STUDIES IN FUNCTIONAL AND STRUCTURAL LINGUISTICS 78

Lexicalization patterns in color naming

A cross-linguistic perspective

EDITED BY
Ida Raffaelli,
Daniela Katunar
and Barbara Kerovec

John Benjamins Publishing Company

Copyright 2019. John Benjamins Publishing Company. All rights reserved. May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.

EBSCO Publishing : eBook Collection (EBSCOhost) - printed on 2/10/2023 1:40 AM via
AN: 2294306 ; Ida Raffaelli, Daniela Katunar, Barbara Kerovec.; Lexicalization Patterns in Color Naming : A Cross-linguistic Perspective
Account: ns335141

Lexicalization patterns in color naming

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Studies in Functional and Structural Linguistics (SFSL)

ISSN 1385-7916

Taking the broadest and most general definitions of the terms functional and structural, this series aims to present linguistic and interdisciplinary research that relates language structure – at any level of analysis from phonology to discourse – to broader functional considerations, whether cognitive, communicative, pragmatic or sociocultural. Preference will be given to studies that focus on data from actual discourse, whether speech, writing or other nonvocal medium.
The series was formerly known as *Linguistic & Literary Studies in Eastern Europe (LLSEE)*.

For an overview of all books published in this series, please see
benjamins.com/catalog/sfsl

Founding Editor

John Odmark

Honorary Editor

Eva Hajičová
Charles University

General Editors

Yishai Tobin
Ben-Gurion University of the Negev

Bob de Jonge
Groningen University

Editorial Board

Alexandra Y. Aikhenvald
La Trobe University

Joan L. Bybee
University of New Mexico

Ellen Contini-Morava
University of Virginia

Nicholas Evans
University of Melbourne

Victor A. Friedman
University of Chicago

Anatoly Liberman
University of Minnesota

James A. Matisoff
University of California, Berkeley

Jim Miller
Emeritus, University of Edinburgh

Marianne Mithun
University of California, at Santa Barbara

Lawrence J. Raphael
CUNY and Adelphi University

Olga Mišeska Tomić
Leiden University

Olga T. Yokoyama
UCLA

Volume 78

Lexicalization patterns in color naming. A cross-linguistic perspective
Edited by Ida Raffaelli, Daniela Katunar and Barbara Kerovec

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lexicalization patterns in color naming

A cross-linguistic perspective

Edited by

Ida Raffaelli
Daniela Katunar
Barbara Kerovec
University of Zagreb

John Benjamins Publishing Company
Amsterdam / Philadelphia

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/sfsl.78

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2019027957 (PRINT) / 2019027958 (E-BOOK)

ISBN 978 90 272 0403 5 (HB)
ISBN 978 90 272 6212 7 (E-BOOK)

© 2019 – John Benjamins B.V.
No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Company · www.benjamins.com

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table of contents

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Introduction

Ida Raffaelli, Daniela Katunar and Barbara Kerovec
University of Zagreb

1. Colors in the universalist and relativist framework

Color is one of the most often mentioned and cited lexical domains in linguistics and other sciences thanks to the revolutionary work done by Berlin and Kay (1969) and E. Rosch (1972, 1973).

As emphasized by Regier, Kay, and Khetarpal (2007), the recent debate about color naming has been dominated by two major views. One is that color categories are organized around universal foci, while the other is that color categories are determined at their boundaries by linguistic convention. The first view is universalist and the second one is relativist. Proponents of the universalist approach argue for the existence of universal focal colors. According to this view, focal colors are linguistically and/or culturally independent, and are thus not related to the differences between languages or the cultures in which these languages are spoken (cf. Regier, Kay, & Khetarpal, 2007).

The main goal of the universalist approach is to determine the origin of the categorization principles in the demarcation of color terms (e.g., neurological, visual, or memory based).

On the other hand, the proponents of the relativist approach claim that the language structure dictates the partition of the color spectrum, as well as that the boundaries between color areas are connected to local linguistic conventions and can differ in relation to language diversity. One of the major relativist arguments that is sometimes obscured in the experimentally-based universalist approach is the question of what really constitutes an experientially-based prototype of a color, as well as the knowledge of the concept of 'color'. Such a view is advocated by Wierzbicka (2005) on examples from diverse languages, where she argues that 'color' itself is not a universal concept. According to Wierzbicka, the search for color universals should be framed around the concept of 'seeing' and the common mechanisms used to describe what is seen (such as "X looks like Y", high and low visibility, and daytime and nighttime vision). In the examples in this volume, such

https://doi.org/10.1075/sfsl.78.01raf
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

things are often lexicalized systematically by using hue or brightness modifiers, or the OBJECT FOR COLOR metonymy. Furthermore, what the relativist view emphasizes is the need to add an "emic perspective" in the description of the color vocabulary. In Wierzbicka's terms (2005, p. 226), this means that what matters from the linguistic point of view is how people think about what they see and anchor it conceptually, in something that can be shared with others – with socially shared prototypes.

Though the universalist-relativist debate is far from resolved, one way of approaching it is presented in this volume via the notion of *lexicalization patterns*. As Levin and Rappaport Hovav (2019) point out, the study of lexicalization patterns has provided fruitful testing ground with respect to the challenges presented by Sapir-Whorf's view of linguistic relativity, e.g., in psycholinguistic research. Lexicalization patterns are relevant in this respect because they paint a picture of constrained relativity, i.e., an opportunity to establish a somewhat finite number of variants in a particular domain.

2. Lexicalization patterns in the relativist approach to color terms

Although many studies have been conducted showing different perspectives and understandings of how colors are categorized and named, there are still many gaps left in the description of the color lexicon and the aim of this book is to fill some of them. A common interest of various studies collected in this volume is to point to the diversity of language mechanisms (phonological, morphological, syntactic, and semantic) used in color naming, i.e., the formation of both focal and non-focal color terms. This is the reason all the studies in the book are oriented towards lexicalization patterns as language structures operative in color naming. As pointed out by Brdar (2017, p. 3), lexicalization *per se* is a notoriously polysemous notion with several layers of more general or more specific meanings. In one of its specific senses, it means something like clothing one's concepts with linguistic labels. A very broad and inclusive understanding of lexicalization is presented by Blank (2001, p. 1596), where it is closely tied to the notion of conventionalization, which on the level of the lexicon includes word formation and idioms, but also semantic change, borrowings, and some other types of lexical change.

The term *lexicalization patterns* as such has been by and large tied to the work of Leonard Talmy and has a more specific definition. Talmy (1985, p. 57) defines lexicalization patterns as a relation between meaning and surface expressions, and investigates which semantic elements are expressed by which surface elements. Semantic elements of different types can be expressed by the same type of a surface element, and conversely, the same type of semantic element can be expressed by

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

several different linguistic surface elements. A range of typological patterns and universal principles can be found by such an approach. Talmy's notion of lexicalization patterns is extensively described on the example of a very specific semantic domain – the event structure. His investigation of lexicalization patterns used in naming the event structure is limited primarily to syntactic constructions. Talmy's notion of lexicalization patterns is foremost oriented towards syntactic understanding of lexicalization patterns. As Levin and Rappaport Hovav (2019) point out, in the current literature lexicalization patterns are attributed to (a) properties of a language's lexical items and generalizations concerning the structure of its lexicon and (b) parametric differences among languages manifested in grammatical mechanisms available to some but not all languages.

This book adopts Talmy's basic notion of lexicalization patterns, but extends it by pointing to similarities and diversities of grammatical mechanisms pertaining to genetically and typologically close and distant languages. Firstly, in accordance with Talmy we consider that the analysis of lexicalization patterns is comprised of defining various lexical forms that encapsulate the way speakers conceptualize perceptual experiences. Secondly, in contrast to Talmy, lexicalization patterns are conceived as language structures produced by diverse grammatical mechanisms, foremost morphological or syntactic, but also phonological, regarding them as equally important and enabling lexicalization of different or similar meanings. Lexicalization patterns refer primarily to diverse realizations of linguistic forms that are available to speakers in the process of naming a certain concept. Every lexicalization pattern construes the lexical meaning in a specific and diverse manner.

In linguistic literature the term *lexicalization strategies* appears as well, as a notion covering diverse processes responsible for linguistic labeling of different concepts (cf. Brdar, 2017). Although linguists use the terms patterns and strategies as interchangeable and covering more or less the same notion, often without a particular definition of the terms, in this book the two terms are considered as related, but different. We consider lexicalization strategies as a broader notion referring to diverse processes speakers use in the process of naming a certain concept. Lexicalization patterns are structural units that represent a more or less conventionalized parts of a language system. For example, lexical borrowing is a lexicalization strategy (cf. Brdar, 2017) used for naming concepts by using foreign lexical units. Borrowed units undergo structural (mostly phonological and morphological) adjustments to the target language structure. Thus, they often adopt a certain lexicalization pattern more or less conventionalized in a target language. For example, the Croatian color term *tirkizna* 'turquoise' is a borrowed unit adjusted to the lexicalization pattern used in the formation of color terms derived from the nouns referring to precious metals or stones such as *zlato* 'gold' : *zlat-na* 'golden', *srebro* 'silver': *srebr-na* 'silver'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the context of applied linguistics, *lexicalization strategies* include various strategies that speakers apply in their use of a second language, such as paraphrase, description, circumlocution, or even avoidance (Duff, 2014, p. 194). Such a definition of lexicalization strategies clearly underlines intrinsic differences between lexicalization strategies and lexicalization patterns. Lexicalization strategies represent a broad scope of mechanisms speakers use in communication and they may not have an effect on the formation of language structure, neither are they necessarily underpinned by existing linguistic forms. In line with Levin and Rappaport Hovav (2019), we consider lexicalization patterns as enabling lexical units to become more or less conventionalized items of a certain lexicon providing a systematic relation to other units of a lexicon. Since lexicalization patterns are more or less conventionalized and pertain to language structure, from a diachronic point of view, they can be more or less stable and, accordingly, replaced by some other patterns over time (cf. Brinton & Closs Traugott, 2005, p. 154).

The reason why we felt it necessary to elaborate upon the terminology and potential differences and relations between the aforementioned terms is to explain the choice of the volume's title, as well as the editors' understanding of the complex notions tied to lexicalization. Though some contributions in this volume do not necessarily follow the same terminological distinctions made here, they can, nevertheless, inform the reader on the complexity and inter-relatedness of the processes behind the phenomenon of lexicalization. Furthermore, whatever the term that particular chapters use, all of them present data on conventional structural units as products of grammatical and semantic mechanisms.

3. Aims and motivation

The book *Lexicalization patterns in color naming: a cross-linguistic perspective* is foremost motivated by the research and results from the project *Evolution of Semantic Systems* (EoSS), 2011–2014, Max Planck Institute for Psycholinguistics, Nijmegen (Ida Raffaelli, leader of the Croatian team). The main interest of the project was how meanings vary over space and change over time and, to that end, data from 50 IE languages were collected. One of the project tasks was a color elicitation task that consisted of 84 Munsell color chips arranged in a single fixed random order (materials developed by Majid & Levinson, 2007). Before the EoSS project, research on color terms was conducted with regards to some IE subfamilies, like Slavic (Comrie & Corbett, 1993), but no large-scale investigation has been undertaken. Therefore, there was room for a more integrative study that would take into account data from a large number of IE languages. Such an endeavor was the EoSS project. The project was grounded on linguistic, psychological, and anthropological

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

theoretical frameworks. One of the basic goals of the project was to investigate color terms speakers use in the partition of the color spectrum. Research on color terms within the EoSS project consisted of several different trials with adult participants. Although the focus of the project was primarily on the demarcation of the color domain, the empirically-based results from the project enabled the beginning of an investigation of lexicalization patterns speakers use in color naming to convey different meanings. Language data sets presented during the *Workshop on Balto-Slavic languages* (Lyon, 2014) inspired the editors of this volume to address questions of similarities and differences with respect to lexicalization patterns of color terms in these languages. This meeting further motivated the organization of the workshop *Lexicalization patterns in color naming: a cross-linguistic perspective* (17th International Morphology Meeting, Vienna, Austria, 18–21 February 2016, convenor Ida Raffaelli) that gathered speakers presenting 11 papers. The aim of the workshop was to provide a forum about lexicalization patterns in color naming, mostly in Indo-European (IE) languages but also in non-IE languages. Thus, not all the papers presented at the Vienna workshop came from the EoSS project and many addressed the topic from different viewpoints than the EoSS project. Therefore, the workshop provided an insight into similar and different language devices that genetically and typologically close and distant languages use in color naming. The theoretical and methodological frameworks of the presented papers demonstrated the great potential of the topic. Furthermore, based on the workshop results, it has become evident that the prospective volume should not be a mere collection of the papers presented there, but that it should represent a coherent, self-sufficient and expanded publication that would cover a variety of languages which reflect the diversity of lexicalization patterns used in color naming.

Thus, the main aim of the present volume is to showcase similarities and diversities in processes and devices that genetically related and typologically diverse language use in color naming.

There are four main questions that the volume aims to address, which emphasize the novel viewpoints in the field of study of color terms. These can be summarized as follows:

a. the focus on language structure per se in the study of lexicalization data,
b. a typological perspective on color terms, especially in terms of inter- and intra-language structural variation,
c. focus on both focal and non-focal color terms,
d. culture and language contact as being reflected in language structure.

With regards to these points (particularly (a), (b) and (c)), the aim of the volume is to approach color data in a novel way. This is particularly tied to the fact that most of the previous studies regarding color terms have dealt with their categorization and,

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

specifically, the perceptual salience of focal colors, whereas few studies have been consistently conducted in order to study the phonological, morphological, syntactic, and semantic devices various languages use to form both focal and non-focal color terms. Some of the main grammatical processes related to word-formation in this volume are:

1. derivation, e.g., Sk. *fialka* 'violet' + *-ový* > *fialový* 'purple, violet-like', Cr. *ljubica* 'violet' + *-asta* > *ljubičasta* 'purple, violet-like', *maslina* 'olive' + *-asta* > *maslinasta* 'olive-like', Po. *fiołek* 'violet' + *-owy* > *fioletowy* 'purple', *oliwka* + *-owy* > *oliwkowy* 'olive-like', Ar. بنفسج */banafsaj/* 'violets' + ـي *'-iyy'* > بنفسجي */banafsajiyy/* 'purple', Hi. *banafśā* 'violet' + *-ī* > *banafśaī* 'violet (color)',
2. reduplication, Hi. *pīlā* 'yellow' > *pīlā-pīlā* 'yellowish, softened color', Ko. *pwulkus* ***pwulkus*** (불긋불긋) [pulgɯtpulgɯt] Adv. 'reddish here and there'
3. conversion, En. *orange*, *violet*, Tu. *horoz ibiği* (lit. 'rooster crest') 'amaranth pink'
4. compounding, Cr. *zeleno-plava* 'green-blue', Ic. *sítrónugulur* 'lemon yellow', Sk. *staroružový* 'old rose'
5. syntactic constructions, e.g., [$N+N_{GEN}$] > Cr. *boja trule višnje* 'color of rotten sour cherry', *boja lavande* 'lavender color', Tu. *fes rengi* (lit. 'fez color') crimson, *türbe yeşili* 'turbe green', Ar. لون الصحراء */lawn aṣ-ṣaḥrā'-i/* (lit. 'color of the desert') 'desert-like color'
6. sound symbolism, e.g., Ko. *nolah-* (노랗-) [norɑt-] 'be (bright) yellow' (positive /o/), *nwuleh-* (누렇-) [nurʌt-] 'be dark yellow' (negative /u/)

Grammatical processes underlying lexicalization patterns are often accompanied by two semantic processes – metonymy and metaphor. These are also attested in cases of colexicification and/or polysemy, e.g., Liberian Kpelle *kpɔlu* (or *kpɔlɔ*) 'to be red, glowing, yellow, ripe', *mɔ* (or *mɔŋ*) is 'to be red, ripe, done; to roast' (for more, see Segerer & Vanhove, this volume).

Though various languages described in the chapters of this book show how these processes vary in their use depending on the language described, they deal with an inventory of diverse phonological, morphosyntactic, and semantic processes that are, as such, universal to language structure in general. Also, examining the relationship between the domain-specific categorization of color terms and the way it is shaped by processes which are not domain specific (e.g., compounding, approximation, reduplication) provides valuable insight into the way languages intertwine broader linguistic categories and processes (e.g., reduplicatives or vowel polarity) with domain-specific ones (e.g., lexicalization of color terms). The focus on language structure per se, being studied consistently through one basic domain of human experience – that of color – is, therefore, a novel investigation into the relationship between language form and language meaning.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Furthermore, the cross-linguistic variation of these processes in between particular languages is a novel theme with regards to a typologically relevant discussion of the relationship between lexicalization and grammatical processes. Hence, the studies described in the present book could be of great interest to the field of linguistic typology and, specifically, the discipline(s) of semantic typology and lexical typology, as interdisciplinary emergent fields of study concerned with the relationship of language form and language meaning cross-linguistically. As some authors emphasize (e.g., Evans, 2010), colors have been at the forefront of the study of cross-linguistic regularity of lexical subsystems and are one of the most studied and well-documented lexical domains across languages. Thus, it can be said that colors are one of the longest typologically studied parts of the lexicon, along with a few other domains such as kinship or body parts. This is also reflected in the *World Atlas of Language Structures* (*WALS*), where out of ten chapters dedicated to the lexicon, four of them deal with color (*Number of Non-Derived Basic Color Categories*, *Number of Basic Color Categories*, *Green and Blue*, *Red and Yellow*). However, the main focus in the investigation of color categories has been on the problem of demarcation of the color domain, i.e., the number of (basic) colors used to segment the color spectrum. This is quite clearly illustrated in the *WALS* data as well. On the other hand, aside from demarcation of lexical categories, a prominent line of investigation posited by both semantic and lexical typology has been the typological systematization of the interaction between lexicalization patterns and grammatical patterns, put forth as one of the basic research goals of lexical typology (Behrens & Sasse, 1997, p. 6). Koptjevskaja-Tamm (2008, pp. 5–6) places lexicalization and lexicalization patterns, "universal" vs. language-specific lexicalizations, carving up of lexical fields/semantic domains and their architecture among the main foci of lexical typology, as well as the investigation of cross-linguistic patterns in the lexicon-grammar interaction. In terms of the main research questions posited by lexical and semantic typology research, this volume addresses them by examining the architecture of the semantic domain of color, especially including data on non-basic colors systematically, as this is the one part of the color domain which can be of great interest in the lexicon-grammar interaction, and the way the color vocabulary is shaped and expanded cross-linguistically. In this respect, the present volume tends to build upon such research and address some of the main questions of the regularities and variations present in the formation of the color lexicon.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4. Language diversity

The language sample in the book is broad enough to ensure the representation of both Indo-European and non-Indo-European languages, from diverse cultural environments, which extends some of the descriptions of color terms into anthropologically, sociologically, and psychologically relevant data. The areal distribution of the languages is comprehensive as it covers languages of Eurasia from western Europe to eastern Asia and Africa, i.e., languages of the Old World. Languages described in the book belong to the following language families or subfamilies:

a. Indo-European, furthermore divided into five relevant subfamilies: (1) Slavic, (2) Germanic, (3) Romance, (4) Indo-Aryan, and (5) Iranian
b. Turkic
c. Koreanic
d. Uralic
e. Afro-Asiatic
f. Niger-Kongo
g. Nilo-Saharan
h. Khoisan
i. isolated (Basque)

Additionally, Icelandic sign language is dealt with in the prospective book, thus expanding the data sets in novel directions.

5. Main topic sections of the volume

The volume is divided into three major sections. The first section, *Lexicalization patterns in and over time*, presents in-depth studies of lexicalization patterns of one language, with a focus on intra-linguistic structural variability. Studies in this section also reexamine some of the major theoretical tenets of color term analysis, such as the notion of basic color term or brightness. The second section, *Color terms in a genealogical and typological perspective*, presents contrastive and/or comparative studies of two or more languages, as well as large-scale typological studies of language families or areas. Studies of typologically relevant languages, such as the isolated Basque, are also presented in this section. A running theme throughout the volume is also the question of language contact, as well as cultural influences on color term formation. However, chapters in the third section, *Languages in culture and languages in contact*, contrast and investigate these parameters in more detail. Each chapter, furthermore, emphasizes and foregrounds various aspects

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

of lexicalization patterns in color naming, in addition to the common theme of presenting structural regularities in and among the languages under study. The thematic division of the volume underlines a wide scope of topics and linguistic disciplines that could consider the lexicalization pattern approach to colors (or any other semantic domain) relevant for further investigation. Although it was not intended at the beginning of this book project to cover different disciplinary subfields of linguistic analysis, the contributions made its relevance evident and, thus, opened new perspectives on the further development or implementation of the lexicalization pattern approach.

5.1 Lexicalization patterns in and over time

Studies collected in this section point to the structural stability and dynamics of lexicalization patterns from a synchronic and diachronic point of view. Since lexicalization patterns represent more or less conventionalized language structures, their stability may change over time – more conventionalized patterns may replace less conventionalized ones.

The lexicalization patterns approach to color terms points to some novel perspectives and understandings of some well-grounded models of color description. In this respect, Berlin and Kay's category of basic color terms (BCTs) could be viewed from a different angle, especially regarding their stipulation of monolexemic criterion as the only (strictly speaking) linguistic characteristic of basic color terms. The question of the concept of "basic color terms", as defined by Berlin and Kay (1969), is discussed in the contribution of **Réka Benczes** and **Erzsébet Tóth-Czifra**, titled *Rethinking the category of "basic color term": Evidence from Hungarian lexicalization patterns*. Based on the empirically-driven language analysis, the authors point to the two Hungarian terms (*piros* and *vörös*) that do not fit the BCTs, according to Berlin and Kay's criteria for identifying such lexical items. The authors provide a critical review of the relevant literature, arguing that the question can be approached by analyzing linguistic data through the investigation of the salience of color terms in language use. Accordingly, by analyzing etymological data and the lexicalization patterns of [N/Adj+color term] constructions, this chapter challenges Berlin and Kay's (1969) original definition of "basic color term" by calling into question the rigidity of this category. Instead, the authors propose a radial category structure for "basic color terms", where certain terms are more typical (central) members than others. Such a flexible approach is able to account for more problematic cases of color terminology as, for example, those exhibited by Hungarian.

Another contribution which challenges the concept of "basic colors", but from a different point of view, is *Lexicalization patterns in color naming in Gbaya, a*

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ubanguian language of CAR by Paulette Roulon-Doko. It deals with color naming strategies in Gbaya, a language spoken in the northwest part of the Central African Republic. In the first part of the chapter, the author shows which lexical terms and grammatical devices are used in Gbaya for color naming, and in the second part, how Gbaya speakers put them to use in a daily task. This second part of the chapter is based on very rich and interesting empirical data, i.e., on the corpus of spontaneous utterances collected during the author's research, including a total of five years of fieldwork. Unlike many of the chapters in this book, this one does not follow Berlin & Kay's (1969) theory of universals of color naming; rather, it adheres to the critical arguments developed by Lucy (1997), Wierzbicka (2008), Dimmendaal (2015), and Moñino (2004) showing that Gbaya does not lexically distinguish color from some other characteristics of visual aspect. Gbaya does not even have a generic term for 'color', and there is no way to directly ask about a color. This is because color is just one element (among others) of the visual appearance of an object.

In the contribution by **Katarina Dudová**, *Lexicalization patterns in Slovak color naming*, the BCT system is approached from a synchronic, as well as from a diachronic point of view by presenting developmental stages of certain color terms. The example of the Slovak language clearly shows that certain lexicalization patterns can be traced back in time and are in accordance with Berlin and Kay's cultural developmental stages of color terms. Referring to Berlin and Kay's notion of the evolutionary development of color terms, the author examines the development of Slovak color terms from Old Slavonic to contemporary Slovak. Consequently, she examines lexicalization patterns of color terms in contemporary Slovak in general, focusing on derivation and compounding and the four semantic categories they lexicalize – 'brightness', 'hue', 'cross-sensory experience', and 'ancientness'. A cultural basis of color term formation, both diachronically and synchronically, is given prominence in the investigation of particular color terms. These categories are examined with respect to the results obtained by the elicitation experiment in the EoSS study and form the main part of the analysis in the chapter. Diachronic perspective to lexicalization patterns could thus assess statements about the cultural development of color terms and point to different levels of their structural complexity over time.

As already mentioned, a diachronic perspective to lexicalization patterns could shed new light on the dynamics of co-occurring patterns over time. A dominant lexicalization pattern from a synchronic perspective may have undergone certain diachronic changes or may have been less prominent in earlier stages of language development. Such a question is addressed in the chapter *Compound color terms in Italian* by **Maria Grossmann** and **Paolo D'Achille**, who investigate compound color terms in Italian. The aim of the chapter is to give a descriptive overview of the

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

characteristics of the individual compounding patterns both from the synchronic and the diachronic point of view. According to the preliminary research conducted by the authors, in contemporary Italian, compounding is the most commonly used device for enlarging the inventory of color terms. Based on the corpus-driven data, the authors state that compound color terms are situated at the crossroads between word-formation and syntax, i.e., between compounding proper and asyndetic coordination. From a diachronic point of view, it has been ascertained that compound color terms start to appear with a certain frequency only from the 18th century onwards; in earlier times they are attested only by a few examples and in a limited number of patterns.

It has been pointed out that lexicalization patterns embrace the diversity of mostly morphological and syntactic mechanisms enabling the formation of color terms. However, the fact that phonological features should also be taken into consideration as being operative in color naming is shown in the contribution *Lexicalization patterns in color naming in Korean* by **Seogha Rhee**, who analyzes lexicalization patterns in Korean as a language that has one of the richest color vocabularies in the world. Korean as a language with sound symbolism is an excellent illustration of how phonological processes are effective in the formation of color terms. Unlike many of the languages in this volume, Korean uses mechanisms such as vowel polarity, consonant tensing and aspiration, and suffix reduplication to form or modify meanings of color terms, e.g., intensity, luminosity, extent, solidity, and pleasurability. With a focus on sound symbolism, the chapter points out synesthesia as an important factor in color term formation. The in-depth analysis of Korean color terms expands the notion of lexicalization patterns in this volume, with sound symbolism as an important typological feature.

Since one of the main goals of the lexicalization patterns approach it should be lexicalization patterns approach to color terms is to contribute to the relativist understanding of how language structures shape the color spectrum, this approach could reveal some linguistic and, thus, cultural specificities within the same language family or cross-linguistically. Moreover, the lexicalization patterns approach could challenge some of the generally acknowledged notions such as 'brightness' that could be cross-linguistically conceived in a different manner. This is the main focus of the chapter *"Brightness" in color linguistics. New light from Danish visual semantics* by **Carsten Levisen** in which the author introduces the Natural Semantic Metalanguage approach to the category of brightness in Danish. The chapter calls for a renewed focus on the category of brightness, pointing out the problematic aspects of this notion in cross-linguistic studies, as well as emphasizing the need for an emic perspective and culture-sensitive analysis of the notion itself. These main points are exemplified by a fine-grained semantic analysis of *lys* 'light' and its compounds in Danish.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

5.2 Color terms in a genealogical and typological perspective

A genealogical account of color terms tends to highlight shared and inherited lexical items. However, innovations within a language family tend to diversify the languages under observation, and sometimes important differences may arise within subfamilies. Therefore, the lexicalization patterns approach can be used to highlight divergences, as much as inherited similarities. A typological perspective can be used to a similar end as well, to highlight and present commonalities and specificities of lexicalization patterns in a language, not only according to a genealogical, but a typological classification as well. The chapters presented in this section deal with both perspectives in their investigation of color terms, some based on an in-depth discussion of one language, others using contrastive and large-scale studies of language families or areas.

Firstly, some of the chapters are dedicated to presenting novel and typologically relevant data on color terms in one language. Some, such as Hindi, are very prominent in their language subfamily, while others, such as Basque, are isolated languages with important typological differences from their neighboring languages.

The chapter by **Andrea Drocco** and **Orsola Risato**, *Lexicalization patterns in color naming: the case of Modern Hindī*, explores several types of lexicalization patterns in this language. The first is the "metonymy-type", i.e., OBJECT COLOR FOR COLOR pattern formed by the suffix *ī* (a common type found in other languages in this volume). A second pattern is the "approximation-type" formed by the suffix *-sā* (e.g. *kālā-sā* 'blackish'). Finally, a third pattern formed via color adjective reduplication is examined in detail and four semantic categories are proposed: (1) emphasized color, (2) softened color, (3) situational color, and (4) moving/changing color. These categories systematize and examine fine-grained contextual cues in basic color variation with respect to the shared mechanism of reduplication, not discussed in other Indo-European languages in the volume. By investigating these patterns in Modern Hindī, the authors emphasize the need of the speakers to lexicalize "shades" that cannot be enclosed in mere softened or emphasized versions of the prototypical color.

In the chapter *Color terms in Basque: lexicalization and categorization*, **Iraide Ibarretxe-Antuñano** explores the domain of color terms in Basque from a descriptive as well as a categorization perspective. The first part provides an overview of the color system in this language: its color lexicalization processes and the meanings these color labels cover. The second part focuses on the results of a categorization study of color in Basque. The data come from 20 native speakers of Basque and were elicited using the color elicitation task as part of the EoSS project at the Max Planck Institute for Psycholinguistics. The chapter concludes with a discussion on the similarities and differences found in the descriptive part and the categorization part.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Especially interesting is the assimilation process of new loanwords used by native speakers of the Basque language who omit original Basque words in their usage. Moreover, the author points to the process of Romanization in color description accounting for the use of left-headed compound loanwords. Basque compounds are right-headed, whereas Romance languages, such as Spanish, are left-headed (e.g., *azal kolore* (skin color) vs. *color carne* (color flesh) 'skin color').

Secondly, there are studies dedicated to investigating the question of divergence between genetically related and typologically similar languages, especially since they are often viewed as part of the same or similar cultural circle and share some important morphosyntactic properties which emphasize their structural commonalities as well. Therefore, investigating differences in the degree of conventionalization of lexicalization patterns is relevant to the study of linguistic forces that drive the divergent color term creation.

In the chapter *Lexicalization patterns in color naming in Croatian, Czech, and Polish*, **Ida Raffaelli**, **Jan Chromý**, and **Anetta Kopecka** define and describe strategies speakers of Croatian, Czech, and Polish use in color naming. The findings are based on the data from the cross-linguistic EoSS project (Max Planck Institute for Psycholinguistics, Nijmegen). The chapter presents the results of the color naming task for the three languages. The study identified the main lexicalization patterns that are productive in the formation of Croatian, Czech, and Polish color terms. They are the results of different grammatical mechanisms used in the lexicalization process. However, the languages differ with respect to the degree of conventionalization of these mechanisms in the domain of color terms. For example, the lexicalization pattern [stem + suffix] is less frequently used (and thus less productive in the formation of color terms) in Croatian when compared to Czech and Polish, whereas the lexicalization pattern [Adj [o] Adj] is frequently used in all the three languages, however, in Polish, less productively in comparison with Croatian and Czech.

Another study dealing with the question of divergence between genetically related and typologically similar, but also areally close languages is the contribution of **Xavier Bach**, **Anetta Kopecka**, and **Benjamin Fagard**. In the chapter *Complex color denomination in French and Occitan*, they investigate the color naming in French and Occitan. While the use of primary terms is common in these languages, both have secondary color terms derived by way of suffixation, such as, for instance, the suffixes *-astre* in Occitan and *-âtre* in French, which carry an approximative meaning comparable to *-ish* in the English term *greenish*. However, it is well known that French, compared to other Romance languages, has a tendency to be very analytic. This raises the following question: do speakers of French and other Romance languages (here, Occitan) differ in morphological strategies used to express meaning in the semantic domain of color? To answer this question, the authors compare the

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

color descriptions provided by speakers of French ($N = 20$) and speakers of Occitan ($N = 20$). The results show some striking differences between the two languages in lexical and morphosyntactic strategies used to name colors. Speakers of Occitan employ secondary color terms using a variety of derivational suffixes, which are absent from the French data, despite their existence in the language.

Finally, all of the aforementioned questions, from microvariations among languages and sources of color terms to similarities and divergences between genetically or areally related languages are addressed in two large-scale studies. The first of them is the study by **Maria Bulakh**, titled *Innovations in Semitic color term systems*. It explores the origins of Semitic color terms for yellow, green, and blue. These fundamental color categories are missing from the reconstructed proto-Semitic basic color term system, but their designations were added into the basic color term systems of many daughter languages. The chapter focuses on the derivations from designations of "referent objects" (objects typically characterized by a certain color and serving as referents for the designation of this color). The referent objects are further classified into natural objects (wax for yellow; vegetation, grass, or leaf for green; sky for blue) and dyes (saffron and turmeric for yellow; kohl and indigo, or indigo-colored objects, for blue; the chapter also considers ink as a source for designations of blue-with-green, violet, or black). In the present contribution, the examples of names of new basic color categories and/or non-basic terms, which can eventually become basic color terms for new basic color categories, are discussed.

The second large-scale study is *Color naming in Africa* by **Guillaume Segerer** and **Martine Vanhove**. Based on the data from 350 languages and from 374 lexical sources, it is the first large-scale typological survey on lexicalization of colors in African languages. It examines morphosyntactic strategies, language-internal semantic sources, and contact-induced terms used to name colors. Most of the analyzed data were taken from the *RefLex* online lexical database (Segerer & Flavier, 2011–2018), and the 350-language sample is relatively balanced, both geographically and genetically. The authors first discuss the issues related to 'basic' and 'polychromatic' color terms, and then provide an overview of their semantic sources, the origin of borrowings, colexifications and metaphorical uses of color terms, main patterns of lexicalization, and, briefly, a particular case of intensifiers and ideophones, an important part of color-related terms in a large number of African languages. The authors' main conclusion is that patterns of lexicalization, metaphors, and colexifications of color terms do not deviate from general cross-linguistic tendencies. Still, a few local patterns emerge, such as the naming of YELLOW after the locust tree (attested throughout West Africa), and borrowing of BLUE from European languages and GREEN from neighboring languages.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

5.3 Languages in culture and languages in contact

One important point made by the relativist view of color conceptualization and categorization is that color is construed in a specific sociocultural context and affects various aspects of human experience. Color plays an important role in society, as seen in conventional connotations of colors in present-day and older societies, the way that color helps to establish the identity of a society or how it is explored and utilized in artwork (for an overview, see Biggam et al., 2011). Furthermore, a cross-linguistic investigation of colors and their lexicalization patterns cannot be easily disentangled from a cross-cultural investigation as well. A similar point is made by Wierzbicka (2008), referring to the fact that not all cultures use the same notion of color to abstract away a common property of various phenomena in the world (e.g., some may use geometrical patterns or lines). Moreover, contact of cultures implies contact of languages as well, and these two phenomena can be jointly observed and contrasted as working towards the structuring of the color domain. From a lexical typology perspective, an area of interest are the differences in borrowability among the different parts of the lexicon, and the ways new words are integrated into a target language and adjusted to new objects and concepts (Koptjevskavja-Tamm, 2008, p. 6). Culture, on the one hand, and language contact on the other, are coded in language structures and drive the formation and conceptualization of lexical units in specific ways, as listed below.

Firstly, there is the question of microvariation among languages or language varieties. This question is addressed in the chapter *Universals and Variability of Color-Naming in Icelandic, Icelandic Sign Language and North American Icelandic* in which **Þórhalla Guðmundsdóttir Beck** and **Matthew Whelpton** deal with the complex issue of examining microvariation in Icelandic, North American Icelandic as a heritage language, and Icelandic Sign Language, also providing data from British English and North American English. The comparison of five language systems from three cultural contexts (Icelandic, British, and American English) paves the way for a multi-level analysis of color categories and formal strategies used to form color terms. Using the methods established by the EoSS research, they provide arguments for a culturally constrained and relativistic view of color variation among languages with different typological, historical, and cultural backgrounds. For example, there are similarities observed in typologically different languages (Icelandic and Icelandic Sign Language) which can be attributed to a common cultural context, while still retaining the importance of typologically based differences as well. Thus, they conclude that the sociocultural relativist view of color variation allows for an examination of differences in the formal strategies used by these languages in color formation.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Secondly, there is the all-important question of color symbolism in language and culture. As is commonly known in lexical studies, one part of the lexicon that shows an intimate relationship between culture and language is the field of idioms. Idioms (or more broadly, phraseological units), as studied in phraseology, are analyzed as being an integral part of linguo-cultural competence, i.e., as cultural-linguistic codes (Teliya et al., 1998). As far as lexicalization patterns go, phraseological units are therefore lexicalization patterns in their own right, with internal morphosyntactic and semantic regularities. In the chapter *Symbolic and Cultural Meaning of Colors in Phraseology A Cross-Linguistic and Cross-Cultural Study of Russian and German Phraseological Units,* **Branka Barčot** and **Anita Hrnjak** contribute to the volume with an innovative investigation of phraseological units with a color component in Russian and German. Phraseological units are often culturally highly specific and the chapter presents arguments for a relativistic perspective of color vocabulary from a phraseological viewpoint. The symbolic and/or cultural meaning of color terms is analyzed as the basis of most phraseological meanings, however, there are cases where denotative lexical meaning is basic as well. Four structural types of phraseological units are examined with respect to 11 basic color terms in Russian and German: (a) collocational idioms, (b) propositional or sentence idioms, (c) lexical (cranberry) idioms and (d) similes. The last category – similes – is analyzed as the only structural type of phraseological unit that can be considered as a base for the lexicalization of certain color terms, particularly in Russian, but German as well, e.g., Ru. *голубой как небо* lit. 'blue as sky' > *небесный* lit. 'sky-like', *небесноголубой* lit. 'skyblue'; Ger. *blau wie der Himmel* lit. 'blue as the sky' > *himmelblau* lit. 'skyblue'.

Thirdly, sources of color terms can be experientially constrained both by the natural and cultural environment. Furthermore, color has the potential to be used in vocabulary formation denoting other domains of experience. In the chapter *From object to color and back: seeing the world in color in Croatian, Turkish, and Arabic*, **Daniela Katunar**, **Barbara Kerovec**, and **Nawar Ghanim Murad** compare lexicalization patterns of color terms in Croatian, Turkish, and Arabic, three languages which are typologically, genetically, and socioculturally quite different. Due to historical, cultural, and civilizational reasons, some (Turkish and Arabic) have been in contact for a longer time and more intensively than others (Croatian), and this is manifest in some shared lexical and morphological items. The authors focus particularly on the relation between colors and other phenomena of experience by examining (a) the way the languages exploit the OBJECT FOR COLOR conceptualization strategy in order to name colors and (b) the way languages use the COLOR FOR OBJECT conceptualization strategy in order to lexicalize experiences from domains other than color itself, e.g., Cr. *crvendać* 'robin', Tu. *kızıl gerdan* 'robin', lit. 'red neck', Ar. أخضَر /akhḍar/ 'greenfinch', lit. 'green'. The results show differences

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

among the three languages in the productivity of lexicalization patterns (derivation, compounding, and multi-word expressions) in both types of conceptualization strategies. By examining the direction 'COLORS → OBJECTS', the authors show that it is those very domains commonly used for color term formation that are enriched by using extant color terms in naming various phenomena.

Finally, color terms in one language can be viewed from many perspectives at once, i.e., by examining areal, historical, and cultural factors as differing tendencies in color term formation. **Arseniy Vydrin**, in his paper *Ossetic Color Terms System*, focuses on color naming in Iron Ossetic, the main dialect of Ossetic as a modern Eastern Iranian language. The paper is mainly based on the EoSS project methodology (Majid, Jordan, & Dunn, 2011), while the Ossetic material is taken from field data collected from native speakers in North Ossetia and data from the Ossetic written and oral corpora. The author first determines Iron Ossetic's basic color terms and the place of Iron Ossetic system in the evolutionary sequence of the development of basic color terms. Then he examines the grammar, semantics, and use of basic and non-basic color terms by explaining the main patterns used to name colors, morphological and lexical color modifiers, synonyms, and connotations of color terms. Very interesting insights are given on the influence of geographical and historical factors on the Ossetic color terms system, especially in relation to borrowings and some peculiarities in comparison with the systems of some other Iranian languages.

Acknowledgements

We would like to thank the authors who accepted the invitation to contribute in this volume. It is their contributions that made the book possible. We express our deepest gratitude to all the anonymous reviewers whose insightful and constructive comments and suggestions helped the authors to improve their contributions. Thanks to their selfless efforts, the overall quality of the present volume was significantly elevated. We would also like to thank Dunja Opatić on her meticulous proofreading of the volume, which was no small feat in itself. Furthermore, we would like to thank the University of Zagreb for providing us with the necessary support to follow through with the publication, as well as John Benjamins Publishing for showing interest and giving us the motivation to compile the volume. Finally, we would like to thank the investigators of the EoSS project, whose work provided us with a pathway to the research questions examined in the volume.

Language abbreviations used in this chapter are: Ar. > Arabic, Cr. > Croatian, En. > English, Ger. > German, Hi. > Hindī, Ic. > Icelandic, Ko. > Korean, Po. > Polish, Ru. > Russian, Sk. > Slovak, Tu. > Turkish.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

References

Behrens, L., & Sasse, H. J. (1997). Lexical typology: A programmatic sketch. *Arbeitspapier* Nr. 30 (Neue Folge). University of Cologne.

Blank, A. (2001). Pathways of lexicalization. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *Language typology and language universals: An International Handbook*, *2* (pp. 1596-1608), Berlin / New York: Walter de Gruyter.

Biggam, C. P., Hough, C. A., Kay, Ch. J., & Simmons, D. R. (Eds.). (2011). *New Directions in Colour Studies*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.167

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

Brdar, M. (2017). *Metonymy and Word-Formation: Their Interactions and Complementation*. Cambridge: Cambridge Scholars Publishing.

Brinton, L. J., & Traugott, E. L. (2005). *Lexicalization and Language Change*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511615962

Comrie, B., & Corbett, G. G. (Eds.). (1993). *The Slavonic Languages*. London: Routledge.

Duff, P. A. (2014). The lexical generation gap: a connectionist account of circumlocution in Chinese as a second language. In G. Kasper & E. Kellerman (Eds.), *Communicative Strategies: Psycholinguistic and Sociolinguistic Perspectives* (pp. 192–215). London / New York: Routledge.

Evans, N. (2010.) Semantic Typology. In J. J. Song (Ed.), *The Oxford Handbook of Linguistic Typology* (pp. 504–533). Oxford / New York: Oxford University Press.

Heider, E. R. (1972). Probabilities, sampling, and ethnographic method: The case of Dani colour names. *Man*, *7* (3), 448–466. https://doi.org/10.2307/2800917

Koptjevskaja-Tamm, M. (2008). Approaching lexical typology. In M. Vanhove (Ed.), *From polysemy to semantic change: Towards a typology of lexical semantic associations* (pp. 3–52). Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/slcs.106.03kop

Levin, B., & Rappaport Hovav, M. (2019). Lexicalization Patterns. In R. Truswell (Ed.), *Oxford Handbook of Event Structure* (pp. 395-425). Oxford, UK: Oxford University Press.

Majid, A., & Levinson, S. C. (2007). The language of vision I: Colour. In A. Majid (Ed.), *Field manual*, *10* (pp. 22–25). Max Planck Institute for Psycholinguistics, Nijmegen. Retrieved February 10th 2015 from http://fieldmanuals.mpi.nl/volumes/2007/language-of-vision-colour/.

Majid, A., Jordan, F., & Dunn, M. (2011). *Evolution of Semantic Systems Procedures Manual*. Nijmegen: Max Planck Institute for Psycholinguistics.

Regier, T., Kay, P., & Ketharpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences of the United States of America 104* (4), 1436–1441. https://doi.org/10.1073/pnas.0610341104

Rosch, E. H. (1973). Natural categories. *Cognitive psychology*, *4* (3), 328–350. https://doi.org/10.1016/0010-0285(73)90017-0

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. *Language typology and syntactic description*, *3* (99), 36–149.

Teliya, V., Bragina, N, Oparina, E., & Sandomirskaya, I. (1998) Phraseology as a Language of Culture: Its Role in the Representation of Collective Mentality. In E. Cowie (Ed.), *Phraseology. Theory, Analysis and Applications* (pp. 55–75). Oxford: Oxford University Press.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wierzbicka, A. (2005). There are no "color universals" but there are universals of visual semantics. *Anthropological linguistics*, *47* (2), 217–244.

Wierzbicka, A. (2008). Why there are no 'colour universals' in language and thought. *Journal of the Royal Anthropological Institute*, *14* (2), 407–425. https://doi.org/10.1111/j.1467-9655.2008.00509.x

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

PART I

Lexicalization patterns in and over time

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rethinking the category of "basic color term"

Evidence from Hungarian lexicalization patterns

Réka Benczes and Erzsébet Tóth-Czifra
Corvinus University of Budapest / Eötvös Loránd University, Budapest

One of the peculiarities of Hungarian is that two of its "basic color terms" (BCTs), as defined by Berlin and Kay (1969), do not conform to the established criteria used to identify such lexical items. By analyzing etymological data and the lexicalization patterns of [N/Adj + color term] constructions, the chapter challenges Berlin and Kay's (1969) original definition of "basic color term" by calling into question the rigidity of this category. The chapter proposes instead a radial category structure for "basic color term", where certain color terms are more typical (central) members than others. Such a flexible approach is able to account for more problematic cases of color terminology, as for example that exhibited by Hungarian.

Keywords: Hungarian, basic color term, corpus linguistics, lexicalization, radial category

1. Introduction[1]

In their seminal study on the universal evolution of color terms, Berlin and Kay (1969, pp. 6–7) establish a number of criteria to distinguish basic color terms (BCTs) from non-basic ones. Thus, in order for a color term to be considered as "basic", it needs to conform to the following list of characteristics: (a) it is monolexemic (i.e., the meaning as a whole is not deducible from the constituents); (b) the color it refers to cannot be contained in another color; (c) it cannot be used for a small number of entities; (d) it must be generally known and well used. Based on these criteria, Berlin and Kay established eleven BCTs and furthermore argued that the basic color terms – of any language – emerge in a universal sequence, which

1. We wish to thank Daniela Katunar, Barbara Kerovec, Ida Raffaelli, and our two anonymous reviewers for the very helpful comments and suggestions.

https://doi.org/10.1075/sfsl.78.02ben
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

is the following: *black*, *white* ⇨ *red* ⇨ *yellow*, *blue*, *green* ⇨ *brown* ⇨ *purple*, *pink*, *orange*, *gray*. Thus, if a language has two BCTs, then those will be *black* and *white*; if a language has three, it will be *black*, *white* and *red*, and so on. In effect, the study offered a novel perspective for an empirical justification of the universalist thesis, and by doing so, it also helped to eradicate the very last remnants of the linguistic relativity hypothesis from mainstream linguistic theorizing.

Needless to say, over the decades Berlin and Kay's (1969) publication has undergone a number of modifications (see, for example, Kay, 1975; Kay & Maffi, 1999; Kay & al., 2009) and has inspired (and still continues to inspire) a host of publications on the BCTs of a rather wide range of languages such as Russian (Corbett & Davies, 1995; Paramei, 2005), French (Forbes, 1979), Turkish (Özgen & Davies, 1998), Japanese (Uchikawa & Boynton, 1987) and Hungarian (Kiss & Forbes, 2001; Benczes & Tóth-Czifra, 2014). Yet one of the common – and most significant – traits of these publications is that they conform to the rather strict and formal criteria of basic color term status, as laid down originally by Berlin and Kay (1969), in order to clearly distinguish and thus separate basic color terms from non-basic ones. The fact that BCT research has more-or-less fully accepted[2] (and has even taken for granted) the criteria established by Berlin and Kay nearly fifty years ago is all the more intriguing in light of the fact that the delineation of basicness in a number of languages can be problematic – at least from the point of view of the original classification. Anomalies concerning Berlin and Kay's (1969) categorization fall into two broad categories; these will be discussed in Section 1.1.

1.1 Anomalies

1.1.1 *Two particularly salient (and rival) color terms share one color domain*

According to Berlin and Kay (1969), and on the basis of the criteria they established for determining what a BCT is, any particular language can be placed onto an evolutionary sequence encompassing eleven basic color terms (see above). Any language that possesses a higher number of BCTs (i.e., higher than eleven) is regarded as exceptional and even unique. Yet such languages do exist and are not that rare; one of the most famous cases of this idiosyncrasy (and which has also been elaborated on by Berlin and Kay, pp. 35–6) is Russian. While English has only one BCT to refer to the blue color domain, Russian has two: *siniy* (referring to a darker hue) and *goluboy* (denoting a lighter hue of blue). The dominant and generally

2. Needless to say, there have been some attempts in the scholarly literature to refine (though not to reassess) Berlin and Kay's (1969) original criteria. See, for example, Forbes' (1979) study on French BCTs (and elaborated on in the following section herein).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

accepted view in the available literature is that both Russian color terms are basic (see Frumkina, 1984; Corbett & Morgan, 1988; Davies & Corbett, 1997; Corbett & Davies, 1995; Paramei, 1999, 2005).[3] A similar division within the blue color domain is observable in Turkish (Özgen & Davies, 1998), Japanese (Uchikawa & Boynton, 1987) and Italian (Paggetti & al., 2016).

One does not need to look far within the Indo-European language family to find yet another language where the number of basic color terms exceeds eleven. As discussed by Forbes (1979), two basic color terms share the brown color domain in French, both of which are psychologically and linguistically salient: *brun* ("brown") and *marron* ("chestnut", "reddish brown"). Forbes was one of the very first scholars to draw attention to the tight correlation between frequency and basicness, claiming that if a color term is basic in a given language, then this will reflect in its overall frequency. In light of this claim, Forbes complemented Berlin and Kay's (1969) original criteria with the frequency data of synchronic and diachronic dictionaries. Thus, Forbes' (1979) study diverges from the vast majority of BCT research in the sense that it brings in a new perspective on the issue by considering frequency as a further, possible factor in delimiting and determining basicness. Importantly, in her concluding remarks she does not establish strict category boundaries between basicness and non-basicness, and – when it comes to *brun* and *marron* – leaves the question definitely open:

> At the moment there appears to be a basic colour category focused in *marron*. The term *brun*, however, still retains all the characteristics of a basic term, although its exclusive use in the domains of hair and skin colour may indicate a specialization not normally associated with basic colour terms.
>
> (Forbes, 1979, p. 304, emphasis as in original)

Last but not least, in their description of the universality of eleven basic color terms, Berlin and Kay (1969) also made note of Hungarian, a language with "ten basic categories exclusive of red and two basic terms for red" (pp. 35–6), i.e., *piros* and *vörös*, and which makes Hungarian in their view "unique" (p. 95) as compared to other languages. While the linguistic literature treats this particular phenomenon as a possible exception to the universalist thesis, the available research on *piros* and *vörös* typically focuses on the semantic disambiguation of the color terms and their respective basic color term status (see, for example, Benczes & Tóth-Czifra, 2014). Consequently, Berlin and Kay's original taxonomy is left untouched; it is the basicness of *vörös* that is typically called into question. In fact, the now consensus view

3. Note, however, that some scholars oppose this view: Taylor et al. (1997), for example, do not consider *goluboy* (the lighter hue) to be a basic color term.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

within the literature is that *vörös* is not a basic color term of Hungarian.[4] See also Uusküla (2008, 2011), who makes note of a similar phenomenon in Czech, where the color domain of red is also split between the more literal *červená* and the more figurative *rudá* (see below for an elaboration).

1.1.2 *Morphologically complex color terms*

Despite the fact that Berlin and Kay (1969) stipulate the monolexemic criterion as the only (strictly speaking) linguistic characteristic of basic color terms (and note that this very feature holds the first position on their list of criteria – see above in the Introduction), the literature happily acknowledges a number of morphologically complex color terms as basic. One such example is Finnish *vaaleanpunainen* ("pink, literally light red", Uusküla 2008), or Hungarian *rózsaszín* ("pink", literally "rose+colour", i.e., "rose-colored") and *narancssárga* ("orange", literally "orange+yellow"). Although neither *rózsaszín* ("pink"), nor *narancssárga* ("orange") conforms to Berlin and Kay's (1969) definition of basicness from a morphological point of view, both color terms have been readily accepted as basic (see Benczes & Tóth-Czifra, 2014, p. 124; Kiefer, 2005, p. 132; Papp 2012a, 2012b). Such laxity with regard to the criterion of morphological simplicity suggests that it might be losing its overall significance – which is exactly what the following passage from Uusküla also implies:

> As I have argued in my article "Basic colour terms of Czech", the first criterion has often been misunderstood, because it contains a "magic" word monolexemic […] The "magic" word monolexemic does not mean morphologically, but rather semantically simple. In fact, Berlin and Kay have never stressed that only monomorphemic words can meet the basic colour term criteria. Nevertheless, it might be useful to redefine the concept in a less confusing way. (2008, p. 27)

Uusküla (2008) thus raises the possibility that the monomorphemic criterion might be a misunderstanding. Nevertheless, Berlin and Kay's (1969) original wording does seem to be rather clear on what "monomorphemic" refers to and rules out the possibility of morphologically complex, i.e., compound words, where the meaning as a whole is indeed motivated by the meaning of the constituents. How the original wording of Berlin and Kay (1969) can or cannot be interpreted is one issue (see also Biggam, 2012, p. 32); Uusküla's (2008) other, consequent point on the need to "redefine the concept" (p. 27) of monomorphemic words is, however, even more relevant, to which the present paper will also return to in the forthcoming sections.

---

4. See Benczes and Tóth-Czifra (2014) for an overview of the Hungarian scholarly debate on *piros* and *vörös*.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

1.2 The foregrounding of psychological salience and the backgrounding of linguistic aspects in basic color term research

The diminishing significance of the criterion of morphological simplicity has gone hand in hand with the increasing importance of psychological salience, which is well documented by the fact that the vast majority of the scholarly literature has adopted Berlin and Kay's (1969) fourth criterion on psychological salience as its starting point. The fourth criterion states the following: "It [the BCT] must be psychologically salient for informants. Indices of psychological salience include, among others, (1) a tendency to occur at the beginning of elicited lists of color terms, (2) stability of reference across informants and across occasions of use, and (3) occurrence in the idiolects of all informants" (p. 7). The evolution of the fourth criterion into a research question becomes apparent in the methodology that these studies have adopted. In line with Berlin and Kay's (1969) original investigations, typically two kinds of empirical studies have been carried out: (1) a color-listing task, in the course of which subjects list as many color terms as they know; (2) a color-naming task, during which subjects provide linguistic data (in the form of color terms) by identifying and naming the color chips that they are presented with.

Such studies have been primarily concerned with the relationship between the physical (perceptual) recognition of color on the one hand and color-naming on the other hand, limiting the interest in BCTs to mostly their physical denotation and psychological salience. In effect, this line of research has focused on the following four questions:

1. To what degree are focal colors identical in various languages?
2. To what degree do the boundaries of the color domains that BCTs refer to diverge in various languages?
3. To what degree is there a consensus (agreement) among native speakers with regard to the referents of color terms?
4. In what order and with what frequency do subjects name color terms in an artificial and decontextualized situation?

The last research question brings one of the most significant aspects of this line of research into the foreground: its negligence of data arising from natural language use. Yet in line with usage-based models (see, for example, Bybee, 2001; Barlow & Kemmer, 2000), it is more than probable that the highly decontextualized utterances that were obtained from subjects do not do justice to the actual usage patterns of basic color terms in the respective languages. This possibility has also been alluded to by Kiss (2004), who complemented his color-listing survey with individual corpus searches for the occurrence of each Hungarian BCT. Interestingly, the results of the individual corpus searches came into stark contrast with those of

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

the color-listing surveys: while *vörös* did not appear in the color-listing task at all, it ranked as the sixth most frequent color term in the corpus-based query.

The individual anomalies of BCTs (as found in a number of languages) and the dominance of psychological salience in basic color term research (and accordingly, the backgrounding and even ignorance of linguistic and semantic aspects) has led a number of researchers to question the original category of BCT as laid out by Berlin and Kay (1969). There have been a number of attempts to redefine the category of "basic color term"; Section 1.3 will provide a brief overview of these efforts.

1.3 Attempts to redefine "basicness"

Generally, it can be claimed that criticism leveled against the theory of basic color terms and the definition of the category itself becomes evident in studies that do not solely rely on psychological aspects but also take linguistic features into consideration (or even shift their focus from a psychological to a linguistic one). Only ten years after the publication of Berlin and Kay's (1969) research, Kay and McDaniel (1978) were the first to argue against the strict boundaries of the category of BCTs and emphasized instead the category's gradedness. Accordingly, they distinguished between primary and secondary BCTs, and also erased the strict boundary between basic and non-basic color terms (see also Bimler & Uusküla 2016 for a similar approach).

The idea of gradedness as a possible alternative to strict category boundaries did not, however, become widely accepted within color term research. For example, Moss (1989), citing Baxilina (1975), claimed that such an approach was not able to account for the phenomenon of salient, yet rival color terms (*piros vs. vörös*, *sinyij vs. golubuy*, *marron vs. brun*, etc.), since the "surplus" color terms emerged relatively early on in the history of the respective languages, and not at the end of the evolutionary sequence (as the twelfth BCT). Moss (1989) instead offers a different solution to the problem of language-specific anomalies. In his view, basicness is, in essence, an interdisciplinary concept, composed of three, partly converging levels. From this perspective, psychological salience is only one aspect of basicness, besides physiological and linguistic aspects.

A further possible aspect is culture. Paramei (2005) has insisted on the role of culture in her analysis of the Russian BCTs *siniy* and *goluboy*, proposing that *goluboy* is a socioculturally embedded BCT of Russian. By doing so, she emphasizes the context-specific nature of BCTs in general, which should not exclude a color term from being considered as basic. In a similar vein, Uusküla (2008, pp. 8–9) has also called for the differentiation of "salience" within the context of basic color terms, suggesting the incorporation of linguistic and cultural aspects. In her analysis of the basic color term for "red" in Czech, she proposes that Czech possesses two BCTs

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

for the same color domain, *červená,* and *rudá*. Nevertheless, there is a clear demarcation with regard to the semantic and pragmatic use of the two terms: the latter term is more "subsidiary" and is also "collocation-specific and connotation-loaded" (p. 9). Thus, an overarching analysis of color terms needs to take into consideration "not only the colour term denotata but also distinct collocations, and as well as positive and negative emotional associations (connotations). It is suggested that the phenomenon could be tackled in terms of cultural salience, a concept which needs further development" (p. 9).

2. Aims of the research

The aim of the present chapter is to tap into the line of color term research that stresses the incorporation of natural language use, on the one hand, and the significance of the cultural context, on the other hand. We wish to focus on four interrelated issues within Hungarian color terminology:

1. Establishing a more nuanced description of Hungarian basic color terms by analyzing the color terms' individual lexicalization patterns[5] in naturally occurring data;
2. Lessening the influence of the results obtained from color-naming tasks within color term research and promoting instead the reliance on the natural occurrence of color terms through the analysis of usage patterns;
3. Abandoning the relative rigidity and universalist attitude toward basic color terms and regarding instead color terms as language- and culture-specific categories;
4. Advocating a radial category structure for "basic color term", where certain color terms are more typical (central) members than others.

2.1 Hypotheses

On the basis of the anomalies discussed above and in light of the results of our previous research on Hungarian color terms (Benczes & Tóth-Czifra, 2014), it can be hypothesized that the category of Hungarian BCTs is far from being homogeneous. On the contrary, what we expect to find is substantial differences in the strength of the color terms' linguistic representations and their lexicalization patterns. Thus, following Kay and McDaniel (1978), we hypothesize the existence of more robustly

5. In line with Raffaelli (2017), we understand lexicalization patterns to subsume word-formation patterns and syntactic patterns employed by speakers for naming concepts.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

represented primary basic color terms and less robustly represented secondary basic color terms. We further hypothesize that the primary and secondary BCTs of Hungarian will be distinguishable along a number of linguistic parameters, which are the following:

1. There will be substantial differences with respect to the frequencies of the individual occurrences of primary and secondary BCTs in language use. Thus, we expect a much higher frequency of primary BCTs than that of secondary BCTs.
2. We also expect primary BCTs to be more productive in language, allowing for a wider range of schemas, which will be made manifest in their higher type frequencies of the most entrenched color term lexicalization pattern, namely [N/Adj + color term] constructions.
3. It can also be hypothesized that the diverging trends of primary and secondary color terms in linguistic representation and lexicalization patterns do not reside in the physiological salience of the color domains denoted by the respective color terms.[6] Rather, reasons for the differences can be sought in the diverging paths of idiomatization that these color terms have taken in Hungarian. At the same time, this latter feature will heavily depend on when a particular color term entered the language (and, accordingly, how much time it had to undergo idiomatization).

3. Methodology

As a first step, we collected the first occurrences of the respective color terms from the *Hungarian Etymological Dictionary* (Benkő, 1967–1984). As a second step, we performed a corpus linguistic analysis by first searching for the individual occurrences of each Hungarian BCT, and then for the respective [N + color term] and [Adj + color term] sequences. Examples of the former are *pulykavörös* ("turkey + red"), or *csokoládébarna* ("chocolate + brown"), examples of the latter include *haragosvörös* ("angry + red"), or *melegbarna* ("warm + brown"). Data were extracted from the Hungarian Gigaword Corpus (HGC, Oravecz & al., 2014).[7] We excluded data that were a combination of colors of any sort – either in a nominal form (such as *piros-fehér-zöld*, "red-white-green") or in an adjectival form (e.g., *fehéreskék*, "whitish blue"). We also excluded the very general [*sötét* ("dark") + color term] and [*világos* ("light") + color term] from among the construction types. Orthographical

6. Nevertheless, we do not wish to claim that the diverging patterns of lexicalization are wholly unrelated to physiological salience – see, for example, Uusküla and Sutrop's (2007) discussion on the physiological, cognitive, and perceptual salience of red.

7. This is the updated version of the Hungarian National Corpus.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

differences (with or without hyphenation, such as *ezüst-kék* and *ezüstkék*, "silver blue") were disregarded.[8]

Needless to say, our approach is very much rooted in a usage-based theory of language (which is also within the framework of the cognitive linguistic enterprise; see e.g. Langacker 1987). This framework considers grammar to be "the cognitive organization of one's *experience with language*" (Bybee, 2006, p. 711; emphasis added). While grammar is still seen as abstract, it is, nevertheless, rooted in the experience that a particular speaker has with language – and which entails that grammar is not static or fixed but constantly shaped by usage from which it also arises (referred to as "emergent grammar"; Hopper, 1987). One of the consequences of a usage-based approach is that it draws on naturally occurring linguistic data,[9] extracted from corpora. Within this spirit, our main priority in this chapter is to show how the dynamics of language, in the form of type frequency (i.e., the number of construction types that are instantiated by a particular lexicalization pattern or schema), can shape category structure. Following Bybee (2001), productivity will be understood here as being dependent on type (and not token) frequency.

In light of the debate surrounding what exactly counts as a BCT (see the Introduction above), it is necessary at this point to define what we consider in this chapter as a basic color term. Taking into consideration the main hypothesis of the present study that the category of BCTs cannot be rigidly defined and can be best described as a radial structure, we have decided to adopt a looser interpretation of what counts as a basic color term in Hungarian and will also include the problematic *vörös* in the study so that its status can be compared to that of the other color terms (we will return to the question of whether it is indeed a BCT or not in the discussion of the corpus-based analysis). Accordingly, the Hungarian BCTs under analysis in this research are the following: *fehér* "white", *fekete* "black", *vörös* "red$_1$", *piros* "red$_2$", *zöld* "green", *sárga* "yellow", *kék* "blue", *barna* "brown", *lila* "purple", *rózsaszín* "pink", *narancssárga* "orange", and *szürke* "gray".[10]

---

8. Due to space limitations, we are unable to provide a semantic analysis of the constructions themselves. The total number of types for all BCTs was 534; the total number of tokens was 45,075. With regard to the latter, the token frequencies (in % of the total) were the following: *vörös* ("red$_1$"): 17%; *zöld* ("green") and *fehér* ("white"): 16%; *kék* ("blue"): 15%; *sárga* ("yellow"): 12%; *barna* ("brown"): 8%; *fekete* ("black"), *piros* ("red$_2$"), and *szürke* ("grey"): 5%; *lila* ("purple") and *rózsaszín* ("pink"): 1%; *narancssárga* ("orange"): 0%.

9. As opposed to color naming or dictionary-based approaches. Other types of naturally occurring data might include recordings of spoken language or interviews; however, as far as we know, a lexically and grammatically annotated large-scale corpus of spoken Hungarian does not yet exist.

10. Note that these are the color terms that are typically accepted and referred to as being "basic" in the literature.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4. The emergence of Hungarian basic color terms

We will begin the investigation of the validity of our claims by focusing on the third hypothesis (which, from a historical point of view, enjoys priority over the other two hypotheses). The first occurrence of the respective BCTs is depicted in Table 1.

Table 1. First occurrence of Hungarian BCTs (source: Benkő, 1967–1984)

Color term	First occurrence	Etymology
fehér "white"	1055	Finno-Ugric
fekete "black"	1055	Finno-Ugric
vörös "red$_1$"	1121	Finno-Ugric
zöld "green"	1215	Origin uncertain; perhaps an Alanic or Iranian borrowing.
kék "blue"	1216	Old Turkic; later on used for greenish and greyish hues as well, but these senses eventually disappeared from Hungarian.
piros "red$_2$"	1237	Finno-Ugric
barna "brown"	1255	Saxon; originally meaning "dark".
szürke "grey"	1258	Finno-Ugric
sárga "yellow"	1332	Old Turkic; originally meaning "pale".
rózsaszín "pink"	1513	Through German mediation; originally referring to a redder hue.
lila "purple"	1807	Through German mediation.
narancssárga "orange"	1845	Source: Hungarian Historical Corpus. The etymological dictionary only mentions *narancs*, which appeared in 1481 (through Italian mediation).

It becomes immediately clear from Table 1 that the color terms defined as "basic" straddle a very wide period when it comes to their first occurrence in the language. There is nearly an eight-hundred-year difference between the appearance of the oldest (*fekete* "black", *fehér* "white") and the youngest (*narancssárga* "orange") color terms. The sequence of emergence of Hungarian color terms can be placed into four larger time frames. The earliest of these periods, lasting from 1055 to the beginning of the 12th century, witnessed the appearance of *fekete* "black", *fehér* "white", and *vörös* "red$_1$", all of which can be found in the oldest written records of Hungarian. The second wave of color terms, from the beginning of the 13th century to the first half of the 14th century, roughly coincides with the middle era of the Old Hungarian period (896–1526). This is when the first occurrences of *zöld* "green", *kék* "blue", *piros* "red$_2$", *barna* "brown", *szürke* "grey", and *sárga* "yellow" appeared in the language. While the first two periods can be relatively well-defined and are also situated quite close to one another, there is a substantial time gap between the second and third waves; *rózsaszín* "pink" appeared nearly two hundred years later than *sárga* "yellow" (the last color term of the second wave), toward the very end of

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

the Old Hungarian period. There is an even more substantial gap between the third and fourth waves; the youngest basic color terms of Hungarian, *lila* "purple" and *narancssárga* "orange" date back to the 19th century and are the linguistic products of the Modern Hungarian period.

Interesting observations can be drawn by contrasting the diachronic data of the emergence of Hungarian BCTs and the evolutionary sequence of color terms as provided by Berlin and Kay (1969) – see Table 2. It seems that the sequences more-or-less coincide with one another. Although the Hungarian sequence does traverse a very wide time frame, it can, nevertheless, be regarded as an evolutionary/diachronic sequence of the emergence of basic color terms in the language.

Table 2. Comparison of the emergence of Hungarian BCTs and Berlin and Kay's (1969) evolutionary sequence of color terms

First occurrence of Hungarian BCTs	Berlin and Kay's (1969) evolutionary sequence
fekete "black" (1055)	black, white
fehér "white" (1055)	
vörös "red$_2$" (1121)	red
zöld "green" (1215)	green, yellow, blue
kék "blue" (1216)	
piros "red$_1$" (1237)	
barna "brown" (1255)	brown
Szürke "grey" (1258)	grey, purple, pink, orange
sárga "yellow" (1332)	
rózsa-szín "pink" (1513)	
lila "purple" (1807)	
narancssárga "orange" (1845)	

Two major differences arise in conjunction with the two sequences. The first of these concerns the emergence of the two color terms for "red", *piros* and *vörös*. Diachronically *vörös* appeared first in the language (and the sequence of its appearance also coincides with the universal evolutionary sequence). It kept its dominance up until the 19th century, after which *piros* became increasingly prominent and productive in the language (the linguistic consequences of this change are discussed in detail in Benczes and Tóth Czifra, 2014). Unfortunately, we cannot provide any explanation with regard to the other main difference in the respective sequences, i.e., the relatively late appearance of *sárga* "yellow" (as compared to the evolutionary sequence), due to the ambiguity and lack of etymological data.[11]

11. Note that the first occurrence of *sárga* is ambiguous in the literature; two dates are provided in Benkő (1967–1984), 1055–1332. However, since the former date is depicted with a question mark, we have decided to accept the second date.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Drawing on the premise that older linguistic units typically have more time to undergo lexicalization than more recent ones, the diachronic sequence of Hungarian BCTs suggests that the lexicalization patterns of the older color terms (i.e., those of the first and second wave, i.e., *fekete* "black", *fehér* "white", *vörös* "red$_1$", *zöld* "green", *kék* "blue", *piros* "red$_2$", *barna* "brown", and *sárga* "yellow") will be distinctively stronger than those of the younger color terms (i.e., the third and the fourth wave, i.e., *rózsaszín* "pink", *lila* "purple", and *narancssárga* "orange"). In other words, what we expect is that the [N/Adj + color term] constructions of the older color terms will have a higher type frequency as compared to the younger color terms. This issue will be investigated in Section 5.

5. Results of the corpus-based analysis

As a first step, we investigated the frequency of the individual occurrences of each BCT in the corpus. This particular variable correlates most strongly with synchronic language use; note that constructions – especially compound forms – typically conserve a former stage of language. We do not wish to claim here that frequency should be regarded as a reliable factor when it comes to delimiting basicness (as suggested by Forbes, 1979), but rather that the individual occurrences of BCTs can shed light on the possible salience of the conceptual category of a particular color term, as represented in current language use. The frequency data are presented in Table 3.

Table 3. Frequency of individual occurrences of Hungarian BCTs, as extracted from the HGC

Color term	No. of hits (n = 789,926)
fekete "black"	161,312 (20.4%)
fehér "white"	155,551 (19.8%)
zöld "green"	99,414 (12.6%)
vörös "red$_1$"	81,143 (10.3%)
piros "red$_2$"	69,811 (8.8%)
kék "blue"	63,189 (8%)
sárga "yellow"	52,969 (6.7%)
barna "brown"	41,006 (5.2%)
szürke "grey"	31,062 (3.9%)
lila "purple"	14,570 (1.8%)
rózsaszín "pink"	13,384 (1.7%)
narancssárga "orange"	6,515 (0.8%)

What can be immediately seen from Table 3 is that the individual frequencies of Hungarian BCTs also correspond to a considerable degree to the evolutionary

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

sequence of Berlin and Kay (1969). Thus, *fekete* ("black") and *fehér* ("white") are the most frequent color terms in the corpus, accounting for nearly 40% of the total number of color terms. The only outlier in the data is *zöld* ("green"), which ranked in the third position, coming before the color terms for "red", *vörös* and *piros*.[12] Comparing the data in Tables 2 and 3, it can be seen that the etymological data do correlate loosely with the frequency patterns. This is especially evident in the case of *fekete* ("black") and *fehér* ("white"), and the color terms of the fourth wave, i.e., *lila* ("purple"), *rózsaszín* ("pink"), and *narancssárga* ("orange"). Note, however, that the frequencies of *lila* ("purple") and *rózsaszín* ("pink") were near-identical, yet the latter preceded the former in the history of the language by approximately three hundred years. Yet the data overall suggest that the date of appearance of a color term in Hungarian does have an effect on the overall frequency of that particular color term in language use, implying a correspondence between the time that a particular color term has been in use and its overall salience in language.

Nevertheless, the question necessarily arises how these data look like when it comes to lexicalization patterns, i.e., constructions containing the basic color terms themselves. We hypothesized that there would be a higher type frequency of [N/Adj + color term] constructions for the color terms belonging to the first and the second wave of the evolution of color terminology in Hungarian, and a lower type frequency of the third and fourth waves. It is schema type frequency – and not token frequency – that correlates with productivity (see Bybee, 2001) since higher type frequency correlates with greater analyzability (i.e., language users are able to identify and generalize patterns more easily). Furthermore, frequently occurring structures "achieve a certain autonomy from related forms" (Bybee, 2001, p. 125). In other words, a higher number of [N/Adj + color term] construction types suggests a linguistically (and possibly cognitively) more salient color term, implying a solid BCT-membership. The data that we have extracted from the corpus are depicted in Figure 1, which shows the percentage of types of [N/Adj + color term] constructions with the respective color terms (the total number of types was 534), as extracted from the HGC. As it can be seen in Figure 1, all of the Hungarian BCTs, with the exception of *narancssárga* ("orange"), were represented in [N/Adj + color term] constructions.[13]

---

12. A brief survey of the hits with *zöld* ("green") indicates that the prominence of this color term can be linked to its figurative use in politics, as in *zöld párt* ("green party") and its metonymy-based meaning extension of "environmental", as in *zöld sziget* ("green island", referring to recycling centers used in cities).

13. *Kék* ("blue") and *zöld* ("green"): 16%; *vörös* ("red1"): 12%; *fehér* ("white"), *sárga* ("yellow"), and *szürke* "grey"): 10%; *barna* ("brown"): 9%; *fekete* ("black") and *piros* ("red$_2$"): 7%; *lila* ("purple") and *rózsaszín* ("pink"): 2%; *narancssárga* ("orange"): 0%.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

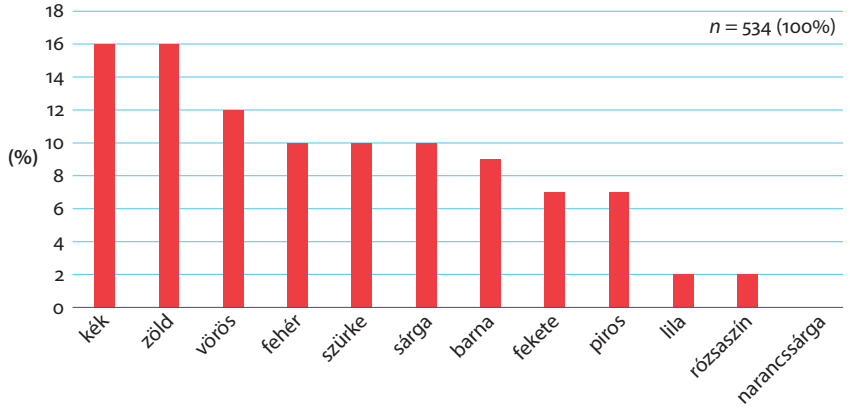


Figure 1. Percentage of [N/Adj + color term] construction types with the respective color terms, as extracted from the HGC

Looking at Figure 1, it also becomes evident that, similarly to the frequencies of individual occurrences of the color terms, substantial differences can be observed in the ratio of the [N/Adj + color term] construction types. As already emphasized above, and following Bybee (2001), we consider type frequency as an indicator of the productivity of a particular lexicalization pattern (or schema). Thus, the type ratio discrepancies among the color terms simply suggest that these schemas are not equally productive – i.e., their linguistic representation is not equally strong.

Here, the order of the color terms (*kék* "blue", *zöld* "green", *vörös* "red$_1$", *sárga* "yellow", *fehér* "white", *szürke* "grey", *barna* "brown", *piros* "red$_2$", *fekete* "black", *rózsaszín* "pink", *lila* "purple", *narancssárga* "orange") fits the evolutionary sequence of Berlin and Kay (1969) only partially. The relatively weak representation of the *fekete* ("black") and *fehér* ("white") lexicalization schemas can be accounted for by their achromatic characteristic (Volbrecht and Kliegl, 1998): i.e., that they are not, in fact, part of the color spectrum, which also implies the difficulty (and even futility) in further defining or partitioning the color domain they denote. Nevertheless, the two color terms show vastly different type ratios (and schema productivity) in comparison to one another: while the [N/Adj + *fehér*] construction instantiates fifty-five types, this number is only thirty-five in case of *fekete* ("black"). (A detailed semantic analysis would be required to establish the possible reasons for this difference – such an investigation exceeds the boundaries of the present study.)

The relationship between the lexicalization schemas of *vörös* ("red$_1$") and *piros* ("red$_2$") also deserves attention, since these exhibit notable discrepancies. *Vörös* ("red$_1$") licensed sixty-three [N/Adj + color term] construction types, ranking as the color term with the third-highest number of construction types. *Piros* ("red$_2$"), in turn, came in only at the eighth position, following *barna* ("barna") and even

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

szürke ("grey"),[14] licensing only thirty-eight construction types. We believe that the differences between the lexicalization patterns of the respective color terms can be easily accounted for by looking at the etymological data. *Vörös* ("red$_1$") appeared much earlier in the language than *piros* ("red$_2$"), allowing for a greater amount of time to undergo lexicalization and idiomatization in the form of [N/Adj + color term] constructions.

Narancssárga ("orange") presented one of the most intriguing results of the corpus-based survey. The color term did not license any lexicalization patterns – or at least not any pattern that formed a part of this research (the color term showed up only in the form of *sötét narancssárga* "dark orange" and *világos narancssárga* "light orange", but we excluded these types from the data – see the Methodology section). This particular result is in line with the fact that *narancssárga* "orange" is the youngest of the Hungarian BCTs and also has the lowest frequency of individual occurrences.

Thus, what the results show is that even though the three variables we have investigated (i.e., etymological data, individual occurrence of color term, and the productivity of lexicalization patterns) do not exactly match with one another, there is an evident correlation of all three with the evolutionary sequence of Berlin and Kay (1969). The relationship of each variable to one another is depicted in Table 4.

While there are obvious discrepancies among the various dimensions we have studied, Table 4 clearly indicates nonetheless that the variables do correlate with one another. Importantly, *rózsaszín* ("pink"), *lila* ("purple"), and *narancssárga* ("orange") fall toward the end of all of the dimensions, *barna* ("brown") and *szürke* ("grey") are somewhere in the middle, while *fekete* ("black"), *fehér* ("white"), *zöld* ("green"), *kék* ("blue"), *vörös* ("red$_1$"), and *piros* ("red$_2$") tend to be situated toward the first half of nearly all four scales. The combined investigation of the three variables has the following three major implications for BCT status:

1. The category of Hungarian BCTs has a heterogeneous structure – i.e., not all members of the category share all the properties. Central or typical members of the category are etymologically older color terms that have high individual frequencies and license productive schemas (in the form of [N/Adj + color term] constructions.
2. The members of the category of Hungarian BCTs are not of equal status. The more typical members of the category fall toward the first half of Berlin and Kay's (1969) evolutionary sequence, while less typical members can be found toward the second half of the evolutionary scale.

14. Note, however, that *grey* can emerge as a so-called "wild card" (Kay, 1975, p. 261) at any stage of the hierarchy.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Correlation of etymological data, individual occurrence, and productivity of lexicalization patterns with the evolutionary sequence of color terms

Productivity of [N/ADJ] + color term] schema	Frequency of individual occurrence	First occurrence of Hungarian BCTs	Berlin and Kay's (1969) evolutionary sequence
zöld	*fekete*	*fekete* (1055)	black
kék	*fehér*	*fehér* (1055)	white
vörös	*zöld*	*vörös* (1121)	red
sárga	*vörös*	*zöld* (1215)	green
fehér	*piros*	*kék* (1216)	blue
szürke	*kék*	*piros* (1237)	yellow
barna	*sárga*	*barna* (1255)	brown
piros	*barna*	*szürke* (1258)	grey
fekete	*szürke*	*sárga* (1332)	purple
rózsaszín	*lila*	*rózsaszín* (1513)	pink
lila	*rózsaszín*	*lila* (1807)	orange
narancssárga	*narancssárga*	*narancssárga* (1845)	

3. The category is by no means rigid in its structure. Thus, the position of the respective color terms can shift from variable to variable, but these changes are not that substantial. What this means is that *rózsaszín* ("pink"), *lila* ("purple"), and *narancssárga* ("orange") will not tend to fall toward the first half of the scales, and *zöld* ("green"), *kék* ("blue"), *fekete* ("black"), and *fehér* ("white") will not tend to fall toward the second half of the scales. Such an approach to BCT category structure is in line with Barsalou's (1983, 1993) proposal, according to whom categories are not necessarily stable mental representations but are context-dependent and have variable structure – and basic color terms are no exception to this idea.

Thus, what we propose is that the category of "basic color term" (at least in Hungarian) is not a rigid set of equal-status members. On the contrary, and in line with the

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

results of this research, we would like to suggest a more flexible – radial – category structure (Lakoff, 1987, p. 65), where not all color terms possess (or need to possess) the same properties (see Kerttula, 2007, for a similar notion of "relative basicness"). We, therefore, propose a differentiation between (a) primary BCTs, which are more typical and central to the category; and (b) secondary BCTs, which include less typical and more peripheral members and which bear family resemblances to the central members. The latter group includes *rózsaszín* ("pink"), *lila* ("purple"), and *narancssárga* ("orange"). The secondary status can also account for the fact that neither *rózsaszín* ("pink"), nor *narancssárga* ("orange") conforms to the criterion of morphological simplicity. With regard to primary BCTs, *fekete* ("black"), *fehér* ("white"), *zöld* ("green"), *kék* ("blue"), and *yellow* ("sárga") definitely fall into this group; the status of *barna* ("brown") and *szürke* ("grey"), however, is ambiguous. The position of these two color terms within the BCT category is dependent on which variable is prominent.

As for the problematic *vörös – piros* distinction of Hungarian, the present analysis and our previous research (Benczes & Tóth-Czifra, 2014) underline the flexibility of the BCT category: not only can the status of respective color terms shift between primary and secondary, depending on which variable it is focused on (as exemplified by *barna* "brown" and *szürke* "grey"), but color terms can also drop out of the BCT category and new ones can enter. We believe that this is the case with *vörös* and *piros*. We uphold our previous claim that *vörös* is no longer a BCT of Hungarian, as this status has now been overtaken by *piros* ("red_2") – see Benczes and Tóth-Czifra (2014) for a detailed analysis. What the present research indicates, nevertheless, is that *vörös* ("red_2") *used* to assume a primary status when it was still a BCT, and *piros*, "red_1", is the *current* primary BCT in Hungarian that refers to the "red" color domain.

