The Development of Color Terms in Shipibo-Konibo Children

Danielle Kellier*[1], Martin Fortier*[2], Maria Fernández Flecha[3], & Michael C. Frank[4]

[1] University of Pennsylvania

[2] PSL Research University

[3] Pontificia Universidad Católica del Perú

[4] Stanford University

Abstract

Enter abstract here. Each new line herein must be indented, like this line.

Keywords: keywords

Word count: X

The Development of Color Terms in Shipibo-Konibo Children

TO BE PASTED FROM GOOGLE DOC

Introduction

Color language is where language and perception meet. Terms like *blue* or *red* draw boundary lines across a perceptually continuous space. In English, there are 11 basic color terms, but this color categorization is not universal. For instance, Russian speakers use two distinct words to describe the colors light blue ("goluboy") and dark blue ("siniy"); and some languages have as few as two words (e.g., the Jalé people only have terms for "light" and "dark"; Berlin & Kay, 1969). Why do languages vary in their color systems? One emerging consensus is that languages categorize the color spectrum in different ways in part due to functional demands (Gibson et al., 2017): both smaller and larger color systems are relatively optimal for suiting different communicative needs (Regier, Kay, & Khetarpal, 2007).

One important component of this hypothesis is the idea that some color systems are easier to learn for children than others; but the actual acquisition of color terms – while well-studied in English (e.g., Wagner, Dobkins, & Barner, 2013) – is extremely under-studied across other populations. Berlin & Kay's seminal World Color Survey (WCS; Kay, Berlin, Maffin, Merrifield, & Cook, 2009) presented adult speakers of over 100 languages with differently colored chips and asked them to produce a label, characterizing the space of color vocabulary in a range of written and unwritten languages. The WCS is an invaluable resource for the cross-linguistic study of color vocabulary, but no comparable resource exists for cross-cultural studies of how this vocabulary is learned across childhood.

In the current project, our goals were (1) to characterize color term knowledge in an indigenous population previously studied by the WCS, the Shipibo-Konibo (SK), and then (2) to build on this foundation to characterize the developmental trajectory of color language acquisition in a group of children raised outside of the WEIRD (Western Educated Industrialized Rich Democratic) populations that are over-represented in behavioral science.

Color in Amazonian languages and Latin American varieties of Spanish

A few studies explore the use of color terms in the varieties of Spanish in Latin America. Berlin and Kay (1969) examine the case of the Mexican dialect of Spanish, which they consider to be in Stage VII of their classification. So, for example, a Stage II system would add the term red to the colors already present in Stage I (*black* and *white*). It wouldn't be possible for a system to have *red* if it doesn't already have *black* and *white*. They identify the following basic color terms: *blanco* (white), *negro* (black), *rojo* (red), *verde* (green), *amarillo* (yellow), *azul* (blue), *café* (brown), *morado* (purple), *rosa* (pink), *anaranjado* (orange) and *gris* (grey). Also, based on their work with forty Tzeltal participants, both Tzeltal monolinguals as well as Tzeltal-Spanish bilinguals, they found that bilingualism did not skew their results regarding the existence of semantic universals in the domain of color vocabulary. Tzeltal has five basic color terms: *ʔihk´* (black), *sak* (white), *cah* (red), *yaš* (green) and *k´an* (yellow). This language is estimated to be transitioning from Stage IV to V, which is reflected in the ambiguity of the focus of *yaš* (grue). The authors posit that a long history of contact with Spanish has probably accentuated this, and suggest that exposure to Spanish in schools will eventually cause *yaš* to be entirely restricted to greens, and *azul* (or some other Spanish term) will be adopted into the Tzeltal color system.

Monroy and Custodio (1989) offers information on Colombian Spanish based on materials collected for the Linguistic-ethnographic Atlas of Colombia. He presents examples of ad hoc color terms referring to colors through objects prototypically instantiating these color: “vegetables”, “animals”, “food”, “metals”, “precious stones”, “fire and its derivatives” and “atmospheric phenomena”.

More recently, Aragón (2016) offers an ethnolinguistic study of color terms in Mexican Spanish: *amarillo* (yellow), *azul* (blue), *blanco* (white), *café* (literally "coffee", but effectively brown), *gris* (gray), *morado* (purple), *naranja* (orange), *negro* (black), *rojo* (red), *rosado* (pink) and *verde* (green). She analyzes the elaboration of these meanings in dictionaries, as well as the references and associations to which informants resort to for their own definitions. Aragón concludes that the local natural and cultural referents constitute a point of consensus among Mexicans when defining terms of color. Although informants also discussed some cultural material referents, these were not salient prototypes in their explanations. A special case that would merit further study in the future is that of *café* in Mexico versus *marrón* in Spain. According to the author, these two color terms are differentiated by the prototype "toasted coffee grain" associated to the Mexican Spanish term. Finally, she reviews the symbolic associations related to some terms, such as the discourses on femininity, especially those centered around the figure of the girl, associated with the term *rosado.*

Gibson et al. (2017) offer some approximations to the case of color terms in Bolivian Spanish, based on their analysis centered on Tsimane, an indigenous language spoken by a group living in the Amazonian piedmont. The authors compare the Tsimane case with Bolivian Spanish and American English. Compared to Bolivian Spanish and English, Tsimane exhibits greater variability in terms of the color terms used for all color chips presented in their study, with the exception of red. Out of a total of 80 color chips, Tsimane exhibits 8 modal color terms while English has 10, and Bolivian Spanish, 11. Also, despite the variability observed, the assignment of modal color terms resulted in a similar partition of the color space in the three languages assessed. The authors also emphasize that the Tsimane color system is less informative than the English and the Bolivian Spanish one. Finally, using the free choice paradigm, they show speakers of Bolivian Spanish extensively use the term *verde* (green) to denominate the color chips displayed, in addition to *celeste* (light-blue) and *azul* (blue), as well as *morado* (purple). Less frequent terms are, for example, *fucsia* (fuchsia), *guinda* (maroon) and *mostaza* (mustard).

