *Language, 85,* 884–892.

Regier, T., Kay, P., Gilbert, A., & Ivry, R. (2010). Language and thought: Which side are you on, anyway? In Malt, B. & Ph. Wolff (eds.), *Words and the mind: How words capture human experience,* pp. 165–182. New York: Oxford University Press.

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, 129,* 369–398.

Roberson, D., Davidoff, J., Davies, I., & Shapiro, L. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50,* 378–411.

Roberson, D., & Hanley, R. (2010). Relatively speaking: An account of the relationship between language and thought in the color domain. In Malt, B. & Ph. Wolff (eds.), *Words and the mind: How words capture human experience,* pp. 183–198. New York: Oxford University Press.

Roberson, D., Pak, H., & Hanley, R. (2008). Categorical perception of color in the left and right visual fields is verbally mediated: Evidence from Korean. *Cognition, 107,* 752–762.

Saunders, B., & van Brakel, J. (1997). Are there non-trivial constraints on color categorization? *Brain and Behavioral Sciences, 20,* 167–228.

Stanulewicz, D. (2010). Polish terms for 'blue' in the perspective of Vantage Theory. *Language Sciences, 32,* 184–195.

Starko, V. (2013). Ukrainian colour concepts for blue. *Slovo: Journal of Slavic Languages and Literatures, 54,* 150–163.

Sternau, S. (1997). *Matisse*. London: Tiger Books International.

Turton, D. (1980). There is no such beast: Cattle and color naming among the Mursi. *Man, 15, 2,* 320–338.

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

Van Leeuwen, T. (2011). *The language of colour: an introduction*. London/New York: Routledge.

Vasilevich, A., Kuznetsova, S., & Mishchenko, S. (2008). *Цвет и названия цвета в русском*; *языке*; [Color and color terms in Russian]. Second ed. Moscow: LKI.

Webster, M., & Kay, P. (2007). Individual and population differences in focal colors. In MacLaury, R., Paramei, G. & D. Dedrick (eds.), *Anthropology of color: Interdisciplinary multilevel modeling,* pp. 29–53. Amsterdam/Philadelphia: John Benjamins.

Webster, M., Webster, S., Bharadwaj, S., Verma, R., Jaikurnar, J., Madan, G., & Vaithilingham, E. (2002). Variations in normal color vision. III. Unique hues in Indian and United States observers. *Journal of the Optical Society of America A, 19,* 1951–1962.

Whorf, B. (1940a). Science and linguistics. *Technology Review, 42,* 6, 229–231, 247–248. Reprinted in Whorf, B. (2012) *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. Second ed. Edited by J. Carroll, S. Levinson, and P. Lee. Cambridge, MA: MIT Press, pp. 265–280.

Whorf, B. (1940b). Gestalt technique of stem composition in Shawnee. Prehistory Research Series, 1, 9. Reprinted in Whorf, B. (2012) *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. Second ed. Edited by J. Carroll, S. Levinson & P. Lee. Cambridge, MA: MIT Press, pp. 205–220.

Wierzbicka, A. (2005). There are no "color universals" but there are universals of visual semantics. *Anthropological Linguistics, 47,* 217–244.