6. Conclusions

At the start of the present research, we had two major aims in mind. On the one hand, we wished to reaffirm (or re-establish) the importance of the linguistic (and cultural) aspect in color term research. On the other hand, we also wished to question the rigidity of the category of "basic color term" and accordingly emphasize the flexibility and heterogeneity of the category by adopting a corpus linguistic analysis of Hungarian BCTs.

Our point of departure was the common observation in the color term literature that a number of languages failed to conform to the rigid criteria of basic color term status as laid down by Berlin and Kay (1969). Nevertheless, the discussion of these anomalies did not extend to the abandonment of the original BCT

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

category; instead, the studies simply focused on the determination of the basic *vs*. non-basic status of a particular (problematic) color term by adopting typically strict formal criteria based on binary distinctions. Here we wished to depart from past approaches by focusing on the *category* of (Hungarian) BCTs itself, by examining its internal structure and flexibility through the analysis of Hungarian basic color terms. Thus, following Kay and McDaniel (1978), we hypothesized the existence of more robustly represented primary basic color terms and less robustly represented secondary basic color terms. We further hypothesized that the primary and secondary BCTs of Hungarian would be distinguishable along a number of linguistic parameters, which were the following: (1) the first occurrence of the BCTs in Hungarian; (2) the frequency of the individual occurrences of Hungarian BCTs; and (3) the type ratio of [N/Adj + color term] constructions with the respective color terms.

According to our results, significant differences can be seen among the BCTs with respect to their linguistic representation, salience, and hence their degree of basicness. Although the three variables that we investigated did not match each other fully and there were discrepancies with regard to the order of the color terms on the respective scales, the sequences did show a loose correlation with the universal evolutionary scale established by Berlin and Kay (1969). Accordingly, *fekete* ("black"), *fehér* ("white"), *zöld* ("green"), *kék* ("blue"), and *yellow* ("sárga") can be considered as primary BCTs, since these color terms tend to be situated toward the first half of all the investigated scales (including Berlin and Kay's evolutionary sequence). In comparison, *rózsaszín* ("pink"), *lila* ("purple"), and *narancssárga* ("orange") are only secondary BCTs of Hungarian, as these typically fall toward the end of all the investigated variables. The status of *barna* ("brown") and *szürke* ("grey"), however, is ambiguous: the position of these two color terms within the BCT category is dependent on which variable is prominent. As for *vörös* and *piros*, both denoting "red", we uphold our previous claim that *vörös* ("red_1") is no longer a BCT of Hungarian, as this status has now been overtaken by *piros* ("red_2") – see Benczes and Tóth-Czifra (2014). What the present research indicates, nevertheless, is that *vörös* ("red_2") *used* to assume a primary status when it was still a BCT, and *piros*, "red_1", is the *current* primary BCT in Hungarian that refers to the "red" color domain.

These results highlight two important characteristics of the category of "basic color term": its heterogeneity and its dynamic nature. In other words, BCT category structure is in line with Barsalou's (1983, 1993) proposal, according to whom categories are not necessarily stable mental representations but are context-dependent and have variable structure – and basic color terms are no exception to this idea. The dynamics of the BCT category – as exemplified by the *piros–vörös* distinction or the relatively late appearance of *rózsaszín* ("pink") and *lila* ("purple") in

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hungarian – is a manifestation of linguistic and cultural changes. These changes, however, are not random and do not redefine the category itself (i.e., *rózsaszín* "pink" will never change places with *fehér* "white" on any of the scales); rather, they allow for a certain interplay of the various linguistic and cultural dimensions that affect the category itself. Due to the limits of the present study, we were able to investigate only three of these dimensions in detail, and our analysis was restricted to only one language, Hungarian. In order to arrive at a more complete picture of the category of "basic color term", further languages and dimensions will need to be added to the analysis.

References

Barlow, M., & Kemmer, S. (Eds.). (2000). *Usage-based models of language*. Stanford, CA: CSLI Publications.

Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition*, *11*(3), 211–227. https://doi.org/10.3758/BF03196968

Barsalou L. W. (1993). Structure, flexibility, and linguistic vagary in concepts: manifestations of a compositional system of perceptual symbols. In A. C. Collins, S. E. Gathercole, & M. A. Conway (Eds.), *Theories of memory* (pp. 29–101). London: Erlbaum.

Baxilina, N. B. (1975). *Istorija cvetooboznacenij v russkom jazyke* [History of color terms in Russian]. Moscow: Nauka.

Benczes, R., & Tóth-Czifra, E. (2014). The Hungarian colour terms piros and vörös: A corpus and cognitive linguistic account. *Acta Linguistica Hungarica*, *61*(2), 123–152. https://doi.org/10.1556/ALing.61.2014.2.1

Benkő, L. (1967–1984). A magyar nyelv történeti-etimológiai szótára [*A historical-etymological dictionary of the Hungarian language*]. Budapest: Akadémiai Kiadó.

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

Biggam, C. P. (2012). *The semantics of colour: A historical approach*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139051491

Bimler, D., & Uusküla, M. (2016). A similarity-based cross-language comparison of basicness and demarcation of "blue" terms. *Color Research and Application*, *42*, 362–377. https://doi.org/10.1002/col.22076

Bybee, J. (2006). *Frequency of use and the organization of language*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301571.001.0001

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511612886

Corbett, G. G., & Davies, I. R. (1995). Linguistic and behavioural measures for ranking basic colour terms. *Studies in Linguistics*, *19*, 301–357. https://doi.org/10.1075/sl.19.2.02cor

Corbett, G., & Morgan, G. (1988). Colour terms in Russian: Reflections of typological constraints in a single language. *Journal of Linguistics*, *24*, 31–64. https://doi.org/10.1017/S0022226700011555

Davies, I. R., & Corbett, G. G. (1997). The basic color terms of Russian. *Linguistics*, *32*, 65–89. https://doi.org/10.1515/ling.1994.32.1.65

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Forbes, I. (1979). The terms *brun* and *marron* in modern standard French. *Journal of Linguistics*, *15*(2), 295–305. https://doi.org/10.1017/S0022226700016406

Frumkina, R. M. (1984). Cvet, smysl, sxodstvo. Aspekty psixolingvisticeskogo analiza [*Color, meaning, and similarity: Aspects of psycholinguistic analysis*]. Moscow: Nauka.

Hopper, P. (1987). Emergent grammar. *Berkeley Linguistics Society*, *13*, 139–157. https://doi.org/10.3765/bls.v13i0.1834

Kerttula, S. (2007). Relative basicness of color terms: Modeling and measurement. In: R. E. MacLaury, G. V. Paramei, & D. Dedrick (Eds.), *Anthropology of Color: Interdisciplinary multilevel modelling* (pp. 151–169). Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/z.137.11ker

Kiefer, F. (2005). Van-e magyar jelentéstan? [Does Hungarian semantics exist?] *Magyar Nyelvőr*, *101*, 129–140.

Kay, P. (1975). Synchronic variability and diachronic change in basic color terms. *Language in Society*, *4*(3): 257–270. https://doi.org/10.1017/S0047404500006667

Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language*, 54, 610–646. https://doi.org/10.2307/412789

Kay, P., & Maffi, L. (1999). Color-appearance and the emergence and evolution of basic color lexicons. *American Anthropolgist*, *101*, 743–760. https://doi.org/10.1525/aa.1999.101.4.743

Kay, P., Berlin, B., Maffi, L., Merrifield, W. R., & Cook, R. S. (2009). *The World Color Survey*. Stanford, CA: Center for the Study of Language and Information.

Kiss, G. (2004). A piros, vörös és más színnevek használata a Magyar Nemzeti Szövegtár alapján [The use of piros, vörös and other colour terms on the basis of the Hungarian National Corpus]. In T. Gecső (Ed.), *Variabilitás és nyelvhasználat* [Variability and language use] (pp. 160–165). Budapest: Tinta Könyvkiadó.

Kiss, G., & Forbes, I. (2001). *Piros, vörös – red, rot, rouge*. In T. Gecső (Ed.), *Kontrasztív szemantikai kutatások [Studies in contrastive semantics]* (pp. 190–199). Budapest: Tinta Könyvkiadó.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago, IL: University of Chicago Press. https://doi.org/10.7208/chicago/9780226471013.001.0001

Langacker, R. W. (1987). *Foundations of Cognitive Grammar. Vol. I: Theoretical Approaches*. Stanford, CA: Stanford University Press.

Moss, A. E. (1989). Basic colour terms: Problems and hypotheses. *Lingua*, *78*, 313–320. https://doi.org/10.1016/0024-3841(89)90027-2

Oravecz, C., Váradi, T., & Sass, B. (2014). The Hungarian Gigaword Corpus. *Proceedings of LREC*. 1719–1723. https://doi.org/10.13140/2.1.3957.1840

Özgen, E. & Davies, I. R. L. (1998). Turkish color terms: Tests of Berlin and Kay's theory of color universals and linguistic relativity. *Linguistics*, *36*, 919–956. https://doi.org/10.1515/ling.1998.36.5.919

Paggetti, G., Menegaz, G., & Paramei, G. V. (2016). Color naming in Italian language. *Color: Research and Application*, *41*(4): 402–418. https://doi.org/10.1002/col.21953

Papp, E. (2012a). *A magyar színnevek terminológiai elemzése [Terminological analysis of Hungarian colour terms]*. PhD diss., University of Pécs, Hungary.

Papp, E. (2012b). A corpus linguistic study of English and Hungarian basic colour terms. In: V. Karabalić, M. A. Varga, & L. Pon (Eds.), *Discourse and dialogue/Diskurs und Dialog*. Frankfurt am Main: Peter Lang

Paramei, G. V. (1999). One basic or two? A rhapsody in blue. *Behavioral and Brain Sciences*, *22*, 967. https://doi.org/10.1017/S0140525X99462211

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Paramei, G. V. (2005). Singing the Russian blues: An argument for culturally basic color terms. *The Journal of Comparative Social Studies*, *39*, 10–38. https://doi.org/10.1177/1069397104267888

Raffaelli, I. (2017). Conventionalized patterns of colour naming in Croatian. In: K. C. Kovačević & S. L. Udijer (Eds.), *Applied linguistics research and methodology* (pp. 171–185). Frankfurt am Main: Peter Lang.

Taylor, J. R., Mondry, H., & MacLaury, R. E. (1997). A cognitive ceiling of eleven basic color terms. In: R. E. MacLaury (Ed.), *Color and cognition in Mesoamerica: Constructing categories as vantages* (Appendix IV, pp. 419–429). Austin, TX: University of Texas Press.

Uusküla, M. (2008). *Basic colour terms in Finno-Ugric and Slavonic languages: Myths and facts*. Tartu: Tartu University Press. https://doi.org/10.13140/RG.2.1.2237.6728

Uusküla, M. (2011). Terms for *red* in Central Europe: An areal phenomenon in Hungarian and Czech. In C. P. Biggam, C. A. Hough, C. J. Kay, & D. R. Simmons (Eds.), *New directions in colour studies* (pp. 147–156). Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/z.167.17uus

Uusküla, M., & Sutrop, U. (2007). Preliminary study of basic colour terms in modern Hungarian. *Linguistica Uralica*, *43*, 102–122. https://doi.org/10.1093/ijl/9.2.102

Uchikawa K., & Boynton R. M. (1987). Categorical color perception of Japanese observers: Comparison with that of Americans. *Vision Research*, *27*, 1825–1833. https://doi.org/10.1016/0042-6989(87)90111-8

Volbrecht, V. J., & Kliegl, R. (1998). The perception of blackness. In: W. G. K. Backhaus, R. Kliegl, & J. S. Werner (Eds.), *Color vision* (pp. 187–206). Berlin: De Gruyter. https://doi.org/10.1515/9783110806984.187

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lexicalization patterns in Slovak color naming

Katarina Dudová
Constantine the Philosopher University in Nitra

This chapter deals with lexicalization patterns in Slovak color naming on the basis of empirical data from the EoSS project (Evolution of Semantic Systems, 2011–2014, Max Planck Institute for Psycholinguistics, Nijmegen). Due to the developmental particularities of the Slovak color naming system, we can find word-formation, and other grammatical (e.g., syntactic) and semantic (e.g., metonymy) patterns featured in the creation of Slovak color terms in the structure of the contemporary Slovak language. For all these patterns, we use the term lexicalization pattern, which is involved in creating different semantic categories. Not all lexicalization patterns are equally productive in the division of the color spectrum. Therefore, we focus on their productivity, as well as on the dominant and peripheral types of meanings that have cultural roots.

Keywords: lexicalization patterns, color terms, semantic categories, Evolution of Semantic Systems, derivation, compounding, West Slavic languages, Slovak

1. Introduction

The diversity and specificity of the world of colors have attracted human attention since time immemorial and it is therefore quite natural that this field of interest spans across several disciplines (biology, zoology, ontology, cultural anthropology, arts, etc.). The aim of our research is to uncover in the structure of present-day Slovak the types of the grammatical (e.g., syntactic) and semantic (e.g., metonymic) patterns that participate in the creation of Slovak color terms. For all these patterns, we use the generic term lexicalization pattern, which is involved in creating different semantic categories (Talmy, 1985, p. 57; Raffaelli & Kerovec, 2017, p. 26). Not all lexicalization patterns and semantic categories are equally productive in the dividing of the color spectrum in Slovak, and therefore we focus on their productivity, as well as on the dominant and peripheral types of meanings that have cultural justification for semantic patterns in particular.

https://doi.org/10.1075/sfsl.78.03dud
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

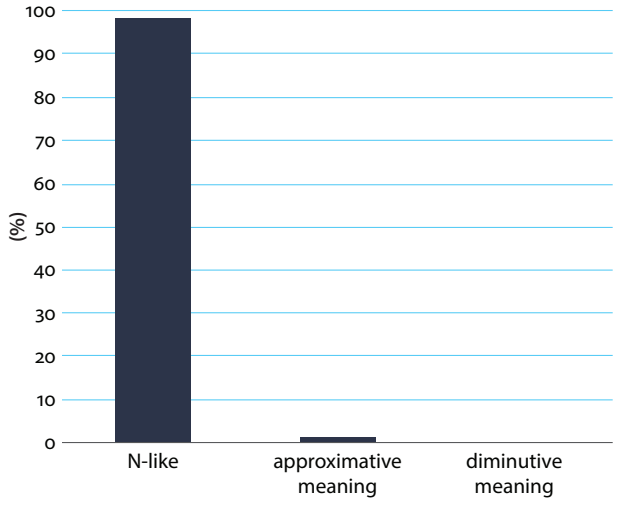


Figure 1. Productivity of meaning categories formed by derivation (according to EoSS results).

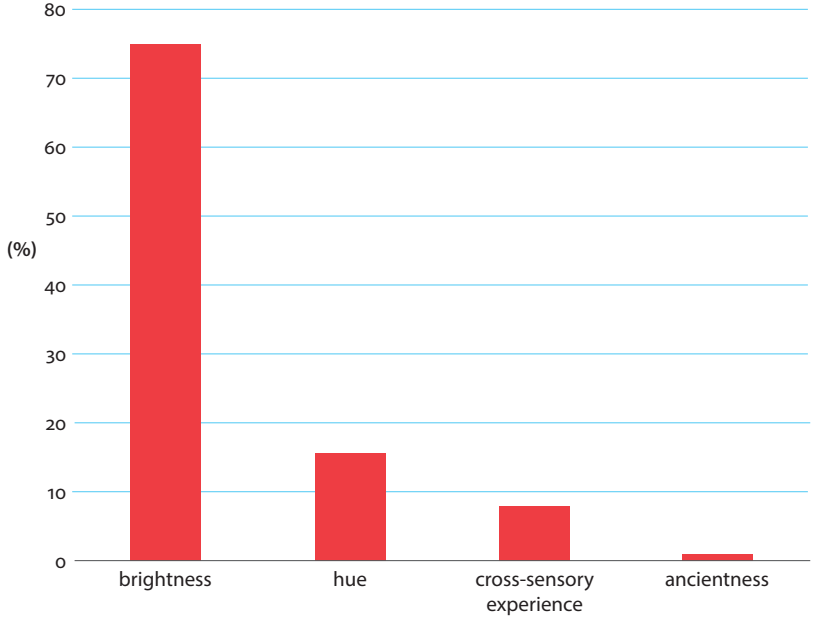


Figure 2. Productivity of semantic categories formed by compounding (according to EoSS results).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The main subject of the paper is the Slovak material acquired over four years of research participation in the project EoSS[1] (Evolution of Semantic Systems, 2011–2014, Max Planck Institute for Psycholinguistics, Nijmegen), which focused on the semantic changes over time and space. Data obtained from more than 50 Indo-European languages were collected and elaborated upon a united methodological basis which concentrated on four different types of categories: containers (kinds of objects); parts of the body (parts of objects); spatial relations (how objects are related to one another); and also colors (attributes of objects).

Empirical results related to the attributes of objects are the basis of the research on lexicalization patterns in Slovak color naming. Three types of analysis are presented in the paper: (a) the diachronic development of color terms, (b) the investigation of lexicalization patterns in Slovak in general and (c) the EoSS experiment investigating color terms and lexicalization patterns that were elicited.

2. The cultural and historical aspects of color naming

2.1 Influence of culture on color spectrum differentiation

Color is a natural feature of the perception of a subject who is able to see the colorfulness of their surroundings almost from birth. The fact that colors are among the most basic perceptual qualities of human beings points to the fact that their visual perception is biologically determined. What is more, however, is the fact that color perception is also conditioned by non-biological, cultural factors (Démuth, 2005, p. 133), which do not only affect attitudes towards individual colors or association with certain emotions (Dorn, Willalva, Giouli, Blanck, Wiebke, Kovalenko & Wandl-Vogt, 2016, pp. 237–243)[2] or the semantic extensity of the lexemes indicating color in the linguistic image of the world (Vaňková, 2003, pp. 69–99; Tokarski, 1995; Waszakowa, 2000, pp. 59–72, etc.), but also the overall division of the color space – the color classification located in grammatical-semantic structures of language.

This subject has so far yielded several more or less different perspectives and approaches, which have only gradually developed since the second half of the 19th century. Comparative studies of the early 20th century (e.g., the Sapir-Whorf

1. However, other material sources were used, and thus not all of the described lexicalization patterns are elicited in the EoSS data.

2. On the basis of the European Network of e-Lexicography (www.elexicography.eu), the authors produce results pointing, for example, to the connection of the red color with anger in different European languages, including Russian.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

relativity hypothesis on the conditionality of the language structure and conceptual processes of an individual) brought some initiative into the field of research on cultural determination of the specific division of the color spectrum, which is realized in individual languages with their own lexicon and semantic-grammatical structure.

However, the pioneering work of B. Berlin and P. Kay, *Basic Color Terms: Their Universality and Evolution* (1969), provided more important findings. On the basis of comparing color naming in 98 languages or dialects, the authors uncovered a universal color list that is conceptualized across cultures and is grouped into seven periods reflecting the evolutionary stage of culture represented by the language. These studies began a discussion that continues between the opinions of relativists on the one hand and universalists on the other. According to the universalists, the color categories are organized around the focal colors, or more precisely the color space is organized around six primary universal colors (black, white, red, green, yellow and blue) to which we easily attach prototypes and which are found in all languages in the division of color space in similar positions (Berlin & Kay, 1969). On the other hand, the relativists argue that language conventions, widely differing across languages, have an influence on the partition of the color spectrum (Roberson, Davies & Davidoff, 2000, pp. 369–398; Lucy, 1997, pp. 320–346).

In addition to these opposing approaches, there is also a third opinion (Jameson & D'Andrade, 1997, pp. 295–319), according to which color terms in individual languages are determined in part by general principles of categorization, but there is also a language convention that can cause irregularities in optimally shaped color space. These perceptual limitations result in a potentially unlimited repertoire of foci, which means every color is "focal" (although some more than others) (Regier, Kay & Khetarpal, 2007, pp. 1436–1441).

The development of the earlier category system in a language is strongly associated with the shaping of the color space that reflects color terms. This is confirmed by the opinion of psychologist Christine Ladd-Franklin (1929), according to which the overall process of color perception is largely influenced by evolutionary development.[3] The developmental particularities of the color naming system in a language show that Berlin and Kay's findings on universal color inventories cannot

3. Unlike in Berlin and Kay´ s notion of color evolution in the Ladd-Franklin´s theory, this is an evolutionary development based on biological evolution. According to Ladd-Franklin the human eye carries fragments of its earlier evolutionary development. Ladd-Franklin observed that the most highly evolved part of the eye is the fovea, where, at least in daylight, visual acuity and color sensitivity are greatest. She assumed that peripheral vision was more primitive than foveal vision because night vision and movement detection are crucial for survival. Such color vision evolved in three stages: 1. achromatic vision (the majority of the population are not affected by black-white color blindness), 2. blue-yellow sensitivity (blue-yellow color blindness affects a

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

be overlooked; on the contrary, they have to be considered significant, however, not to such extent to solely constitute the final results, because the partition of the color spectrum is also determined by culture and its language conventions.

2.2 Developmental particularities of the Slovak color naming system

The system of color naming in a language has its own history – evolving from an earlier system of categories. Based on the opinions of Berlin and Kay (1969), each society divides the color spectrum into at least two basic parts, white color with associated light colors and black color with dark colors. The more highly developed cultures (measured by our Euro-American criterion) differentiate more colors in their respective language (Vaňková, 2005, p. 41). Etymological dictionaries confirm this fact also in the Slovak language.

Derivates of *čierny* 'black'and *biely* 'white', which Berlin and Kay put into the first stage of cultural development are the oldest, documented in the 12th century; but the names of the colors themselves existed and were used earlier (*12th century). In Old Slavonic, these colors have the form of **čъrnъa *bělъ (*bhēlo-)*, from which other lexical units were derived, e.g., *čierňava* 'smuttiness', *černň* 'charcoal black', *černica* 'blackberry', *beľmo* 'sclera', *belasý* 'sky-blue', *belieť sa* 'to be white', *bieliť* 'albicate', *bielok/bielko* 'white', etc. (Králik, 2015, pp. 69, 104). Three centuries later, *červená* 'red' became known (from the 15th century), which belongs to the second evolutionary stage and is derived from the participle of the unexemplified verb **červiť* (in Old Slavic **čъrviti*, to colorize by coccid–to obtain red dye from a special kind of insect (Králik, 2015, p. 103; Machek, 1968, p. 99), as well as *zelený* 'green' (from the 15th century) from the participle **zel-ti* 'zelenieť sa, to be green' (Králik, 2015, p. 684), which Berlin and Kay place in the third development stage.

The other color terms, which according to Berlin and Kay represent the 4th–7th evolutionary degrees, have been documented since the 17th century. *Žltý* 'yellow' (since the 17th century), in Old Slavonic *žъltъ*, comes from the Indo-European expression **ĝhel-* with the meaning 'gleam' (Králik, 2015, p. 698). The explanation of the origin of the term *modrý* 'blue' (from the 17th century), in Old Slavic **modrъ*, is not coincidental: (a) it may be from the Indo-European **mad-* with the meaning 'wet, climb, drop'; (b) it is related to the English *madder* (Králik, 2015, p. 365). The Old Slavonic **gnědъ*, now *hnedý* 'brown' (from the 17th century), is related to the Old Slavic verb **gnětiti* 'heat' and with the semantic shift 'burn'> brown (Králik, 2015, p. 200). Old Slavonic (**sivъ, *šědъ*) has the names *sivý/šedý* 'gray' (from the

small population), 3. red-green sensitivity (many people suffer from red-green color blindness). (Hergenhahn, 2009, pp. 243–244).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

17th century), which originate from the common Indo-European designation of various, especially darker shades, which expanded with the Indo-European **kie-, *ki-* 'darkness' (Králik, 2015, pp. 532, 576).

In the following color terms (*fialový* 'purple', *ružový* 'pink', *oranžový* 'orange') that Berlin and Kay also include in the 7th stage, an Old Slavonic origin has not been reconstructed, and derivatives from nouns designating the entities *ruža* 'rose' (from the 16th century), *pomaranč* 'orange' (from the 17th century), and *fialka* 'violet' (from the 19th century) which developed later. All the colors mentioned above (*biely* 'white', *čierny* 'black', *červený* 'red', *zelený* 'green', *žltý* 'yellow', *modrý* 'blue', *hnedý* 'brown', *sivý/šedý* 'gray', *fialový* 'purple', *ružový* 'rose', *oranžový* 'orange'), which are considered in the Berlin-Kay theory as universal (focal) colors, proved in our research[4] to be basic colors. The colors emerging in the 7th stage of development have two distinct suffixes *-ý, -á, -é* (*siv-ý* 'gray') and *-ový, -ová, -ové*[5] (*fial-ový* 'lilac', *ruž-ový* 'pink', *oranž-ový* 'orange'), which reveal roots (e.g., Latin, Greek, German, French, etc.) other than Old Slavonic origin. In the inventory of basic colors, the EoSS research on the Slovak language also added the colors *tyrkysový* 'turquoise' and *béžový* 'beige', which are of non-Slavic origin (loanwords from non-Slavic languages).

It appears that the suffix may be a relevant indicator in the synchronic appearance of the language, referring to the origin of the color term as well as to the differentiation of the color spectrum in the diachronic aspect. *Tyrkysová* 'turquoise' is a derivative of the German *Türkis* taken from the French *turquoise* 'turkish stone' (Králik, 2015, p. 637); *béžová* 'beige' comes from the French *beige*; its origin is unclear (Králik, 2015, p. 68). However, it is also similar in other derivatives: *lilavá* 'lilac' is from the French *lilas/lilac* through the German *lila* (Králik, 2015, p. 328); *bordová* 'claret, wine color' comes from the French 'Bordeaux' (from the region of Bordeaux in southwest France) (Králik, 2015, p. 77); *purpurová* 'purple' is derived from the Latin *purpura*, taken from the Greek *porfyrā* (Králik, 2015, p. 484).

4. In the research, we obtained a list of Slovak basic colors in the following way: from five people (not participants in the research) for whom Slovak is a native language, we requested a list of colors with the following instructions: "Please write/state a list of all the colors you can think of as quickly as possible – you have two minutes". All of these data were centrally processed at the EoSS office according to the frequency criteria and position on the list of each questioned person.

5. The color adjectives change their Slovak grammar gender according to the grammar gender of the superior word (e.g., masculine: *siv-ý dom, fial-ový dom*; feminine: *siv-á brána, fial-ová brána*; neuter: *siv-é okno, fial-ové okno*).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

All color terms are used as adjectives in present-day Slovak. Some of these terms are used with a lower productivity rate[6] in the original uninflected form without a suffix (*lilavá/lila, bordová/bordó, purpurová/purpur*), which indicates the effort to emulate through the suffix the developmentally older color terms with Old Slavonic roots; sometimes they have a synonymous domestic equivalent expressed in composite (*bordová–tmavočervená, vínovočervená; purpurová–sýtočervená, lilavá–svetlofialová, tyrkysová–modrozelená; béžová–žltohnedá*).

The diversity of language constructions participating in the partition and differentiation of the color spectrum is rooted in culture and reflects the developmental particularities of the language. In color naming, languages have different lexical, semantic and grammatical patterns that are involved in the creation of specific meaning categories. Our goal is to identify these patterns and meanings in the Slovak language and to point out how they behave in dividing the color space.

3. Lexicalization patterns in color naming

3.1 Basic meanings expressed by derivation and composition

The term lexicalization pattern includes types of word-formation, and other grammatical (e.g., syntactic) and semantic (e.g., metonymy) patterns used for color naming in language. Research of color naming in Slovak pointed to these basic types of lexicalization patterns (comprising derivation, compounding, as well as other syntactic patterns).[7] These are lexicalization patterns, which, in conjunction with other syntactical patterns, contribute to the formation of different types of meaning in color naming. The meanings expressed primarily through the derivation are different from the meanings formed primarily by compounding, but some of the categories of meaning can also be expressed secondarily by other lexicalization patterns (derivation or compounding).

There are three main types of meaning conveyed via derivation in Slovak: (1) N-like; (2) approximative meaning; (3) expressive (diminutive) meaning.

1. 'N-like, N entity + suffix *-ový* / *-ová* / *-ové* (e.g. *fialový* 'violet-like', *olivový* 'olive-like'), when the noun for the familiar entity is used as a color term. There is a lexicalized metonymy OBJECT COLOR FOR COLOR, which is based on the common relationship between an entity and color. In this case, nouns

6. The Slovak National Corpus (version prim-7.0-public-all) reports, for example, for the color *bordová* 'claret' (1222 frequency) / *bordó* (78 occurrences), *purpurová* 'magenta' (3311 frequency) / *purpur* (815 frequency).

7. This overview of patterns is independent of the EoSS data.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

for entities (stems in the formation of color terms) indicate, e.g., fruit (*broskyňová* 'peach-like'), vegetables (*hrášková* 'pea-like'), fish (*lososová* 'salmon-like'), flowers (*orgovánová/lilavá* 'lilac-like'), etc. The pattern N entity + suffix is expressed by an adjective (adjective + noun) as one word (*tyrkysová + košeľa* 'turquoise-like + shirt') or a construction with sentence constituents in a subordinate relation (*trávová zelená* 'grass-like green'). In Slovak, this syntactic construction can be replaced by a compound Adv [N entity + suffix *-ovo*] + Adj, e.g., *trávovo-zelená* 'grass-like – green'. However, we also encounter color naming where the meaning 'N-like' loses its adjectival form, which is signaled by the suffix and is substantivized (*limetka* 'lime', *losos* 'salmon').

2. 'Approximative meaning' is formed via suffix *-(k)avý, -(a)stý* (*zelen-(k)avý/zelen-(k)astý* 'greenish'), or prefix + suffix *-na-N-(a)stý* (*na-modr-(a)stý* 'bluish'), *na-N-(a)lý* (*na-zelen-(a)lý* 'greenish'), *na-N-(a)vý* (*na-červen-(a)vý* 'reddish'). However, this meaning can also be expressed in Slovak by the prepositional phrase *do modra, do zelena, do červena*, etc. (E.g., *Očko na prsteni má modrastý/modravý nádych. / Očko na prsteni má nádych do modra.* 'The eyelet on the ring has a bluish tint.')
3. 'Diminutive meaning' is expressed by suffixes *-učký, -unký, -ulinký* (*modr-učký/modr-unký/modr-ulinký*); the suffix *-ušký* (*zelenuškýý*) is used marginally.

When it comes to derivation, suffixes are mainly used (very marginally prefix + suffix), which are specific for each meaning category. On the other hand, a different degree of synonymity is observed, because one meaning type can be expressed by several lexicalization (grammatical or semantic) patterns: e.g., 'approximative meaning' is formed via the sufix *zelen-(k)avý*, prefix + sufix *na-zelen-alý*; the prepositional phrase *do zelena;* the meaning N-like is expressed by the adjective (*lososová*) or substantive (*losos*).

Compounding enables more creativity in Slovak color naming than derivation and allows for a more fine-grained naming of the color spectrum. In this word-formation pattern, modifications refer to (1) brightness, (2) hue, (3) cross-sensory experience and (4) ancientness.

1. 'Brightness' is expressed by the adverbs *tmavo-* 'dark', *svetlo-* 'light' and *bledo-* 'pale' in compounds, such as *tmavomodrý* 'dark blue', *svetlomodrý* 'light blue', *bledomodrý* 'pale blue'. However, there are also syntagms with graded adjectives (*tmavá/tmavšia modrá* 'dark/darker blue', *svetlá/svetlejšia modrá* 'light/lighter blue') that the Slovak language uses to express this meaning.
2. 'Hue' is formed via a subordinate relation between two color terms in compounds, e.g., the Slovak *zelenožltý* 'greenish-yellow'. The adverb (as modifier) expressing the color *zeleno-* specifies the basic color term expressed by the adjective *-žltý*.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. 'Cross-sensory experience' is characterized by a similar word-formation pattern to modifiers referring to hue and brightness. However, an adverb indicates: (a) intensity (*silnožltý* 'strong yellow', *slabo-* 'weak', *prenikavo-* 'strident', etc.) connected with (b) acoustic experience (*krikľavozelený* 'garishly green'), (c) consistency (*tuho-* 'fast'), (d) haptic experience (*zamatovo-* 'velvety'). In Slovak, color naming does not apply to the experience of taste and aroma.
4. 'Ancientness' is lexicalized by the adverbs *staro-* 'old' and *šedivo-* 'gray', which form the compounds *staroružový* 'old rose' and *šedivozelený* 'grey green'.

Not all lexicalization patterns in contemporary Slovak color naming are equally productive. We observe the dynamics in the preferences not only of individual lexicalization patterns but also of the meaning categories that are involved in the distribution of the color space by Slovak native speakers.

3.2 Productivity of lexicalization patterns and meaning categories

The productivity of lexicalization patterns and meaning types is monitored through research conducted within the framework of the international EoSS project (Evolution of Semantic Systems). One of the basic goals of the project was to find out what color terms speakers use when dividing the color spectrum. In addition to data collected from more than 50 Indo-European languages, we obtained data from 20 adult Slovak respondents (10 men and 10 women, aged 19–29) on a single methodological basis, asking for color terms according to the Munsell Color Scale (Majid & Levinson, 2007). The research material consisted of 84 gray chips in numerical order on which individual color samples were glued. All the color chips arranged in a box were shown to each participant who had to answer the question: "What color is this?" The respondents answered with the first color term in their native language (Slovak) that came to their mind when looking at a specific color chip; they were not to use long color descriptions. All responses from each respondent were recorded and then rewritten on to the coding sheets that were centrally processed for all languages participating in the comparative research EoSS in 2011–2013.

The results of this empirical research led us to investigate the productivity of the lexicalization patterns used by Slovak speakers in naming colors, expressing different categories of meanings. This research showed that in Slovak, out of a total of 1038 respondents' answers, 424 color terms (40.8%) were formed by derivation and 605 color terms (58.3%) by compounding. Although the difference between the two observed lexical patterns is not significant in Slovak, it shows that compounding offers speakers better opportunities to differentiate the color spectrum and thus has the potential to be a more appropriate lexicalization pattern for more accurate color naming in Slovak. In addition, a small group (9 responses) represented

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

unconventional (idiosyncratic) color terms, which were formed by a combination of several lexicalization patterns (0.9%), e.g., *tmavá modrozelená* 'dark bluish green', *tmavá olivovozelená* 'dark olive-like green', *sýtosvetlomodrá* 'deep light blue', etc.

Significant differences have emerged in meaning categories. Within the derivations, we noted the significantly dominant meaning category 'N-like', i.e., N entity + suffix *-ový/-ová/-ové* (e.g., *horčic-ový* 'mustard-like', *lososový* 'salmon-like'). Each respondent recorded at least 14 (but often more, up to 28) color terms with this meaning type. Of all the color terms formed by derivation, the significant type 'N entity+suffix' represented up to 417 responses (98.35%). They were mostly one-word terms, but in 16 replies, other syntactic patterns competed with each other: a compound word composed of elements in a determining relationship (e.g., *oblačnemodrá* 'cloud-like blue', *nebeskymodrá* 'sky-like blue', *slnečneružová* 'sun-like rose', *orgovánovofialová* 'lilac-like violet', etc.) or a determinative syntagma (e.g., *tehlová červená* 'brick-like red', *trávová zelená* 'grass-like green', *námornícka modrá* 'marine-like blue', etc.).

A peripheral meaning category expressed exclusively by derivation was 'approximative meaning' (e.g., *zelenkastý* 'greenish', *fialovkastý* 'purplish', *hnedastý* 'brownish', etc.), which had a minimal representation, appearing in the answers of only two respondents (1.65% of all terms formed by derivation). In these responses, only the suffix *-(k)astý* was applied and other possibilities of formation (by the suffix *-(k)avý* or prefix + suffix *na-N-(a)stý, na-N-(a)lý*, *na-N-(a)vý* or a prepositional phrase e.g., *do modra*, *do zelena*, *do červena*) were not used. Participants did not attach the diminutive meaning to any colors shown, which can be deduced from the neutral function of their responses.

In the next lexicalization pattern, significant quantitative differences were also revealed between the different meaning categories. The dominant meaning formed via compounding was 'brightness' in Slovak (74.9%) of all the color terms formed by compounding) expressed by the adverbial components of *tmavo-, svetlo-, bledo-* in compound words, e.g. *svetlomodrý* 'light blue', *tmavofialový* 'dark purple', *bledoružový* 'pale rose' and others. This meaning category was encountered in 453 responses of 19 respondents, who mentioned it at a different frequency (6 to 36 responses per respondent). Respondents uniquely (16 times) expressed this meaning category also by multiword expressions, with a graded adjective (*tmavá* 'dark', *tmavšiamodrá* 'darker blue') and, eventually, by a modifier (*veľmi tmavočervená* 'very dark red').

Color terms that indicated hue by compounding had a much lower representation (15.7%). In particular, there were color terms for which there are Slovak color names expressing the meaning of 'N-like' referring to a specific entity: *sivozelená/olivová* 'grayish green/olive-like', *oranžovoružová/lososová* 'orangeish rose/salmon-like', *červenofialová/cyklámenová* 'reddish purple/cyclamen like', *modročervená/orgovánová* 'bluish red/lilac-like', *červenohnedá/škoricová* 'reddish brown/

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

cinnamon-like', *hnedožltá/horčicová* 'brownish yellow/mustard-like', *modrozelená/tyrkysová* 'bluish green/turquoise-like'.

Color terms expressing 'cross-sensory experience' appeared in the replies of 12 respondents and represented only 8.1% (49 responses) of all the terms formed via compounding. This 'cross-sensory experience' was expressed by adverbs, indicating, in particular, power (*silnofialová* 'strong purple', *slabomodrá* 'weak blue', etc.), and acoustic experience (*krikľavozelená* 'garishly (lit. screaming) green'). Other cross-sensory experiences were absent in the replies. This meaning type was expressed peripherally in the determinant syntagma (*sýta červená* 'deep red'), eventually intensified by another proverb (*veľmi slabožltá* 'very weak yellow'). Sporadically (1.3%), we noticed the meaning type 'ancientness' in the compound word *staroružová* 'old rose'.

In summary, the dominant meaning categories formed via the derivation ('N-like') or compounding (brightness) have cultural and cognitive backgrounds. Brightness comes from the oldest partition of the color spectrum into dark and light colors with added shades of color. The meaning of 'N-like' is based on cultural prototypes as preferred psychological impressions of a color, which are an intersubjective matter in the language and in its individual elements. The prototype is the most typical representative and most ideal example of a color that has all the essential features. For example: the ideal representative of the rose color in our cultural context is the wild rose (Rosa canina), which occurs in nature, not crossbred roses; the most typical chocolate color is not white chocolate, but chocolate with cocoa; for skin color, caucasian is the typical skin color, not, e.g., African and others.

3.3 The semantic pattern 'N-like' in color naming

As mentioned above, color terms with the lexical meaning of 'N-like' were highly frequent in the responses of all participants, which suggests that it is a significantly widespread meaning category in the awareness of Slovak native speakers. The results of previous linguistic studies (Conklin, 1973; Wierzbicka, 2005; Malt & Majid, 2013) show that languages regularly use terms for familiar and well-known entities to name colors. This metonymic shift is primarily realized in Slovak via the lexicalization pattern N entity + suffix *-ový/-ová/-ové*.

The name forming the root of a color term can represent entities of the most diverse kind. The examples[8] obtained from the EoSS research (listed before parenthesis) as well as from fashion catalogs (listed in parenthesis) showed the following types of entities:

8. The examples are given in the basic form.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. flowers: *fialový* 'violet-like', *ružový* 'rose-like', *orgovánový/lilavý* 'lilac-like', (*nezábudkový* 'forget-me-not-like', *cyklámenový* 'cyclamen-like', *lipový* 'linden flower-like', *mätový* 'mint-like', *lotosový* 'lotus-like', *vanilkový* 'vanilla flower-like', *šafránový* 'saffron-like', *iskerníkový* 'kingcup-like', *karafiátový* 'maiden pink-like', *levanduľový* 'lavender-like');
2. fruits: *olivový* 'olive-like', *broskyňový* 'peach-like', *marhuľový* 'apricot-like', *limetkový* 'lime-like', (*citrónový* 'lemon-like', *malinový* 'raspberry-like', *jahodový* 'strawberry-like', *čerešňový* 'cherry-like', *jablkový* 'apple-like', *mandarínkový* 'mandarin-like', *černicový* 'blackberry-like', *slivkový* 'plum-like', *morušový* 'cloudberry-like', *banánový* 'banana-like');
3. rocks, metals and gemstones: *tyrkysový* 'turquoise-like', *okrový* 'ochre-like', *antracitový* 'anthracite-like', (*strieborný* 'silver-like', *zlatý* 'gold-like', *smaragdový* 'emerald-like', *akvamarínový* 'aquamarine-like', *jantárový* 'amber-like', *ametystový* 'amethyst-like', *hrdzavý* 'rust-like', *uhľový* 'carbon-like');
4. drinks: *čokoládový* 'chocolate-like', *kávový* 'coffee-like', *bordový* 'wine-colour-like', (*koňakový* 'cognac-like', *šampanský* 'champagne-like');
5. environment: *azúrový* 'azure-like', *oceánový* 'ocean-like', *nebeský* 'sky-like', (*slnečný* 'sun-like', *dymový* 'fume-like', *bahnový* 'mud-like');
6. building materials: *tehlový* 'brick-like', *pieskový*' gravel-like', *kamienkový* 'stone-like', (*kamenný* 'stone-like');
7. gastronomy: *maslový* 'butter-like', *horčicový* 'mustard-like', *škoricový* 'cinnamon-like';
8. human body: *telový* 'body-like', *pleťový* 'skin-like';
9. professions: *námornícky* 'sailor-like', *vojenský* 'soldier-like', (*kráľovský* 'king-like', *kardinálsky* 'cardinal-like', *poľovnícky* 'hunter-like');
10. other fruits: *gaštanový* 'chestnut-like', (*bobuľový* 'berry-like', *pistáciový* 'pistachio-like');
11. 9. vegetables: *hráškový* 'pea-like', (*kukuričný* 'corn-like', *mrkvový* 'carrot-like', *tekvicový* 'cucumber-like', *baklažánový* 'aubergine-like');
12. animals: *lososový* 'salmon-like', (*ťavý* 'camel-like', *korálový* 'coral-like', *slonovinový* 'ivory-like', *perleťový* 'mother-of-pearl-like');
13. colorants: *purpurový* 'magenta-like/purple', (*karmínový* 'carmine-like');
14. trees: *jedľový* 'fir-like';
15. herbs: *trávový* 'grass-like';
16. names of artists: *titianový* 'titian-like';
17. cosmetics: (*púdrový* 'talcum-like', *krémový* 'ointment-like');
18. fuels: (*petrolejový* 'kerosene-like');
19. clothing: (*džínsový* 'jeans-like').

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The most common entities used in Slovak to create color terms are those whose color is not difficult for a typical Slovak native speaker to distinguish, that is, different flowers and fruits; but there is a relatively strong representation (as well as a relatively simple color distinction) of rocks, metals, and gems. Some color names with the lexical meaning of 'N-like' are in a determinative syntagma (as a structurally and linearly organized combination of two words, in which one dependent (subordinate) member determines the second independent (superordinate) member) or a compound word often associated with a basic (meaning superordinate) color term. These are primarily entities whose color attributes are not well-known or are vague, e.g., color naming referring to professions (*kardinálska červená* 'cardinal-like red', *kráľovská – námornícka modrá* 'king-like blue – sailor-like blue', *poľovnícka – vojenská zelená* 'hunter-like green– soldier-like green', etc.), and the environment (*azúrová modrá – oceánová modrá – dymová modrá* 'azure-like – ocean-like – fume-like blue').

Substantivization is regularly present in color terms of entities whose suffix is formed by vowel combinations (*fuksia, aqua*). However, substantivization is also irregularly applied to other color terms that have a metonymic relationship to entities with easily identifiable single color attributes (*smaragd, jahoda, citron* 'emerald, strawberry, lemon', etc.). A specific group of color terms with the meaning 'N-like' includes entities used in phrases: *ružové drevo*, *jahňacia nappa*, *morský rak*. These color terms were only encountered in fashion catalogs where a very detailed perceptual resolution is expected of the reader.

These data show that metonymy in color naming with the meaning 'N-like' is greatly influenced by the knowledge of native speakers about the world around them, their preferences, and that with which they are in immediate visual contact. The cultural context and visual color stimuli from the natural environment play a major role in the process of color naming. As the most colorful, modern Slovaks perceive the world of flowers and fruits which, in the earlier developmental phases, were a source of entities in the partition of the color spectrum (these are the color terms *fialový* 'violet-like/purple', *ružový* 'rose-like/pink', *oranžový* 'orange-like/orange' that Berlin and Kay placed into the 7th stage of development, although the semantic category 'N-like' began to form much earlier, specifically when naming the color red based on a dye obtained from a special kind of insect).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4. Conclusion

Based on the results of the empirical research Evolution of Semantic Systems (EoSS), we explored lexicalization patterns that are involved in Slovak color naming. The Slovak language has word-formation patterns, derivation, and compounding, while compounding offers speakers the possibility to more precisely differentiate the color spectrum and is, therefore, preferable to derivation in communication. However, the difference in frequency between these two basic lexicalization patterns was not so significant in Slovak: 40,8 % of color terms were formed by derivation and 58.3% by compounding.

More significant differences emerged in the productivity of semantic categories, where the formation involves a derivative or compounding in conjunction with syntactic patterns (especially with a prepositional phrase and determinative syntagma). While the derivation is primarily in the formation of three meaning types with a marked representation 'N-like', N entity + suffix and low representation of approximative meaning, compounding forms four semantic types, out of which brightness is the central meaning type; hue and cross-sensory experience are peripheral meaning types. The postcentral semantic categories are ancientness (formed by compounding) and diminutive meaning (formed by derivation), which are collocation- and style-bound.

The lexicalization patterns preferred in contemporary Slovak color naming reveal inherent laws that have their developmental and cultural justification. The meaning category brightness expressed by the adverbial components *tmavo-/temno-* 'dark', *svetlo-* 'light', *bledo-* 'pale' is associated with the distribution of the color spectrum composed of two basic parts (white color with assigned light colors and black color with dark colors), which in Slovak (and also in Old Slavonic) **čъrnъa *bělъ (*bhēlo-)* have been documented since the 12th century. According to Berlin and Kay (1969), this is the start point in the evolution of culture and its language with its own way of conceptualizing the color world.

In the next developmental stages, the semantic category 'N-like' begins to form. Its dominance in contemporary communication is linked to the basic cognitive assumptions of human beings who perceive the properties of entities (e.g., colors) naturally, via their senses, as integral parts of a whole, well-known object. This metonymic relationship is realized toward entities of the most diverse kind (in particular, flowers and fruits in Slovak), which are ideal representatives (prototypes) of a color, which is largely associated with cultural specific (intersubjective) motives. Research of lexicalization patterns has shown that the system of color naming in the present-day language is not accidental, but has evolved from a previous state and is largely culturally determined.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

References

Berlin, B., & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley, CA: University of California Press.

Conklin, H. C. (1973). Color categorization. *American Anthropologist*, *75*(4), 931–942. https://doi.org/10.1525/aa.1973.75.4.02a00010

Démuth, A. (2005). *Čo je to farba?* Bratislava: Iris.

Dorn, A., Willalva, A., Giouli, V., Wiebke, B., Kovalenko, K., & Wandl-Vogt, E. (2016). Displaying Language Diversity on the European Dictionary Portal – COST ENeL – Case Study on Colours and Emotions and their Cultural References. In T. Margalitatze, & G. Meladze (Eds.), *Proceedings of the 17th EURALEX International Congress* (pp. 237–243). Tbilisi: Ivane Javakhishvili Tbilisi University Press.

Hergenhahn, B. R. (2009). *An introduction to the history of psychology* (6th ed.). Australia: Wadsworth Cengage Learning.

Jameson, K., & D'Anadrade, R. (1997). It´s not really red, green, yellow, blue: an inquiry into perceptual color space. In C. L. Hardin, & L. Maffi (Eds.), *Color Categories in Thought and Language* (pp. 295–319). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519819.014

Králik, Ľ. (2015). *Stručný etymologický slovník*. Bratislava: VEDA.

Ladd-Franklin, Ch. (1929). *Colour and colour theories*. New York: Harcourt.

Lucy, J. A. (1997). The linguistics of "color". In C. L. Hardin, & L. Maffi (Eds.), *Color Categories in Thought and Language* (pp. 320–346). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519819.015

Machek, V. (1968). *Etymologický slovník jazyka českého*. Praha: Academia.

Malt, B., & Majid, A. (2013). How thought is mapped into words. *WIREs Cognitive Science*, *4*(6), 583–597. https://doi.org/10.1002/wcs.1251

Majid, A., & Levinson, S. C. (2007). The language of vision I: colour. In A. Majid (Ed.), *Field Manual Volume 10* (pp. 22–25). Nijmegen: Max Planck Institute for Psycholinguistics.

Raffaelli, I., & Kerovec, B. (2017). The concept of 'taste' in formation of Croatian and Turkish lexicon: A contrastive analysis. *Suvremena lingvistika*, *83*(43), 21–48. https://doi.org/10.22210/suvlin.2017.083.02

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*(3), 369–398. https://doi.org/10.1037/0096-3445.129.3.369

Regier, T., Káy, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(4), 1436–1441. https://doi.org/10.1073/pnas.0610341104

Talmy, L. (1985). Lexicalization patterns: semantic structures in lexical forms. In: T. Shopen (Eds.), *Language Typology and Syntactic Description III: Grammatical Categories and the Lexicon* (pp. 57–149). Cambridge: Cambridge University Press.

Tokarski, R. (1995). *Sematyka barw we współczesnej polszczyźnie*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.

Vaňková, I. (2003). Kolory w czeskim językowym obrazie świata: barwa „žlutá". In R. Grzegorczykowa, & K. Waszakowa (Eds.), *Studia z semantyki porównawczej. Nazwy barw. Nazwy wymiarów. Predykaty mentalne. Cz. II*. (pp. 69–99). Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Vaňková, I., Nebeská, I., Saicová-Římalová, L., & Šlédrová, J. (2005). *Co na srdci, to na jazyku*. Praha: Univerzita Karlova v Praze. Nakladatelství Karolinum.

Waszakowa, K. (2001). Struktura znaczeniowa podstawowych nazw barw. Założenia opisu porównawczego. In R. Grzegorczykowa, & K. Waszakowa (Eds.), *Studia z semantyki porównawczej. Nazwy barw. Nazwy wymiarów. Predykaty mentalne. Cz. I*. (pp. 59–72). Warszawa: Wydawnictwo Uniwersystetu Warszawskiego.

Wierzbicka, A. (2005). There are no "color universals" but there are universals of visual semantics. *Anthropological Linguistics*, *47*(2), 217–244.

Wittgenstein, L. (2002). *Modrá a hnedá kniha*. Bratislava: Kaligram.

Internet sources

Slovenský národný korpus – prim-7.0-public-all. Online: http://korpus.juls.savba.sk

Alfatex – slovenský výrobca oblečenia. Online: http://blog.alfatex.sk,cit.30.5.2017

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Compound color terms in Italian

Maria Grossmann and Paolo D'Achille
University of L'Aquila / University of Roma Tre

The present paper is dedicated to compound color terms in Italian. The aim is to give a descriptive overview of the characteristics of the individual compounding patterns both from the synchronic and the diachronic point of view. According to our preliminary research (cf. Grossmann, 1988; D'Achille & Grossmann, 2013), in contemporary Italian compounding is the most commonly used device for enlarging the inventory of color terms. Data from corpora show that we are dealing with constructions at the crossroads between word-formation and syntax, i.e., between compounding proper and asyndetic coordination. From a diachronic point of view, we have ascertained that compound color terms appear with a certain frequency only from the 18th century onwards; in earlier times they are attested only by a few examples and in a limited number of patterns.

Keywords: color terms, Italian, word-formation patterns, compounding

1. Introduction and brief survey of the literature

As we have pointed out in a previous paper (cf. Grossmann & D'Achille, 2016, pp. 21–23, on which this introductory section is mainly based), the "space" of color names, whose referent is the color space, is made up of the semantic structure of a set of lexemes. The linguistic encoding of the perceptual and physical substance of colors implies a generalization process: each name refers to a group of shades regardless of any differences among them. Different languages segment the color space in different ways: the boundaries of individual categories may fall in different places; there may be differences in: (a) the number of available terms, (b) the basis for making distinctions, (c) the weight attributed to the three psycho-sensorial variables of color, i.e., hue (what a layman refers to when talking about color by classifying his/her perception as *green*, *red*, etc.), brightness (variation in intensity, passing from lighter to darker shades), and saturation (percentage of pure hue in a color; saturation is higher when the color appears to be less blended with white). The assessment of differences between languages should also take account of other

https://doi.org/10.1075/sfsl.78.04gro
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

factors. In some languages, color terms can also include information about other factors besides the three aforementioned variables (succulence vs. desiccation, features of surface texture, features of shape and consistency, etc.). Often, as well as identifying a given shade, color names can be used as symbols for social, religious and moral concepts of varying complexity. We also find intralinguistic differences based on variables such as age, gender, education and social status of the speaker.

The various aspects of color-naming (linguistic, perceptive, cognitive, cultural) have been a privileged area of research for linguists, psychologists and anthropologists since the 19th century (for a survey of the literature see Grossmann & D'Achille, 2016, pp. 22–28). The most recent studies on color terms can be framed in the more general debate regarding linguistic relativism vs. universalism (cf. Grossmann, 1988, pp. 8–27; Kay & Regier, 2006).

The seminal work for the color-naming debate is the 1969 book by Brent Berlin and Paul Kay which aimed to demonstrate both the existence of universals in color lexicons and the presence of an evolutionary connection between these universals and the historical development of languages. According to these two scholars, there exists a universal inventory of eleven perceptual categories that act as psychophysical referents for basic color terms (two to eleven) in each language. These terms have a number of characteristic features: from the morphological point of view they are not complex; from the semantic point of view they are not transparent, they are not hyponyms of other terms and their application is not restricted to specific classes of entities; from the psychological point of view, they are salient to speakers. Non-basic color terms, on the other hand, are morphologically analyzable, semantically transparent and generally more recent than basic terms. When a language encodes fewer than eleven categories, according to Berlin and Kay, restrictions apply to the categories that are actually encoded. As for the historical evolution of the color lexicon, it is assumed that the encoding of the eleven universal perceptual categories takes place in a partially fixed chronological order of seven evolutionary stages.

Berlin and Kay's book (1969) was a revolution in color-naming studies. It attracted vast consensus, but also widespread disagreement, and literally boosted research aimed at validating or invalidating the theory of the two scholars. There have been many partial revisions and successive elaborations of Berlin and Kay's theses, mainly concerned with problems of the number of basic terms and how to interpret the evolutionary sequence.

The question of whether the number of basic terms could be increased past the eleven mark has constituted a major challenge for color vocabulary studies in individual languages. The disputed status of the Russian terms *sinij* 'blue/dark blue' and *goluboj* 'light blue' was the occasion for a vast number of specific studies on the BLUE area also in Italian and in Italo-Romance dialects. Grossmann and

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

D'Achille (2016, 2017) focus on historical evolution and contemporary use of the nearest Italian equivalents of *sinij* and *goluboj*, i.e., *azzurro*, *blu*, *celeste* and *turchino*, and review the previous literature. The authors claim that in contemporary Italian *azzurro* still denotes a shade between *celeste* (light) and *blu* (dark), but gradually hands over the basic term role to *blu* and comes closer to *celeste* inasmuch as it indicates only light shades; *turchino* is nowadays in disuse. The same authors carried out research also on basic terms in the BROWN area, another area which has been radically restructured over time. D'Achille and Grossmann (2017) focus on changes in semantic relations between *bruno*, *marrone*, *castano,* and *moro* from Old to present-day Italian, and observe that in the course of the 20th century *marrone* has replaced *bruno* as the basic term in this area.

Patterns for forming derived and compound color terms in Italian have been studied by Grossmann and Mazzoni (1972, 1976), Grossmann (1988), Elwert (1989), Koura (1992), Timmermann (2002), D'Achille and Grossmann (2013), D'Achille (2014) and Rainer (2017). Masini (2017; 2019a; 2019b) analyzed the presence and use of basic color terms within multiword expressions and focused on the demarcation and competition issues between compounds and phrasal lexemes.

2. Color terms formed by means of morphological devices

The basic color terms in Italian are *bianco* 'white', *nero* 'black', *rosso* 'red', *verde* 'green', *giallo* 'yellow', *azzurro*/*blu* 'blue', *marrone* 'brown', *grigio* 'grey', *viola* 'purple'.[1] The majority of the other color terms are formed by means of various morphological and syntactical devices. Productive word-formation patterns used in color-naming comprise derivation, compounding and conversion. In this study, we shall focus on compounding, the most commonly used device for enlarging the inventory of color terms in contemporary Italian. Our analysis will be carried out both from the synchronic and the diachronic point of view.

First, we shall briefly illustrate the other word-formation patterns relevant to the analysis of some features of the constituents of compound color terms, i.e., derivation by means of suffixes and conversion.

1. The status of *rosa* 'pink' and *arancione* 'orange' needs more research; on the problem of the basicness of *azzurro* and/or *blu* 'blue', see above, Section 1.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.1 Derivation

2.1.1 *Deadjectival adjectives*

Basic color terms, and also some non-basic ones, can be bases for the derivation of other color adjectives (Grossmann & Mazzoni, 1972; Kristol, 1978; Grossmann, 1988, pp. 63–73, 175–181, 200–209; Merlini Barbaresi, 2004, pp. 444–450). They share a set of suffixes expressing approximative/attenuative/intensifying and/or evaluative meaning. As for evaluative meaning, the positive or negative value judgment of the speaker may regard, depending on the context, either the quality denoted by the adjectival base or the referent of the head noun. The majority of deadjectival color adjectives of this kind belong to the first of the following two types:

1. Derivatives denoting approximation – in terms of hue, brightness and saturation – to the focal point denoted by the base (e.g., *bianco* 'white' → *bianchiccio* 'whitish', *giallo* 'yellow' → *giallognolo* 'yellowish', *rosso* 'red' → *rossastro* 'reddish', *verde* 'green' → *verdino* 'greenish').
2. Derivatives with the elative suffix *-issimo* denoting the highest degree of brightness and saturation compared to an implicit standard (cf. Rainer, 1983 and Section 2.3.1.1). Featured in the corpora or on the internet are, among others: *azzurro* 'blue' → *azzurrissimo* 'very blue', *verde* 'green' → *verdissimo* 'very green'. In combination with color terms, augmentative suffixes with an intensifying function, such as *-one*, have low productivity (e.g., *giallo* 'yellow' → *giallone* 'intense yellow', *verde* 'green' → *verdone* 'intense green'). It should be noted that Italian can express intensification of a given quality also by means of the prefix *stra-*, which is, however, seldom known to modify color terms (Google-searches yield rare instances of derivatives such as *strarosso* 'very red', *straverde* 'very green').

Many deadjectival derivatives are characterized by their availability for nominal conversion (e.g., *bionda* 'blonde' → *biondona* (F).SG 'hot blonde'). Some of them developed non-compositional meanings (e.g., *rosso* 'red' → *rossetto* 'lipstick').

2.1.2 *Denominal adjectives*

Denominal derivatives specify the shade by means of a comparison with the color attributed *par excellence* to the base-noun *designatum*. They are formed with suffixes that have the function of forming resemblance adjectives (e.g., *ambra* 'amber' → *ambrato* 'amber (A)', *cera* 'wax' → *cereo* 'pale, waxen', *corvo* 'raven' → *corvino* 'raven (A)', *oliva* 'olive' → *olivastro* 'olive (A)', *rame* 'copper' → *ramato* 'auburn, copper (A)').

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.1.3 *Deverbal adjectives*

Verbs derived from color terms (cf. Timmermann, 2002) can also be bases for derivation of adjectives denoting a shade resulting from an agentive or non-agentive event (e.g., *annerito* 'blackened' ← *annerire* 'to blacken' ← *nero* 'black', *ingiallito* 'yellowed' ← *ingiallire* 'to yellow' ← *giallo* 'yellow'), or a state or condition (e.g., *verdeggiante* 'verdant, somewhat green' ← *verdeggiare* 'to be verdant/ green/greenish' ← *verde* 'green').

2.2 Conversion

Another productive device for creating color terms is recourse to names of entities for denoting a color by metonymy (cf. Grossmann, 1988, pp. 182–199; an analysis of the morphological characteristics of this type of lexeme is given in Thornton, 2004, pp. 529–530). Among the nouns used as color terms we find bases of denominal adjectives (see Section 2.1.2) and nominal constituents of A-N compounds (see Section 2.3.1.2) that can be used independently to denote a particular shade (such as *antracite* in *abito antracite* 'anthracite dress', *carta da zucchero* in *parete carta da zucchero* 'sugar paper [blue][2] wall'), and also other nouns generally used on their own for color-naming (such as *acquamarina* in *occhi acquamarina* 'aquamarine eyes', *seppia* in *effetto seppia* 'sepia effect', *turchese* in *lago turchese* 'turquoise lake'). This kind of term is probably generated by ellipsis of multiword expressions (such as (*di*) *color*(*e*) + (*di*) + N 'lit. (of) color (of) N'; color adjective + *come* 'as/like' + (ART) + N), or of A-N compounds. The more frequent the use of this kind of color-naming pattern, the greater the probability of a noun being used independently as a color term.

2.3 Compounding

A commonly used device for enlarging the inventory of color terms is that of forming compound words made up of two adjectives or an adjective and a noun. These constructions, frequently nonce-formations in both literary and journalistic texts, confront scholars with the same problems as all other adjectival compounds in Italian, i.e., spelling,[3] allomorphy, inflectional characteristics, relationship between the constituents, etc. (cf. Tollemache, 1945, pp. 60–73, 234–240; Rohlfs,

2. Until the 1950s, sugar and other groceries were frequently sold in dark blue paper wrappers.

3. The orthography of Italian compounds varies depending on the degree of lexicalization and the inflectional properties of the construction (cf. Grossmann & Rainer, 2009). Constituents can be written as one word or with a hyphen, but we can also find a blank in lieu of the hyphen.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

1966–1969, § 998; Dardano, 1978, pp. 193–194; Tekavčić, [2]1980, §§ 1134–1140; more recently, Bisetto, 2004, pp. 47–49; D'Achille & Grossmann, 2009, 2010; Dardano, 2009, pp. 236–237; Forza, Guevara & Scalise, 2009; Gaeta & Ricca, 2009; Grossmann & Rainer, 2009; Masini & Scalise, 2012; Micheli, 2016; and, with particular reference to color compounds, Grossmann, 1988; D'Achille & Grossmann, 2013; D'Achille, 2014; Grossmann & D'Achille, 2016; Rainer, 2017; Masini, 2017; 2019a; 2019b). In compound color terms these problems are even more complex. The main questions concern whether these constructions are of morphological or syntactic nature, whether they and their constituents are nouns or adjectives, whether their origin is popular or learned, ancient or modern.

As we shall see in greater detail, compound color terms come into widespread use in Italian only after the 17th–18th centuries. Before that, they were found in technical terminologies (painting, dyeing, natural history); subsequently, and especially during the 19th and 20th centuries, in literary language. Finally, they spread into journalism and, at least to some extent, the language of everyday use. In more remote periods, adjectival compounding is attested only by a few examples and in a limited number of patterns. It is worth recalling that also in Classical Latin compound adjectives occur only rarely: referring specifically to color compounds, André (1949, pp. 229–232) notes that those that do occur are mainly borrowings from Greek or late nonce-formations (e.g., *albogiluus* 'whitish yellow', *ignicolor* 'fire color', *lacticolor* 'milk color'). On the other hand, Neo-Latin, namely the scientific Latin used in Europe from the 16th century onwards, makes widespread use of constructions of this type, constructions which have no classical precedents and which are characterized from a morphological point of view by the presence of the linking vowel *-o* at the end of the first adjective (the earliest examples of color compounds are *atrovirens* [1620] 'black-green', *nigro-purpureum* [1641] 'black-purple', *cinereo-fuscus* [1657] 'grey-dark', *luteo-viridis* [1659] 'yellow-green', *nigro-luteum* [1676] 'black-yellow'). It is, therefore, Neo-Latin that provides the model for the formation of compound adjectives in Italian (cf. Hatcher, 1951; Grossmann & Rainer, 2009; Rainer, 2017).

The availability of large and searchable corpora now makes it possible to study color compounds in greater depth, from both a diachronic and a synchronic perspective. The contemporary Italian data we will be discussing are taken from the *la Repubblica* corpus (newspaper texts dating from 1985 to 2000, roughly 380M tokens) and from searches run on Google (between July and September 2017). To investigate the diachronic evolution of compounding as a color-naming strategy, we searched the following corpora (listed in chronological order): OVI (Old Italian texts dating before 1375, roughly 23M tokens), BIZ (literary texts from early Italian to the first decades of the 20th century), MIDIA (various types of texts dating from

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

the 13th century to 1947, roughly 7.5M tokens), DiaCORIS (different kinds of prose texts dating from 1861 to 2001, roughly 25M tokens), PTLLIN (literary prose texts dating from 1947 to 2006). Further examples have been extracted from historical dictionaries such as TB and GDLI, from Savoca's (1995) concordances of 20th century poetry, and from searches run on Google Books.[4]

Let us now observe the characteristics of the individual compounding patterns.

2.3.1 *Relationship between the constituents*

The classification of compounds from the point of view of the relationship between constituents is the subject of a wide range of literature using different morphological frameworks (see Bisetto & Scalise, 2005, for an overview). Here we shall follow the classification proposed in Bisetto and Scalise (2005), Masini and Scalise (2012), and operate a distinction between coordinate and attributive adjectival color compounds.

Coordinate compounds are made up of two or more constituents, semantically at the same taxonomic level, which modify the head noun both independently and simultaneously. Constructions with an attributive relationship between constituents display a head-modifier structure and indicate a given shade of the color denoted by the left-hand constituent.[5] It must be stressed that color compounds usually have a compositional meaning and that cases of lexicalized construction are very rare.

2.3.1.1 *Coordinate compounds*

Among coordinate compounds we can further identify three subtypes of which the first two spread in usage from the 18th century onwards (although we also have isolated examples of the second subtype in earlier times), while the third is already well-attested in Old Italian.

1. Constructions with an additive reading that qualify two- or multi-colored entities such as flags, apparel (e.g., *sciarpa bianco-nera* 'white-black scarf', *insegna rosso-nero-gialla* 'red-black-yellow sign', *bandiera rosso-verde-bianco-nera* 'red-green-white-black flag'), or that denote the colors of a team jersey (e.g., *maglia gialloverde* 'yellow-green jersey'). Compounds of this kind are endocentric and have two or more semantic heads.

---

4. We give no references to sources for contemporary examples; as already indicated, they are taken from the *la Repubblica* corpus and from searches run on Google. For the examples taken from diachronic corpora, we give the name of the author or the title of the text, and date by year or century. The English glosses of the examples are literal translations.

5. In Italian the canonical position of the head is on the left.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

This pattern spread in usage from the 18th century onwards (e.g., *vessil nero-eburno* 'black-white ensign', Baruffaldi, 1758; *insegne giallonere* 'yellow-black signs', E. Praga, 1875); 20th century literary texts yield numerous nonce-formations, especially constructions with a binary structure, but also formations showing ternary structure, or even playful constructions formed by concatenating several constituents (e.g., [*sciantosa*] *giallolillarosablù* 'lit. yellow-lillac-pink-blue [chanteuse]', Pirandello, 1932; *tuniche* […] *biancorossogialloviolarosazzurre* 'lit. white-red-yellow-purple-pink-blue tunics', Palazzeschi, 1968).

In Old Italian an additive relationship between two or more color terms was indicated by a syntactic construction with the conjunction *e* 'and' between the two adjectives (e.g., *lombardese verde e vermigl*[*i*]*o e nero* 'green and vermilion and black fabric', *Documenti pratesi*, 1247; *pan*[*n*]*o giallo e nero* 'yellow and black cloth', *Documenti fiorentini*, 1274–1310; *bandiera* […] *mezza bianca e mezza vermiglia* 'flag half white and half red', Boccaccio, 1336–1338).

In compounds of this type, we also find color terms with metonymical meanings which, in this case, have a classifying and not a qualifying function: the so-called "club colors" (cf. D'Achille, 2014) also denote the players, fans, managers, executives (e.g., *manager bianconero* '[lit. white-black manager] Juventus manager', *società rossonera* '[lit. red-black club] Milan club'); "political colors" are used symbolically to represent a political party, its members, constituency, etc. (e.g., *governi rosso-verdi* '[lit. red-green governments] governments formed of Communists and Greens') and in certain contexts allow a relational reading (*alleanze rosso verdi* '[lit. red green alliances] alliances between Communists and Greens'); some color terms denominate narrative genres (e.g., *romanzo giallo-rosa* '[lit. yellow-pink novel] detective-romantic novel').

2. Constructions with an intersective reading that indicate a hue intermediate between the colors denoted by the constituents (such as *occhi azzurro-grigi* 'blue-grey eyes', *capelli biondo-castani* 'blonde-brown hair', *divisa grigioverde* 'grey-green uniform', *pelame rosso-bruno* 'red-brown fur'). Compounds of this kind are exocentric and identify a color which does not correspond to any of the individual constituents. Note, however, that in many cases we need to know the context and the head noun's referent in order to decide whether a construction is to be read as additive or intersective (e.g., *tunica nerazzurra* 'a tunic of a color between blue and black' vs. *maglia nerazzurra* 'a black and blue striped jersey').

This kind of compound, which is modeled on Neo-Latin formations and thus of learned origin (cf. Hatcher, 1951; Grossmann & Rainer, 2009; Rainer, 2017), occurs with a certain frequency only from the 18th century onwards, when it is found especially in writings on natural history (e.g., *liquor giallo-verde* 'yellow-green liquor', Landriani, 1783). According to Rainer (2017), the rare examples of intersective

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

compounds found in previous centuries are probably to be interpreted as the result of an endogenous process of univerbation. The earliest example occurs in Boccaccio's *Corbaccio* (1354–1355): *viso verdegiallo* 'green-yellow face'.[6] The Boccaccio example, which is picked up by Ariosto (*scudo nero* […] *fregiato a color verdegiallo* 'black shield decorated in a green-yellow color', 1532), was evidently the model for other such compounds, all of which have *verde* 'green' as their first element (e.g., *verdebruno* 'green-brown', Franco Sacchetti, 14th cent.; *verde azzurro* 'green blue', *Ricettario fiorentino*, 1567; *verdenegro* 'green-black', Bartolomeo Crescenzio, c. 1600). Folena (1951), on the other hand, argues that the *verdegiallo* type is modeled on attributive compounds, such as *verde chiaro* 'light green', which are of popular origin and already attested in Old Italian (see Section 2.3.1.2).[7] In support of Folena's view, we might hypothesize that compounds with *bruno* as their second constituent had a 'bridging' role between the two patterns in the sense that in Old Italian it meant both 'brown' and 'dark' (cf. D'Achille & Grossmann, 2017).

Apart from these rare examples, Old Italian prefers syntactic strategies to compounding for expressing intersective meaning. Some of these, such as *(in)tra* 'between' + color term$_1$ + *e* 'and' + color term$_2$, are still in use (e.g., *colore mischiato intra verde e giallo* 'mixed color between green and yellow', Zucchero Bencivenni, 1310; *faccia* […] *tra bianca e gialla* 'face between white and yellow', Dante, 1321; *pel tra bigio e nero* 'skin between gray and black', Ariosto, 1532).

3. Reduplicative constructions, more syntactic in nature, of the type *occhi azzurri azzurri* '[lit. blue blue eyes] very blue eyes', *ragazza bionda-bionda* '[lit. blonde blonde girl] very blonde girl', are used with an intensifying function to describe the presence of a color at the highest degree of brightness and saturation (as is the case with the derivatives in *-issimo* mentioned in Section 2.1.1). Elatives of color terms formed with the suffix *-issimo* are also found as reduplicants (e.g., *è davvero rosso rossissimo* '[lit. it is really red very red] it is really very red'*)*.

Reduplicative constructions are attested right from Old Italian (e.g., [*acqua*] *bruna bruna* '[lit. brown brown water] very brown water', Dante, 1321; *pappagalli* […] *verdi verdi* '[lit. green green parrots] very green parrots', Sassetti, 1588; *una testuggine nera nerissima* '[lit. black very black tortoise] very black tortoise', Baruffaldi, 1758).

Reduplication of adjectives has an intensifying function also in contexts in which this type of construction has a figurative meaning (*uomo grigio grigio* '[lit.

---

6. The compound appears as *verde*, *giallo* in the edition used by the OVI, but as a single word in the 3rd edition of the *Vocabolario della Crusca*, 1691.

7. On the origins of *verdegrigio* 'green-grey' and *grigioverde* 'grey-green', see the recent study by Vitale-Brovarone (2014).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

grey grey man] very boring, mediocre man', *segnali neri neri* '[lit. black black signs] very negative signals').

The order of constituents in coordinate compounds is, in principle, free (e.g., *occhi azzurro-verdi* 'blue-green eyes' and *occhi verde-azzurri* 'green-blue eyes'). However, there appears to be, on the one hand, a pragmatically motivated preference for using the adjective designating the more important property as the first constituent, and on the other, a prosodically motivated preference for placing the shorter adjective in initial position, according to Behagel's law (cf. Grossmann & Rainer, 2009). Often the choice seems to be mainly usage-based. In the case of colors of flags, the order is fixed and can become a distinctive feature; for example, *bandiera rosso-bianco-verde* 'red-white-green flag' indicates the Hungarian flag, while *bandiera bianco-rosso-verde* 'white-red-green flag' indicates the Italian one.

2.3.1.2 *Attributive compounds*

The majority of compounds found in the corpora are left-headed attributive constructions describing a particular shade of the color denoted by the head. The first constituent is the hyperonym of the whole construction while the second constituent acts as a modifier; modifiers can be adjectives or nouns. The following subsections offer an overview of the patterns classified according to the types of modifiers. The first subtype is already attested in Old Italian, while the second and third are documented only from the 18th century onwards (with some isolated examples from earlier times).

1. Modifiers can be adjectives referring to the degree of brightness and saturation of the color denoted by the head constituent (e.g., *tuta grigio-scura* 'dark-grey overalls', *occhi azzurro-chiari* 'light-blue eyes', *cieli blu intensi* 'intense blue skies', *abito rosso acceso* 'bright red dress'). These constructions are very frequent in the corpora and the array of adjectival modifiers is vast.[8] The head constituent can also be a color term of nominal origin (such as *montature avana scuro* 'dark Havana frames', *occhi nocciola scuro* 'dark hazel eyes'; see Section 2.2).

Constructions of this type are already present in Old Italian (e.g., *panni* […] *due verdi kiari e due verdi bruni* 'cloths … two light green and two dark green', *Documenti fiorentini*, 1278–1279; *cielo azzurro chiaro* 'light blue sky', Restoro d'Arezzo, 1282; *colore bigio chiaro* 'light grey color', Niccolò da Poggibonsi, 1345; *colore giallo smorto* 'dull yellow color', Pegolotti, 14th cent.).

---

8. In addition to the adjectives cited, other constructions include *brillante* 'brilliant', *carico* 'deep', *cupo* 'dark', *fosco* 'sombre', *luminoso* 'bright', *opaco* 'opaque', *pallido* 'pale', *pastello* 'pastel', *profondo* 'deep', *sbiadito* 'faded', *sgargiante* 'gaudy', *slavato* 'faded', *smorto* 'dull', *spento* 'dull', *squillante* 'shrill', *tenue* 'soft', *vivo* 'vivid', etc.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. Modifiers can be derivatives of other color adjectives (see Section 2.1.1) that indicate a hue to which the color denoted by the head comes close (e.g., *chiazza rosso-bluastra* 'bluish-red stain', *buccia marrone giallognola* 'yellowish brown peel', *occhi azzurro grigiastri* 'greyish blue eyes', *pelo biondo-rossiccio* 'reddish-blond fur').

This pattern is attested only quite late, and almost exclusively in literary texts (e.g., *color grigio azzurrognolo* 'bluish grey color', Leopardi, 1813; *cavaliere* […] *biondo-rossigno* 'reddish-blonde horseman', Faldella, 1877; *pàgine giallorossastre* 'reddish-yellow pages', Dossi, 1881; *linee nero-giallastre* 'yellowish-black lines', Fogazzaro, 1895). One could, however, also attribute to this type one isolated 14th century example made up of a denominal adjective and a deadjectival one: *cianbellotto rosato biadecto* 'bluish rosy fabric', Ranieri Sardo, 1354–1399.

3. Modifiers can be denominal adjectives (see Section 2.1.2) that specify the shade by comparing it with the quintessential color attributed to the referents of their nominal bases (e.g., *occhi castano ambrati* 'amber brown eyes', *capelli nero-corvini* 'raven-black hair', *colorito bianco-latteo* 'milky-white complexion', *cartoncino giallo paglierino* 'straw yellow card', *chioma biondo-ramata* 'copper-blonde head of hair').

This pattern is attested continuously only from the 18th century (e.g., *striscia di color bianco latteo* 'streak of milky white color', *Gazzetta toscana*, 1780; *dita bianco-rosate* 'rosy-white fingers', Monti, 1801; *frondi bigio-argentee* 'grey-silver fronds', Fogazzaro, 1881). Early examples, such as *giallo cetrino* 'citron yellow', Pegolotti, 14th cent.; *fronte bianca lattata* 'milky white brow', Mazzeo Bellebuoni, 1333, are extremely rare. In fact, the preferred pattern in Old Italian is the syntactic construction with the conjunction *e* 'and' between the two adjectives (e.g., *bianca e lattata* […] *dea* 'white and milky goddess', *Detto d'amore*, 13th cent.; *palafreni lattati e bianchi* 'milky and white palfreys', Antonio Pucci, 1388).

4. The modifier can be a noun specifying the shade by comparing it with the quintessential color attributed to its referent (e.g., *capelli biondo-paglia* 'straw-blonde hair', *pantaloni blu notte* 'night blue pants') or, metonymically, with the color of an object intimately associated with, or typical of the referent (such as that of *aviazione* 'air force' uniforms in *giacca blu aviazione* 'air-force blue jacket'). From a semantic point of view, these nouns have an adjectival function, but from a morphological point of view they do not share the features typical of adjectives (cf. Thornton, 2004, pp. 529–530; for some types of "noun-clad adjectives" in Italian see Grandi, Nissim & Tamburini, 2011).

A-N compounding is the most productive pattern in today's Italian. A vast array of nouns with an adjectival function can act as modifiers of color terms (cf. Kristol, 1978; Grossmann, 1988, pp. 73, 182–192). There are nouns denoting:

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

a. metals/minerals (e.g., *acciaio* 'steel', *cobalto* 'cobalt', *ferro* 'iron', *lavagna* 'slate');
b. stones/jewels (e.g., *corallo* 'coral', *perla* 'pearl', *smeraldo* 'emerald', *zaffiro* 'sapphire');
c. coloring agents (e.g., *indaco* 'indigo', *lacca* 'lacquer', *ocra* 'ochre', *porpora* 'purple');
d. animals (e.g., *aragosta* 'lobster', *canarino* 'canary', *salmone* 'salmon', *topo* 'mouse');
e. associated with the plant kingdom (e.g., *bosco* 'wood', *paglia* 'straw', *sottobosco* 'undergrowth', *tabacco* 'tobacco');
f. flowers (e.g., *amaranto* 'amaranth', *ciclamino* 'cyclamen', *fucsia* 'fuchsia', *lavanda* 'lavender');
g. fruits/vegetables (e.g., *carota* 'carrot', *limone* 'lemon', *mela* 'apple', *pisello* 'pea');
h. foods and beverages (e.g., *panna* 'cream', *senape* 'mustard', *vino* 'wine', *zafferano* 'saffron');
i. others (e.g., *biliardo* 'billiards', *bottiglia* 'bottle', *cenere* 'ash', *cielo* 'sky', *fumo* 'smoke', *fuoco* 'fire', *notte* 'night', *petrolio* 'petroleum').

In our corpora we also find some examples of semantically opaque compounds made up of color terms and nouns with an intensifying function which denote abstract entities, such as *giallo tradimento* 'betrayal yellow', *rosso passione* 'passion red', *verde speranza* 'hope green'.

While the A-N compounds discussed so far denote a shade typical of the noun's referent, others indicate colors of objects intimately associated with, or typical of the referent. The nominal constituent is generally a common noun (e.g., *blu tranviere* 'tram-driver blue', *rosso-cardinale* 'cardinal-red'); however, there are also a few instances of personal names (e.g., *azzurro Tiepolo* 'Tiepolo blue', *blu Savoia* 'Savoy blue', *castano-isabella* 'Isabella-brown'), names of institutions (e.g., *azzurro Europa* 'Europe blue', *blu Nazioni Unite* 'United Nations blue'), and place names (e.g., *rosso magenta* 'Magenta red', *verde Padania* 'Padania green'). Similar in nature from the semantic point of view are compounds made up of a color term and a denominal adjective such as *blu ministeriale* 'lit. ministerial blue', *blu berlusconiano* 'lit. Berlusconian blue', *rosso-pompeiano* 'lit. Pompeian-red'.

In the vast majority of A-N compounds the head constituent is a simplex color term, but Google searches also yield occasional examples of constructions made up of a derived adjective denoting approximation (see Section 2.1.1) and a noun (e.g., *bluastro notte* 'lit. night bluish', *rossiccio fuoco* 'lit. fire reddish', *verdino mela* 'lit. apple greenish').