Several indigenous Amazonian color systems have been studied in the WCS. One of them, Candoshi, has been further examined by Surrallés (2016). In this thought-provoking study, Surrallés suggests that no proper color term exists in this language. If the fieldworkers of the WCS found otherwise, it is only because they misidentified the elicited terms as color terms while they are nothing more than a series of ad hoc terms referring to objects or animals of the surrounding environment. For example, in Candoshi, the word for yellow is "*ptsiyaromashi*" ("like the feathers of a milvago bird"), the word for red is "*chobiapi*" ("ripe fruit"), the word for green is "*kamachpa*" ("unripe fruit"), etc. These findings lead Surrallés to argue that the Candoshi do not have a proper color system. When they use "color terms" they are not trying to subsume objects of the world under abstract color categories, but they are rather establishing horizontal and ad hoc comparisons between similar objects of the world.

A similar criticism of the WCS approach had been previously developed by Everett (2005, pp. 627–628) based on his study of Pirahã, another Amazonian language. Everett also rejects the idea that there are basic color terms in this language. He argues that the four color terms identified as basic in the WCS are not such. For example, the word identified as the basic color term for "red" and "yellow" (*bi i sai*) means nothing more than "bloodlike". Here again, color terms seem to be ad hoc comparisons rather than proper basic terms.

As mentioned earlier, SK color terms have been thoroughly studied in the WCS. It is worth mentioning that two anthropological studies (Morin, 1973; Tournon, 2002) have also investigated the color terms used in this Amazonian language. However, these two studies contain some serious methodological pitfalls: a very limited number of chips were tested with only a few participants. As a result, we will not further discuss these studies in the remaining of this article and will only focus on a comparison with the WCS data.

In sum, while some dialectical differences can be noticed across varieties of Spanish, these slight variations are consistent with the general framework proposed by the WCS. Less consistent, however, is the recurrent finding that ad hoc terms seem to play a central role in Amazonian color systems – and possibly also in some South-American varieties of Spanish (such as Colombian Spanish). More broadly, it seems that Amazonian color systems are characterized by fewer color terms than dialectical Spanish systems.

The Development of Color Vocabulary

The Current Study

In the last two decades, cross-cultural research aiming to go beyond North-American "convenience samples" has mainly focused on the study of East Asian children and adults. This endeavor has proved very fruitful (Kitayama & Cohen, 2007) but is still limited because of its almost exclusive focus on North-American vs. East-Asian samples. The current study contributes to the general effort to go beyond such samples and study the development of human cognition in a non-North American and non-East Asian context. The SK people are an indigenous group located within the Peruvian Amazon. They are mainly horticulturalists, fishermen, occasionally hunters but are noted for their strong display of tradition despite increasingly regular interactions with the western world. Their children receive formal schooling for 4 hours a day and begin formal Spanish lessons closer to adolescence. Most SK adults have some grasp of Spanish but younger adults show more proficiency than elders.

The SK indigenous people are particularly interesting for at least two reasons: They differ from samples usually studied by cross-cultural evolutionary psychologists (Apicella & Barrett, 2016). Indeed, evolutionary psychologists are particularly interested in the study of contemporary hunter-gatherers because they are believed to be a good model of our Pleistocene ancestors. By contrast, like most riverine Amazonian cultures, the SK culture is not based on hunting and gathering, but on horticulture, fishing, and to a limited extent, hunting.

Because of their location on the Ucayali River, one of the main tributaries of the Amazon, the SK culture has always been enmeshed in rich trading networks involving other indigenous groups of the Andes and the Lowlands (in pre-conquest times) as well as Mestizos and Westerners (in post-conquest times) (Lathrap, 1970). It would thus be mistaken to think of this culture as an "isolated" or "preserved" one. On the contrary, having been extensively exposed to numerous cultural influences, the SK culture has been constantly reworked and reshaped through the centuries. This was especially true in the second half of the 20th century with intense contact with the Spanish-speaking Mestizo populations established along the Ucayali River. As a result, today's SK culture straddles two worlds.

Study 1

Before we could assess the developmental trajectory of color term knowledge in SK children, our goal was to replicate and update the characterization of the adult SK color system given by the World Color Survey. As the WCS study took place generations prior, we could not assume the SK color term mappings had remained static especially through years of industrialization and exposure to the Spanish language and its own color term system. As such, Study 1 used a modified version of the original WCS protocol, with an identical color chip set. The goals were to characterize the current SK vocabulary and to generate a standard of adult knowledge against which subsequent child participants could be scored.

Methods

Participants. Our protocol for Study 1 and all subsequent studies received ethical approval from Pontificia Universidad Católica del Perú's Institutional Review Board. We recruited 39 adult participants (7 men). We experienced difficulty recruiting male participants as many of the men were away from the village during the day, resulting in a sample that is predominantly female. Most participants (31, 4 men) were from SK communities of the Middle Ucayali region (Yarinacocha, San Francisco, and Nueva Betania), with a subset from communities of the Lower (Paoyhan) and Upper (Puerto Belén) Ucayali region. Within the small town of Yarinacocha (in the vicinity of Pucallpa), we recruited participants (9, 2 men) from Bena Jema, a predominantly SK neighborhood. All the other recruitment sites were native community villages with exclusively SK residents. Overall, the sample included SK adults who could be characterized as more urban (Yarinacocha and San Francisco sites) or more traditional and in regular contact with the surrounding rainforest (Nueva Betania, Paoyhan, and Puerto Belén sites).

The median age for participants was 38 years (IQR = 26-48) ranging from 20 to 64 years. Regarding occupations, 41% of the 32 female participants were homemakers or housewives (33% of the overall sample) and another 41% were artisans (33%). Three of the 7 male participants (43%, 8% overall) were horticulturalists. Across both sexes, 5 women (16%, 13% overall) and 3 men (43%, 8% overall) identified as students, comprising a total of 21% of the population. Although all adult participants were required to be native SK speakers, all were introduced to the Spanish language prior to adolescence (median age = 8yo, IQR = 5-10).

Materials and procedure. Similar to the original WCS, we used a set of 330 Munsell color chips and asked participants to name these (Berlin & Kay, 1969). We made a number of changes to the procedure, however. In the WCS, every participant provided terms for all 330 chips. Due to fear of participant fatigue, we split up color chips based on their ID numbers (even or odd) and participants were randomly assigned work with either even- or odd-numbered color chips. As a result, each participant was presented with only 165 chips. All 330 hues within the set are visualized in Appendix 1. Dimensions of the chips were 2 cm × 2.5 cm.

First, the experimenter explained the general procedure and goals of the study to a participant. The experimenter would then present a single color chip to the participant and ask in SK: “What is the color of this chip?” The study was conducted solely in SK language with the assistance of a bilingual SK- and Spanish-speaking research assistant. It should be noted that although the experiment was conducted in SK, the SK word for color used is identical to the Spanish word *color* (an example of SK speakers adopting Spanish words into their lexicon), which might have encouraged Spanish language use.

Besides the reduction in set size, our procedure also differed from that of WCS (see Kay et al., 2009, pp. 585–591) in other aspects. Participants sat in front of the experimenter. To manage changes in natural light intensity between participants, the experiment took place indoors near a window or a door instead of outdoors. Another difference between our study and the WCS procedure is in our approach for encouraging participants to describe chips using BCTs. In the WCS, the experimenter would instruct participants to only provide BCTs during the task (e.g., describing a chip as “blue” as opposed to “navy blue” or “sky-like”). However, we had difficulties concisely explaining the concept of a BCT compared to other terms[1]. We decided to allow participants to describe a chip with any term they wished, and to ask further questions to elicit a BCT when they did not do so on their first try. For example, when presented with a red color chip, the participant might use the term “blood-like” (a non-BCT). The experimenter would ask: “Do you know of any other word to refer to the color of this chip?”. Should the participant subsequently respond with “dark red” (another non-BCT), the experimenter would further

[1] Indeed, as Kay et al. (2009, pp. 587–589) acknowledge, there is no straightforward necessary and sufficient criteria for the basicness of a color term (cf Levinson, 2000).

ask: “How would you refer to this color with only one word?” Eventually, the participant might use the term “red” (a BCT). For some chips, participants provided a BCT as their first description. For others, a BCT might be preceded by 1 or 2 non-BCTs. When participants failed to provide a BCT after 3 attempts (i.e., two follow-up questions), no further questions were asked, and the experimenter moved on to the next chip. All responses, BCT or not, were recorded in the order produced by the participant.

Results and Discussion

All participants used the following set of color terms to describe at least once during their session: “joxo” (light/white), “wiso” (dark/black), “panshin” (yellow), “joshin” (red), and “yankon” (green/blue). Given the widespread use of this term set and their interpretations, we will refer to these as SK-language BCTs. Most (79%) participants also described at least 1 chip as “manxan” (faded), referring to a chip’s saturation. In terms of overall popularity, participants described a median of 32% of chips as “yankon” ($IQR =$ 26-39%) followed by “joshin” ($Mdn =$ 10%, $IQR =$ 7-16%), “joxo” (9%, 6-15%), “panshin” (10%, 6-12%), “manxan” (6%, 1-10%), and “wiso” (5%, 3-8%). We failed to find any significant sex differences in the overall spread of color term usage across chip set ($t(59) = 0.00$, $p > .999$) or in the proportion of subjects who used a term at least once during their session ($t(117.95) = -0.38$, $p = .706$).

Compared to the WCS dataset which only reported SK language terms, 59% of our participants used a Spanish-language color term to describe at least 1 chip, which accounted for 4% of all responses (Figure @ref(fig:study1-corrected_entropy)). Across chips, Spanish use peaked at 55% when participants were asked to label chips that English speakers would consider to be orange. However, Spanish use varied greatly between subjects ($Mdn =$ 1%, $IQR =$ 0-4%) with one participant responding in Spanish in 71% of the time despite being prompted solely in SK.