The earliest example of this pattern goes back to the 16th century (*color verde-porro* 'leek green color', *Ricettario fiorentino*, 1550), but this is an isolated case. A-N color compounds begin to be attested with greater continuity in the 18th and early 19th centuries, and come into widespread use probably as a result of the influence of French. Some terms are certainly unadapted loans or calques from previously

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

attested French compounds: thus, the French *gris de fer* 'iron grey' (1690) is found in the 18th century both as a calque (*grigioferro*, Fagiuoli, 1734) and in semi-adapted forms (*gridefer*, Fagiuoli, 1733; *grisdeferro*, Bergantini, 1745); the French *vert bouteille* 'bottle green' (which, thanks to Google Books is datable to 1800–1805) is the model for *verde bottiglia*, and appears in the 19th century fashion magazine, *Il Corriere delle Dame*, in the semi-adapted form *verde bouteille* (1824) (cf. Sergio, 2010). The frequency of A-N compounds increases appreciably in the second half of the 19th century (e.g., *criniera biondo-lino* 'linen blonde mane', D'Azeglio, 1866; *penne color verdemare o verdepisello* 'sea green or pea green color quills', P. Savi, 1871; *guanti grigio-perla* 'pearl grey gloves', Oriani, 1878); significantly, several terms are recorded in TB (e.g., *bianco-neve* 'snow-white', *bianco-perla* 'pearl-white', *verde bronzo* 'bronze green'); annotations to these terms in TB (*T*[*ermine*] *de' tintori* 'dyers' term') frequently describe them as belonging to jargons related to the world of fashion. The proliferation of A-N compound color terms in the 20th century is mainly attributable to the fashion industry and to a number of crafts and businesses that kept introducing new hues into the market and giving them coined names that were subsequently disseminated by the mass media. An interesting case in point is represented by the names coined for car body colors (cf. Caffarelli, 2014). The apparently arbitrary selection of nominal modifiers can often be accounted for by the cultural salience of their referents in the Italian linguistic community, and in some cases is influenced by foreign models.

To denote a shade by comparing it with a color typical of a noun's referent Old Italian resorted to multiword expressions which remained in use and are still current, such as:

a. (*di*) *color*(*e*) *di* N 'lit. (of) color of N' (e.g., *nebula di colore di fuoco* 'cloud the color of fire', Dante, 1293; *sopravesta di color di sabbia* 'over-garment the color of sand', Ariosto, 1532; *piume* […] *di color di noce* 'feathers the color of walnut', Tassoni, 1624; *acqua* […] *colore di foglia morta* 'water the color of dead leaves', Spallanzani, 1765);
b. color adjective + *come* 'as/like' + (Art) + N (e.g., *acqua* […] *verde come erba* 'water green as grass', *Il Milione*, 14th cent.; *garzoni* […] *neri come carbone* 'servant boys black as coal', Frescobaldi, 14th cent.; *occhi* […] *rossi come fuoco* 'eyes red as fire', Agazzari, 1422; *camice bianco come neve* 'shirt white as snow', Straparola, 1532; *peli* […] *biondi come lo oro* 'hair fair as gold', Aretino, 1534).

As for the competition between color compounds and multiword expressions with *come* 'as/like' in contemporary Italian, Masini (2017, 2019a, 2019b) has carried out corpus-based research and shown that even though the two types of construction denote the same shade, multiword expressions have an additional intensification meaning and a higher degree of expressiveness.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.3.2 *Morphosyntactic characteristics*

Inflectional properties of color compounds, as of all other adjectival compounds in Italian, are a complex issue (cf. D'Achille & Grossmann, 2009, 2010, 2013; Grossmann & Rainer, 2009). We will here illustrate the main patterns, discussing separately A-A and A-N compounds.

A-A compounds waver between various patterns: a pattern with inflectional markers added to the second adjective, an asyndetic pattern with double inflection, and a pattern with lack of agreement.

If the first constituent belongs to the first class of adjectives, which have a distinct inflectional suffix for each feature combination,[9] it can occur with an internal readjustment consisting of the addition of the linking element *-o*, an inheritance from the Neo-Latin pattern. Constructions with a linking *-o* have a single morphological locus for agreement at the end of the sequence and are characterized by a high degree of integration. Some examples of coordinate and attributive A-A compounds displaying this inflectional pattern are provided in (1a) for contemporary data and in (1b) for data from diachronic corpora:[10]

(1) a. *sciarpa bianco-nera* 'white-black scarf', *tuta grigio-scura* 'dark grey overalls', *capelli castano biondi* 'brown blonde hair', *pietra bianco-giallastra* 'yellowish-white stone', *parrucca biondo-argentea* 'silver-blonde wig'

b. *poltiglia giallopallida* 'pale-yellow pulp', Vallisneri, 1730; *rosso-fosca nuvola* 'somber-red cloud', Cesarotti, 1808; *divise biancorosse* 'white-red uniforms', Rovani, 1869; *orecchini* […] *bianco-verdastri* 'greenish-white earrings' Serao, 1891; *lanetta azzurro-cupa* 'dark-blue light-weight wool', Oriani, 1899

A-A sequences with the first adjectives of the *-o* class that allow more patterns usually show agreement on both adjectives (2a) or on neither of them (2b):

---

9. Italian adjectives agree with their head noun in gender (masculine/feminine) and number (singular/plural). There are five adjective classes. The largest class is that of adjectives with a distinct inflectional suffix for each feature combination: $-o_{\text{M.SG}}$, $-a_{\text{F.SG}}$, $-i_{\text{M.PL}}$, $-e_{\text{F.PL}}$ (e.g. *bianco* 'white'). The second largest class shows no gender distinction: for both genders, the singular is realized as *-e*, the plural as *-i* (e.g. *verde* 'green'). The remaining set of adjectives is divided between: a group which has a threefold declension, with *-a* for singular, both masculine and feminine, *-i* for M.PL and *-e* for F.PL (e.g. *belga* 'Belgian'); a group which has *-e* for M.SG and F.PL, *-a* for F.SG and *-i* for M.PL (e.g. *sornione* 'sly'); and a group of indeclinable adjectives (e.g. *blu* 'blue').

10. Phrases with masculine singular controllers have been discarded, since in their case the form of the first constituent in *-o* coincides with the M.SG form of the adjective and the *-o* can in principle be interpreted either as a linking vowel or as an inflectional ending.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(2) a. *bandiera rossa bianca* 'red white flag', *occhi azzurri-grigi* 'blue-grey eyes', *luna rossa accesa* 'bright red moon'
tinta rosea dorata 'golden rosy shade', C. Boito, 1883; *nerebianche viaggiatrici* [*rondini*] 'black-white travelers [swallows]', D'Annunzio, 1900; *monti grigi rossigni* 'reddish grey mountains', Pirandello, 1928

b. *magliette giallo vivo* 'vivid yellow t-shirts', *cravatta rosso cupo* 'dark red tie', *busta giallo-paglierino* 'straw-yellow envelope'
faille grigio bianco 'grey white faille', *Corriere delle Dame*, 1875; *cotonina rosso-vivo* 'vivid-red calico', Serao, 1891

The frequency with which these inflectional patterns occur was assessed through a survey of use on the internet. The results obtained by means of a Google search run on September 15, 2017, for compounds qualifying *capelli* 'hair' are set out in Table 1. These data are not to be taken as completely reliable as they are raw and need manual revision, but they are still useful as a rough guide to overall tendencies. Table 1 also includes numbers for examples in which only the first constituents exhibit agreement markers, examples absent from the corpora and of dubious acceptability. The results show some preference for the pattern without agreement.

Table 1. Agreement patterns of A-A compounds with first constituent of *-o* class

Head noun: *capelli* (M).PL 'hair'	Agreement on first constituent	Agreement on second constituent	Agreement on both constituents	Lack of agreement
biondo chiaro 'light blonde'	387	1040	1130	13100
castano chiaro 'light brown'	1480	43000	16500	47700
biondo scuro 'dark blonde'	7140	1180	10900	115000
castano scuro 'dark brown'	5660	39800	6100	36500
biondo castano 'blonde-brown'	494	2370	9630	2320
biondo ramato 'copper blonde'	3	788	2040	17800

A-A sequences with first adjectives of the *-e* class[11] do not allow insertion of the linking vowel *-o*. They follow one of three possible options: (a) agreement at the end of the sequence, (b) on both adjectives, (c) on neither of them. As exemplified in (3a–c), we find all three types of inflectional behavior in the corpora:

11. In fact, this class is represented only by compounds with *verde* 'green'. A-A sequences with *arancione* 'orange', *celeste* 'light blue' and *marrone* 'brown' as their first constituents are very rare.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(3) a. *monti verde-bruni* 'green-brown mountains', *seta verde chiara* 'light green silk', *toni verde-chiari* 'light-green hues'
veste verdebruna 'green-brown garment', Franco Sacchetti, 14th cent.; *foglioni verdecupi* 'dark-green leaves', Manzoni, 1840; *muschi verdegialli* 'green-yellow mosses', D'Annunzio, 1893

b. *occhi verdi azzurri* 'green blue eyes', *ombre verdi-turchine* 'green-blue shadows', *piastrelle verdi-azzurre* 'green-blue tiles'
bande verdigialle 'green-yellow banks', Gozzano, 1911; *colli verdineri* 'green-black hills', Slataper, 1912

c. *tute verde-chiaro* 'light-green overalls', *siepi verde scuro* 'dark green hedges', *valigia verde chiaro* 'light green suitcase'
cappellini […] *verde-americano* '[lit. American green caps] intense green caps', *Corriere delle Dame*, 1825; *verdechiaro berlina* 'light-green saloon', Dossi, 1881; *foglie verde chiaro* 'light green leaves', Berto, 1947

According to Google frequencies (Table 2), in contemporary Italian lack of agreement is the preferred option. Table 2 also includes numbers of occurrences of constructions in which only the first constituent agrees, these too being of dubious acceptability.

Table 2. Agreement patterns of A-A compounds with first constituent of *-e* class

Head noun	Compound	Agreement on first constituent	Agreement on second constituent	Agreement on both constituents	Lack of agreement
foglie (F).PL 'leaves'	*verde scuro* 'dark green'	1880	236	4000	61000
pantaloni (M).PL 'pants'	*verde scuro* 'dark green'	141	364	1460	79400
occhi (M).PL 'eyes'	*verde scuro* 'dark green'	2110	5960	572	17700
foglie (F).PL 'leaves'	*verde-grigiastro* 'greyish green'	0	72	306	319
occhi (M).PL 'eyes'	*verde-azzurro* 'green-blue'	288	16400	10100	1350
occhi (M).PL 'eyes'	*verde-grigio* 'green-grey'	629	324	546	2880

If the first constituent belongs to the class of invariable adjectives like *blu* 'blue', the second constituent may agree with the head noun (4a) or not (4b):

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(4) a. *cristalli blu-verdastri* 'greenish-blue crystals', *completi blu presidenziali* 'presidential blue suits'
stoffa blu scura 'dark blue material', C. Levi, 1951; *capelli blu-neri* 'blue-black hair', C. Levi, 1957

b. *abiti blu-chiaro* 'light-blue dresses', *completi blu-scuro* 'dark-blue suits', *occhi blu-verde* 'blue-green eyes'
calzoni blu chiaro 'light blue pants', Arbasino, 1960

Examples in Table 3, followed by their Google frequencies, show a certain preference for usage without agreement.

Table 3. Agreement patterns of A-A compounds with invariable first constituent

Head noun	Compound	Agreement on second constituent	Lack of agreement
camicia (F).SG 'shirt'	*blu scuro* 'dark blue'	4840	33400
pantaloni (M).PL 'pants'	*blu scuro* 'dark blue'	5470	1370000
cravatta (F).SG 'tie'	*blu scuro* 'dark blue'	610	1060000
occhi (M).PL 'eyes'	*blu intenso* 'intense blue'	1980	13700
colorazione (F).SG 'coloration'	*blu-violaceo* 'violet blue'	1230	242

It should be noted that in reduplicative constructions the pattern with a double inflection is the only possibility (e.g., *occhi azzurri azzurri* 'lit. blue blue eyes', *ragazza bionda-bionda* 'lit. blonde-blonde girl'), and that it is documented *ab antiquo* ([*acqua*] *bruna bruna* '[lit. brown brown water] very brown water', Dante, 1321).

In the case of compounds with more than two adjectival constituents, the number of morphological loci increases accordingly. We find several examples of sequences with a linking vowel *-o* at the end of the first and second adjective (such as *insegna rosso-nero-gialla* 'red-black-yellow sign'), with agreement on all constituents (such as *bandiera nera-bianca-rossa* 'black-white-red flag'), and with lack of agreement (such as *bandiere bianco-rosso-verde* 'white-red-green flags').

A-N constructions are invariable (5a). There are also rare instances of agreement on the head adjective (5b) that are definitely non-standard in modern Italian, if not at the limits of acceptability:

(5) a. *divise giallo senape* 'mustard yellow uniforms', *pomodori rosso sangue* 'blood red tomatoes', *piume verde smeraldo* 'emerald green feathers'
criniera biondo-lino 'linen-blonde mane', D'Azeglio, 1866; *camionetta grigio-piombo* 'lead-grey truck', Calvino, 1950

b. *seta rossa papavero* 'poppy red silk', *occhi verdi nocciola* 'hazel green eyes'
materia rossa fuoco 'fire red matter', *Diz. scienze naturali*, 1848; *colori verdi pisello* 'pea green colors', *Emporium*,1900

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Google data (Table 4) show a preference for invariability.

Table 4. Agreement patterns of A-N compounds

Head noun	Compound	Agreement on first constituent	Lack of agreement
occhi (M).PL 'eyes'	*azzurro cielo* 'sky blue'	1470	14000
occhi (M).PL 'eyes'	*azzurro-ghiaccio* 'ice blue'	1290	10200
capelli (M).PL 'hair'	*biondo-cenere* 'ash blonde'	8260	17300
capelli (M).PL 'hair'	*biondo-platino* 'platinum blonde'	2720	46000
occhi (M).PL 'eyes'	*verde smeraldo* 'emerald green'	1970	8930
pantaloni (M).PL 'pants'	*verde militare* 'military green'	4040	13000

The prevalence of invariability in A-N compounds and their productivity in today's Italian may well have determined the analogous tendency of A-A compounds to avoid agreement.

Finally, the analysis of our data also shows that, besides resorting to the addition of the linking element *-o*, there are occasional examples of other strategies for obtaining A-A and A-N sequences with greater formal integration and external inflection:

a. elision of the last unstressed vowel of the first adjective in cases where the second begins with a vowel (e.g., A-A: *nerazzurro* [← *nero* + *azzurro*] 'black-blue', *verdelettrico* [← *verde* + *elettrico*] 'electric-green'; A-N: *giallovo* [← *giallo* + [*u*] *ovo*] 'egg-yellow', Raimondi, 1957; *verdacqua* [← *verde* + *acqua*] 'water-green');
b. apocope of the last unstressed vowel of the first adjective when preceded by /n/ (e.g., *capelli castan scuri* 'dark brown hair'; *color turchin chiaro* 'light blue color', Bartoli, 1661);
c. haplology (e.g., *verdorato* [← *verde* + *dorato*] 'golden-green', Marinetti, 1944).

The picture that emerges from the analysis of inflectional properties of color compounds is a heterogeneous one. They lie along a continuum from constructions with external inflection showing morphological compoundhood to asyndetic constructions with a double agreement which are more syntactic in nature.

Finally, let us look briefly at how A-A and A-N sequences behave with respect to two other criteria often used to distinguish compounds from phrases. For a construction to exhibit compoundhood it should: (a) denote a unitary concept; (b) represent a syntactic atom.[12] The constructions analyzed respond positively

12. Stress assignment is not considered relevant in Italian since compounds behave differently as regards stress properties depending on the degree of lexicalization and frequency of usage.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

to criterion (a). Semantically they behave like adjectives: they describe and name a property of the head noun's referent. They are characterized by large numbers of nonce-formations with a low degree of lexicalization, as is demonstrated by the limited number of established formations in the dictionaries. As for criterion (b), most of the constructions examined do not allow for insertion of lexical material. Insertion of a modifier is only allowed in some attributive constructions whose non-head constituent refers to the degree of brightness and saturation (i.e., an adjective such as *chiaro* 'light', *pallido* 'pale', *scuro* 'dark'; see Section 2.3.1.2). Constructions such as *occhi azzurri, molto chiari* 'lit. blue eyes, very light' show double agreement and should be bracketed as follows: [[*occhi azzurri*] *molto chiari*]). Internal modification appears to be unacceptable in constructions belonging to the other patterns discussed in Section 2.3.1.2: for example, the insertion of *molto* 'very' between a color term and a derivative of another color term (**pelo biondo molto rossiccio*), between a color term and a denominal adjective (**chioma bionda molto ramata*), between a color term and a noun (**gonna verde molto bottiglia*).

3. Conclusion

The color lexicon of Italian is a system in continuous movement, with terms disappearing and new ones being adopted. In this regard, it is typical of technologically advanced societies in which colors can be produced artificially and constantly in a huge number of shades and on any material. Our research shows that the recourse to compounding as a means of codifying different values of hue, brightness and saturation has been especially productive in recent times. The analysis of the morphosyntactic characteristics of these constructions shows that they lie at the crossroads between word-formation and syntax, i.e., between compounding proper and asyndetic coordination.

From a diachronic point of view, the data collected allow us to state that Old Italian used mainly syntactic strategies. There are, however, some ancient examples of coordinate and attributive sequences (both univerbated and not) which might be considered morphological objects. It seems certain that, as is the case with other adjectival compounds, the Neo-Latin compounding model was decisive in the spread of these constructions. Color compounds dating from before the 17th and 18th centuries are rare. They are documented initially in technical terminologies (painting, dyeing, natural history); then, especially in the course of the 19th century and at the beginning of the 20th in literary texts, and subsequently in journalism; finally, at least to some extent, they begin to appear in everyday language.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

References

André, J. (1949). *Étude sur les termes de couleur dans la langue latine*. Paris: Klincksieck.

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

Bisetto, A. (2004). Composizione con elementi italiani. In M. Grossmann, & F. Rainer (Eds.), *La formazione delle parole in italiano* (pp. 31–51). Tübingen: Niemeyer.

Bisetto, A., & Scalise, S. (2005). The classification of compounds. *Lingue e Linguaggio*, 4(2), 319–332.

BIZ = Stoppelli, P. (Ed.) (2010). *Biblioteca italiana Zanichelli* (DVD-ROM). Bologna: Zanichelli.

Caffarelli, E. (2014). Per un approccio al labirinto onimico dei modelli automobilistici italiani e stranieri. *Rivista Italiana di Onomastica*, 20, 481–534.

D'Achille, P. (2014). Per una storia delle parole del calcio: i nomi dei giocatori, i composti 'bicolori' e il caso di *blucerchiato*. *Lingua nostra*, 75, 112–126.

D'Achille, P., & Grossmann, M. (2009). Stabilità e instabilità dei composti aggettivo + aggettivo in italiano. In E. Lombardi Vallauri, & L. Mereu (Eds.), *Spazi linguistici: Studi in onore di Raffaele Simone* (pp. 143–171). Roma: Bulzoni.

D'Achille, P., & Grossmann, M. (2010). I composti aggettivo + aggettivo in italiano. In M. Iliescu, H. M. Siller-Runggaldier, & P. Danler (Eds.), *Actes du XXV^e^Congrès International de Linguistique et de Philologie Romanes (Innsbruck, 3–8 septembre 2007)* (VII, pp. 405–414). Berlin: de Gruyter.

D'Achille, P., & Grossmann, M. (2013). I composti ‹colorati› in italiano tra passato e presente. In E. Casanova Herrero, & C. Calvo Rigual (Eds.), *Actas del XXVI Congreso Internacional de Lingüística y de Filología Románicas (Valencia, 6–11 septiembre 2010)* (III, pp. 523–537). Berlin: de Gruyter.

D'Achille, P., & Grossmann, M. (2017). I termini di colore nell'area BRUNO-MARRONE in italiano: sincronia e diacronia. *Lingua e stile*, 42, 87–115.

Dardano, M. (1978). *La formazione delle parole nell'italiano di oggi*. Roma: Bulzoni.

Dardano, M. (2009). *Costruire le parole. La morfologia derivativa dell'italiano*. Bologna: il Mulino.

DiaCORIS = DiaCORIS. *Corpus diacronico dell'italiano scritto*. http://corpora.dslo.unibo.it/DiaCORIS/.

Elwert, W. T. (1989). Appunti sui termini di colore in italiano e spagnuolo. In R. Antonelli, F. Beggiato, A. Ferrari, & A. Solimena (Eds.), *Miscellanea di studi in onore di Aurelio Roncaglia a cinquant'anni dalla sua laurea* (II, pp. 537–549). Modena: Mucchi.

Folena, G. (1951). Chiaroscuro leonardesco. *Lingua Nostra*, 12, 57–63.

Forza, F., Guevara, E., & Scalise, S. (2009). Compounding adjectives. *Lingue e Linguaggio*, 8(1), 137–162.

Gaeta, L., & Ricca, D. (2009). *Composita solvantur*: Compounds as lexical units or morphological objects?. *Italian Journal of Linguistics / Rivista di Linguistica*, 21(1), 35–70.

GDLI = Battaglia, S. (Ed.) (1961–2002). *Grande dizionario della lingua italiana*. Torino: Utet.

Grandi, N., Nissim, M., & Tamburini, F. (2011). Noun-clad adjectives: On the adjectival status of non-head constituents of Italian attributive compounds. *Lingue e Linguaggio*, 10(1), 161–176.

Grossmann, M. (1988). *Colori e lessico. Studi sulla struttura semantica degli aggettivi di colore in catalano, castigliano, italiano, romeno, latino ed ungherese*. Tübingen: Gunter Narr.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Grossmann, M., & D'Achille, P. (2016). Italian colour terms in the BLUE area: Synchrony and diachrony. In J. P. Silvestre, E. Cardeira, & A. Villalva (Eds.), *Colour and colour naming: crosslinguistic approaches* (pp. 21–50). Lisboa: Centro de Linguística da Universidade de Lisboa – Universidade de Aveiro.

Grossmann, M., & D'Achille, P. (2017). I termini di colore nell'area AZZURRO-BLU in italiano: sincronia e diacronia. *AION-Linguistica*, 6, 109-143.

Grossmann, M., & Mazzoni, B. (1972). Motivato /v/ immotivato: tipi di denominazione dei termini di colore in italiano. In *Beiträge zur Romanistik und Allgemeinen Sprachwissenschaft: Festschrift Wilhelm Giese* (pp. 65–69). Hamburg: H. Buske.

Grossmann, M., & Mazzoni, B. (1976). Analisi semantico-strutturale dei termini di colore in italiano. In *Actes du XIII^e Congrès International de Linguistique et Philologie Romanes* (I, pp. 869–890). Québec: Les Presses de l'Université Laval.

Grossmann, M., & Rainer, F. (2009). Italian adjective-adjective compounds: Between morphology and syntax. *Italian Journal of Linguistics / Rivista di Linguistica*, 21(1), 71–96.

Hatcher, A. G. (1951). *Modern English word-formation and Neo-Latin: A study of the origins of English (French, Italian, German) copulative compounds*. Baltimore: Hopkins.

Kay, P., & Regier, T. (2006). Language, thought and color: Recent developments. *Trends in Cognitive Sciences*, 10(2), 51–54. https://doi.org/10.1016/j.tics.2005.12.007

Koura, T. (1992). Studio contrastivo dei vocaboli cromatici tra il giapponese e l'italiano con speciale riferimento ai due tipi 'colore + nome' e 'colore + aggettivo cromatico + nome'. *Studi Italiani di Linguistica Teorica ed Applicata*, 21(1–3), 199–218.

Kristol, A. M. (1978). *Color: Les langues romanes devant le phénomène de la couleur*. Berne: Francke.

la Repubblica = "la Repubblica" Corpus. http://sslmitdev-online.sslmit.unibo.it/corpora/corpus.php?path=/name=Repubblica.

Masini, F. (2017). Polirematiche 'di colore' in italiano: uno studio quantitativo. In R. D'Alessandro, G. Iannàccaro, D. Passino, & A. M. Thornton (Eds.), *Di tutti i colori: Studi linguistici per Maria Grossmann* (pp. 203–215). Utrecht: Utrecht University. <https://www.ris.uu.nl/ws/files/34667598/Di_tutti_i_colori.pdf>

Masini, F. (2019a). Compounds and multiword expressions in Italian. Demarcation and competition issues. In B. Schlücker (Ed.), *Complex lexical units: compounds and phrases* (pp. 153-188). Berlin: de Gruyter.

Masini, F. (2019b). Competition between morphological words and multiword expressions. In F. Rainer, F. Gardani, W. U. Dressler, & H. C. Luschützky (Eds.), *Competition in inflection and word-formation* (pp. 283-305). Cham: Springer.

Masini, F., & Scalise, S. (2012). Italian compounds. *Probus*, 24, 61–91.

Merlini Barbaresi, L. (2004). Aggettivi deaggettivali. In M. Grossmann, & F. Rainer (Eds.), *La formazione delle parole in italiano* (pp. 444–450). Tübingen: Niemeyer.

Micheli, M. S. (2016). Limiti e potenzialità dell'uso di dati empirici in lessicografia: il caso del plurale delle parole composte. *Ricognizioni*. *Rivista di Lingue, Letterature e Culture Moderne*, 3(6), 15–33.

MIDIA = MIDIA. *Morfologia dell'Italiano in DIAcronia*. www.corpusmidia.unito.it/.

OVI = *Corpus OVI dell'Italiano antico*. *Opera del Vocabolario Italiano*. http://gattoweb.ovi.cnr.it/.

PTLLIN = De Mauro T., (Ed.) (2007). *Primo Tesoro della Lingua Letteraria Italiana del Novecento* (DVD-ROM). Torino: Utet / Fondazione Maria e Goffredo Bellonci.

Rainer, F. (1983). *Intensivierung im Italienischen*. Salzburg: Institut für Romanistik der Universität Salzburg.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rainer, F. (2017). On the origin of Italian adjectival colour compounds of the type *grigioverde* 'grey-green'. In R. D'Alessandro, G. Iannàccaro, D. Passino, & A. M. Thornton (Eds.), *Di tutti i colori: Studi linguistici per Maria Grossmann* (pp. 247–255). Utrecht, Utrecht University. <https://www.ris.uu.nl/ws/files/34667598/Di_tutti_i_colori.pdf>

Rohlfs, G. (1966–1969). *Grammatica storica della lingua italiana e dei suoi dialetti*. Torino: Einaudi.

Savoca, G. (1995). *Vocabolario della poesia italiana del Novecento*. Bologna: Zanichelli.

Sergio, G. (2010). *Parole di moda. Il "Corriere delle Dame" e il lessico della moda nell'Ottocento*. Milano: FrancoAngeli.

TB = Tommaseo, N., & Bellini, B. (1861–1879). *Dizionario della lingua italiana*. Torino: Unione Tipografico-Editrice.

Tekavčić, P. (^{2}1980). *Grammatica storica dell'italiano*.Bologna: il Mulino.

Thornton, A. M. (2004). Conversione in aggettivi. In M. Grossmann, & F. Rainer (Eds.), *La formazione delle parole in italiano* (pp. 526–533). Tübingen: Niemeyer.

Timmermann, J. (2002). La verbalisation des adjectifs de couleur en français, espagnol et italien. *Vox Romanica*, 61, 1–31.

Tollemache, F. (1945). *Le parole composte nella lingua italiana*. Roma: Edizioni Rores di Nicola Ruffolo.

Vitale-Brovarone, A. (2014). Due etimi facili*: grigioverde* e *cane*. *Bollettino dell'Atlante Linguistico Italiano*, 38, 141–149.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

"Brightness" in color linguistics

New light from Danish visual semantics

Carsten Levisen
Roskilde University

This chapter scrutinizes the discourse of "brightness" in color linguistics. Drawing on insights from visual semantics and linguistic anthropology, and challenging the universal applicability of "brightness", the study provides new evidence from Danish. The chapter provides a new analysis of the lexicogrammar and linguaculture of *lys* 'light, brightness' in relation to color. The Natural Semantic Metalanguage approach is used to provide detailed semantic explications for three grammatical devices based on *lys* (*lys*, *lys-*, and *lyse-*), along with an analysis of three Danish *lys* + color compounds *lyserød*, *lysegrøn*, and *lyseblå*. Based on the evidence from Danish and other studies in visual semantics, the chapter calls for a renewed focus on the non-chromatic aspects of visual meanings, and for a metalinguistic reform in color linguistics.

Keywords: color linguistics, brightness, visual semantics, Natural Semantic Metalanguage, Danish

1. Introduction

In the post-Berlin and Kay era of color studies, researchers have increasingly paid attention to non-chromatic aspects of visual semantics, such as "brightness" (Biggam, 2007), "glossiness" (Uusküla & Eessalu, 2018), "visual conspicuousness" (Goddard & Wierzbicka, 2014), "iridescence and illuminance" (Hill, 2011). This shift towards what could be called the "color + " paradigm has challenged the traditional paradigm of color linguistics and, in particular, its notions of "basic color terms" and "color proper".[1] Wierzbicka and colleagues have demonstrated that the "color" concept in itself is a modern, European conceptual construct, rather than a global human concern, and that many of the world's languages have rich visual

1. Key texts in the universalist "basic color term" tradition include Berlin and Kay (1969), Kay (2006), Kay et al. (2009).

https://doi.org/10.1075/sfsl.78.05lev
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

vocabularies without "color proper", and, in fact, without "color" at all (Wierzbicka, 2005, 2006, 2008; Goddard & Wierzbicka, 2014). Based on cross-linguistic evidence, Wierzbicka concluded that "SEEing, not color, must be the starting point, and the cornerstone, of our investigations" (Wierzbicka, 2006, p. 20). Historical semantics has seen similar developments. Casson (1997) documented a shift from "brightness to hue" in the visual vocabulary of English. The rise of color and the fall of brightness in the history of English speaks of "color" as (a) non-universal and (b) as a modern concept. Furthermore, postcolonial-semantic studies in color and visuality have shown how European color-based visual semantic systems have been imported and imposed on many speakers, and that many languages of the world have been semantically and visually "Europeanized" in recent times as a consequence of colonialism and globalization (Levisen, Sippola, & Aragón, 2016). Thus, in the context of modern European visual semantics, "color" remains a force to be reckoned with. At the same time, it is clear that European visual systems have much more to offer than the so-called "color proper", for instance, "brightness"-related visual meanings, and meanings in which "hue" and "brightness" are conceptually combined.

To progress the European "color + " paradigm, we need an "emic" turn that takes local lexicalizations seriously both at the level of semantic analysis and at the level of theory formation. This means that "brightness" should be scrutinized in the same way as "color". And "brightness" turns out to be a very slippery term. As discussed in detail by the color linguist Carole Biggam, there is no consensus in expert discourse about what "brightness" means, let alone what the terms stand for in a semantic analysis (Biggam, 2007). Unlike Biggam, who attempted to "save" the discourse of brightness by providing new definitions, this chapter seeks to deconstruct the discourse of "brightness" through a cross-linguistic reality check. As a case study, I have selected the Danish co-construction of *lys* 'light' and *farve* 'color', with special reference to three compounds *lyserød*, *lysegrøn*, and *lyseblå* – literally 'light red', 'light green' and 'light blue'. This Danish compound construction has been flagged by previous studies as unsystematic and inconsistent: *lyserød* does not relate light to *rød* 'red' in the same way as *lyseblå*, and *lysegrøn* relates light to *blå* 'blue' and *grøn* 'green' (Levisen, 2012b; Vejdemo et al., 2015; Vejdemo, 2017). This (un)systematicity of the Danish compound morphology of *lys* 'light' needs to be studied in greater detail, and the result has potentially far-reaching consequences for "brightness studies", for European color linguistics as well as for global visual semantics.

This chapter is structured in the following way: (a) Section 2 summarizes key findings in visual semantics and discusses these insights in the light of the case study on "brightness"; (b) Section 3 accounts for the methods, a combination of database/corpus analysis and the NSM interpretive methodology (for a similar two-pronged approach, see Levisen, 2015); (c) Section 4 introduces the linguaculture and

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

lexicogrammar of Danish *lys* 'light' and provides semantic explications for *lyserød*, *lysegrøn*, and *lyseblå*; (d) Section 5 brings back the discussion of "brightness" in color linguistics; (e) the final Section offers some concluding remarks.

2. Visual semantics

"Visual semantics" is a new paradigm launched by Anna Wierzbicka and colleagues (Wierzbicka, 2005, 2006, 2008, 2018; Goddard & Wierzbicka, 2014; Levisen, Sippola & Aragón, 2016; Aragón, 2016, 2017). Goddard & Wierzbicka say:

> It was widely assumed for a long time that all languages have "color words", and this is a myth that dies hard. Even so, it is clear that the domain of "visual semantics" exhibits many commonalities and recurring features, such as the contrast between light and the absence of light, or the prominence of bodily and environmental prototypes such as blood, vegetation, sky, and earth.
>
> (Goddard and Wierzbicka, 2014, p. 79)

Although the universal color myth was exposed decades ago by color-critical philosophers and anthropologists (such as e.g., Saunders, 1992; Saunders & van Brakel, 1997; and Lucy, 1997), only a few alternatives for doing empirical, linguistically grounded studies have emerged. Visual semantics presents itself as such an alternative. Contributing to the reframing of color studies into a more globally oriented discipline, visual semantics has made particular advances in descriptive, interpretative, and metalinguistic practices. The key insights of visual semantics can be summarized as follows:

a. "Color" is not a universal human concept or concern. It is, in essence, a modern European visual idea. Most languages in the history of the world have had no visual category corresponding to "color", "red", "yellow" "blue", etc. and the question "what color is this?" has not been expressible in many languages. By contrast, other questions such as "what does it look like?" appear to be translatable across languages. The European "color" concept (English *color*, Spanish *color*, German *Farben*, Danish *farve*, etc.) functions as a collective superordinate and is a constituting part of the meaning of "color terms", such as English *yellow*, *blue*, *purple*, Spanish *amarillo, azul,* and *morado*, German *gelb, blau, lila*, or Danish *gul*, *blå*, *lilla*, and so on.
b. There is a rich diversity in the visual semantics of the world's languages, but this richness has often been excluded from color linguistics because it did not live up to the status of "color proper" as defined by the universalists in the Berlin & Kay school. Nuancing and paying attention to new aspects of visual semantics which are not (exclusively) chromatic, e.g., lexicalizations of visual meaning without

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

chroma, or lexicalizations in which color is combined with other aspects of visual meaning-making, is therefore now a priority for the "color + " paradigm.

c. The primacy of 'color' is a feature of European visual semantics, but the European system has spread with colonialism and globalization. In other words, the European idea of "color" and the whole system it stands for is currently being imposed on many speakers with radically different visual vocabularies. With the general influx of Eurocentric and Anglocentric ideas and semantic systems of the so-called "modern world", middle-class school children everywhere are now exposed to crayon socialization, which, according to Zimmermann et al. (2015), is one of the hotbeds for developing "socio-visual ideologies".

d. The metalanguage of color studies needs to be reformed. There is a widespread tendency to endorse Anglocentric interpretative biases, grounded in the use of English categories, and to take for granted words like "color", "hue", "blue", "bright", etc. While these might seem natural to English-speaking theorists, they turn out to be folk concepts from a very particular visual universe of meaning, namely that of "Anglo English",[2] and thus, they cannot readily be used as metalinguistic tools for the study of all languages.

In order to contextualize these insights for the present study, I will now discuss some caveats and consequences. While the deconstruction of "color" as a visual category has indeed proven to be a breakthrough for the study of global visual semantics, it is clear that all modern European languages are highly dependent on the "color" concept. In Danish, for instance, the word *farve* 'color' acts as a collective superordinate and plays a vital part in the semantic configuration of words like *blå* 'blue', *gul* 'yellow', and *lilla* 'purple'. This socialization into chromatocentric, hue-isolating visual thinking is established from an early age (see Zimmermann et al., 2015).

The other important point is that "brightness" needs to undergo the same critical analysis and cross-linguistic confrontation as "color". In English, "brightness" might seem natural and self-evident because of the centrality of the word *bright* in the English visual vocabulary. But unlike *color*, which is shared across European and Europeanized visual systems, *bright* does not need to travel far in order to be deconstructed. Danish, for instance, does not have any obvious translational equivalents for *bright* and *brightness*. This means that a speaker of Danish will find it

2. "Anglo English" is used in cultural linguistics, postcolonial semantics, and the semantics of World Englishes (see e.g., Wierzbicka, 2013; Wong, 2014; Levisen & Jogie, 2015). Levisen (2016) summaries these traditions, saying that: "'Anglo English' … is a short hand term for the standardized, prestigious kinds of Englishes associated with the historical Anglosphere", and that "'Anglo Englishes' (prestigious standards of English English, American English, Australian English, etc.) despite [their] internal differences [are] rather similar, and different from, say, Singapore English … Aboriginal English … Trinidadian English Creole … or Bislama and Tok Pisin" (p. 37).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

difficult, (or unnecessary) to distinguish "bright blue" from "light blue". In practical translations, both would most likely come across as *lyseblå* '*lys* blue'. In fact, it would seem that "light" – in its relevant visual sense – is a more translatable concept than "bright". Thus, the inverted commas around "brightness" in this chapter, signals this awareness that "brightness" is an English word, and that "brightness" is not a valid analytical category for studying Danish visual semantics or the semantics of any language. I will return to this point in Section 5.

3. Methods

In all types of semantic-conceptual analysis, it is essential not to conflate "emic" and "etic" perspectives. As memorably phrased by Wierzbicka: "Neurophysiology or physics may explain how people see colors, but not what they mean by *red* and *blue*" (2006, p. 8). The primary job of a semanticist is to study the "emic" perspective and to analyze how people make sense of their worlds through words. Most speakers are not scientists, and it would be a scientistic fallacy to assume that the folk perspectives and naïve worldviews reflected in natural semantics would align with the categories established by "etic" color physics. In my view, linguists can contribute to etic theory, mainly in the following two ways: (1) by providing cross-linguistic evidence and developing accounts of universality, areality, and culture-specificity through careful comparative, semantic analysis, (2) and by offering a metalanguage that is not locked into the framework of a particular language; in other words, a metalanguage that is fit to be used across linguacultures,[3] without imposing any culture-specific interpretative bias.

The methods section of this chapter reflects both concerns. I will take a two-pronged approach, relying on corpus-assisted discourse methods and semantic explication techniques. The corpus-assisted discourse methods utilize two databases: (1) the Danish dataset in the Evolution of Semantic Systems (EoSS) program and (2) DaTenTen, a large corpus of Danish texts hosted by SketchEngine.

The Danish dataset in the EoSS database was compiled by the author, as a part of the EoSS project. The EoSS project explored four different semantic domains, and collected empirical evidence from 50 Indo-European languages, based on psycholinguistic experiments and stimuli-based techniques. One of the EoSS domains was color. The design of the color stimuli followed the spirit of the Berlin-Kay design,

3. In linguistic anthropology, "linguaculture" is a replacement term for the traditional conceptual split into "language" and "culture". Friedrich (1989), who coined the term, elaborated on the notion in the following way: "the many sounds and meanings of what we conventionally call 'language' and 'culture' constitute a single universe of its own kind" (p. 306).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

but the interpretation of the data has been combined with a number of newer approaches and techniques (see Majid, Jordan and Dunn, 2015, and the introduction of this volume). The EoSS color dataset has the advantage of providing a narrow and focused take on Danish color. DaTenTen, by contrast, allows for a wider picture. This corpus reflects naturally occurring discourse, including visual discourse in the form of color words and related visual words. The differences in designs of the two corpora allow for a discourse study that balances the specific with the general. I will make use of these two corpora in an exploratory way and as a way of reading signals and testing analyses. The EoSS dataset was designed to answer questions about "color", not about "brightness", and this has consequences for the way I will use the EoSS evidence. What I find extremely valuable in the EoSS dataset is the variety of elicited words and the opportunity to compare lexical evidence gathered in an experimental setting, with naturally occurring words.

In my view, lexical labels are names for concepts, not for things in the world, and my reading of the EoSS data will reflect this conceptualist, post-referential view of semantics (see also Levisen, 2015). To account for the conceptually complex configuration of visual meanings in natural language, I will make use of the Natural Semantic Metalanguage (NSM) technique of reductive paraphrase. The advantage of this method is that it allows for a high-resolution analysis on conceptual grounds.

NSM explications consist of semantic primes and semantic molecules. Semantic primes are ultimately simple building blocks of meaning, which are believed to have lexical exponents across the world's languages. I have listed below the exponents of semantic primes in English (based on Wierzbicka and Goddard, 2014) and Danish (based on Levisen, 2012b, 2017).

Table 1. Semantic primes, Anglo English exponents

I ~ ME, YOU, SOMEONE, SOMETHING ~ THING, PEOPLE, BODY, KIND, PART
THIS, THE SAME, OTHER ~ ELSE ~ ANOTHER
ONE, TWO, SOME, ALL, MUCH ~ MANY, LITTLE ~ FEW
GOOD, BAD, BIG, SMALL
KNOW, THINK, WANT, DON'T WANT, FEEL, SEE, HEAR
SAY, WORDS, TRUE
DO, HAPPEN, MOVE
BE (SOMEWHERE), THERE IS, BE (SOMEONE/SOMETHING), IS MINE
LIVE, DIE
WHEN ~ TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT
WHERE ~ PLACE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE, TOUCH
NOT, MAYBE, CAN, BECAUSE, IF, VERY, MORE, LIKE

Exponents of primes exist as the meanings of lexical units, not at the level of lexemes. Exponents of primes may be words, bound morphemes or phrasemes. Exponents of primes can be formally complex. Exponents of primes can have language-specific combinatorial variants (allolexes indicated with ~). Exponents have well-specified syntactic (combinatorial) properties.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Semantic primes, Danish exponents

JEG-MIG, DU-DIG, NOGEN, NOGET, FOLK-MENNESKER, KROP, SLAGS, DEL
DEN HER, DEN SAMME, EN ANDEN
ÉN, TO, NOGET, ALLE, MEGET ~ MANGE, LIDT ~ FÅ
GOD, DÅRLIG, STOR, LILLE
VED, TÆNKER, VIL HA', VIL IKKE HA', HØRER, SER, FØLER
SIGER, ORD, DET PASSER
GØR, SKER, BEVÆGER SIG
ER (ET STED), DER ER, ER (NOGEN/NOGET), ER MIN
LEVER, DØR
TID-NÅR-DA, NU, FØR, EFTER, LÆNGE, KORT TID, NOGET TID, ET ØJEBLIK
STED, HER, OVER, UNDER, LANGT VÆK, TÆT PÅ, SIDE, INDENI, RØRER
IKKE, MÅSKE, KAN, FORDI, HVIS, MEGET, MERE, SOM

Exponents of verbs are given in their present tense forms.

Semantic molecules are relatively low-complexity concepts, which can be dissolved into semantic primes. They serve an important function as building blocks in a number of concepts. For instance, the semantic molecule *money* [m] functions as a semantic molecule in *bank*, and *children* [m] as a semantic molecule in *kindergarten*. Molecules can be domain-specific and language-specific. Some molecules have been found to occur across domains, and also across all, or the majority of the world's languages (for a recent exposition of molecule theory, see Goddard, 2016). In visual semantics, certain molecules appear to be highly productive. For instance, *the sun* [m] functions as a molecule in English visual words such as *shine*, *shiny*, and *yellow*; *fire* [m] is a molecule in *red* and *flame*; *grow* [m] and *ground* [m] are molecules in the English word *green*, and *the sky* [m] and *the sea* [m] are molecules in the English word *blue*. The molecules *during the day* [m], and *at night* [m] are needed in both "brightness"-related and "color"-related words. The English molecules *red* [m], *white* [m], *black* [m], *blue* [m] are used in the configuration of English words like *pink* (*red* [m] and *white* [m]), *grey* (*black* [m] and *white* [m]), *purple* (*blue* [m] and *red* [m]). As mentioned, some molecules appear to be universal. Apart from *the sun* [m], and *fire* [m], molecules like *during the day* [m], and *at night* [m], are candidates for universality. The same applies to *look like* [m], a highly productive molecule in visual analogies (Goddard, 2016). Other molecules are highly local. For instance, Aragón (2016) shows that *grano de café tostado* [m] 'roasted coffee grain' is needed for the analysis of Mexican "brown". Levisen, Sippola, and Aragón (2016) found that *foroba* 'locust bean' [m], *kamaron* 'shrimp' [m], among other locally salient visual "anchor concepts", were needed in the analysis of visual words in Guinea Bissau Kryol. Semantic molecules, both the universal, the areal, and the highly culture-specific ones can be lexically decomposed into semantic primes.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

In NSM explications of visual meanings, the analytical idea is to outline a hypothesis of what the word means. To exemplify this, I have proposed the following explications for the Danish words *rød*, *blå*, and *grøn*[4] in the frame 'something Z is rød/blå/grøn'. The explications below are modeled on Wierzbicka's work on the English terms *red*, *blue* and *green*.

[A] Explication for Danish *rød* 'red'

something Z is *rød*

a. people can think like this about the color of Z:
"this color is like the color of blood"
b. at the same time people can think like this:
"when people see fire, they can sometimes see this color"

[B] Explication for Danish *blå* 'blue'

something Z is *blå*

a. people can think like this about the color of Z:
"the sky is often this color during the day"
b. at the same time people can think like this:
"the sea is often this color"

[C] Explication for Danish *grøn* 'green'

something Z is *grøn*

a. people can think like this about the color of Z:
"many things that grow in the ground can be this color"

Color words tend to follow the same template, based on one, two or more components, which in turn are based are on prototypical anchor concepts or scenarios (see also Aragón, 2016, 2017). These conceptual anchors are often molecular in structure. The molecules in the [A]-[C] are singled out under the explications.

Table 3. Semantic molecules in *rød*, *blå*, and *grøn*

Semantic molecules	
Rød 'red'	*Farve* 'color'; *blod* 'blood'; *ild* 'fire'
Blå 'blue'	*Farve* 'color'; *himlen* 'the sky', *havet* 'the sea'
Grøn 'green'	*Farve* 'color'; *jorden* 'the ground'; *vokser* 'grow'

To my knowledge, no explications for "brightness"-related terms have previously been proposed within the NSM literature. Below, I will suggest three working

4. Danish NSM versions of the explications are attached in the Appendix.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

hypotheses in the form of explications for English *bright* and *light*, and for Danish *lys*. All three terms, *bright*, *light*, and *lys*, are polysemous, so it should be clarified that my explications for these words are only meant to cover the color frames of the words. In other words, I attempt to explicate the senses of *bright* and *light* in relation to *color*, and for *lys* in relation to *farve*.

[D] Explication for English *light* (in relation to color)

The color of thing Z is *light*

a. when someone sees this thing Z, they can think about Z's color like this:
"something can look like this during the day,
if the sun is shining on this thing"

[E] Explication for English *bright* (in relation to color)

The color of Z is *bright*

a. when people see this thing Z, they can think about Z's color like this:
"the color of Z is shining,
like when the sun is shining during the day"
b. because of this, it is like this: if people are near this thing Z, they can't not see this color
c. people can feel something good when they see a color like this

[F] Explication for *Danish lys* (in relation to *farve*)

The color of Z is *lys*

a. when people see this thing Z, they can think about Z's color like this:
"something can look like this during the day,
if the sun is shining on this something"
b. people can feel something good when they see a color like this

Table 4. Semantic molecules in *light*, *bright*, and *lys*

Semantic molecules	
Light	*Sun, shining, look like, during the day*
Bright	*Sun, shining, during the day*
Lys	*Sol 'sun', skinner 'shine'*

The initial hypothesis put forward in these explications is that English *light* and *bright* are both configured around a prototypical scenario that involves *the sun* [m] and a scenario that happens *during the day* [m]. In my analysis, *bright* is "active", and based on a visual analogy with *the sun* and its quality of *shining*. *Light*, on the other hand, conceptualizes a more "passive" visual scenario, based on a "sunlit" prototype, in which the color to some degree is "washed out". *Bright*, moreover

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

seems to have two additional features. The semantics of *bright* includes a component of "visual conspicuousness", cf. the component 'if people are near this thing Z, they can't not see this color'. Also, there is an emotional component. *Bright colors* somehow incite "good feelings" (cf. e.g., the *bright* side of life). Danish *lys* seems fully equivalent with the "passive" scenario in *light*,[5] but *lys* shares with *bright*, the positive emotional component that 'people can feel something good when they see something like this', a point to which I turn in Section 4 (cf. e.g., *de lyse sider af livet* 'the *lys* sides of life').

4. Linguaculture and lexicogrammar

The discourse of *lys* 'light' plays a very important role in Danish linguaculture and lexicogrammar. Danish academia has responded to this culturally constituted emphasis on *lys* 'light' by studying *lyskultur* 'lys culture' (see Bille, 2015, for studies in the anthropology of light and luminosity in the Danish context). The linguistics of *lys* is less well studied (but see Levisen, 2012b). In the following, I will provide a short overview of the ways in which *lys* 'light' been elaborated in Danish linguaculture. After that, I will explore three lexicogrammatical frames through which *lys* 'light' can combine with *farve* 'color' in Danish.

Lys is a polysemous word, with a scope that covers both visual and material meanings. The noun *lys* 'light' is an antonym of *mørke* 'darkness', and the adjective *lys* 'light' is an antonym of *mørk* 'dark'. Of course, such antonymic duality is common across languages, and the conceptual basis for this dualistic discourse is, predictably, two semantic molecules: *om dagen* 'during the day' [m], and *om natten* 'during the night' [m]. However, evidence suggests that *lys* has penetrated Danish discourse, concept formation, and speech routines in a way that *mørk* 'dark' and *mørke* 'darkness' have not. Apart from the mass noun *lys* 'light', and various general count nouns, the most culturally salient lexical unit is that of *lys* 'candle'.[6] *Lys* 'candle' is linked with the Danish cultural keyword *hygge* and is often effectively framed as *levende lys* 'living lys' (Levisen, 2012, p. 108). Apart from its link

---

5. Danish has a number of other *lys*-related adjectives that, like English *bright*, share a more "active" scenario. These include *lysende* 'shining, bright, luminous', *skinnende* 'shining, bright,' and *strålende* 'radiant, bright'. Note that *lysende*, *skinnende*, and *strålende*, are not typically used in combination with color words.

6. *Stearinlys* 'candles' are often conceptualized as *levende lys* 'living lights', as opposed to electric lights, and are valued as facilitators of atmospheric agents of a positive sociality (Levisen, 2012b, p. 108).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

to sociality constructs such as *hygge*, *lys* is prominent in phraseology and ritual communication. For instance, collocations such as *dejligt lyst* 'nice and *lys*' and extended versions such as *hvor er her dejligt lyst!* 'it so nice and *lys* in here!', are typical in discourses of design, architecture and real estate talk. In discourses of personality, *lys* is used in combination with the hard-to-translate Danish personhood construct *sind* 'disposition, inclination, mind' (Levisen, 2012a; Levisen, 2017). A person with a *lyst sind* is "inclined to think positively". While the evidence from polysemy, phraseology, and discourse is revealing, the "keyness" of *lys* in Danish linguaculture also has ramifications for lexicogrammar.

The combinatorics of *lys* 'light' with *farve* 'color' is complex. There are three lexicogrammatical combinations, with three different semantic profiles. Using *lilla* 'purple' as an example, these three combinations are:

(1) a. *lys lilla* 'light purple'
b. *lyselilla* 'light-purple'
c. *lyslilla* 'light-purple'

All three combinations would be rendered in English translation as "light purple", but there are subtle semantic and indexical meaning differences, which I will account for in the following. The construction in (1a) is an ad-hoc generated phrase consisting of an attributive adjective and a color word. This construction differs formally from (1b) and (1c) which are both lexically fused words in the form of compounds. The compound construction in (1b) uses the interfix -e for the combination of *lys* + color, whereas the (1c) construction is interfix-free. In spoken Danish, stress patterns and vowel length distinctions make the forms of the three constructions discernable.[7]

Consider now the EoSS dataset, and the different distributions of (a) the ad-hoc or "free" combinations with the attributive adjective *lys*, (b) the *lyse*-compound, and (c) the *lys*-compound.

7. As in all types of communication, and especially communication that involves vocoid structures, schwa, and schwa reductions, there could of course be certain instances of genuine doubt on how to classify and interpret these words and strings of words. The point that I am making here is a general one, namely that there are three distinguishable types of the "grammar of lys", and that these are recognizable in speech, as well as in writing.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(2) "Free" attributive forms + color words attested in EoSS

lys grå	'light grey'
lys lilla	'light purple'
lys violet	'light violet' (alt: purple, mauve)
lys mintgrøn	'light mint-green'
lys blå	'light blue'
lys turkis	'light turqoise'
lys orange	'light orange'
lys laksefarvet	'light salmon-colored'
lys mintblå	'light mint-blue'
lys blågrøn	'light blue-green'
lys brun	'light brown'

(3) Compound words based on *lyse-* attested in EoSS

lysegrå	'light-grey'
lysegrøn	'light-green'
lyseblå	'light-blue'
lyselilla	'light-purple'
lyserød	'light-red'
lysebrun	'light-brown'

(4) Compound words based on *lys-* attested in EoSS
[No examples attested].

From EoSS we can learn at least three things: (1) The free *lys*-construction appears to attract a variety of color words, both core color words such as *blå* 'blue', *brun* 'brown', and *grå* 'grey', and more peripheral words such as *laksefarvet* 'salmon-colored' and *blågrøn* 'blue-green'; (2), the *lyse*-compound attracts only core color words, and (3) the *lys-* compound is not attested in this dataset. Interestingly, speakers do not seem to use this construction at all in the tasks of the EoSS experiment. As the EoSS data was based on young Danish speakers (aged 20–30) a possible explanation for this could be the age factor of the participants. The DDO (Den Danske Ordbog) dictionary provides a different clue, suggesting that *lys-* is a *faglig* 'professional' variant of *lyse-* (Lyserød, DDO). If the latter hypothesis is correct, *lys-* indexes "profession", rather than "age".

Also in the DaTenTen corpus, we find that the *lyse*-compound occurs much more frequently than the *lys*-compounds. At the same time, we find examples of the *lys*-compounds for all the major color categories.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(5)

lyse-	*lys-*
lyserød **32259**	*lysrød* 110
lyseblå **9430**	*lysblå* 183
lysegrøn **4680**	*lysgrøn* 143
lysebrun **4111**	*lysbrun* 140
lysegrå **2119**	*lysgrå* 182
lysegul **1517**	*lysgul* 106

Furthermore, the DaTenTen also shows that for at least five colors, there are more *lys*-based than *lyse*-based compounds:

(6)

lyse-	lys-
lyslilla **745**	*lyselilla* 279
lysviolet **81**	*lyseviolet* 2
lysorange **30**	*lyseorange* 7
lysbeige **16**	*lysebeige* 4
lysturkis **9**	*lyseturkis* 0

Of the *lys*-favoring colors, only *lilla* has a substantial number of occurrences. Looking at this list, it is obvious that formal constraints could explain in part the distribution: the *lyse*-compound appears to attract monosyllabic words, and frequently occurring Danish color words such as *rød* 'red', *blå* 'blue', *grøn* 'green', *brun* 'brown', *grå* 'grey', and *gul* 'yellow'. *Lys*-, on the other hand, attracts polysyllabic words which are also less frequent and more "French-and-foreign-sounding" words in Danish such as *lilla* 'purple', *violet* 'violet', *orange* 'orange', *beige* 'beige', and *turkis* 'turquoise'. From a semantic point of view, it would seem that *lyse-* is a default, which attracts the most frequent and common color words, and that *lys-* attracts, and/or creates more of an "expert" ring to the words. As we have seen, none of EoSS participants used the *lys*-compound, not even in the case of *lilla* 'purple'. They all used *lyselilla*, but never *lyslilla*. Based on the hypothesis that *lys-* is a *faglig* 'professional, expert' variant of *lyse-*, the evidence points toward an indexically-grounded difference. This takes us to the following, almost identical, explications:

[G] Semantic explication for the '*lyse* + color' compound

e.g., thing Z is *lyselilla* (light-purple)

a. when people see the color of Z, they can think like this:
   "the color of Z is one kind of purple"
b. at the same time these people can think like this:
   "this kind of purple is a *lys* purple"

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

[H] Semantic explication for the '*lys* + color' compound

e.g., thing Z is *lyslilla* (light-purple)

a. when people see the color of X, they can think like this:
   "the color of Z is one kind of purple"
b. at the same time these people can think like this:
   "this kind of purple is a *lys* purple"

   [if people know many things about colors, when they say something about this color, they will say it with this word *(lysgul lyslilla, lysgrøn, etc.)*, not with another word]

Both of these two compound constructions construe color and *lys* by way of an integrated category, as 'a kind of purple, yellow, etc.', namely the *lys* kind. By my analysis, the two compound forms differ only in one parameter: the *lys*-compound has an additional metapragmatic annotation (marked in brackets). The indexicality of *lys*-based words can be paraphrased as 'knowing many things about colors'. By contrast, the free attributive construction simply juxtaposes the two ideas: 'the color of Z is purple' + 'the color of Z is a *lys* color', i.e., there is no 'kind'-based integration of the two components.

[I] Semantic explication for the '*lys* + color' construction

e.g., thing Z is *lys lilla*

a. when people see the color of Z, they can think like this:
   "the color of Z is purple"
b. at the same time these people can think like this:
   "the color of Z is a *lys* color"

The three grammatical frames combining *lys* with *farve* are all productive. But as previously noted, there is more to the story. The most frequent and salient *lyse*-based lexicalizations have achieved a life on their own – they have, so to speak, developed "post-systemic" meanings, and this means that they do not, or at least do not completely conform to the grammatical semantics described above. *Lyserød* is the most obvious "defiant" lexicalization, but the two other compounds *lysegrøn* and *lyseblå* also show signs of drifting away from the three-way system accounted for with the explications [C], [D], [E]. I will now review these cases one by one.

4.1 The meaning of *lyserød*

The EoSS-based study on Germanic color words (Vejdemo et al., 2015) concluded that Danish *lyserød* is "not a kind of *rød*", but a separate color category in Danish. This makes Danish *lyserød* dissimilar to the Swedish compound *ljusröd* 'light-red'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

precisely for the reason that Swedish 'light-red' stands for a kind of *röd* 'red'. *Lyserød* and *ljusröd*, then, are examples of cryptodiversity – when the forms are similar, but the meanings differ (cf. Levisen & Jogie, 2015; Levisen, 2016). Instead, the comparative cross-Germanic analysis found that *lyserød* is an equivalent of English *pink,* and Swedish *rosa* 'pink'. The study also found that Danish discourse – seen from an English perspective – operates with "two pinks": *lyserød,* roughly '*pink* in general', and *pink*, roughly, 'conspicuous *pink*'.[8] Additionally, a third more marginal Danish word *rosa*[9] signifies a 'mild pink'.

Lyserød is a high-frequency word in both EoSS and DaTenTen. The sheer frequency of the word suggests that it differs from other *lyse*-compounds (compare e.g., *lyserød* 32259 with *lysegul* 'light yellow' 1517). *Lyserød* can also be modified. In the EoSS dataset, we find examples such as:

(7) *pastellyserød* 'pastel-lyserød'
Miss-Piggy-lyserød 'Miss-Piggy-lyserød'
mørk lyserød 'dark lyserød'

Of these examples, *mørk lyserød* 'dark lyserød' is perhaps the most revealing. *Lyserød* stands out from other *lyse*-based compounds in this respect. To compare, constructs like **mørk lysegul* 'dark light yellow', or **mørk lyselilla* 'dark light purple' would sound contradictory or humorous. Apart from frequency and modifiability, *lyserød*'s independence is also evident from its combinatorial properties. One can say "*lyserød og rød*", (or "*rød og lyserød*"), whereas "**lysegul og gul*" sounds odd. To conclude, we can say that *lyserød* is 'a color of one kind', and not 'a kind of *rød*'. This is not to say that there is no semantic component of *rød* 'red' in *lyserød*. Wierzbicka suggests that English *pink* – with *orange*, *purple*, and *grey* – is conceptualized as an admixture. She explains: "This does not mean that we do not think of orange, pink, purple, and grey as unitary colors […] but at some level of our linguistic consciousness we can also relate each of them to two other color concepts" (Wierzbicka, 1996, p. 325). This analysis is helpful for the *lyserød* case. *Lyserød* is, semantically speaking, a conceptual admixture, but the catch here is that *lyserød* is not an admixture of *rød* 'red' and *lys* 'light'. *Lyserød* is, just like English *pink*, an admixture of *rød* 'red' and *hvid* 'white'. In NSM, this implies that for *lyserød* the explication will rely on both *color* [m], as well as *red* [m] and *white* [m]:

---

8. A similar case is found in Finnish where *vaalean-punainen* is formally composed of 'light' + 'red', but means 'pink' (on *vaalean-punainen* see Uusküla, 2007, for a cross-Finno-Ugric study, see also Uusküla, Hollman & Sutrop, 2012, p. 61).

9. Danish *rosa*, roughly, 'mild pink' adds another layer of cryptodiversity in European color words, where *rosa* in both German, Spanish, Swedish and several other languages simply stand for 'pink in general'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Z is lyserød (partial explication)
a. when people see the color of Z, they can think like this:
"it is a color of one kind,
b. at the same time, they can think like this:
"it looks like red, it looks like white"

There seems to be one additional feature of *lyserød* that needs to be accounted for: there is a very strong "girly signal" in DaTenTen. Consider the following high-frequent collocates of *lyserød*:

(8)	*hun, hende, hendes,*	'she, her (ACC), hers'
	piger, prinsesser,	'girls, princesses'
	kjoler, tøj, hår, sløjfe, neglelak	'dresses, clothes, hair, bow, nail polish'
	blomster, roser	'flowers, roses'
	sky, elsker, svæver, hjerter	'cloud, love, floating, hearts'
	farve, glimmer, smuk, skær, sød,	'color, glitter, beautiful, gleam, sweet'

A traditional account of color might consider such signal to be of an associative nature and thus not a part of the 'color proper'. This, to me, is a very problematic position, reflecting a referentialist bias. The signal clearly indicates that there is another semantic molecule in *lyserød*, namely *pige* 'girl', and that this is not simply a loose association or private connotation, but a defining element in the conceptual configuration of *lyserød*.

[J] Semantic explication for *lyserød*

Z is lyserød (full explication)
a. when people see the color of Z, they can think like this:
"it is a color of one kind"
b. at the same time, they can think like this:
"it looks like something red, it looks like something white"
c. many small girls can feel something good when they see this color,
many small girls can feel something good when they think about this color

Table 5. Semantic molecules in *lyserød*

Semantic molecules	
Lyserød 'pink' (lit. light red')	*farve* 'color' [m], *rød* 'red' [m], *hvid* 'white' [m], *piger* 'girls' [m]

The independent and elaborate semantics of *lyserød* is a distinct feature of *lyserød*, which is not found in the other grammatical frames described in 4.1. For instance, *rød* 'red' can also be modified with *lys*, as in the construction *lys rød,* which

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

essentially means, 'this color is red – I want to say something more about it: it is *lys*'. The "professional" word *lysrød* does not appear to have made the move to 'pinkness', and thus, except for its indexical meanings, it seems to be the equivalent of Swedish *ljusröd*. Also, neither *lys rød* nor *lysrød* relies on *girl* [m] as a semantic molecule.

4.2 The meaning of *lysegrøn*

Lysegrøn is a prominent word in Danish color discourse. According to a modern Danish proverb, *håbet er lysegrøn* 'hope is *lysegrøn*'. While *grøn* 'green' is believed to be *godt for øjnene* 'good for the eyes', *lysegrøn* is associated with the optimism of springtime. A modern hymn synesthetically combines *lysegrøn* with smell: *det dufter lysegrønt af græs* 'there is a *lysegrøn* smell of grass' (Johansen, 2002). This aura of positive discourse surrounding *lysegrøn* must be accounted for in the semantic analysis, and the linguacultural elaboration of *lysegrøn* bears witness to the relative prominence of the word. At the same time, it seems clear that its independence from *grøn* 'green' is of a different kind than that of the *lyserød-rød* relation. For instance, "*lyserød* and *rød*" is more much common than "*lysegrøn* and *grøn*", although the latter is possible.

(9)	'*Lysegrøn og grøn*'	+ '*grøn og lysegrøn*'	
	1	+ 4	= 5
	'Lyserød og rød'	+ '*rød og lyserød'*	
	50	+ 95	= 145

The EoSS dataset has no examples of *lysegrøn* being modified, but rather features numerous examples of *grøn* modified either through a compound construction, attributive constructions, or derivational morphology:

(10)	*avocadogrøn*	'avocado-green'
	flaskegrøn	'bottle-green'
	limegrøn	'lime green'
	mørk grøn	'dark green'
	fesen grøn	'shitty green'
	grønlig	'greenish'

On the surface, it would seem that *lysegrøn* is a kind of *grøn*, just like the general *lyse*-based compound morphology would predict. As proposed earlier, the explications for *grøn* involves the molecules *grow* [m] and *ground* [m] – and a prototypical scenario suggestive of something like 'vegetation', but without a particular conceptual anchor. Now consider the following collocations of *lysegrøn* in DaTenTen.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(11)	*blade, løv, græs, skov, træerne,*	'leaves, leafage, grass, woods, trees'
	bøgeskov, bøgeblade, bøgen	'beech wood, beech leaves, the beech'
	foråret, skud, nyudsprunge, spæde,	'spring, shoot, new shoots, tender,
	spirer, knopper	sprout, buds'
	farve, lys, dufter, smuk, flot, fin,	'color, light, smell nice, beautiful, mag-
	frisk	nificent, fine, fresh'
	håbet, håb	'hope (determinate form), hope'

This evidence suggests two things. Firstly, that *lysegrøn* has developed a more specific prototypical scenario within the "vegetation" theme of *grøn* 'green', namely a prototype based on a scenario of sunlit leaves in spring. Secondly, the evidence from phraseology and collocations suggests that *lysegrøn* is an inherently "positive" color. In NSM:

[K] Semantic explication for *lysegrøn*

Z er lysegrøn
a. when people see the color of Z, they can think like this:
"this color is one kind of green
it is like the color of the leaves of trees in spring
when the sun shines on the leaves during the day"
b. when people see this color, they can feel something good

Semantic molecules	
Lysegrøn (lit. light green)	*farve* 'color' [m], *green* 'grøn' [m], *blade* 'leaves' [m], træer 'trees', *forår* 'spring', *sol* 'sun', *skinner* 'shines' [m], *om dagen* 'during the day' [m]

4.3 The meaning of *lyseblå*

Lyseblå resembles the *lysegrøn* case in many ways. *Lyseblå* is also a discursively salient and frequent color word in Danish discourse. It is one kind of *blå* 'blue' and thus, we see no radical departure from the meaning of *blå*. Instead, we find a parallel development to that of *lysegrøn*, namely the development of a specific prototype, with a specific and integrated "lightness" scenario.

In EoSS, there is only one type of modification of *lyseblå*. It is modified by the intensifier *meget* 'very', in both double and triple ways:

(12)	*meget lyseblå*	'very lyseblå'
	meget meget lyseblå	'very, very lyseblå'
	meget meget meget lyseblå	'very, very, very lyseblå'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

These examples are interesting because *meget* 'very' in these constructions must clearly be modifying *lys* 'light', not *blå* 'blue. What does it mean for the analysis of *lyseblå*? To recap, the double prototype proposed for the explication for *blå* 'blue' was based on the semantic molecules *havet* 'the sea' and *himlen* 'the sky'. Out of the two prototypes in *blå*, DaTenTen shows that *lyseblå* invokes only one of these prototypes: *himlen* 'the sky'.

(13) *himmel, himlen, sky, skyer, himmelen* 'sky, the sky, cloud, clouds, the sky'
farve, smuk, lys, flot 'color, beautiful, light, nice'

Interestingly, the collocational profile for lyseblå[10] is not nearly as extensive as that of *lysegrøn*. But the corpus signal is clearly sky-oriented. Let us take a closer look at some common examples in the corpus.

(14) *Solen skinner fra en smuk lyseblå himmel*
'The sun is shining from a beautiful *lyseblå* sky'

(15) *Vejret var rigtig godt, lyseblå himmel og omkring de 30 grader.*
'The weather was really good, *lyseblå* sky and around 30 degrees'.

(16) *Og så solen, der omfavner de hele fra en lyseblå himmel. Hvor er det smukt herude.*
'And then the sun that embraces it all from a *lyseblå* sky. How beautiful it is out here'.

(17) *Til grill i haven, sand og strand og* lyseblå *himmel.*
'For a barbeque in the garden, sand and beach and *lyseblå* sky'

(18) *Så er min mentale himmel blevet lyseblå igen*
'Now my mental sky has become *lyseblå* again'

It is clear from the examples that *lyseblå* shares with *lysegrøn* a "positive" semantics. Like *lysegrøn*, it has an integrated prototype based on 'the color of the sky during the day, when the sun is in the sky':

[L] Semantic explication for *lyseblå*

Z er lyseblå
a. when people see the color of Z, they can think like this:
"this color is one kind of blue
it is like the color of the sky
during the day, when the sun is in the sky"
b. when people see this color, they can feel something good

10. *Lyseblå* also collocates with the two football clubs 'Manchester City' (UK) and 'Sønderjyske' (Denmark), as well as clothing items: *skjorte, kjole, slips, bluse, jeans.* Note that there is no built-in "gendered" meaning in *lyseblå:* both *kjole* 'dress' and *slips* 'tie' are collocates.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Semantic molecules for *lyseblå*

Semantic molecules	
Lyseblå (lit. light blue)	*farve* 'color' [m], *blue* 'blå' [m], *himlen* 'the sky', *om dagen* 'during the day', *solen* 'the sun' [m]

4.4 Summary

To conclude, this section has put forward several new claims about color and "brightness" in Danish. Firstly, that there is a grammar of *lys*, with three possible ways of bringing light and color together: *lys*, *lyse-*, and *lys-*, all with different semantic (or indexical) meanings. Secondly, there are three exceptions to this general grammar of color:

- *Lyserød* is not a kind of *rød* 'red', but has developed into an independent color. It has departed from its lexical origin and does not incorporate *lys* as such in its semantics. Instead, it relies on a conceptual admixture of *hvid* 'white' and *rød* 'red', by way of visual similarity, rather than being a subordinate of *rød* 'red'.
- *Lysegrøn and lyseblå* also show signs of independence, but in a less radical way than *lyserød*. *Lysegrøn* and *lyseblå* rely in part on the concepts of *grøn* 'green' and *blå* 'blue', but they also developed their own conventional meaning which cannot be deduced from the grammar of *lys* + *farve* alone, and therefore, they are worth studying in their own right, and from a post-systemic perspective. While they are still "kinds of *grøn/blå*", they have (a) developed their own prototype, and (b) their own solution to the integration of light and color. Both are based on a natural prototypical scenario, and both are discursively salient color words with a "positive" semantics.

5. Back to "brightness"

In her essay "The ambiguity of brightness", Carole Biggam laments the lack of agreement about what "brightness" means in the linguistic and anthropological literature on color, and she describes at length what she means by the term. I sympathize with Biggam's attempt to arrive at an unambiguous metalanguage, and her plea for conceptual clarity, but I am not so convinced about her attempt to save "brightness" as a valid analytical category. Surely, "brightness" may have some role to play in English color linguistics, but as this study shows, the "bright/brightness" theme in English semantics is not applicable to global visual studies. In other words, "brightness" should not be saved. "Brightness" should be abandoned as an analytical category. The reason for this is that "brightness" is an English-specific concept that will most likely cause

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

confusion when imposed on other languages where such a category does not exist. This does not mean that we should completely stop doing researching on *brightness*. Firstly, the words *bright* and *brightness* are important words in English(es), and from an Anglo "emic" perspective, there is much more to say about these terms than what I have discussed here. Also, there may be semi-equivalents in a number of languages, such as Danish *lys*, and these can be studied comparatively with *bright/ness*.

It will, hopefully, be crystal clear at this point, that my initial use of "brightness" in the title of this chapter is not an endorsement of the concept, but an attempt to raise awareness of the non-universality of the category. Based on the encounter with Danish, one might object that a better term for comparative studies could be "lightness", rather than "brightness", given that the Danish *lys* shares its basic scenario with "light", and not "bright". Color physics pays great attention to the "brightness" vs. "lightness" distinction, and some color linguists seem to follow in their footstep. But adding "-ness" to emic English words will hardly transform it into terminology that is fit for cross-linguistic purposes. The tendency to take ordinary words from the English lexicon of visuality such as *bright*, *light*, *shiny*, *pale*, and turn them into "etic-sounding" words like *brightness*, *lightness*, *shininess*, *paleness*, and so on, may add a rhetorical air of gravitas to the terms, but this does not solve the crisis in color linguistics, which in my view has one name: Anglocentrism (see also Levisen, 2018). The naturalization of English "color talk" has caused tremendous confusion in global visual semantics, and by adding "brightness talk", we might solve the issue of hue-centrism, but not cut the roots of Anglocentrism. The problem is not that most works about color and visuality are written in English, or that people use English words heuristically in comparative studies. The problem emerges when whole frameworks of interpretation, including terminologies and theoretical notions, rely on Anglocentric concepts. Until the problem of metalanguage is solved, such Anglocentric stranglehold cannot be dismantled, and we are prevented from making real progress in visual semantics. The problem runs deep – it is not only key terms such as "color", "brightness" and semi-technical terms such as "saturation" that cause the problem. Numerous English-specific words are taken for granted in color studies. A good example is the word "hue", a mystery category to many non-Anglo linguists, and a word without equivalents in many European languages. Thus, the difference between "hue" and "color" in English is lost on many non-English speakers, and as a non-Anglo color theorist, I find it difficult to accept "hue" and "brightness" as baseline categories in any "global theory of color" if these cannot even be translated readily into a language like Danish, which is so closely related English.

Nevertheless, there is cause for optimism. Progress in analyzing culture-specific meanings has been made through the globally "shared emic" vocabulary of semantic primes. When color meanings and other visual meanings are explicated in the Natural Semantic Metalanguage, we can begin to compare and contrast

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

meanings across cultural communities and appreciate the semantic diversity and complexity of natural language categories. The obscurity and Anglocentrism of words like "brightness", "saturation", "hue", and similar, can be circumvented by the NSM technique, because it allows the researcher to go emic, i.e., to use only words that exist for the speakers of the language in question, and words which, at the same time, are cross-translatable. By phrasing research questions and results around molecules like "light", "sun", "fire", "during the day", "at night", etc., rather than English-specific terms such as "brightness" would improve the comparability of the analysis, and thus, help setting up a more global visual semantics in which the problem of Anglocentrism can be circumvented.

6. Concluding remarks

In this chapter, "brightness" was denaturalized and deconstructed with the help of Danish visual semantics and the Natural Semantic Metalanguage. Studying the Danish compounds *lyserød*, *lysegrøn*, and *lyseblå,* and the intricate Danish system of combining *lys* 'light' with *farve* 'color', the chapter demonstrated how emic analysis of local lexicogrammar can help progress visual semantics, and the "color + " paradigm. Based on the evidence presented in the chapter, a discussion of emerging insights of the visual semantics paradigm was provided in which three points were advanced: (1) "brightness" is a problematic term in cross-linguistic studies, because of its status as an "Anglo" and hard-to-translate term. Instead, a metalinguistic reform was recommended in which universal semantic molecules such as *the sun* [m], *during the day* [m], *fire* [m], and other more locally important molecules were proposed as alternatives to the "brightness discourse"; (2) emic concepts – as exemplified by the intricate Danish system of *lys* and its post-systemic developments – are the key to any ethnotheory of visuality. It is important to explicate the meanings of local words and grammars if we are to make headway in understanding the complexities of natural semantics in the visual domain; (3) the Natural Semantic Metalanguage method allows for a culture-sensitive analysis. Future studies in "color + " can benefit from using this method in comparative analysis, and eventually, semantic typologies of color and visuality can be built from the high-resolution semantics that the NSM method allows.

Acknowledgment

I have greatly benefited from discussing the aspects of the analysis presented in this paper with Sophia Waters, Mari Uusküla, Susanne Vejdemo, and Karime Aragón. I would also like to thank the two anonymous reviewers for their comments and suggestions. Needless to say, the analysis and interpretations presented in this chapter, including any mistakes or misunderstandings, remain my own.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

References

Aragón, K. (2016). Mexican Colors and Meanings: An Ethnolinguistic Study of Visual Semantics in Oaxaca. In G. Paulsen, M. Uusküla, & J. Brindle, (Eds.), *Color Language and Color Categorisation* (pp. 302–332). Newcastle Upon Tyne: Cambridge Scholars Publishing.

Aragón, K. (2017). Visuality, Identity and Emotion: *Rosa Mexicano* as a Mexican Spanish keyword. In C. Levisen, & S. Waters (Eds.), Cultural Keywords in Discourse [Pragmatics and Beyond New Series] (pp. 132–156). Amsterdam / Philadelphia: John Benjamins.

Berlin, B., & Kay, P. (1969). *Basic Color Terms*. Berkeley: University of California Press.

Biggam, C. (2007). The ambiguity of brightness (with special reference to Old English) and a new model for color description in semantics. In R. E. Maclaury, G. V. Paramei, & D. Dedrick (Eds.), *The Anthropology of Colour: Interdisciplinary Multilevel Modelling* (pp. 171–188). Amsterdam: John Benjamins. https://doi.org/10.1075/z.137.12big

Biggam, C. P. (2012). *The Semantics of Colour: A historical approach*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139051491

Bille, M. (2015). Lighting up cosy atmospheres in Denmark. *Emotion, Space and Society*, 15, 56–63.

Casson, R. W. (1997). Colour shift: evolution of English colour terms from lightness to hue. In C. L. Hardin, & L. Maffi (Eds.) *Colour Categories in Thought and Language* (pp. 224–239). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519819.010

Friedrich, P. (1989). Language, Ideology, and Political Economy. *American Anthropologist*, 91(2), 295–312. https://doi.org/10.1525/aa.1989.91.2.02a00010

Goddard, C., & Wierzbicka, A. (2014). *Words and Meanings: Lexical Semantics across Domains, Languages and Cultures*. New York: Oxford University Press.

Goddard, C. (2016). Semantic molecules and their role in NSM lexical definitions. *Cahiers de Lexicologie*, 2(109), 13–34.

Hill, C. (2011). Named and Unnamed Places. Color, Kin, and the Environment in Umpila. *The Senses and Society*, 6, 57–67. https://doi.org/10.2752/174589311X12893982233759

Hjorth, E., Kristensen, K. et al. (2003). Den Danske Ordbog. Copenhagen: Det Danske Sprog- og Litteraturselskab. (DDO)

Johansen, J. (2002). Det dufter lysegrønt af græs. In *Den Danske Salmebog. Det Kgl. Vejsenhus Forlag*.

Jones, W. J. (2013). *German Colour Terms: a Study in Their Historical Evolution from Earliest Times to Present*. Amsterdam: Benjamins. https://doi.org/10.1075/sihols.119

Kay, P. (2006 [1999]). Methodological Issues in Cross-Language Color Naming. In C. Jourdan & K. Tuite (Eds.), *Language, Culture and Society* (pp. 115–134). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511616792.006

Kay, P., Berlin, B., Maffi, L., Merrifield, W. R., & Cook, R. (2009). *The World Colour Survey*. Stanford: CSLI.

Lucy, J. (1997). The linguistics of "color". In C. L. Hardin & L. Maffi (Eds.), *Color Categories in Thought and Language* (pp. 320–346). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519819.015

Levisen, C. (2012a). *Cultural Semantics and Social Cognition: A Case Study on the Danish Universe of Meaning*. Berlin: De Gruyter. https://doi.org/10.1515/9783110294651

Levisen, C. (2012b). "Reds" and "Pinks" in Scandinavia and beyond: Language as the key to the Ethnotheory of Colour. Handout, Departmental Lecture, Aarhus University, 6th June.

Levisen, C. (2015). Scandinavian Semantics and the Human Body: An ethnolinguistic study in diversity and change. *Language Sciences*, 49, 51–66. https://doi.org/10.1016/j.langsci.2014.05.004

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Levisen, C. (2016). Postcolonial Lexicography: Defining Creole Emotion Words with the Natural Semantic Metalanguage. *Cahiers de Lexicologie*, 109(2), 35–60.

Levisen, C. (2017). Personhood Constructs in Language and Thought: New Evidence from Danish. In Z. Ye (Ed.), *The Semantics of Nouns* (pp. 120–144). Oxford: Oxford University Press.

Levisen, C. (2018). Biases we live by. Anglocentrism in Linguistics and Cognitive Sciences. *Language Sciences*, Special Issue on Biases in Linguistics, Ed. by S. Borchmann, C. Levisen, and B. Schneider.

Levisen, C., & Jogie, M. R. (2015). The Trinidadian Theory of Mind: Personhood and Postcolonial Semantics. *International Journal of Language and Culture* 2(2), 169–193. https://doi.org/10.1075/ijolc.2.2.02lev

Levisen, C., Sippola, E., & Aragón, K. (2016). Color and Visuality in Iberoromance Creoles. Towards a postcolonial semantic analysis. In G. Paulsen, M. Uusküla, & J. Brindle (Eds.), *Color Language and Color Categorization* (pp. 270–301). Newcastle upon Tyne: Cambridge Scholars Publishing.

Majid, A., Jordan, F., & Dunn, M. (2015). Semantic systems in closely related languages. *Language Sciences*, 49, 1–18. https://doi.org/10.1016/j.langsci.2014.11.002

Saunders, B. (1992). *The invention of basic colour terms*. Utrecht: ISOR.

Saunders, B. A. C., & van Brakel, J. (1997). Are there nontrivial constraints on colour categorization? *Behavioral and Brain Sciences*, 20 (2), 167–228.

Saunders, B., & van Brakel, J. (Eds.). (2002). *Theories, technologies, instrumentalities of colour: Anthropological and historiographic perspectives*. London: University Press of America.

Uusküla, M. (2007). The basic colour terms of Finnish. *SKY Journal of Linguistics*, 20, 367–397.

Uusküla, M., Hollman L., and Sutrop, U. (2012). Basic color terms in five Finno-Ugric languages and Estonian Sign Language: A comparative study. *Journal of Estonian and Finno-Ugric Linguistics* 3(1), 47–86.

Uusküla, M., Liivi, H., & Urmas S. (2012). Basic colour terms in five Finno-Ugric languages and Estonian Sign Language: A comparative study. *Journal of Estonian and Finno-Ugric Linguistics*, 3(1), 47–86.

Uusküla, M., & Eessalu, M. (2018). Glossy Black is not actually 'Black': Evidence from Psycholinguistic Colour-Naming Studies in 14 European Languages. *Cultura e Scienza del Colore – Color Culture and Science Journal*, 09, 39–44.

Vejdemo, S., Levisen, C., Beck, T. G., von Scherpenberg, C., Næss, Å., Zimmerman, M., Stockall, L., & Whelpton, M. (2015). Two Kinds of Pink: Development and Difference in Germanic Colour Semantics. *Language Sciences*, 49, 19–34. https://doi.org/10.1016/j.langsci.2014.07.007

Vejdemo, S. (2017). Triangulating perspectives on lexical replacement. From predictive statistical models to descriptive color linguistics. (Doctoral dissertation). Stockholm: Stockholm University.