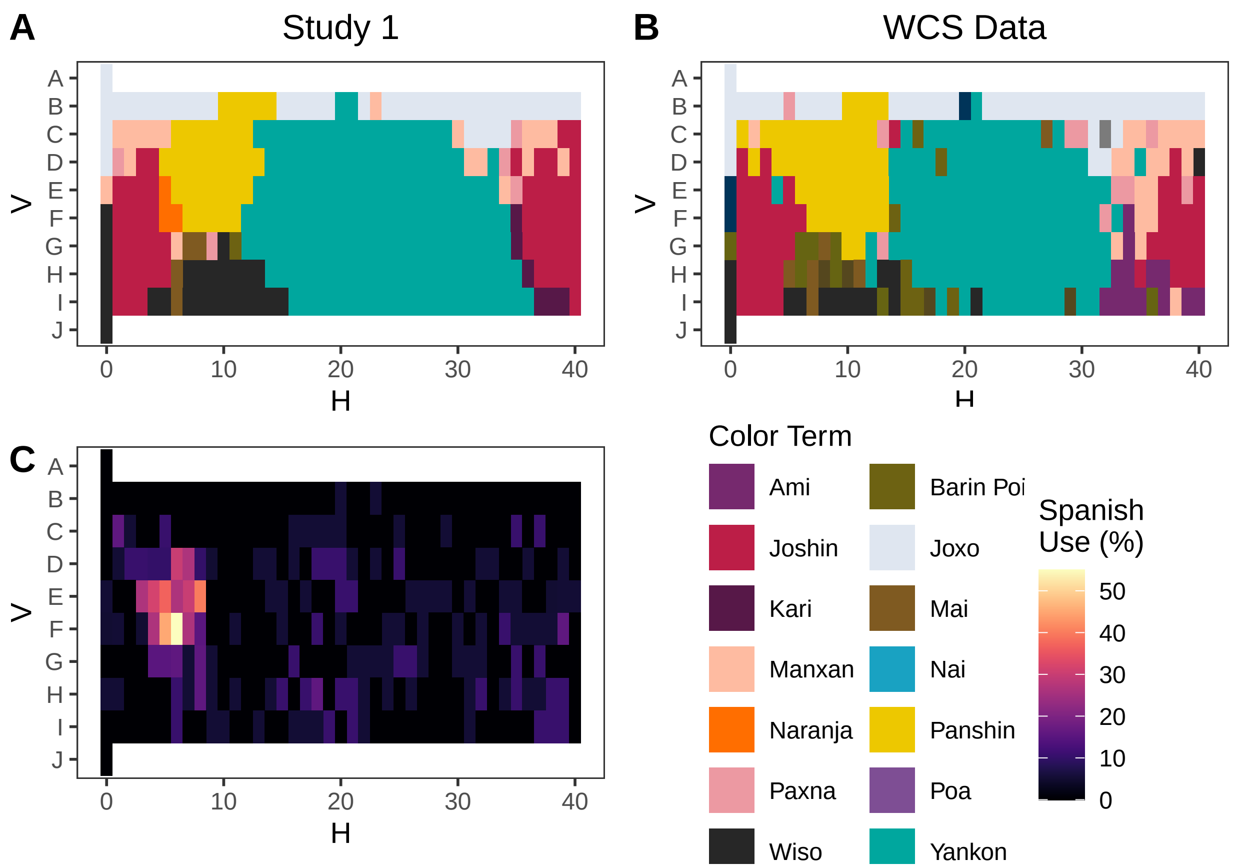
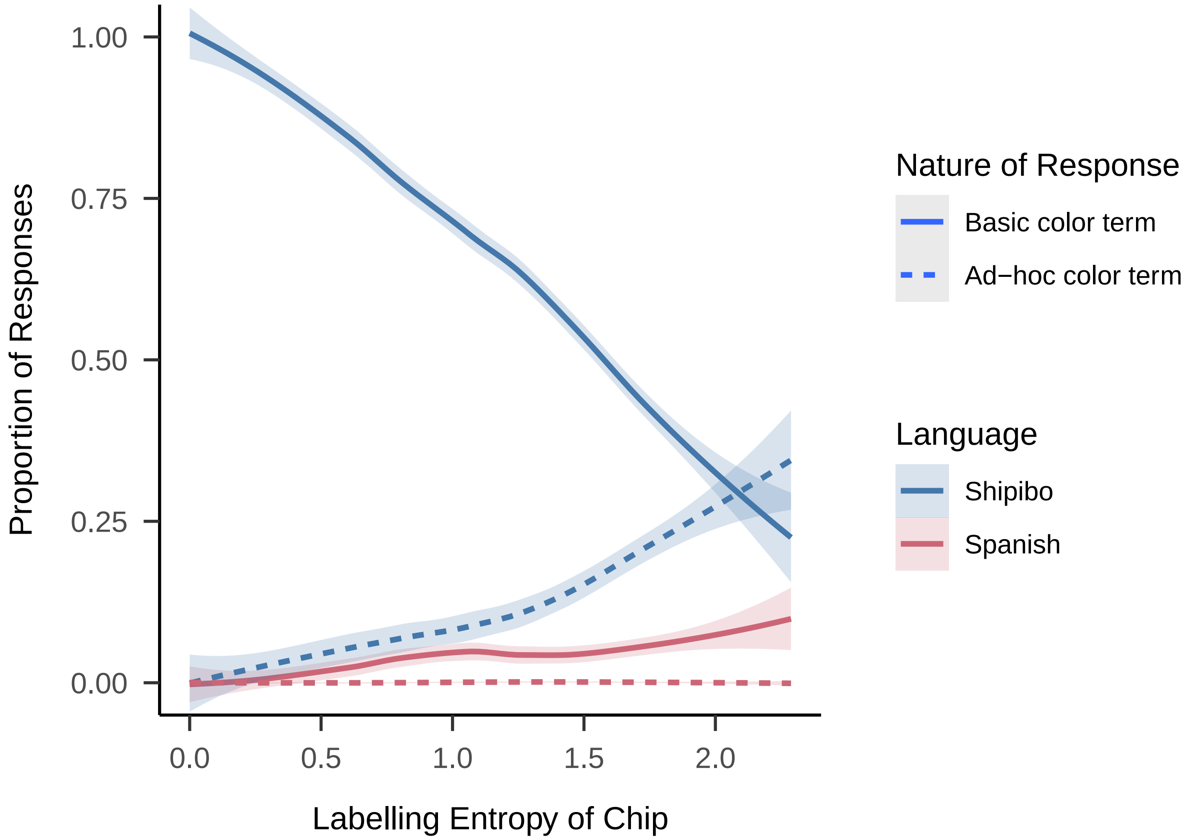


Figure 1. (A and B) Plots of the modal term given for a particular chip. Color coordinates were represented in 2-D Munsell space. Modal responses were given by SK adults during (A) the original World Color Survey and during (B) our Study 1. (C) Heat map of prevalence of Spanish-language responses during Study 1. Legends for all three subplots located in the bottom-right quadrant.

Participants used an SK-language BCT (i.e., "yankon") to describe a median of 65% of chips (IQR = 53-90%). Besides BCTs, 59% of participants used SK-language ad hoc hue terms (i.e., "nai" or sky for blue chips) for an overall median of 6% of chips (IQR = 0-19%). SK-language terms referring to saturation or luminosity of a chip, such as "manxan" (faded) were used for an overall median of 13% of chips (IQR = 6-20%). Most instances (91%) of Spanish use involved a Spanish BCT such as "rojo" (overall Mdn = 1%, IQR = 0-4%). In other words, participants responded in Spanish to label chips with basic color categories but mostly relied on SK for all other descriptor types.



Study 2

After generating an updated SK color term map using the responses from adult participants in Study 1, we created Study 2 to assess child participants' production and comprehension of SK color terms. Because we did not think that we could feasibly ask children across a range of ages about more than 100 color chips, we selected a subset of chips representing the prototypical instances for prominent SK terms from Study 1.

Methods

Participants. Fifty-seven children (23 boys) ages 5- to 11-years-old were recruited in predominantly SK neighborhoods in Yarinacocha (Nueva Era and Bena Jema) and in Bawanisho, a native community settled along the Ucayali River, more than 500 kilometers

Table 1

Demographics of participants in Studies 2 and 3.

Age Group	n	Boys
	Study 1	
5	3 (5% of overall sample)	1
6	8 (14%)	3
7	12 (21%)	4
8	15 (26%)	5
9	10 (18%)	5
10	4 (7%)	2
11	5 (9%)	3
	Study 2	
5	2 (4% of overall sample)	1
6	2 (4%)	0
7	11 (24%)	4
8	9 (20%)	1
9	11 (24%)	4
10	8 (17%)	3
11	3 (7%)	3

southeast of Pucallpa. Recruitment occurred either through direct contact with interested parents or through their local school. If recruited via school, consent for participation had to be given by both teacher and parent. Outside of the school environment, consent was only given by the parent.

Materials and procedure. Based on the findings of Study 1, we chose 8 color chips from our original set of 330 to serve as prototypical instances of major SK color terms. These color chips were blue (WCS n°1), green (n°234), red (n°245), white (n°274), yellow (n°297), black (n°312), greeny-yellow (WCS n°320), and purple (WCS n°325) (see Appendix 1). Study 2 was conducted entirely in SK and participants were explicitly instructed to give responses in SK as opposed to Spanish. In the production and comprehension tasks, children sat at a table across from the experimenter with color chips arranged between them. The production task was always performed before the comprehension task.

Production task. Similar to Study 1, the experimenter introduced a participant to the general procedure and the goals of the study. The experimenter would then ask: "What is the color of this chip?" As in Study 1, we used follow-up questions to elicit a BCT when the child's initial response was not a BCT. In a departure from Study 1, we were more explicit in soliciting an SK-language response. When a participant provided a Spanish-language term, the experimenter would record their response but further ask: "What is the name of this color in SK?" If a participant could not respond with an SK term, the experimenter would not ask further questions and would move forward to the next chip. As a result, some children could only produce SK non-BCTs or Spanish-language terms for particular chips.

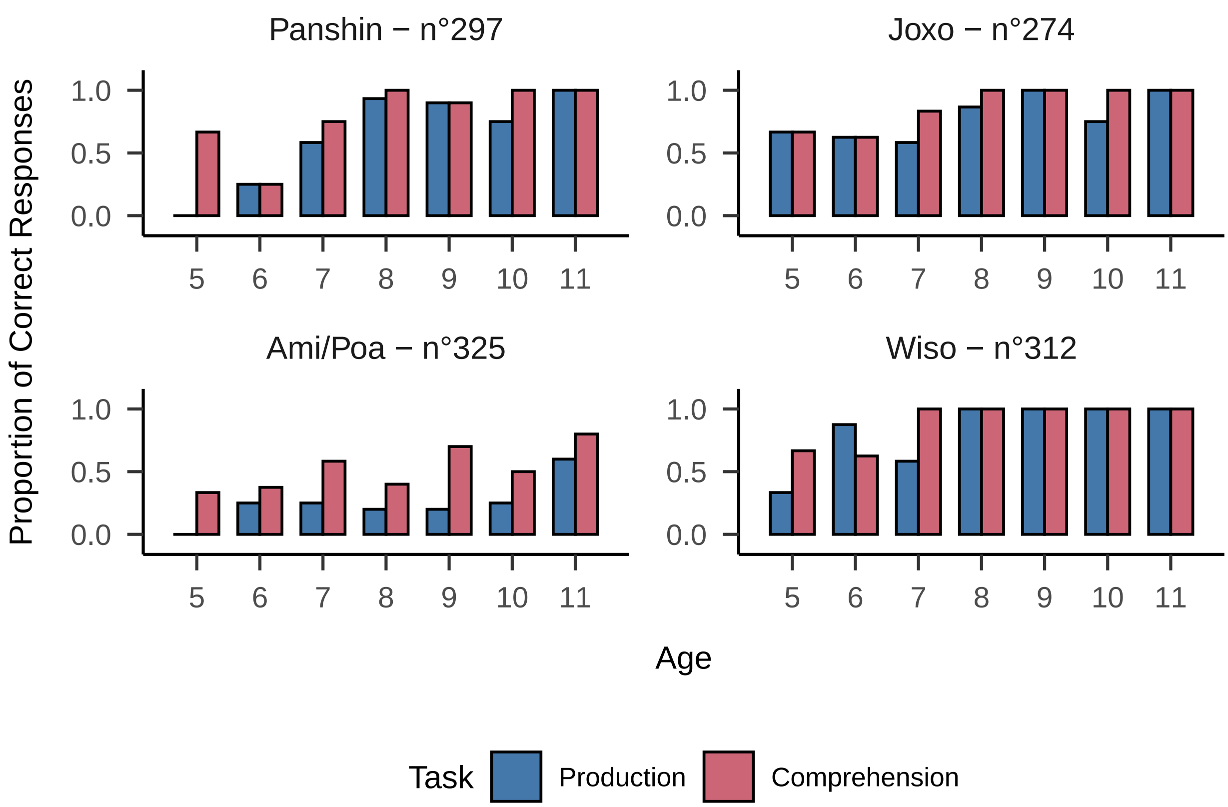
Comprehension task. The comprehension task had a notably different procedure compared to the preceding production task or that of Study 1. We tested the comprehension of 9 SK color terms. The choice of these terms was based on common responses given by adult participants in Study 1. The color term prompts included BCTs:

yankon ("green/blue"), *joshin* ("red"), *panshin* ("yellow"), *joxo* ("white/light"), *wiso* ("black/dark"). We also included non-basic but prominent terms as prompts which were *nai* ("blue/sky"), and *barin poi* ("greenish-yellow") and two dyads of non-basic terms *pei/xo* ("green") and *ami/pua* ("purple"). Children sat at a table across from the experimenter with the 8 color chips of the production task displayed between them. The experimenter asked: "Can you give me the [*color term*] chip?" Participants chose one of the 8 chips and their response was recorded.