Wierzbicka, A. (1996). *Semantics: Primes and Universals*. Oxford: Oxford University Press.

Wierzbicka, A. (2005). There are no 'color universals', but there are universals of visual semantics. *Anthropological Linguistics*, 47(2), 217–244.

Wierzbicka, A. (2006). The semantics of colour: A new paradigm. In: C. P. Biggam & C. J. Kay (Eds.), *Progress in Colour Studies*. Volume 1: *Language and Culture*. (pp. 1–24). Amsterdam: John Benjamins. https://doi.org/10.1075/z.pics1.05wie

Wierzbicka, A. (2008). Why there are no 'colour universals' in language thought. *Journal of the Royal Anthropological Institute*, 14, 407–425.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wierzbicka, A. (2013). *Imprisoned in English: The Hazards of English as a Default Language*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199321490.001.0001

Wierzbicka, A. (2018). Deconstructing "colour", exploring indigenous meanings. In D. Young (Ed.), *Rematerializing Colour*: *From concept to substance* (pp. 67–90). Wantage, UK: Sean Kingston.

Wong, J. O. (2014). *The Culture of Singapore English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139519519

Zimmermann, M., Levisen, C., van Scherpenberg, C., & Beck, T. G. (2015). Please pass me the skin colored crayon! Semantics, socialization, and folk models of race in contemporary Europe. *Language Sciences*, 49, 35–50. https://doi.org/10.1016/j.langsci.2014.07.011

Appendix. Explications in Danish NSM

[A] Semantic explication for Danish *rød*
(noget Z er *rødt*)
a. folk kan tænke sådan om Z's farve
"den her farve er som farven på blod"
b. samtidigt kan folk tænke sådan:
"når folk ser ild kan de sommetider se den her farve"

[B] Semantic explication for Danish *blå*
(noget Z er *blåt*)
a. folk kan tænke sådan om Z's farve:
"himlen er tit den her farve om dagen"
b. samtidigt kan folk tænke sådan:
"havet kan være den her farve"

[C] Semantic explication for Danish *grøn*
(noget Z er *grønt*)
a. folk kan tænke sådan om Z's farve:
"mange ting der vokser i jorden kan være den her farve"

[D] Semantic explication for English *light* (in relation to color)
(The color of thing Z is *light*)
a. når folk ser den her den her ting Z, kan de tænke sådan om Z's farve:
"noget kan se sådan ud om dagen
hvis solen skinner på den her ting"

[E] Semantic explication for English *bright* (in relation to color)
(the color of Z is *bright*)
a. når folk ser den her ting Z kan de tænke sådan om Z's color:
"Z's farve skinner
som når solen skinner om dagen"
b. derfor er det sådan: hvis folk er tæt på den her ting Z, kan de ikke undgå* at se den her farve
c. folk kan føle noget godt når de ser sådan en farve
(**undgå* is an allolex of *ikke* in the double negotiation construction *ikke* + *ikke*)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

[F] Semantic explication for *Danish lys* (in relation to *farve*)
(Z's farve er *lys*)
a. når folk ser den her ting Z, kan de tænke sådan om Z's farve:
"noget kan se sådan ud om dagen,
hvis solen skinner på den her ting"
b. folk kan føle noget godt når de ser sådan en farve

[G] Semantic explication for the '*lyse* + color' compound
(Z er *lyselilla*, etc.)
a. når folk ser Z's farve kan de tænke sådan:
"Z's farve er én slags *lilla*"
samtidigt kan de her folk tænke sådan:
b. "den her slags *lilla* er en lys *lilla*"

[H] Semantic explication for the '*lys* + color' compound
(Z is *lyslilla*, etc.)
a. når folk ser Z's farve kan de tænke sådan:
"Z's farve er én slags *lilla*"
b. samtidigt kan de her folk tænke sådan:
"den her slags *lilla* er en *lys lilla*"
[hvis folk ved mange ting om farver vil de, når de siger noget om farver sige det med det her ord (*lyslilla*, etc.), ikke med et andet ord]

[J] Semantic explication for *lyserød*
(Z is lyserød)
a. når folk ser Z's farve kan de tænke sådan:
"det er én slags farve"
b. samtidigt kan de tænke sådan:
"den ligner noget rødt, den ligner noget hvidt"
c. mange små piger kan føle noget godt når de ser den her farve
mange små piger kan føle noget godt når de tænker på den her farve

[K] Semantic explication for *lysegrøn*
(Z er lysegrøn)
a. når folk ser Z's farve kan de tænke sådan:
"den her farve er én slags grøn"
den er som farven på træers blade om foråret, når solen skinner på på bladene om dagen"
c. når folk ser den her farve kan de føle noget godt

[L] Semantic explication for *lyseblå*
(Z er lyseblå)
a. når folk ser Z's farve kan de tænke sådan:
"den her farve er én slags blå
den er som himlens farve om dagen,
når solen er på himlen"
c. når folk ser den her farve kan de føle noget godt

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lexicalization patterns in color naming in Korean

Seongha Rhee
Hankuk University of Foreign Studies

Korean has a large inventory of color terms lexicalized through diverse strategies. Sound symbolism is one of the most determinative factors. Sound symbolism operative in color naming includes vowel polarity, consonant tensing and aspiration, prefixation of intensifiers, phonetic extension, and reduplicative suffixation, through which diverse aspects of color perception, e.g., intensity, luminosity, extent, solidity and pleasurability, are encoded. Color naming patterns in Korean reveals an iconic relationship between the perceived state of affairs in the world and the linguistic forms in an intricate application of elaborate sound symbolism, which suggests synesthesia with other related perceptual domains such as auditory perception, distance perception, opacity perception, and value judgment.

Keywords: color lexicalization, sound symbolism, synesthesia, vowel polarity, consonant tensing, intensity, luminosity, extent, solidity, pleasurability

1. Introduction

Considering the sheer number of color terms found in Korean dictionaries, Korean seems to have one of the richest color-term inventories across languages. Lexicographers list a large number of native terms expressing colors across nominal, adjectival and adverbial categories. For instance, Park (1989) lists 362 adjectives/verbs and 98 adverbs that describe colors and hues, and Nam (1992) lists 41 nouns that denote color names. Hong (2015) analyzes 429 adjectives in the 'black', 'white', 'red', 'yellow', and 'grue' categories, and Kim (1994) analyzes as many as 560 color terms. The total number of adjectives in these five color categories in an authoritative dictionary, *Phyocwun Kwuke Taysacen,* by the National Institute of the Korean Language (1999), is 752, inclusive of 'standard' and 'non-standard' forms (Hong, 2015, p. 13). The size of the inventory seems extraordinary, but more

https://doi.org/10.1075/sfsl.78.06rhe
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

intriguingly, the listing is not exhaustive, since the inventory is not self-contained but is systematically expandable, as shall be made clear in the following discussion.[1]

In contemporary Korean, there are five focal colors in the category of nouns, and in terms of color typology, Korean belongs to the 'grue' language, i.e., a single native term denoting either green or blue. These color nouns are in fact derived terms with the nominalizing suffix *-ang/-eng* (allophonic variants per vowel harmony), as shown in (1):[2]

(1) a. 'black': *kkamang* (까망), *kemceng* (검정)
b. 'white': *hayang* (하양)
c. 'yellow': *nolang* (노랑), *nwuleng* (누렁)
d. 'red': *ppalkang* (빨강), *kemyang* (거먕)[3]
e. 'grue': *phalang* (파랑)

The nominalizing suffix *-ang/-eng* seems to have originated from the noun *aki* 'baby' as a diminutive suffix (Rhee, 2001a, 2001b). The fact that the focal color nouns are suffixed forms suggests that the corresponding adjectival forms are primary color terms. However, it is not clear from the available sources whether it is the nominal forms or the adjectival forms that are historically primary. Some historical linguists hypothesize object names as the etymological sources of certain color terms, e.g., 'white' from *hAy* (해) 'sun', 'red' from *pul* (블) 'fire', 'grue' from *phul* (플) 'grass', 'black' from **ket* (걷) 'fire/smoke' (Yoo, 1978, pp. 357–362; Suh, 2003[2000], p. 42). Since an investigation into historical development is beyond the scope of this research, our discussion will focus on the synchronic state of affairs.

---

1. In addition to the complex derivational strategies discussed in this research, there are other strategies such as compounding of two or more color terms into single terms to denote mixed colors. This strategy is very productive and warrants more in-depth investigation, but it is not addressed here (see Lee, 2006; Koo, 2009, among others).

2. Korean words are presented in *Hangeul*, the Korean writing system (simplified, if involving defunct characters), and transliterations following the Yale Romanization System (Martin, 1992) and the Extended Yale Romanization System (Rhee, 1996), presented in italics. When the sound values are focused, the words are given either in phonemic representation between virgules (sometimes simplified) or in phonetic representation between square brackets (also sometimes simplified), as deemed appropriate in the context of discussion. The conventions of phonemic notation largely follow Shin et al. (2013) and Shin (2015).

3. *Kemyang* is a native term for crimson red, but in contemporary Korean it is recognized as such by very few.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. Preliminaries: Sound symbolism

In order to look into the system of color lexicalization we need to first examine the sound symbolism since it is very productive in Korean (Jung, 1938; Kim, 1976; Kim-Renaud, 1976; Martin, 1962; Lee, 1993; Sohn, 1999; Koo, 2007; Larsen & Heinz, 2012; Kwon, 2018) and is systematically reflected in color lexicalization. Among many aspects of sound symbolism, we will look at vowel polarity, consonant tensing and aspiration, and reduplication, since these are deeply involved in color lexicalization.

2.1 Vowel polarity

Korean has an elaborate system of sound symbolism for vowels.[4] The most prominent aspect of vowel sound symbolism is based on the notion of vowel polarity, i.e., 'positive', 'negative' and 'neutral' vowels.[5] The membership of individual sounds in these polarized categories shows minor variations by researchers, but a version of the vowel chart based on the oral space can be shown as in Figure 1.

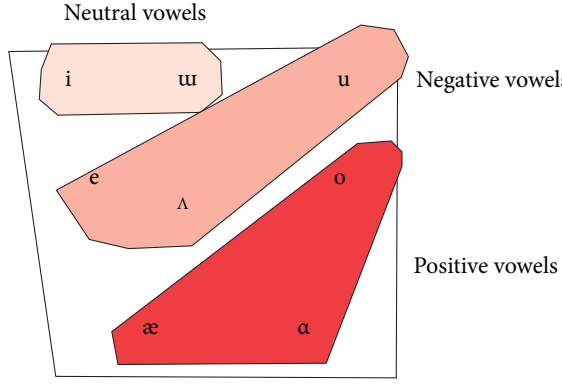


Figure 1. Vowel polarity in Korean

4. For a general introduction to Korean vowel system see Lee (1993, pp. 58–63), Chang (1996, pp. 11–13), Sohn (1999, pp. 156–159), Song (2005, pp. 29–32), Yeon & Brown (2011, pp. 12–15), Shin et al. (2013, pp. 95–116), Shin (2015, pp. 5–6), among others.

5. The 'positive' and 'negative' vowels are often referred to as 'light' and 'dark' vowels (e.g., Martin, 1962; Kim-Renaud, 1976; McCarthy, 1983; Lee, 1993; Sohn, 1999; Kwon, 2018) or as 'yang' and 'yin' vowels, respectively (Sohn, 1999), and the category membership is debated (see Martin, 1992; Lee, 1992; Larsen & Heinz, 2012; Kroeger, 2016). The systematic alternation of vowels for vowel harmony is also called 'isotopes' in Martin (1992) and Chang (1996), among others.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The function of the vowel polarity is to encode certain aspects of an object or event being described in terms of appearance, manner, etc. The polarity modulation is pervasive in language use in Korean and usually involves ideophones that describe the sound or appearance and color terms that describe diverse aspects of color sensation. There are also other areas of grammar to which it applies. The meaningful contrasts of polarity are generally between positive and negative vowels, as briefly summarized in (3):

(3) a. 'positive' vowels: /ɑ/, /o/, /æ/, and diphthongs involving them; describe small, delicate, and bright objects, and movements arousing such sensation
b. 'negative' vowels: /ʌ/, /e/, /u/, and diphthongs involving them; describe big, crude, and dark objects, and movements arousing such sensation

Since there are multiple vowels in vowel polarity classification, the contrasts can be created with many different combinations. In order to present the general idea of vowel polarity contrast, two types are exemplified with some possible instances of such use in (4):

(4) a. Ideophone examples for the 'positive' /ɑ/ vs. 'negative' /ʌ/ contrast
(i) /kɑntɯlkɑntɯl/ (간들간들):
e.g., a small flower swaying in the breeze
/kʌntɯlkʌntɯl/ (건들건들):
e.g., a tall tree being shaken in the wind
(ii) /sɑkɑksɑkɑk/ (사각사각):
e.g., sand grains crackling under feet
/sʌkʌksʌkʌk/ (서걱서걱):
e.g., dried bamboo leaves bustling in the wind
b. Ideophone examples for the 'positive' /o/ vs. 'negative' /u/ contrast
(i) /k*okitk*okit/ (꼬깃꼬깃):
e.g., a piece of paper crumpled with many tiny wrinkles
/k*ukitk*ukit/ (꾸깃꾸깃):
e.g., a shirt with big, long wrinkles
(ii) /moŋkɯlmoŋkɯl/ (몽글몽글):
e.g., furry patches of a cuddly rabbit
/muŋkɯlmuŋkɯl/ (뭉글뭉글):
e.g., clouds growing into giant columns

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.2 Consonant tensing and aspiration

Another aspect of sound symbolism involves consonant tensing and aspiration, an iconic representation of force dynamics in linguistic coding. For instance, forces involved in tensing and aspiration in articulatory gestures are systematically represented in the semanticization of word derivation and new word coinage (Koo, 2007, pp. 201–202). In particular, Korean consonantal system has a tripartite contrast among the plain, aspirated, and tensed counterparts.[6] This contrast is also well represented in the ideophones, as exemplified in (5) and (6):

(5) Ideophone examples for the plain /p/ vs. tensed /p*/ contrast
   a. /pokɯlpokɯl/ (보글보글):
      e.g., water boiling in a small pot
   b. /p*okɯlp*okɯl/ (뽀글뽀글):
      e.g., thick stew boiling in a pot (modified from Koo, 2007, p. 201)

(6) Ideophone examples for the plain /tɕ/ vs. aspirated /tɕʰ/ contrast
   a. /tɕolaŋtɕolaŋ/ (졸랑졸랑):
      e.g., a colt walking behind its mother
   b. /tɕʰolaŋtɕʰolaŋ/ (촐랑촐랑):
      e.g., a colt trotting along, behind its mother
      (modified from Koo, 2007, p. 202)

As shown in (5b), tensed consonants strongly suggest the presence of localized and intensified force causing friction among the elements being described, e.g., cut vegetables and meat pieces turning in the boiling stew. This friction encoding by means of a tensed counterpart exhibits the iconic relationship between articulatory gesture (i.e., making the muscles of the articulators tensed, contra 'laxed') and the state of affairs (presence of a strong force in friction among event participants). Likewise, in (6b), the aspirated consonants strongly suggest the presence of a strongly exerted but diffused force as, for example, the colt trotting with its limbs and tail rhythmically moving beyond some reasonable operational space. This is a direct reflection of the articulatory gesture involved in aspiration, i.e., delayed onset of voicing with a strong release of air puffs, as opposed to the 'unaspirated' or 'unreleased'.[7]

---

6. This tripartite contrast is orthographically represented in Korean writing system, *Hangeul*. For instance, k-kh-kk /k-kʰ-k*/ (ㄱ-ㅋ-ㄲ), t-th-tt /t-tʰ-t*/ (ㄷ-ㅌ-ㄸ), p-ph-pp /p-pʰ-p*/ (ㅂ-ㅍ-ㅃ), c-ch-cc / tɕ-tɕʰ-tɕ*/ (ㅈ-ㅊ-ㅉ), etc., are formed through a systemic operation, i.e., adding a stroke for the aspirated counterpart, and doubling the plain character for the tensed counterpart. For this reason, some grammatologists classify Korean *Hangeul* as a 'featural system' (Sampson, 1985) in addition to being an alphabetical system.

7. Syllable-final consonants should be pronounced as unreleased sounds after central closure in Korean (see Shin et al., 2013, p. 181).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.3 Reduplication: Iconicity of multiplicity

Korean onomatopoeic words often involve phonological (total/partial) reduplication. Reduplication is typically used in describing the multiple occurrences of sound or certain aspects of an object or motion appearing in repetition. The following examples show the operation of reduplication and its semantic effect:

(7) a. /pʰuŋtʌŋ/ (풍덩)
e.g., a big object falling into deep water
b. /pʰuŋtʌŋpʰuŋtʌŋ/ (풍덩풍덩)
e.g., big objects falling into deep water in succession
c. /pʰuŋt͡ʌtʌŋ/ (풍더덩)
e.g., a big object with an irregular surface falling into deep water, thus creating multiple splashes

As shown in (7), /pʰuŋtʌŋ/ describes a single occurrence of a big object, e.g., a big rock, thrown into water is making a splash upon contact with the surface. The native speaker's intuition is such that the word describes a single event of a splash made by a bulky object even though the word involves two sub-events, i.e., the forced opening of the water surface described with /pʰuŋ/ and the closing of the broken surface described with /tʌŋ/. This onomatopoeic word is a good descriptor of a skilled diver falling into water in a streamlined fashion. The second one /pʰuŋtʌŋpʰuŋtʌŋ/ is a full reduplication of the first. The instance describable with this word may be two (or more) divers falling into the water in succession at short intervals, or one person without swimming skills floundering in the water, thus making splashes in succession. The word describes an event with multiple sub-events emitting the same sound each time. The last one /pʰuŋtʌtʌŋ/ is an instance of partial reduplication, i.e., /tʌ/ is reduplicated from /tʌŋ/. The event which it describes might be the falling of a large object with an irregular surface, such as a chair, into deep water. Since the object has an irregular surface, its entry into water may be a bit prolonged and its contact with the water's surface, if viewed microscopically, may occur as multiple sub-events as well. This word may also describe someone falling into water not in a streamlined fashion but with stretched limbs.

In this section, we have briefly looked at three elements of sound symbolism, vowel polarity, consonant tensing and aspiration, and reduplication. As will be seen in the following discussion, these elements play important roles in semanticizing the various aspects of color perception and evaluation.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. Lexicalization patterns

The lexicalization patterns of the color terms in Korean differ according to their grammatical categories. We will look at the color terms in the nominal, adverbial/adjectival, and verbal categories.

3.1 Nominal category

The nominals in the color lexicon may be further divided into several subcategories based on their source characteristics. The first is a set of focal color terms that Koreans treat as basic native terms, as illustrated in the previously mentioned example (1). Analytically, these terms are derived forms as they carry the nominalizing suffix *-ang/-eng*.

The second set is a group of words which are lexicalized forms from the morpho-syntactic operation. These words are built on an adjective, followed by an adnominalizer and a noun which denotes 'color'. Incidentally, the reason the adjectives are combined with an adnominalizer is that adjectives in Korean inflect, just like verbs, and thus are often treated as stative verbs. True adjectives, those that can directly modify a noun, are very small in number. Therefore, color adjectives need to be headed by an adnominalizer (ADN) in order to modify a noun. The noun recruited in this operation is a Sino-Korean word *sayk* (색) 'color'. The morphosyntactic fusion of the participating forms in this derivation has proceeded to such an extent that native speakers of contemporary Korean largely regard the final product as a single word rather than a phrase with analyzable morphosyntactic components. Color nominals thus derived are as shown in (8) with their source constructions, glosses, and the final products, i.e., color nominals denoting 'x color' or 'color of x':

(8) a. 'black': *kkamah-n-sayk* ('be.black'-ADN-'color') > *kkamansayk* (까만색)
b. 'white': *hayah-n-sayk* ('be.white'-ADN-'color') > *hayansayk* (하얀색)
c. 'yellow': *nolah-n-sayk* ('be.yellow'-ADN-'color') > *nolansayk* (노란색)
d. 'red': *ppalkah-n-sayk* ('be.red'-ADN-'color') > *ppalkansayk* (빨간색)
e. 'grue': *phalah-n-sayk* ('be.grue'-ADN-'color') > *phalansayk* (파란색)

Another subcategory of color nominals is those derived from a compounding operation. The derivation involves the name of a natural object, whose characteristic color is representative of the color category, and the native Korean word *pich* (빛), denoting 'light', as exemplified in (9). Since this compounding is either an asyndetic combination of the object name and *pich*, as in (9a) through (9c), or one with a genitive (GEN) marker *-s* between the two, as in (9d) through (9g), there is no complex

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

syntactic operation involved.[8] As there are many natural objects with characteristic colors, this compounding, in principle, may produce many color nominals. But the color nominals in this category that have gained wide circulation are limited in number. Some of these examples are as shown below:

(9) a. 'yellow': *kum-pich* ('gold'-'light') > *kumpich* (금빛)
b. 'green': *phwul-pich* ('grass'-'light') > *phwulpich* (풀빛)
c. 'blue': *hanul-pich* ('sky'-'light') > *hanulpich* (하늘빛)
d. 'crimson': *phi-s-pich* ('blood'-GEN-'light') > *phispich* (핏빛)
e. 'reddish brown': *kwuli-s-pich* ('copper'-GEN-'light') > *kwulispich* (구릿빛)
f. 'rosy': *cangmi-s-pich* ('rose'-GEN-'light') > *cangmispich* (장밋빛)
g. 'milky': *wuyu-s-pich* ('milk'-GEN-'light') > *wuyuspich* (우윳빛)

Another category of color nominals is the borrowings from foreign languages. Due to the extensive contact with English, Koreans occasionally use English color terms (Tyson, 1994), such as *pullayk* (블랙, black), *hwaithu* (화이트, white), *pullwu* (블루, blue), *oleynci* (오렌지, orange), *peyici* (베이지, beige), *koltu* (골드, gold), *phel* (펄, pearly white), etc., but their use is usually genre-specific, e.g., in descriptions of cosmetics, cars, furniture, etc. These have not yet acquired the full membership in the Korean color lexicon and they are regarded as borrowings and treated as such. However, there are other borrowings which, though small in number, have a longer history and are thus considered as 'regular' color terms by contemporary speakers of Korean. They are as listed, in part, in (10)[9]:

(10) a. 'crimson red': *tahong* (다홍) (< Chinese *duōhóng* (多紅) 'much.red')
b. 'purple': *pola* (보라) (< Mid. Mongolian *boro* 'purple') (Baek, 2012)
c. 'purple': *cacwu* (자주) (< Chinese *zīzhū* (紫朱) 'purple.vermilion')
d. 'dark blue': *yacheng* (야청) (< Chinese *yǎqīng* (雅青) 'elegant.grue'; Manchu *yacin*)

8. An anonymous reviewer questions the status of *-s* as a genitive marker and suggests the possibility of the so-called *sai-sios*, an epenthetic linker used in certain compounding, which is indeed plausible. However, many scholars (e.g., Ramstedt, 1939; Ahn, 1968, Kwon, 2001) hypothesize that the epenthetic *-s* is historically related to the genitive marker (for more recent discussion see Kim, 2017).

9. Borrowing from Chinese has been extensive, and one subcategory of color nominals from Chinese consists of a Chinese color term compounded with the Sino-Korean word *sayk* 'color', as in *huksayk* 'black-color', *payksayk* 'white-color', *chengsayk* 'blue-color', *noksayk* 'green-color', *ceksayk* 'red-color', *hwangsayk* 'yellow-color', etc. In Chinese the first syllable of these words may constitute a color term by itself. Compounding with this *sayk* 'color' is also found in other nominals such as *kumsayk* 'gold-color', *unsayk* 'silver-color', *oksayk* 'jade-color', *catwusayk* 'plum-color', etc.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.2 Adjectival/adverbial categories

Nominal color terms seem to be less than impressive since the same or similar lexicalization patterns are widely attested across languages. The lexicalization patterns in the adjectival/adverbial categories, however, present a very different picture in terms of their elaborateness and systematicity. There are diverse strategies involved in the lexicalization of color adjectives/adverbs.

3.2.1 *Prefixation*

The first morphological operation for the derivation of color adjectives/adverbs is prefixation. This operation changes the color semantics with respect to the level of saturation. The prefixes recruited for that purpose are *say-* /sæ-/, *says-* /sæt-/, *si-* /si-/, and *sis-* /sit/.[10] When these prefixes are affixed to a color term, the color denoted by the term becomes of a higher saturation, i.e., one with a stronger and often darker hue (except for the 'white' words). This is illustrated with the examples in (11):

(11) a. Plain: *phalah-* (파랗-) [pʰɑrɑt-] 'be grue (blue/green)'
Prefixed: ***says****phalah-* (샛파랗-) [sæppʰɑrɑt-] 'be strongly grue'
b. Plain: *ppalkah-* (빨갛-) [p*algɑt-] 'be red'
Prefixed: ***say****ppalkah-* (새빨갛-) [sæp*algɑt-]'be very red'
c. Plain: *hayah-* (하얗-) [hɑjɑt-] 'be white'
Prefixed: sayhayah- (새하얗-)[sæhɑjɑt-] 'be very white'
d. Plain: *nwuleh-* (누렇-) [nurʌt-] 'be yellow'
Prefixed: sisnwuleh- (싯누렇-)[ʃinnurʌt-] 'be very yellow'
e. Plain: *kkemeh-* (꺼멓-) [k*ʌmʌt-] 'be black'[11]
Prefixed: ***si****kkemeh-* (시꺼멓-) [ʃik*ʌmʌt-] 'be very black'

3.2.2 *Vowel polarity modulation*

Another linguistic device for modulation of color is the use of vowel polarity. As was noted in 2.1, vowel polarity modulates the sensory perception with respect to size, brightness, and the emotion caused by the perception event. Thus, the vowel polarity modulates the luminosity of the color being described. The use of positive

---

10. The choice of these prefixes also depends on vowel harmony, i.e., /i/ for words containing negative vowels and /æ/ for words containing positive vowels. The origin of these intensifying prefixes, however, is unknown. The distribution of these prefixes with and without coda, e.g., /sæt/ vs. /sæ/, does not seem to be phonologically controlled, since same words may be prefixed with either one of them, e.g., *saysppalkah-* and *sayppalkah-*, both for 'be very red'.

11. As shall be discussed in 3.2.3, *kkemeh-* is more strongly black than *kem-* by virtue of having a tensed consonant, which signals deeper saturation. This shows that even the intensity of a color hue can be expressed as having multiple levels.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

vowels, i.e., /a/, /o/, /æ/, etc., results in denoting the color as having bright hue, whereas the use of negative vowels, i.e., /ʌ/, /e/, /u/, etc., suggests that the color has a darker hue. The brightness modulation is directly related to the evaluation, i.e., the speaker's evaluative attitude toward the object (see 4.3). This contrast is illustrated in the following examples:

(12) a. Positive /o/: *nolah-* (노랗-)[norɑt-] 'be (bright) yellow'
Negative /u/: *nwuleh-* (누렇-)[nurʌt-] 'be dark yellow'
b. Positive /a/: *ppalkah-* (빨갛-) [p*ɑlgɑt-] 'be (bright) red'
Negative /ə/: *ppelkeh-* (뻘겋-) [p*ʌlgʌt-] 'be dark red'
c. Positive /a/: *phalah-* (파랗-) [phɑrɑt-] 'be (bright) grue'
Negative /ə/: *pheleh-* (퍼렇-) [phʌrʌt-] 'be dark grue'

3.2.3 *Consonant tensing/de-tensing*

Just as the vowels can be modulated, so can the consonants, with tensification as a prominent strategy. When a consonant is tensed, the color term carries the meaning that the hue is saturated, i.e., the color is intense. The color term with a non-tensed consonant signals weaker hue. The direction of tensification can be upward or downward, i.e., if the base form contains a non-tensed consonant, tensification will increase the intensity of the color hue, as in (13a), and if the base form already contains consonants that are tense, de-tensification will decrease the intensity, as in (13b). This is exemplified in the following:

(13) a. Plain (base form): *kem-* (검-) [kʌm-] 'be black'
Tensed: *kkem-* (껌-) [k*ʌm-] 'be dark black, pitch black'
b. Tensed (base form): *ppalkah-* (빨갛-) [p*ɑlgɑt-] 'be red'
De-tensed: *palkah-* (발갛-) [pɑlgɑt-] 'be reddish'

3.2.4 *Reduplication*

Another linguistic device to modulate color description is reduplication. Reduplication in color lexicalization makes reference to the distribution pattern of the color hue. In other words, color terms containing reduplication signal multiple occurrences of saturated coloration over a space of a colored object, i.e., uneven distribution, thus making the colored object look 'spotty' due to the lack of coloration constancy. This is illustrated in the following examples:

(14) a. Plain: *pwulk-* (붉-) [pulk-] adj. 'be red'
Redup.: *pwulkus* ***pwulkus***(불긋불긋) [pulgɯtpulgɯt] adv. 'reddish here and there'
b. Plain: *nolah-* (노랗-) [norɑt-] adj. 'be yellow'
Redup.: *nolus* ***nolus*** (노릇노릇) [norɯnnorɯt] adv. 'yellowish here and there'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.2.5 *Phonetic extension*

The last linguistic device for the modulation of color description is phonetic extension by means of suffixation. There are three subtypes within this derivational strategy. The first one involves phonetic extension by using special suffixes such as *-ukkeyha-* [-ɯkˀeha-], *-uleyha-* [-ɯreha-], *-usulumha-* [-ɯsɯrɯmha-], *-usuleyha-* [-ɯsɯreha-], *-ukkulumha-* [-ɯkˀɯrɯmha-], *-umwuleyha-* [-ɯmureha-], *-ukkwuleyha-* [-ɯkˀureha-], etc. When these suffixes are affixed to a color word, thus phonetically extended, the sense of extended surface and/or subdued hue is invoked. In other words, this type of suffixation modulates the distributional patterns and the saturation. For this reason, objects with extremely small colored space (such as the tip of a pen, etc.), or objects with larger space covered with solid coats of paint (such as cars, etc.) cannot be described with these color terms. This is exemplified by the following:

(15) a. Plain: *nolah-* (노랗-) [norat-] 'be yellow'
Suffixed: *nol**ukkeyha*-** (노르께하-) [norɯkˀeha-] 'be slightly yellowish'
Suffixed: *nol**usulumha*-** (노르스름하-) [norɯsɯrɯmha-] 'be slightly yellowish'

b. Plain: *phalah-* (파랗-) [pʰarat-] 'be grue'
Suffixed: *phal**ukkeyha*-** (파르께하-) [pʰarɯkˀeha-] 'be slightly grue'
Suffixed: *phal**usulumha*-** (파르스름하-) [pʰarɯsɯrɯmha-] 'be slightly grue'

The second subtype involves phonetic extension by means of reduplicative suffixation. Reduplicative suffixation invokes the sense of extended surface, uneven distribution, spottiness, and the evaluation of the color as being less pleasurable. In other words, this derivational strategy modulates distributional patterns, saturation, and the perceiver's attitude toward the object being described. There are numerous suffixes for this operation, e.g., *-uswukswuk* [-ɯsuksuk], *-ucapcap* [-ɯʤaptʃap-], *-ucepcep* [-ɯʤʌptʃʌp], *-ucokcok* [-ɯʤoktʃok], *-ucwukcwuk* [-ɯʤuktʃuk], *-ucekcek* [-ɯʤʌktʃʌk], *-uchikchik* [-ɯtʃʰiktʃʰik], *-uchwukchwuk* [-ɯʤuktʃuk], *-uchukchuk* [-ɯtʃʰɯktʃʰɯk], *-uchwungchwung* [-ɯtʃʰuŋtʃʰuŋ], *-uteytey* [-ɯdede], *-utaytay* [-ɯdædæ], *-utweytwey* [-ɯdwedwe], *-uthweythwey* [-ɯtʰwetʰwe], *-uthwithwi* [-ɯtʰwitʰwi], *-uteyngteyng* [-ɯdeŋdeŋ], *-utingting* [-ɯdiŋdiŋ], *-utwungtwung* [-ɯduŋduŋ], *-uthwungthung* [-ɯtʰuŋtʰuŋ], etc. The suffixed examples are adverbs, which can be converted into adjectives by attaching the light verb *ha-* 'do' (see 3.3 below).

(16) a. Plain: *kem-* (검-) [kʌm-] adj. 'be black'
Suffixed: *kemwucwukcwuk* (거무죽죽) [kʌmuʤuktʃuk] adv. 'unevenly, unpleasingly black'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

b. Plain: *phwulu-* (푸르-) [pʰurɯ-] adj. 'be grue'
Suffixed: *phwuluthwungthwung* (푸르퉁퉁) [pʰurɯtʰuŋtʰuŋ] adv. 'unevenly, unpleasingly grue'

Yet another subtype of phonetic extension is the use of discontinuous reduplicative suffixes. When these suffixes are attached, the meaning changes with respect to the distribution pattern of coloration, i.e., the discontinuous reduplicative suffixation invokes the sense of extended surface and uneven distribution. There are a few suffixes used in this operation, e.g., *-wus -wus-* [-ut -ut-], *-us -us-* [-ɯt -ɯt-], *-swung -swung-* [-suŋ -suŋ-], *-cek -cek-* [-ʧʌk -ʧʌk-], *-sil -sil* [-ʃil -ʃil-], etc. This pattern is illustrated in the following examples:

(17) a. Plain: *kkam-* (깜-) [kˀam-] adj. 'be strongly black'
Suffixed: *kkamwuskkamwus* (까뭇까뭇) [kˀamutkˀamut] adv. 'spotty black'
b. Plain: *nolah-* (노랗-) [norɑt-] adj. 'be yellow'
Suffixed: *nolusnolus* (노릇노릇) [norɯnnorɯt] adv. 'spotty yellow'

3.3 Verbal category

In terms of the grammatical categories of color naming, nouns, adjectives, adverbs, and verbs are all closely related. Nouns, as exemplified in 3.1, are either focal color names (which still contain the nominalizing suffix *-ang/-eng*), or univerbated forms from syntactic constructions involving a noun *sayk* 'color', or compounding with *pich* 'light', or borrowings from foreign languages. Adjectives often end with *-ha-*, which is the light verb *ha-* 'do/be/say', and some adjectives end with *-h-*, which is the historical vestige of the light verb. There are other adjectives that do not contain either *-ha-* or *-h-*. Most color terms that do not contain *-ha-* or *-h-* are the words to which suffixes are attached (as shown in 3.2 above), and these are classified as adverbs. As briefly noted, all these forms derived from suffixation can be converted into adjectives by adding the light verb *ha-* 'do'.[12]

Color terms in the (genuine) verbal category are those that denote inchoation. These are derived by way of inchoative suffixes *-eci-* or *-keytoy-*, largely translatable as 'become color x'. Incidentally, *-e* (in *-eci*) is a non-finite marker commonly used for verb serialization with the 'consolidation' meaning (Koo, 1987; Rhee, 1996), and *-key* (in *-keytoy-*) is a non-finite marker with the 'mode' meaning (Rhee, 1996).[13]

12. Many adjectives in Korean are also classifiable as stative verbs since they carry much of the verbal characteristics including tense inflection.

13. The verb-category membership of these color terms is primarily due to the morpho-syntactic operations involving the verbs *ci-* (< 'fall') and *toy-* 'become' and linkers *-e* and *-key* (see Rhee, 1996; Rhee & Koo, 2014).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Both inchoative markers are grammaticalized markers as a result of the serialization of a verb/adjective with (another) verb, but their lexicalization strengths or the levels of constructional entrenchment are different, i.e., the *-eci-* derived forms are regarded as more fully lexicalized (as single forms) than the *-keytoy-* derived forms. One of the manifestations of such a variable perception is reflected in that most Korean speakers rarely use inter-lexical spacing for *-eci-* forms, whereas they often do for *-keytoy-* forms (e.g., *kemeci-* vs. *kemkey toy-*, both denoting 'become black').

In short, all adverbs can be converted into an adjective with a light verb, which, in turn, can be converted into a verb of inchoative meaning. The interrelatedness among the grammatical categories can be illustrated with the following examples of 'black' color words (note that there are many other forms available through diverse derivational operations described above), with a brief note at the end, in which (i) through (l) are derived verbs:

(18) a. *kkamang* 'black' n. focal color name with *-ang*
b. *kkamansayk* 'black' n. *sayk* 'color'
c. *kem-* 'be black' adj. Focal color adjective
d. *kkem-* 'be dark black' adj. Intensified with tensing
e. *kkemeh-* 'be dark black' adj. *-h-* (< the light verb *ha-*)
f. *kkemwusha-* 'be blackish' adj. *-ha-* (light verb)
g. *kkemwuskkemwus* 'spotty black' adv. Reduplicative suffix
h. *kkemwuskkemwusha-* 'be spotty black' adj. *-ha-* (light verb)
i. *kemeci-* 'become black' v. *-eci-* inchoative
j. *kkemwuskkemwushayci-* 'become spotty black' v. *-eci-* inchoative
k. *kemkeytoy-* 'become black' v. *-keytoy-* inchoative
l. *kkemehkeytoy-* 'become very black' v. *-keytoy-* inchoative

4. Semantics-morphology interface

Now that we have seen the diverse mechanisms involved in color lexicalization, we turn to a discussion of significant aspects exhibited by color lexicalization patterns. Due to potentially numerous issues, we choose four of them, i.e., iconicity, synesthesia, attitude, and productivity, since these comprise noteworthy aspects in color lexicalization in Korean.

4.1 Iconicity

As shown in the previous discussion, color lexicalization patterns in Korean attest to an iconic relationship between the perceived state of affairs in the world, i.e., color perception, and the linguistic forms, i.e., color terms, through the application of

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

elaborate sound symbolism. Color terms are modulated with respect to the color's saturation (intensity), luminosity (brightness), the extent of distribution (coverage), solidity (constancy), and pleasurability (emotion). These color variables can be summarized as follows:

(19) a. saturation (intensity): intensifying prefixes (stronger), tensed consonants (stronger)
b. luminosity: positive vowels (brighter)
c. extent: reduplicative suffixes (larger covered area)
d. solidity: reduplicative suffixes (uneven, spotty)
e. pleasurability: reduplicative suffixes (less pleasurable), positive vowels (more pleasurable)

4.2 Synesthesia of multiple senses

The second issue is about synesthesia, which is considered an extraordinary perceptual phenomenon from a neurological point of view (Safran & Sanda, 2015, p. 36; see also Cytowic, 1989). There are examples of commonly used synesthetic expressions such as English *cool dress, hot girl, loud color, cold words, sweet voice,* etc. However, unlike these rather isolated and local instances of synesthetic operations, color term lexicalization in Korean borders on, and sometimes blends into, other related perceptual domains such as auditory perception, distance perception, opacity perception, etc., in a more systematic way.[14] For instance, the relation between the visual perception (color perception) and auditory perception (sound perception) is prominent. This is well illustrated by the fact that many of the suffixes are sound-onomatopoeic, suggesting that the extended color distribution is similarly conceived as the sound ringing out and lingering in the air. Some of these ideophones, i.e., those that can be used in the description of color as well as sound, are exemplified in (20):

(20) Ideophones of sound perception
a. *-usulum* [-ɯsɯrɯm]: sound coming from minimal friction of a gliding motion (e.g., of a sliding door smoothly opening along a rail)
b. *tululu* [tɯrɯrɯ]: sound coming from a round or lubricated object rolling on a smooth surface (e.g., of a bead rolling on a platter)

14. For a discussion on synesthesia in taste terms in Korean, see Rhee and Koo (2017) who illustrate the lexicalization of gustatory terms intricately and nearly inseparably related to olfactory, tactile, visual and auditory terms.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

c. *chikchik* [ʧʰikʧʰik]: sound of a wet stubby object brushing on a surface repeatedly (e.g., of a paint brush on canvas, or of an aerosol paint sprayed on canvas)
d. *swuk* [suk]: sound of an object moving into an object with minimal friction for certain depth (e.g., a well-sharpened knife driven deep into a pumpkin)
e. *teyngteyng* [teŋdeŋ]: sound of a metal object hitting another metal object (e.g., a metal club hitting a bell)
f. *cwukcwuk* [ʧukʧuk]: sound of a sharp object scratching on a surface (e.g., a chalk drawing long lines on a blackboard)
g. *thwungthwung* [tʰuŋtʰuŋ]: sound of a wooden container hit by an object (e.g., an oak-barrel hit by a stick)

Another type of synesthesia relates to color and distance perception. The relation seems to be based on the fact that the color perception is strongly influenced by the distance, i.e., if something is far off, its color perception acuity is reduced to the point of perceiving it black or dim. Some of these cases are illustrated in (21):

(21) Color & distance perception

a.	*kkam-* [k*ɑm-]	(i) 'be black'
	kkamatukha-[k*ɑmɑdɯkhɑ-]	(ii) 'be faraway, hazy'
		(cf. *atukha-* 'faraway')
b.	*asulaha-* [ɑsɯrɑhɑ-]	(i) 'dim, gray, shadowy'
		(ii) 'be far off'

There is yet another type of synesthesia that relates to color and opacity/thickness perception. The thickness may be of depth (for solids) or concentration (for liquids).

(22) Color & opacity/thickness perception

a.	*pwuyeh-* [puyʌt-]	(i) 'be whitish, milky'
		(ii) 'be thick, turbid, impure'
b.	*huypwuyeh-* [hibuyʌt-]	(i) 'be milky white'
		(ii) 'be murky'
c.	*malk-* [mɑlk-]	(i) 'be clear, clean, lucid, bright'
		(ii) 'be limpid, pure, transparent'
d.	*malkah-* [mɑlgɑt-]	(i) 'be clear, clean, lucid, bright'
		(ii) 'be limpid, thin (of liquid)'
e.	*haymalkah-* [hæmɑlgɑt-]	(i) 'be very white'
		(ii) 'be pure, uncontaminated'
f.	*yelp-* [jʌlp-]	(i) 'be light in intensity'
		(ii) 'be thin in thickness of solid object'
g.	*huymelkeh-* [himʌlgʌt-]	(i) 'be too weakly white in intensity'
	melkeh- [mʌlgʌt]	(ii) 'be too thin in thickness of liquid'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.3 Color and evaluation (attitude)

As has been noted in the previous exposition, color term formation incorporates the describer's value judgment on the object being described. Color terms derived from suffixation of certain suffixes, e.g., *-cwukcwukha-* [-ʧukʧukhɑ-], *-ukkiliha-* [-ɯkˀirihɑ-], *-teyngteyngha-* [-teŋdeŋhɑ-], *-kkwumuleyha-* [-kˀumɯrehɑ-], *-kkumuleyha-* [-kˀɯmɯrehɑ-], *-twungtwungha-* [-tuŋduŋhɑ-], *-thwungthwungha-* [-tʰungtʰunghɑ-], *-thepwunha-* [-tʰʌbunhɑ-], *-thepthepha-* [-tʰʌptʰʌphɑ-], etc., tend to carry negative connotation. The negativity is even more strongly signaled when these suffixes occur with negative-polarity vowels in the stem. Thus derived color terms signal the presence of an undesirable, unpleasurable quality of the coloration. For example, in the list of objects that can be described with variable color terms, the objects with the suffixed color terms are negatively evaluated by the speaker. If they are food names, they are not appetizing. While the non-suffixed color terms may not always be positive, even though they are by default, they are at least neutral. This is well illustrated, in part, by the following examples:

(23) a. Black objects
kkamah-: pupil (eye), cherry, bean, etc. (positive)
kkemwucwukcwukha-: irregularly tanned face, face with patches of freckles, discolored skin after contusion, etc. (negative)

b. Yellow objects
nolah-: melon, tulip, chrysanthemum, banana, etc. (positive)
nwulikkiliha-: discolored face with jaundice, discolored teeth, dust-covered land, etc. (negative)

c. Blue/green objects
phalah-: clear sky, ocean, grass, leaves, berries, etc. (positive)
phwulutwungtwungha-: blackened spot from contusion, decomposing body, skin affected by frostbite, unripe pumpkin, etc. (negative)

4.4 Productivity and novel coinage

Equipped with a derivation system, both rich and systematic, Korean has a paradigm of extraordinarily fine-grained color terms that defy faithful translation to other languages. As a matter of fact, there are many blog postings on the Internet that make queries about how to translate Korean color terms into other languages.[15]

15. Such blogs are commonly found in Korean Internet blog portals and query sites (e.g., https://blog.naver.com/, http://cafe.daum.net, http://www.ybmbooks.com/reader/question/, http://befly.yoons.com/, etc.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

This is problematic on both sides: for Koreans since their Korean color terms cannot be effectively translated, and for foreigners learning Korean as a foreign language, mastering the color lexicalization patterns with all variables that are absent in their first language, can be a formidable task.

An important aspect, however, is that even though the paradigm of color terms is overwhelmingly large in size, the speakers of the language can create new color terms and understand the color neologisms with ease since the paradigm exhibits a considerable degree of internal systematicity. In other words, since the sound symbolism has a robust representation in the minds of native speakers, when new color terms are encountered, they can conceive of the colors and hues denoted by the novel expressions without trouble. For instance, there are 127 color terms for the 'red' category, taken from the lexicons of Park (1989), Nam (1992), and elsewhere (see appendix). However, by applying the lexicalization mechanisms a nonce color term can be coined, e.g., *ppwulkumwuliha-* to mean, roughly, 'reddish'. Any native speaker can grasp the color quality of this newly coined adjective. This term can be readily analyzed by using the analytic variables discussed above as follows:

(24) *ppwulkumwuliha-* [p*ulgɯmuriha-] 'reddish'
  (i) red hue < base *pwulk-*
  (ii) dark hue < negative vowel *wu* [u]
  (iii) has a point of concentration < tensed consonant *pp* [p*]
  (iv) extended surface < suffix *-umwuli-*
  (v) displeasurable < suffix *-umwuli-* & negative vowel *wu* [u]

As shown above, the word *ppwulkumwuliha-* denotes dark red hue that is spread over some extended surface with darker spots scattered here and there, which the speaker finds displeasurable. At the time of writing this chapter, the Google search found two hits for this newly coined term. One was found in a blog where the blogger is describing a politician whom the blogger hates. The blogger is describing the ruddy face of the politician (incidentally, the cyber-policing authorities blocked access to the full text for the use of indecent and explicit language).[16] The other instance is a transcript of an orally transmitted story in an endangered Daegu dialect, as part of a language documentation project.[17] The speaker is narrating a story where a man, who failed a government civil service exam, describes the reddish color of the sky as the sun sets, evidently not in a good mood.

16. http://www.badkiller.kr/bbs/.

17. http://yoksa.aks.ac.kr/jsp/ur/.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

5. Summary and conclusion

The color lexicalization patterns in Korean show the following noteworthy aspects. The systematic operation of vowel polarity, consonantal tensification, and suffixation modulate the semantics of color terms with respect to saturation, luminosity, constancy, solidity, and pleasurability. A notable observation is that there exists an elaborate iconic relationship between the perceived states of affairs and linguistic coding via sound symbolism. Color lexicalization also exhibits synesthesia, since the perception of color borders with, and sometimes blends into, other perceptual domains such as auditory, distance and opacity perceptions. In addition, color lexicalization patterns incorporate value judgment as a semantic component, i.e., whether the speaker regards the colored object as pleasurable or unpleasurable. Furthermore, the derivational processes in color lexicalization are highly productive, and thus the color vocabulary is systematically expandable.

Acknowledgment

This research was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2014S1A5B1065578) and the Research Grant of Hankuk University of Foreign Studies. My special thanks go to So-Yeon Ahn and two anonymous reviewers for their insightful comments.

References

Ahn, B. -H. (1968). Cwungseykwuke sokkyekemi *s*-ey tayhaye [On the genitive *-s* in Middle Korean]. *Iswungnyengpaksa songswu kinyemnonchong* [Festschrift for Dr. Lee, Sung-Nyong] (pp. 3–11). Seoul: Eulyoo Moonhwasa.

Baek, M. -S. (2012). *Wulimal hyengthayso sacen* [Korean morphology dictionary]. Seoul: Pagijong Press.

Chang, S. -J. (1996). *Korean*. Amsterdam: John Benjamins. https://doi.org/10.1075/loall.4

Cytowic, R. E. (1989). *Synesthesia: A union of the senses*. New York: Springer-Verlag. https://doi.org/10.1007/978-1-4612-3542-2

Hong, S. -J. (2015). A lexico-morphological study on color adjectives in the Korean language. Ph.D. dissertation, Seoul National University, Korea.

Jung, I. S. (1938). Moum sangtay pepchikkwa caum kasey pepchik [Principles of vowel relativity and consonant reinforcement]. *Hangeul*, 60, 10–25.

Kim, J. -I. (2017). Some issues regarding *sai-siot* of 'N1+s+N2' in the 15th century Korean. *Journal of Korean Linguistics* 83, 183–219. https://doi.org/10.15811/jkl.2017.83.007.

Kim, K. -O. (1976). Sound symbolism in Korean. *Journal of Linguistics*, 13, 67–75. https://doi.org/10.1017/S0022226700005211

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kim, Y. -C. (1994). Kwukeuy koyuekyey saykchayeey kwanhan yenkwu [A study on native Korean color terms]. MA thesis, Kunsan National University, Korea.

Kim-Renaud, Y. -E. (1976). Semantic features in phonology: Evidence from vowel harmony in Korean. In S. A. Mufwene, C. A. Walker, & S. B. Steever (Eds.) *Papers from the Twelfth Regional Meeting of the Chicago Linguistic Society* (pp. 397–412). Chicago, IL: Chicago Linguistic Society.

Koo, B. -K. (2009). A study on meaning of color-complex words in Korean – concentrating on definitions in the dictionary. *Journal of Korean Linguistics*, 55, 173–211.

Koo, H. -J. (1987). Ssikkuth {*-a, -key, -ci, -ko*}-uy ssuimkwa uymi [The usage and semantics of suffixes, *-a, -key, -ci,* and *-ko*]. *Konkuk Emwunhak*, 11/*12*, 167–188.

Koo, H. -J. (2007). Force dynamics as a variational factor: A case in Korean. *LACUS Forum*, 33, 201–210.

Kroeger, R. (2016). Folk definitions of Korean ideophones. *San Diego Linguistic Papers*, 6, 66–121.

Kwon, N. (2018). Iconicity correlated with vowel harmony in Korean ideophones. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 9(*1*), 1–18. https://doi.org/10.5334/labphon.53.

Kwon, Y. -K. (2001). A historical study of insert *-s* in Korean. Ph.D. dissertation, *Seoul National University*, Seoul, Korea.

Larsen, D, & Heinz, J. (2012). Neutral vowels in sound-symbolic vowel harmony in Korean. *Phonology*, 29, 433–464. https://doi.org/10.1017/S095267571200022X.

Lee, S. -Y. (2006). Lexical characteristics and meaning of color terms in the Late Middle Korean. *Journal of Korean Linguistics*, 47, 235–263.

Lee, Y. (1993). Topics in the vowel phonology of Korean. Ph.D. dissertation, *Indiana University*, Bloomington, Indiana.

Martin, S. E. (1962). Phonetic symbolism in Korean. In N. Poppe (Ed.), *American studies in Altaic linguistics* (Uralic and Altaic Series 13) (pp. 177–189). Bloomington, IN: Indiana University Publications.

Martin, S. E. (1992). *A reference grammar of Korean*. Rutland & Tokyo: Charles E. Tuttle Publishing.

McCarthy, J. J. (1983). Phonological features and morphological structures. In J. Richardson, M. Marks, & A. Chukerman (Eds.), *Papers from the parasession on the interplay of phonology, morphology, and syntax* (pp. 135–161). Chicago: Chicago Linguistic Society.

Nam, Y. (1992). *Wulimal pwunlyu sacen* [Korean lexicon by semantic categories]. 4 vols. Seoul: Hangang Publishing.

(The) National Institute of the Korean Language (1999). *Phyocwun kwuke taysacen* [A comprehensive standard Korean dictionary]. Seoul: Doosan Dong-A Publishing.

Park, Y. (1989). *Wulimal kallay sacen* [Korean lexicon by semantic categories]. Seoul: Hankil Publishing.

Ramstedt, G. J. (1939). *A Korean grammar.* Helsinki: Suomalais-Ugrilainen Seura.

Rhee, S. (1996). *Semantics of verbs and grammaticalization: The development of Korean from a crosslinguistic perspective*. Ph.D. dissertation, The University of Texas at Austin, Seoul: Hankook Publisher.

Rhee, S. (2001a). Forms and functions of diminutives. *The Journal of Linguistic Science*, 19, 115–146.

Rhee, S. (2001b). Grammaticalization of diminutives in Korean. *Language and Linguistics*, 28, 133–149.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rhee, S., & Koo, H. J. (2014). Grammaticalization of causatives and passives and their recent development into stance markers in Korean. *Poznań Studies in Contemporary Linguistics* 50(3), 309–337. https://doi.org/10.1515/psicl-2014-0018.

Rhee, S., & Koo, H. J. (2017). Multifaceted gustation. *Terminology*, 23(1), 38–64. https://doi.org/10.1075/term.23.1.02rhe.

Safran, A. B., & Sanda, N. (2015). Color synesthesia: Insight into perception, emotion, and consciousness. *Current Opinions in Neurology*, 38, 36–44. https://doi.org/10.1097/WCO.0000000000000169

Sampson, G. (1985). *Writing systems: A linguistic introduction*. Stanford, CA: Stanford University Press.

Shin, J. (2015). Vowels and consonants. In Lucien Brown & Jaehoon Yeon (Eds.), *The handbook of Korean linguistics* (pp. 3–21). Oxford: Wiley Blackwell. https://doi.org/10.1002/9781118371008.ch1

Shin, J., Kiaer, J., & Cha, J. (2013). *The sounds of Korean*. Cambridge: Cambridge University Press.

Sohn, H.-M. (1999). *The Korean language*. Cambridge: Cambridge University Press.

Song, J. J. (2005). *The Korean language: Structure, use and context*. London: Routledge. https://doi.org/10.4324/9780203390825

Suh, C. S. (2003[2000]). *Kwuke ewen sacen* [Korean etymological dictionary]. Seoul: Bogosa Books.

Tyson, R. E. (1994). Korean color naming and Korean-English language contact: A study in linguistic variation and semantic change. Ph.D. dissertation, University of Arizona, Tucson, Arizona.

Yeon, J., & Brown, L. (2011). *Korean: A comprehensive grammar*. London: Routledge.

Yoo, C. D. (1978). *Ehwisa yenkwu* [Study of historical lexicology]. Seoul: Iwoo Publishing.

Appendix. 127 RED-words in Korean (not exhaustive)

BRT: bright; **DRK**: dark; **EXT**: extended surface; **INT**: intense; **SPT**: spotty; **UND**: undesirable; **WK**: weak

Korean	Transliteration	Pronunciation	Meaning (not exhaustive)
Noun			
빨강	ppalkang	p*algaŋ	Red
빨간색	ppalkansayk	p*algansæk	red BRT
붉은색	pwulunsayk	pulgɯnsæk	red DRK
Adjective (most derived with the light verb *ha-*)			
발가야드르르하-	palkayatululuha-	palgajadɯrɯrɯrɯha-	reddish BRT WK EXT
발가우리하-	palkawuliha-	palgauriha-	reddish BRT WK EXT
발갛-	palkah-	palgat-	reddish BRT WK
발그대대하-	palkutaytayha-	palgɯdædæha-	reddish BRT WK EXT
발그댕댕하-	palkutayngtayngha-	palgudæŋdæŋha-	reddish BRT WK EXT
발그레하-	palkuleyha-	palgɯreha-	reddish BRT WK
발그름하-	palkulumha-	palgɯrɯmha-	reddish BRT WK

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Korean	Transliteration	Pronunciation	Meaning (not exhaustive)
발그무레하-	palkumwuleyha-	pɑlgɯmurehɑ-	reddish BRT WK EXT
발그속속하-	palkusoksokha-	pɑlgɯsoksokhɑ-	reddish BRT WK EXT
발그스름하-	palkusulumha-	pɑlgɯsɯrɯmhɑ-	reddish BRT WK EXT
발그족족하-	palkucokcokha-	pɑlgɯʤoktʃokhɑ-	reddish BRT WK EXT
발긋발긋하-	palkuspalkusha-	pɑlgɯtpɑlgɯthɑ-	reddish BRT WK EXT SPT
발긋하-	palkusha-	pɑlgɯthɑ-	reddish BRT WK
벌거여드르르하-	pelkeyetululuha-	pʌlgʌjʌdɯrɯrɯhɑ-	reddish DRK WK EXT UND
벌거우리하-	pelkewuliha-	pʌlgʌuriha-	reddish DRK WK EXT UND
벌겋-	pelkeh-	pʌlgʌt-	reddish DRK WK UND
벌그데데하-	pelkuteyteyha-	pʌlgɯdedehɑ-	reddish DRK WK EXT UND
벌그뎅뎅하-	pelkuteyngteyngha-	pʌlgɯdeŋdeŋhɑ-	reddish DRK WK EXT UND
벌그레하-	pelkuleyha-	pʌlgreha-	reddish DRK WK EXT UND
벌그름하-	pelkulumha-	pʌlgɯrɯmhɑ-	reddish DRK WK EXT UND
벌그무레하-	pelkumwuleyha-	pʌlgɯmurehɑ-	reddish DRK WK EXT UND
벌그숙숙하-	pelkuswukswukha-	pʌlgɯsuksukhɑ-	reddish DRK WK EXT UND
벌그스름하-	pelkusulumha-	pʌlgɯsɯrɯmhɑ-	reddish DRK WK EXT UND
벌그죽죽하-	pelkucwukcwukha-	pʌlgɯʤuktʃukhɑ-	reddish DRK WK EXT UND
벌긋벌긋하-	pelkuspelkusha-	pʌlgɯtpʌlgɯthɑ-	reddish DRK WK SPT UND
벌긋하-	pelkusha-	pʌlgɯthɑ-	reddish DRK WK UND
볼그대대하-	polkutaytayha-	polgɯdædæhɑ-	reddish BRT WK EXT
볼그댕댕하-	polkutayngtayngha-	polgɯdæŋdæŋhɑ-	reddish BRT WK EXT
볼그레하-	polkuleyha-	polgɯrehɑ-	reddish BRT WK EXT
볼그름하-	polkulumha-	polgɯrɯmhɑ-	reddish BRT WK EXT
볼그무레하-	polkumwuleyha-	polgɯmurehɑ-	reddish BRT WK EXT
볼그속속하-	polkusoksokha-	polgɯsoksokhɑ-	reddish BRT WK EXT
볼그스름하-	polkusulumha-	polgɯsɯrɯmhɑ-	reddish BRT WK EXT
볼그족족하-	polkucokcokha-	polgɯʤoktʃokhɑ-	reddish BRT WK EXT
볼긋볼긋하-	polkuspolkusha-	polgɯtpolgɯthɑ-	reddish BRT WK SPT
볼긋하-	polkusha-	polgɯthɑ-	reddish BRT WK
불그데데하-	pwulkuteyteyha-	pulgɯdedehɑ-	reddish DRK WK EXT UND
불그뎅뎅하-	pwulkuteyngteyngha-	pulgɯdeŋdeŋhɑ-	reddish DRK WK EXT UND
불그레하-	pwulkuleyha-	pulgɯrehɑ-	reddish DRK WK EXT UND
불그름하-	pwulkulumha-	pulgɯrɯmhɑ-	reddish DRK WK EXT
불그무레하-	pwulkumwuleyha-	pulgɯmurehɑ-	reddish DRK WK EXT UND
불그숙숙하-	pwulkuswukswukha-	pulgɯsuksukhɑ-	reddish DRK WK EXT UND
불그스름하-	pwulkusulumha-	pulgɯsɯrɯmhɑ-	reddish DRK WK EXT UND
불그죽죽하-	pwulkucwukcwukha-	pulgɯʤuktʃukhɑ-	reddish DRK WK EXT UND
불긋불긋하-	pwulkuspwulkusha-	pulgɯtpulgɯthɑ-	reddish DRK WK SPT
불긋하-	pwulkusha-	pulgɯthɑ-	reddish DRK WK
붉-	pwulk-	pulk-	red DRK
붉으락푸르락하-	pwulkulakphwululakha-	pulgɯrɑkp^{h}urɯrɑkhɑ-	reddish SPT (mixed)
빨갛-	ppalkah-	p*algat-	Red
빨그대대하-	ppalkutaytayha-	p*algɯdædæhɑ-	reddish BRT EXT

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Korean	Transliteration	Pronunciation	Meaning (not exhaustive)
빨그댕댕하-	ppalkutayngtayngha-	pˀalgɯdæŋdæŋha-	reddish BRT EXT
빨그레하-	ppalkuleyha-	pˀalgɯreha-	reddish BRT EXT
빨그름하-	ppalkulumha-	pˀalgɯrɯmha-	reddish BRT EXT
빨그무레하-	ppalkumwuleyha-	pˀalgɯmureha-	reddish BRT EXT
빨그속속하-	ppalkusoksokha-	pˀalgɯsoksokha-	reddish BRT EXT
빨그스름하-	ppalkusulumha-	pˀalgɯsɯrɯmha-	reddish BRT EXT
빨그족족하-	ppalkucokcokha-	pˀalgɯʤoktʃokha-	reddish BRT EXT
빨긋빨긋하-	ppalkusppalkusha-	pˀalgɯtpˀalgɯtha-	reddish BRT EXT SPT
빨긋하-	ppalkusha-	pˀalgɯtha-	reddish BRT
뻘겋-	ppelkeh-	pˀʌlgʌt-	reddish DRK UND
뻘그뎅뎅하-	ppelkuteyngteyngha-	pˀʌlgɯdeŋdeŋha-	reddish DRK EXT UND
뻘그레하-	ppelkuleyha-	pˀʌlgɯreha-	reddish DRK EXT UND
뻘그름하-	ppelkulumha-	pˀʌlgɯrɯmha-	reddish DRK EXT UND
뻘그무레하-	ppelkumwuleyha-	pˀʌlgɯmureha-	reddish DRK EXT UND
뻘그숙숙하-	ppelkuswukswukha-	pˀʌlgɯsuksukha-	reddish DRK EXT UND
뻘그스름하-	ppelkusulumha-	pˀʌlgɯsɯrɯmha-	reddish DRK EXT UND
뻘그죽죽하-	ppelkucwukcwukha-	pˀʌlgɯʤuktʃukha-	reddish DRK EXT UND
뻘긋뻘긋하-	ppelkusppelkusha-	pˀʌlgɯkpˀʌlgɯtha-	reddish DRK EXT SPT UND
뻘긋하-	ppelkusha-	pˀʌlgɯtha-	reddish DRK
뽈그름하-	ppolkulumha-	pˀolgɯrɯmha-	reddish BRT
뽈그스름하-	ppolkusulumha-	pˀolgɯsɯrɯmha-	reddish BRT
뽈그족족하-	ppolkucokcokha-	pˀolgɯʤoktʃokha-	reddish BRT EXT
뽈긋뽈긋하-	ppolkusppolkusha-	pˀolgɯtpˀolgɯtha-	reddish BRT EXT SPT
뽈긋하-	ppolkusha-	pˀolgɯtha-	reddish BRT
뿔그름하-	ppwulkusulumha-	pˀulgɯrɯmha-	reddish DRK EXT UND
뿔그스름하-	ppwulkusulumha-	pˀulgɯsɯrɯmha-	reddish DRK EXT UND
뿔그죽죽하-	ppwulkucwukcwukha-	pˀlgɯʤuktʃukha-	reddish DRK EXT UND
뿔긋뿔긋하-	ppwulkusppwulkusha-	pˀulgɯtpˀulgɯtha-	reddish DRK EXT SPT UND
뿔긋하-	ppwulkusha-	pˀulgɯtha-	reddish DRK
뿕-	ppwulk-	pˀulk-	reddish DRK
새빨갛-	sayppalkah-	sæpˀalgat-	red BRT INT
샛빨갛-	saysppalkah-	sæppˀalgat-	red BRT INT
시뻘겋-	sippelkeh-	ʃipˀʌlgʌt-	red DRK INT UND
싯뻘겋-	sisppelkeh-	ʃippˀʌlgʌt-	red DRK INT UND
울긋불긋하-	wulkuspwulkusha-	ulgɯtpulgɯtha-	reddish DRK (mixed)
Verb (derived from the causative affix *-ci-*)			
발개지-	palkayci-	pɑlgæʤi-	become reddish BRT WK
빨개지-	ppalkayci-	pˀalgæʤi-	become red
벌개지-	pelkayci-	pʌlgæʤi-	become reddish DRK WK
붉어지-	pwuleci-	pulgʌʤi-	become red DRK
뻘개지-	ppelkayci-	pˀʌlgæʤi-	become reddish DRK UND
새빨개지-	sayppalkayci-	sæpˀalgæʤi-	become red INT

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Korean	Transliteration	Pronunciation	Meaning (not exhaustive)
샛빨개지-	saysppalkayci-	sæpp*ɑlgædʒi-	become red INT
시뻘개지-	sippelkayci-	ʃip*ʌlgædʒi-	become red DRK INT UND
싯뻘개지-	sisppelkayci-	ʃipp*ʌlgædʒi-	become red DRK INT UND
Adverb ('in such a manner of being adjective')			
발그스름히	palkusulumhi	pɑlgɯsɯrɯmhi	reddish BRT WK EXT
발긋발긋	palkuspalkus	pɑlgɯtpɑlgɯt	reddish BRT WK SPT
벌그름히	pekulumhi	pʌlgɯrɯmhi	reddish DRK WK UND
벌그스름히	pelkusulumhi	pʌlgɯsɯrɯmhi	reddish DRK WK EXT UND
벌긋벌긋	pelkuspelkus	pʌlgɯtpʌlgɯt	reddish DRK WK SPT UND
볼그름히	polkulumhi	polgɯrɯmhi	reddish BRT WK EXT
볼그스름히	polkusulumhi	polgɯsɯrɯmhi	reddish BRT WK EXT
볼긋볼긋	polkuspolkus	polgɯtpolgɯt	reddish BRT WK EXT SPT
볼긋이	polkusi	polgɯʃi	reddish BRT WK
불그름히	pwulkulumhi	pulgɯrɯmhi	reddish DRK WK UND
불그스름히	pwulkusulumhi	pulgɯsɯrɯmhi	reddish DRK WK EXT UND
불긋불긋	pwulkuspwulkus	pulgɯtbulgɯt	reddish DRK WK EXT SPT UND
불긋이	pwulkusi	pulgɯʃi	reddish DRK WK UND
붉으락푸르락	pwulkulakphwululak	pulgɯrɑkp^{h}urɯrɑk	reddish DRK WK (mixed)
빨그스름히	ppalkusulumhi	p*algɯsɯrɯmhi	reddish BRT EXT
빨긋빨긋	ppalkusppalkus	p*algɯtp*algɯt	reddish BRT EXT SPT
뻘그름히	ppelkulumhi	p*ʌlgɯrɯmhi	reddish DRK EXT UND
뻘그스름히	ppelkusulumhi	p*ʌlgɯsɯrɯmhi	reddish DRK EXT UND
뻘긋뻘긋	ppelkusppelkus	p*ʌlgɯtp*əlgɯt	reddish DRK EXT SPT UND
뽈그름히	ppolkulumhi	p*olgɯrɯmhi	reddish BRT EXT UND
뽈그스름히	ppolisulumhi	p*olgɯsɯrɯmhi	reddish BRT EXT
뽈긋뽈긋	ppolkusppolkus	p*olgɯtp*olgɯt	reddish BRT EXT SPT
뽈긋이	ppolkusi	p*olgɯʃi	reddish BRT EXT
뿔그름히	ppwulkulumhi	p*ulgɯrɯmhi	reddish DRK EXT
뿔그스름히	ppwulkusulumhi	p*ulgɯsɯrɯmi	reddish DRK EXT
뿔긋뿔긋	ppwulkusppwulkus	p*ulgɯtp*ulgɯt	reddish DRK EXT SPT
뿔긋이	ppwulkusi	p*ulgɯʃi	reddish DRK EXT
올긋볼긋	olkuspolkus	olgɯtpolgɯt	reddish BRT (mixed)
울긋불긋	wulkuspwulkus	ulgɯtpulgɯt	reddish DRK (mixed)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lexicalization patterns in color naming in Gbaya, a Ubanguian language of CAR

Paulette Roulon-Doko
LLACAN (CNRS, INALCO, Université Sorbonne Paris-Cité)

In this chapter, I show that Gbaya, a Ubanguian language of the Central African Republic, does not lexically distinguish color from the visual aspect resulting from a variety of parameters. I first analyze in detail the lexical terms and the grammatical devices that Gbaya uses for color naming. Then, I discuss how Gbaya speakers put these terms into actual use in a daily task. Gbaya illustrates the main role played by grammar to organize the lexicon according to properties expressed by adjectives, processes expressed by verbs, and visual conspicuousness express by nouns. Beyond language idiosyncrasies, the case of Gbaya illustrates, once more (Wierzbicka, 2008), a possible universal tendency among hunter-gatherers in the partition of the visual aspect.

Keywords: color, visual aspect, Niger-Congo, Ubanguian, Gbaya, grammatical categories, grammar

1. Introduction

This study deals with Gbaya kara ɓodoe, a Ubanguian language (Niger-Congo phylum) spoken in the northwest part of the Central African Republic. The corpus was collected between 1970 and 1995, then again in 2011 and 2012, in the village of Ndongué in the course of research including a total of five years of fieldwork.

My study does not follow Berlin & Kay's (1969) theory of universals in color naming based on elicitation, and challenges the universality of "basic" color terms, adhering to the critical arguments developed by Wierzbicka (2008), Dimmendaal (2015, pp. 104–112), and, namely, Moñino (2004) who refers especially to Gbaya ɓodoe. It is based on Lucy's (1997) relativist approach that recommends "balancing the emphasis on form and substance (language and reality) in establishing the semantic system of a language" (p. 340). I will first present the lexical terms and

https://doi.org/10.1075/sfsl.78.07rou
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

the grammatical devices that Gbaya uses for color naming (in Section 2). Then, on the basis of a corpus of spontaneous utterances, I will discuss how Gbaya speakers put these terms into actual use in daily tasks (in Section 3).[1]

2. Lexical and grammatical devices for color naming

Among the sixteen grammatical categories of Gbaya, five are relevant for color-related terminology: nouns, verbs and three categories of adjectives.

2.1 Nouns

While noun compounding is a frequent procedure yielding almost half of all nouns (48%), derivation is a procedure of a lesser importance (2.7%) producing only 134 deverbal nouns,[2] none related to color. There are also 10 deadjectival human nouns. The set of simple nouns is formally or semantically unrelated to verbs.

2.1.1 *The generic noun* dàp

The Gbaya noun *dàp* is a simple term whose meaning relies on its various uses. When used by itself, this noun means 'tattoo, drawing' and functions as a direct object (DO) of various verbs that each specify the technique used: *gɔn* 'cut', *gba* 'split', *pasi* 'slit' for skin scarifications; *ʔa* 'put several things', *yɔr* 'draw (lines)' for pottery patterns.[3] In a genitive phrase, *dàp* refers to the pattern that is specific to the relevant item, such as *dàp té gɔ̀* (pattern.L / body.L / leopard) 'the patterns of the leopard's fur' or *dăp gòè* (pattern.L/cloth) 'the patterns of the cloth'.[4] The following four compound nouns can be formed using this structure.

dàp-nú-ʔér	(pattern.L/tip/hand)	'fingerprints'
dăp-sɔ̀	(pattern.L /divinity-ancestor)	'stretch marks'
dăp-kɔ̀-náŋ	(pattern.L /inside/foot)	'foot wrinkles'
dăp-kɔ̀-ʔér	(pattern.L /inside/hand)	'lines of the palm'

1. Thanks to Martine Vanhove and to the two anonymous reviewers for their relevant comments and suggestions on this chapter, to Julienne Doko for the English translation and Raymond Boyd for his re-reading.

2. These deverbal nouns have an abstract meaning (*tàà* "thought" from V *ta* "think") or refer to an element considered as a cultural prototype (*tòì* "luggage" from V *toi* "carry on the head").

3. The lexical form of verbs is written without tones because verbs do not have lexical tones. Their tones are always TAM markers.

4. A tonal mark links the elements of these phrases, glossed L. When a connective *kɔ́* appears between the two elements of a genitive phrase, it expresses possession.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

All these examples illustrate that *dàp* is a generic noun that refers to a type of visual conspicuousness, namely, any pattern as opposed to a unified appearance or solid color that have no hyperonyms in the language. It is an abstract term related to a unified concept in Gbaya that does not match a unique English term and needs to be translated by numerous terms such as 'tattoo' on human skin, 'marks' or 'designs' to be incised or painted on pottery, 'pattern' on animal fur, liana, cloth, etc. It cannot refer only to color, as exemplified in Section 3. From the point of view of Gbaya speakers, *dàp* is a salient and unitary concept but it has no equivalent in English.

Moñino (2004), though not adhering to Berlin and Kay's theory and claiming that there is no specific term for 'color' in Gbaya, goes on to say: "However, one can translate the French word 'color' by the generic noun *dàp-* 'pattern, drawing, setting, color'"[5] (pp. 245–246). This claim points out the adverse effects of any investigation using color stimuli and exemplifies Lucy's assertion: "What is crucial to recognize is that 'an adequate knowledge' of the system would never have been produced by restricting the stimuli to color chips and the task to labeling" (1997, p. 326). Indeed, Gbaya has no generic term for 'color', which cannot be considered as a unified concept in this language, nor as a "natural categorization of nature", as mentioned by Dimmendaal for many Nilotic Languages (2015, p. 165). There is no way in Gbaya to directly ask about color like in English "what is the color of this?". The Gbaya question *mɔ̀ hɛ̰̀ hégè* (thing/PROX/how) 'how is it?' triggers an answer that is a global description of the visual appearance where color is one element among others, as noted by Wierzbicka (2008, p. 412) for Warlpiri, an Australian language.

2.1.2 *Color nouns*

Like many other languages, Gbaya uses familiar and well-known entities to name some colors. The lexical pattern that realizes an 'N-like' meaning is a genitive phrase with the noun *tɛ̀* 'body, entity' as the first component, and as the second component, a noun referring to an entity which constitutes a prototype for the color it represents: [tɛ̀ (body) + N]. But, contrary to the regular construction (see Section 2.1.1. above), the head of the NP does not carry the tone marker (L) which would be required in a phrasal construction and this structure is not syntactic, highlighting that these terms are compound nouns.[6]

5. "On peut cependant traduire le mot français 'couleur' par le terme générique nominal dàp- 'motif, dessin, décor, couleur'."

6. The regular phrase is [tɛ̀.H N] lit. body.L N.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

There are nine compound nouns of this kind which use animals or plants as prototypical elements.[7] Four animals and three plants are prototypes for a solid color and two animals are prototypes for patterns out of which some are not related to a specific color. Thus, the Gbaya N-like pattern is not specific to color but relates globally to visual patterns. In addition, one caterpillar's name *nàà-gbèngbè* (lit. The one badly shaved)[8] is used by itself to refer to cross stripes. Table 1 provides the full list of these color nouns.

Table 1. Color nouns in the lexicon

	Color nouns: [tɛ̀ N] 'body/N' or N		
	color	**term**	**prototypical reference**
solid	gingery red	*tɛ̀-bàn*	red duiker
	brown	*tɛ̀-bìò*	bush duiker
	dark brown	*tɛ̀-mbòò*	yellow duiker
	bright green	*tɛ̀-kènà*	green pigeon
	soft green	*tɛ̀-tɔ́r-zɔ́*	grass regrowth (raw grass)
	yellow	*tɛ̀-tárà ou tɛ̀-tárà*	*Anogeissus leiocarpus*
	purple	*tɛ̀-kèlú*	*Syzygium guineense*
pattern	mottled black and white	*tɛ̀-gbànù*	guinea fowl
	with lengthwise stripes	*tɛ̀-gòɗá*	striped grass mouse
	with cross stripes	*nàà-gbèngbè*	caterpillar sp.

These color nouns are systematically used in predicative constructions. They are either placed after the copula *né* or introduced by the preposition *nɛ̀* 'as' after the verb *ʔɔ* 'be'.

(1) *dùà hɛ̰̀ né tɛ̀-bàn*
goat PROX COP body-red duiker[9]
'This goat is gingery red'

(2) *làà hɛ̰̀ né nàà-gbèngbè*
garment PROX COP caterpillar sp.
'This garment has horizontal stripes'

7. Moñino (2004) reports some comparable constructions which are not attested in my corpus. I suspect that this structural frame [tɛ̀.N] allows the creation of immediately understandable prototypical references in the framework of a survey regarding color samples. Only the examination of a larger corpus would make it possible to confirm whether they also need to be included in the lexicon.

8. So named because its stripes are reminiscent of the lines that remain after shaving.

9. Glosses follow the Leipzig glossing rules with the addition of L = tonal linker, AUG = augmentative, SIM = similative, DO = direct object, and Fr = borrowed from French.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(3) *dùà hɛ̰̀ ʔɔ́ nɛ̀ tɛ̀-bàn*
goat PROX IPFV.be as body-red duiker
'This goat is gingery red'

This way of referring to color is well attested but only applies to ten occurrences, seven of which are solid colors, emphasizing, however, that the reference to an animal or a plant remains a more complex visual aspect than a simple color.

2.1.3 *Deadjectival human nouns*

Some verbal adjectives can be combined with the suffix *-í*, which is a reduced form of *wí* 'person', to create ten color related nouns referring to human beings. Two of them are based on color adjectives *tṵ́ḭ́* 'black person' from the verbal adjective *tṵ́* 'black' and *búí* 'white person' from the primary adjective *bú* 'white'. When used with the plural marker *ʔó*, they express a collective noun: *ʔó búí* 'Whites' and *ʔó tṵ́ḭ́* 'Blacks'. Interestingly, the verbal adjective *gbɛ̰́* 'red' does not produce such a deadjectival noun. It is only the NP *gbɛ̰́ wí* (red/person) that is used to refer to a particular skin color which characterizes, among others, the Mbororo, Fula transhumant pastoralists.

2.2 Verbs

The Gbaya verb is always an expression of a process since there are no quality verbs. Gbaya verbs have a double diathesis – the internal versus the external – without any morphological marking (Roulon-Doko, 2006, p. 319). When there is a DO after a verb (a noun or the term *mɔ̀* 'thing' by default), the construction is transitive and expresses an active voice. When the verb has no DO, it is an intransitive construction that expresses a middle voice. From a notional abstract meaning, each verb develops an important polysemy depending on the constructions. There is no hierarchical ranking between the meanings of a verb and the distinction between literal and figurative meanings is not relevant. When referring to color, there are three verbs that express the process of change. They are presented in Table 2.

Table 2. Verbs associated with color

V	Intransitive (middle voice)	Transitive (active voice)
tṵ	become dark, strengthen	darken, enhance
gbɛ	ripen, become red	redden
fɛŋ	bleach, become white, fade	depreciate, humiliate, soil, whiten

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

From the Gbayas' cultural point of view, the changing process has a positive connotation for the verb *tu̱* 'become dark' and a negative one for the verb *fɛŋ* 'become white', while the verb *gbɛ̱* 'become red' is related to many situations such as ripeness, skin inflammation, or eye irritation, and does not have a unique cultural connotation. It is the change of color that is positively or negatively perceived and not the color itself. For instance, *sùkà* 'cassava leaves' which are dark green when cooked on their own, are used to cook a dish by adding sesame paste, which makes its color much lighter (like spinach with cream, for example). Nevertheless, the verb used for this process is *tu̱* 'darken', that is, 'enriched': *sùká tu̱ǎ* (cassava leaves/PFV. become dark). It is clear that, in this case, the verb *tu̱* does not refer to the color of the leaves but to this highly appreciated dish.

2.3 Adjectives

Adjectives (A) are widely used to express a part of the visual perception of what English refers to as 'color'. The adjectival category comprises a small group of 20 primary adjectives (PA), a group of 171 adjectives derived from verbs (VA), and a group of 2263 adjective-adverbs (AA) (Roulon-Doko, 2001), known in the literature as ideophones or as expressive adverbs. These three categories of adjectives share some properties: all three are used as attributives, placed before the noun they modify[10] and derive a noun by means of the suffix *-à* (which is assimilated to the previous vowel with the CV terms). In all the other positions (as a postnominal modifier or in a predicative or adverbial construction), the adjective-adverb remains unchanged, whereas the adjective and the verbal adjective must be nominalized and introduced by the preposition *nɛ̀* 'as'.

Table 3. The construction types of primary adjectives and verbal adjectives

Term	Category	Meaning	Construction	Literal meaning
tú̱	VA	'black'	*nɛ̀ tú̱ù̱*	'as black.NMLZ'
gbɛ̱́	VA	'red'	*nɛ̀ gbɛ̱́ɛ̱̀*	'as red.NMLZ'
bú	PA	'white'	*nɛ̀ búù ~ búà*	'as white.NMLZ'

Thus, in a predicative function, the AA directly follows the verb *ʔɔ* 'be',[11] whereas the VA and the PA need to be placed within a prepositional phrase [*nɛ̀* VA. NMLZ ~ PA.NMLZ] to have the same function.

10. In this position, the A and the VA carry the tonal marker L and the AA undergoes a tone rise TR.

11. Otherwise, the basic meaning of this verb is 'stand, stretch out'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(4) *wáɲáà*[12] *ʔɔ́* *fàtàfùtù*
leave.DEF IPFV.be light-beige
'The leaves are beige.' [AA]

(5) *wáɲáà* *ʔɔ́* *nɛ̀* *gbɛ́ɛ̀*
leave.DEF IPFV.be as red.NMLZ
'The leaves are red.' [VA]

Furthermore, this prepositional phrase can also directly follow a noun or a noun phrase forming the second part of a juxtapositive predication, whereas there are no instances in the corpus of the use of an AA in such a construction.

(6) *dóà* *nɛ̀* *búù*
flower.DEF as white.NMLZ
'The flowers are white.' [PA]

I will now present the specificity of each of these adjective types.

2.3.1 *The primary adjective* bú *(PA)*

bú is one of the twenty primary adjectives designating "a fundamental quality – an original natural state – that cannot result from a process" (Roulon-Doko, 2008, p. 89). It means 'rough, natural'. The raw condition expressed by *bú* refers to the light beige color of clay for pottery, the natural brown color of bark cloth before dyeing or the application of a design, and the dark green color of cassava leaves. There is no reference to a specific color.

bú kpánà	(~ / pottery)	'unfired pottery'
bú túí	(~ / bark cloth)	'raw bark cloth'
bú sùkà	(~ / cassava leaves)	'cassava leaves cooked on their own'

Applied to flowers whose colors are various,[13] the use of bú refers only to whiteness which matches with their natural color.

bú dò	(~ / flower)	'white flower'

Consequently, this adjective *bú* is polysemous, and the meaning 'white' is one of the manifestations of its 'rough, natural' meaning, as the previous examples show.

---

12. Nouns may refer to a singular or plural, with no marker.

13. Flowering is expressed by the verb *do* 'ferment, bloom' which does not necessarily involve whiteness. Thus depending on the tree, this verb *do* can be follow by AA as *bóŋ-bóŋ* 'white' to refer to the white flowers of the *Terminalia glaucescens*, or *zóŋ-zóŋ* 'earth-reddish' to refer to the red flowers of the *Parkia clappertoniana*, and *kpʒ́ʒ́-kpʒ́ʒ́* 'very dark' to refer to the dark brown inflorescence of some grass.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.3.2 *Verbal adjectives (VA)*

The three verbs *fɛŋ* 'become white', *tu̱* 'become black', and *gbɛ̱* 'become red' derive an adjective (VA). Only the last two VAs designate a color as a resulting state, *tú̱* 'black' and *gbɛ̱́* 'red', respectively. The VA that stems from *fɛŋ* 'become white', *fɛ́ŋá* has the resulting meaning 'faded, humiliated' (*fɛ́ŋá béí* 'a person without shame') or 'whitened' (*fɛ́ŋá náŋáà* 'dusty legs' or *fɛ́ŋá gèɗà* 'dried cassava'), but never designates the color 'white' which is expressed by the adjective *bú*.

It is important to note that the meaning of the corresponding VAs conveys a value that is no longer resultative. So the meaning conveyed via the verbs *tu̱* and *gbɛ̱*, and those conveyed by the corresponding VAs are quite different. Thus, the verbs primarily expresses a process and the VAs a property, a state, in this case, a color. For example, the verb *gbɛ̱* 'ripen, redden' is used even when ripeness is not linked to red. In contrast, the VA *gbɛ̱́* which only designates the color 'red' is used to talk about a ripe plant only if it is red. The Gbaya case highlights that "grammatical properties have implications on meanings of […] terms", as in the Zuni case (Lucy, 1997, p. 337).

The two VAs (*tú̱* and *gbɛ̱́*) and the PA (*bú*) are used exclusively in contrast to each another. They express a partition of the color spectrum into three areas: 'black' *tú̱*, 'red' *gbɛ̱́* and 'white' *bu*', but do not refer to a precise visual reference unlike the adjective-adverbs I will present now.

2.3.3 *Adjective-adverbs (AAs)*

AAs often have a reduplicated structure (identifiable by the use of a hyphen: *kpúŋ-kpúŋ* 'white').[14] This category is formally and semantically unrelated to other word categories. Properties expressed by AAs can be classified into 25 semantic domains, such as odor (15 AAs), opening (33 AAs), quantity (93 AAs), size-shape (352 AAs) or movement (383 AAs). There are 82 AAs for the color domain, comprising 67 AAs only dedicated to color and 15 dedicated to patterns which, from the native point of view, cannot be separated from color AAs, all of them being part of the visual aspect of an entity.

So AAs have their own specific meaning and can be used alone after the verb *ʔɔ* 'be'.

(7) *làà* *kɔ̱́ à̱* *ʔɔ́* *zèngèlè*
garment of.3SG IPFV.be glowing_red
'His garment is glowing red.'

(8) *zân ʔɔ́* *fàràfùrù*
sky IPFV.be whiten
'The sky has white clouds.'

14. There is no AA *kpúŋ* and *kpúŋ-kpúŋ* is not based on a reduplicative process. (Roulon-Doko, 2001).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

It is difficult to identify the typical referential value of each AA and it is also very hard to match it with an English term. In a few cases, Gbaya speakers specify a particular source of referential value, such as the color of the underside of *Solanum nigrum*'s leaves for *míkɔ́ɛ́-míkɔ́ɛ́* 'dark green', the blue duiker's color for *tu̱ù̱* 'slate gray',[15] and the palm kernel's color for *báŋá-báŋá* 'fiery red'.[16] However, they most commonly list a number of possible sources such as the flower of *Thonningia sanguinea*, a mushroom similar to the chanterelle, and lips for *ngbòlòè* 'orange-brown'; ashes, head with ringworm, and body stained with flour for *dùkùyùkù* 'very light gray'. In these cases, the referential value of the AA is the common denominator that, from the natives' point of view, suggests a "visual conspicuousness" (Wierzbicka, 2008, p. 412). There are some other systematic devices in the way AAs express property and color. First, the notion of degree is integrated lexically into AAs, such as *kábá-kábá* 'very white' or *kpɔ́ɔ́-kpɔ́ɔ́* 'very white'.[17] Secondly, the way the color is produced (modification, addition, dyeing, etc.) is taken into account in case of AAs whose meaning is complex, such as *mbùyèlɛ̀* 'speckled with white by milk or flour' or *kító-kító* 'darken by smoke'. Thirdly, some AAs combine color with another feature which can be another property, such as a tactile modality *ɓùlàɛ̀* 'white and soft' or the way the color stands out in the environment (see also Wierzbicka, 2008, for Warlpiri), such as *bùbùbù* 'large white space', describing a cotton field. Finally, a few AAs are contextualized color terms, "that is, terms which are invariably linked to specific natural or cultural objects and which cannot be used as indiscriminable [*sic*] or abstract way" (Lucy, 1997, p. 342). Most of them are linked to human beings: skin (4 AAs) (see 3.2.c), new-born baby skin (*yèngbèrè* 'light pink'), and two body parts, teeth (*bíí́* 'darkened') and lips (*ɗólé-ɗólé* 'deep pink'). Two AAs are linked to animals: *vɛ̰́ɛ̰́* 'red' for a baboon's buttock and *bèrìèŋ* 'red mass' for the buffalo. The last one, *ɗéréré* 'orange-yellow-golden', is linked to fluids, mainly urine, some decoctions and, more recently, to tea.

Without going into the details of the values of the adjective-adverbs presented in the annexed tables (8, 9, and 10), I will now discuss just a few examples. Fire can be described by several AAs depending on its range of colors combined with the way it stands out in the environment and the time when it is observed: *zèngèlè* 'glowing', *zɛ̰́ɛ̰́* 'fiery red', *ngbèzèlɛ̀* 'blazing', *bèzèrè* 'bright red', *ŋmɛ̀ɛ̀* 'dull red' (remote fire), *pàì* 'orange red' (fire in the night). However, none of these AAs are specific to

---

15. This AA is probably linked to the verb *tu̱* 'darken', the only case of a formal link between a verb and an AA. It does not constitute a derivational process, rather it shows motivation behind the AA meaning.

16. To express color, *báŋá-báŋá* is the only case in which the form of the AA is related to the name of the element it refers to: *báŋá* 'palm tree'.

17. There is no adjective intensifier in Gbaya.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

fire: *ngbèzèlè* is also used for embers and oozing wounds, *pài* for the setting sun and the red moon, and *zɛ̰̀ɛ̰̀* for the bird *ndùlù* 'Dybowski's twinspot' and the red duiker whose back is also described as *ŋàà* 'gingery', like a kind of a ripe banana. So it is clear that each AA has a specific meaning that relates to a color combining different elements such as hue, shape, luminosity, intensity, and even movement, distance or time, and corresponds to a concept that is salient and unitary in the Gbaya view of the world, but which has no equivalent in English. So it is difficult to translate them by a single color term, and the suggested translations in the tables or in the examples are approximations since the Gbaya values do not match the colors that are commonly used in European languages.

An AA can be used on its own in a predicative construction (cf. Example 6 and 7), but it can also be used as an adverb modifying any verb as mentioned previously. Each AA can only modify one of the three verbs of color: *tṵ*, *gbɛ̰* or *fɛŋ*.[18] In the same way, an AA can follow only one of the contrastive color adjectives *tṵ́* 'dark', *gbɛ̰́* 'red', and *bú* 'white'. The AA *kpúŋ-kpúŋ* 'white' is attested in the following sentence pointing to its capacity to modify the verb *fɛŋ* and the PA *bú*.

(9) *dó sùnù ʔɔ́ kpúŋ-kpúŋ*
flower sesame IPFV.be white
'The flowers of sesame are white.' [AA = Predicative]

(10) *kɔ̀rá hɛ̰̀ ʔɔ́ nè búà kpúŋ-kpúŋ*
chicken PROX IPFV.be as white.NMLZ white
'This chicken is white.' [AA = post predicative PA]

(11) *làmbù sá̰ kpúŋ-kpúŋ*
birdlime IPFV.clean white
'Birdlime is clean and white.' [AA = Adverb]

(12) *gèdà fɛ̀ŋà kpúŋ-kpúŋ*
cassava PFV.become_white.L white
'Cassava turns white.' [AA = Adverb]

So one can classify AAs depending on the color verb or adjective they can follow, pointing out their distribution into three areas, black, red, and white as shown in Table 4. The lines in the characterization of the referential value refer to the five semantic and formal types presented above.

18. In my corpus the AA *nèsì-nèsì* means 'not quite red' or 'not quite black', depending on whether it is used with the verb *gbɛ̰* 'become red' or *tṵ* 'become dark'. It is not considered as a color adjective but expresses an unachieved property.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. The semantic organization of the 67 adjective-adverbs of color

Characterization of the referential value	Spectrum partition		
	white	red	black
Basic (one or multiple sources)	6	19	4
+ Degree	2	0	2
+ Way of color production	5	10	4
Combination of properties	4	0	0
Contextualized colors	1	6	4
Total 67	18	35	14

In addition, there are fifteen AAs that express patterns (see Table 10). Five of them include a color in the definition of their meaning such as *làɓù-làɓù* 'black and white' and *làk-làk* 'with black spots'. The other ten do not include any specific color, but refer to alternating colors, striations, stripes or spots, with specific variations based on size, layout, and shape, one of which applies only to human skin (*dìr-dìr* 'with two-colored big spots' due to leprosy or oil on the body).

It may be relevant for certain AAs of color to differentiate a component that is more or less bright, shiny or dull. However, for luminosity, a distinct group of 29 AAs, which do not refer to a specific color, is used (Table 11). They refer to a cline of luminosity, ranging from the darkest to the lightest, distinguishing between shining, glowing, glimmering, dazzling or sparkling. They also indicate the nature of the light – flickering, punctual, translucent – and even bring out the specificity of some materials, like the eyes or fire. I will not deal further with these as they do not specifically express color even if they are part of the visual aspect.

2.3.4 *Summary*

This section showed that there is no generic noun for 'color' but a generic noun for 'pattern' *dàp*, a few N-like nouns and one noun which refer specifically to a color and two deadjectival human nouns, while three polysemous verbs referring to color deal primarily with a process of change. Otherwise, Gbaya widely uses adjectives to express a part of the visual perception of what English refers to as 'color'. Within these adjectives, the PA *bú* 'white' and the two VAs *tṵ́* 'dark' and *gbɛ́* 'red' do not refer to a precise visual reference but express a partition of the color spectrum into three areas. Finally, it is the AAs that are in Gbaya dedicated to express color property, but they do not correspond to English colors and are instead related to a salient and unitary visual concept significant for Gbaya speakers.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. When and how do the Gbaya use color terms in a daily task?

In this section, I study the way Gbaya speakers usually use all the terms previously presented. I will first discuss the use of color in compound nouns. Then, I will discuss the way they use color to characterize elements in their environment and in diverse life circumstances through tales.

3.1 Colors in compound nouns

I will first describe how some color adjectives become parts of compounds (3.1.1), then how an entity can become the visual reference to a color within a compound noun (3.1.2).

3.1.1 *Color adjectives*

The verbal adjectives *tú̱* 'black', *gbɛ̱́* 'red', and the primary adjective *bú* 'white' are used attributively in compound nouns. The referred color is a relevant criterion for identifying the entity that the compound noun denotes. This is the case for the names of 18 plants and 13 animals. In 60% of the cases, the adjective modifies a generic term such as liana: *bú-ɲàk* (lit. white / liana) '*Triclisia macrophylla*', tree, monkey, giant African centipede, lizard, etc. The color then has a contrastive value for identifying the entity in question. In other instances, color is a specific part of the entity identified: the 'monkey's red tail' *gbɛ̱́-dòm* (red / tail) or the 'grasshopper's black mouth' *nàà-tú̱-nú-ngírí* (this one / black / mouth / grasshopper sp.). Lastly, the names of two mushrooms contain the VA *gbɛ̱̀* 'red' and the AA *lòè* (for *ngbòlòè*) 'reddish orange': *yàà-gbɛ̱́-lòè* (this one / red / reddish-orange), and the AA of texture *sàngbà-sàngbà* 'fibrous'*: yàà-gbɛ̱́-sàngbà* (this one / red / fibrous). In ;, there are only six other compound nouns containing a color term: three refer to body parts, one to a disease and two to climatic phenomena.[19]

gbɛ̱́-tɛ̀	(red/body)	'children's hookworm'[19]
bú-wèsé	(white/sun)	'beaming sun'
bú-wèé-bɛ̀ɛ̀	(white/fire/dry season)	'months of dry season'
bú-yík	(white/eye)	'whites of the eye'
tú̱-yík	(black/eye)	'eye pupil'
tú̱-nú-bér	(black/tip/breast)	'breast areola'

The inclusion of color adjectives in compounds is relatively rare and represent only 2.1% of this lexical group.

19. A disease which first causes red feces, then causes redness of the scalp and eyes.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.1.2 *Name of an entity referring to color*

Table 5 lists the entities used as parts of compound nouns and suggested translations for the color of the entity it relates to, since these entities which are significant for the Gbaya are not significant for non-Gbaya speakers.

Table 5. Entities used in compound nouns as a reference to a color

	Entity		**> Color**
PLANTS	*gàú*	rattan	dark green trunk
	mbírò	kind of liana	black trunk
	dàrɛ̀	*Acacia ataxacantha*	juice from black leaves
	báŋá	oil palm tree	fiery red palm nut
	kùì	*Caloncoba welwitschii*	redwood
	tárà	*Anogeissus leiocarpus*	greenish-yellow leaves
	mbɛ́ɛ́	kind of squash	brown-green skin
ANIMALS	***bàn***	red duiker	red skin
	bíó	grey duiker	maroon-brown skin
	bíá	cane rat	dark gray skin
	sàí	*Tatera* rat *cf.* nigrita	light gray skin
	gɔ̀	leopard	mottled skin
	súmbùlá	village weaver	bright yellow plumage
	mbòm	kind of termite	dark brown body
BODY	*gɔ́nɔ̀*	leucorrhea	dirty yellow discharge
	wíkɔ́yá	impetigo	brown-yellow scabs
	dàì	wound	blood speckled
	tɔ̀k	blood	red
	mbùí	white hair	white
	ndáyá	birthmarks	lightened skin
OTHER	*wɛ̀ɛ̀*	fire	red
	mbɔ́rí-gèdà	pool-retting	dark gray stagnant water
	pùdú	flying ashes	light gray
	bùk	ashes	gray

These entities are well known and easily recognized by native speakers. I have noted 24 entities most of which are used in just one compound. Only two, *bàn* 'red duiker' and *kùì* 'red wood, *Caloncoba welwistchii*', appear in two compounds.[20] It is interesting to note that even though 'kaolin' *yémbé*, 'chalk' *dáŋá* (red or yellow extracted from a 'dead stone' *fìó tà*) and 'graphite' *ɗáŋ* are widely used to paint, they are not used in compound nouns. In addition, among the ten entities used as prototypes for compound color nouns (Table 1) only three are used as parts of compound nouns: two animals *bàn* 'red duiker' and *bìò* 'bush duiker', and one tree *tárà Anogeissus*

20. Marked in bold italics in Table 5.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

leiocarpus. All these entities used for color reference in compound nouns are not usually prototypes, they are just a link to the source of the color.

By using familiar entities to convey a visual impression, compounds can incorporate a color designation that is essentially linked to its source material. Gbaya speakers know to which striking visual effect they refer when using the name of a tree. Depending on the tree, the visual effect can refer to leaves, trunk, fruit, and even sometimes to only a part of it, but nothing of this is mentioned; it comes from a shared knowledge of the environment. Using *wèè* 'fire', *tɔ̀k* 'blood' or *dàl̰* 'wound' to refer to 'red' shows the complexity of these references. 'Fire' is associated with the shrub *làà-wèè* (? / fire) *'Mussaenda erythrophylla*' whose flower has a fiery red sepal used to play with babies. 'Blood' is associated with the blood red sap of the tree *tè-tɔ̀k* (tree / blood) '*Harungana madagascariensis*'. 'Wound' is associated with the bird *nɔ́ɛ́-dàl̰* (bird / wound) '*Lybius vieilloti*' which has a red head and a chest speckled with red spots. In addition, the 'red duiker' *bàn* is the color reference for the 'red-winged warbler' *ndèí-bàn* (*Cisticola* / red duiker) in the *Cisticola* genus and for the liana *súí-bàn* (swell /red duiker) which forms a red mass, while the 'red wood' *kùì* characterizes the color of a bird *zèzèŋ-kùì* (striped / red wood) and a snake *gɔ́k-kùì* (snake / red wood). It would be pointless to try to determine a specific hue for each entity because what matters is the complex visual effect in which the hue is one element among others.

This procedure generates 26 compounds (1.4% of all compounds), even less than the compounds with color adjectives.

In addition, I will mention the compound noun *sàdî-fɛ́ŋ* (animal / IPFV.whiten) which clusters 'blue duiker' and 'cane rat' in consideration of their whitish belly.

3.2 Colors in actual use

Unlike pastoralists who have a 'cattle-color terminology' to describe visual experiences in terms of color (Dimmendaal, 2015, p. 105), Gbayas who are hunter-gatherer-cultivators use extensively all the elements of their natural environment, especially animals, insects, and plants and do not have any specific terminology restricted to one of these domains. In this section, following Lucy when he notes and deplores the fact that "the actual grammar of the language being examined plays almost no role in the analysis" (Lucy, 1997, p. 333), I will pay special attention to the Gbaya grammar by looking "at the procedure in its most usual form" (Lucy, 1997, p. 332). I will first consider the plants for which I had the opportunity to write down what Gbayas say about them, during the time I had been making a herbarium. The analyzed corpus was not obtained by elicitation and shows the actual interest that the speakers have for the visual aspect, in particular, for color. I will then consider caterpillars (83 items) which are an important part of

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

their daily food and constitute a corpus where the visual aspect is essential. Finally, I will present the way in which the Gbaya talk about colors in oral production and tales which are not a specific register but use the language of everyday life.

3.2.1 *Plants*

In order to describe plants, particularly to specify the color of the different plant parts (stems or trunks, roots, flowers, leaves, seeds or pits, sap, latex, resin, inner part of the wood, fiber, or thorns), the speakers do not use verbs or nouns but systematically use adjectives in predicative constructions. It is worth pointing out that all the adjectives are used as predicates placed after the verb *ʔɔ* 'be', directly for the AAs (Example 4, 7, 8, and 9), and within a noun phrase for the nominalized PA and VAs introduced by the preposition *nɛ̀* 'as' (Table 3, Example 5 & 10).

Out of a corpus of 224 plants whose description involves color, there are only six adjectives referring to color. As shown in Table 6, the three adjectives widely used are *bú* 'white' (PA), *gbɛ́* 'red' (VA) and *tú* 'black' (VA) illustrating their role to highlight the Gbaya's partition of the spectrum in three areas, whereas references to more specific colors specified by AAs such as *fàtàfùtù* 'light beige' and *ɓàtàbùtù* 'light gray' are much less frequent.

Table 6. The predicative constructions used to describe plants

Juxtaposition	With verb *ʔɔ* 'be'	Cat.	Term		No.*
dóà nɛ̀ búù	*zàŋ téà ʔɔ́ nɛ̀ búù*	PA	*bú*	white	101
The flowers are white	The wood is white				
tɔ̀káà nɛ̀ gbɛ́ɛ̀	*wáŋáà ʔɔ́ nɛ̀ gbɛ́ɛ̀*	VA	*gbɛ́*	red	137
The sap is red	The leaves are red				
wáŋáà nɛ̀ túù	*zàŋáà ʔɔ́ nɛ̀ túù*	VA	*tú*	black	38
The leaves are black	The inside is black				
	wáŋáà ʔɔ́ fàtàfùtù	AA	*fàtàfùtù*	light beige	3
	The leaves are beige				
	zàŋ téà ʔɔ́ ɓàtàbùtù	AA	*ɓàtàbùtù*	light gray	1
	The wood is gray				

* The total number is larger than 224 since a color can be used to describe several parts of the same plant.

The contrasting use of the three adjectives PA 'white' *bú* and VAs 'red' *gbɛ́* and 'black' *tú,* satisfies most of the needs of Gbaya speakers to describe plants. However, there are some cases where the abovementioned adjectives are followed by an AA to express a more specific value (11 for 'red', 9 for 'black', and 1 for 'white'). Thus, for instance, referring to 'black' *tú*, the AA *kpɔ́ɔ́-kpɔ́ɔ́* 'very dark' is attested six times, and there are three other AAs used only once, *míkɔ́ɛ́-míkɔ́ɛ́* 'dark green', *díríríí* 'dyed black', and *bím-bím* 'dark purple'. The following AA can refer to color as previously seen, but it can also refer to another property as it is said about savannah grass:

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(13) *tàréà ʔá̰ sènέ nè búù ʔòk-sòk*
urticating.DEF BE.LOC inside as white.NMLZ skinny+noise
'There are urticating hairs white and skinny making noise.'

Finally, the term *dàp* 'pattern' appears five times. It can be associated with an AA which describes the nature of these patterns (see the list in Table 10), but not an AA of color. In this set, we have *làp-làp* 'dark with round white spots' (2 cases), *màtà-màtà* 'with big patches' (2 cases) and *làk-làk* 'with black patches' (1 case). For two lianas, we find a specification as in (14):

(14) *dàp tέ ɲàkáà ʔɔ́ màtà-màtà*
pattern.L body.L liana. L.3SG IPFV.be with big patches
'The stem of the liana has big patches.' (lit. The pattern… is)

There is also the possibility to mention a pattern by using *dàp* after the mention of a contrastive color as in (15):

(15) *zàŋ téà ʔɔ́ nè búù nè dàpáà*
inside body. L.3SG IPFV.be as white.NMLZ with pattern. L.3SG
làp-làp
dark_with_round_white_spots
'The wood is white with patterns black with round white spots.'

3.2.2 *Caterpillars and larvae*

To describe the 'caterpillars and larvae' *dɔ̀k*[21] they eat, the Gbayas use the same predicative constructions as for plants. In the description of 39 caterpillars, there are 26 occurrences of the VA *tɥ́* 'black', 12 of the PA *bú* 'white', 6 of the VA *gbɛ́* 'red', and only 2 occurrences of an AA: *zàrùà* 'vermilion red' and *fàtàfùtù* 'light beige'. Most often (30 cases) when the color is mentioned, it is also specified whether there is a pattern or not. Unlike caterpillars which have a variety of features, larvae are always white because they live inside wood, and are described as such:

(16) *ʔà̰ nè búà*
3SG as white.NMLZ
a. *dàpáà bé ná*
pattern. L.3SG IPFV.can be NEG
b. *mà dàp téà bé ná*
a certain pattern.L body. L.3SG IPFV.can be NEG
'It is white (a) with no pattern, (b) with no pattern on its body'

21. *dɔ̀k* is the generic term for 'caterpillars and larvae', but 83 species have specific names.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Caterpillars can have various colors but they also have 'patterns' *dàpáà ʔá̰ sènέ* (pattern.L.3SG / COP.LOC / inside). The caterpillar called *dɔ̌k-nàà-tá-kɔ̀ngà* is described as follows (17):

(17) *ʔà̰ nè tṵ́ṵ̀ dàpáà ʔá̰ ká zàŋáà*
3SG as black.NMLZ pattern.L.3SG cop-LOC side belly.TM.3SG
nè búù
as white.NMLZ
'It is black with white patterns on its sides.'

To express that a caterpillar has two equally important colors, the term *dàp* is not used, as shown in the description of the caterpillar *nàà-sáŋ-kàɗáŋ* in (18).

(18) *ʔà̰ nè búà ʔín tṵ́à̰*
3SG as white.NMLZ and.L black.NMLZ
'It is black and white'

Nàà-wílí-wèlɛ̀ which is predominantly black is described as in (19):

(19) *ʔà̰ nè tṵ́ṵ̀ nè búù kpá-nè tɛ́à fét*
3SG as black.NMLZ with white.NMLZ all over body.L.3SG all
'It is black with white all over its body.'

The comparison between the examples 18 and 19 shows that to associate two colors of equal importance the relator is the comitative morpheme *ʔín* 'and', while to associate one predominant color to another one, the relator is the instrumental *nè* 'with'.

Very often, to refer to the nature of 'patterns', Gbaya uses AAs that may also be combined with a color term, as in the case of the caterpillar *nàá-dà-mùr* (20):

(20) *ʔà̰ nè tṵ́ṵ̀ tɛ́à*
3SG as black.NMLZ body.L.3SG
ʔɔ́ lɛ́k-lɛ́k nè búù
IPFV.be with tiny marks with white.NMLZ
hé bòyò gá nè gɔ̀n-sèèáà
SIM iron SIM in back.L.3SG
'It is black and its body has fine white marks like iron on its back'

In the description of the caterpillar *ɓòkòyòm*, it is the color of its patterns that is mentioned:

(21) *dàpáà ʔɔ́ nè búà ʔín tṵ́à̰*
pattern.L.3SG IPFV.be as white.NMLZ and.L black.NMLZ
'Its patterns are black and white'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

As in all the previous examples, 'white' and 'black' have a contrastive value and do not express specific hues, which are instead expressed by adjective-adverbs or color names.

The caterpillar *kpòdò* is described using the N-like construction as being of the same color as the tree '*tárà*' (*Anogeissus leiocarpus* Combretaceae), which has yellow leaves.[22]

(22) *tɛ́à* *né* *tè-tárà* *dàpáà* *ʔá̰* *sènɛ́*
body.L.3SG COP yellow pattern.L.3SG COP-LOC inside
kpá *nè* *tɛ́à* *fét*
INF.PFV.find in body. L.3SG all
'It is yellow with lines all over its body'

3.2.3 *In oral productions*

Referring to colors can be relevant in a game, for example, where boys and girls sing songs featuring characters identified by their differences in skin color: *ʔó mbúlú gbɔ́ŋɔ́* 'jet-black people', *ʔó ɓɔ̀rɔ̀dîk* 'dark-skinned people', *ʔó ɓɔ́tɔ́tɔ́* 'brownish-red people', *ʔó ɓɔ́tí-ɓɔ́tí* 'dark-brown people', etc. The first of these terms is the noun phrase *mbúlú gbɔ́ŋɔ́* that translates literally as '*Cussonia barteri* soot'[23] and refers to a entity easily identified by the speakers. The other three are contextualized AAs specific to human skin that are used, by conversion, as nouns and bear the same plural marker *ʔó* as the deadjectival human nouns in Section 2.1.3.

In tales, color adjectives are used in an attributive position, preceding or following the noun. In a tale about a baboon who is ridiculed when a dog steals its clothes thus revealing its buttocks, a contextualized AA specific to the color of the baboon's buttocks is used: *vɛ̰̀ɛ̰̀* 'red'. In another tale, a hero finds one horn, which provides food, and is described as *mà gbɛ̰́ zák zóŋ-zóŋ* (some / red / horn / earth-reddish) 'a horn of reddish-earth color'. Then he finds the second one that whips him, described as *mà zɔ̀m tṵ́ṵ̀ kpɔ̰́ɔ̰́-kpɔ̰́ɔ̰́* (some / AUG / black.NMLZ / really dark) 'big really dark black one'. In the latter case, the head of the NP is the nominalized VA *tṵ́ṵ̀* 'the black' followed by the AA *kpɔ̰́ɔ̰́-kpɔ̰́ɔ̰́*.

Color adjectives are also used as predicates, such as in the tale of the picky Gbaya girl who refuses all her suitors on the grounds of their physical appearance. She mocks the blue duiker, the mongoose, and the cane rat for their 'slate gray' *tṵṵ̀*̀color; the red duiker for its 'fiery red' *zɛ̰̀ɛ̰̀* color; the bushbuck for its 'speckled' *kpùŋùŋùŋù* coat; the civet cat for its 'mottled' *ŋùŋɛ̀ŋ* fur; the buffalo for its 'gingery

---

22. *tè-tárà* < *tè tárà* (body.L *Anogeissus leiocarpus*), this procedure has been discussed in Section 2.1.2.

23. The pounded bark of the *Cussonia barteri* is used to make a black dye.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

bulk' *bèrèŋ;* and the baboon for the 'red' buttocks *vɛ̀ɛ̀*. These descriptions, all using AAs, evoke in the mind of the audience an image which is evidently of special interest to the Gbaya speakers.

Finally, color verbs are used to express a process of change such as *Canarium schweinfurthii* fruits – a kind of deep purple plum – which 'darken the ground [that becomes] very dark' *bírí tṵ́ nù kpɔ́ɔ́-kpɔ́ɔ́-kpɔ́ɔ́* (plant sp / IPFV.darken / earth / very dark.INTENSIVE), or a woman who dyes her hair 'really dark' *kpɔ́kɪ̰ɔ̰́-kpɔ́kɪ̰ɔ̰́* to attract men.

All these examples capture the circumstances under which the Gbaya put into words the visual aspects, including color, and confirm the grammatical organization that supports the distinct roles of each word category.

4. Conclusion

As this study shows, Gbaya has no generic term to refer to color but one for 'pattern' *dàp*. The visual aspect that includes color, pattern, and a lot of other elements (shape, luminosity, intensity, and even movement, distance, or time) is mainly expressed by adjectives which are either attributes or part of a predicative construction and can also function as adverbs. Two verbal adjectives 'black' *tṵ́* and 'red' *gbɛ́*, and one primary adjective 'white' *bu'* are used exclusively to express a partition of the color spectrum into three areas. They contrast with one another rather than designate specific hues or colors. On the other hand, adjective-adverbs are dedicated to the expression of property. Sixty-seven of them express a salient and unitary visual concept significant for the Gbaya speakers that can be grouped into three areas in accordance with the tripartite distribution defined by the contrastive adjectives. The other fifteen adjective-adverbs refer to patterns where color may be involved or not. Furthermore, three verbs, one for each color area identified, are polysemous: *gbɛ̰* 'redden, ripen', *tṵ* 'darken, reinforce' which has a positive meaning, and *fɛŋ* 'whiten, bleach' which has a negative one. The latter derives the adjective *fɛ́ŋá* whose resultative meaning is 'humiliated, faded', and not 'white'. *bú* is the primary adjective whose basic meaning is 'rough, natural' and it is used to denote 'white'. The fact that verbal adjectives and verbs are functionally and semantically different shows that the latter can only incidentally express color, whereas this is the basic role of the adjective category in general. Lastly, a few common nouns are used as color names. For this purpose, Gbaya uses a specific construction (body.N) to refer to an entity that functions as a color prototype. The Gbaya use of color argues for a conceptualization of the visual aspect where color is only one element among others and confirms the importance of grammatical organization to support distinct semantic domains.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

This description of the Gbaya "color system" is highly reminiscent of the notional system described by Wierzbicka for Warlpiri, a language also spoken by a hunter-gatherer society and which probably exhibits a true universal in color naming, not limited to color proper.

References

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

Dimmendaal, G. J. (2015). *The Leopards Spots. Essays on Language, Cognition and Culture*. Leiden, Boston: Bril.

Lucy, J. (1997). The linguistic of "color". In C. L. Hardin & M. Luisa (Eds.), *Color categories in thought and language* (pp. 320–346). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519819.015

Moñino, Y. (2004). Une autre conception des lumières: Sur les noms des couleurs en gbaya. In E. Motte-Florac & G. Guarisma (Eds.), *Du terrain au cognitif, Linguistique, Ethnolinguistique, Ethnosciences. A Jacqueline M. C. Thomas* (pp. 241–264). Paris: Peeters.

Roulon-Doko, P. (2001). Le statut des idéophones en Gbaya. In E. Voeltz & C. Kilian-Hatz (Eds.), *Ideophones* (pp. 275–301). Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.44.23rou

Roulon-Doko, P. (2006). Le traitement de la polysémie verbale du gbaya dans le cadre d'un dictionnaire gbaya français. In T. Szende (Ed.), *Le français dans les dictionnaires bilingues* (pp. 317–328). Paris: Honoré Champion éditeur.

Roulon-Doko, P. (2008a). Le Gbaya. In T. Holger, E. Rothmaler, & K. Winkelmann (Eds.), *La qualification dans les langues africaines*, *Qualification in African Languages* (pp. 87–101). Cologne: Rüdiger Köppe.

Roulon-Doko, P. (2008b). *Dictionnaire Gbaya-Français (République Centrafricaine), suivi d'un dictionnaire des noms propres et d'un index français- Gbaya*. Paris: Karthala.

Tornay, S. (1973). Langage et perception. La dénomination des couleurs chez les Nyangatom du sud-Ouest éthiopien, *L'Homme*, 13 (4), 66–94.

Wierzbicka, A. (2008). Why There Are No 'Color Universals' in Language and Thought, *The Journal of the Royal Anthropological Institute*, 14/ 2, 407–425. https://doi.org/10.1111/j.1467-9655.2008.00509.x

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Appendix

Table 7. The 14 AAs for 'black'

AA	Value	Used for
bìfiŋìrì	dark brown	
míkɔ̰ɛ̰-míkɔ̰ɛ̰	dark green	underside of *Solanum nigrum*'s leaves
bìm-bìm	dark purple	*Tephrosia vogelii* leaves, fly sp.
tṵṵ̀	slate gray	blue duiker
kpɔ̰kiɔ̰-kpɔ̰kiɔ̰	very dark	natural or dyeing
kpɔ̰ɔ̰-kpɔ̰ɔ̰	really dark	natural, dyeing or by rain
kárí-kárí	darken + dirtiness	body, hand, face
ndíŋ-ndíŋ	darken + opaque black	rain, hands
díríríí	darken + dyeing	
kító-kító	darken + smoke, fire	
bḭḭ̀	darkened + dirtiness	tooth
ɓótí-ɓótí	dark brown	human skin
ɓótótó	brown red	human skin
ɓɔ̀rɔ̀dìk	dark-skinned	human skin

Table 8. The 19 AAs for white

AA	Value	Used for
bóŋ-bóŋ	white	European person, *Terminalia* flower, kaolin, moldy kernels, teeth, cotton, cassava
kpúŋ-kpúŋ	white	European person, inside of the ear, head + ringworm, cassava chips, sesame flowers, chicken, birdlime
ndùyɛ̀ɛ̀	pale gray	fog, mist, ashes
fùyɛ̀ɛ̀	grayish white	fog, mist, *Imperata cylindrica*'s flowers
ɗùkùyùkù	pale yellowish gray	ashes, head + ringworm, body + flour
ɓàtàɓùtù	light gray	milk, kaolin, blue duiker, soiled by mud, badly made cassava dough
fàtàfùtù	light beige	underside of *Crotalaria pallida* leaves, striped ground squirrel, sand, white clouds, blurry vision
káɓá-káɓá	very white	foaming saliva, teeth
ndál-ndál	real white	cotton, fresh cassava flour
fàràfùrù	whiten	< sp. white clouds, blurry vision
kúlɛ́-kúlɛ́	whiten	< flour or ashes (body, animals, hands)
mbùyèlɛ̀	speckled with white	< kinds of flour or milk
vàm-vùm	whiten	< burn injury or illness (eyes or skin)
ɗúk-ɗúk	whiten	< fermentation: germinated corn, kaolin, vagina, curdled milk, smegma, some mushrooms, top of the foot

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

AA	Value	Used for
ɓùlàè	white and soft	termite queen, termites, wound, badly made cassava dough, upturned eyes
ɓàlùà	white and flabby	upturned eyes, fresh wound, badly made cassava mush
bùbùbù	large white space	with many squashes of cassava dough, cotton field
vúɗúɗú	light + negative value	sauce, beef, palms of a sick person, insult
ɓálá-ɓálá	discolored, faded, pinkish brown	human body [European person, drowned person, mycodermatitis, palms of the hands, body after bathing]

Table 9. The 35 AAs for red

AA	Value	Used for
kɔ̀sɔ̀sɔ̀	plain red	
báŋá-báŋá	fiery red	palm kernel (báŋá)
hàu̯u̯	bright red	sun, *Landolphia*
há̰t-há̰t	fire engine red	red chechia
yúngú-yúngú	blood red	horse-fly belly, *Landolphia*
zóŋ-zóŋ	red brick	foot, water + red earth, eyes, kind of *Landolphia*, iron
bèzèrè	shiny red	fire, candling an egg, eyes
ŋmèè	dull red	(spot of color without sheen) remote fire, raw sore
pài	orange-red	the setting sun, fire in the night, red moon
ngémbé-ngémbé	rusty red	ferruginous water, garment
zɔ̀rɔ̀gbɔ̀tɔ̀	dark red, reddish brown	chicken, kind of horse fly
zɛ̰̀ɛ̰̀	glowing, blazing	*Senegalia sp*., red duiker, fire
zèngèlè	glowing	fire, garment
ngbèzèlè	blazing	fire, embers, oozing wound
ngbòlòè	orange-brown	*Thonningia sanguinea* flower, mushroom-like chanterelle, lips
zàrùà zàrùwà	vermilion	fresh wound, embers, *Landolphia* / banana / ripe mangos
háŋá-háŋá	ochre	mango tree red ants, feet + red earth
ŋàà	gingery	red duiker's back, ripe banana
bàm-bòm	pinkish brown	banana, *Landolphia*, hot iron
káŋá-káŋá	reddened	+ red earth: body, tortoise
kpísí-kpísí	reddened	+ blood, tears: bloody wound, sore red eyes
kpósó-kpósó	reddened	+ plant dye: 'redwood' *Caloncoba welwistchii*, *Chrysanthellum americanum*
ŋàs	golden-colored	+ cooked or fried meat or chicken
gàà	yellowed	< modification: overcooked *Amaranthus*, dried leaves during the dry season

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

AA	Value	Used for
ngàò	scorched	< modification: grass
dùà	browned	< modification: inside of a damaged cassava tuber
ɗólóló	reddish	< modification: red instead of an original black, diseased hair, green vegetables [poor land or poorly cooked]
ngàɗàɔ̀	pink beige	< modification: healing burns, dried grasses
ngùwèŋ	pink ochre	+ external agent: healing wound, garment + red earth
yèngbèrè	pale pink	newborn baby
ɗéréré	orange-yellow, golden	fluids (urine, tea, some decoctions)
ɗólé-ɗólé	deep pink	lips
vɛ̀ɛ̀	red	baboon's buttocks
bèrìèŋ	red mass	buffalo
sóndóró	madder red	imported fabrics (military belt < F *ceinturon*, mattress ticking)

Table 10. The 15 AAs for patterns

AA	Value	Used for
làɓù-làɓù	white and black	impetigo, spotted with scabies, fire scar, raised eyebrows, flash of lightning
ŋàŋàwàŋà	beige with small dots	some leaves, leopard
ŋàŋàŋùnù	gingery with small white dots	leopard, serval
làp-làp	dark with round white spots	*Trema orientalis* leaves and trunk, civet cat, wounded body, *Haemantus* stem, guinea fowl
làk-làk	with black spots	inside of *Landolphia* sp, leopard, civet cat
mbálò	plain color +1 strip of another color	goat
tɔ̀ɓɔ̀	color alternation	banana, cassava on a red slab
mìsì-mìsì	two-colored pattern	mushroom (black and white), buffalo (black and red)
rɔ̀tè-rɔ̀tè	striated	heart fibers, garment, vines of some parasites, cornrows on head
zèrèŋ-zèrèŋ	striped	garment, lizard on cassava slab
ŋùŋèŋ	spotted	squirrel's tail, serval, leopard
kpùŋùŋùŋù	speckled	bushbuck, leopard
métɛ́-métɛ́	with very small pieces	dots, stains, holes (sieve, banana, letters)
màtà-màtà	with big spots	liana, banana, leopard, genette, *Acanthus montanus*, poule, iris des Blancs
dìr-dìr	with two-colored big spots	leprosy, oil on the body

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 11. The 29 AAs for brightness

AA	Value	Used for
ɓɛ̀ɛ̀	faint light	at dawn and dusk
mbɛ̀lɛ̀lɛ̀	first or last lights	at dawn and dusk
bùù	clear	a bright place (daylight, bared forehead, guinea folk neck)
mbáó	cleared	sky, head (after being drunk)
bòòŋ	crystal clear	open door, hole in a roof, clear sky
hàlòò	blurred	vision (anger, nervousness, alcoholic drink)
tùr-fùndùŋ	obscured, clouded	vision (fog, glare), foggy sky (rain, clouds)
ndììŋ	darken	vision, early night, evening
bìr-fìndìŋ	beginning of darkness	vision, evening
hìrìp	overshadowed	clouds, rain
fi̱i̱	dark	evening, inside the house, darkness
kɛ̀ɗù-kɛ̀ɗù	glittering (punctual)	stars, sparkling stone in sun
kpàl-kpàl	glittering (wide area)	moon, top of cassava slab
bɛ̀ɗɛ̀ŋ-bɛ̀ɗɛ̀ŋ	uniformly bright	no shadow (forehead, moon, sun)
kòl-kòl	very bright	mirror, stone
pàr-pàr	dazzling	blinding light (water, sun, steel roof, shaven head)
kɛ̀n-kɛ̀n	with metallic luster	kind of fly, cane rat
làì-làì	with specks of light	wink, raised eyebrows, flash of lightning, mirror, headache
zɛŋɛ̀ì-zɛŋɛ̀ì	with glints or bright spots	headache, garment with golden thread, vision through a mat
zɛ̀r-zɛ̀r	glowing, shiny, gleaming	because of sun or grease on a flat surface (water, cheek)
ʔɛŋɛ̀ì-ʔɛŋɛ̀ì	strongly shining	because of sun or fat on a flat surface (water, cheek)
vạ̀ạ̀	bright beam of light	fire, headlight, flashlight
ŋàì-ŋàì	flickering	fire (be periodically visible)
zàr	bright	fire, itching eyes, glazed eyes with fire
kàlò-kàlò	shining	eyes with sun or fire
zòl-zòl	shiny	eyes
zàl-zàl	sparkling, twinkling	eyes
ɓɛ́nɛ́nɛ́	clear	eyes, water
ngúlúlú	translucent, crystal clear	blind person, glance, water

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

PART II

Color terms in a genealogical and typological perspective

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Innovations in Semitic color term systems

Maria Bulakh
Russian State University for the Humanities & National Research University Higher School of Economics

The article explores the origins of Semitic color terms for yellow, green, and blue. These fundamental color categories are missing from the reconstructed proto-Semitic basic color term system, but their designations were added into basic color term systems of many daughter languages. This chapter focuses on derivations from designations of "referent objects" (objects typically characterized by a certain color and serving as referents for the designation of this color). Referent objects are further classified into natural objects (wax for yellow; vegetation, grass, or leaf for green; sky for blue) and dyes (saffron and turmeric for yellow; kohl and indigo, or indigo-colored objects for blue; the chapter also considers ink as a source for designations of blue-with-green, violet, or black).

Keywords: Semitic, Aramaic, Ethio-Semitic, Modern South Arabian, referent object, basic color terms, evolution of basic color term systems

1. Introduction

In their seminal work on the typology of color naming, B. Berlin and P. Kay (1969, p. 2) have introduced the terms Basic Color Categories (BCCs) and Basic Color Terms (BCTs). While definitions of these terms and criteria of their establishment have been elaborated and partially revised in subsequent years, these terms remain essential for research on color naming, and can and should also be applied when reconstructing the color terms of a protolanguage (cf. Biggam, 2012, pp. 20–23; 2014, p. 4). BCCs can be defined as "subdivisions of the colour space which are recognized by a particular society at a particular time", while Basic Color Terms (BCTs) are lexemes used to denote BCCs (Biggam, 2014, p. 4, fn. 1). The reconstruction of a BCT system of a protolanguage is thus a task closely related to the reconstruction of the BCC system of the speakers of this language, and the comparison between semantics of BCTs of the daughter languages may help establish the

https://doi.org/10.1075/sfsl.78.08bul
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

ranges of the respective subdivision of the color spectrum in the ancestor language, even if it is not possible to reconstruct the name of the involved BCC.

For proto-Semitic, four basic color categories can be reconstructed, namely, white, black ("macro-black"), red, and yellow-with-green. This reconstruction is based both on etymological evidence and typological data. Only for two of these categories, one can offer reliable reconstructions of the respective BCTs: **ṯ̣lm* 'to be (macro-)black', **wrḳ* 'to be yellow-with-green' (cf. Section 2). The typology of BCC systems and their evolution teaches us that yellow-with-green BCC presupposes the existence of white and red BCCs (MacLaury, 2001, p. 1240).[1] This suggests the existence of such BCCs and, consequently, of BCTs for 'white' and 'red' in proto-Semitic – despite the fact that we do not know which roots were employed to form these BCTs. The plausibility of the reconstructed four-member BCC system with a yellow-with-green category still needs to be verified (cf. Bulakh, 2007, p. 256).

In some of the daughter languages, the structure of the proto-Semitic BCT system is preserved intact: while some of the proto-Semitic roots denoting a certain color are replaced with innovative ones, their semantics did not undergo substantial changes. In the majority of Semitic languages, however, the innovations involve both the structure of the system, the color meanings and their lexical exponents. As predicted by the evolutionary typology of BCT systems, new BCTs have emerged and, as a result, the semantics of the old BCTs has been modified.

In the present investigation, I will explore the sources for some innovative color terms in Semitic languages. I will focus on the designations of fundamental color categories which are missing in the reconstructed proto-Semitic basic color term system (but were added into BCT systems of many daughter languages), that is, on terms for green, yellow, and blue.

While I am primarily interested in BCTs, for quite a few terms discussed in this study the basic status is uncertain (furthermore, some clearly non-basic terms are included in the discussion as well). For most of the languages discussed in the present contribution, comprehensive descriptions of BCT systems are entirely lacking, whereas, for the remaining few ones, only preliminary research has been conducted. For some languages, my only sources were lexicographic works (which often give no or very scanty information on the employment of the terms). Furthermore, as soon as a certain strategy of forming an innovative color term is identified, I found it useful to include it into the investigation other – even if clearly non-basic – CTs formed after the same principle.

Besides, establishing the BCTs of a given language is a complex procedure. In a recent study on the methodology of color research in a historical perspective by

1. For the elaboration on the evolutionary path offered in Berlin and Kay (1969), see also Kay et al. (1997, p. 33), Dedrick, 1998, p. 92), Kay and Maffi, 1999, p. 748), MacLaury, 2001, p. 1232), Kay et al. (2009, p. 2).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

C. P. Biggam, one finds a detailed discussion of criteria for basicness suggested by B. Berlin and P. Kay and their followers (Biggam, 2012, pp. 21–41), together with an evaluation of their validity and degree of importance. The author, however, stresses that "none of the criteria are foolproof by themselves and, also, that there is no guarantee that what appears to work for one language will work for another" (Biggam, 2012, p. 41). Furthermore, criteria for basicness involve a number of graduated parameters, i.e., a CT can be more basic and less basic in comparison to other BCTs (cf. Kerttula, 2007).

In the present study, I have made use of only those criteria collected in Biggam (2012) for which I could collect the necessary data (and some of these criteria could only be applied to some languages or some terms). Obviously, my evaluation of basicness of the CTs under scrutiny is preliminary and requires further verification. Below, a brief summary of the relevant criteria is given.

1. Non-predictable character: the meaning of a BCT is not expected to be synchronically predictable from the semantics of its parts.
2. Hyponymy: the meaning of a BCT is not fully covered by the meaning of another color term.
3. Contextual restriction: a BCT cannot be applied to a limited group of objects only. This criterion has been further elaborated by Rakhilina and Paramei (2011), who show that a BCT must be applicable to natural phenomena, as well as human artifacts (whereas newly coined color terms are often restricted to descriptions of artifacts).
4. Elicited lists: a BCT is expected to appear among the words listed by native speakers when asked to give color names of their language.
5. Derivational morphology: a BCT is expected to have more derivatives than a non-basic color term.
6. Polysemy/homonymy: a color term which is also used as a name of a concrete object with the same color is less likely to be a BCT than a color term which exhibits no such homonymy/polysemy. In Biggam's elaboration, one should distinguish here between homonymy and polysemy: a BCT can be homonymous with an object name (as in English, *orange* 'color orange' vs. *orange* 'a kind of fruit'), but if the association between the name of the object and its color is strong enough synchronically, one should speak of a polysemous lexeme which is then unlikely to be a BCT. Polysemy of the type 'related to X' and 'X-colored' (cf. Rakhilina & Paramei, 2011, p. 126) must also belong here.
7. Recent loanwords: borrowings which have recently entered the language are less likely to be BCTs (this criterion is to be used with a great deal of caution since it remains unclear how much time a loanword needs to become fully integrated and enter the BCT system of a given language).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

8. Morphological complexity: a BCT is more likely to be morphologically simple than morphologically complex. Biggam observes that this is a rather vague criterion on Berlin and Kay's list, but she offers practically no additional explanations how this criterion should be applied. Apparently, a BCT is expected to consist of the root and the inflectional morphemes, whereas a non-basic term can have a more complex structure.
9. Expression length: long terms are less likely to be BCTs than short terms.
10. Frequency of occurrence in texts and in speech: BCTs are expected to occur more often in texts and in speech than non-basic terms.

In addition to these, I tentatively used one more criterion: if a language contains two or more color terms, with more or less the same range of reference and combinability, they are all suspect as non-basic.

All considered terms can, at best, be regarded as new members of BCT systems and are thus likely to be less basic than the terms for 'white', 'black', and 'red'. They are, therefore, not expected to meet *all* the criteria of basicness. In particular, all terms considered here are etymologically transparent and hence, some problems arise concerning the criterion of non-predictable character.

This contribution is based on my Ph.D. thesis (2005), which was never published (although various parts of it were published as separate articles). Some of the conclusions drawn in the 2005 work need revision, and the present investigation provides me with this opportunity.

BCT systems of modern Arabic vernaculars deserve a special investigation, going beyond the scope of the present contribution. Much progress in this domain has been made thanks to A. Borg, whose works (1999, 2007, 2011a, 2011b, 2014) contain a rich collection of relevant data, supported by a thorough synchronic and diachronic analysis.

2. Proto-Semitic basic color terms and subsequent developments of Semitic BCT systems

In my Ph.D. thesis (Bulakh 2005; cf. Bulakh, 2007, pp. 253–257), the proto-Semitic basic color term system is reconstructed as consisting of four members: **lbn* 'to be white', **ṯlm* 'to be dark, (macro-)black', **ʔdm* 'to be red', **wrḳ* 'to be yellow-with-green'. As observed in Section 1, only two of these terms–**ṯlm* and **wrḳ*–are safely reconstructed on the proto-Semitic level. The evidence supporting the proto-Semitic basic status of **ʔdm* 'to be red' is much less convincing, whereas **lbn* lacks cognates in Akkadian and hence, can be reconstructed as a proto-West-Semitic BCT (but not as a proto-Semitic one).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ascribing the basic status and the particular color semantics to *ṯ̣lm and *wrḳ and hence, ascribing the basis status to the corresponding categories, is, of course, disputable. As C. Biggam (2014, p. 14) rightly observes, "since the tests for basic status depend on questioning native speakers and/or investigating texts… assessing potential prehistoric BCTs will never be a securely based procedure".

In L. Kogan's discussion (2015) on the proto-Semitic Swadesh wordlist, only *wrḳ is mentioned as a reliable proto-Semitic BCT (pp. 27–46). However, for Kogan, the meaning of this term in proto-Semitic is simply 'green', despite the fact that the semantics of 'yellow-with-green' is supported both by the evidence within Semitic and by the typology of BCT systems (cf. Bulakh, 2003, pp. 8–10; 2006a, pp. 204–211; Borg, 2014, pp. 40, 46–47). Furthermore, the possibility that *wrḳ was a brightness term that spread over the yellow and light-green parts of the spectrum cannot be ruled out (cf. Bulakh, 2007, p. 256). However, the fact that its reflexes in several branches of Semitic function as BCTs for 'yellow-with-green' speaks in favor of the reconstruction of *wrḳ as a hue term covering both yellow and green parts of the spectrum: it seems improbable that the yellow-with-green category emerged in several genetically related (but geographically not adjacent) languages independently.

While the proto-Semitic status of *ṯ̣lm is indisputable (Kogan, 2015, p. 237; cf. Bulakh, 2003, pp. 5–7; 2006b, pp. 738–740), the fact that in several daughter languages its reflexes are not members of the BCT system and exhibit the semantics of 'dark' rather than 'black', makes Kogan reluctant to reconstruct the latter meaning as proto-Semitic. However, it is essential for the discussion of the original semantics of *ṯ̣lm that a BCT for 'black' at an early stage of the BCT system evolution was used to denote a wider color range than simply black (at first being applied to green and blue hues as well). Such a term is better defined as "macro-black", or "cool-dark" (cf. Kay & McDaniel, 1978, p. 637; MacLaury, 2001, p. 1236). The proto-Semitic *ṯ̣lm (in the absence of a blue category) most likely included dark blue and probably also dark green (while light green was covered by *wrḳ; on the meaning of 'lightness, paleness' in the semantics of the proto-Semitic *wrḳ cf. Bulakh, 2007, p. 256).[2] In principle, Kogan's "dark" can also be understood as "macro-black"; the essential point, however, is that the function of the proto-Semitic term differed drastically from the modern English *dark* – the former

2. Incidentally, the reconstruction of such a meaning for the proto-Semitic *ṯ̣lm makes it semantically close to Old Arabic *ʔaḥḍar-* (cf. Borg, 2014, pp. 40–44) and makes it probable that it was *ʔaḥḍar-* that ousted the reflex of the proto-Semitic *ṯ̣lm as a designation of the "macro-black" category, followed by the introduction of *ʔaswad-* 'black' and splitting of the category into 'green-with-blue' (denoted by *ʔaḥḍar-*, which later narrowed its meaning to 'green') and 'black' (denoted by *ʔaswad-*).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

was the basic term covering the dark hues and the achromatic black, whereas the latter is a modifier of proper color terms.

Within this approach, **ṯlm* was a part of the proto-Semitic BCT system, which well agrees with its functioning as a BCT in two branches of Semitic – Akkadian and Ethio-Semitic. In the rest of Semitic languages, it was ousted by innovative terms, often being preserved in derived nouns and verbs with the meaning 'darkness'/'to be dark (of night)' (cf. especially the Arabic reflex of **ṯlm*: *ẓaluma*/*ʔaẓlama* 'to become dark (of night)', Lane, 1863–93, p. 1921; for darkness of the night as a prototype for 'dark' cf. Wierzbicka, 1996, pp. 301–303).

All known BCT systems containing 'yellow-with-green' have also BCTs for '(macro-)black', 'white', and 'red' (cf. MacLaury, 2001, p. 1240). It was this consideration, coupled with the thought that the original proto-Semitic terms could not have disappeared without a trace from all the daughter languages, that forced me to ascribe proto-Semitic status to the terms **lbn* 'to be white' and **ʔdm* 'to be red' in my 2005 work. In reality, while the inclusion of 'white' and 'red' categories into the proto-Semitic BCC system remains the most plausible solution, it is uncertain which terms designated these categories.

What happened in the transition from proto-Semitic to daughter languages? Predictably, in most branches, the BCT systems became more elaborate, with new terms added and meanings of the old terms readjusted. In particular, the "green-with-yellow" category was preserved only in a few daughter languages (Akkadian; probably also Ugaritic; the evidence on the continental Modern South Arabian languages remains to be verified), whereas, in most groups of Semitic languages, separate "yellow" and "green" categories emerged. Interestingly, in Geez, the old 'green-with-yellow' term was replaced by a BCT for 'green' (which was not used to refer to yellow hues), while no special BCT for 'yellow' was employed (cf. Bulakh, 2007, pp. 251–252; a similar situation is observed in a number of modern Ethio-Semitic languages).[3] In ancient Arabic, the "yellow-with-green" category was abandoned, but another composite category emerged, namely, "green-with-blue", denoted by *ʔaḥḍar-* (which in Classical Arabic became a BCT for 'green'; cf. Fischer, 1965, p. 306; Borg, 2007, pp. 282–283; Borg, 2014, pp. 36–37; cf. also above, fn. 2).[4]

---

3. In Bulakh (2007, pp. 251–252) an attempt was made to explain this narrowing of the original four-member BCT system into a three-member system by means of reconstructing a primarily brightness system for proto-Semitic. A simple solution is to attribute this change to the influence of Cushitic substratum: the Cushitic speakers, shifting to Ethio-Semitic languages, could take over their own BCT systems.

4. Contrary to Borg (2007, p. 282), Arabic *ʔaḥḍar-* is hardly a reflex of the proto-Semitic term for 'green-with-blue', since its comparison with Akkadian *ḫaṣartu*/*ḫaṣertu* is unwarranted (cf. criticism in Kogan, 2001, pp. 272–273).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The "green-with-blue" category is also observed in speech of some Soqotrans,[5] although a separate term for 'blue' is apparently on its way to being introduced into the BCT system of Soqotri (cf. Section 4.3).[6]

3. Sources for innovative CTs

A diachronic study of CTs in a given language family provides a fairly rich collection of innovative CTs of various origins. They can be classified as follows: (1) transparent morphological and semantic derivations from the source lexemes which are designations of "referent objects" (objects typically characterized by a certain color and serving as referents for the designation of this color), (2) results of other semantic shifts,[7] (3) borrowings, (4) terms of uncertain origin.

Due to spatial restrictions imposed by this chapter, I will consider only the terms of the first type. In the two subsequent sections, I will discuss color terms for 'yellow', 'green', and 'blue', etymologically going back to names of objects typically characterized by these colors (Section 4), or to names of dyes yielding these colors (Section 5). For completeness's sake, I will also discuss a few borrowed CTs which are derived from the names of referent objects in the donor languages.

Color terms derived from names of objects associated with certain colors are, of course, widespread cross-linguistically (cf. Kerttula, 2002, pp. 242–251 for English; Rakhilina & Paramei, 2011, p. 125 for Russian; cf. also Wierzbicka, 1996, p. 289). Such a derivation usually leads to the emergence of a non-basic color term whose range of denoted hues and focal color are determined by the typical color(s) of

---

5. Denoted by *šéžhar*, possibly a cognate of Arb. *ʔaḥḍar-* (on phonological difficulties of comparison between the Arabic and MSA color terms cf. Bulakh, 2004, pp. 276–277).

6. For the rest of MSA, various lexicographic sources differ in defining various color terms: for instance, Mhr. *həḍáwr* is defined as 'yellow-with-green' in Johnstone, 1987, p. 163 and as 'green-with-blue' in Nakano, 1986, p. 125. Obviously, more fieldwork is necessary. The existence of a cover term for yellow, green and blue in Harsusi (*heźōr* 'yellow, blue, green', cf. Johnstone, 1977, p. 54) likewise remains to be verified (this development reminds the extension of the original 'yellow-with-green' BCT into the blue hues observed in some modern Aramaic vernaculars, which also requires a special investigation (cf. Borg, 2014, pp. 44–45).

7. E.g., brightness term > proper color term. Cf. Kerttula (2002, pp. 242–244) on English color terms derived from brightness terms. Cf. Borg (2007, pp. 269–271) on the etymology of the Aramaic BCTs for 'white' (**ḥwr*), and Borg (2011a, pp. 85–87, 2011b, pp. 26–27) on the etymology of the Classical Arabic BCT for 'blue', *ʔazraq-*. In both cases, the original semantics of brightness is convincingly reconstructed. The semantic shift can also be observed within the lexical field of color proper, cf. *ʔašhab* 'grey, blue' in the Arabic dialect of Negev Bedouin, with the original semantics of 'grey' extended into 'blue' (Borg, 1999; 2011, p. 14).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

the referent object. However, it may eventually acquire the status of a basic one (cf. Rakhilina, 2007, pp. 374–375; Rakhilina & Paramei, 2011, pp. 129). This process can be accompanied with broadening of its semantics and loss of the original close connection with the name of the referent object. Thus, the Russian BCT for 'brown', *koričnevyj*, etymologically derived from *korica* 'cinnamon', is not anymore synchronically perceived as a designation of the specific color of cinnamon (cf. Rakhilina & Paramei, 2011, pp. 126–127). This is why it is important to distinguish between a diachronic referent object and a synchronic prototype, which may coincide or not: for instance, blood is one of commonest prototypes for red color, whereas BCTs for 'red' may be etymologically traceable to terms for 'blood' or not (cf. Wierzbicka, 1996, p. 314).

The association of a certain object with a certain color may be universal or regional (cf. Wierzbicka, 1996, p. 289). Furthermore, the same color of the same object can be described with different color terms in different languages (for instance, the color of sky, obviously blue for the Europeans, is described with the term for '(bright) white' by the Bedouins of Negev, speakers of an Arabic vernacular, cf. Borg, 1999, p. 133; the typical color of sea in Soqotri language is black according to p. c. from Dr. L. Kogan). Hence, the semantics of a newly derived color term is not fully predictable, and the object names with more or less identical meanings may yield designations of different colors in different languages. For instance, the BCTs for 'red' in the Modern South Arabian languages are traceable to the noun **ʕapar-* 'earth' (cf. Bulakh, 2004, pp. 274–276). One may hypothesize that red was perceived as the typical color of earth in the region in question (a confirmation can be seen in such expressions as *mátfəl' ʕáfer*, lit. 'red plastering', used as an euphemistic designation of grave in Naumkin et al., 2014, p. 206, Text 10, line 9; on the semantics of non-basic color terms derived from the same noun elsewhere in Semitic cf. Bulakh, 2004, p. 274, fn. 3). Furthermore, the derived color term can change its semantics in the course of time (cf. e.g. the application of *samāwi* in Harari discussed under 4.3).

As for the names of dyes, they are usually viewed as a subtype of referent objects (cf. e.g. Rakhilina & Paramei, 2011, p. 125). However, a subtle difference exists: the color is named not after the color of the dye itself, but rather after the color of the objects dyed with it. Obviously, the strategy of the formation of a color term is different: the original meaning of the adjective is 'dyed with the dye X', which by contiguity is extended into 'resembling (in its color) an object dyed with the dye X' (for instance, color terms derived from designations of *Kermes vermilio*, an insect used for production of a red dye, are consistently used to denote the type of vivid red produced by this dye, not the color of the insect). A color term can also be formed from a name for an object typically dyed with a certain dye (or with a dye of a certain color), as shown in 5.2.3.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4. Derivations of color terms from names of natural objects

4.1 'Wax' > 'yellow'

In some modern Aramaic dialects, one finds terms for 'yellow' (their basic status remains to be verified) derived from the name of 'wax':

> Tur. *šaʃūʁo* 'yellow' (Kuipers, 1983, p. 59, p. 292; Borg, 2004, p. 84; 2011a, p. 77; *shâʃûthâ* 'yellow', Maclean, 1901, p. 310); Tlk. *šaʔūʁa* 'blonde; yellow' (Kuipers, 1983, pp. 59, 262). Cf. Tur. *šʃûtâ* 'beeswax; balm' (Kuipers, 1983, p. 59).

Among other Aramaic languages, we find a similar strategy in Syriac, where several non-basic CTs for 'yellow' are derived from *šəʃōtā* 'wax': *šəʃōtā* 'colore cerae, flavus', *šəʃōtāyā* 'flavus', *šəʃōtānāyā* 'flavus', *šəʃōtwatā* 'color flavus' (Brockelmann, 1928, p. 792, Sokoloff, 2009, p. 1582).

4.2 Various types of vegetation (grass, vegetable, leaf) as sources for 'green'

4.2.1 *Derivatives from* *ḥaml *'vegetation' in Ethio-Semitic*

The BCT for 'green' in Classical Ethiopic (Geez) is *ḥamalmil*, *ḫamalmil* 'green, verdant' (Leslau, 1987, p. 233; Dillmann, 1865, p. 71; cf. Bulakh, 2006b, pp. 741–742). Morphologically, it takes the pattern with partial reduplication of the root ($C_1aC_2aC_3C_2iC_3$), which is a productive adjectival pattern in Geez (cf. Tropper, 2002, p. 68). Note, however, that the rest of Geez BCTs take the pattern $C_1aC_2iC_3$ from which the reduplicated pattern can be derived (according to A. Dillmann's lexicon of Geez, such derivatives have the semantics of a low degree of a certain color, cf. Dillmann, 1865, pp. 457, 1307). The morphology of the terms for 'white', 'black', and 'red' is thus simpler than that of the term for 'green', which suggests a less basic status of the latter (for the criterion of morphological complexity cf. Section 1). The feminine is expressed by the replacement of *i* with *ā*: *ḫamalmāl*.

Comparable color terms are found in other Ethio-Semitic languages. In Tigre and Tigrinya, modern Ethio-Semitic languages linguistically the closest to Geez, one finds non-basic color terms which exhibit the same morphological structure: Tgr. *ḫamalmil* (Littmann & Höfner, 1962, p. 59), Tna. *ḫamälmal* (Kane, 2000, p. 170). Since the reduplicated pattern is productive neither in Tigre, nor in Tigrinya, one has to assume early loans from Geez (the alternative and less plausible solution is to treat the Tigre and Tigrinya color terms in question as the sole relics of a morphological pattern once shared by the North Ethio-Semitic languages, which eventually fell out of use in modern languages).

The Tigre term does not appear on the lists of color terms in various Tigre dialects collected by Saleh Mahmud Idris (2015, pp. 241, 259). In the dictionary

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

by E. Littmann and M. Höfner, based on the Publications of Princeton Expedition by E. Littmann (1910–1913), the term is rendered as 'brown' (Littmann & Höfner, 1962, p. 59), while in the dictionary of modern Tigre by the Eritrean scholar Musā ʔĀron (2005, p. 21) it is defined as the color of vegetables and grass. Both definitions find confirmation in the textual employment. For the first meaning, compare the two descriptions of brown-skinned girls in two poems collected by E. Littmann (1910–1913, vol. III, L. 339, lines 32–35 and L. 61, line 2; in the latter verse compared with a brown cow; the feminine form of the adjective, *ḥamalmāl*, can be used as a cow name, lit. 'brown', cf. Littmann, 1910–1913, vol. II, p. 216). The meaning 'grass-colored' likely appears in Littmann's Publications (1910–1913, vol. III, p. 471, line 13): *ḥamalmāl sirā ramādi raʔasā fāgər mən ʕəčč̣ay* 'the dark-brown *sirā*-grass, *ramādi*-grass, her head is going out over the trees' (the translation 'brown' given in Littmann's publication may be imprecise, perhaps to be corrected to 'dark-green' or just 'green').

The meaning 'brown' is connected with the employment of the term as a member of the subset of human skin colors, where it designates an intermediate color between light brown and black. This subset, employing color terms with the original semantics of 'red', 'green', and 'black' (cf. Munzinger, 1864, p. 143, referring either to Tigre, Tigrinya, or both), apparently emerged under the influence of Arabic, where three main types of human complexion are designated by BCTs for red (of highest lightness), green, and black: *ʔaḥmar-*, *ʔaḥḍar-* and *ʔaswad-* (see Fischer, 1965, p. 338). One can suspect that the Tigre terms were relatively widespread in the past, hence a loanword from Tigre appears in Bilin: *hamalmíl* 'braun' (Reinisch, 1887, p. 195). The employment of a BCT for 'green' as a member of a subset of cattle color (with a shift from designation of green hue into brown or achromatic colors possible in such a context) may be an internal development, parallel to similar shifts elsewhere in Semitic (for Soqotri, cf. Naumkin – Porkhomoski, 1981, pp. 28–29).

The Tigrinya color term *ḥamälmil* was used, according to Th. Kane's dictionary (Kane, 2000, p. 170), similarly to its Tigre cognate, to describe verdant plants and skin color ('green, e.g., plant; brown (skin'), cf. *ḥəmälmälay* 'non troppo nero, di colorito di persona' in Bassano, 1918, p. 38). However, it is also said to have the meaning 'spotted, speckled (animals' coat)', and only this meaning appears in the Täkkəʔä Täsfay's dictionary of modern Tigrinya (Täkkəʔä Täsfay, 1999, p. 62). Furthermore, the term seems to be of limited use in modern Tigrinya: it is practically unknown to my informant from Eritrea, Täsfaldät Ḥadgämbäs (born in ʕAddigrat resident of Asmära).

On the contrary, the adjective *ḥamlay* (Kane, 2000, p. 170; Täkkəʔä Täsfay, 1999, p. 64, alongside the variant *ḥamläway*), which is another derivative from the root *ḥml*, is actively employed as a term for 'green' in Tigrinya (although its basic status is questionable). In Zelealem Leyew's list (2016, p. 63; based on field data collected from two informants, region unknown), this is the only term for 'green'. My informants from Eritrea (Dr. Abbäbä Kəfläyäsus from Asmära; Täsfaldät Ḥadgämbäs) named

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

the same term alongside the more widespread *ḳäṭälya* 'green' (cf. Section 4.2.3). The adjective *ḥamlay* seems to be used typically in contexts referring to abundant vegetation (e.g., *ʔab gize ḥaggay mosko ḥamlay təḵ̆əyyär* 'In summer Moscow becomes green'), but it is can also refer to artifacts (*ḥamlay kədan* 'a light-green dress'). For Dr. Abbäbä Kəfläyäsus, *ḥamlay* denotes the lighter shades of green, whereas *ḳäṭälya* is used for the darker hues. My informant from Tigray, Kidanä Gäbräḥəywät (born in Aksum), did not name *ḥamlay* among the color terms for green.

In Amharic, the only comparable terms are Geez borrowings *hamälmal* 'green, verdant' (Kane, 1990, p. 7; Täsämma Habtä Mikaʔel & Käśśate Bərhan Täsämma, 1958/59, p. 55), *hamälmil* 'green, verdant' (Kane, 1990, p. 7), 'spotted, striped, multicolored' (Täsämma Habtä Mikaʔel & Käśśate Bərhan Täsämma, 1958/59, p. 55; Kane, 1990, p. 7). In modern language, according to Dr. Getie Gelaye (a native speaker of Amharic), these terms are practically unused, with *hamälmal* being known as a female name only.

In the East Gurage dialects Selti and Wolane, we find color terms for green derived from the same root: Sel. *aml dänä* (Wedekind & Wedekind, 2002, p. 6), Wol. *ʔaməl dänä* (Meyer, 2006, p. 137). Both terms are collocations of the name of the referent object, 'cabbage' (Sel. *aml*, Wol. *aməl*) and the noun *dänä* 'aspect, appearance, color' (Leslau, 1979, vol. III, pp. 48, 210). Both CTs are most probably non-basic (they are semantically predictable, morphologically complex and, at least for the Wolane term, a competing term with the same semantics is attested, *ḳuṭäl ʔaynät*; see Section 4.2.3).

The Selti and Wolane terms for 'cabbage' go back to the proto-Ethio-Semitic **ḥaml* 'vegetation; cabbage':

Gez. *ḥaml* 'vegetation, vegetables, herbs, shrubs' (Leslau, 1987, p. 233; Dillmann, 1865, p. 70);

Tgr. *ḥamle* 'vegetables, herbs, grass' (Littmann & Höfner, 1962, p. 59; Musā ʔĀron, 2005, p. 22);

Tna. *ḥamli* 'edible herbs, vegetables, cabbage or mustard greens' (Kane, 2000, p. 169; Täkkəʔä Täsfay, 1999, p. 63), 'brassica campesteris var rapa, *B. carinata*' (Wolde Michael Kelecha, 1987, p. 173);

Arg. of Aliyu Amba *haməl* 'cabbage' (Leslau, 1997, p. 205; Girma Demeke, 2013, p. 447), Arg. of Ṭollaha *ḥaməl* id. (Girma Demeke, 2013, p. 447);

Har. *ḥūl* 'kind of cabbage' (Leslau, 1963, p. 82);

Gaf. *amlä* 'chou' (Leslau, 1956, p. 177);

Zay Wol. *aməl*, Sel. *aml*, Sod. *ambəl*, Msq. Eža Čah. Gyt. Enm. *ambər*, Gog. Muh. *ambi* 'cabbage' (Leslau 1979, vol. III, p. 48, cf. also Gutt & Hussein Mohammed, 1997, p. 404 for Sel. *aml*/*haml*).

Amh. *haməl* 'various kinds of leaves; vegetables' (Kane, 1990, p. 7) is a borrowing from Geez.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

There is little doubt that the BCT *ḥamalmil* in Geez, the apparently non-basic color term *ḥamlay* in Tigrinya, and the terms for 'green' in Wolane and Selti are independently derived from the same root. While vegetation *per se* is of course a typical prototype of green color (cf. Wierzbicka, 1996, p. 306), one has to remember that the source of derivation in East Gurage (Wolane and Selti) has a narrower meaning 'cabbage' (whose functioning as a prototype for green color is a less trivial feature). The fact that the reflexes of the same noun **ḥaml*, although with different denotata, function as referent objects for the same color in geographically remote languages, suggests a strong association between the term **ḥaml* and the green color already in proto-Ethio-Semitic. The semantic narrowing from 'vegetation' to 'cabbage' in South Ethio-Semitic probably reflects the perception of cabbage as a typical representative of vegetation in the region, hence its continuing function of the referent object for green. In modern South Ethio-Semitic languages, cabbage can indeed be used as a referent object for (a certain shade of) green, as illustrated by the Amharic expression *gomän aynät* 'cabbage-like' (i.e., 'green'), adduced by Zelealem Leyew (2016, p. 64).

4.2.2 *Derivatives from the referent object *śaʕar 'grass'*

In Tigre, the term *saʕarsaʕaro* is practically the only term for 'green' appearing in the dictionary by Littmann & Höfner (1962, p. 194, cf. also Musā ʔĀron, 2005, p. 81). However, it is absent from the corpus of the Publications of the Princeton Expedition (in Littmann & Höfner [1962, p. 194], it is only quoted as a term occurring in Tigre wordlists collected by the Swedish missionary G. Sundström). According to Saleh Mahmud Idris's survey of Tigre dialects (2015, pp. 241, 259), this term (or its variants) is used in the varieties of Mansaʕ, Beni Amer (cf. also Nakano, 1982, p. 122), Māryā Ṣallām, and in Samhar (including Gindaʕ), in some of them alongside a concurrent term with the same meaning, such as the Arabic borrowing *ʔakdar* (< Arb. *ʔaḥḍar-*). Thus, several criteria (predictable character, morphological complexity, non-occurrence in the available corpus of Tigre oral literature, restrictedness to some dialects only) suggest its non-basic status.

The source of this color term is obvious: it is the reduplication of the well-known noun *saʕar* 'grass, hay' (Littmann & Höfner, 1962, p. 194; Musā ʔĀron, 2005, p. 81). The cognates of this noun are used to derive terms for 'green' in a number of Gunnän-Gurage languages, all of them being compound adjectives (formed as relative constructions with the verb **masala* 'to resemble, be like, look like' [cf. Leslau, 1979, vol. III, p. 428], governing the noun 'grass', lit. 'resembling grass'):

Muh. *sär yəmäsyu* (Leslau, 1979, vol. I, p. 930);
Eža *särimäsər* (Leslau, 1979, vol. III, p. 560);
Čah. *särimäsər* (Leslau, 1979, vol. III, p. 560);

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gyt. *säʔär yəm̱äsər* (Leslau, 1979, vol. I, p. 715).

The respective terms for 'grass' are reflexes of proto-Ethio-Semitic **śaʕar* 'grass (green or dried)':

Gez. *śāʕr* 'grass; straw' (Leslau, 1987, p. 525; Dillmann, 1865, p. 259);

Tgr. *saʕar* 'grass, hay' (Littmann & Höfner, 1962, p. 194; Musā ʔĀron, 2005, p. 81);

Tna. *saʕri*, *saʕari* 'grass, hay, straw' (Kane, 2000, p. 763; Täkkəʔä Täsfay, 1999, p. 250);

Amh. *sar* 'grass (either green or dried, i.e., hay or straw)' (Kane, 1990, p. 482);

Arg. *səʔər* 'grass' (Leslau, 1997, p. 219), Arg. of Aliyu Amba *sir* id., Arg. of Ṭollaha *säʕər* id. (Girma Demeke, 2013, p. 151);

Har. *säʔar*, *sār* id. (Leslau, 1963, p. 136);

Zay *sār*, Sel. *sār*, Wol. *sar* id. (Leslau, 1979, vol. III, p. 530; cf. also Gutt – Hussein Muhammed, 1997, p. 144, for Sel. *sār*);

Gaf. *sar* id. (Leslau, 1956, p. 232);

Sod. *sar*, Gog. Muh. Msq. Eža Čah. *sär*, Gyt. *säʔär*, Enm. End. *saʔar* id. (Leslau, 1979, vol. III, p. 530).

Beyond Ethio-Semitic, comparable terms are found in Arabic and Modern South Arabian:

Arb. *šaʕar-*, *šuʕr-* 'plants and trees' (Lane, 1863–1893, p. 1560; Biberstein Kazimirski, 1860, vol. I, p. 1238), *šaʕār-* 'trees; tangled, luxuriant, abundant and dense trees' (Lane, 1863–1893, p. 1561; Biberstein Kazimirski, 1860, vol. I, p. 1238);

Mhr. *śēr* 'straw' (Johnstone, 1987, p. 370);

Jib. *śáʕər* 'dry grass, straw' (Johnstone, 1981, p. 244).

Note that the meaning 'hay, straw' in MSA seems to be the narrowing of the broad semantics present in the Ethio-Semitic cognates, which can denote both fresh and dried grass. Nevertheless, the Ethio-Semitic **śaʕar* obviously typically referred to fresh, green grass, which explains its employment as a referent object for green in several Ethio-Semitic languages. Since grass as a prototype and a referent object for green is a widespread phenomenon (cf. Wierzbicka, 1996, p. 306), there is no need to explain the formation CTs for 'green' derived from this noun in different branches of Ethio-Semitic as a retention from the proto-Ethio-Semitic stage. They may well be independent of each other (although the factor of language contact may have played a certain role in the diffusion of this strategy).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.2.3 *Derivatives from the referent object* *ḳaṣl *'leaf'*

According to my informants for Eritrean Tigrinya (Dr. Abbäba Kəfläyäsus, Täsfaldät Ḥadgämbäs), the most typical designation of green is *ḳäṭälya* (Kane, 2000, p. 1044), despite the existence of the alternative term *ḥamlay* (cf. Section 4.2.1). In Täkkəʔä Täsfay's dictionary (1999, p. 312), *ḳäṭälya* and *ḥamlay* are treated as synonymous, although according to Dr. Abbäba Kəfläyäsus, *ḳäṭälya* is used to denote darker hues of green, and *ḥamlay* the lighter hues.

My informant from Tigray (Kidanä Gäbräḥəywät, born in Aksum) named the terms *ḳoṣli* and *ḳoṣṣal* as equivalents of English *green*, although the basic status of these terms remains unclear (cf. *ḳoṣṣal* 'green', Täkkəʔä Täsfay, 1999, p. 338; *ḳoṣlay*/*ḳoṣläway* 'leafy, green'; Kane, 2000, p. 1059; the variant *ḳuṣlay* is also mentioned as a non-basic term for 'green' in Zelealem Leyew, 2016, p. 64).

Comparable color terms elsewhere in Ethio-Semitic include Amh. *ḳəṭäläyya* 'green, olive-green, greenish, leaf-like', *ḳəṭälämma* 'forest-green', *ḳəṭəlṭal* 'green, olive-green, chartreuse; spotted, piebald (horse)' (Kane, 1990, p. 830; cf. also Dästa Täklä Wäld, 1970, p. 145; Kidanä Wäld Kəfle, 1955–1956, p. 449), Gog. *ḳəṭäläyä* 'green, kind of snake' (Leslau, 1979, vol. III, p. 508), Wol. *ḳuṭäl ʔaynät* 'green' (Meyer, 2006, p. 138). None of them seems to function as a basic color term: the Amharic BCT for 'green' is *aräng^{w}ade* of uncertain origin (cf. Vassilevich, 1988, pp. 144, 146; Zelealem Leyew, 2016, p. 63); the Wolane term is semantically predictable and morphologically complex, with a recorded competing term for 'green', *ʔaməl dänä* (cf. Section 4.2.1); the term in Gogot may well be an Amharism, with a recorded competing term for 'green', *aräng^{w}ade* (cf. Leslau, 1979, vol. III, p. 89). However, Amh. *ḳəṭäläyya* must have been rather widespread, being the source of borrowing for Tna. *ḳäṭälya* (contrast the above-mentioned genuine Tigrinya terms *ḳoṣṣal*, *ḳoṣlay*/*ḳoṣläway*, *ḳoṣli*, which exhibit *ṣ* rather than *ṭ*) and probably also for Gog. *ḳəṭäläyä*.

These color terms are clearly derived from the terms for 'leaf', going back to proto-Ethio-Semitic **ḳVṣl* with the same meaning (Kogan, 2015, p. 447):

Gez. *ḳ^{w}aṣl* (Leslau, 1987, p. 450; Dillmann, 1865, p. 472);
Tna. *ḳ^{w}äṣli* (Kane, 2000, p. 1059; Täkkəʔä Täsfay, 1999, p. 338);
Amh. *ḳəṭäl* (Kane, 1990, p. 829);
Arg. of Aliyu Amba *ḳəṭäl* (Leslau, 1997, p. 217; Girma Demeke, 2013, p. 198), Arg. of Ṭollaha *ḳəṭäl* (Girma Demeke, 2013, p. 198);
Zay *ḳəṭe*, Sel. Wol. *ḳuṭäl* (Leslau, 1979, vol. III, p. 508);
Gaf. *ḳəṭälä* (Leslau, 1956, p. 226);
Muh. *ḳəṭe*, Sod. Gog. Msq. *ḳəṭäl*, Eža Čah. Gyt. *ḳəṭär*, Enm. End. *käʔär* (Leslau, 1979, vol. III, p. 508).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.3 'Sky' > 'blue'

In Ethio-Semitic languages, one encounters a number of color terms for 'blue' derived from the referent object 'sky'. This strategy is, of course, well-known cross-linguistically (cf. Wierzbicka, 1996, p. 309); it is also widespread in non-Semitic languages of Ethiopia (cf. Zelealem Leyew, 2016, pp. 67–68). The source noun, in most cases, is the reflex of proto-Semitic **šamāy-* 'sky', preserving the original meaning throughout Ethio-Semitic (for the cognates elsewhere in Semitic cf. Kogan, 2011, p. 192):

Gez. *samāy* (Leslau, 1987, p. 504; Dillmann, 1865, p. 341);
Tna. *sämay* (Kane, 2000, p. 663);
Amh. *sämay* (Kane, 1990, p. 474);
Arg. of Aliyu Amba *səmay* (Leslau, 1997, p. 220), *sumay*, *säma?* (Girma Demeke, 2013, p. 136), Arg. of Ṭollaha *sämay*, *sumaye* (Girma Demeke, 2013, p. 136);
Har. *sämi* (Leslau, 1963, p. 140);
Zay Wol. *sämay* (Leslau, 1979, vol. III, p. 549);
Gaf. *sämay* (Leslau, 1956, p. 231);
End. Sod. *sämay*, Gyt. *säm̰ay*, Enm. *sām̰āy*, Čah. Eža Muh. Msq. Gog. Sod. *säme* (Leslau, 1979, vol. III, p. 549).