Our findings from Study 1 suggested that color terms varied in their degrees of specificity. For example, *wiso* best describes a narrow range of very dark to black. By contrast, *yankon* could encompass blue, green, greenish-yellow, and purple; *joshin* could describe red, purple, and orange; *pei* or *xo* to green or greenish-yellow. In cases where a term could apply for more than one chip (i.e., *yankon*), the initially selected chip would be removed from the table, leaving 7 remaining chips. The experimenter would then ask: "Can you give me another [*color term*] chip?" The participant would then pick another one of the 7 chips, have their response recorded, and so on. We prompted participants 4 times for *yankon* and 2 times each for *joshin* and *pei/xo*; every other term only received a single prompt. Due to the inherent ambiguity in term-hue pairings, accuracy for a child participant was coded based on adult responses given during Study 1. If at least 15% of adult participants in Study 1 associated a chip with a particular term, we coded a similar term-chip pairing from a child participant as correct. Some trials could have multiple pairings, accuracy was scored as an average, rather than dichotomous. For instance, if a child correctly chose 3 out of 4 chips for the "yankon" trial, instead of 1 (correct) or 0 (incorrect) they would receive a score of 0.75.

Table 2

Term	$\hat{\beta}$	95% CI	z	p
Intercept	-2.87	[-3.70, -2.05]	-6.85	$<$ .001
Age years c	-0.52	[-1.01, -0.03]	-2.09	.036
Entropy	1.54	[1.02, 2.06]	5.82	$<$ .001
Entropy	0.34	[0.03, 0.66]	2.13	.034



Results and Discussion

Older children were more accurate in both production and comprehension compared to younger children (Figure 2). This produced a significant developmental projection for

Table 3

Term	$\hat{\beta}$	95% CI	z	p
Intercept	0.67	[-0.12, 1.46]	1.67	.095
Age years c	-0.02	[-0.49, 0.45]	-0.08	.939
Entropy	-1.49	[-2.07, -0.92]	-5.10	< .001

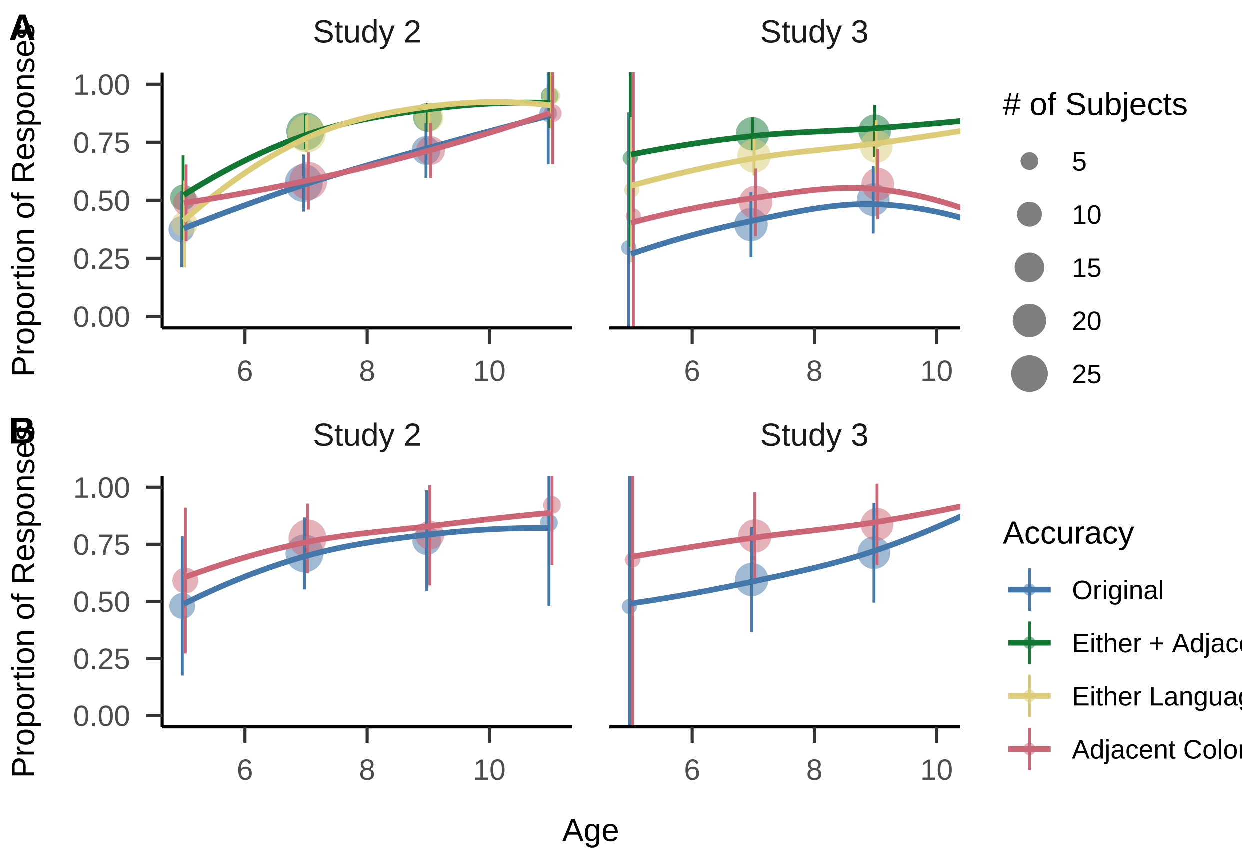


Figure 2. Proportion of accurate responses when applying different accuracy criteria, by age and study. Points show the mean for a 2-year age group (chosen arbitrarily for visualization) with 95% confidence intervals. Lines show a loess smoothing function. Points show the mean for a 2-year age group (chosen arbitrarily for visualization) with 95% confidence intervals.