The Ethio-Semitic color terms for 'blue' derived from this noun are as follows (cf. also Zelealem Leyew, 2016, p. 67): Tna. *sämayawi* (Kane, 2000, p. 664), Amh. *sämayawi*, *sämawi*, *sämayay* (Kane, 1990, p. 475), Arg. of Ṭollaha *sämayawi* (Girma Demeke, 2013, p. 136; Wetter, 2010, p. 265), Arg. of Aliyu Amba *säma?ido* (Girma Demeke, 2013, p. 136), Har. *sämāwi* (Leslau, 1963, p. 140; Beniam Mitiku Cherinet, 2013, p. 288), Zay *sämāyāwi* (Meyer, 2005, p. 222), Wol. *sämay dänä* (Meyer, 2006, p. 138). The Amharic term is likely to be treated as a BCT since it is among the words most frequently named by the Amharic speakers when asked to list color terms of their language (cf. Vassilevich, 1988, p. 144). At the same time, its semantic predictability, morphological complexity, polysemy, and the status of borrowing from Geez suggest that it is less basic than terms for 'red', 'white', 'black', 'yellow', and probably 'green'.

The terms derived by means of the suffix **-āwi* in Tigrinya, Argobba of Ṭollaha, Harari, and Zay are likely formed under the influence of Amharic. In speech of my Tigrinya informants, the respective Tigrinya term is apparently used as a BCT since it appears among the top color terms in the elicited lists (although, not unlike Amharic, quite a few other criteria suggest lesser basicness than in terms for 'red', 'white', 'black').

The employment of the Harari term requires additional investigation. In my informant Baḥar Ali's speech (interviewed in Harar, November 2012), the term was used to refer to both blue and green parts of the spectrum: *gēy alämbāḥ näčiḥ*

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

wā ḳēḥ, samāwi wā näčiḥ intā 'The Harari flag is white and red, green (sic!) and white'. This extension of the original semantics of 'blue' may result from the influence of Oromo, a language with which Harari is in intensive contact (cf. Wagner, 1997, pp. 486–487). For the existence of "green-with-blue" terms *bulee* and *magarisa* in Oromo, see Zelealem Leyew's investigation (2016, p. 67). Note that a separate term for 'green', *wäriḳ*, also exists in Harari: *ḳuṭiyač wäriḳ intäyu* 'The leaves are green' (cf. also the example provided in Section 5.1.2).

The status of Argobba, Zay, and Wolane terms is uncertain (the Wolane term is a compound noun, lit. 'color of sky'). In Geez, we find a comparable adjective, derived from the same noun – Gez. *samāyāwi* 'heavenly, divine, celestial being' (Leslau, 1987, p. 504; Dillmann, 1865, p. 341) – but it was hardly used as a color term (the meaning 'blue' [Lat. caeruleus] is adduced in Dillmann [1865, p. 341], but it is not supported by textual evidence). The same applies to Sod. *sämayawi* 'heavenly' (Leslau, 1979, vol. III, p. 549), likely a recent Amharism. The meaning 'heavenly, related to the sky' was obviously the original meaning of the color terms in Tigrinya and Amharic (which probably represent early Geez borrowings) and is still in use, as the dictionaries of these languages attest: Tna. *sämayawi* 'celestial, heavenly; inhabitant of heaven; deep blue' (Kane, 2000, p. 664), Amh. *sämayawi*, *sämawi*, *sämayay* 'heavenly, celestial, blue' (Kane, 1990, p. 475). For a similar instance of two co-existing meanings ('related to X' and 'X-colored'), within one adjective in Russian at a certain stage of development of a CT, see Rakhilina & Paramei (2011, p. 126).

In Soqotri, according to the author's own data (work with Ahmed ʕIsa al-Daʃrhi, ʕIsa Gumʃan al-Daʃrhi, Soqotra, October 2017), the modern equivalent of 'blue' is *samáwi* (f. *samawíya*, pl. *samawiyín*): *siyyára samawíya* 'a blue car' (the variant *sawámi* was employed by the female informant, Maysun Muḥammad). This term, undoubtedly a recent innovation, is in all probability non-basic. It is preferably used to describe artifacts. It is absent from Leslau's dictionary (and thus, from the corpus of Soqotri texts collected by the Vienna expedition at the beginning of the twentieth century), in which the term *šéḍar* (*šíḍehor*) is given the meaning 'green, blue' (Leslau, 1938, p. 420).[8] The term *samáwi* does not occur in the first two volumes of Corpus of Soqotri Oral Literature (Naumkin & al., 2014; Naumkin & al., 2019). The morphology of *samáwi* points to a recent Arabic borrowing, although

8. The meaning 'blue' is once attested for *šéḍar* in the corpus of Soqotri texts collected by the Vienna expedition (Müller, 1905, p. 288, line 8, where the color term describes a man's blue eyes), in the rest of the contexts it refers to green color of vegetation or is used in the transferred meaning 'to be beautiful' (Müller, 1905, p. 266, line 23). In speech of our informants, the respective BCT *šéẓhar* was usually used with the meaning 'green' (Naumkin, et al. 2014, p. 695), although its employment was admitted when referring to the blue color of sea or sky (for a discussion of the various employments of this BCT in the speech of highlanders and coastal fishermen of Soqotra cf. further Naumkin – Porkhomovskij, 1981, pp. 28–29).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

pthe exact source of this borrowing remains unclear. Comparable terms exist in various Arabic vernaculars (cf. Borg, 1999, p. 133, for *samāwi* 'sky-blue' in speech of younger speakers of Negev Bedouin Arabic; Al-Rasheed & al., 2011, pp. 54, 57, on a comparable non-basic term used by Saudi Arabic speakers). Still, the ultimate derivation from the term for 'sky', going back to the same proto-Semitic term **šamāy-*, is not to be doubted.

In Semitic, I have detected only one color term for 'blue' derived from a different term for 'sky': Tgr. *ʕastarāy* 'heavenly, celestial; blue' (Littmann & Höfner, 1962, p. 465; Musā ʔĀron, 2005, p. 217). The term occurs in the Publications of Princeton Expeditions only once, on the list of colors of cows (Littmann, 1910–1913, vol. I, p. 181, No. 40), and its exact semantics, compatibility, and usage restrictions remain unclear (moreover, in Littmann (1910–1913, vol. II, p. 214), this term is said not to be used for cattle). There are no reasons to include it into the BCT system of Tigre.

5. Derivation of color terms from the names of dyes

5.1 Names of dyes as sources for 'yellow'

Among the color terms considered in the present investigation, one finds quite a number of color terms for 'yellow' derived from names of the plants used to produce yellow dyes, such as saffron (*Crocus sativus*), turmeric (*Curcuma longa*), safflower (*Carthamus tinctorius*) (Löw, 1922, pp. 103–114; Wood & Roberts, 2005, pp. 308–311). It is not always possible to say the name of which plant was the source for a given color term since the same nouns are often applied to different plants in different languages or even in one and the same language. Thus, Amharic *hərd/ərd* can denote both turmeric (*Curcuma longa* or *Curcuma domestica*) and saffron (*Crocus sativus*) (Kane, 1990, pp. 10, 1157). Syr. *kurkəmā* was apparently used to refer both to saffron (as in Merx, 1885, p. 278, line 17, explaining the *qrwqws* 'crocus') and turmeric (*ʕiṣr d-krkma* '*Curcuma longa*' as recorded in Löw, 1881, p. 219). The Arabic term for saffron, *zaʕfarān-* (Lane 1863–93, p. 1231; Biberstein Kazimirski, 1860, vol. I, p. 990; Wehr, 1979, p. 438), is also applied to safflower (cf. Musselman, 2007, pp. 248–250).

Below, several designations of the above-mentioned plants and of the dyes produced from them are discussed, which have served as referent objects for yellow in some Semitic languages.

5.1.1 *CTs going back to Sanskrit* kuṅkuma

A number of terms ultimately going back to Sanskrit *kuṅkuma* 'saffron' (Monier-Williams, 1899, p. 287), via Middle Persian *kurkum* (cf. Ciancaglini, 2008, p. 194; Löw, 1881, p. 215), are applied to saffron (*Crocus sativus*) and turmeric (*Curcuma longa*):

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hbr. *karkōm* 'saffron, *Curcuma longa*, *Crocus sativus*' (Koehler & Baumgartner, 2001, p. 498; Brown & al., 1951, p. 501; cf. further Zohary, 1982, pp. 206–207, where also the identification with safflower [*Carthamus tinctorius*] is considered);
JBA *kurkəmā* 'saffron' (Sokoloff, 2002, p. 566; Jastrow, 1926, p. 625);
Syr. *kurkəmā* '*Crocus sativus*' (Brockelmann, 1928, p. 346; Sokoloff, 2009, p. 613), *ʕqr d-kwrkm?* '*Curcuma longa*' (Löw, 1881, p. 219);
Arb. *kurkum-* 'safran indien, curcuma; henna' (Biberstein Kazimirski, 1860, vol. II, p. 888)[9];
Mhr. *kərkēm* 'orange-yellow dye' (Johnstone, 1987, p. 214);
Hrs. *kerkēm* / *kerōkem* 'yellow dye' (Johnstone, 1977, p. 69);
Jib. *kɛrkúm* 'yellow dye' (Johnstone, 1981, p. 134; 'turmeric' according to Miller & Morris, 1988, p. 182), *ekɛrkím* 'to dye yellow', *ənkɛrím* 'to be dyed, become yellow' (Johnstone, 1981, p. 134);
Soq. *kérkham* 'safran' (Leslau, 1938, p. 225; it is difficult to say whether the term, used in Müller, 1905, p. 187, lines 12, 13, denotes saffron or turmeric – the latter is suggested by the employment of the corresponding terms in Jibbali and Arabic, cf. above).

Predictably, in some Semitic languages, color terms are derived from the above-mentioned designations of yellow dyes: Syr. *kurkəmūtā* 'color crocinus', *məkarkəmat gawnā* 'crocicolor', *kurkəmānā* 'crocicolor; pl. vestes crocicolores' (Brockelmann, 1928, p. 346; Sokoloff, 2009, pp. 613–614, 655), JNA *kurkmāna* 'yellow, pale' (Sabar, 2002, p. 160), Beni Amer dialect of Tigre *kurkum* 'yellow' (Nakano, 1982, p. 122), Mhr. *kerkmín* 'yellow' (Nakano, 1986, p. 125) or *karkamī* (Watson, 2012, p. 105), Hrs. *mekárkem* 'dyed yellow, yellow' (Johnstone, 1977, p. 69), Hob. *karkmí* 'yellow' (Nakano, 2013, p. 238), Soq. *kérkɛm* 'yellow' (author's own field data; cf. *kírkam* 'jaune', Leslau, 1938, p. 225); *kárkam* 'yellow', Nakano, 1986, p. 125). In Jibbali, Johnstone's dictionary contains only a verb with the color meaning: Jib. *ekɛrkím* 'to dye yellow', *ənkɛrkím* 'to be dyed, become yellow' (but notably, the verb is not restricted to artificially dyed objects: *ənkɛrkím ēgəhš* 'his face turned yellow', Johnstone 1981, p. 134).

In the absence of special investigations, it is difficult to evaluate the functions of the above-mentioned color terms. Certainly not all of them are to be treated as basic. According to Rubin, 2010, p. 80, the BCC system of Mehri of Oman does not have a separate yellow category, employing rather a composite category 'yellow-with-green', denoted by *həźawr*. As for Yemeni Mehri, *karkamī* is not the basic term for yellow at least in the dialect of Rēhän, where the Arabic loanword

9. Cf. also Arb. Daṭ. *kurkum* 'safran des Indes' (Landberg, 1942, p. 2569); Arb. Yem. *kurkum* 'Curcuma' (Behnstedt, 1992–2006, p. 1065), Arb. Ḍfr. *kerkem* 'turmeric (*Curcuma longa*)' (Miller & Morris, 1988, xxiii, p. 182).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

aṣfar is consistently used to denote yellow color in one of the available texts (cf. Sima, 2009, p. 456, line 4, p. 458, line 9, etc.).

In case of Soqotri, the adjective *kérkɛm* is already present in the Vienna corpus as an epithet of human skin (Müller, 1905, p. 168, line 2, p. 167, line 17) and dried lumps of frankincense resin (ibid. p. 44, line 4). The Vienna corpus also contains its derivatives *kérkam* 'jaune d'oeuf' and *inkórkim* 'être jaune' (Leslau, 1938, p. 225). Although it does not occur in the first two volumes of the Corpus of Soqotri Oral Literature (Naumkin & al., 2014; Naumkin & al., 2019), it is well known to our informants, who use it mostly to describe artefacts (*sirwál kérkɛm* 'yellow trousers', *ḳamís kérkim* 'yellow shirt') and never apply it to human hair or cattle colors. All in all, the evidence from the Vienna corpus suggests the basic status for this CT.

5.1.2 *CTs going back to Arabic* hurd-

In Arabic, the term *hurd-* 'certain roots with which one dyes, *kurkum-*' (Lane, 1863–93, p. 2891; Biberstein Kazimirski, 1860, vol. II, p. 1411) is typically used as a designation of turmeric (*Curcuma longa*).[10] The Arabic *hurd-* is borrowed from Skr. *hāridrá* 'colored with turmeric, yellow' (Monier-Williams, 1899, p. 1292; cf. Leslau, 1990, p. 14).

Both the Arabic term for turmeric and the Arabic adjective *hurdiyy-* 'dyed yellow with *hurd-*' (Lane, 1863–93, p. 2891; Biberstein Kazimirski, 1860, vol. II, 1411), derived from the same term, are borrowed into a number of Ethio-Semitic languages (Leslau, 1990, pp. 136, 276, 340; cf. also the Arabic borrowing in Jibbali, *herd* 'turmeric', Miller & Morris, 1988, p. 152).

Borrowings from Arb. *hurd* 'turmeric':

Amh. *hərd*, *ərd* 'saffron (*Crocus sativus*); turmeric (*Curcuma longa*)' (Kane, 1990, pp. 10, 1157);

Tna. *hərud* 'turmeric (*Curcuma longa*, *Curcuma domestica*)' (Kane, 2000, p. 17).

Borrowings from Arb. *hurdiyy-* 'dyed with turmeric' (all with the meaning 'yellow'):

Tgr. of Samhar (Zəla) *hərdi* (Saleh Mahmud Idris, 2015, p. 259);

Southern Arg. *hurdi* (Leslau, 1997, p. 206), Arg. of Aliyu Amba *hurdi* (Girma Demeke, 2013, p. 215);

Har. *ḥurdi* (Leslau, 1963, p. 86; Beniam Mitiku Cherinet, 2013, p. 288; Zelealem Leyew, 2016, p. 65).

10. Cf. Arb. Ḍfr. *herd* '*Curcuma longa*' (Miller & Morris, 1988, pp. 174, 185).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The color terms in Tigre and Argobba are apparently non-basic. In each of the concerned languages, alternative terms for 'yellow' exist. In Tigre, different dialects employ different terms for yellow, such as *šagrāy* (borrowed from Arabic),[11] *ʔasfar* (from Arabic *ʔaṣfar-* 'yellow'), *bəčˌa* and its variants (likely from Tigrinya *bəčˌa* 'yellow'), cf. Saleh Mahmud Idris, 2015, pp. 241, 259). In Argobba of Aliyu Amba, the term *hurdi* is used alongside *bəčˌa* (likely from Amharic) and *ʔasfär* (from Arb. *ʔaṣfar-*), whereas in Argobba of Ṭollaha, only *bəčˌa* is employed (Girma Demeke, 2013, p. 215; cf. Wetter, 2010, p. 265).

In Harari, however, there are good reasons to treat the term *ḥurdi* as a BCT. It was adduced by my informant Baḥar Ali (Harar, November 2012) as the main term for 'yellow', applicable to both artificial and natural objects: *ityoṗyā ʔalämbāḥ ḳēḥ ḥurdi wa wäriḳ inta* 'The Ethiopian flag is red, yellow, and green',[12] *aḳoḥāč lāyle näčˌiḥ intäyu usṭube ḥurdi intäyu* 'The eggs are white from the outside and yellow inside'.

5.2 Names of dyes and dyed artifacts as sources for 'blue'

5.2.1 *'Kohl' > 'blue'*

The mineral dye, known as kohl and primarily used as an eye paint, was produced from lead (cf. Stol, 1989, pp. 165–166) and, more recently, from antimony, or stibium (found in a mineral called stibnite, or antimonite). The term "kohl" (or "kuhl") in European languages is borrowed from Arb. *kuḥl-*. Comparable terms are present in most West Semitic languages as a result of a chain of borrowings (cf. Nöldeke, 1910, p. 40; see also Leslau, 1990, pp. 17, 74, 151, 169, 203). The relevant terms for the cosmetic dye and the process of dying with it are as follows:

Hbr. *kḥl* (qal) 'to paint (the eyes)' (Koehler & Baumgartner, 2001, p. 469), pB. *kāḥal* 'to paint the eyelids' (Jastrow, 1926, p. 629), *kōḥāl* '*koḥl*, a powder used for painting the eyelids, stibium' (Jastrow, 1926, p. 618);

JBA *kuḥlā* 'stibnite, kuhl' (Sokoloff, 2002, p. 558), *kḥl* 'to paint the eyes' (Sokoloff, 2002, p. 568);

JPA *kḥl* '*kohl* (eye paint)', *kḥl* 'to paint the eyes' (Sokoloff, 1992, p. 255; Jastrow, 1926, p. 629);

---

11. Leslau, 1990, p. 178. In the corpus of Publications of Princeton Expedition *šagrāy* and *ʔašgar* occur only as colors of cattle (cf. Littmann & Höfner, 1962, p. 229); cf. Arb. Sud. *ašgar* 'yellowish brown, golden brown', used only as a color of sheep (Reichmuth, 1981, p. 61). The term ultimately goes back to Arb. *ʔašqar-* (fem. *šaqrāʔ-*) 'applied to a man of a ruddy complexion; applied to a horse: sorrel' (Lane, 1863–93, p. 1581).

12. The adjective *burtukān* (borrowing from Amh. *bərtukan* 'orange (fruit, color)', cf. Zelealem Leyew, 2016, p. 72) was mentioned as a possible substitute for *ḥurdi* in this context.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Syr. *kāḥlā* 'nigris oculis praeditus', *kuḥlā* 'stibium' (Brockelmann, 1928, p. 324; Sokoloff, 2009, p. 604);
Mnd. *kahla* 'kohol, antimony, collyrium' (Drower & Macuch, 1963, p. 195);
JNA *kixla* 'antimony, kohl, mineral eye-paint' (Sabar, 2002, p. 186);
Arb. *kuḥl-* 'collyre; tout cosmétique ou remède employé pour les yeux' (Biberstein Kazimirski, 1860, vol. II, p. 871; Ibn Manḍūr, 2003, vol. 11, p. 695; az-Zabīdī, 1986–1998, vol. 30, p. 317; cf. also Lane, 1863–93, pp. 2998–2999);
Gez. *kʷəḥl* 'antimony, dye for the eyelid' (Leslau, 1987, p. 279; Dillmann, 1865, pp. 823–824)[13];
Tgr. *kəḥl* 'antimony', *kaḥala* 'to apply antimony (to the eyelids)' (Littmann & Höfner, 1962, p. 393);
Tna. *kʷəḥli* 'antimony, kohl', *kʷäḥalä* 'to daub the eyelashes, eyebrows with kohl' (Kane, 2000, p. 1572);
Amh. *kul*, *kʷəl* 'antimony dust, kohl', *kʷalä* 'to put kohl on the eyelids; to plaster a house; to write on parchment; to paint a picture, daub (with color)' (Kane, 1990, p. 1365);
Har. *kuḥul* 'antimony' (Leslau, 1963, p. 91);
Wol. *kul* 'antimony' (Leslau, 1979, vol. III, p. 341);
Mhr. *kēḥəl* 'kohl' (Johnstone, 1987, p. 206);
Hrs. *kéḥel* 'kohl, antimony' (Johnstone, 1977, p. 67);
Jib. *kúḥəl* 'kohl', *kḥal* 'to apply kohl (to one's eyes)' (Johnstone, 1981, p. 129);
Soq. *kéḥel* 'collyre', *kḥl* 'enduire de collyre' (Leslau, 1938, p. 216).

Akk. *guḫlu* is often thought to be a name for kohl and a source of borrowing for the West Semitic terms (cf. Gelb & al., 1956, p. 125; 'antimony (used as eye paint)'; cf. also von Soden, 1965–1981, p. 296; Koehler & Baumgartner, 2001, p. 469). However, an alternative interpretation exists (cf. Potts & al., 1996), according to which *guḫlu* denoted bdellium and is totally unrelated to the West Semitic *kḥl*.

The name for kohl is used in a few Semitic languages to derive (mostly non-basic) color terms with the semantics of 'black-with-blue' (sometimes extended into 'dark-green').

Evidence for a color term with the meaning 'black, dark' derived from the name for 'antimony' may be present already in Hebrew epigraphy: *kḥl* 'dark coloured' (in a collocation *yyn kḥl* 'wine of K.' (?) = 'dark wine') (Hoftijzer & Jongeling, 1995, p. 456; for the interpretation of *kḥl* as 'black, dark', see Demsky, 1972).

---

13. In the dictionaries by Leslau (1987, p. 279) and Kidanä Wäld Kəfle (1955–1956, p. 528), a Geez color term derived from this term is adduced, namely, *kʷəḥli* 'blue'; it is absent, however, from Dillmann's lexicon.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

In Classical Arabic, a non-basic color term derived from the same root refers to black and dark blue, but can also be applied to fresh vegetation (thus it can be defined as a non-basic macro-black): *ʔakḥal-* 'qui a des yeux d'un noir foncé; qui a les paupières enduites de collyre; qui a les paupières naturellement d'une nuance brune comme si elles étaient enduites de collyre; noir; qui a le corps blanc et les yeux noirs (mouton)' (Biberstein Kazimirski, 1860, vol. II, p. 871; Ibn Manḏūr, 2003, vol. 11, p. 696; az-Zabīdī, 1986–1998, vol. 30, p. 317), *kaḥl-* 'ciel, voûte azurée du ciel' (Biberstein Kazimirski, 1860, vol. II, pp. 870–71; Ibn Manḏūr, 2003, vol. 11, p. 697; az-Zabīdī, 1986–1998, vol. 30, p. 317), *ʔikḥālla* 'commencer à se couvrir de verdure (se dit du sol)' (Biberstein Kazimirski, 1860, vol. II, p. 870; Ibn Manḏūr, 2003, vol. 11, p. 697; az-Zabīdī, 1986–1998, vol. 30, p. 319; cf. also *kuḥl-u l-ʕušb-i* in Ibn Manḏūr, 2003, vol. 11, p. 697; az-Zabīdī, 1986–1998, vol. 30, p. 363; cf. also Fischer, 1965, pp. 60, 284). The inclusion of (dark-)green hue into the reference range of this root is preserved in Maltese Arabic (Borg, 2011a, p. 83).

In some modern Arabic vernaculars one finds new derivatives (with the adjectival suffix **-iyy*) from the same term, designating dark-blue (cf. Borg, 2011b, pp. 11, 22; Al-Rasheed & al., 2011, p. 55; Reichmuth, 1981, p. 57), all of them non-basic. (Cf. also JNA *kiḥli* 'azure-colored', borrowed from Arb. *kuḥliyy-*; Sabar, 2002, p. 184).

The only Semitic language where a BCT is derived from the name of this dye is Modern Hebrew (e.g., Brenner, 1982, p. 211): the adjective *kāḥōl* came to be used as the main term for 'blue', probably under the influence of Arabic (cf. Borg, 2011b, p. 22).

5.2.2 *'Ink' > 'black', 'blue', 'blue-with-green'*

In Jewish Neo-Aramaic (dialects of Amidya, Dihok, Nerwa, and Zakho, Northwestern Iraq), the only color term for 'blue' (but also for 'green') is *mīlāna* 'blue (Zakho JNA), green (Nerwa Texts)' (Sabar, 2002, p. 216).

Comparable terms, notably also with the semantics of 'blue' and 'green', are reported for Neo-Assyrian Aramaic dialect (northern Iran, near Lake Urmia): *mīlāna* 'green' (Kuipers, 1983, p. 119) and *millāna* 'blue' (Kuipers, 1983, p. 59), which are probably two variant pronunciations of the same color term, 'green-with-blue'.

An undoubtedly related term in Jewish Palestinian Aramaic is a non-basic CT 'black': JPA *mylny* 'black' (Sokoloff, 1992, p. 305; Jastrow, 1926, p. 775). Note the usage *mylny ʔymyrʔ* 'black (luckless) day' in Avoda Zara (see Sokoloff, 1992, p. 305; Jastrow, 1926, p. 775), obviously a transcription of Gr. *mélaina ^{h}ēméra*. Cf. also ChrPA *mylʔnyn* 'black color' (Sokoloff, 2014, p. 220). For other comparable terms in modern Aramaic vernaculars see Borg (2014, p. 46).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

All these terms are clearly related to Gr. *mélas*, gen. *melanós* 'black', *mélan* 'black dye, ink'.[14]

Since the semantics of the Greek and Aramaic adjectives are rather different, it is unlikely that the Aramaic color terms are direct borrowings from Greek. Rather, the Aramaic terms for 'black' and 'blue' (with a further extension into 'green') are derived from the nouns with the meaning 'ink, blue dye' and borrowed from Greek *mélan* or *melánion* (where these substantives are derived from BCT 'black', cf. Liddell & Scott, 1996, p. 1095). While I failed to find borrowings from these Greek nouns in Jewish Palestinian Aramaic and Neo-Assyrian Aramaic, such loanwords are present elsewhere in Aramaic, namely, ChrPA *myl'n, mylwn* 'ink' (Sokoloff, 2014, p. 220), Syr. *mlnyn* 'atramentum' (Brockelmann, 1928, p. 392; Sokoloff, 2009, p. 774), and JNA *mīla* 'liquid blue dye' (Sabar, 2002, p. 216).

5.2.3 *'Indigo' > 'blue', 'indigo-dyed cloth' > 'blue'*

Indigo (*Indigofera tinctoria*) was well-known and widely used as a dyestuff in the Middle East since ancient times (e.g., Löw, 1922, pp. 122–135; Balfour-Paul, 1997, pp. 3–29; Borg, 2007, p. 276). Some non-basic terms for 'blue' in modern Aramaic dialects, such as Tlk. *nīli*, Tur. *nīlo* 'blue (navy)' (Kuipers, 1983, p. 59), are obvious loans from Arb. *nīl-* 'plante dont ón tire l'indigo (indigifera tinctoria); isatis, plante' (Biberstein Kazimirski, 1860, vol. II, p. 1376), in turn borrowed from Sanskrit (Skr. *nī́la* 'dyed with indigo', Monier-Williams, 1899, p. 566). The Arabic term is also the source of designations of indigo in South Arabia and in Ethiopia (note that the Tigrinya and Amharic loans are used as the designations of a specific hue rather than dye):

Gez. *nil* 'indigo' (Leslau, 1987, p. 398);
Tgr. *nil* id. (Littmann & Höfner, 1962, p. 323);
Tna. *nil* 'blue (color)' (Kane, 2000, p. 1303);
Amh. *nil* 'blue, blue vitriol', *näyyälä* 'to dye blue' (Kane, 1990, pp. 1018, 1054);
Mhr. *nayl* 'indigo' (Johnstone, 1987, p. 309);
Hrs. *nīl* id. (Johnstone, 1977, p. 100);
Jib. [*nyl*] *nuź* id., *enbél* 'to put indigo on one's face (to cool it); to dye with indigo' (Johnstone, 1981, p. 200).

For a detailed etymology see Powels (1999, p. 157 f.; cf. also Löw, 1922, pp. 133–134).

14. Contrary to Borg (2014, pp. 45–46), where a phonologically unconvincing comparison with Arb. *nīl-*, *nīlat-* 'indigo' is proposed.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

In his vocabulary of Hobyot, an endangered Modern South Arabian language, A. Nakano (2013, p. 238) lists a term for 'blue', *ṭowḳí*, *ṭowḳiyóot*, pl. *ṭowḳó*. The basic status of this term is far from certain, and more fieldwork is needed to establish whether we deal with a non-basic term or a nascent BCT. The referent object is, in this case, clearly an artifact (cloth, a piece of textile) dyed with indigo. Cf. the comparable terms in the South Arabian and Ethiopian area:

Arb. *ṭāḳ-* 'a certain sort of garment, having sleeves; a garment such as *ʔaḥḍar* (properly meaning green, but when applied to a garment commonly meaning 'of a dark, or an ashy, dust-color'); a woman's muffler, or head-covering' (Lane, 1863–93, p. 1895; Ibn Manḍūr, 2003, vol. 10, p. 270; az-Zabīdī, 1986–1998, vol. 26, pp. 107–108)[15];
Gez. *ṭāḳ*, *ṭāḳā* 'bolt of cloth' (Leslau, 1987, p. 595, not in Dillmann, 1865);
Tgr. *ṭāḳat* 'cloth, stuff (for clothes)' (Littmann & Höfner, 1962, p. 614);
Tna. *ṭaḳ̌a*, *ṭaḳ̌at* 'haze; roll, bolt (of cloth), piece goods' (Kane, 2000, p. 2441);
Amh. *ṭaḳa* 'bolt of cloth (ca. 27 m.)' (Kane, 1990, p. 2129);
Har. *ṭāḳa* 'roll (of cloth)' (Leslau, 1963, p. 155);
Sel. *ḳāḳa* 'woven cloth before being made into a garment' (Leslau, 1979, vol. III, p. 492);
Mhr. *ṭōḳat* 'ingigogefärbter Tuch' (Sima, 2009, pp. 334–335, line 31);
Jib. *ṭíḳət*/*ṭéḳ* 'indigo cloth for a woman's dress; indigo-dyed woman's dress' (Johnstone, 1981, p. 282);
Soq. *ṭáḳeḥ* 'drap noir' (Leslau, 1938, p. 201).

Both MSA and Ethio-Semitic terms are borrowed from Arabic (cf. Leslau, 1990, p. 369).

Thus, in Hobyot we clearly deal with a color term derived from a name of an artifact typically dyed with a certain dye. A similar strategy (a term for a porcelain cup, typically dyed blue, as a source of derivation for *ṣīni* 'blue'/'black'/'green') in several Arabic vernaculars has been extensively discussed by A. Borg (2007, pp. 277–278; cf. also Borg, 1999, p. 133; 2014, p. 46).

15. Cf. also Yem. Arb. *ṭawḳ* 'breadth of cloth', *ṭāḳah* 'measure of cloth; woman's undershirt' (Piamenta, 1990, p. 310).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

6. Conclusions

Employment of names of referent objects to derive non-basic color terms is widespread in the languages of the world. Eventually, some of these terms may enter the BCT systems. In the present contribution, examples of names of new BCCs and/or of non-basic terms which can eventually become BCTs for new BCCs have been discussed.

As far as the terms for 'yellow' are concerned, their etymological sources of derivation include wax (see Section 4.1) and names of yellow dyes: saffron (see Section 5.1.1) and turmeric (see Sections 5.1.1 and 5.1.2).

The diachronic referent objects for 'green' considered here are various types of vegetation: plants (possibly cabbage), grass, leaf (see Section 4.2).

The terms for 'blue' in Semitic can be derived from designations of sky (see Section 4.3), kohl (a mineral-based cosmetic dye, see Section 5.2.1), indigo and indigo-colored objects (see Section 5.2.3). Derivatives from the designation of 'ink' (see Section 5.2.2) are used to form color terms for 'black' and 'green-with-blue' (or 'violet', cf. Borg, 2014, p. 46).

A few borrowed terms are present among the considered CTs, such as Har. *ḥurdi* 'yellow' (Section 5.1.2), Tna. *ḳäṭälya* 'green' (4.2.3), Soq. *samáwi* 'blue' (Section 4.3). At least in the case of Har. *ḥurdi*, there is good reason to suppose its basic status (although its source word is a non-basic color terms in the donor language).

Acknowledgment

Russian State University for the Humanities, Miusskaya pl. 6, 125993, Moscow, Russia – National Research University Higher School of Economics, Myasnitskaya ul. 20, 101000, Moscow, Russia. My work on this article was funded by RFBR/РФФИ (grant #17-06-00391).

Abbreviations

Arb.	Arabic	End.	Ǝndägañ
Arb. Daṭ.	Arabic dialect of Daṭīna	Enm.	Ǝnnämor
Arb. Ḍfr.	Arabic dialect of Ḍofar	Gez.	Geez
Arb. Sud.	Arabic dialect of Sudan	Gog.	Gogot
Arb. Yem.	Arabic dialect of Yemen	Gr.	Greek
Arg.	Argobba	Gyt.	Gyeto
Čah.	Čaha	Har.	Harari
ChrPA	Christian Palestinian Aramaic	Hbr.	Biblical Hebrew

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hob.	Hobyot	Sel.	Səlṭi
Hrs.	Harsusi	Skr.	Sanskrit
JBA	Jewish Babylonian Aramaic	Sod.	Soddo
Jib.	Jibbali	Soq.	Soqotri
JNA	Jewish Neo-Aramaic	Syr.	Syriac
JPA	Jewish Palestinian Aramaic	Tgr.	Tigre
Mhr.	Mehri	Tlk.	Neo-Aramaic of Telkappe
Msq.	Mäsqan	Tna.	Tigrinya
Muh.	Muhər	Tur.	Neo-Aramaic of Turoyo
pB.	post-Biblical Hebrew	Wol.	Wolane

References

Al-Rasheed, A. S., Al-Sharif, H. H., Thabit, M. J., Al-Mohimeed, N. S., & Davies, I. R. L. (2011). Basic colour terms of Arabic. In C. B. Biggam, C. A. Hough, C. J. Kay, & D. R. Simmons (Eds.), *New Directions in Colour Studies* (pp. 53–58). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.167.10alr

Balfour-Paul, J. (1997). *Indigo in the Arab World*. London/New York: Routledge.

Bassano, F. da. (1918). *Vocabolario tigray-italiano e repertorio italiano-tigray*. Roma: Casa Editrice Italiana di C. de Luigi, 1918.

Behnstedt, P. (1992–2006). *Die nordjemenitischen Dialekte*. II. *Glossar*. Wiesbaden: Reichert Verlag.

Beniam Mitiku Cherinet. (2013). *Harari – A Descriptive Grammar*. Ph.D. thesis, Addis Ababa University.

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

Biberstein Kazimirski, A. de. (1860). *Dictionnaire arabe-français*. Vols. *I-II*. Paris: Maisonneuve.

Biggam, C. (2012). *The Semantics of Colour: A Historical Approach*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139051491

Biggam, C. (2014). Prehistoric colour semantics: A contradiction in terms. In W. Anderson, C. P. Biggam, C. Hough, & C. Kay (Eds.), *Colour Studies: A broad spectrum* (pp. 2-28). Amsterdam/Philadelphia: John Benjamins.

Borg, A. (1999). The color categories of the Negev Bedouin. In A. Borg (Ed.), *The Language of Color in the Mediterranean* (pp. 121–147). Stockholm: Almkvist/Wiksell.

Borg, A. (2004). *A Comparative Glossary of Cypriot Maronite Arabic*. Leiden: Brill.

Borg, A. (2007). Towards a History and Typology of Color Categorization in Colloquial Arabic. In R. E. MacLaury, G. V. Paramei, D. Dedrick (Eds.). *Anthropology of Color. Interdisciplinary multilevel modeling* (pp. 263–293). Amsterdam/Philadelphia: John Benjamins.

Borg, A. (2011a). Towards a diachrony of Maltese basic colour terms. In C. P. Biggam, C. A. Hough, C. J. Kay, & D. R. Simmons (Eds.), *New Directions in Colour Studies* (pp. 73–90). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.167.12bor

Borg, A. (2011b). The Basic Colour Terms in Maltese. *Zeitschrift für Arabische Linguistik*, 53, 5–35.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Borg, A. (2014). Towards a historical and cultural atlas of colour terms in the Near East: Remarks on GREEN and BLUE in some Arabic and Aramaic vernaculars. In W. Anderson, C. P. Biggam, C. Hough, C. Kay (Eds.), *Colour Studies: A Broad Spectrum* (pp. 31–52), Amsterdam/Philadelphia: John Benjamins.

Brenner, A. (1982). *Colour Terms in the Old Testament*. Sheffield: JSOT Press.

Brockelmann, C. (1928). *Lexicon Syriacum*. Hallis Saxonum: Niemeyer.

Brown, F., Driver, S. R., & Briggs, Ch. A. (1951). *A Hebrew and English Lexicon of the Old Testament*. Oxford: Clarendon Press (repr. with correction of the original 1907 edition).

Bulakh, M. (2003). Etymological Notes on the Akkadian Colour Terms. In: L. Kogan (Ed.), *Studia Semitica* (Orientalia: Papers of the Oriental Institute 3) (pp. 3–17). Moscow: Russian State University for the Humanities.

Bulakh, M. (2004). Color Terms of Modern South Arabian Languages: a Diachronic Approach. *Babel und Bibel*, 1, 269–282.

Bulakh, M. (2005). *Cvetooboznačenie semitskih jazykov v etimologičeskom aspekte* [Color Naming of Semitic Languages in Diachronic Aspect]. Ph.D. thesis, Russian State University for the Humanities, Moscow.

Bulakh, M. (2006a). Basic Color Terms of Biblical Hebrew in Diachronic Aspect. *Babel und Bibel*, 3, 181–216.

Bulakh, M. (2006b). Basic Colour Terms in Gəʕəz: Synchronic and Diachronic Aspects. In S. Uhlig (Ed.), *Proceedings of the XVth International Conference of Ethiopian Studies, Hamburg July 20–25, 2003* (pp. 737–745). Wiesbaden: Harrassowitz.

Bulakh, M. (2007). Basic color terms from Proto-Semitic to Old Ethiopic. In R. E. MacLaury, G. V. Paramei, D. Dedrick (Eds.). *Anthropology of Color. Interdisciplinary multilevel modeling* (pp. 247–261). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.137.16bul

Ciancaglini, C. A. (2008). *Iranian Loanwords in Syriac*. Wiesbaden: Dr. Ludwig Reichert Verlag.

Dästa Täklä Wäld. (1962 A.M. [1970 A.D.]). *Addis yamarəñña mäzgäbä ḳalat* [The New Amharic Dictionary]. Addis Abäba: Artistik mattämiya bet.

Dedrick, D. (1998). *Naming the Rainbow. Colour Language, Colour Science, and Culture*. Dordrecht/Boston/London: Kluwer Academic Publishers.

Demsky, A. (1972). 'Dark wine' from Judah. *Israel Exploration Journal*, *22*(4), 233–234.

Dillmann, A. (1865). *Lexicon linguae aethiopicae cum indice latino*. Lipsiae: T. O. Weigel.

Drower, E. S., & Macuch, R. (1963). *A Mandaic Dictionary*. Oxford: Clarendon Press.

Fischer, W. (1965). *Farb- und Formbezeichnungen in der Sprache der altarabischen Dichtung*. Wiesbaden: Harrassowitz.

Gelb, I. J., Jacobsen, T., Brinkman, J. A., Civil, M., & Biggs, R. D. (1956). *The Assyrian Dictionary of the Oriental Institute, the University of Chicago*. Vol. 5: G. Chicago: Oriental Institute.

Girma A. Demeke. (2013). *Amharic-Argobba Dictionary*. Trenton, NJ: Red Sea Press.

Gutt, E. H. M., & Hussein Mohammed. (1997). *Silt'e-Amharic-English Dictionary (with Concise Grammar by Ernst-August Gutt)*. Addis Ababa: Addis Ababa University Press.

Hoftijzer, J., & Jongeling, K. (1995). *Dictionary of the North-West Semitic Inscriptions*. (Handbuch der Orientalistik 21). Leiden/New York/Köln: Brill.

Ibn Manḍūr. (2003). *Lisān-u l-ʕarab-i*. Beyrouth/Liban: Dar Al-Kotob Al-Ilmiyah.

Jastrow, M. (1926). *A Dictionary of the Targumim, the Talmud Babli and Yerushalmi, and the Midrashic literature*. New York: Chorob.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Johnstone, T. M. (1977). *Ḥarsusi Lexicon and English-Ḥarsūsi Word-list*. London et al.: Oxford University Press.
Johnstone, T. M. (1981). *Jibbāli Lexicon*. Oxford/New York: Oxford University Press.
Johnstone, T. M. (1987). *Mehri Lexicon and English-Mehri Word-list*. London: School of Oriental and African Studies, University of London.
Kane, Th. L. (1990). *Amharic-English Dictionary*, 2 vols. Wiesbaden: Harrassowitz.
Kane, Th. L. (2000). *Tigrinya-English Dictionary*, 2 vols. Springfield: Dunwoody Press.
Kay, P., & Maffi, L. (1999). Color Appearance and the Emergence and Evolution of Basic Color Lexicons. *American Anthropologist*, *101*(4), 743–760. https://doi.org/10.1525/aa.1999.101.4.743
Kay, P., & McDaniel, Ch. K. (1978). The Linguistic Significance of the Meanings of Basic Color Terms. *Language* 54(3), 610–646. https://doi.org/10.1353/lan.1978.0035
Kerttula, S. (2002). *English Colour Terms: Etymology, Chronology, and Relative Basicness*. Helsinki: Société néophilologique.
Kerttula, S. (2007). Relative basicness of color terms. Modeling and measurement. In R. E. MacLaury, G. V. Paramei, & D. Dedrick (Eds.). *Anthropology of Color. Interdisciplinary multilevel modeling* (pp. 151–169). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.137.11ker
Kidanä Wäld Kəfle. (1948 A.M. [1955–1956 A.D.]). *Mäṣḥafä säwasəw wä-gəss wä-mäzgäbä ḳalat ḥaddis* [The Book on Grammar and Verb, and a New Dictionary]. Addis Abäba: Artistik mattämiya bet.
Koehler, L., & Baumgartner, W. (2001). *The Hebrew and Aramaic Lexicon of the Old Testament*, subsequently revised by Walter Baumgartner and Johann Jakob Stamm. Leiden/Boston/Köln: Brill.
Kogan, L. (2001). *ɣ in Akkadian. *Ugarit-Forschungen*, 33, 263–298.
Kogan, L. (2011). Proto-Semitic Lexicon. In St. Weninger in collaboration with G. Khan, M. P. Streck, J. C. E. Watson (Eds.), *The Semitic Languages: An International Handbook* (pp. 179–258). De Gruyter Mouton: Berlin – Boston. https://doi.org/10.1515/9783110251586.179
Kogan, L. (2015). *Genealogical Classification of Semitic: The Lexical Isoglosses*. Boston/Berlin: De Gruyter. https://doi.org/10.1515/9781614515494
Kuipers, F. G. (1983). *A Comparative Lexicon of Three Modern Aramaic Dialects*. Ph.D. thesis, Georgetown University.
Landberg, C. de. (1942). *Glossaire Daṯînois*. Vol. III. Leiden: Brill.
Lane, E. W. (1863–93). *An Arabic-English Lexicon*. 8 vols. London: Williams and Norgate.
Leslau, W. (1938). *Lexique Soqotri (Sudarabique moderne)*. Paris: Librairie C. Klincksieck.
Leslau, W. (1956). *Étude descriptive et comparative du Gafat (éthiopien méridional)*. Paris: Klincksieck.
Leslau, W. (1963). *Etymological Dictionary of Harari*. Berkeley – Los Angeles: University of California Press.
Leslau, W. (1979). *Etymological Dictionary of Gurage (Ethiopic). Vol I: Individual Dictionaries; Vol II: English-Gurage Index; Vol III: Etymological Section*. Wiesbaden: Harrassowitz.
Leslau, W. (1987). *Comparative Dictionary of Geʕez (Classical Ethiopic)*. Wiesbaden: Harrassowitz.
Leslau, W. (1997). *Ethiopic Documents: Argobba. Grammar and Dictionary*. (Aethiopistische Forschungen, 47). Wiesbaden: Harrassowitz.
Leslau, W. (1990). *Arabic Loanwords in Ethiopian Semitic*. Wiesbaden: Harrassowitz.
Liddell, H. G., & Scott, R. (1996). *A Greek-English Lexicon*. Oxford: Clarendon Press.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Littmann, E. (1910–1913). *Publications of the Princeton Expedition to Abyssinia.* Vol. I. *Tales, Customs, Names and Dirges of the Tigrē Tribes: Tigrē text*. Leyden: Brill, 1910; Vol. II. *Tales, Customs, Names and Dirges of the Tigrē Tribes: English translation*. Leyden: Brill, 1910; Vol. III. *Lieder der Tigrē-Stämme*. Leyden: Brill, 1913. Vol. IV. *Lieder der Tigrē-Stämme. Deutsche Übersetzung und Commentar*. Leyden: Brill, 1913.

Littmann, E., & Höfner, M. (1962). *Wörterbuch der Tigrē-Sprache. Tigrē-Deutsch-English*. Wiesbaden: Franz Steiner.

Löw, I. (1881). *Aramäische Pflanzennamen*. Leipzig: Verlag von Wilhelm Engelmann. https://doi.org/10.5962/bhl.title.30293

Löw, I. (1922). Semitische Färberpflanzen. *Zeitschrift für Semitistik und verwandte Gebiete*, *1*(2), 97–162.

MacLaury, R. E. (1991). Exotic Color Categories: Linguistic Relativity to What Extent? *Journal of Linguistic Anthropology*, *1*(1), 26–51. https://doi.org/10.1525/jlin.1991.1.1.26

MacLaury, R. (2001). Color terms. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (eds.), *Language Typology and Language Universals. Sprachtypologie und sprachliche Universalien. La typologie des langues et les universaux linguistiques. An International Handbook / Ein internationales Handbuch / Manuel international* (pp. 1227–1251). Berlin/New York: Walter de Gruyter.

Maclean, A. J. (1901). *Dictionary of the Dialects of Vernacular Syriac*. Oxford: Clarendon Press.

Merx, A. (1885). *Proben der syrischen Übersetzung von Galenus' Schrift über die einfachen Heilmittel. Zeitschrift der Deutschen Morgenländischen Gesellschaft*, 39, 237–305.

Meyer, R. (2005). *Das Zay. Deskriptive Grammatik einer Ostguragesprache (Äthiosemitisch)*. Köln: Rüdiger Köppe Verlag.

Meyer, R. (2006). *Wolane. Descriptive Grammar of an East Gurage Language (Ethiosemitic)*. Köln: Rüdiger Köppe Verlag.

Miller, A., & Morris, M. (1988). *Plants of Dhofar. The Southern Region of Oman: Traditional, Economic and Medicinal Uses*. [Muscat]: Office of the Advisor for Conservation of the Environment, Diwan of Royal Court.

Monier-Williams, M. (1899). *A Sanskṛit-English Dictionary*. Oxford: Oxford University Press.

Müller, D. (1905). *Die Mehri- und Soqoṭri-Sprache*. Bd. II: *Soqoṭri-Texte*. *Wien*: Hoelder.

Munzinger, W. (1864). *Ostafrikanische Studien*. Schaffhausen: Fra. Hurtersche Buchhandlung.

Musā ʔĀron. (2005). *Kəbət-ḳālāt həgyā təgre* [A Tigre Dictionary]. Asmara: Hdri Publishers.

Musselman, L. J. (2007). *Figs, dates, laurel, and myrrh: plants of the Bible and the Quran*. Timber Press: Portland.

Nakano, A. (1982). *A vocabulary of Beni Amer dialect of Tigre*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Nakano, A. (1986). *Comparative Vocabulary of Southern Arabic – Mahri, Gibbali and Soqotri*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Nakano, A. (2013). *Hobyōt (Oman) Vocabulary with example texts*. Edited by Robert Ratcliffe. Tokyo: Research Institute for Languages and Cultures of Asia and Africa (ILCAA).

Naumkin, V., Kogan, L., ʕIsa Gumʃan al-Daʃrhi, Ahmed ʕIsa al-Daʃrhi, Cherkashin, D., Bulakh, M., & Vizirova, E. (2014). *Corpus of Soqotri Oral Literature*, Volume 1. Leiden – Boston: Brill.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Naumkin, V., Kogan, L., ʕIsa Gumʃan al-Daʃrhi, Ahmed ʕIsa al-Daʃrhi, Maysoon Mohammed al-Daʃrhi, Cherkashin, D., Bulakh, M., & Vizirova, E. (2019). *Corpus of Soqotri Oral Literature*, Volume 2. Leiden/Boston: Brill.

Naumkin, V. V., & Porkhomovskij, V. Ja. (1981). *Očerki po etnolingvistike Sokotry* [Essays on Ethnolinguistics of Soqotra]. Moskva: "Nauka".

Nöldeke, Th. (1910). Lehnwörter in und aus dem Äthiopischen. In *Th. Nöldeke, Neue Beiträge zur Semitischen Sprachwissenschaft*, pp. 31–66. Strassburg: Verlag von Karl J. Trübner.

Piamenta, M. (1990). *A Dictionary of Post-Classical Yemeni Arabic*. Leiden et al.: Brill.

Potts, D. T., Parpola, A., Parpola, S., & Tidmarsh, J. (1996). Guḫlu and Guggulu. *Wiener Zeitschrift für die Kunde des Morgenlandes*, 86 [*Festschrift für Hans Hirsch zum 65. Geburtstag gewidmet von seinen Freunden, Kollegen und Schülern*], 291–305.

Powels, S. (1999). The Historical and Cultural Background of Two Non-Basic Terms for 'Blue' in the Arabic Dialect of the Negev Bedouin. In A. Borg (Ed.), *The Language of Color in the Mediterranean* (pp. 152–166). Stockholm: Almkvist/Wiksell.

Rakhilina, E. (2007). Linguistic construal of colors. In R. E. MacLaury, G. V. Paramei, & D. Dedrick (Eds.). *Anthropology of Color. Interdisciplinary multilevel modeling* (pp. 363–377). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.137.24rak

Rakhilina, E., & Paramei, G. (2011). Colour terms. Evolution via expansion of taxonomic constraints. In C. B. Biggam, C. A. Hough, C. J. Kay, & D. R. Simmons (Eds.), *New Directions in Colour Studies* (pp. 121–131). Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/z.167.15rak

Reichmuth, S. (1981). Die Farbbezeichnungen in Sudanesisch-Arabischen Dialekten. *Zeitschrift für Arabische Linguistik*, 6, 56–68.

Reinisch, L. (1887). *Die Bilin-Sprache. Zweiter Band: Wörterbuch der Bilin-Sprache*. Wien: Alfred Hölder.

Rubin, A. D. (2010). *The Mehri Language of Oman*. Leiden – Boston: Brill.

Sabar, Y. (2002). *A Jewish Neo-Aramaic Dictionary. Dialects of Amidya, Dihok, Nerwa and Zakho, Northwestern Iraq*. Wiesbaden: Harrassowitz.

Saleh Mahmud Idris. (2015). *A Comparative Study of the Tigre Dialects*. Semitica et Semitohamitica Berolinensia 18. Aachen: Shaker Verlag.

Sima, A. (2009). *Mehri-Texte aus der jemenitischen Šarqīyah*. Transkribiert unter Mitwirkung von ʕAskari Ḥugayrān Saʕd, bearbeitet und herausgegeben von Janet C. E. Watson und W. Arnold. Wiesbaden: Harrassowitz Verlag.

Soden, W. von. (1965–1981). *Akkadisches Handwörterbuch*. 3 vols. Wiesbaden: Harrassowitz.

Sokoloff, M. (2009). *A Syriac Lexicon: A Translation from the Latic, Correction, Expansion, and Update of C. Brockelmann's Lexicon Syriacum*. Winona Lake, Indiana: Eisenbrauns; Piscataway, New Jersey: Gorgias Press.

Sokoloff, M. (1992). *A Dictionary of Jewish Palestinian Aramaic of the Byzantine Period*. Ramat-Gan: Bar Ilan University Press.

Sokoloff, M. (2002). A Dictionary of Jewish Babylonian Aramaic of the Talmudic and Geonic Periods. Ramat-Gan: Bar Ilan University Press; Baltimore: Johns Hopkins University Press.

Sokoloff, M. (2014). *A Dictionary of Christian Palestinian Aramaic*. Leuven: Peeters.

Stol, M. (1989). Old Babylonian Ophthalmology. In M. Lebeau, & Ph. Talon (Eds.), *Reflets des deux fleuves* (pp. 163–166). Leuven: Peeters.

Täkkəʔä Täsfay. (1999). Zämänawi mäzgäbä ḳalat təgrəñña [Modern Tigrinya Dictionary]. Asmara: ʔAḥtämti ḥədri.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Täsämma Habtä Mikaʔel & Käśśate Bərhan Täsämma. 1951 A.M. [1958/59 A.D.]. *Yäʔamarəñña mäzgäbä ḳalat* [Amharic Dictionary]. Addis Abäba: Artistik Mattämiya Bet.

Tropper, J. (2002). *Altäthiopisch. Grammatik des Geʿez mit Übungstexten und Glossar*. Münster: Ugarit-Verlag.

Vassilevich, A. P. (1988). Colour Names in Amharic. In A. Gromyko (Ed.), *Proceedings of the Ninth International Congress of Ethiopian Studies* (pp. 142–148). Vol. 4. Moscow: Nauka Publishers.

Wagner, E. (1997). Harari. In R. Hetzron (Ed.), *The Semitic languages* (pp. 486–508). London: Routledge.

Watson, J. C. E. (2012). *The Structure of Mehri*. Wiesbaden: Harrassowitz Verlag.

Wedekind, K., & Wedekind, C. (2002). Initial SLLE Survey of the Zway Area. SIL International 2002, www.sil.org/silesr/2002/SILESR2002-063.pdf

Wehr, H. (1979). *A Dictionary of Modern Written Arabic (Arabic-English)*. Ed. by J. M. Cowan. Fourth Edition. Wiesbaden: Harrassowitz.

Wetter, A. (2010). *Das Argobba: eine deskriptive Grammatik der Varietät von Shonke und T'ollaha (Zentraläthiopien)*. Köln: Rüdiger Köppe Verlag.

Wierzbicka, A. (1996). *Semantics. Primes and Universals*. Oxford/New York: Oxford University Press.

Wolde Michael Kelecha. 1980 A.M. [1987 A.D.]. *A Glossary of Ethiopian Plant Names*. Addis Ababa: [publisher not identified].

Wood, F. A., & Roberts, G. A. F. (2005). Natural Fibers and Dyes. In Sir Gh. Prance, & M. Nesbitt (Eds.), *The Cultural History of Plants* (pp. 287–313). New York/London: Routledge.

az-Zabīdī. (1986–1998). *Tāǧ-u l-ʕarūs-i*. Maṭbaʕat-u l-ḥukūmat-i: Kuwait.

Zelealem Leyew. (2016). The Description of Colour Terms in Ethiopian Languages. *Eastern Africa Social Science Research Review*, *32*(2), 53–86. https://doi.org/10.1353/eas.2016.0008

Zohary, M. (1982). *Plants of the Bible: A complete handbook*. Cambridge et al.: Cambridge University Press.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lexicalization patterns in color naming

The case of Modern Hindī

Andrea Drocco and Orsola Risato
Ca' Foscari University, Venice

The aim of this chapter is to draw attention to the various lexicalization patterns which are used in Modern Hindī to define colors, starting from basic color terms and including the lexicalization strategies which are employed to describe different shades. The following analysis, grounded in a corpus-based research, as well as in the most authoritative grammars and dictionaries, will outline both the "metonymy"-type of Hindī color lexicalization patterns and the "approximation"-type mechanisms. We will show that in the first case, Hindī color terms are usually created by means of the suffix *-ī* [N_{entity} + suffix], while the "approximation" is conveyed by (1) the suffix *-sā*, and (2) the reduplication of the color term, meant as the juxtaposition of the same repeated color adjective.

Keywords: lexicalization, color, Hindī, suffixes, reduplication

1. Introduction

Since the 1960s, starting with Berlin and Kay's study (1969), much attention has been paid to the cross-linguistic categorization and development of color terms. As a consequence, thanks to the works of various scholars, we possess numerous investigations into color terms of IE and non-IE languages. However, it is rather surprising that no significant research has been conducted on Indo-Aryan, and especially Hindī color terms (that is, Modern Standard Hindī). Starting from those which are considered Hindī basic color terms, as recorded in the few studies found in the literature (see Caskey-Sirmons & Hickerson, 1977, pp. 362, 363; see also Berlin & Kay, 1969, about basic color terms in Urdū), the aim of this chapter is to draw attention to a topic that has remained relatively unexplored thus far: the various "lexicalization patterns" that are used in Modern Hindī to define colors and, in particular, to convey different shades of meaning of basic colors. The following analysis, grounded in a corpus-based research and in the most authoritative grammars (see Gurū 1978, p. 73; McGregor, 1977; Platts, 1990) and dictionaries

https://doi.org/10.1075/sfsl.78.09dro
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

(see Śyāmasundardās (Ed.), 1965–75; McGregor, 2002; Platts, 1884; Shakespear, 1834) of this language, is divided into two parts. In the first part, we will focus on the "metonymy"-type of Hindī color lexicalization patterns, whereas in the second part the "approximation"-type mechanisms will be examined. In the second part of the analysis, data will be presented to show the two Hindī strategies which are employed to convey the "approximation"-type meaning in regard to color terms: (a) the suffix *-sā* and (b) the reduplication of the color term.

2. The "metonymy"-type of Hindī colors lexicalization patterns

In this part, we will show that the majority of Hindī non-basic color terms are actually the result of derivation by means of the suffix *-ī*. The function of this suffix is exactly to realize the metonymic shift "OBJECT COLOR for COLOR", as the suffix leads to the creation of a color term starting from the noun of the object that is characterized by the very same color.[1] Therefore:

$$N_{entity} + \text{ -}\bar{\imath} \text{ suffix} = \text{color of } N_{entity}$$

In many cases, the suffix *-ī* is already present in Persian color terms, which are used as direct and genuine loanwords in Hindi[2]; the best-known examples of this kind are:

kiśmiśī 'of the color of *kiśmiś*' (McGregor, 2002, p. 198b) < *kiśmiś* 'raisin' (< Persian)

ābnūsī 'black' (McGregor, 2002, p. 89a) < *ābnūs* 'ebony' (< Persian)

gaṃdumī 'wheat-colored' (McGregor, 2002, p. 249a) < *gaṃdum* 'wheat' (< Persian)

gulābī 'pink' (McGregor, 2002, p. 272b) < *gulāb* 'rose' (< Persian)

pyāzī 'reddish', 'pinkish' (McGregor, 2002, p. 651a) < *pyāz* 'onion' (< Persian)

bādāmī 'light brown' (McGregor, 2002, p. 723b) < *bādām* 'almond (tree nut)' (< Persian)

Only in a handful of instances, the color term or the original word is an Arabic loan, as shown in the example of the two following words:

---

1. Another Hindī strategy to realize the metonymic shift OBJECT COLOR FOR COLOR is via the polysemous suffix *-vālā*. Even though this function of *-vālā* is apparently not mentioned in any Hindī grammar/dictionary, perhaps because it is only used in a more colloquial register, it is possible to find a few examples in our corpus (e.g. Hindī *ret vālā* "sand (colour)").

2. This suffix can be considered as a form in part developed from the Sanskrit secondary suffix *-in* (*–ī* in nominative), that forms adjectives which convey the sense of possession. However, the history of this suffix has yet to be thoroughly studied, especially in regard to its productivity in Hindī. However, the first study on this topic was conducted in Chatterjee (1926).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

zāfrānī 'yellow' (McGregor, 2002, p. 369a)	< *zāfrān* 'saffron' (< Arabic)
qirmizī 'crimson', 'scarlet' (McGregor, 2002, p. 197b)	< *qirmiz* 'cochineal' (< Arabic)

In Hindī, this suffix seems to be very productive in the formation of color terms, regardless of the etymology of the word to which it is attached. We can certainly point some examples where the new color term is a genuine Hindī word derived, by means of *-ī*, from an originally Persian loan:

banafśaī 'violet' (McGregor, 2002, p. 705b)	< *banafśā* 'violet (*Viola odorata*)'
fālsaī 'purple' (McGregor, 2002, p. 681b)	< *fālsā* 'the plant *Grewia asiatica* and its fruits'

However, in most cases, the etymology of the word is Hindī:

katthaī 'dark red in color' (McGregor, 2002, p. 162b)	< *katthā* 'a vegetable extract from the plant or tree *Acacia catechu* (eaten in betel leaf with lime, which it turns red)'
kāhī 'greenish' (McGregor, 2002, p. 195a)	< *kāī* 'scum on stagnant water', 'moss'
kesrī 'saffron-colored' (McGregor, 2002, p. 214a)	< *kesar* 'saffron'
baiṃganī 'purple' (McGregor, 2002, p. 747b)	< *baiṃgan* 'eggplant'
jāmunī 'purple'[3]	< *jāmun* 'the rose-apple tree (*Eugenia jambolana*)' (McGregor, 2002, p. 369b)
nāraṃgī 'orange' (McGregor, 2002, p. 554a)	< *nāraṃg* 'orange (fruit)'
dhānī 'light green' (McGregor, 2002, p. 527b)	< *dhān* 'the rice plant'
gājrī 'of the color of carrot'[4]	< *gājar* 'carrot' (McGregor, 2002, p. 262b)
vasaṃtī 'yellow (as spring flowers)'	< *vasaṃt*/*basaṃt* 'spring' (McGregor, 2002, p. 713b)

As we can see from the lists above, the suffix *-ī* is typically used with fruits and vegetables. However, the lexicalization mechanism related to color naming by means of this suffix is not restricted to these object categories. By contrast, it is also adopted with other kinds of entities:

---

3. This color term is not mentioned in McGregor (2002) or in other dictionaries (Caturvedi 1970; Pathak, 1989). Nowadays, it appears in many websites in Hindī, for example: https://www.bbc.com/hindi/specials/1055_purple_mb (Retrieved 20 July 2018); https://hi.wikipedia.org/wiki/जामुनी (Retrieved 20 July 2018).

4. This color term is not mentioned in McGregor (2002) or in other dictionaries (Caturvedi, 1970; Pathak, 1989). Nowadays, it appears in many websites in Hindī, for example: https://www.patrika.com/bhilwara-news/carrot-and-red-sherwani-do-show-312886/ (Retrieved 20 July 2018); https://www.amarujala.com/uttar-pradesh/meerut/41500041820-meerut-news (Retrieved 20 July 2018).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

āsmānī 'azure, sky-blue'	< *āsmān* 'the sky'
surmaī 'dark-grey', 'blackish'	< *surmā* 'powdered sulphide of antimony'
siṃdurī 'vermilion-colored'	< *siṃdur* 'vermilion'
śarabtī 'orange-yellow'	< *śarabat* 'a typical sweet drink'
lākhī 'dark-reddish'	< *lākh* 'lac'
morpaṃkhī 'deep-blue'	< *morpaṃkh* 'peacock-feather'
fāk̲h̲taī 'reddish-grey colored'	< *fāk̲h̲tā* 'dove', 'wild pigeon'
nāk̲h̲ūnī 'nail-colored'	< *nāk̲h̲ūn* 'nail'
k̲h̲ākī 'dust- or earth-colored'	< *k̲h̲āk* 'dust', 'earth'
qirmizī 'crimson', 'scarlet'	< *qirmiz* 'cochineal'
abīrī 'red'	< *abīr* 'a powder, usually red'
zaṃgārī 'green'	< *zaṃgār* 'verdigris' 'of the color of verdigris'

As already highlighted above, although the original word to which *-ī* is attached is Persian or Hindī, a good number of color terms cited above are genuine Hindī words. Indeed, even though the history of the productivity of this suffix is yet to be thoroughly researched (see Footnote 3), it seems that its adoption to make color terms is especially related to the New Indo-Aryan period, probably due to Persian influence. Moreover, the same suffix in Hindī is not only adopted to form adjectival color terms from object-nouns, but also other kinds of adjectival terms such as, for example:

jaṃglī 'wild (a region)'	< *jaṃgal* 'jungle, forest'/'a wild or uninhabited region'/'wild (animals, plants)'
videśī 'foreign'	< *videś* 'a foreign country'
qānūnī 'legal'	< *qānūn* 'law, regulation'
dehātī 'rural'	< *dehāt* 'country', 'countryside'

This is most likely the result of the fact that the original meaning of this suffix is 'having to do with'. Indeed, in the majority of Hindī dictionaries the following meanings are also reported in regard to the previous four Hindī words:

jaṃglī	'having to do with forest'
videśī	'having to do with a foreign country'
qānūnī	'having to do with law'
dehātī	'having to do with countryside'

The same is true for color terms. In different Hindī dictionaries (i.e., Caturvedi, 1970; Pathak, 1989; McGregor, 2002), for some of the aforementioned terms,

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

besides the color terms, we also find the meaning 'having to do with the OBJECT COLOR', even though for some color terms this original meaning is not attested, perhaps because these terms are now deemed as basic color terms due to their salience. For example:

baiṃganī	'purple'
gulābī	'pink'
nāraṃgī	'orange'
pyāzī	'reddish', 'pinkish'

However, for many others, this original meaning is reported. For example:

śarabtī 'having to do with sherbet', 'of orange-yellow color' (McGregor, 2002, p. 944a); 'of the color of *śarabat*' (Caturvedi, 1970, p. 736a) < *śarabat* 'a typical sweet drink'

surmaī 'of the color of *surmā*', 'dark-grey', 'blackish' (McGregor, 2002, p. 1030a); 'light blue' (Pathak, 1989, p. 763a); 'of the color of *surmā*', 'dark-grey' (Caturvedi, 1970, p. 823b) < *surmā* 'powdered sulphide of antimony'

vasaṃtī/*basaṃtī* 'yellow (as spring flowers)' (McGregor, 2002, p. 713b); 'pertaining to spring', 'of yellow color' (Pathak, 1989, p. 676b); 'pertaining to *vasaṃt*', 'light yellow' (Caturvedi, 1970, p. 692a) < *vasaṃt*/*basaṃt* 'spring'

siṃdurī 'vermilion-colored' (McGregor, 2002, p. 1011b); 'very red' (Pathak, 1989, p. 751a); 'vermilion-colored' (Caturvedi, 1970, p. 806b) < *siṃdur* 'vermilion'

bādāmī 'almond colored', 'light brown', 'almond-shaped', 'made of almonds' (McGregor, 2002, p. 723b); 'of the color or form of almond' (Pathak, 1989, p. 539a); 'almond colored', 'light yellow', 'prepared from almonds' (Caturvedi, 1970, p. 515b) < *bādām* 'almond (tree and nut)'

kiśmiśī 'made with raisins', 'of the color or taste of *kiśmiś*' (McGregor, 2002, p. 198b); 'raisin-colored', 'of raisin' (Caturvedi, 1970, p. 136a) < *kiśmiś* 'raisin'

morpaṃkhī 'deep-blue' (McGregor, 2002, p. 837a); 'of the color of a peacock's feathers, resembling such a color' (Pathak, 1989, p. 627a); 'of the color of a peacock-feather' (Caturvedi, 1970, p. 622a) < *morpaṃkh* 'peacock-feather'

abīrī 'having to do with *abīr*', 'of the color of *abīr*', 'red' < *abīr* 'a powder, usually red'

3. The "approximation"-type of Hindī color lexicalization patterns

Thus far, we have focused on the "metonymy"-type of Hindī color lexicalization patterns; in Section 3, "approximation"-type mechanisms will be examined, that is, two particular Hindī strategies that convey the "approximation"-type meaning in regard to color terms: (a) the suffix *-sā*, as in *lāl-sā* = 'reddish, red-like' from *lāl* = 'red', and (b) the reduplication of the color term, meant as the juxtaposition of the same repeated adjective concerning colors, as in *nilī-nilī ṭopī* (F) = 'bluish cap'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

from *nilā* = 'blue'. When we speak about "approximation", we mean the creation of a new color term that denotes an "attenuation" or an "intensification" of the basic color. It is important to highlight that both the suffix *-sā* and the reduplication process in Modern Hindī are not limited to color lexicalization. On the contrary, they are rather productive and affect different lexical items (nouns, verbs, adjectives, according to the cases). What is of great significance to our analysis, though, is the usage of both these strategies in relation to adjectives. As we will see in Section 3.2, this is very important especially for the reduplication process.[5]

Given these guidelines, what follows is an attempt to outline the productivity and the semantics of (a) the suffix *-sā* (Section 3.1) and (b) the reduplication of the color term (Section 3.2) as lexicalization strategies in Modern Hindī.

3.1 The use of the comparative suffix *-sā*

The suffix *-sā* is a morpheme of Sanskrit derivation, which primarily means 'like'. It is suffixed to adjectives, so it bears gender, number and case agreement markers according to the adjective's inflection (McGregor, 1977; Caracchi, 2002; Kachru, 2006; Montaut, 2012).

As underlined by previous extended studies, whenever *-sā* is attached to an adjective "The semantic notional contours [of the same adjective] get blurred, and this weakening of the distinctive limits is particularly clear with color adjectives" (Montaut 2004, p. 155). As previously noted, this 'weakening' does not necessarily mean attenuation/softening of the basic color. Rather, it seems that the speaker chooses this particular form to distance himself from the basic, universally recognized standard version of it. This distance is determined, or even required, by a specific context in which the standard color is not sharp enough to convey a particular worldview which is strictly intrinsic to the individual speaker. It is not our concern here to determine, once and for all, if this mindset is due to cultural elements, the individual's biography, or community beliefs. Suffice it to say that it is probably the outcome of all these factors.

We chose four sentences to exemplify this lexicalization strategy (Examples 1, 2, 3, and 4). They intentionally include modifications to the same basic colors, so as to better analyze the variable of context. The first one is given in (1) below:

---

5. Most commonly, apart from color adjectives, these two linguistic strategies affect adjectives of taste. We could assert that color and taste are two of the most subjective perceptions that the speaker needs to convey. In this sense, it is even clearer what we mean by "context-dependent" denotative patterns.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(1) *agar āp-k-e dāṁt yā kāl-e paṛ*
If HON-GEN-M.PL tooth.M.PL or black-M.PL appear
āp ham-ār-e die sujhāv-oṁ
HON GEN-M.PL give.PAST.PART.M.PL advice-M.PL.OBL
*par nazar **pīle-se** rahe haiṁ*
on gaze.M.SG yellow-*se* (= yellowish) PRES.CONT-M.PL
to kuch ḍāl sak-t-e.
so some put can-PRES-M.PL
'If your teeth appear *yellowish* or dark, you can take some of our advice.'
(adapted from http://hindi.kyakyukaise.com)

This sentence, taken from a dental clinic advertisement, clearly concerns the social context in which healthy teeth are primarily revealed by their color. Defining them "yellow" would have been inappropriate, as the standard color is not apt in this context. At the same time, choosing a "softened" version of the adjective is not a judgment-free statement. It implies that the teeth could be dirty, maybe even decayed. It is important to highlight that the speaker (or in this case, better to say the impersonal speaker) does not focus on this particular light shade due to some kind of love for chromatic details, but rather because this color conveys a less pleasant meaning than, for example, a bright white would. The second sentence intentionally including modifications to the same basic colors (i.e., *pīlā* 'yellow') is exemplified in (2):

(2) *chāl ke andar k-ī tah*
bark.M.SG inside GEN.F.SG layer.F.SG
tāz-ī avasthā meṃ lāl-sī
fresh-F.SG state.F.SG in red-*sī*
*yā **pīlī-sī** safed aur*
or yellow-*sī* (= yellowish) white and
meṃ bahut kaḍv-ī hot-ī hai
in very bitter-F.SG be.PRES-F.SG
'The inner layer of the bark, when it is fresh, appears as dark, *reddish* or light, *yellowish* and it has a very bitter taste.' (adapted from Bedi, 2004, p. 44)

Actually, here we find two basic colors, *pīlā* ('yellow') and *lāl* ('red'), subjected to the same lexicalization strategy. Unlike in Example (1), no judgment about pleasantness is expressed, while it is noticeable that the speaker needs to describe a single color restricted to a very specific field (in this case, botany). Not only does he resort to the adjectival suffix, but he also adds more color terms (*bhūrī*, 'dark' and *safed,* 'white') to give the most precise description of the object.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(3) *ek badsūrat* ***kālā-sā*** *ādmī*
One ugly black-*sā* (= 'blackish') man.M.SG
Zukām k-ī śikāyat le-kar
cold.M.SG GEN-F.SG complaint.M.SG take-CNV
ḍoctor ke pās ga-y-ā.
doctor by go-PERF-M.SG
'An ugly man with a *quite dark* complexion caught a cold and went to the doctor's to complain.'

(adapted from http://www.amarujala.com/news-archives/entertainment-archives/amar-ujala-joke-hindi-jn)

In Example (3), we have the first line of a (politically incorrect) joke. To really understand the value of *kālā-sā* 'blackish' mentioned in the first line, we need to keep in mind what the aesthetic opposition between a dark complexion and a fair one means in contemporary Indian society. Without examining the potentially racist connotations of this aesthetic, which is not the point of our analysis, it is undeniable that fairness and brightness have been the criteria for social approval for decades (see Mishra, 2015). The speaker is being ironic here, as the joke requires, and the choice of the color adjective is not coincidental. Had he been talking about hair color, this *kālā-sā* could have been considered as a neutral definition for a specific shade, with no further connotations. As it is a person's skin which is described as "blackish, quite dark", we need to consider the huge cultural implication that lies in the choice of this adjective.

On the contrary, in Example (4), as in Example (2), the adjectival suffix has a more neutral connotation. There is no judgment involved in this specific color pattern, but rather a sense of vagueness. The adjectival suffix literally describes a "blackish thing" and, since the color is blurred in the speaker's eyes, it is also difficult to understand what the very thing is (this led us to propose the translation as 'something').

(4) *kinār-e par pahuṃcn-e par*
bank-M.SG.OBL on arrive-INF on
surendra bābū k-ī nigāh
Surendra.M.SG mister.M.SG GEN-F.SG gaze.F.SG
ek ***kālī-sī*** *cīz k-ī*
one black-*sī* (= 'blackish') thing.F.SG GEN-F.SG
taraf ga-ī
direction.F.SG go-PERF.F.SG
'Once he reached the river bank, Mr Surendra's gaze was captured by *something blackish*.' (adapted from Cattopādhyāy, 2011, p. 68)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.2 Reduplication in south Asian languages: The case of modern Hindī

Before proceeding with the analysis of the reduplication models in color lexicalization, we need to explicate which particular reduplicative mechanisms are significant in the context of Modern Hindī, starting from the very categorization of reduplicative forms. Therefore, on the basis of previous studies on reduplication across the world's languages, we will first assume that the reduplication is a specific pattern which occurs whenever systematic repetition of phonological material happens within a word for specific and context-dependent semantic or grammatical purposes (Rubino, 2005, p. 11).[6] Taking this into account, one could focus then on establishing whether the reduplication should be considered in terms of mere phonology or morphology. Our analysis does not intend to provide a definite answer to this question (as it could also differ from one language to another), but rather to define which reduplication patterns are involved and significant in color lexicalization as concerns Modern Hindī. In order to achieve this aim, we need to restrict the reduplication definition to some specific cases. For this purpose, we will follow what has already been thoroughly presented by other scholars (see, among others, Singh, 2005, and Moravcsik, 1978): the reduplication consists of the repetition of phonological material that marks individual and meaningful morphemes. This further condition that is both necessary and sufficient, allows the study of other "reduplicative" patterns, like onomatopoeias and the so-called *echo-words*, as a distinct and peculiar category.

Since we are talking about the repetition of morphologic material, it is extremely important to stress that the repetition itself is never randomly conjugated by the speaker. The repetition patterns are always codified according to phonemic and grammatical rules of the language in question: the alternation of vowel-consonants, the number of syllables and the position of the repeated items in the sentence are not objects of the individual speaker's creativity. We must remember that the reduplicated form is treated as individually meaningful; it becomes an indexical unity just like any other part of a language's lexicon (see references above).

What is left to analyze is the semantics of the reduplication patterns. Why does a speaker choose the modified, reduplicated version of a term, instead of the standard one? Are there any universally valid meanings that could be traced back to

6. According to this first definition, the mere repetition of the *semantic referent* is, of course, to be excluded from the category of reduplication. In the sentences like "She was looking at herself in the mirror" or "He blamed himself for the accident", the semantic referent is, indeed, repeated, but this is not considered as an occurrence of reduplication.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

reduplication? As this still appears to be a *vaexata questio* for Linguistics,[7] we will adhere to what is commonly accepted as a sufficiently comprehensive description of the process: the reduplication expresses the augmented quantity of either emphasis or semantic referents (or, often, of both emphasis and referents). The referent augmentation could concern the multiplicity of referents in the same space – time continuum, as well as the multiplicity of referents in a sequence (Moravcsik, 1978).

According to the seminal works by Abbi (1980, 1985) and Montaut (2008), in the context of the South Asia Linguistic Area,[8] the reduplication pattern appears to be a typical feature of non-Indo-Aryan languages of India and, therefore, Modern Hindī probably maintains this characteristic due to linguistic contamination. In addition, it is worth noting that, most commonly, both the standard form and the reduplication result are equally used in the same language on a diastratic level. If it is also true for Modern Hindī that the reduplication conveys an augmentation of emphasis/referents, this mechanism totally depends on the morpheme involved: in this language, the reduplication happens for nouns, pronouns, adjectives, verbs, and adverbs. Some studies suggest a precise semantic classification of the reduplicated forms. In accordance, we can, therefore, identify some properties of reduplication, like exclusiveness, a variation of emphasis, temporal relation, distributiveness (see Abbi, 1985). Since this classification is, however, very prone to exceptions, for the purpose of our analysis we prefer to comply with a wider consideration: reduplication of a lexical entity modifies its relation with the constituents of the final reduplicated form, and this new relation gives the final meaning to the whole expression (see Montaut, 2008, p. 23). Moreover, it is also worth noting that in Modern Hindī different semantic properties are sometimes distinctive and exclusive of different parts of speech. For example, exclusiveness and distributiveness are idiosyncratic properties of reduplicated numerals and nouns, as shown in Example (5).

(5) *bacc-oṁ* *ko* ***ek-ek*** *miṭhāi* *d-o.*
child-M.PL.OBL DAT one – one sweet.F.SG give-IMPER.SG
'Give a sweet to *every* child'

In this case, the reduplication of the numeral "one" allows to understand that the action of giving is both distributive and exclusive: each child in this individuality will receive one and only one sweet.

7. For a deeper investigation on reduplication as a widespread linguistic pattern, see Bloomfield (1970), Moravcsik (1978), Hurch (2005), and Rubino (2005).

8. For a definition of India as a "linguistic area", see Masica (1976, 2001) and Emeneau (1980).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

On the other hand, when it comes to pronouns, the distinctive property is rather the augmentation of referents, as in Example (6), where the speaker is expecting a list, precise information about the many things the interlocutor saw[9]:

(6) *tumne* ***kyā-kyā*** *dekh-ā?*
2.PL-ERG what-what see-PAST.M.SG
'*What* did you see?'

In the case of verbs, it is more complicated to identify a single semantic property which could widely include many different occurrences of reduplication. Nevertheless, we can identify in "relation of time" the distinctive properties for most reduplicated verb forms. As "relation of time" we mean either the description of a fragmented process or the description of a process in continuity: each single element of the reduplicated form is meant as a slide or a segment. In fact, this particular example could also be considered as a sample of the wider semantic category that we previously called the "augmentation of referents", since "continuance is simply distributivity over time" (Singh, 2005, p. 269). This case is exemplified in (7):

(7) *kha-te-kha-te* *mat* *bol-o*
eat-PRES.PART-eat-PRES.PART no speak-IMP.SG
'Don't speak *while eating*.' (adapted from Montaut, 2008, p. 28)

The analysis of reduplicated adjectives considerably complicates our reduplication pattern scenario, due to a seemingly very prosaic cause: in Modern Hindī adjectives agree with nouns. This means that not only does the reduplicated adjective have its semantic specificity, but it can also convey a certain meaning according to the noun it refers to. For example, a reduplicated adjective preceded by a plural noun will also convey the "distributiveness" given by the plural itself. Therefore, by just recalling our basic concepts of the "augmentation of emphasis" and the "augmentation of referents", without even going into a more detailed examination of the reduplication's semantic weight, it is immediately clear that reduplicated adjectives are affected by both these variations of meaning (mostly, at the same time). As we noticed, the augmentation of emphasis and referents can inflect in exclusiveness, distributiveness, time-specific relation another semantic specificity according to the main constant, that is, the communicative context. Consequently, at this point, one question spontaneously arises: given all the aforementioned semantic properties, the subjectivity of the single speaker and the variable of context, is it really possible to find a few fixed, productive patterns in reduplicated adjectives? This study intends to focus on very confined records, as the reduplication patterns in the

9. Where no reference is given, that means that the example is a direct quotation from a native speaker (see Footnote 1).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

lexicalization of color adjectives are. For our purposes, then, this question will here remain unanswered, as it has been thoroughly examined before (Abbi, 1980, 1985; Montaut, 2004, 2008, 2009; Singh, 2005). Trying to understand, though, which semantic properties could be involved in the process, there are some incontrovertible points to take account of. Talking about the variation of emphasis, it is correct to state that the reduplicated adjective is more emphasized than the basic one "rather for its adequateness to the customer's expectations than for its objective degree" (Montaut, 2008, p. 33). This also means that, as opposed to the simple adjective, the reduplicated one is not descriptive, but strictly subject (speaker) - and situation - dependent (Montaut, 2008, p. 33). As was the case of the suffix *-sā*, reduplicated adjectives lose their standard semantic connotations and their semantic contours get blurred, as shown in Example (8):

(8) *Rāvan* *k-ī* ***baṛ-ī-baṛ-ī*** *āṁkh-eṁ* *th-ī-ṁ*
Rāvan.M.SG GEN-F.SG big-F-big-F eye-DIR.F.PL be.IMPF-F-PL
'Rāvan had *enormous* eyes.' (adapted from Abbi, 1980, p. 86)

In this sentence, the reduplication of the adjective is a strategy used to convey both the emphasis on its single meaning and distributiveness. In reference to the latter, the speaker wants to stress that both of Rāvan's eyes were very big. Knowing the context (Rāvan is a mythological demoniac figure), we can also say that his eyes are big and frightening. By contrast, in Example (9), it becomes clear why it should be better to speak about the "variation" rather than the "augmentation" of emphasis. According to the context, the reduplicated form could convey two meanings that are almost polar opposites. On the one hand, the speaker could find the mango in Example (9) very pleasant, ripe and sweet. On the other hand, the reduplication can mean that the sweetness of this specific, unique mango, is different from the "standard", universally recognized idea of sweetness. The speaker finds it "a kind of sweet", and this specific kind can only be defined in relation to the specific communicative context.

(9) *is* *ṭokrī* *meṁ* *ek* *ām* *h-ai*
3.SG.OBL basket.F.SG in one mango.M.SG be-SG-PRES
'In this basket there's a *sweetish/pleasantly sweet* mango.'
(adapted from Abbi, 1980, p. 93)

3.2.1 *Reduplication of color terms*

If quantity adjectives shape adjectives and quality adjectives can be taken back into a limited spectrum of meanings, although subordinated to the variable of the communicative context, it is clear how taste adjectives are deeply related to the speaker's subjectivity and can, therefore, describe an almost infinite range of meanings. This

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

is exactly what happens to color adjectives and how the reduplication becomes a lexicalization pattern of color terms.

Although any categorization always has some limits, in terms of exceptions, simplicity, and, perhaps, approximation, the purpose of this analysis is to provide some guidelines capable of identifying the productiveness of color adjective reduplication and the linguistic accuracy of this process in creating lexical entities. What is going to be analyzed is a record of reduplicated color adjectives, consisting of authentic linguistic material. As spoken Modern Hindī is quite susceptible to local varieties and English influences, the source of our record has been the language of the media: advertising, TV shows, Bollywood movies, as well the 20th-century literature. Therefore, this study attempted to limit the variation of basic color terms, conveyed by reduplication, by sorting them into four categories: (1) emphasized color; (2) softened color; (3) situational color; and (4) moving/changing color.

3.2.1.1 *Emphasized color*

As it should be clear at this point, "emphasis" is a critical word in the context of reduplication. It would be misleading to consider it as a series of points on a scale, the lowest of them being the basic grade of the adjective. This criticality, besides, had already been discussed in the first studies on reduplication.[10] What is actually emphasized is the quality of the standard color adjective. Reduplication indicates something near the prototypical shade and, at the same time, clarifies that the speaker implicitly appreciates the object. This is illustrated in Example (10):

(10) ***harā-harā*** *lazīz* *panīr* *svād* *meṁ* *yamī yamī…*
green-green appetizing.M.SG. *panīr*.M.SG. taste.M.SG. in yummy yummy
'Tasteful, *appetizing green panīr*, with a "yummy yummy" taste…'

Context:
har mauke par svād kā lutph uṭhāie aur ghar meṁ āsānī se banāīye ***harā-harā*** *svādiṣṭ panīr.*
Enjoy good taste any time and make a delicious easily home-made, appetizing green *panīr*. (adapted from http://www.aapkisaheli.com/articles)

As concerns this sentence, the communicative context is a food advertisement, specifically, a cottage cheese advertisement: *panīr* – a typical Indian cheese – is a common ingredient in Northern Indian cuisine. This cheese is commonly cooked in a spinach-based gravy, which explains the green color of the final preparation. In Example (10), the message conveyed by the reduplication of "green", then, does

10. In 1929, Bailey noticed that in a sentence like "*gorī gorī bālikā kī lāl lāl gālhẽ* 'the rosy cheeks of the prettily fair girl' […] if the idea of emphasis were present, [it would have meant] the hectic cheeks of the deathly pale girl" (p. 512).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

not refer merely to the color, but to a kind of involvement by the audience: this particular vivid green appears to be very pleasant and tempting as it is related to a (potentially) delicious food.

The Example (11) comes from a Bollywood movie soundtrack and relates to a serenade scene.

(11) ***nīle-nīle*** *ambar par, cāṁd jab ā-y-e.*
blue-blue sky in moon when rise-PAST-M.PL
'When the moon rises in the *deep blue* sky.'

Context:
pyār barsāe, hamko tarsāye.
love rained, we pined for
aisā koī sāthī ho, aisā koī premī ho.
a companion like this to be here, a lover like this to be here
pyās dil kī bujhā jāe.
to quench heart's thirst. (adapted from http://www.lyricsindia.net/songs/769)

In this sentence, the "blue" shade of the sky, deep and intense as it can be in the moonlight of a full moon (the moon itself is mentioned in the verse, plus the full moon is a typical feature of love scenes in Indian tradition), is described as such through the lover's eyes, not only as a shade that differs from the basic one. The same seems to be true for the example given in (12), where another love song is under examination:

(12) *ye* ***kāl-ī-kāl-ī*** *āṁkh-eṁ*
3.PL.DIR **black**-F-**black**-F eye-F.PL.DIR
ye *gor-e-gor-e* *gāl.*
3.PL.DIR **white**-M.PL-**white**-M.PL cheek.M.PL
'These *beautiful black* eyes, these *gently fair* cheeks.'

Context:
ye tīkhī-tīkhī nazareṁ ye hiranī jaisī cāl
these sharp glances, this deer-like gait
dekhā jo tujhe jānam/ huā hai burā hāl.
when I saw you my darling/ my condition got bad.
(adapted from http://www.lyricsindia.net/songs/5482)

Here, the reduplication of "black" and "white" is not representative of mere emphasis, as they would mean "very black" and "very white" which would be quite unusual in any description of a lover. Rather, shiny black hair and fair complexion are two essential elements in the traditional definition of a beautiful woman.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.2.1.2 *Softened color*

Following the process that is almost opposite to the "emphasis" one, speakers can also use reduplicated color terms to describe something with a paler or more subdued color than the prototypical one. However, this "softness" characterization is caused by an emphatic relationship between the speaker and the object, as we can observe in Example (13):

(13) *Qai kar-t-e hue uske*
vomit do-PART-M.SG.OBL 3.SG.OBL-GEN-M.OBL
nāk-muṁh se ***pīlā-pīlā*** *saṛe*
nose-mouth from **yellow-yellow** rotten
aṁḍ-e k-ī jardī-sā kuch
egg-M.PL GEN-F.SG paleness-F.SG something
nikl-ā aur āṁkh-oṁ ke āge
come out-PAST.M.SG and eye-M.PL.OBL in front of
aṁdherā chā ga-y-ā.
darkness.M.SG bedim-PAST-M.SG
'He vomited and his face turned into a *yellowish*, rotten egg-like paleness, and darkness bedimmed his eyesight.'

Context:
"auk…auk" karke vah nālī par jhuk gayā.
He bent over the drain, moaning.
ghar ke agvāṛe kī nālī bū mār rahī thī, lekin mannān kī sūṁghne kī quvvat jātī rahī thī.
The house frontage drain was killing the odor, but Mannān's smelling skill was still alive. (adapted from Srimjay, 2001, p. 92)

This passage clearly shows that the character involved in the situation is in discomfort. "Paleness" and reference to "rotten eggs" evoke a chromatic range that instills unease in the reader. In this sense, the reduplicated color term does not indicate a particular lighter shade of yellow, but an unpleasant version of the basic color.

The contrast between the pleasant and the unpleasant shade of the basic color is even clearer in Example (14), where the difference between the previously analyzed *emphasized color* and *softened color* can be noticed:

(14) a. *hare patt-e k-ā*
green leaf-M.PL GEN-M.SG
dukh ***pīlā-pīlā*** *h-ai /*
sorrow.M.SG **yellow-yellow** be-PRES.SG
'Green leaves' sorrow is *yellowish*'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

b. *pīle* *patt-e* *k-ā*
yellow leaf-M-PL GEN-M.SG
sapnā ***harā-harā*** /
dream.M.SG **green-green**
'Yellow leaves' dream is *bright green*'

Context:
is lie jab naujavān haṃs-te haiṃ / surajmukhī khil-te haiṃ /
For this reason, when youngsters smile, sunflowers blossom
aur… hāṃ, is lie būḍhoṃ / ke āṃsū jhilmilā-te haiṃ / maṇiyoṃ kī tarah…
And…yes, for this reason, old peoples' tears glitter/ as precious stones…
(adapted from http://kavitakosh.org)

It is important here to stress the fact that "yellowish" leaves express a negative feeling, while the "bright green" ones convey positive emotions.

3.2.1.3 *Situational color*

According to what has previously been noted, it seems that any case of a reduplicated color adjective could be labeled as "situational". This is indeed true if we agree that the semantic connotation of the reduplicated adjectives is strictly context-related and communicative situation-dependent. Nevertheless, we noticed that the previous two categories describe, respectively, a brighter and a more subdued shade of the basic color, regardless of context. There are, instead, some particular cases in which the speaker has a cultural/situational mental image of a color and he recognizes it in the particular shade he sees at a particular moment. The specific shade is quite close to the prototypical, universally recognized basic color, but it is closer to a strictly personal idea that the speaker has of that color. This is what we mean by "situational" and a clear example of this specific use of a reduplicated color term is given in Example (15):

(15) *khiṛk* *se* *ātīṁ*
window-F.SG.OBL from come.PART-F-PL.
sūraj *kī* ***lāl-lāl*** *kirṇ-eṁ*
sun.M.SG GEN-F-SG. **red-red** ray-F.PL
sīdhe *muṁh* *par* *paṛ*
directly face.M.SG on pass
rah-ī *haiṁ.*
PRES.CONT-F PL
'The *red* sun rays pass through the window and get directly to my face.'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Context:
barsātoṁ ke bād kī khopṛī caṭkānevāle garmī kā takāzā lie.
I need the heat that comes after the monsoon season and sticks to one's head.
(adapted from Śrī, 2001, p. 22)

Let us consider the context of this sentence. It is probably sunset since the speaker who is in the room describes the sun rays passing through the window as red. They could also be defined as "reddish", as we imagine the typical sunset light. The element that gives a specific connotation to the reduplicated color term is the following passage: the speaker has clearly in mind how the evening light appears in the specific season following the monsoon season. Keeping in mind the totality of sensations (the hotness, the sticky feeling), he/she compares the color that appears in this precise moment to the color that is fixed in his/her memories, and they correspond.

In Example (16), this "situational" element is even clearer: the speaker sees in the present specific shade of blue (that is also implicitly pleasant) a prototypical sky-blue color that belongs to his/her personal mental record:

(16) *yah* ***nīlā-nīlā*** *ākāś* *dekh-kar*
This **blue blue** sky.M.SG see-CNV
mujh-e *pahāṛ* *meṁ* *bacpan*
1.SG.OBL-DAT mountain in childhood
k-ī *yād* *ā-ī.*
GEN-F memory.F.SG come-PAST.F.SG
'Seeing this *bright blue* sky, I remembered my childhood in the mountains.'
(adapted from Montaut, 2008, p. 34)

Again in Example (17), the reduplication of the color term "green" describes a particular shade that the speaker clearly associates with a personal mental image of that same color:

(17) *har-e-har-e* *ām /* *peḍ*
green-M.OBL-**green**-M.OBL mango.M.PL tree
meṁ *lag-e* *h-aiṁ /*
in hang-PAST.PART.M.PL be-PRES.M.PL
'Green mangoes / are hanging on the tree /'

Context:
kacce kacce ām / yah peḍ par lage hue /
Still unripe mangoes/ are hanging on this tree/
kacce ām / har koī dekh rahā hai.
Everyone is seeing/ the unripe mangoes.
(adapted from http://vinodbissa.blogspot.it/2010/04)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Unripe mangoes are characterized by a particular type of "green" which is evidently distinct and unique. The speaker has a well-defined pre-construction of this color, which can only be associated with a specific fruit at a specific time of the year. When he sees it in reality, the reduplication is the linguistic pattern that can best describe what the speaker is perceiving.

3.2.1.4 *Changing color*

In some cases, the speaker makes an iconic effort to indicate a color that is not static but changes to a multi-shade variation at the moment of speaking. It happens, for instance, when the speaker needs to describe a person who turns pale or flushes, as in the following Example (18):

(18) *āp-k-ā* *cehrā* ***pīlā-pīlā***
2.HON-GEN-M.SG face.M.SG **yellow-yellow**
ho *rahā hai!*
look PRES CONT-M.SG
'Your face looks *so yellowish*!'

Context:
nihāl ne jhaṭpaṭ savāl hal kie aur kofī lekar māssāb ke pās jā pahuṁcā.
Nihāl immediately solved the question and went to bring Māssāb a coffee.
māssāb savāl jāṁcne lage to nihāl dhīre se bolā, "māssāb, kyā āpkī tabiyat ḳharāb hai?"
Māssāb started to investigate, so Nihāl slowly said: "Māssāb, are you feeling sick?
(adapted from Śrīvāstava, 2015, p. 23)

In many other instances, the particularity of reduplication is its power to reflect the gradual transition from one color shade to another (examples 19 and 20):

(19) *Torch* *k-ī* *rośnī*
torch.F.SG GEN-F.SG light.F.SG
us-k-e *gāl-oṁ* *se*
3.SG.OBL-GEN-M.PL cheek-M.PL.OBL through
lāl-lāl *phūṭ-t-ī.*
red-red filter-PRES-F.SG
'The torch light filters are *becoming red* through her pinkish cheeks.'

Context:
Bahanjī bistar se uṭh jātī. Torch uṭhā lātī. Bāl chitarā ke mere cehre ke pās ātī.
Bahanjī gets out of bed. She brings a torch. She gets close to my face with dishev-eled hair.
muṁh meṁ jaltā torch ghusāe mujhe jagātī.
She wakes me up, holding the torch in her mouth. (adapted from Śrī, 2001, p. 100)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(20) *Tīn part-oṁ vale* ***lāl-lāl***
three layer-M.PL.OBL SUFF **red-red**
sik-e parāṭh-oṁ māṁ
toast-PART.M.PL parāṭhā-M.PL.OBL Mum
golsā kar-ke banā-kar
beautifully round shape-CNV make-CNV
Tilak Rāj k-e hāth-oṁ
Tilak Rāj GEN-M.PL.OBL hand-M.PL.OBL
meṁ de de-t-ī
In give (= put)- RAD give-F.SG-PRES
'His mum makes some *well toasted*, beautifully round shaped, three-layered *parāṭhā* and puts them in Tilak Rāj's hands.'

Context:
pahlī sīṛhī se hī parāṭhoṁ ke banne se nikalne vālī ghī ke dhueṁ kī sugandh nāk meṁ ātī.
do-tīn parāṭhoṁ kā nāśtā kartā tilak rāj.
Just from the first stair you could smell the fragrance of *ghī* that comes out from *parāṭhā* that are going to be baked. Tilak Rāj eats two or three *parāṭhā* for breakfast.
(adapted from http://www.abhivyakti-hindi.org/kahaniyan/2002/satah_se_uper/satah1.htm)

The translation here means to suggest that the "parāṭhā" are now done to a turn, so they have turned a hazel brown, kind of reddish color.

4. Conclusion

The purpose of this chapter was to describe two main "lexicalization patterns" adopted in Modern Hindī to define colors and, in particular, to convey different shades of meaning of basic colors. Section 2 concerns the "metonymy"-type of Hindī color lexicalization patterns, while Section 3 pertains to the two "approximation"-type Hindī strategies by means of (1) the suffix *-sā* and (2) the reduplication of the color term. When we addressed the "approximation" type of the lexicalization process of color naming, we referred to the creation of a new color term describing an "attenuation" or an "intensification" of the starting color. However, we had the opportunity to show that it would be imprecise and simplistic to limit this "variation" mechanism to these two polar opposites. Rather, we would define the two lexicalization patterns at issue as "color shade lexicalization strategies": the semantic limits of the basic color term are disputed according to a previous mental image existent in the speaker himself. In this sense, we can understand the creation

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

of "shades" that cannot be enclosed in mere softened or emphasized versions of the prototypical color. Moreover, if the "metonymy"-type concerns color terms understood as both color adjectives and color nouns (that is, color indexicals), then the "approximation"-type is a pattern which arises from a very specific need of the speaker: describing an object or, in a wider sense, a phenomenal being, in its specificity, in its uniqueness. This explains the context-related nature of the "approximation"-type color lexicalization pattern.

Acknowledgement

This chapter is the result of a productive collaboration between the two authors. Andrea Drocco bears primary responsibility for Sections 1, 2, the introduction of Section 3 and conclusions, whereas Orsola Risato for Sections 3.1 and 3.2. We are grateful to the audience of the workshop 'Lexicalization patterns in colour naming: A cross-linguistic perspective' in the context of the *7th International Morphology Meeting*, 18–21 February 2016, Vienna, where an earlier version of this study was presented, for their critical comments and constructive suggestions. We would also like to thank our native Hindī informants: Neha Tiwari, Chandra Bhushan, and Mohd Raza Ali. All remaining weaknesses are due to our own shortcomings.

Abbreviations

CNV	converb	NIA	New Indo-Aryan
DIR	direct	OBL	oblique
ERG	ergative	OIA	Old Indo-Aryan
F	feminine	PART	participle
HON	honorific	PAST	past
IA	Indo-Aryan	PAST.PART	past participle
INF	infinitive	PERF	perfect
IMPER	imperative	PL	plural
IMPF	imperfective	PRES	present
M	masculine	SG	singular
MIA	Middle Indo-Aryan	SUFF	suffix

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

References

Abbi, A. (1980). *Semantic grammar of Hindi. A study in reduplication*. New Delhi: Bahri Publications.

Abbi, A. (1985). Reduplicative Structures: A Phenomenon of the South Asian Linguistic Area. *Oceanic Linguistics Special Publications*, 1(20), 157–171.

Bailey, T. G. (1929). Repetition of Words in Urdu, Hindi, and Panjabi. *Bulletin of the School of Oriental and African Studies*, 5(3), 511–513. https://doi.org/10.1017/S0041977X00084676

Bedi, R. (2004). *Jungle ke upyogī vrkś*. New Delhi: Rājkamal Prakāśan.

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

Bloomfield, L. (1970). *Language*. London: George Allen & Unwin Ltd.

Caracchi, P. (2002). *Grammatica Hindī* (4th ed.). Torino: Magnanelli.

Caskey-Sirmons, L. A., & Hickerson, N. P. (1977). Semantic shift and bilingualism: variation in the color terms of five languages. *Anthropological Linguistics*, 19(8), 358–367.

Cattopādhyāy, Ś. (2011). *Śubhdā*. New Delhi: Kauṃpal Books.

Caturvedi, M. (1970). *A practical Hindi-English dictionary*. Delhi: National Publishing House.

Chatterjee, S. K. (1926). *The Origin and Development of Bengalī Language*. Calcutta: Calcutta: Calcutta University Press.

Emeneau, M. B. (1980). *Language and Linguistic Area: Essays*. Stanford: Stanford University Press.

Gurū, K. (1978). *Hindī vyākaraṇ* (12th ed.). Vārāṇasī: Nāgarīpracāriṇī Sabhā.

Hurch, B. (2005). *Studies on Reduplication*. Berlin/New York: Mouton de Gruyter. https://doi.org/10.1515/9783110911466

Kachru, Y. (2006). *Hindi*. Amsterdam/Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/loall.12

McGregor, R. S. (1977). *Outline of Hindi Grammar, with exercises* (2nd ed.). Delhi: Oxford University Press.

McGregor, R. S. (Ed.) (2002). *The Oxford Hindi-English Dictionary* (4th ed.). Oxford: Oxford University Press.

Masica, C. P. (1976). *Defining a linguistic area: South Asia*. Chicago: University of Chicago Press.

Masica, C. P. (2001). The Definition and Significance of Linguistic Areas: Methods, Pitfalls, and Possibilities (with Special Reference to the Validity of South Asia as a Linguistic Area). In R. Singh (Ed.), *The Yearbook of South Asian Languages and Linguistics* (pp. 205–267). New Delhi – Thousand Oaks – London: Sage Publications.

Mishra, N. (2015). India and Colorism: The Finer Nuances. *Washington University Global Studies Law Review*, 14(4), 731–748.

Montaut, A. (2004). *A Grammar of Hindi*. Munchen: Lincom.

Montaut, A. (2008). Reduplication and echo-words in Hindi. In R. Singh (Ed.), *Annual Review of South Asian Languages and Linguistics* (pp. 22–91). Berlin: Mouton de Gruyter.

Montaut, A. (2009). Formes et Valeurs de la Réduplication en Hindi/Ourdou. *Faits de Langue (Série Les Cahiers de l'Inalco)*, 1(29), 9–44.

Montaut, A. (2012). *Le Hindi*. Paris: Leuven Peeters.

Moravcsik, E. (1978). Reduplicative constructions. In J. H. Greenberg et al. (Eds.), *Universals of Human Language. Volume 3: Word Structure* (pp. 297–334). Stanford: Stanford University Press.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pathak, R. Ch. (1989). *Bhargava's Standard Illustrated Dictionary of the Hindi Language (Hindi-English Edition)* (5th revised and enlarged ed.). Varanasi: Bhargava Book Depot.

Platts, J. Th. (1990 [1878]). *A Grammar of the Hindustani or Urdu Language* (2nd ed.). Delhi: Munshiram Manoharlal.

Platts, J. Th. (1884). *A dictionary of Urdu, classical Hindi, and English*. London: W. H. Allen & Co.

Rubino, C. (2005). Reduplication: Form, Function and Distribution. In B. Hurch (Ed.), *Studies on Reduplication* (pp. 11–30). Berlin/New York: Mouton de Gruyter. https://doi.org/10.1515/9783110911466.11

Shakespear, J. (1834). *A dictionary, Hindustani and English: with a copious index, fitting the work to serve, also, as a dictionary of English and Hindustani*. 3rd ed., much enl. London: Printed for the author by J. L. Cox and Son: Sold by Parbury, Allen, & Co.

Singh, R. (2005). Reduplication in Modern Hindi and the Theory of Reduplication. In B. Hurch (Ed.), *Studies on Reduplication* (pp. 263–282). Berlin/New York: Mouton de Gruyter. https://doi.org/10.1515/9783110911466.263

Śrī, G. (2001). *Tirohit*. Nayī Dillī: Rājkamal.

Śrīvāstava, M. (2015). *Māster Sāhab*, Sī. Bī. Tī. Prakāśan.

Srimjay. (2001). *Naṃgā*. New Delhi: Radhakrishna Prakashan Private Limited.

Śyāmasundardās, D. (Ed.). (1965–75). *Hindī śabd-sāgar, arthāt hindī bhāṣā kā ek bṛhat koś* (2nd ed.). Banaras: Nāgarīpracāriṇī sabhā.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Complex color denomination in French and Occitan

Xavier Bach, Anetta Kopecka and Benjamin Fagard
University of Oxford / University of Lyon / Lattice, CNRS, ENS & Université Sorbonne Nouvelle

In this chapter, we investigate color naming in French and Occitan. It is well known that French, compared to other Romance languages, has a tendency to be very analytic. This raises the following question: do speakers of French and other Romance languages (here, Occitan) differ in morphological strategies used to express meaning in the semantic domain of color? To investigate this question, we compare the color descriptions provided by speakers of French ($N = 20$) and speakers of Occitan ($N = 20$). The results show some striking differences between the two languages in lexical and morphosyntactic strategies used to name colors. Speakers of Occitan employ secondary color terms using a variety of derivational suffixes, which are absent from the French data despite their existence in the language.

Keywords: color naming, Romance, French, Occitan, experimental studies, morphological complexity

1. Introduction

The conceptual domain of color and color naming has been widely studied across languages and within disciplines as diverse as anthropology, linguistics, psycholinguistics, and cognitive science. Driven by a very influential study conducted by Berlin and Kay (1969) on *basic color terms* (i.e., morphologically non-derived, not restricted to a particular class of objects, cognitively salient, and present in each speaker's idiolect), many researchers have sought to examine universal dimensions of color naming and investigate cognitive invariants in color categorization (e.g., Heider & Olivier, 1972; Kay, Berlin, & Merrifield, 1991). Others, on the contrary, sought to assess the impact of inter-linguistic differences observed in the repertoire of color terms on visual perception (Roberson, Davies, & Davidoff, 2000; Roberson,

https://doi.org/10.1075/sfsl.78.10bac
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

Davidoff, Davis, & Shapiro, 2005). This debate is by no means new, as can be seen by the following comment opening Ott's (1899) book on colors in Old French: 'But is language a faithful mirror, do the terms it uses give back in minute detail what the organs of sight have perceived? These questions have sparked off a debate which has gone on for a number of years, and is not yet over' (p. ix).[1]

In Romance languages, as in many other Indo-European languages, research on color terms is abundant. Researchers have examined various aspects of the color lexicon such as, for example, the semantics of color terms (see e.g., Grossmann (1988) for Catalan, Castilian, Italian, Romanian, and Latin); their lexical properties and the constructions in which they occur (e.g., Molinier, 2001 and 2006, for French); their derivational morphology (see e.g., Mora & Lécrivain, 2004, for a comparative study of French and Spanish, and Molinier, 2005, for French); the use of color terms and syntactically complex constructions (Dubois & Grinevald, 1999 and 2003, for French); the diachrony of color terms (e.g., Ott, 1899; Kristol, 1978; Dworkin, 2016a,b) or, from a more philological perspective, their use in earlier stages of the language in prose or poetry (see e.g., Doebler, 2010, for Occitan and Italian), to mention just a few studies. We did not find studies dedicated to the color vocabulary in Occitan, though.

In this chapter, we address the use of simple vs. complex terms for color naming in two Gallo-Romance languages, French and Occitan.[2] While the use of primary (non-derived) terms is common in these languages, both have secondary color terms derived from primary terms by way of suffixation, such as, for instance, the suffixes *-astre* in Occitan (*verdastre* 'greenish') and *-âtre* in French (*verdâtre* 'greenish'), which carry an approximative meaning comparable to *-ish* in *greenish*.

---

1. "Mais la langue est-elle un miroir fidèle, les désignations qu'elle emploie rendent-elles, avec une minutieuse exactitude, ce que les organes de la vue ont perçu? Ces questions ont donné lieu à une controverse qui dure depuis nombre d'années et n'est point encore terminée" (Ott, 1899, p. ix).

2. Occitan is generally described as a group of Gallo-Romance dialects spoken in the South of France, and in limited areas in Monaco and the North of Spain and Italy. These dialects differ quite widely in all respects, but are generally considered to be linked to one and the same language, with the possible exception of Gascon (South-Western France). While there were possibly as many as ten million native speakers in the early 20th century (Tesnière, 1928, p. 387), the situation of Occitan today is rather precarious. This remains true despite the attempts towards revitalization (which remains a complex issue, see Costa, 2017), with roughly 100,000 speakers (Bernissan, 2012), all bilingual. There are little or no young native speakers: children learning Occitan mostly do so at school, in specialized schools called *calandretas*. It has recently gained some official recognition, as one of the official languages of the Pyrenees-Mediterranean Euroregion (Pirenèus-Mediterranèa, in Occitan). In this chapter, we describe mainly the variety spoken in and around the city of Toulouse, also known as Lengadocian.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Finally, both languages allow for adjective-adjective (as well as noun-adjective and noun-noun) compounds (*(le) bleu vert* '(the) green blue'),[3] as well as prepositional compounds (*(un) vert d'eau* '(lit.) (a) green of water'). Thus, one might expect to find some similarities between these two languages in the use of linguistic means to name colors.

However, it is well known that French, compared to other Romance languages, has a tendency to be very analytic, a trait that has led some authors to classify it as "more grammaticalized", on the whole than other Romance languages (see Schwegler, 1990; Lamiroy, 2011; De Mulder & Lamiroy, 2012). As has been noted in previous research, grammaticalization does not affect homogeneously all linguistic categories within a given language (Hopper & Traugott, 2003; Marchello-Nizia 2006), and this is also true for languages within families, e.g., German / Dutch / English (König & Gast, 2007), where some languages seem to be more affected by processes of grammaticalization than others, and thereby form a cross-linguistic grammaticalization cline. Similarly, French stands out among Romance languages for a series of morphosyntactic features (Lamiroy, 1999, 2001, 2011; Lamiroy & Pineda, 2017; Carlier, 2007; Carlier & Lamiroy, 2018; De Mulder & Lamiroy, 2012; Fagard, 2010, 2011; Fagard & Mardale, 2012). More specifically, there seems to be a cline that goes from less grammaticalized languages, such as Portuguese or Romanian, to more grammaticalized languages such as French, with in-between cases represented by Spanish on the less grammaticalized side, and Italian, Catalan, and Occitan on the more grammaticalized end of the continuum, as shown in Figure 1.

Figure 1. (Hypothetical) grammaticalization cline of the Romance languages.

3. Depending on the context, French and Occitan color terms can be realized either as adjectives (*bleu* 'blue', *vert* 'green', *bleu vert* 'green blue') or as nouns (*le bleu* 'the blue', *le vert* 'the green', *le bleu vert* 'the green blue'). In this study, the speakers of French did not use the article (neither the definite *le* nor the indefinite *un*) in response to the question 'What color is it?', hence, we consider them as adjectives. In contrast, speakers of Occitan sometimes used determiners, which shows that they were using color nouns at least part of the time. We coded as color *nouns* only occurrences in which the color term is preceded by a determiner.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Thus, when comparing French with other Romance languages, we notice that French has developed a number of analytic constructions that are absent from other Romance languages and that show a higher degree of grammaticalization. This is, for example the case for the expression of the near future shown in (1): while Italian uses a synthetic form *verrò* 'I will come', French typically uses an analytic construction *je vais venir* to convey the same meaning. Though a synthetic form also exists (1c), it tends to become less frequent in French (e.g., going down from five occurrences per million in the 18th c. to three in the 20th and 21st c., in the Frantext database).

(1) a. IT *verrò domani*
'I'll come tomorrow'
b. FR *je vais venir demain*
'I'll come tomorrow (lit. I go.PRS.1SG come.INF tomorrow)'
c. FR *je viendrai demain*
'I'll come tomorrow'

Similarly, for the expression of the recent past, Portuguese makes use of a synthetic form, such as *fiz* 'I have done' in (2a), while French typically uses an analytic construction formed of the present tense of the auxiliary verb *avoir* 'have' together with the past participle of the lexical verb, such as *j'ai fait* in (2b). As in the case of simple future, a synthetic form also exists (2c), though it is seldom used in Colloquial French, and also tends to disappear from Standard French (e.g., *je fis* decreased in frequency, from a relative frequency of 92 occurrences per million in the 18th c. to 32 in the 21st c., in the Frantext database).

(2) a. PT *sempre fiz isto*
b. FR *j'ai toujours fait cela*
'I always did that (lit. I have always done that)'
c. FR *je fis cela*
'I did that'

For indefinites and partitives,[4] the French constructions involve the presence of a special determiner, the grammaticalized form of the preposition *de* 'of' as in (3b and 4b), while Spanish uses bare indefinites (3a) and Romanian uses bare partitives (4a). The determiner can be omitted only in specific constructions, as in (3c) or (4c).

4. See e.g., Carlier and Lamiroy (2014).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(3) a. Sp *tengo amigos en Lisbona*
b. Fr *j'ai des amis à Lisbonne*
'I have friends in Lisbon'
c. Fr *j'ai amis et famille à Lisbonne*
'I have friends and family in Lisbon'

(4) a. Rm *mie îmi place să fac pâine*
b. Fr *j'aime faire du pain*
'I like to make bread'
c. Fr *j'aime faire peur aux gens*
'I like scaring (lit. to make fear) people'

French also shows less derivation than other Romance languages. This is particularly noticeable in evaluative morphology, which is no longer productive in French. For instance, while Portuguese and Italian (among other Romance languages) use affixation for diminutives and augmentatives as shown in examples (5a) and (6a), French tends to use adjectival modification in the same contexts, as shown in examples (5b) and (6b); derivation seems often impossible (in the case of *café*), lexicalized (as in *fouet* 'whip', for which the non-suffixed equivalent is not available in Modern French, or *tablette* 'shelf, pad, tablet PC', for which the link with *table* 'table' might not be transparent) or marginal (as in 5c).[5]

(5) a. Pt *um cafezinho*
b. Fr *un petit café*
'(lit.) a small coffee'

(6) a. It *un filmaccio*
b. Fr *un film pourri*
'a bad movie'
c. Fr ?*un filmounet*, *un filmouille*
'a small/bad movie'

The central hypothesis we want to investigate in this chapter is whether the Romance grammaticalization cline might have implications for color naming. Derivational suffixes are present in both French (7) and Occitan (8), and in both languages, the suffixes (*-âtre* in French and *-astre* in Occitan) convey an approximative meaning, similar to the English suffix *-ish*.

5. For instance, *filmounet* and *filmouille* scored respectively 1,140 and 144 hits in a simple web search , while the frTenTen corpus (accessed via SearchEngine) shows only 6 occurrences of *filmounet* (and none of *filmouille*) for over 2 million occurrences of *film* 'movie'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(7) Fr *verd**âtre***
'greenish'

(8) Oc *verd**astre***
'greenish'

The question that arises then is whether speakers of French and Occitan differ in linguistic strategies used to express meaning in the semantic domain of color. The hypothesis, in line with the theory of the Romance grammaticalization cline, is that French should be less derivational than other languages of the family. It is thus expected that speakers of French will use some derivation (color terms bearing derivational affixes), but that Occitan speakers will use it more frequently. Conversely, we might also expect that Occitan will show some constructional hedging, but that more of it will occur in French (examples 9 to 12). Finally, one might expect the difference between the two languages to be most obvious when participants are naming colors which are further from focal colors and, thus, at the borderline between two categories, since focal colors seem to be more stable across languages and speakers alike (see e.g., Berlin & Kay, 1969; Kay & McDaniel, 1978).

(9) Oc *jaune verd-enc*
yellow green-ish
'a greenish yellow'

(10) Oc *blau verd, blau-astre*
blue green blue-ish
'a blue green, bluish'

(11) Fr *bleu **assez** pâle*
blue quite light
'a rather light blue'

(12) Fr *bleu **très très** pâle*
blue very very pale
'a very very pale blue'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. Methodology

The methodology used for collecting the data in French and Occitan is based on the description of color chips designed for the cross-linguistic project *Evolution of Semantic Systems across Indo-European languages* (EoSS project, Majid & al., 2015). The task consists in naming and describing 84 color chips displayed on a neutral grey background, in random order (but always the same order across participants). In Occitan, there were 20 participants, 10 women and 10 men aged 18 to 82, all native speakers from the Languedoc region (Southern France). The wide age range is due to the fact that, on account of the socio-linguistic situation of Occitan, it is rather difficult to find native speakers of the language. In French, there were also 20 participants, 15 women and 5 men, aged 18 to 26, all living in Lyon (South-Eastern France). For a total of 84 color chips, the task thus yielded 1680 descriptions in each language.[6]

Each of the descriptions was annotated at three levels. The first level is concerned with semantics, for which we distinguish primary color terms such as *gris* 'grey', and analogical color terms such as *marron* 'chestnut, brown'. The second level is concerned with the morphology of the color terms, distinguishing primary terms which are synchronically simple and underived, such as French *blanc* 'white', and secondary terms which bear one or more affixes, such as in Occitan *marron-às* 'brown-ish'. At the level of morphosyntax, we distinguish compounding, for example French *vert bleu* 'blue green' (Noailly, 2005), from adverbial modification such as French *bleu clair* 'light blue' (Kleiber 2007), and from verbal constructions as in French *blanc tirant sur le gris clair* 'white verging on light grey'.

3. Color naming strategies in French and Occitan

Let us first examine what kinds of linguistic strategies French and Occitan speakers tend to select for naming colors, and to what extent they are similar or different in their reliance on different constructions. As Table 1 shows, in the EoSS corpus, French and Occitan display diverging strategies for the naming of color chips. In particular, when naming colors, French speakers have a stronger preference for using simple color terms than Occitan speakers: 70.1% of the descriptions involve a simple term in French, while only 44.3% do so in Occitan. Interestingly, both

6. Data on Occitan were gathered by Xavier Bach, transcribed and annotated by Xavier Bach and Benjamin Fagard. Data on French were gathered and transcribed by Camille Frouin (PhD student, DDL, University of Lyon), and annotated by Camille Frouin and Anetta Kopecka.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

languages make comparable use of compounding and modifying expressions: in both French and Occitan, the use of compounds accounts for 7.9% of color descriptions; as for modifying expressions, they account for 19.8% in French and 21% in Occitan.

Table 1. Types of color naming strategies in French and Occitan EoSS data.

Naming strategies *	**French (*N* = 1680)**	**Occitan (*N* = 1680)**
Simple terms	70.1% (1178)	44.3% (744)
Compounds	7.9% (133)	7.9% (132)
Modifying expressions	19.8% (333)	21.0% (353)
of which derived terms	–	*1.0% (17)*
Complex descriptions	2.1% (36)	26.8% (451)

* We define modifying expressions as those in which a color term (adjective or noun) is modified by another term, generally an adjective, an adverb, a suffix or a combination thereof, as in *a light blue*, *bluish*, or *very blue*. Complex descriptions are those in which participants describe a color with more complex syntactic strategies, as in *a kind of light blue that almost seems grey*. Complex descriptions may include cases of derivation, compounding or modifying expressions, as in *a light bluish grey* (counted *only* as a complex description).

The difference between the two languages is statistically significant both for simple ($\chi^2 = 229.0$, $p < .01$), and for *complex* descriptions ($\chi^2 = 413.6$, $p < .01$). On the other hand, there is no significant difference in the number of *compounds* ($\chi^2 = 0.0$, $p > .05$) or *modifying expressions* ($\chi^2 = 0.7$, $p > .05$). With respect to derivation, the number of occurrences is too limited for statistical purposes, but there is a clear qualitative difference, since they are absent from the French data.

Thus, one of the main differences between French and Occitan concerns the use of secondary color terms (i.e., those which show derivational affixes), and even more noticeably the use of complex descriptions for naming colors. As regards derived terms, although the numbers in Table 1 are very low, such terms are relatively common in the Occitan data: the use of such terms alone accounts for 1% only (Table 1), but, as we will see later, they are also used in other, more complex, constructions such as compounds, modifying expressions, and complex descriptions, with a total of 76 occurrences (4.5% of all descriptions). Besides, the derivational suffixes used to form these terms vary: ten different suffixes were identified in the data. Conversely, as we will see below, such derived terms are nearly absent from the French data (two occurrences or 0.1% of all descriptions).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.1 Simple terms: Primary terms and analogy

As we have shown in Table 1, among different strategies used in color naming, the use of simple terms is the most pervasive in the French and Occitan EoSS data. Such simple terms are of three types: (a) primary color terms, (b) conventionalized terms formed by analogy, and (c) non-conventionalized terms used by analogy.[7] We give below the inventory of such terms in the EoSS data.

a. primary color terms

(13) Fr *beige* 'beige' (2), *blanc* 'white' (23), *bleu* 'blue' (194), *gris* 'grey' (40), *jaune* 'yellow' (45), *noir* 'black' (28), *rouge* 'red' (35), *vert* 'green' (282), *violet* 'purple' (155)

(14) Oc *verd* 'green' (149), *blau* 'blue' (102), *negre* 'black' (41), *jaune* 'yellow' (40), *gris* 'grey' (31), *roge* 'red' (27), *blanc* 'white' (24), *brun* 'brown' (1)

b. conventionalized terms formed by analogy

(15) Fr *bordeaux* 'bordeaux' (6), *caca d'oie* 'poo-of-goose' (3),[8] *indigo* 'indigo' (1), *kaki* 'kaki' (14), *marron* 'brown, (lit.) chestnut' (88), *mauve* 'mauve' (13), *orange* 'orange' (66), *parme* 'parma' (3), *pourpre* 'purple, crimson' (4), *rose* 'rose' (138), *turquoise* 'turquoise' (9), *violine* 'violine' (1)

(16) Oc *violet* 'purple' (106), *ròse* 'pink' (91), *irange* 'orange' (41), *marron* 'brown' (38), *malva* 'mauve' (10), *fuchsia* 'fuchsia' (4), *turqués* 'turquoise blue' (4), *caquí* 'khaki' (3), *salmon* 'salmon pink' (3), *cacà d'auca* 'greenish yellow (lit. poo-of-goose)' (2), *lillà* 'lilac' (2), *òcre* 'ochre' (1), *crèma* 'cream' (1), *caramèl* 'caramel' (1)

c. non-conventionalized terms used by analogy

(17) Fr *anis* 'anise' (1), *aubergine* 'aubergine' (1), *fuchsia* 'fuchsia' (10), *magenta* 'magenta' (1), *paprika* 'paprika' (2), *prune* 'plum' (2), *saumon* 'salmon' (11)

(18) Oc *aubergina* 'eggplant' (4), *grenat* 'garnet' (4), *beige* 'beige' (3), *framboèsa* 'raspberry' (2), *parme* 'Parma violet' (2), *brica* 'brick red' (1), *citron* 'lemon yellow' (1), *kiwi* 'kiwi green' (1), *prune* 'plum purple' (1), *tèrra* 'earth' (1)

7. In our view, a *conventionalized* color term is one whose *primary meaning*, as listed by dictionaries, is *not* color, but for which dictionaries (i.e., not just *one* dictionary) also provides a *color meaning*, for instance French *marron* '(a) horse chestnut; (b) brown' or Occitan *ròse* '(a) rose; (b) pink'.

8. Though this item is structurally complex, it has lexicalized and we believe it should be considered a simple color term, on a par with other conventionalized analogies.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2 below shows the number of types and the proportion of tokens of the simple terms. We may note that French and Occitan show both similarities in terms of general tendencies and differences regarding the diversity and the frequency of use of the three types of simple terms.

Table 2. Primary terms vs. analogy in the French and Occitan EoSS data.

Simple terms	**French (*N* = 1178)**		**Occitan (*N* = 744)**	
	Type	**Token**	**Type**	**Token**
Color (primary) terms	9 Adj	68.3% (804)	8 Adj/N	55.8% (415)
Conventionalized analogy	11 Adj & 1 N	29.4% (346)	15 Adj/N	41.4% (308)
Non-conventionalized analogy	7 Adj	2.4% (28)	11 Adj	2.8% (21)

First, among the different types, primary color terms are the most frequent, slightly more so in French (68.3%) than in Occitan (55.8%); this constitutes a statistically significant difference ($\chi^2 = 31.9$, $p < 0.01$). Note also that the four most frequent color terms in Occitan – *verd* 'green' (149), *violet* 'purple' (106), *blau* 'blue' (102) and *ròse* 'pink' (91) – account for 60.2% of all simple descriptions; interestingly, the second and fourth most frequent terms, *violet* 'purple' and *ròse* 'pink', are not primary color terms but conventionalized analogies.

The data displayed in Table 2 reveal a tendency toward the use of conventionalized analogy in Occitan: this strategy is more frequent than in French, both in terms of types and of tokens, since it accounts for 41.4% of all occurrences with a simple term in Occitan, against 29.4% in French – again, a statistically significant difference ($\chi^2 = 29.4$, $p < .01$). For non-conventionalized analogy, there is no marked difference between the two languages in the proportion of use, which is comparable (and not statistically significant, with $\chi^2 = 0.4$). However, there are slightly more terms in Occitan, since Occitan speakers draw on ten different color terms, while French speakers used seven different color terms only.

Note also that a further difference between French and Occitan speakers in our data lies in the fact that Occitan speakers, unlike French participants, sometimes used nominal constructions, i.e., a color term with a (definite or indefinite) determiner, to describe colors – see e.g., *un verd marronàs* 'a brownish green' in example (22) below. However, this remains a marginal phenomenon, with 57 occurrences in total: nominal constructions are found for five primary color terms, and seven conventionalized analogies.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.2 Derived terms

The most remarkable difference between the French and Occitan EoSS data lies in the use of morphological (derivational) strategies. Table 3 below shows the proportion of use of derived terms, including their use alone and in various other constructions found in the data. As we can see, the use of derived terms is almost non-existent in French (two derived terms were found only in complex expressions) contrary to Occitan where 76 derived terms have been found; such morphologically derived terms can be used alone, but are mainly found in compounds, modifying expressions as well as more complex descriptions.

Table 3. Derived terms in the French and Occitan EoSS data.

Derived terms	French (*N* = 2)	Occitan (*N* = 76)
In modifying expressions	–	32.3% (17)
In complex descriptions	100% (2)	23.1% (59)

The difference between the two languages regarding the use of derived terms is particularly noticeable at the type level. In French, we found only one derivational morpheme carrying an attenuative meaning: *-(t)é* in *jaune orangé* 'orangish (lit. *oranged*) yellow' and *vert bleuté* 'bluish (lit. *blued*) green'. In contrast, in Occitan, we found a rather complex derivational paradigm, for both basic and analogical color terms: they display a series of suffixes, as shown in Table 4 below. Besides, these suffixes are not specific to one color term: some, such as *-et* (diminutive suffix) and *-às* (evaluative suffix), are quite productive in terms of combinability with basic and analogical color terms; others, such as *-astre*, though they are less productive, can still be used with more than one term. And while some may be primarily associated with some color terms in our data, for instance *-enc*, in *rosenc* 'pinkish', they can be associated with other colors, as shown by dictionaries and corpora, for instance *rossenc* 'reddish', *verdenc* 'greenish', etc.[9]

9. These suffixes are not limited to color terms, but also have diminutive uses (*-et*), approximative uses (*-às*) as well as other uses (for instance, *tolosenc* 'from Toulouse') (*Dictionari General de la Lenga Occitana*).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. The diversity and use of derivational suffixes in the EoSS Occitan data.

Terms	Root	*-et*	*-às*	*-èl*	*-enc*	*-ós*	*-it/-at*	*-astre*	*-ada*	*-òt*
Primary terms	*jaune* 'yellow'	*jaunet*	*jaunàs*	*jaulinèl*	*jaunenc*					
	blanc 'white'	*blanquet*	*blancàs*	*blanquinèl*						
	verd 'green'		*verdàs*		*verdenc*	*verdós*				
	roge 'red'					*rojós*				
	blau 'blue'	*blauset*						*blauastre*		
	gris 'grey'		*gridàs*							
Analogical color terms	*irange* 'orange'						*iranjat*			
	malva 'mauve'	*malvet*								
	marron '(horse chestnut) brown'	*marronet*	*marronàs*					*marronastre*		
	ròse 'pink'	*roset*	*rosàs*		*rosenc*		*rosit*			
	ros 'red (of hair or fur)'	*rosset*		*rossèl*						
	vin 'wine'		*vinassa*						*vinada*	
Other	*clar* 'clear, light'	*claret*								*claròt*
	mocin 'a bit'			*mocinèl*						

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.3 Compounds

Both French and Occitan speakers make use of compounds to name colors and, as we have seen earlier (see Table 1), the proportion of compound terms used shows no significant difference: it makes up 7.9% of all the linguistic strategies found in French and Occitan.

Table 5 lists the types of compounds found in the French and Occitan EoSS data and indicates their proportion. Three main types of compounds were identified, in which a color adjective is combined with either another color adjective ([ADJc ADJc]) or a noun, and with or without a preposition ([ADJc (Prep) N]). Together, these three subtypes make up the vast majority of all compounds found in our data, except for further modified compounds, which are quite frequent in the Occitan data and were coded as 'complex' descriptions (see Table 5). The [ADJc N] construction alone accounts for the majority of compounds in French (51.9%, against 32.6% in Occitan), while the [ADJc ADJc] construction is slightly less frequent (43.8% in French and 30.3% in Occitan). As for the [ADJc Prep N] construction, its use is significantly more frequent in Occitan (22.7%) than in French (4.4%). The data, especially in Occitan, also contain a few other subtypes, for instance the [Nc Prep N] construction, in which the first term is a color noun. However, these other subtypes display very low frequencies.

Table 5. Types of compounds found in French and Occitan EoSS data.

Compound types	Example	French ($N = 133$)	Occitan ($N = 132$)
[ADJc N]	Fr *vert pomme* 'apple green' Oc *roge brica* 'brick red'	51.9% (70)	32.6% (43)
[ADJc ADJc]	Fr *rose orange* 'orange pink' Oc *verd blau* 'green blue'	43.8% (57)	30.3% (40)
[ADJc Prep N]	Fr *jaune d'or* 'yellow-of-gold' Oc *verd d'aiga* 'green-of-water'	4.4% (6)	22.7% (30)
[Nc Prep N]	Oc *un verd d'aiga* 'a water green (lit. a green of water)'	–	6.8% (9)
[Nc ADJc]	Oc *es un blancròse* '(it) is a pink white'	–	3.0% (4)
[Nc N]	Oc *lo blau rei* 'the king blue'	–	4.5% (6)

In the most frequent pattern, [ADJc N], four color terms in French and six in Occitan are combined with a series of nouns, generally specific to a given color, such as *sky* for blue, *water* for green, etc. In this compound type (as in others), *blue* and *green* are the most common and diversified in terms of associations with different types of nouns, in both languages, as shown in (19) and (20).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(19) [ADJc N] compounds in French
   a. *bleu* 'blue': *bleu ciel* 'sky blue' (15), *bleu marine* 'marine blue' (6), *bleu nuit* 'night blue' (4), *bleu cyan* 'cyan blue' (2) *bleu eau* 'water blue' (1);
   b. *vert* 'green': *vert pomme* 'apple green' (13), *vert anis* 'anise green' (3), *vert bouteille* 'bottle green' (3), *vert caca d'oie* 'goose-poo green' (1), *vert eau* 'water green' (3), *vert émeraude* 'emerald green' (1), *vert prairie* 'meadow green' (1), *vert sapin* 'fir tree green' (1);
   c. *rose* 'pink': *rose saumon* 'salmon pink' (7), *rose fuchsia* (5), *rose barbie* 'Barbie doll pink' (2), *rose bonbon* 'candy pink' (1);
   d. *jaune* 'yellow': *jaune moutarde* 'mustard yellow' (1)

(20) [ADJc N] compounds in Occitan
   a. *blau* 'blue': *blau cèl* 'sky blue' (10), *blau petròle* 'petroleum blue' (5), *blau marina* 'marine blue' (4), *blau pastèl* 'pastel blue' (2), *blau gendarmariá* 'police blue' (1), *blau sala de bains* 'bathroom blue' (1), *blau carreta* 'car blue' (1)
   b. *verd* 'green': *verd sapin* 'fir tree green' (3), *verd pastèl* 'pastel green' (2), *verd anís* 'anise green' (1), *verd oliva* 'olive green' (1), *verd pistache* 'pistachio green' (1), *verd tilhul* 'lime tree green' (1), *verd veronesa* 'Veronese green' (1)[10]
   c. *roge* 'red': *roge sang* 'blood red' (3), *roge brica* 'brick red' (1)
   d. *ròse* 'pink': *ròse pastèl* 'pastel pink' (2)
   e. *jaune* 'yellow': *jaune palha* 'straw yellow' (1), *jaune solelh* 'sun yellow' (1)
   f. *irange* 'orange': *irange saumon* 'salmon orange' (1)

The [ADJc ADJc] pattern comprises a greater diversity of terms, as shown by Tables 6 and 7 below. Both slots display a high degree of variability, with a large paradigm of possibilities in both languages – from 8 to 15 different terms. The first slot is slightly less subject to variability, with 10 terms in French and 8 in Occitan, while we found, in the second slot, 15 terms in French and 14 in Occitan. In both languages, the most frequent color terms used in compounds are *green*, *blue* and *khaki*; *green* and *blue* can occur in both the first and second slot of the compound, while *khaki* typically occurs in the second slot. While the number of occurrences is similar in terms of *types*, it differs in terms of tokens, since – unlike Occitan – the French data display several cases of seemingly well-established compounds, for instance *vert kaki* 'khaki green' (12 occ.), *bleu vert* 'green blue' (7 occ.), and *bleu turquoise* 'turquoise blue' (5 occ.). In total, in the French data, more than ten compounds appear at least twice. In Occitan, on the other hand, only three constructions are found four times: *verd blau* 'blue green', *verd jaune* 'yellow green',

10. In this case, *verd veronesa*, the nature of the second term is unclear. We coded it as a noun, considering that it stands for the painter Veronese, but it could be an adjective. The case of *basc* 'basque' in *verd basc* and *roge basc* has been coded as a modifying adjective.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

and *blau turqués* 'turquoise blue', while *ròse violet* 'purple pink' (3 occ.) and *blau verd* 'green blue' (2 occ.) are the only remaining compounds which appear more than once in our dataset.

Table 6. Types and tokens of [Adjc Adjc] compounds in the French EoSS data.

TERM2 / TERM1	kaki	vert	bleu	orange	turquoise	rouge	violet	beige	blanc	marron	rose	gris	jaune
vert	**12**		**4**							1			1
bleu		7			**5**				1				
rose				3		2	2	2					
jaune		3		1									
rouge				2						1			
marron		1				1	1						
gris			2										
violet									1		1		
blanc												1	
orange											1		

Table 7. Types and tokens of [Adjc Adjc] compounds in the Occitan EoSS data.

TERM2 / TERM1	verd	blau	caquí	violet	jaune	turqués	ròse	fuchsia	azur	burèl	irange	malva	marron	roge	vermilhon
verd		4	4		4					1					
blau	2	1				4			1						
rose				4				2						1	
jaune	1		1				1				1		1		
violet				1								1			
brun	1														
gris	1														
roge							1								
irange															1
kiwi	1														

In Occitan, there are a number of marginal cases in which the second term of the compound is either suffixed or further modified. We did not count these occurrences as legitimate compounds, considering, on account of the modification, that they were closer to a syntactic strategy. This includes descriptions which are relatively simple and straightforward, as in (21a), as well as more complex descriptions, as in (22). Though such descriptions are not altogether absent from the French data, as shown by (21b), they are quite rare.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(21) a. Oc *blanc rosenc* 'pinkish white', *verd jaunenc* 'yellowish green',
b. Fr *jaune orangé* 'orangish (lit. *oranged*) yellow', *vert bleuté* 'bluish (lit. *blued*) green'

(22) Oc *aquò es un verd marronàs es un marron verdàs tanben que tira a cacà d'auca* 'this is a brownish green (it) is a greenish brown too which seems like (lit. pulls to) goose poo'

3.4 Modifying expressions

Another type of construction which is quite frequent in our data, both in French and Occitan, is what we call *modifying expressions*, in which a color adjective or a color noun is modified, generally by another adjective, an adverb, a suffix or a combination thereof.

Table 8 presents the types of modifying strategies found in the French and Occitan data together with their frequencies. The most common modifying expression used by the speakers in both languages consists of a color adjective and a modifying adjective ([ADJc + ADJ]), which typically specifies the intensity of a given color, either light or dark. In French, this construction type accounts for 89.5% of all the modifying expressions and in Occitan for 66.6%. Such modifying expressions can be further reinforced by the intensifying adverb *très* 'very' (sometimes reduplicated) as in the construction [ADJc (+ADV) + ADV + ADJ] which corresponds to 9.9% in French and 8.2% in Occitan.

Table 8. Types of modifying expressions in French and Occitan EoSS data.

	Example	**French** (*N* = 333)	**Occitan** (*N* = 336)
ADJc+ADJ	Fr *bleu clair* 'light blue' Oc *blau clar* 'light blue'	89.5% (298)	66.6% (235)
Nc + ADJ	Oc *un gris clar* 'a light grey'	–	13.3% (47)
ADJc+ADV(+ADV) + ADJ	Fr *vert très foncé* 'very dark green', *bleu très, très clair* 'very very light blue' Oc *verd pro fonçat* 'very dark green'	9.9% (33)	8.2% (29)
Nc + ADV(+ADV) + ADJ	Oc *un blau pro clar* 'a very light blue', *un ròse plan plan clar* 'a very very light pink'	–	7.1% (25)
ADJc-suff	Oc *verdàs* 'greenish'	–	4.0% (14)
Nc-suff	Oc *un jaunàs* 'a dirty yellow (lit. *a yellowish*)', *blancàs* 'withish'	–	0.6% (2)
N-suff	Oc *vinada* 'wine-colored (lit. *wine-ade*)'	–	0.3% (1)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Other types of modification, found only in Occitan, include (a) color nouns which can be modified by an adjective reinforced by one or two intensifying adverbs, (b) color adjectives or nouns combined with a modifying suffix, and (c) a noun combined with a modifying suffix. Although these modifying expressions are not very frequent in the EoSS data, they set Occitan clearly apart from French.

3.5 Complex descriptions

So far we have observed that there are clear differences between the two datasets, mainly the importance of simple descriptions in French, and that of derived terms in Occitan. However, the most remarkable difference lies elsewhere, namely in the frequency of use of complex descriptions in the Occitan data.

In the French data, complex descriptions are very rare: we found 36 occurrences of complex constructions, representing only 2.1% of the data. These include mainly verbal forms such as *tirant sur* 'drawing on' as in (23), and cases in which participants rephrase their initial description, with or without negation, as in (24) and (25).

(23) FR *blanc tirant sur le gris clair*
'white drawing on light grey'

(24) FR *noir? Non c'est plus marron*
'black? No, that's rather brown'

(25) FR *rouge, bordeaux*
'red, bordeaux'

In Occitan, we found 441 complex descriptions, which account for 26.3% of all descriptions, and include a variety of strategies. The main feature of complex descriptions, as in (26), is that they very often involve some kind of reformulation (325 occurrences, or 73.7% of complex descriptions), frequently with various hedging strategies, i.e., strategies by which the speaker introduces some uncertainty into his utterance, as in (27) and (28): hedging appears in 212 occurrences, almost half (48.1%) of complex descriptions. Another feature is that complex descriptions often include the strategies we described in the previous sections, as well as others, including verbal modification (29) and (30).

(26) OC *blanc blauset, o gris blauset, oè, blanc blauset*
'bluish white, or bluish grey, yeah, bluish white'

(27) OC *un verd me sembla*
'a green, seems to me'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(28) Oc *ròse puslèu un ròse*
'pink rather a pink'

(29) Oc *un jaune que peta*
'flashy yellow (lit. a yellow that explodes)'

(30) Oc *blancàs aquò mas tira sul ròse un mocin*
'whitish this but draws on (the) pink a bit'

An interesting subtype of reformulation is that in which the participant reformulates without changing the color name, as in (31). But the most remarkable feature of these complex constructions in Occitan is perhaps the fact that they generally combine strategies: for instance, they often display a combination of compounding and modification (32), or compounding and derivation (33). We even found highly complex descriptions which seem to offer a *pot-pourri* of such strategies, as in (34). The importance of such combinations is illustrated in Table 9, which clearly shows that – except for modification – the combination of strategies is the most common scenario.

Table 9. Main strategies in complex color descriptions in Occitan.

Strategy		Occurrences
Modification		27.9% (126)
Compounding	(alone or with hedging and/or reformulation)	15.1% (68)
Verbal constructs		12.0% (54)
Derivation		1.8% (8)
Combinations of the above strategies		23.3% (105)
Hedging and/or reformulation only		20.0% (90)
Total		451

(31) Oc *negre, pas negre negre, gaireben negre*
'black, not black black, almost black'

(32) Oc *blau verd fonçat*
'dark green blue'

(33) Oc *jaune verdenc*
'greenish yellow'

(34) Oc *blanc amb un pauc de quicòm d'una autra color, que tira sul violet un pauc*
'white with a bit of something of another color, which draws on (the) violet a bit'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4. Naming strategies and variation among participants

A striking result of our study is that there is a wide variation across speakers, in both datasets. In French, seven participants almost exclusively use simple descriptions, while the thirteen others can be placed along a continuum from 20% to 80% of simple descriptions; these participants use a mean of roughly 10% of compounds and 40% of modifications (see Figure 2).

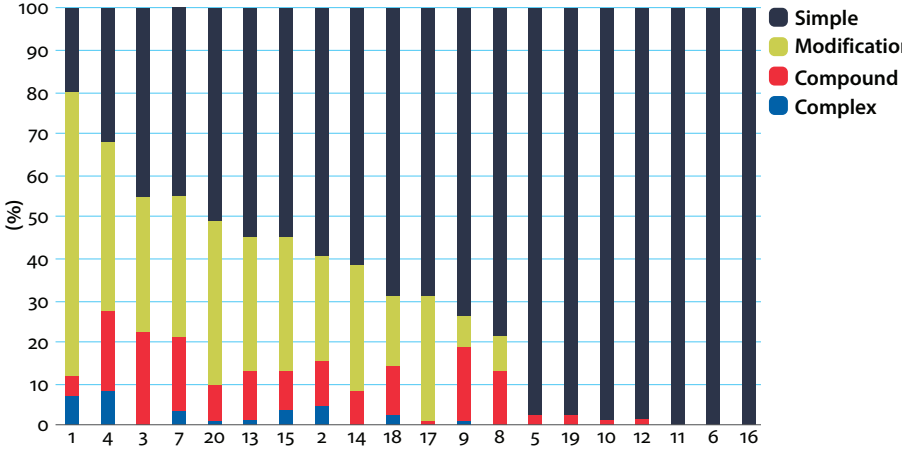


Figure 2. Variation in the use of naming strategies across participants, in French.

The variation across speakers seems much more important for the Occitan dataset. Two facts may account for this variation. The first one is that the sociological span is much wider for Occitan participants than for French participants: while French participants were all students at the same faculty, Occitan participants offer a wide spectrum of age, activity and schooling. The second one is that Occitan is not a standardized language, and is a minority language: thus, participants (a) could not refer to a common, standard language system and (b) were probably influenced to different degrees by French. This is obvious in some descriptions, in which there is clearly some code-shifting, as in (35) and (36). Though there are very few cases of this kind, they are a good indication of the existing asymmetry between French and Occitan.

(35) *ròse, mas fuchsia, sai pas se se dis en occitan, donc ròse, ròse viu, ròse viu*
'pink, but [*switches to French*] fuchsia, [*switches back to Occitan*], I don't know whether one can say this in Occitan, so pink, flashy pink, flashy pink'

(36) *caca d'oie, pas kaki, caca d'oie… d'auca*
'[French] *greenish-yellow* (lit. *poo-of-goose*), [Occitan or French] not khaki, [French] greenish-yellow… [Occitan] of goose'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3 below illustrates the impact of this divergence on participants' responses in Occitan: while some participants display patterns which are quite akin to what we found in French – namely participants 16, 9, 19, 18 and 12 – others deviate from this pattern to various degrees. Some participants provided complex color descriptions more than half the time.

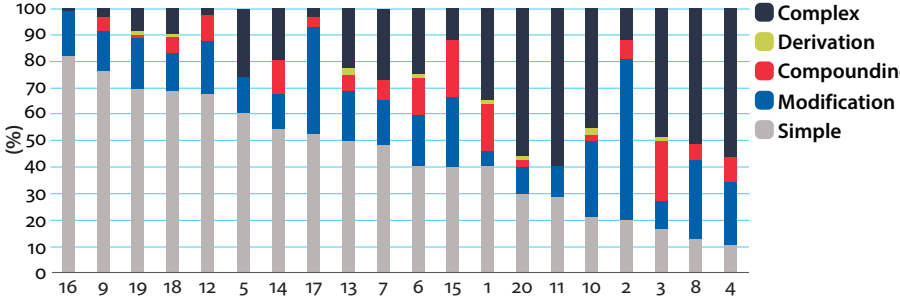


Figure 3. Variation in the use of naming strategies across participants, in Occitan.

For instance, concerning the main structural difference between French and Occitan, namely the use of derivation, there is a wide divergence among Occitan participants, with eight participants who never used suffixes, nine participants who used suffixes from 1 to 5 times, and three participants who used them more than 10 times.

The three participants with the highest frequency of simple constructions considered French as their first language. Besides, the true outlier, participant 16, is much younger than all other participants, and one of the least fluent in Occitan, having French as his first language and speaking Spanish at home. Conversely, the three participants with the lowest frequency of simple constructions mentioned Occitan as their first language; this is also true of the three participants (not the same ones, however) with the highest frequency of suffixation (though one of the participants actually gave French *and* Occitan as his first languages).

Table 10. Number of suffixed words per participant, in the Occitan data.

Number of participants	Number of suffixed words
1	18
1	15
1	11
1	5
2	4
1	3
3	2
2	1
8	0
Total	**68**

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

5. Discussion and concluding remarks

We expected our results to illustrate the structural similarity between French and Occitan, as well as differences, with a presumable preference for analytic constructions in French and synthetic constructions in Occitan. In other words, we predicted that participants speaking French would tend to favor analytic (multi-word) descriptions for naming colors, and participants speaking Occitan would favor synthetic descriptions with color term and derivational suffix.

Concerning the use of various color-naming strategies, there is indeed a large overlap between the two languages, at least as far as we can tell on the basis of the EoSS data investigated in this study. Both groups of participants made use of simple color naming, compounding, modification, derivation and more complex strategies; yet, there is a difference between the two languages in the frequency of each strategy.

However, as far as the importance of analytic constructions is concerned, our data did not quite yield the results we expected – i.e. a clear contrast between 'Analytic French' and 'Synthetic Occitan'. As we have shown in Sections 2 and 3, the contrast is actually much more complex. The strategies used by participants in both languages are mostly similar, but structurally complex strategies seem to be frequent in Occitan, not French. Besides, derivation, which is a synthetic strategy, is found in Occitan more than in French, as expected, but not very frequent even in Occitan: even though we expected to find *some* derivation in French and much more in Occitan, what we found was rather *some* derivation in Occitan, and *none* in French.

On the other hand, we observed a stark contrast between the prolific descriptions in the Occitan dataset, with participants combining various strategies and describing colors with a mean of 3.13 words, and the very straightforward descriptions in French, with a mean of 1.36 words. However, this contrast might not be linked to differences in language structure, but to socio-linguistic features of French and Occitan. This includes the fact that, for French, participants were students used to experimental set-ups, thus prone to provide simple, straightforward answers. Conversely, for Occitan, participants were not used to such experimental protocols, and also possibly delighted to have an opportunity to use a language which they all considered as beautiful, giving it a mean of 4.9 out of 5, while French score a measly 4.2. This shows how difficult it can be to compare a standardized language such as French with a non-standardized, endangered language such as Occitan.

Further research shall compare the EoSS data with corpus-based data to investigate the actual language use of color terms and the evolution of the different strategies used by participants of the two languages. Our first attempts in this direction seem to show that there is indeed a clear evolution, with a much higher frequency of suffixation in older stages of the language, at least for French.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Acknowledgment

The study that we report in this chapter was conducted within the framework of the cross-linguistic project *Evolution of Semantic Systems*, carried out at the Max Planck Institute for Psycholinguistics, the Netherlands (Majid, Jordan, & Dunn). We wish to thank the editors and anonymous reviewers for their comments.

Abbreviations

FR:	French
INF:	infinitive
IT:	Italian
OC:	Occitan
PRS:	present
PT:	Portuguese
RM:	Romanian
SG:	singular
SP:	Spanish

References

Berlin, B., & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.

Carlier, A. (2007). From Preposition to Article. The grammaticalization of the French partitive. *Studies in Language*, *31*(1), 1–49. https://doi.org/10.1075/sl.31.1.02car

Carlier, A., & Lamiroy, B. (2014). The Romance partitive. In Silvia Luraghi & Tuomas Huumo (eds.), *Partitive Cases and Related Categories* (pp. 477–518). Berlin: Mouton De Gruyter.

Carlier, A., & Lamiroy, B. (2018). The emergence of the grammatical paradigm of nominal determiners in French and in Romance: Comparative and diachronic perspectives. *Canadian Journal of Linguistics*, *63*(2), 141–166. https://doi.org/10.1017/cnj.2017.43

De Mulder, W. & Lamiroy, B. (2012). Gradualness of grammaticalization in Romance. The position of French, Spanish and Italian. In K. Davidse, T. Breban, L. Brems, & T. Mortelmans (Eds.), *Grammaticalization and Language Change: New reflections* (pp. 199–226). Amsterdam / Philadelphia: John Benjamins. https://doi.org/10.1075/slcs.130.08mul

Doebler, W. M. (2010). *Ciascun saria di color vinto: The Philology and Philosophy of Terms for Color and Light in Occitan and Italian Poetry from the Troubadours through Petrarch*. (Ph.D. dissertation). University of California, Los Angeles.

Dubois, D., & Grinevald, C. (1999). Pratiques de la couleur et dénominations. *Faits de langues* 14, 11–25. https://doi.org/10.3406/flang.1999.1261

Dubois, D., & Grinevald, C. (2003). En voir de toutes les couleurs : processus de dénomination des couleurs et constructions cognitives. In C. Vandeloise (Ed.), *Langues et* cognition (pp. 79–113). Paris: Lavoisier.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dworkin, S. N. (2016a). A diachronic overview of color terms in the Romance languages: the lexical stability of the Latin color vocabulary. In J. P. Silvestre, E. Cardeira, A. Villalva (Eds.), *Colour and colour naming: crosslinguistic approaches* (pp. 9–20). Centro de Linguística da Universidade de Lisboa, Universidade de Aveiro.

Dworkin, S. N. (2016b). La estabilidad léxica: la pervivencia de los nombres de los colores del latín en español y en las otras lenguas romances. In A. L. Serena, A. Narbona & S. del Rey Quesada (Eds.), *El español a través del tiempo: Estudios ofrecidos a Rafael Cano Aguilar* (pp. 245–257). Seville: Ediciones de la Universidad de Sevilla.

Fagard, B., & Mardale, A. (2012). The pace of grammaticalization and the evolution of prepositional systems: Data from Romance. *Folia Linguistica*, *46*(2), 303–340. https://doi.org/10.1515/flin.2012.011

Fagard, B. (2010). *É vida, olha*…: Imperatives as discourse markers and grammaticalization paths in Romance – A diachronic corpus study. *Languages in Contrast*, *10*(2), 245–267. https://doi.org/10.1075/lic.10.2.07fag

Fagard, B. (2011). Conjonctions et grammaticalisation : le cas des langues romanes. *Mémoires de la Société de Linguistique de Paris*, 19, 79–102.

Grossmann, M. (1988). *Colori e lessico. Studi sulla struttura semantica degli aggettivi di coloure in catalano, castigliano, italiano, romeno, latino ed ungherese*. Tübingen: Gunter Narr.

Heider, E. R., & Olivier, D. C. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology*, *3*(2), 337–354. https://doi.org/10.1016/0010-0285(72)90011-4

Hopper, P., & Traugott, E. (2003 [1993]). *Grammaticalization*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165525

Kay, P., Berlin, B., & Marrifiend, W. (1991). Biocultural implications of systems of colour naming. *Journal of Linguistic Anthropology*, 1, 12–25. https://doi.org/10.1525/jlin.1991.1.1.12

Kay, P., & McDaniel, C. K. (1978). The Linguistic Significance of the Meanings of Basic Color Terms, *Language*, 54, 610–646. https://doi.org/10.1353/lan.1978.0035

Kleiber, G. (2007). Adjectifs de couleur et gradation : une énigme… « très » colourée. *Travaux de linguistique*, *55*(2), 9–44. https://doi.org/10.3917/tl.055.0009

Kristol, A. M. (1978). *Color. Les langues romanes devant le phénomène de la couleur*. Zurich: Francke Berne.

König, E., & Gast, V. (2007). *Understanding English-German contrasts*. Berlin: Schmidt.

Lamiroy, B. (1999). Auxiliaires, langues romanes et grammaticalisation. *Langages*, 135, 33–45. https://doi.org/10.3406/lgge.1999.2201

Lamiroy, B. (2001). Le syntagme prépositionnel en français et en espagnol: une question de grammaticalisation? *Langages*, 143, 91–106. https://doi.org/10.3406/lgge.2001.892

Lamiroy, B. (2011). Degrés de grammaticalisation à travers les langues de même famille. *Bulletin de la Societe de Linguistique de Paris*, 19, 167–192.

Lamiroy, B., & Pineda, A. (2017). The pace of grammaticalization in Romance. The position of Catalan. *Linguisticae Investigationes*, *40*(2), 304–331. https://doi.org/10.1075/li.00007.lam

Majid, A., Jordan, F., & Dunn, M. (2015). Semantic systems in closely related languages. *Language Sciences*, 49, 1–18. https://doi.org/10.1016/j.langsci.2014.11.002

Marchello-Nizia, C. (2006). *Grammaticalisation et changement linguistique*. Brussels: De Boeck-Duculot.

Molinier, C. (2001). Les adjectifs de couleur en français. Eléments pour une classification. *Revue romane*, 36(2), 193–206.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Molinier, C. (2005). Sur la forme et le statut des adjectifs de couleur suffixés. *Le français moderne*, *LXXXIII*(2), 145–156.

Molinier, C. (2006). Les termes de couleurs en français. Essai de classification sémantico-syntaxique. *Cahiers de Grammaire*, 50, 259–275.

Mora, L, & Lécrivain, C. (2004). Les adjectives suffixés colorés : entre la péjoration restrictive et la péjoration exhaustive. *Estudios de Lengua y Literatura Frascesas*, 15, 23–42.

Noailly, M. (2005). L'intensité de la couleur. Remarques sur l'emploi de *très* devant bleu, rouge, vert, jaune. In I. Choi-Jonin, M. Bras, A. Dagnac, & M. Rouquier (Eds.), *Questions de classification en linguistique : méthodes et descriptions* (pp. 263–273). Berne: Peter Lang.

Ott, A. G.. (1899). *Étude sur les couleurs en vieux français*. Paris: É. Bouillon.

Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, *50*(4), 378–411. https://doi.org/10.1016/j.cogpsych.2004.10.001

Roberson, D., Davies, I. R., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a Stone-age culture. *Journal of Experimental Psychology: General*, *129*(3), 369–398. https://doi.org/10.1037/0096-3445.129.3.369

Schwegler, Armin. 1990. *Analyticity and Syntheticity: A Diachronic Perspective with Special Reference to Romance Languages*. Berlin, New York: Mouton de Gruyter.

Tesnière, L. (1928 [1918]). *Statistique des langues de l'Europe*. Paris: Payot.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Color terms in Basque

Lexicalization and categorization

Iraide Ibarretxe-Antuñano
Universidad de Zaragoza

This chapter explores the domain of color terms in Basque from a descriptive, as well as a categorization perspective. The first part provides an overview of the color system in this language: its color lexicalization processes and the meanings these color labels cover. The second part focuses on the results of a categorization study of color in Basque. The data were retrieved from 20 native speakers of Basque and elicited using the color elicitation task as part of the EoSS project at the Max Planck Institute for Psycholinguistics (Majid et al., 2015). The chapter concludes with a discussion on the similarities and differences found in the descriptive and the categorization part.

Keywords: color, Basque, lexicalization, categorization

So ist es klar, daß die Farbenharmonie nur auf dem Prinzip
der zweckmäßigen Berührung der menschlichen Seele ruhen muß.
Diese Basis soll als Prinzip der inneren Notwendigkeit bezeichnet werden.
W. Kandinsky (*Über das Geistige in der Kunst*, Munich 1911)

1. The Basque color system

1.1 Ancient and not so ancient color labels

The answer to the question "how many color terms are there in the Basque language today?", is not a simple one. Not only because Basque is a language with several distinctive dialects (Zuazo, 2013) and a rich morphology that creates new color terms ad hoc (e.g., composition, derivation, see Section 1.2), but also because the list of color labels has been increasing throughout centuries due to the language's own resources, as well as different loans from Latin and surrounding Romance languages at different stages (see, Gorrochategi et al., 2018; Lakarra, 1995, 2002;

https://doi.org/10.1075/sfsl.78.11iba
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

Martínez Areta, 2013; Michelena, 1990). As a result, old and brand new loanwords of color can be distinguished, and thus, the most appropriate answer to the above question would be "it depends".

According to the literature (Azkue, 1923–25 [1969]; Larzabal, 1975; Ortiz-Osés & Garagalza, 2006; Perurena, 1992; Preciado Sáez de Ocáriz, 1992, 2009; Txillardegi, 1975), the color system in Basque was originally organized around three main colors[1]:

- ***Zuri/txuri*** encoded the color 'white' but also meant 'without color' and 'cleanliness' (cf. the verb *zuritu* 'to peel'). It also referred to the color of 'light', which metaphorically did not bring positive connotations. On the contrary, it was related to laziness (e.g., *zuri-zuri bizi* (white-white live) 'to live without working'), fakery, lying, and untrustworthiness (e.g., *zuri hitz egin* (white word make) 'to flatter'). It was also considered the color of the 'spirit, between earth and sun'. Etymologically, *zuri* may have come from *zur* 'wood' plus the participial ending *-i* and meant '?the color of wood, peeled' (Lakarra, 2002, p. 434).
- ***Beltz***/*blatz*/*beltx* encoded both 'dark' and 'black', as well as 'dirty' and 'sad'. It was considered the color of 'earth'. Etymologically, *beltz* may have come from **bel* 'black' plus the adjectival ending *-tz* and meant 'blackish' (Lakarra, 2002, p. 432).
- ***Gorri*** generally meant 'with color' (cf. the verb *gorritu* 'to blush') and included colors such as 'red', 'rose', 'pink', 'yellow', 'blond', 'orange', and 'brown'. It was also the color of flesh, and therefore, it also meant 'bare, naked' in the sense of 'without protection', which in turn came to have negative connotations as in 'terrible, fierce, ghastly, extreme' (e.g., *gorriak ikusi* [red.ABS.PL see] 'to go through hell'), and 'inexperienced' as in *umegorri* (child.red) 'green hornet'. In opposition to *zuri*, *gorri* also meant 'energetic worker'. It was considered the color of 'life in action, the sun'. Etymologically, *gorri* may have come from *gor* 'deaf, mute' plus the participial ending *-i* and meant 'muted color' (Martínez Areta, 2003, p. 70).

1. For each color label, I include diatopic alternatives (in bold main lemma in dictionaries), a full description of the color meanings, as well as their explanations and metaphorical meanings together with possible etymologies, if available. This information may be seen as irrelevant for the discussion in the second part. However, my intention is to provide a brief overview of these colors, since most of the literature available is not accessible to a general audience, either because it is written in Basque or because it is published locally. The information is drawn from the relevant literature but also from the *Orotariko Euskal Hiztegia* (OEH, Mitxelena, 1985–2005). *The General Dictionary of the Basque Language* is a sixteen-volume dictionary that compiles all the information available for each entry including not only diachronic and synchronic meanings, but also attributed quotations. The OEH can be consulted online at [www.euskaltzaindia.eus/oeh]. The OEH is also used to date the first written appearance of color labels.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Besides these three main colors, there were:

- ***Urdin***, a GRUE category color, covered colors such as 'grey' (cf., *ileurdin* 'grey hair'), 'blue', 'old, rotten color', 'color of water', and 'bluish green, moldy' (e.g., *orgian urdina* [bread.loc blue.abs] 'moldy bread'). Etymologically, *urdin* may have come from *ur* 'water', the suffix *-di* 'possessor' plus the ending *-n*. Its meaning could have been 'dirty water' (Michelena, 1990, p. 52) or '?what has turned into water > the color of water' (Martínez Areta, 2008, p. 9). This color has two other alternatives, the neologism *oztin* (1897) 'blue' and the Roncalese dialectal word *dundu* 'blue', 'dark'.
- ***Hori****/hoi/gor(i)/kori* '(canary) yellow'. Etymologically, *hori* may have come from *hor* 'hound' plus the participial ending *-i* and meant '?the color of a hound' (Lakarra, 2002, p. 434). This color has another two subtypes. *Beilegi/be(ll)(e)(i)gi* 'lively yellow' and *Laru* 'pale yellow' (also 'faded color') which is a very early Latin-Romance loan (Latin CLARUS 'clear, bright, gleaming').
- ***Ubel****/ugel* covers colors such as 'purple', 'violetish', 'dunish', 'dark grey', 'livid, pale', 'pale near blue'. It is the word used for a 'bruise' and also refers to any kind of a 'defect, fault, imperfection'. Etymologically, *ubel* may have come from *ur* 'water' and **bel* 'black' with the meaning 'the colour of black water' (Martínez Areta, 2003, p. 69).
- ***Arre*** means 'muddy-brown', 'dun', 'dark grey', 'close to black, but also close to red and brown', 'earth color', 'dirty'. Etymologically, Van Eys (1873, p. 27) suggests that it may have come from *harri* 'stone', that is, 'the color of stone' (see, Martínez Artea, 2008, p. 10).
- ***Uher****/uger* denotes 'dun', 'mixed', 'brown', 'grey', and 'rotten'. It developed the metaphorical meaning 'evil, malicious'.

Apart from these basic color terms that cover the (a)chromatic spectrum, another two groups of colors should be mentioned.[2] One of them covers the labels for "multicolor/mixed colors":

2. For the sake of completeness, in this brief tour around Basque color terms, it is worth mentioning a special set of labels, mainly used with animals (esp. cattle), that refer to both patterns and color. These are: *Pint(t)o* for animals that have well-defined color patches, 'dappled', *Pikart* for animals that have many colored spots, *Izpil* for animals with different colored spots on one single color, 'dappled', *Bertzin*/*bertxin* for animals (cattle) with different colored strips, *Maskaro*/*ma(x)(z)karo* for animals with a lot of black and white spots (also animals with spots near the eyes), and *Betroi* (> *Breton*) for animals (cattle) with mixed white and black colors but more white than black.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

- ***Nabar/ñabar*** refers to 'a mixed dark and lighter color', 'brown-grey', 'dun', 'motley', and 'multicolored' (cf. *nabar egin* (nabar make) 'to inlay'). It developed metaphorical meanings such as 'laziness', 'a two-faced person', as well as 'daring, adventurous, bold' and 'cheeky'. It also refers to something that is 'prominent', 'that captures attention' (e.g., *nabaritu* 'to realize', 'to be evident').
- ***Ñibirri-ñabar****/nibirri-nabar/nibirri-ñabar*, an ideophonic formation from *nabar*, also means 'multicolored' but in the sense of 'loud, garish, gaudy colors'.
- ***Kikirriki***, an ideophonic word, means 'multicolored' but in the sense of 'clashing, motley colors'.