accuracy of term-chip pairings in both production ($z = 3.74$, $p < .001$) and comprehension ($z = 3.27$, $p = .001$). For some term-chip pairings such as ami/pua and pei/xo, children performed failed to produce the correct term in the production task but improved significantly during the comprehension task (Figure **??**). It is possible that for children's color term knowledge, comprehension precedes production. It is also possible that, given that the comprehension task always followed the production task, children were able to pick up on their errors and update their color term mapping in real-time. However, given that the experimenter did not provide feedback on accuracy during sessions, the former explanation seems more likely.

Language switching. Over a quarter (28%) of all responses were given in Spanish, despite children being prompted solely in SK (i.e., labeling a panshin chip as "amarillo"). The distribution of Spanish responses was non-random, with median use in 2/8 trials (*IQR* = 0-5). We failed to find a significant correlation between age and overall Spanish use throughout the production task ($t(55) = -1.13$, $p = .263$).

As an exploratory analysis, we attempted to quantify low naming consensus using naming entropy (following Gibson et al., 2017). We computed the naming entropy for each chip by computing the probabilities for each chip c to be named with a particular label l ($p(l \mid c)$) and then taking $H(c) = -\sum p(l \mid c) \log[p(l \mid c)]$ (see inset entropy values by chip in Figure **??**). To assess the hypothesis that naming entropy in adults was related to Spanish use in children, we fit a mixed effects model predicting Spanish responses as a function of child age, entropy of the chip's naming distribution for adults in Study 1, and their interaction. Despite age not being significantly correlated with overall frequency of Spanish responses, when added to a mixed-effects model, we found that older children were less likely to respond in Spanish (Table 2). Older children were also more likely to respond in Spanish when presented with a chip with high entropy (low naming consensus) among adult participants in Study 1. We found a positive interaction between age and entropy.

Overextensions. One reason to use Spanish would be if children fail to recall the proper SK color term but do know the proper mapping in the Spanish. But another possibility is that children may have more imprecise representations and choose to respond with a same-language but adjacent color term (i.e., labeling a *panshin*-colored chip as "joshin"). Following Wagner et al. (2013), we aggregated across color chips and examine the pattern of children's first responses, categorizing them as same-language, adjacent, and different-language. Using a mixed-effects model, we found that younger children were more likely to respond with SK-language but adjacent terms ($\hat{\beta} = -1.16$, 95% CI $[-1.96, -0.35]$) but chip entropy did not have a significant factor in this strategy ($\hat{\beta} = -1.23$, 95% CI $[-3.00, 0.54]$).

If children fail to recall the proper color term in SK, but do know the proper mapping in the Spanish color system, language-switching may be an appropriate strategy. Should they lack knowledge of the corresponding Spanish color term, they may also choose to respond with a same-language but adjacent color term. If we allow for more leniency in scoring–accepting same-language but adjacent or different-language but corresponding responses–we can check for more subtlety surrounding color term mapping (Figure 2). Using a mixed-effects model, we found a significant improvement in accuracy scores when we allowed different-language but corresponding responses ($z = 6.98$, $p < .001$) but no significant change when allowing for same-language but adjacent responses ($z = 0.86$, $p = .390$). This increase in accuracy with language-switching strengthened with age ($z = 2.74$, $p = .006$).

Study 3

Noting the apparent strategy of language switching seen in Study 2, we designed Study 3 as its complement. Here, we tested children's production and comprehension of Spanish color terms with a similar protocol to Study 2, albeit with a subset of chips representing prototypical colors for the Spanish color system.

Participants. Similar to Study 2, 46 children (16 boys) ages 5- to 11-years-old were recruited from the neighborhood of Bena Jema in Yarinacocha and from Bawanisho. Recruitment occurred either through interested parents or a local school. With consent collected from parents and, if in a school environment, teachers as well.

Materials and procedure. Based on Study 1 and on previous studies of Spanish color systems (XXX et al), we selected 11 color chips to serve as prototypical instances of prominent Peruvian Spanish color terms. These color chips included 6 also used during Study 2: green (n°234), red (n°245), white (n°274), yellow (n°297), black (n°312) , and purple (n°325). Five additional chips were selected: gray (WCS n°46), pink (n°65), orange (n°121), brown (n°266), and blue (n°291) (see Appendix 1). The blue chips differed between Studies 2 and 3 as we decided that the prototypical hues for *yankon* and *azul* differed enough to warrant the use of a different chip.

As SK children are not very fluent in Spanish (formal Spanish education occurs in adolescence), the production and comprehension tasks were both conducted in SK, and Spanish was only used for color terms (i.e., Spanish color terms were embedded within otherwise SK sentences). In both tasks, a participant would sit at a table across from the experimenter with 11 color chips in front. As in Study 2, the production task was always performed prior to the comprehension task.

Production task. The procedure was similar to that of both Studies 1 and 2. The experimenter would introduce a participant to the general procedure and aims of the study. Despite much of the study being conducted in SK, the experimenter would specify that participants would be expected to provide color terms in Spanish. The experimenter would then ask: “What is the color of this chip?”. If the participant responded in SK, the experimenter would record their response but further ask: “What is the name of this color in Spanish?”. If a participant responded with “I don’t know” to this prompt, the experimenter would not prompt any further and would move forward to the next chip. As a result, some responses lack Spanish-language BCTs and only consist of non-basic and/or SK color terms. In total, we collected production data for 11 color chips. For each chip, the data include either one response (when children provided a Spanish basic color term in the first trial) or two or three responses (when children's initial responses were either non-basic and/or in SK).

Comprehension task. The procedure was similar to that of Study 2. The experimenter would ask: Can you give me the [*color term*] chip? For 11 Spanish color terms. The choice of these terms was based on both previous studies examining Spanish color terms as well as responses given by adult participants in Study 1 (as some adult participants used Spanish color terms to label particular color chips). The 11 terms used as prompts were *blanco* ("white"), *verde* ("green"), *rojo* ("red"), *amarillo* ("yellow"), *azul* ("blue"), *negro* ("black"), *naranja* ("orange"), *gris* ("grey"), *morado* ("purple"), *marrón* ("brown"), and *rosa* ("pink"). Since each color term was best instantiated by a single color chip and lacked the ambiguity seen with certain SK color terms, accuracy was graded one-to-one for term-chip pairings with less leniency compared to Study 2.

Results and Discussion

Contrasting Study 2, we found age to be a significant predictor of accuracy in the comprehension task ($z = 2.90$, $p = .004$), but not in the production task ($z = 1.61$, $p = .108$, see Figure 2). Similar to Study 2, over a quarter (30%) of all responses were given in SK, despite being prompted to respond in Spanish. There was significant variation in language-switching with some children responding solely in Spanish while others responded to upwards of 9/11 trials in SK ($Mdn = 5$ trials, $IQR = 1.25$-6). We found only a marginal correlation between age and label accuracy ($t(44) = 1.91$, $p = .063$) and an insignificant correlation between age and language-switching ($t(44) = 0.44$, $p = .663$). Still, when accounting for age, we found that participants tended to respond in SK when presented with items that had low entropy in Study 1 ($z = -5.10$, $p < .001$). This suggests that child participants across Studies 2 and 3 preferred to respond in SK when presented with a high-consensus chip, and in Spanish with a low-consensus chip.

Similar to Study 2, we adopted alternative scoring to accommodate language-switching from Spanish to SK (different-language) but corresponding and same-language but adjacent responses. Using a mixed-effects model, we failed to find age as a significant predictor for accuracy even with this more lenient scoring ($z = 1.45$, $p = .146$), in concordance with earlier analyses. However, we did find that participants improved accuracy through both mapping strategies of either providing different-language but corresponding responses ($z = 10.51$, $p < .001$) or same-language but adjacent responses ($z = 3.29$, $p = .001$). We find frequent use of language switching in both Studies 2 and 3, but only Study 3 exhibits significant use of same-language but adjacent terms.

It is possible that early but informal Spanish language exposure can explain the discrepancies seen in Studies 2 and 3. Children may be exposed to Spanish at a young age but likely do not receive any formal Spanish education until adolescence. With limited knowledge of Spanish color terms, children may spontaneously supplement their color term knowledge with Spanish terms during SK-language Study 2 but struggle to succeed during Spanish-language Study 3. More generally, we see children relying on a mixture of strategies to communicate colors even in the absence of mastery in either language.

General Discussion

TO BE PASTED FROM GOOGLE DOC

Summary of study. Adult data - bilingualism and relation to WCS data When we turned to the children's data, two important generalizations emerged. First, we observed a much longer developmental trajectory for color than is observed in modern US populations (cf. Bornstein, 1985). Second, we found evidence for competition between the Shipibo and Spanish color systems, implying the potential for functionally-driven language change.

Gibson analysis of optimality: children use spanish words for low-consensus chips, SK words for high consensus chips. These support the optimality hypothesis Limitations of our work. Cross-sectional Limited number of chips for kids (limits entropy analyses)

In sum, these data further support a model of color word knowledge and acquisition that is driven by communicative need. Need for more developmental work on other languages. Huge role for environmental input

References

Apicella, C. L., & Barrett, H. C. (2016). Cross-cultural evolutionary psychology. *Current Opinion in Psychology*, *7*, 92–97.

Aragón, K. (2016). Color language and color categorization. In G. Paulsen, M. Uusküla, & J. Brindle (Eds.). Cambridge Scholars.

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

Everett, D. L. (2005). Cultural constraints on grammar and cognition in pirahã another look at the design features of human language. *Current Anthropology*, *46*(4), 621–646.

Gibson, E., Futrell, R., Jara-Ettinger, J., Mahowald, K., Bergen, L., Ratnasingam, S., … Conway, B. (2017). Color naming across languages reflects color use. *Proceedings of the National Academy of Sciences*, *114*(40), 10785–10790.

Kay, P., Berlin, B., Maffin, L., Merrifield, W. R., & Cook, R. (2009). *The world color survey*. Stanford, CA: Center for the Study of Language; Information.

Kitayama, S., & Cohen, D. (2007). *Handbook of cultural psychology*. Guilford Press.

Lathrap, D. W. (1970). *The upper amazon*. Thames; Hudson.

Levinson, S. C. (2000). Yeli dnye and the theory of basic color terms. *Journal of Linguistic Anthropology*, *10*(1), 3–55.

Monroy, M., & Custodio, S. (1989). Algunos usos de los términos del color en el español de colombia. *Thesaurus; Bogotà*, *44*(2), 441.

Morin, E. (1973). *Le paradigme perdu: La nature humaine*. Éditions du Seuil.

Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, *104*(4), 1436–1441.

Surrallés, A. (2016). On contrastive perception and ineffability: Assessing sensory experience without colour terms in an amazonian society. *Journal of the Royal Anthropological Institute*, *22*(4), 962–979.

Tournon, J. (2002). *La merma magica: Vida e historia de los shipibo-conibo del ucayali.* Centro Amazonico de Antropologia Yaplicacion.

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