Another group includes what I call "fading colors":

- ***Hits*** means (Southern dialects) 'pallid', 'dull', 'drab', 'near grey', 'muted', 'shadowy', 'dark', and it is also used to refer to a sad person or a person with no character.
- ***Beluri*** (Gipuzkoan dialect) refers to a 'faded color', 'pallid'.
- ***Margul*** means 'discolored', 'devoid of color'.
- ***Zurbil****/(x)(z)urpil/xurpail/xurbil/zurrupil/zurbel* encodes 'pale', 'pallid', and 'wan'.

Besides these sets of color words, Basque has been assimilating to its color system other labels at different stages in its history: first, from Latin-Romance (e.g., *laru* > Lt. CLARUS), and later, mainly from Spanish and French, the two Romance languages also spoken in the same area. Some of these later loans are relatively recent. For example, according to the OEH, the first written mention of *kaki* 'khaki' dates back to 1937 and the first one for *lila* 'lilac' to 1960. Labels such as *fuksia* 'fuchsia' or *zian* 'cyan' are brand new loans[3] and they are not even registered in the OEH. There are, however, some other labels that, despite being loanwords, are quite

3. Interestingly enough, this tendency to include "new" specific color terms is not restricted to Basque, but very much present in other languages. *Khaki*, a word originally from Persian *khāki* 'dust', is also found in Italian *cachi* or Dutch *kaki* for khaki, or *International Klein Blue* (*IKB*), a new color patented by artist Yves Klein in 1960, is also used in Spanish *azul klein* 'blue klein'. Some of these labels, first restricted to specific contexts, have been introduced in the language by different means, one of them being the garment industry or the world of fashion. An interesting issue that deserves more research in the future is the scope of these "international" color loanwords and the reasons why the semantic field of color is more susceptible to the incorporation of new loanwords than others, such as body parts or kinship terms, just to name two fields usually explored in semantic typology.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

old. Some of them have been registered in the first written texts in Basque according to the OEH (see dates in parenthesis).[4] Some of these old loans are:

- *Berde*/*f(h)erde* 'green' (14th century). This color has two other alternative labels, the neologism *orlegi* and the metonymy-based *musker* 'lizard'. As in the source languages, *berde* also means 'immature', 'childish', 'healthy, glowing', 'obscene, libertine'. It is also the color of anger (cf. *berde-berde* 'very green' but also 'very angry, furious').
- *Doratu*/*dauratu* (1562) 'gold'.
- *Gris*/*grix* (1620) 'grey'.
- *More* (1660) 'purple', 'from blue to purple, including violet'.
- *Brioleta*/*bioleta* (1696) 'violet'.
- *Blu*/*azul* (1672/1800) '(dark) blue'.
- (*L*)*arroxa*/*(l)arrosa* (1853 [as flower, 1580]) 'pink, rose'.
- *Laranja*/*liranja* (20th century [as fruit,1620]) 'orange'.

These color labels are fully integrated into the actual basic color system in Basque (see Section 2). They have taken part in the color spectrum of native color words and, as a result, the meaning of native color words has narrowed. For example, some of the meanings of *gorri* are covered by *laranja* 'orange' and (*l*)*arrosa* 'pink', and *gorri*, nowadays, is mainly used for 'red' (e.g., blood is *gorri*); the color 'grey', previously included in the GRUE color *urdin* is, nowadays, encoded with *grisa*, and only remains in some specific lexical items such as *ileurdin* 'grey hair' or *urdindu* 'to go grey'. An interesting case of how some color labels have been established in the language is the label *marroi* 'brown'. Nowadays, it is the common and standardized option to describe the color brown. In fact, a search for the word *brown*/*marrón*/*marron* in bilingual dictionaries offers *marroi* as their sole translation (the other possibility is to say *gaztain kolore* 'chestnut color' documented in the 19th century). However, it is a recently acquired word. The Statistical Corpus of Twentieth Century Basque, which compiles written texts from 1900 to 1999 ([http://xxmendea.euskaltzaindia.net/Corpus/]), retrieves no examples before the last third of the 20th century.

Table 1 summarizes the whole range of color labels described in this section and the (a)chromatic colors they describe.

4. Due to the scarcity of written texts in Basque before the 20th century, color labels attested before the 20th century are included in the old loanword category.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. One-word color labels in Basque

Basic native colors	*Zuri/txuri* 'white', 'without color', 'cleanliness'
	Gorri 'red, rose, pink, yellow, blond, orange, brown', 'flesh color', 'with color'
	Beltz 'black', 'darkness', 'dirty'
	Hori '(canary) yellow' (*beilegi* 'lively yellow')
	Urdin 'blue', 'grey', 'old, rotten color', 'dun', 'water color', 'bluish green' (*dundu* (R) 'blue'; *oztin* (neologism) 'blue')
	Ubel 'purple' 'violetish', 'dunish', 'greyish', 'livid, pale', 'pale near blue'
	Arre 'muddy-brown', 'dun', 'dark grey', 'close to black, but also close to red and brown', 'earth color', 'dirty'
	Uher 'mixed', 'dun', 'brown', 'grey', 'rotten'
	(*Orlegi* [neologism] 'green')
"Mixed" native colors	*Nabar/ñabar* 'mixed dark and lighter color', 'dun', 'motley', 'multicolored'
	Kikirriki 'multicolored', 'clashing, motley colors'
	Ñibirri-ñabar/nibirri-nabar 'multicolored', 'loud, garish, gaudy colors'
"Fading" native colors	*Hits* 'pallid', 'dull', 'drab', 'near grey', 'muted', 'shadowy'
	Beluri (G) 'faded color'
	Margul 'discolored', 'devoid of color'
	Zurbil 'pale', 'pallid', 'wan'
Old loans	*Laru* 'pale yellow', 'faded color'
	Berde/f(h)erde 'green'
	Doratu/dauratu 'gold'
	Gris/grix 'grey'
	More 'purple', 'from blue to purple including violet'
	Brioleta 'violet'
	Blu/Azul '(dark) blue'
	(L)arroxa/arrosa 'pink, rose'
	Laranja 'orange'
Recent loans	*Kaki* 'khaki', *lila* 'lilac'
	Fuksia 'fuchsia', *zian* 'cyan'
	Marroi 'brown'

1.2 Lexico-morphological devices for color

The labels included in Section 1.1 are mostly single-morph words,[5] but Basque has several morphological devices to create new color labels. One of those devices is compounding. This morphological process is very productive in Basque in general

5. Excluding ideophonic color terms which are reduplicated. Basque is a language with a large ideophone inventory (see Ibarretxe-Antuñano, 2017, for a typological description). Both *kikirriki* and *nibirri-nabar* can be considered ideophones, but their structure is different. In *kikirriki*, the members of the reduplicated form do not have meaning if taken separately (*kiki* and *riki*), whereas in *nibirri-nabar*, the second element *nabar* does have a meaning 'mixed dark and light' which

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(see de Rijk, 2008, chap. 31; Hualde, 2003a, and Wälchli, 2005, for a typological perspective), as well as in the color domain, where color labels can be freely combined.

Compounding is a very well-known morphological process, but equally complex and problematic (Lieber & Stekauer, 2009). The relevant literature discusses a wide range of topics related to compounding: types of compounds, categories used in compounding, syntactic dependencies within compound members, differentiation between compounding and other syntactical processes such as appositions, etc., just to name a few relevant for Basque. All these complexities are multiplied in the case of color terms since, starting from their category (adjectives but with some nominal properties), they are freely combined beyond some of the constraints applicable to other compounds in Basque. A thorough discussion/justification of why some of these color formations are treated as different types of compounds/appositions lies beyond the scope of this chapter. The use of the concept of compounding in this chapter is, in a way, "instrumental". The goal is to provide an overview of the linguistic mechanisms (not their formal basis) available in Basque to codify color and the semantic differences encoded by these mechanisms. This classification will be the basis to explore how these means are exploited by Basque young adult native speakers (Section 2). Therefore, the following classification of color compounds in Basque aims at capturing both the morphosyntactic and the semantic relationship between the two compound elements, as well as the resulting compound, as theoretically neutral as possible. This classification is mainly built on previous accounts of compounding in general (e.g., Lieber & Stekauer, 2009; Scalise et al., 2009; Wälchli, 2005) and Basque compounding in particular (e.g., Azkarate, 1990; de Rijk, 2008; Etxeberria et al., 2015).[6] I will divide these compounds into two main groups: co-compounds and sub-compounds.

---

seems to be intensified by the first element. This latter form could be considered an example of what Wälchli (2005, p. 147) classifies as an "imitative co-compound". Basque has many expressive reduplicative co-compounds (see de Rijk, 2008, pp. 864–866) which may structurally resemble an ideophone without actually being one. In these cases, it is the first element of the reduplicative form that has meaning and the second element is a repetition with an initial consonant alternation (usually *m*-). For example, *zalantza-malantza* (doubt-*malantza*) 'many doubts' or *inguru-minguru* (around-*minguru*) 'meander, outskirts, detour'. It is interesting to note that the order in *nibirri-nabar* differs from these other compounds.

6. The names for each of these compounding types, as well as their characteristics, are drawn and adapted from these references. The interested reader in formal discussions on Basque morphological compounding and syntactical apposition (especially from the generative perspective) can consult Azkarate, 1990, 1991, 2002; Azkarate & Pérez Gaztelu, 2014; Azkue, 1923–25 [1969]; Euskaltzaindia, 1987; Irurtzun, 2013; Etxeberria et al., 2015; Jacobsen, 1982; Martínez Areta, 2003, 2006; Odriozola & Pérez Gaztelu, 2002; Pérez Gaztelu et al., 1996; Villasante, 1974, just to name a few.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

In this context, co-compounds are generally described as [color label A + color label B].[7] Co-compounds are further subdivided into two types. On the one hand, there are "coordinated co-compounds (dvanda)", that is, semantically endocentric compounds with two heads (Scalise et al., 2009) and with an additive meaning ([A color + B color] = A + B color). They refer to two (or three) colors put together but not mixed. For example, *zuri-nabar* 'white and dun', *gorri-hori* 'red and yellow', and *beltz-nabar* 'black and white'. On the other hand, there are "intermediate denoting co-compounds", that is, semantically exocentric headless compounds with an intersective meaning ([A color + B color] = C color, where C is a hybrid or intermediate color hue). They refer to one single color hue resulting from mixing two different colors, e.g., *hori-gorri* 'auburn', *ubel-beltz* 'bluish pale', and *beltz-hori* 'sooty'.

Some of these compounds, especially those with native color labels, can cover both functions, a different hue and two colors together. For instance, *gorri-urdin* encodes both 'purplish' and 'red and blue', *beltz-gorri* 'coppery', 'violet' and 'black and red', and *zuri-arre* 'off-white' and 'white and dun'.

This morphological strategy is fully productive, that is, any color combination is possible; however, some of these compounds are already lexicalized and compiled in dictionaries as subentries. Table 2 alphabetically summarizes those found in the OEH, the General Dictionary of the Basque Language.

Sub-compounds, the second group of compounds in this context, are generally described as cases where one of the elements in the compound modifies the other element. In Basque, sub-compounds are usually right-headed as in English; therefore, the general structure is [color modifier A + color head B]. Within this category, there are two subtypes. On the one hand, there are "possessor sub-compounds", that is, cases where the head B has A. They have the structure [object noun + *kolore* 'color'] that renders 'the color B that A has'. They are used to describe different color hues, e.g., *urre kolore* (gold color) 'the color of gold' or *azal kolore* (skin color) 'the color of skin.'[8] On some occasions, the second part of the compound could be omitted, especially in highly lexicalized cases such as

7. The order in this type of compounds is somewhat fixed; sometimes it follows rhythmical patterns (shorter word comes first), but many times what comes first is culturally determined (see de Rijk, 2008, p. 861). In color compounding, these rules are not that clear since both orders are possible, e.g., *zuri-gorri* and *gorri-zuri* 'rose' and 'red and white'.

8. As suggested by one of the reviewers, this type could be also considered an apposition with the token-type order; appositions in Basque are also right-headed. Either classification could be used in this study.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Color label+color label co-compounds in the OEH

Zuri	*Zuri-gorri* 'rosy'
	Zuri-arre 'white and dun', 'off-white'
	Zuri-berde 'white and green'
	Zuri-nabar 'white and dun'
	Zuri-hori-berde 'white, yellow, and green'
	Zuri-urdin 'white and blue'
Gorri	*Gorri-arroxa* 'dusk color'
	Gorri-beltz 'dark red', 'blond and black'
	Gorri-nabar 'blond'
	Gorri-hori 'red and yellow'
	Gorri-ubel 'violetish red'
	Gorri-urdin 'purplish', 'red and blue'
	Gorri-zuri 'blond'
Beltz	*Beltz-gorri* 'coppery', 'violet', 'black and red'
	Beltz-nabar 'black and white'
	Beltz-hori 'sooty'
	Beltz-urdin 'bluish black'
	Beltz-zuri 'black and white'
Urdin	*Urdin-arre* 'grey', 'blue and dun'
	Urdin-beltz 'dark blue'
	Urdin-zuri 'grey', 'light blue'
	Urdinabar 'grey'
Hori	*Hori-beltz* 'dark greenish-yellow'
	Hori-gorri 'auburn'
	Hori-gorrixka 'reddish'
	Hori-ubel 'buff-colored', 'pale blue'
	Hori-urdin 'green'
	Hori-zuri 'yellowish white'
Nabar	*Nabar-gorri* 'white red'
	Gorri-nabar 'blond'
	Nabar-beltz 'black and white pattern'
	Beltz-nabar 'black and white'
	Zuri-nabar 'white and dun'
Ubel	*Ubel-beltz* 'bluish pale'
Berde	*Berde-nabar* 'green and dun'
	Berde-hori 'yellowish green', 'green and yellow'
	Berde-urdin 'green and blue'
Azul	*Azul-zuri* 'blue and white'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

salmoi 'salmon'.[9] On the other hand, there are "subsumptive compounds", that is, hyponymic compounds where the resulting compound describes a color subtype. They have either the structure [color label A + color label B] or the structure [label A + color label B] where 'A (modifier) is a logical subtype of B (head)', e.g., *urdin berde* 'a type of green with some hints of blue (bluish green)' or *belar berde* (grass green) 'a type of green similar to the color of grass'.

What is interesting about these compound categories is that the same color compound label in isolation may trigger different interpretations as illustrated in (1).

(1) *Txuri* 'white' and *urdin* 'blue'
   a. *txuri-urdin* as a coordinated co-compound → 'white and blue'.
   b. *txuri-urdin* as an intermediate denoting co-compound → 'a white-blue blended color'
   c. *txuri-urdin* as a subsumptive sub-compound → 'a whitish type of blue'

This ambiguity disappears when these forms are actually used in specific contexts and their referents are clear. This is what Wälchli calls "contextual semantic sharpening", that is, the selection of "a single contextual meaning for words and morphemes from the variety of their possible conventionalized and non-conventionalized meanings" (2005, p. 32). For example, if the referent is a two-colored striped T-shirt, the interpretation corresponds to the example (1.a).[10] As discussed in Section 2, disambiguation between (1.b) and (1.c) is not always that easy. What is more, in the (1.c) pattern, there are two further possible interpretations. On the one hand, in regard to the modifier-head compound given in (1.c), where *urdin* 'blue' on the right is the head, the compound renders the meaning 'a type of blue'. However, this kind of structure is also open to a different interpretation. Instead of being regarded as compounds, they could be analyzed as syntactic N + Adj combinations,[11] and therefore, the meaning of this structure (*txuri urdin*) changes to 'a bluish type of white'. Orthographically, the latter case would not be written with a hyphen.

---

9. This also applies to other well-established color labels, such as *laranja* 'orange' from the fruit orange and *(l)arroxa* 'pink' from the rose flower. These are not included since they were assimilated in the language very early. Nevertheless, there is still a tendency to include the word *kolore* after these labels, i.e., *laranja kolore* and *arroxa kolore*; labels that are lexicalized in the language and included as subentries of these colors in the OEH.

10. In fact, this type of a compound in Basque also renders a semantically exocentric figurative co-compound reading. *Txuri-urdin* is also the metonymically based name given to a local football team (and its supporters). They wear white and blue striped jerseys.

11. Notice that case in Basque is assigned to the whole phrase and not to the head (e.g., *mutil handi-ari* is [boy big]-dative). Therefore, compounds and these NPs also coincide with respect to this formal property.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

In sum, what all these possibilities reveal is that [color label A + color label B] structures are problematic and cannot be rightly interpreted without context. Therefore, these structures will be discussed separately in Section 2.

Finally, there are another two possible compound forms. One is the compound[12] formed by the color label+*begi* 'eye'. Here, the body part *begi* 'eye' takes one of its metaphorical meanings as 'a bit', e.g., *zuri begi* (white eye) 'a little bit white' and *gorri begi* (red eye) 'a little bit red'. The second one is the compound color label+*aire/aide* 'similar' to *gorriaire* 'reddish' and *urreaire* 'shiny copper, similar to gold' (cf. *urraida* 'copper'; *zirraida* 'tin').

Another morphological process is derivation, also very common in Basque morphology (see Hualde, 2003b; de Rijk, 2008, chap. 32, 33). Basque has several suffixes related to color. The suffix *-kara* 'color' is used to create new color labels based on the object that has that color. For example, *lur* means 'earth' and, therefore, *lurkara* means 'earth color', *hauts* means 'dust' and *hauskara* 'dust color'. The two suffixes *-xka*/*-ska* and *-xko*/*-sko* denote an approximate color as in *gorrixka* 'reddish' and *berdexko* 'greenish'. The suffixes *-tsu* 'approximation' and *-txo* 'small' can also be used to describe an approximate color as in *berdetsu* 'greenish' and *gorritxo* 'reddish'.[13]

These devices can be combined. For example, in the structure color label+color label, the second element can also be derived by means of any of the abovementioned suffixes such as *-tsu* or *-txo*. On these occasions, the suffix refers to 'a bit of, a hint of' rather than to an approximation. For instance, *urdin berdetsu* and *urdin berdetxo* mean 'blue with a bit/a hint of green'.[14]

Reduplication is another morphological resource for color and a common device in Basque. It is used to express a distributive meaning (*banan-banan* 'one by one') and, even more frequently, to intensify or make more precise the meaning of

---

12. The status of *begi* in these formations is problematic. For some Basque grammarians *begi* is a noun and, therefore, these formations are taken as compounds; for others, on the other hand, *begi* has suffixal properties ("parasuffix", see Múgica, 1978; de Rijk, 2008) and, as such, these formations are derivational.

13. In general, these affixes have distributional restrictions in Basque: *-tsu* usually forms denominal adjectives, *-xka* deadjectival adjectives, and *-txo* both denominal and deadjectival adjectives. However, there are some exceptions (or problems, depending on the theoretical stand one takes about the category of color words) with color terms (e.g., *gorri* 'red' → *gorritsu* 'reddish'). See Martínez Areta, 2003, for more information.

14. Notice that in these cases the order reflects the phrasal structure noun+adjective in Basque (*mutil handi bati* ([boy big one]DAT) 'to a big boy'). That is why, *urdin berdetsua* ([blue greenish] ABS) means 'blue with a bit of green' and not 'a color similar to blue-green' (intermediate denoting co-compound) or 'a greenish color towards blue' (subsumptive sub-compound).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

the reduplicated morph (*handi-handi* 'very big'). In the latter cases, the morph can be repeated several times (triplication, etc.) in an iconic way; the more times it is repeated the more intense the meaning it conveys is (see de Rijk, 2008, pp. 877–882; Hualde, 2003a, pp. 360–362). Reduplication in the color domain serves one main purpose: to describe intensity; either the intensity of color as in *urdin-urdin* ('blue blue') 'very blue' or one of the qualities of color as in *urdin argi-argi* ('blue clear-clear') 'very light blue'.

As far as color value is concerned, that is, the lightness (tints) or darkness (shades) of color, Basque has two main adjectives, *argi* 'light' and *ilun* 'dark' as in *urdin argi* (blue light) 'light blue' and *urdin ilun* (blue dark) 'dark blue'. With respect to color saturation or intensity, that is, the brightness or dullness of color, they can be expressed by quantifiers such as *oso* 'very', as well as by reduplicated forms. As mentioned above, these adjectives and quantifiers can be also reduplicated to intensify their meanings (color *ilun-ilun-ilun* 'very dark color', *oso oso* color 'very, very color').

1.3 Basque lexicalization patterns for color

The previous sections have briefly explained the main lexicalization patterns available in Basque to express color (chromatic and achromatic), as well as to describe different hues, values, and saturation.[15] Table 3 provides a list of these color lexicalization patterns (hereafter, CLPs) together with an illustrative example and the main function of each CLP in the encoding of color in Basque. Section 2 explores how many and how often Basque speakers use these CLPs in present-day Basque.

Table 3. Basque lexicalization patterns for color

Color Lexicalization Pattern (CLP)		Example	Function
CLP1	Native single morph	*gorri* 'red'	Name color hue/ achromatism
CLP2	Old loan single morph	*berde* 'green'	Name color hue/ achromatism
CLP3	Recent loan single morph	*fuksia* 'fuchsia'	Name color hue/ achromatism
CLP4	Compounding: coordinated co-compound	*gorri-hori* 'red and yellow' *txuri-urdin* 'white and blue'	Name color hue/ achromatism

15. These are the main components of color. In this chapter, these components are used in the following sense: hue refers to the whole range of color, value to the lightness or darkness of a color, and saturation to the brilliance and intensity of color.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. (*continued*)

Color Lexicalization Pattern (CLP)		Example	Function
CLP5	Compounding: intermediate-denoting co-compound	*ubel-beltz* 'bluish pale' *txuri-urdin* 'a white-blue blended color'	Name color hue/ achromatism
CLP6a	Compounding: possessor sub-compound [object noun+*kolore*]	*azal kolore* 'skin color'	Name color hue/ achromatism
CLP6b	Compounding: possessor sub-compound [object noun+(*kolore*)]	*salmoi* (*kolore*) 'salmon (color)'	Name color hue/ achromatism
CLP7a	Compounding: subsumptive sub-compound [color label A + color label B]	*txuri-urdin* 'whitish type of blue'	Name color hue subtype
CLP7b	Compounding: subsumptive sub-compound [label A + color label B]	*belar berde* (grass green) 'grass green'	Name color hue subtype
CLP8	Compounding: color+*begi*	*gorri begi* 'a little bit red'	Describe color saturation
CLP9	Compounding: color+*aire/aide*	*gorriaire* 'reddish'	Describe hue similarity
CLP10	N + Adj. syntactic combination	*txuri urdin* 'bluish green'	Name color hue subtype
CLP11	Affixation: *-kara* 'color'	*hauskara* 'dust color'	Name color hue/ achromatism
CLP12	Affixation: *-xka/-ska*	*gorrixka* 'reddish'	Describe hue similarity
CLP13	Affixation: *-xko/-sko*	*berdexko* 'greenish'	Describe hue similarity
CLP14	Affixation: *-tsu* 'approximation'	*berdetsu* 'greenish'	Describe hue similarity
CLP15	Affixation: *-txo* 'small'	*gorritxo* 'reddish'	Describe hue similarity
CLP16	Combination of devices: color+color+ *-tsu/-txo*	*urdin berdetsu/berdetxo* 'blue with a bit of green'	Name color hue/ achromatism + Describe color saturation
CLP17	Reduplication: color-color; color adjective-adjective…	*urdin-urdin* 'very blue' *urdin argi-argi* 'very light blue'	Describe color saturation
CLP18	Adjectives: *argi* 'light', *ilun* 'dark'	*urdin argi* 'light blue' *urdin ilun* 'dark blue'	Describe color value
CLP19	Quantifiers: *oso* 'very', *oso oso* 'very, very'	*urdin oso iluna* 'very dark blue' *urdin oso oso iluna/urdin ilun-iluna* 'very, very dark blue'	Describe color saturation

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. The use of Basque lexicalization patterns for color

This section explores how Basque native speakers categorize and exploit the nineteen color lexicalization patterns described in Table 3. The main goal is to establish how often and for which purpose these lexicalization patterns are used.

2.1 Methodology

2.1.1 *Stimuli, procedure, and coding*

Data have been collected using the methodology proposed in the "Evolution of semantic systems" project developed at the Max Planck Institute for Psycholinguistics. The EoSS project aims at describing how meanings vary, how they are categorized and expressed by speakers of different languages in four main semantic domains: body parts (parts of objects), containers (kinds of objects), spatial relations (relations between objects), and color (attributes of objects) (see Majid et al., 2015, for an overview).

The color semantic domain was tested by means of two different tasks. The color elicitation task consisted of 84 Munsell color chips arranged in a single fixed random order and presented individually on a numbered grey box as shown in Figure 1a (see Majid & Levinson, 2007; Majid et al., 2010). Color chips varied on the basis of hue, saturation, and brightness. 80 chips showed 20 hues at four degrees of brightness and four chips were achromatic. The participants had to name the color in the chip. Responses were codified according to "full response" (the whole answer; e.g., *light blue-green*) and "main response" (answer without modifiers or hedges; if two basic colors were mentioned, both were coded; e.g., *blue-green*). The second task, the focal color, consisted of the same 84 Munsell color chips. These chips were individually identified with a letter (A – D) and a number (1–20), and

Figure 1a. Color elicitation task.

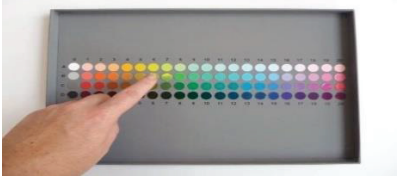


Figure 1b. Focal color task.

Figure 1. Tasks in the color semantic domain (adapted from Majid et al., 2010, pp. 25, 29)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

arranged on a two-dimensional array on the basis of hue and brightness with achromatic colors situated on the left-hand side, as shown in Figure 1b (Majid, 2008; Majid et al., 2010). The participants were asked to point to the color chip that best corresponded to a previously elicited list of twelve main colors in language. The participants had to identify the best example by letter and number. All participants were screened for color-blindness (Waggoner, 2002).

2.1.2 *Participants*

For the Basque language, 20 participants took part in the elicitation tasks: twelve women and eight men. Their mean age is 20.8 (age range 19–24). Basque (mainly Gipuzkoan variety) was their first language from birth but they were all early native speakers of Spanish (some from birth, some from school age, 3–5). They reported knowledge of other languages such as English, French, and Portuguese. This chapter reports results from the color elicitation task (1680 answers in total; full and main responses).

2.2 Analysis

Section 1 described the different color lexicalization patterns (CLPs) that the Basque language offers to encode color in language. These patterns serve to fulfill five main functions: to name color hue or achromatism, to name a color hue subtype, to describe color hue similarity, to describe color value, and to describe color saturation. In the following, each of these functions is examined in order to determine which lexicalization patterns Basque speakers prefer to codify each function. As explained in Section 1.2, color label+color label structures may correspond to four different color lexicalization patterns (that is, CLP4, CLP5, CLP7a, and CLP10); it is for this reason that these structures will be analyzed separately in Section 2.2.6.

2.2.1 *Color hue and achromatism*

Color hue refers to the whole range of color, and achromatism to colors that lack hue, i.e., white, grey, and black. Basque can codify hue/achromatism by means of single morphs (CLP1, CLP2, CLP3), compounding (CLP6), and affixation (CLP11). Table 4 summarizes the use of these CLPs by Basque informants.

Table 4 shows that all the CLPs for naming color hue and achromatism are used but one, the suffix *-kara* (CLP11). Primary colors (blue, red, yellow) and white and black are mainly described by means of native words (CLP1), whereas secondary colors (violet, orange, green) and grey are encoded by means of old loanwords (CLP2). The rest of the color hues are encoded by recent loanwords (CLP3) and compounds (CLP6). Most of the types included in these two latter categories have few tokens but for a few exceptions: *marroi* 'brown', *lila* 'lilac' and *granate* 'maroon'

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Usage of CLPs to describe color hue and achromatism in Basque

CLP	#Type	*Hues/achromatism* and (#tokens)
CLP1 Native single morph	5	*urdin* 'blue' (349), *hori* 'yellow' (94), *beltz* 'black' (67), *zuri* 'white' (54), *gorri* 'red' (41)
CLP2 Old loan single morph	10	*berde* 'green' (423), *more* 'purple' (173), *arroxa* 'pink' (154), *laranja* 'orange' (87), *gris* 'grey' (54), *azul* '(dark)blue' (5), *purpura* 'purple' (3), *anil* 'indigo' (1), *bioleta* 'violet' (1), *doratua* 'golden' (1)
CLP3 Recent loan single morph	11	*marroi* 'brown' (115), *lila* 'lilac' (53), *granate* 'maroon' (30), *kaki* 'khaki' (10), *fuksia* 'fuchsia' (7), *magenta* 'magenta' (7), *okre* 'ochre' (4), *marengo* 'dark grey' (3), *pardo* 'dun' (3), *fosforito* 'fluorescent' (1), *zian* 'cyan' (1)
CLP6a Compounding: object noun+*kolore*	8	*haragi kolore* 'flesh col.' (30), *pistatxo kolore* 'pistachio col.' (6), *azal kolore* 'skin col.' (4), *urre kolore* 'gold col.' (4), *txerri kolore* 'pig col.' (2), *belar kolore* 'grass col.' (1), *brontze kolore* 'bronze col.' (1), *zeru kolore* 'sky col.' (1)
CLP6b Compounding: object noun(+*kolore*)	5	*salmoi* 'salmon' (5), *turkesa* 'turquoise' (5), *koral* 'coral' (1), *mostaza* 'mustard' (1), *oliba* 'olive' (1)
CLP11 Affixation: *-kara* 'color'	0	–

in CLP 3 and *haragi kolore* 'flesh color' in CPL 6. Appendix 1 offers a summary of the distribution of the labels Basque speakers chose to encode the 84 Munsell color chips used in the color elicitation task.

2.2.2 *Color hue subtype*

There were two main lexicalization patterns in naming a subtype of a color hue: subsumptive sub-compounds formed by [color label A + color label B] (CLP7a) or by [label A + color label B] (CLP7b), and CLP10, i.e., N + Adj. syntactic combinations. Both CLP7a and CLP10 are discussed in Section 2.2.6.

As far as CLP7b is concerned, there is only one type with one token in the data: *belar berde* (grass green) 'grass green'. This does not exactly mean that Basque speakers hardly ever use this type of lexicalization pattern where the color label is modified by another element. What Basque speakers seem to do is to calque or adapt this lexicalization pattern straight from Spanish. Table 5 compiles these cases. Spanish original terms are within square brackets.

Sub-compounds are right-headed in Basque as in *belar berde* (grass green) but left-headed in Spanish as in *azul cielo* (blue sky) 'sky blue'. As shown in Table 5, Spanish color sub-compounds have been used by Basque speakers to describe color hue subtypes. In some cases, calques are used, as in *gris marengo* from *gris marengo*

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Spanish color sub-compound loanwords in Basque

Color hue	Hue subtypes	#Types	#Tokens
arroxa 'pink'	*arroxa palo* [*rosa palo*] (pink tipuana tree) 'rosewood' (5), *arroxa pastel* [*rosa pastel*] (pink cake) 'pastel pink' (2)	2	7
azul 'blue'	*azul marino* [*azul marino*] (blue marine) 'navy blue' (3), *azul turkesa* [*azul turquesa*] (blue turquoise) 'turquoise blue' (2)	2	5
berde 'green'	*berde botil* [*verde botella*] (green bottle) 'bottle green' (6), *berde kaki* [*verde caqui*] (green khaki) 'khaki green' (10), *berde lima* [*verde lima*] (green lime) 'lime green' (3), *berde limoi* [*verde limón*] (green lemon) 'lemon green' (1), *berde militarra* [*verde militar*] (green army's) 'military green' (2), *berde pastel* [*verde pastel*] (green cake) 'pastel green' (2), *berde pistatxo* [*verde pistacho*] (green pistachio) 'pistachio green' (17), *berde turkesa* [*verde turquesa*] (green turquoise) 'turquoise green' (4)	8	45
gorri 'red'	*gorri burdeos* [*rojo burdeos*] (red bordeaux) 'bordeaux red' (1)	1	1
gris 'grey'	*gris marengo* [*gris marengo*] (grey dark grey) 'dark grey' (3)	1	3
hori 'yellow'	*hori fosforito* [*amarillo fosforito*] (yellow fluorescent(N)) 'fluorescent yellow'(1), *hori txilio* [*amarillo chillón*] (yellow loud(adj.)) 'loud yellow' (2)	2	3
laranja 'orange'	*laranja melokotoi* [*naranja melocotón*] (orange peach) 'peach orange' (1), *laranja pastel* [*naranja pastel*] (orange cake) 'pastel orange' (1), *laranja teja* [*naranja teja*] (orange tile) 'brick red' (1)	3	3
urdin 'blue'	*urdin cielo* [*azul cielo*] (blue sky) 'sky blue' (1), *urdin marino* [*azul marino*] (blue marine) 'navy blue' (4), *urdin palo* [*azul palo*] (blue tipuana tree) 'pale blue' (1), *urdin pastel* [*azul pastel*] (blue cake) 'pastel blue' (5), *urdin turkesa* [*azul turquesa*] (blue turquoise) 'turquoise blue' (6)	5	17

'dark grey'. On the other hand, some other cases contain a Basque "adapted" element as in *urdin cielo* from *azul cielo* 'sky blue' and *gorri burdeos* from *rojo burdeos* 'bordeaux red'. In general, the use of these loanwords is not that salient. GRUE colors are the ones that produce more different kinds of hue subtypes and tokens, especially *berde* 'green' with seven types, and within it, *berde pistatxo* 'pistachio green' with 17 tokens.

2.2.3 *Color hue similarity*

The color hue similarity refers to those cases where an approximate color is described. Hue similarity is encoded by means of compounding (CLP9) and, mainly, by affixation (CLP12, CLP13, CLP14, and CLP15). Table 6 summarizes the results.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Usage of CLPs to describe hue similarity in Basque

CLP	#Types	Cases and #Tokens
CLP9 Compounding: color+*aire/aide*	0	---
CLP12 Affixation: *-xka/-ska*	14	*berde horixka* (green yellowish) 'yellowish green' (1), *berde marroixka* (green brownish) 'brownish green' (1), *berde urdinxka* (green bluish) 'bluish green' (6), *berde zurixka* (green whitish) 'whitish green' (1) *gris beltzexka* (grey blackish) 'blackish grey' (1) *hori berdexka* (yellow greenish) 'greenish yellow' (5) *laranja horixka* (orange yellowish) 'yellowish orange' (1) *marroi horixka* (brown yellowish) 'yellowish brown' (2) *pistatxo horixka* (pistachio yellowish) 'yellowish pistachio' (1) *urdin berdexka* (blue greenish) 'greenish blue' (6), *urdin morexka* (blue purplish) 'purplish blue' (1), *urdin zurixka* (blue whitish) 'whitish blue' (1) *zuri arrosaxka* (white pinkish) 'pinkish white' (1), *zuri morexka* (white purplish) 'purplish white' (1)
	4	*berdexka* 'greenish' (1), *gorrixka* 'reddish' (1), *horixka* 'yellowish' (5), *marroixka* 'brownish' (3)
CLP13 Affixation: *-xko/-sko*	1	*berde ilunxko* (green darkish) 'darkish green' (1)
CLP14 Affixation: *-tsu* 'approximation'	2	*urdin berdetsu* (blue greenish) 'greenish blue' (1), *laranja haragi koloretsu* (orange flesh colorish) 'flesh-colored orange' (1)
CLP15 Affixation: *-txo* 'small'	1	*hori laranjatxo* (yellow orangish) 'orangish yellow' (1)

The most frequent CLP by far is the diminutive *-xka/-ska*. It was used with 18 types of hues. It is interesting to notice that most of these strategies applied to color compounds but for the diminutive *-xka/-ska*; this suffix was used with four one morph color labels (*berde*, *gorri*, *hori*, and *marroi*).

Apart from these lexicalization patterns, Basque speakers used other means to express hue similarity usually with the general meaning of 'towards', 'between', or 'similar'. These are: the directional allative case *-rantz* 'towards', the adverb *ia* 'almost', the spatial noun *alde* 'side' + allative case *-ra*: *aldera* 'towards', the spatial noun *(t)arte* 'space between' + locative case *-an*/relational suffix *-ko*: *(t)artean/arteko* 'in between', the noun *erdi* 'half', the noun *antz* 'appearance, likeness' + relational suffix *-ko*: *antzeko* 'similar', the color label+nominalized participial form *-tua*, the allative case *-ra* and the verb *hurbildu* 'approach, get closer', the inclusive

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

disjunctive conjunction *edo* 'or', and even one calque, the Spanish approximation suffix *-oso* '-ish'. Table 7 summarizes these strategies ordered by usage.

Table 7. Usage of alternative CLPs for color hue similarity in Basque

Other CLPs	#Types	Cases and #Tokens
ia 'almost'	5	*urdin ia zuri* 'blue almost white' (4), *marroi ia beltz* 'brown almost black' (2), *more ia zuri* 'purple almost white' (2), *ia beltz* 'almost black' (1), *lila ia zuri* 'lilac almost white' (1)
(t)artean/arteko 'in between'	5	*hori urre tartean* 'between yellow and golden' (1), *more eta lila artean* 'between purple and lilac' (1), *zuri gris tartean* 'between white and grey' (1) *berde eta urdin arteko* 'between green and blue' (1), *urdin eta lilaren arteko* 'between blue and lilac' (1)
aldera 'towards'	4	*laranja zuri aldera* 'orange towards white' (1), *marroi beltz aldera* 'brown towards black' (1), *urdin zuri aldera* 'blue towards white' (1), *horixka laranja aldera* 'yellowish towards orange' (1)
color+ *-tua*	4	*hori berdetua* (yellow greened) 'greenish yellow' (1), *horixka laranjatua* (yellowish oranged) 'orangish yellowish' (1), *hori marroitua* (yellow browned) 'brownish yellow' (1), *zuri berdetua* (white greened) 'greenish white' (1)
edo 'or'	2	*beltz edo* 'black or so' (1), *marroi edo beltz* 'brown or black' (1)
directional allative *rantz* 'towards'	2	*arroxa gorrirantz* 'pink towards red' (1), *lila morerantz* 'lilac towards purple' (1)
erdi 'half'	1	*erdi zuri erdi urdin* 'half white half blue'(1)
antzeko 'similar'	1	*urdin antzeko* 'similar to blue' (1)
allative *-ra* + *hurbildu* 'approach'	1	*urdin lilara hurbiltzen* 'blue towards lilac' (1)
Spanish *-oso* '-ish'	1	*hori berdoso* (yellow-greenish) 'greenish-yellow' (1)

2.2.4 *Description of color value*

The color value refers to the degree of lightness and darkness of color. Color value could be encoded by means of the *argi* 'light' and *ilun* 'dark' adjectives (CLP18). Table 8 summarizes the usage of this CLP.

As shown in Table 8, light colors are mostly expressed by the adjective *argi* 'light' (a total of 234 occurrences for 10 hue types) and dark colors by the adjective *ilun* 'dark' (a total of 202 occurrences for 13 hue types). Lightness and darkness can also be expressed by other adjectives, some of which are loanwords from Spanish, but their use is quite marginal in comparison to these options. Lightness is also expressed by adjectives such as *klaru* 'light' (7), *suabe* 'soft' (5), *palido* 'pale' (2), and *itzalia* 'muted' (1), while darkness is expressed by adjectives such as *fuerte* 'strong' (18), *intentxo* 'intense' (2), and *ez oso argia* 'not very light' (1).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 8. Usage of CLP18 to describe color value in Basque

CLP	#Types	Modified colors and (#tokens)
CLP18 Adjectives: *argi* 'light'	10	*urdin argi* 'light blue' (98), *berde argi* 'light green' (66), *more argi* 'light purple' (25), *laranja argi* 'light orange' (19), *arroxa argi* 'light pink' (15), *marroi argi* 'light brown' (11), *lila argi* 'light lilac' (8), *gris argi* 'light grey' (4), *hori argi* 'light yellow' (2), *gorri argi* 'light red' (1)
CLP18 Adjectives: *ilun* 'dark'	13	*berde ilun* 'dark green' (86), *urdin ilun* 'dark blue' (60), *marroi ilun* 'dark brown' (21), *more ilun* 'dark purple' (21), *arroxa ilun* 'dark pink' (9), *gris ilun* 'dark grey' (4), *hori ilun* 'dark yellow' (4), *lila ilun* 'dark lilac' (4), *gorri ilun* 'dark red' (3), *granate ilun* 'dark maroon' (3), *magenta ilun* 'dark magenta'(1), *haragi kolore ilun* 'dark flesh color' (1), *pistatxo kolore ilun* 'dark pistachio' (1)

2.2.5 *Description of color saturation*

Color saturation refers to the brilliance and intensity of color. There were three main CLPs for this function: color+*begi* 'eye (a bit)' compounds (CLP8), tri/reduplication (CLP17), and the quantifier *oso* 'very' with or without reduplication (CLP 19). Table 9 summarizes the usage of these CLPs.

Table 9. Usage of CLPs to describe color saturation in Basque

CLP	#Types	Color saturation and (#tokens)
CLP8 color+*begi*	0	–
CLP17 Reduplication: color-color	1	*urdin-urdin* 'blue-blue' (1)
CLP17 Reduplication: color adjective-adjective	7	*arroxa argi-argi* 'light-light pink' (3), *berde argi-argi* 'light-light green' (2), *lila argi-argi* 'light-light lilac' (1), *more argi-argi* 'light-light purple' (2), *urdin argi-argi* 'light-light blue' (10) *berde ilun-ilun* 'dark-dark green' (1), *marroi ilun-ilun* 'dark-dark brown' (1)
CLP17 Triplication	1	*lila argi-argi-argi* (lilac light light light) 'extremely light lilac' (1)
CLP19 Quantifiers: *oso* 'very'	13	*berde oso argi* 'very light green' (1), *berde oso klaru* 'very light green' (1), *laranja oso argi* 'very light orange' (1), *lila oso argi* 'very light lilac' (1), *marroi oso argi* 'very light brown' (1), *more oso argi* 'very light purple' (6), *urdin oso argi* 'very light blue' (15), *urdin oso suabe* 'very soft blue' (3) *berde oso ilun* 'very dark green' (12), *gris oso ilun* 'very dark grey' (1), *marroi oso ilun* 'very dark brown' (3), *more oso ilun* 'very dark purple' (3), *urdin oso ilun* 'very dark blue' (6)
CLP19 Quantifiers: *oso oso* 'very, very'	3	*lila oso oso argi* 'very, very light lilac' (1), *urdin oso oso argi* 'very, very light blue' (1) *berde oso oso ilun* 'very, very dark green' (1)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

As shown in Table 9, most cases of color saturation have to do with the intensification of color value, lightness, and darkness, especially in the case of GRUE colors (blue and green). There is only one instance of reduplication to encode the intensity of a color hue (*urdin* 'blue'). Apart from these CLPs, Basque informants also employed other strategies to describe intensity: on the one hand, the quantifier *pixka bat* 'a bit' as in *arroxa haragirekin pixka bat* 'pink with a bit of flesh (color)' (only 1 token), and on the other, the suffix *-kada* 'a heap of' together with the diminutive suffix *-xka*/*–ska*. There are three different types in the data with only one token each: *berde marroixkada* 'green with a heap of brownish', *berde urdinxkada* 'green with a heap of bluish', and *urdin berdexkada* 'blue with a heap of greenish'.

2.2.6 *[Color label A + color label B] structures*

The discussion in Section 1.2 about [color label A + color label B] structures in Basque reveals that these structures, albeit problematic, can be interpreted as four different color lexicalization strategies with two main functions:

- CLP4 and CLP5 → to name color hue/achromatism
- CLP7a and CLP10 → to name a color hue subtype

Table 10 compiles all the [color label A + color label B] structure types found in the data. For the sake of convenience and an objective analysis, these types (i) are organized by the first color label, i.e., *hori berde* goes into the *hori* + color label category (independently from its meaning), (ii) are not written with a hyphen, and (iii) are only glossed not translated.

Table 10 shows that there are a total of 45 different types and 117 tokens.[16] However, many of these are only used once or twice, with the exception of GRUE colors, especially *berde urdin* (green blue) with 22 tokens and *urdin berde* (blue green) with 17 tokens.

The major problem that one faces with these data is how to interpret the [color label A + color label B] structures. The experimental design itself only discards one color lexicalization pattern, CLP4: coordinated co-compounds, that is to say, cases where there are two (or three) colors put together but not mixed. There is no single Munsell color chip like this.

However, the experimental elicitation procedure does not provide further clues for the other three CLPs since the participants only had to name each color chip without any context. In other words, there was no follow-up session where the experimenter could ask them for further explanations about their label choices.

---

16. The reader should bear in mind that the use of these structures is very low. There is a total of 1680 answers in this color naming task (see Section 2.1); this means that [color A + color B] structures only represent 6.96% of the participants' responses and, therefore, they do not undermine the results presented in previous sections.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 10. [Color label A + color label B] structures in Basque data

Color label A + color label B	Cases	#Types	#Tokens
arroxa + color label	*arroxa azal kolore* (pink skin col.) (1), *arroxa gorri* (pink red) (1), *arroxa haragi kolore* (pink flesh col.) (5), *arroxa lila* (pink lilac) (1), *arroxa more* (pink purple) (1), *arroxa zuri* (pink white) (2)	6	11
berde + color label	*berde beix* (green beige) (1), *berde beltz* (green black) (1), *berde hori* (green yellow) (6), *berde marroi* (green brown) (5), *berde pardo* (green dun) (2), *berde urdin* (green blue) (22), *berde zuri* (green white) (1)	7	38
gorri + color label	*gorri arroxa* (red pink) (2), *gorri laranja* (red orange) (2), *gorri marroi* (red brown) (2)	3	6
granate + color label	*granate marroi* (maroon brown) (1)	1	1
gris + color label	*gris beltz* (grey black) (1)	1	1
hori + color label	*hori berde* (yellow green) (8), *hori laranja* (yellow orange) (1), *hori marroi* (yellow brown) (2)	3	11
laranja + color label	*laranja arroxa* (orange pink) (1), *laranja gorri* (orange red) (1), *laranja haragi kolore* (orange flesh color) (1), *laranja hori* (orange yellow) (2), *laranja marroi* (orange brown) (1), *laranja zuri* (orange white) (1)	6	7
marroi + color label	*marroi beltz* (brown black) (2), *marroi berde* (brown green) (1), *marroi granate* (brown maroon) (1), *marroi hori* (brown yellow) (3), *marroi laranja* (brown orange) (1)	5	8
more + color label	*more arroxa* (purple pink) (1), *more lila* (purple lilac) (1)	2	2
pistatxo + color label	*pistatxo hori* (pistachio yellow) (1)	1	1
urdin + color label	*urdin berde* (blue green) (17), *urdin gris berde* (blue grey green) (1), *urdin more* (blue purple) (2), *urdin zuri* (blue white) (4)	4	24
urre kolore + color label	*urre kolore pistatxo* (gold color pistachio) (1)	1	1
zuri + color label	*zuri arroxa* (white pink) (1), *zuri berde* (white green) (1), *zuri gris* (white grey) (2), *zuri more* (white purple) (2)	4	6
Total		44	117

For instance, explicit questions such as: "For this color chip, you used 'orange red', do you consider this chip as a type of 'orange' or a type of 'red' or something in between, both at the same time?", or specific contexts for these terms such as, "Would you describe the color of the sunset in a summer day as 'red', 'orange' or 'orange red'? Why?", etc.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

One way to avoid this problem might be to examine all the participants' answers for each individual Munsell color chip where this structure is used and to decide the CLP on the basis of the most frequent answer. There are twenty answers per color chip and these structures are not used by all participants at all times. Therefore, if most of the participants chose the term 'red', for example, and one uses 'orange red', one could assume that the participant's strategy is CLP7a, 'a type of red'. This might work for some cases as illustrated in Table 11. Full responses correspond to the participants' literal transcriptions and the main responses to answers without modifiers or hedges (color terms are always kept, see Section 2.1.1). CLPs are included in the main responses. Nineteen speakers have categorized 53 N7 as 'grey'; therefore, the experimenter can consider the structure *zuri gris* as an example of CLP7a, 'a type of grey', i.e., 'whitish grey'.

Table 11. [Color label A + color label B] structures as CLP7a

Color chip	**Participants' responses**			
	Full responses	**1st main resp.**		**2nd main resp.**
	Type (#Token)	**Type (#Token)**	**CLP**	
C0	*grisa* (grey.ABS) 'grey' (16)	*gris* 'grey' (16)	CLP2	
	gris argia (grey clear.ABS) 'light grey' (3)	*gris* 'grey' (3)	CLP18	
(53) N7	*zuri gris bat* (white grey one) '**???**' (1)	*gris* 'grey' (1)	CLP7a	*zuri* 'white' (1)
Total	***gris* 'grey' (20)**			

However, these assumptions are intrinsically ambiguous and problematic. Firstly, the researcher is inferring (and not demonstrating) from the participant's own categorizations. Secondly, it does not hold for most cases. Table 12 illustrates the participants' responses to a couple of Munsell color chips. CLPs are added to the main responses.

For the first color chip in Table 12, the participants used two [color label A + color label B] structures, namely, *gorri laranja* and *laranja marroia*, and three different colors to name it: *gorria* 'red', *laranja* 'orange', and *marroia* 'brown'. Given these three choices, either of the CLPs make sense: they could be taken as a strategy to name a color hue subtype (a type of orange and a type of brown (CLP7a) or a type of red and a type of orange (CLP10)), and also as an intermediate denoting color compound (a red-orange blended color and an orange-brown blended color (CLP5)). The second color chip in Table 12 yields similar results. The participants mostly used two main color categories: *berdea* 'green' and *urdina* 'blue'. The *berde urdina* structure is also open to all the aforementioned possibilities: a type of blue (CLP7a), a type of green (CLP10), and a green-blue blended color (CLP5).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 12. [color label a + color label B] structures with multiple CLPs.

Color chip	Participants' responses				
	Full responses	**1st main resp.**		**2nd main resp.**	
	Type (#Token)	**Type (#Token)**	**CLP**	**Type**	**CLP**
C2	*gorria* (4)	*gorria* (4)	CLP1		
	gorrixka (1)	*gorria* (1)	CLP12		
(12) 10R 4/12	*gorria laranja puntu batekin* (1)	*gorria* (1)	CLP1	*laranja*	hue sim.
	gorri laranja (1)	*laranja*	CLP2	*gorria*	CLP7a
		or	**or**	**or**	**or**
		gorri-laranja	CLP5	*laranja*	CLP10
	laranja marroia (1)	*marroia*	CLP3	*laranja*	CLP7a
		or	**or**	**or**	**or**
		laranja-marroia	CLP5	*marroia*	CLP10
	laranja (5)	*laranja* (5)	CLP1		
	laranja gorri aldera (1)	*laranja* (1)	CLP2	*gorria*	hue sim.
	laranja iluna (2)	*laranja* (2)	CLP18		
	laranja ilun gogorra (1)	*laranja* (1)	CLP18		
	marroia (3)	*marroia* (3)	CLP3		
Total	***gorria* (6), *laranja* (9), *marroia* (3),*??* (2)**				
B11	*berdea* (4)	*berdea* (4)	CLP2		
	berde klarua (1)	*berdea* (1)	CLP18		
(78) 5BG 6/10	*berde lima* (1)	*berdea* (1)	CLP7b		
	berde urdina (3)	*urdina*	CLP1	*berde*	CLP7a
		or	**or**	**or**	**or**
		berde-urdina	CLP5	*urdina*	CLP10
	berde urdinxkada (1)	*berdea* (1)	CLP2	*urdina*	saturat.
	berde urdin *argia* (1)	*urdina*	CLP1	*berde*	CLP7a
		or	**or**	**or**	**or**
		berde-urdina	CLP5 (+CLP18)	*urdina*	CLP10
	turkesa 1	*turkesa* (1)	CLP6b		
	urdina 4	*urdina* (4)	CLP1		
	urdin suabea 1	*urdina* (1)	CLP18		
	urdin turkesa 2	*urdina* (2)	CLP7b	*turkesa*	
Total	***berdea* (7), *turkesa* (1), *urdina* (7),*??* (4)**				

In sum, it is impossible to interpret these structures with the current data, and therefore, further research is needed in the future. There are at least a couple of possible ways to expand this investigation. One could be to undertake an acoustic analysis of these data. The prosodic characteristics of these structures could be examined in order to unveil different/similar patterns depending on the type of CLP

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

used. This is not such an easy task in Basque since suprasegmentals work differently depending on the language variety, and neither the nature of Basque compounds nor their accentual patterns are completely understood yet (Hualde, 1997, 1999; see Irurtzun, 2013, for Central Basque). Another way could be to conduct a follow-up study to measure the distance from the colors involved in these [color A + color B] structures with respect to their corresponding focal colors. This type of study could reveal whether these colors are considered "types of" or "blended colors" depending on how close or distant they are categorized with respect to focal color chips.

3. Conclusions

This chapter has offered an overview of the encoding of color in Basque. The first part has discussed the linguistic mechanisms employed in Basque to describe (a) chromatic colors and their main components, hue, value, and saturation. Eighteen main lexicalization patterns were proposed. These include the use of native and loanword labels, morphological processes such as compounding, affixation, and reduplication, as well as the use of adjectives.

The second part has focused on Basque speakers' categorization of color on the basis of a color elicitation task. In general, results show that most of the color lexicalization patterns proposed in Section 1.3 are frequently used. Some CLPs, nevertheless, were not or hardly ever used by the speakers: compounds with *begi* 'eye' to describe color saturation, compounds with *aire/aide* 'similar' to describe hue similarity, and the suffix *-kara* to name color hue/achromatism. Others, such as the suffixes *-(s)/(x)ko*, *-tsu*, and *-txo* describing hue similarity, were hardly ever used.

The color value is mainly described by means of adjectival complementation (*argi* 'light' and *ilun* 'black'). Reduplication, as well as quantifiers (*oso* 'very'), are the main CLPs to encode color saturation. As far as the naming of color hues and achromatism goes, there are several possibilities. Basque speakers widely use native single morph labels (e.g., *zuri* 'white'), as well as old and recent loanwords (*berde* 'green', *marroi* 'brown') to encode primary and secondary colors, and compounding (object noun+*kolore* 'color', *azal kolore* 'skin color') to express the rest of the color palette.

The set of basic color names in Basque, nowadays, includes both ancient and recent labels. The assimilation of new color words has narrowed the range of color hues covered by single labels, e.g., the use of *gris* instead of *urdin* for 'grey', the use of *laranja* for 'orange', and *marroi* 'brown' instead of *gorri*. Data confirm that these "new" colors are fully integrated into the Basque color system, *gris*, *laranja*, and *marroi* are the labels used by most speakers to describe these colors. As a consequence, native colors such as *urdin* and *gorri* are almost exclusively used for

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

'blue' and 'red', respectively (see Appendix 1). Another interesting result of this color loanword assimilation is that some native color words are not used at all by Basque speakers in this study. An illustrative example is *arre* 'muddy-brown, dun, earth color'. Basque speakers prefer to use *marroi* 'brown', and even *pardo* 'dun' instead of *arre*. The same applies to other labels in these groups such as *nabar* 'dun, motley' and *ubel* 'purple, violetish, greyish' (the latter is mostly restricted to the meaning 'bruise').

This tendency to use non-native means in the domain of color is not restricted to the incorporation of new loanwords. The first example, as observed in the data, would be the use of the object noun+*kolore* 'color'. The Basque suffix *-kara* covers exactly that function. Another example would be the use of foreign words to describe color components (hues, value, and saturation) instead of native words, e.g., Spanish *klaru*, *suabe*, and *palido* instead of *argi* 'light'. This also affects cases where, although the words are Basque, the structure does not reflect a Basque pattern. An illustrative example is the abovementioned use of different adjectives to describe tints (either *palido* or *argi*). Basque has a whole set of "fading" colors such as *hits* or *subil* that could be used on these occasions, but none of them turns up in the data. Basque speakers prefer descriptions such as *hori suabe*, *hori palido*, and *hori argi* 'light yellow' instead of the native one morph label *laru* 'pale yellow'.

Several conflating reasons could explain these preferences in the data. One could be the restricted stimuli. It is true that the 84 Munsell chips covered a wide color spectrum; however, it is also true that several specific color hues were either not included or hardly represented. Another reason, perhaps the crucial one, is related to the informants' traits. They are all young speakers (aged 19–24), and literate in Standard Basque (all schooled in Basque). The list of basic colors taught in Standard Basque at schools encourages the uses of these labels instead of other options.

This issue was already pointed out by Preciado Sáez de Ocáriz (1992) in his early study on Basque basic colors. This author compared the color system of very young speakers (aged 6–15, schooled in Basque) with that of old speakers (aged 65-onwards, Basque illiterate farmers). Similar results were found. Colors encoded by new words in younger speakers, native color names identified with single hues in younger speakers and not with the whole range of colors as in older speakers, etc. What is more, Preciado Sáez de Ocáriz also found that, beyond different labels and a narrower semantic scope, the conceptualization and perception of colors was different. The colors used to describe objects and ordinary elements were different. For example, whereas the color of earth for the younger speakers was *marroi* 'brown', for the older people was *beltz* 'black' and *gorri* 'red' (see Preciado Sáez de Ocáriz, 1993, for more examples). In other words, it seems that two different color systems, a romanized and an ancient, coexist.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The coexistence of two color systems is a well-known fact in Basque (see, Txillardegi, 1975; Sarrionandia, 1985; Azurmendi, 1991; Preciado Sáez de Ocáriz, 1997). However, on the basis of both Preciado Saéz de Ocáriz's work and the results in this study, the romanized model seems to have gained ground, especially in the last half century. Another finding in this chapter that seems to corroborate this tendency towards a romanization of color description is the use of left-headed compound loanwords to describe color hue subtypes. Basque compounds are right-headed whereas Romance languages such as Spanish are left-headed (e.g., *azal kolore* (skin color) vs. *color carne* (color flesh) 'skin color'). When describing color hue subtypes, Basque participants mostly prefer to calque or adapt Spanish color compound loanwords (*gris marengo* (grey dark grey) 'dark grey' or *urdin marino* [*azul marino*] (blue marine) 'navy blue') rather than use Basque-like formations such *belar berde* (grass green) 'grass green', which is used once.

Authors such as Preciado Sáez de Ocáriz (2003, 2006a, 2006b, 2007, 2009) have often claimed that it is necessary to bring knowledge about the ancient native Basque system back to schools. This is important not only to enrich the younger generations' vocabulary of Basque colors, but also to enable them to understand a way of thinking and conceptualizing the world. After all, languages are the accretion of a culture's knowledge, with words and structures as their main carriers.

Acknowledgements

This research has been supported by grants FFI2013–45553–C3–1–P and FFI2017–82460–P from the Spanish Ministry of Economy, Industry, and Competitiveness and by the Government of Aragon (Psylex H11-17R). I would like to thank Koldo J. Garai, the research assistant for Basque at the EoSS project, for his work in data collection and transcription, as well as the editors for their enormous patience and the two anonymous reviewers for the their helpful and insightful comments. Special thanks to Bárbara Marqueta, Beatriz Fernández, José Luis Mendívil, and Julio Santiago.

References

Azkarate, M. (1990). *Hitz elkartuak euskaraz*. Donostia: Deustuko Unibertsitatea.

Azkarate, M. (1991). *Hitz-elkarketa /3: LEF batzordearen lanak*. Bilbo: Euskaltzaindia.

Azkarate, M. (2002). Hitz-elkarketa, argumentu-egitura eta theta rolak. In X. Artiagoitia, P. Goenaga, & J. A. Lakarra (Eds.), *Erramu Boneta. Festschrift for Rudolf P. G. de Rijk, ASJUren Gehigarriak XLIV* (pp. 91–104). Bilbo: UPV/EHU Press.

Azkarate, M., & Pérez Gaztelu, E. (2014). *Hitz-elkarketa /2: LEF batzordearen emaitza*. Bilbo: Euskaltzaindia.

Azkue, R. Mª. (1923–25 [1969]). *Morfología vasca*. Bilbao: La gran enciclopedia vasca.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Azurmendi, M. (1991). Palabras en color y cosas en color en la cultura vasca tradicional. In Arteleku (Ed.), *Arte y estética. Metáforas de la identidad* (pp. 55–72). Donostia: Arteleku.

de Rijk, R. P. G. (2008). *Standard Basque. A Progressive Grammar*. Cambridge, MA: MIT Press.

Etxeberria, U., Etxepare, R., & Irurtzun, A. (2015). Euskarazko hitz-elkartu kopulatiboen egitura eta interpretazioaz. In B. Fernández, & P. Salaburu (Eds.), *Ibon Sarasola, Gorazarre. Homenatge, Homenaje* (pp. 189–200). Bilbao: EHU Press.

Euskaltzaindia. (1987). *Hitz-elkarketa /1: LEF batzordearen lanak*. Bilbo: Euskaltzaindia.

Gorrochategui, J., Igartua, I., & Lakarra, J. (Eds.). (2018). *Historia de la lengua vasca*. Gasteiz: Eusko Jaurlaritza.

Hualde, J. I. (1997). *Euskararen azentuerak. Supplements of ASJU XLII*. Donostia: UPV/EHU & Gipuzkoako Foru Aldundia.

Hualde, J. I. (1999). Basque accentuation. In H. van der Hulst (Ed.), *Word Prosodic Systems in the Languages of Europe* (pp. 947–994). Berlin: Mouton de Gruyter.

Hualde, J. I. (2003a). Compounds. In J. I. Hualde, & J. Ortiz de Urbina (Eds.), *A Grammar of Basque* (pp. 351–362). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110895285.351

Hualde, J. I. (2003b). Derivation. In J. I. Hualde, & J. Ortiz de Urbina (Eds.), *A Grammar of Basque* (pp. 328–351). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110895285.328

Ibarretxe-Antuñano, I. (2017). Basque ideophones from a typological perspective. *The Canadian Journal of Linguistics / Revue canadienne de linguistique, 62*(2), 1–21. https://doi.org/10.1017/cnj.2017.8

Irurtzun, A. (2013). Focal NSR and Stress Placement in Basque Phrases and N+N Compounds. *Linguistic Analysis, 38*(3–4), 207–242.

Jacobsen, W. H., Jr. (1982). Basque copulative compounds: A problem in irreversible binomials. *Proceedings of the Eight Annual Meeting of the Berkeley Linguistics Society, 18*, 384–397.

Lakarra, J. (1995). Reconstructing the Pre-Proto-Basque Root. In J. L. Hualde, J. Lakarra, & R. L. Trask (Eds.), *Towards a History of the Basque Language* (pp. 189–206). Amsterdam: John Benjamins.

Lakarra, J. (2002). Etymologiae (proto)uasconicae LXV. In X. Artiagoitia, P. Goenaga, & J. Lakarra (Eds.), *Erramu boneta: Festschrift for Rudolf P. G. de Rijk (ASJU Supplement 44*) (pp. 425–442). Bilbo: UPV/EHU Press.

Larzabal, P. (1975). Euskaldunak eta koloreak. *Euskera*, *XX*, 401–405.

Lieber, R., & Stekauer, P. (2009). *The Oxford Handbook of Compounding*. Oxford: Oxford University Press.

Majid, A. (2008). Focal colors. In A. Majid (Ed.), *Field Manual Volume 11* (8–10). Nijmegen: Max Planck Institute for Psycholinguistics.

Majid, A., Jordan, F. M., & Dunn, M. (2010) *Evolution of Semantic Systems: Procedures Manual*. Max Planck Institute for Psycholinguistics, Nijmegen.

Majid, A., Jordan, F. M., & Dunn, M. (2015). Semantic systems in closely related languages. *Language Sciences*, *49*(1), 1–18. https://doi.org/10.1016/j.langsci.2014.11.002

Majid, A., & Levinson, S. C. 2007. The language of vision I: color. In A. Majid (Ed.), *Field Manual Volume 10* (22–25). Nijmegen: Max Planck Institute for Psycholinguistics.

Mitxelena, K. (1985–2005). *Orotariko Euskal Hiztegia*. Bilbo: Euskaltzaindia.

Martínez Areta, M. (2003). Hitz-konposakera euskaraz eta aitzin-euskaraz. *Anuario del Seminario de Filología Vasca "Julio de Urquijo"*, *XXXVII* (1), 1–84.

Martínez Areta, M. (2006). Adjektiboa aitzin-euskaraz. *Anuario del Seminario de Filología Vasca "Julio de Urquijo"*, *XL*, 687–722.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Martínez Areta, M. (2008, September 10). Sobre la formación de palabras antigua en euskera. Workshop "Monumenta Linguae Vasconun proiektua (2002–2008). Sei urte ikerketan". *Retrieved from* https://www.ehu.eus/documents/6174861/6203039/Martinez-Sobre_la_formacion_de_palabras_antigua_en_euskera.pdf

Martínez Areta, M. (2013). *Basque and Proto-Basque. Language-internal and typological approaches to linguistic reconstruction*. Frankfurt: Peter Lang. https://doi.org/10.3726/978-3-653-02701-3

Michelena, L. (1990). *Fonética histórica vasca*. Donostia: Gipuzkoako Foru Aldundia.

Múgica, L. M. (1978). *Origen y desarrollo de la sufijación euskérica*. Donostia/San Sebastián: Ediciones Vascas.

Odriozola, J. C., & Pérez Gaztelu, E. (2002). Aposizioa euskal hitz-elkarteetan. In X. Artiagoitia, P. Goenaga, & J. A. Lakarra (Eds.), *Erramu Boneta. Festschrift for Rudolf P. G. de Rijk, ASJUren Gehigarriak XLIV* (pp. 467–478). Bilbo: UPV/EHU Press.

Ortiz-Osés, A., & Garagalza, L. (2006). *Euskal mitologia: izena duen guztia omen da*. Donostia: Fundación Kutxa.

Perurena, P. (1992). *Koloreak euskal usarioan*. Donostia: Erein.

Preciado Sáez de Ocáriz, J. Mª. (1992). *Kolorea herri euskaldun batean*. Leioa, E.H.U.: Argitarapen Zerbitzua.

Preciado Sáez de Ocáriz, Tx. (1993). Bizitakoa versus ikasitakoa: horia eta urdina gatazka honen adibideak. *Tantak*, *9*, 103–120.

Preciado Sáez de Ocáriz, Tx. (1997). Unibertsalismoaren arrapadalan. *Revista de Psicodidáctica*, *3*, 93–112.

Preciado Sáez de Ocáriz, Tx. (2003). Eskolak ongi erakusten du koloreak ikusten?. *Hik hasi*, *83*, 24–29.

Preciado Sáez de Ocáriz, Tx. (2006a). Euskal ikusteko era ez ezik, koloreen artean ere badago bazterreraturik. *Tantak*, *36*, 73–103.

Preciado Sáez de Ocáriz, Tx (2006b). *Euskaldunon kolore mundua atsotitzetan*. Donostia: Gaiak.

Preciado Sáez de Ocáriz, Tx. (2007). Enseñanza-color-País Vasco: claves de una reflexión (ávida de esperanza)/Teaching-color-Basque Country: keys to reflection (keen for hope). *Arte Individuo y Sociedad*, *19*, 155–178.

Preciado Sáez de Ocáriz, Tx. (2009). *Euskarak erakutsi koloreak ikusten*. Gasteiz: Eusko Jaurlaritza.

Sarrionandia, J. (1985). *Ni ez naiz hemengoa*. Donostia: Pamiela.

Scalise, S., Fábregas, A., & Forza, F. (2009). Exocentricity in compounding. *Gengo Kenkyu*, 135, 49–84.

Txillardegi. (1975). La structuraction du champ semantique de la couleur en Basque, *Euskera*, *XX*, 221–238.

Van Eys, W. J. (1873). *Dictionnaire Basque-Français*. Paris/London: Maisonneuve/Williams & Norgate.

Villasante, L. (1974). *Palabras vascas compuestas y derivadas*. Oñati: Editorial Franciscana Aránzazu.

Wälchli, B. (2005). *Co-compounds and Natural Coordination*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199276219.001.0001

Waggoner, T. L. (2002). *Quick six color vision test pseudoisochromatic plates. From Color vision testing made easy* by Good-Lite Company.

Zuazo, K. (2013). *The dialects of Basque*. Reno: the Center for Basque Studies, University of Nevada.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Appendix 1. Basque speakers' labels and usage for the description of color hue and achromatism

Color hue	Types	CLP	Tokens
A14 B14 B15 B13 C13 C14 C15 D14 B12 A15 D13 A13 D15 A12 A11	*Urdin*	CLP1	349
A4 A5 A6	*Hori*	CLP1	94
D0	*Beltz*	CLP1	67
A0	*Zuri*	CLP1	54
C1	*Gorri*	CLP1	41
B8 B9 C9 C8 C10 C7 B10 AB D8 D10 D9 A7 A9 C6 A10	*Berde*	CLP2	423
D19 D17 D16 B16 B17 C16 C17	*More*	CLP2	173
A19 B19 B20 A20 A18 A1	*Arroxa*	CLP2	154
B2 B3	*Laranja*	CLP2	87
B0 C0	*Gris*	CLP2	54
D15 D13 C11 C10	*Azul*	CLP2	5
D16 D17 D18	*Purpura*	CLP2	3
D15	*Anil*	CLP2	1
D20	*Bioleta*	CLP2	1
1	*Doratu*	CLP2	1

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Color hue	Types	CLP	Tokens
D3 C3 C4	*Marroi*	CLP3	115
B10 B11	*Lila*	CLP3	53
C19	*Granate*	CLP3	30
C5	*Kaki*	CLP3	10
C19	*Fuksia*	CLP3	7
C19 C20	*Magenta*	CLP3	7
B4 B5 C4	*Okre*	CLP3	4
1 1 1	*Marengo*	CLP3	3
1 1 1	*Pardo*	CLP3	3
B7	*Fosforito*	CLP3	1
B12	*Zian*	CLP3	1
A2 A1	*haragi kolore*	CLP6	30
A7	*pistatxo kolore*	CLP6	6
A1 A2	*azal kolore*	CLP6	4
C4 C5 C6	*urre kolore*	CLP6	4
B1 A3	*txerri kolore*	CLP6	2
1	*belar kolore*	CLP6	1
C4	*brontze kolore*	CLP6	1
B12	*zeru kolore*	CLP6	1
B11 C12	*Turkesa*	CLP6	6
A1 B1 A2 A20	*Salmoi*	CLP6	5
A2	*Koral*	CLP6	1
B4	*Mostaza*	CLP6	1
B5	*Oliba*	CLP6	1
	39 types		**1804 tokens**

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lexicalization patterns in color naming in Croatian, Czech, and Polish

Ida Raffaelli, Jan Chromý and Anetta Kopecka
University of Zagreb / Charles University, Prague / CNRS & Université Lumière Lyon 2

The goal of this chapter is to identify and describe strategies speakers of Croatian, Czech, and Polish use in color naming. The findings are based on the data from the cross-linguistic project Evolution of Semantic Systems (Max Planck Institute for Psycholinguistics, the Netherlands). Croatian, Czech, and Polish were among more than 50 Indo-European languages included in the project. In this study, we present the results of the color naming task for the three languages. The study identified the main lexicalization patterns that are productive in the formation of Croatian, Czech, and Polish color terms. They are the results of different grammatical mechanisms used in the lexicalization process. However, the languages differ with respect to the degree of conventionalization of these mechanisms in the domain of color terms.

Keywords: color terms, lexicalization patterns, word-formation, Croatian, Polish, Czech

1. Introduction[1]

The color domain has been one of the most researched, as well as one of the most propulsive areas in the field of lexical semantics. The study that had a significant impact on the investigation of color terms is the influential Berlin and Kay's research (1969) of color terms in various languages. It has become the basis for other types of research on color terms (e.g., Conklin, 1973; Comrie & Corbett, 1993; Lucy, 1997) and further developed since its publication. As Majid, Jordan, and Dunn (2015)

1. This chapter is based on the article by Raffaelli (2017) which contains a detailed analysis of Croatian EoSS color terms data. The same theoretical and methodological framework has been implemented in this research in order to provide comparable data that could yield relevant insights into lexical structures of the three languages.

https://doi.org/10.1075/sfsl.78.12raf
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

have pointed out, Berlin and Kay's methodology was refined in the World Color Survey (Kay et al., 2009) – the largest empirical study of semantics ever to have been undertaken, encompassing 110 languages spoken mostly in small, preliterate, and non-industrialized communities.[2] The World Color Survey enabled researchers to examine the inventory of basic color terms (i.e., terms which are morphologically simple, not restricted to a class of objects, cognitively salient, and present in each speaker's idiolect) and to investigate cognitive principles underlying color naming and categorization (e.g., Heider & Olivier, 1972; MacLaury, 1987; Kay, Berlin, & Merrifield, 1991).

As emphasized by Majid, Jordan, and Dunn (2015), while Berlin and Kay's work has been an inspiration for many types of research, it has also been criticized for over-sampling Indo-European (IE) color terms (Lucy, 1997; Wierzbicka, 2005). Furthermore, it has also prompted some researchers to consider cross-linguistic differences and their effects on perception (e.g., Roberson, Davies & Davidoff, 2000; Roberson, Davidoff, Davis & Shapiro, 2005). The research on basic color terms was conducted in regard to some IE sub-families, including the Slavic one (Comrie & Corbet, 1993). Yet, despite numerous studies on basic color terms in individual languages, color naming has not been the subject of a large-scale investigation in IE languages, and we still lack a systematic account of the diversity of constructions IE languages use to name colors. Therefore, there was a need for a more integrative research that would take into account data from a large number of IE languages. Such an endeavor was the project called the Evolution of Semantic Systems (EoSS). The project was conducted at the Max Planck Institute for Psycholinguistics (Nijmegen) from 2011 to 2014 and involved research on more than 50 IE languages. The aim of the project was to investigate how meanings vary across Indo-European languages and, in particular, how similar or different the semantic categories of colors (attributes of objects), containers (kinds of objects), body parts (parts of objects), and spatial relations (relations between objects) are.

In this chapter, we examine three of the Slavic languages investigated in the context of the EoSS project: Croatian (South Slavic), Czech, and Polish (both West Slavic). These languages constitute an interesting case study, since they show many typological (including morphological, syntactic, and lexical) similarities. Yet, the question arises: do speakers of these languages use similar types of constructions when naming colors?

The main goals of this chapter are threefold: (a) to gain insight into strategies Croatian, Czech, and Polish speakers use in naming colors; (b) to determine

2. Majid, Jordan and Dunn (2015).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

similarities and differences between lexicalization patterns[3] that are conventionalized and thus productive in the formation of color terms in the three languages; (c) to gain insight into the color terms speakers use in the partition of the color spectrum and that can be categorized either as conventionalized or novel expressions. All findings stem exclusively from the data collected within the EoSS project. More precisely, we investigate the data collected from native speakers of the three languages (20 speakers of Croatian and Czech, and 21 speakers of Polish). The data were collected with a color elicitation task that consisted of 84 Munsell color chips (materials developed by Majid & Levinson, 2007).[4] For both Croatian and Czech, we collected 1680 full responses; for Polish, it was 1764 responses.

The chapter is structured as follows. After the introductory part, we provide the theoretical background for the analysis and the explanation of the notion of lexicalization pattern adopted in this study. In the third section, we describe the methods used in the research. In the fourth section, the results are presented and discussed. Some concluding remarks are given at the end of the chapter.

2. Theoretical background

Three main theoretical tenets relevant for the analysis of language data provide the basis for the analysis of the lexicalization patterns that the speakers of Croatian, Czech, and Polish use in color naming. These tenets are: (1) the debate between the universalist and relativist approach to colors, (2) the statement that the grammar of a language – seen from the cognitive linguistic perspective – is a characterization of an established linguistic convention (Langacker, 1987, p. 62), and (3) the claim that the lexicalization of different meanings is underpinned by different grammatical processes (morphological and syntactic).

2.1 Universalist vs relativist approach to the study of colors

As emphasized by Regier, Kay and Khetarpal (2007), the recent debate about color naming has been dominated by two major views. One is that color categories are organized around universal foci, while the other is that color categories are

3. There is a difference between strategies and lexicalization patterns. Strategies are procedures speakers use in the process of naming. Strategies are based on the speakers' knowledge of language. Lexicalization patterns are, more or less, conventionalized languages structures that are, more or less, frequently used in the process of naming. Not all strategies equal lexicalization patterns.

4. Details on the methods employed are presented in the third section.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

determined at their boundaries by linguistic convention. The first view is universalist. Proponents of the universalist approach argue for the existence of universal focal colors. According to this view, focal colors are linguistically or culturally independent, and are thus not related to the differences between languages or the cultures in which these languages are spoken (cf. Regier, Kay & Khetarpal, 2007).

The second view is relativist. The proponents of this approach claim that the language structure dictates the partition of the color spectrum, as well as that the boundaries between the color areas are connected to local linguistic conventions and can differ in relation to language diversity. The findings presented in Regier, Kay and Khetarpal (2007) showed that neither of these two opposing approaches is fundamentally accurate. According to their findings, the position in the spectrum and color categories are governed and shaped, respectively, by universal perceptual constraints. However, their results attest to the important role that linguistic convention might play in determining color categories.

This chapter adopts the relativist view, considering its primary objective is to investigate the kind of strategies speakers use in color naming in order to gain insight into the differentiation and partition of the color spectrum.

2.2 Linguistic conventionalization

One of the main interests of the analysis of the Croatian, Czech, and Polish EoSS data is to reveal various lexicalization patterns used in color naming and determine the extent to which these patterns are conventionalized in the three languages. According to Langacker (1987), linguistic units become conventionalized when they become part of the speakers' shared knowledge, thus pertaining to a speech community and being frequently used in communication.[5]

However, we consider that the process of conventionalization cannot be explained exclusively in terms of the speakers' frequent use of a certain expression or as a 'shared knowledge' in a speech community. Conventionalization is a much more complex process that cannot be defined without relation to language structure. Žic Fuchs and Tuđman Vuković (2000) point out that the conventionalization of a certain linguistic expression should be regarded with respect to its syntactic-semantic integration into the language structure. In Raffaelli (2009), this is further developed through the notion of the conceptual and structural embeddedness of a language expression. It means that conventionalization subsumes structural integration of a language expression on the syntagmatic and the paradigmatic level, as well as

5. Langacker (1987, pp. 56–69) elaborates the genesis of linguistic units as conventionalized expressions.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

on the conceptual level, capturing a certain conceptual structure and encoding a certain meaning.

The EoSS data allow for identifying several conventionalized patterns used frequently in color naming by Croatian, Czech, and Polish speakers. However, as stressed by Langacker (1987, p. 59), conventionalization is a matter of degree. Regarding conventionalization with respect to the structural integration of a language expression, the notion of degree or continuum is of considerable importance as well. Thus, some linguistic expressions are more structurally integrated than others, as we shall see based on the data from Croatian, Czech, and Polish discussed in this chapter. Hence, a distinction between more conventionalized and less conventionalized lexicalization patterns and terms will be made and the continuum of patterns used in color naming will be presented in the following sections. Moreover, it is expected that Croatian, Czech, and Polish will show similarities (since all are Slavic languages), but also certain differences (since they are part of different Slavic sub-families) with respect to the degree of conventionalization of determined lexicalization patterns.

2.3 Lexicalization patterns in color naming

Before proceeding to the analysis and the results, the term *lexicalization pattern* and the importance of examining lexicalization patterns in color naming should be, respectively, defined and explicated.[6]

The term *lexicalization pattern*, as presented in this chapter, is akin to Talmy's eponymous concept.[7] However, our approach and Talmy's differ to an extent. In comparison to Talmy, we primarily define lexicalization patterns on the morphosemantic level. This means that, in addition to the word-formation patterns, the lexicalization patterns also comprise other grammatical (e.g., syntactic) patterns used in naming concepts. In line with Talmy, we believe that the analysis of lexicalization patterns is comprised of defining various lexical forms that encapsulate the way speakers conceptualize perceptual experiences. Moreover, we consider lexicalization patterns as a language structure underpinned by diverse grammatical

6. The model is presented in more details also in Raffaelli (2013, 2017) and Raffaelli and Kerovec (2017).

7. Talmy (1985, p. 57) defines lexicalization patterns as a relation between meaning and surface expressions, and investigates which semantic elements are expressed by which surface elements. Semantic elements of different types can be expressed by the same type of a surface element, as well as the same type of semantic element can be expressed by several different linguistic expressions. A range of typological patterns and universal principles can be found by such an approach.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

processes, either morphological or syntactic, regarding them as equally important and enabling lexicalization of meanings.[8]

Hence, in this study, the model of lexicalization patterns is used to describe linguistic strategies speakers use in color naming that become more or less conventionalized. For the three Slavic languages, which are morphologically rich, it is of the utmost significance to define the word-formation patterns speakers use in the lexicalization of different colors.[9] Moreover, besides derivation, it is necessary to define the role of other grammatical processes such as compounding and diverse syntactic processes in the formation of new terms.

3. Method and data

To collect the data, we used 84 Munsell color chips selected for the purpose of the EoSS project (cf. Majid & Levinson, 2007). Color chips were arranged in a single fixed random order and they differed with respect to brightness, hue, and saturation. 20 native speakers have been interviewed in Croatian and in Czech, and 21 in Polish; they were all students in the humanities and/or social sciences aged between 18 and 26. The participants were asked to name the color chips presented to them individually (one at a time). The data sample consists of 1680 responses in Croatian and Czech respectively, and 1764 in Polish.

Table 1 shows the diversity of linguistic strategies (lexicalization patterns)[10] speakers of Croatian, Czech, and Polish used for naming color chips. For each language, the table gives (1) the construction type (i.e., a specific lexicalization pattern) and (2) the number of types and tokens (both number of occurrences and their proportion).

The results show both differences and similarities among the three languages. It can be observed that there are important differences between the three languages with regard to the diversity of construction types observed in the data: in Croatian and Polish, 163 and 195 construction types were found respectively, while in Czech there are as many as 263 types; these include simple and derived terms, as well as more complex constructions, and comprise both recurrent patterns and

8. See also Raffaelli and Kerovec (2017).

9. In morphologically reach languages word-formation processes are integrated in lexicalization of new lexemes. Therefore, word-formation and lexicalization are regarded as complementary and unseparable.

10. Note that the simple color terms are integrated as well. This shows the ratio between the usage of simple terms versus terms formed via certain lexicalization pattern.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Lexicalization patterns used by Croatian, Czech, and Polish speakers in color naming

	Croatian		Czech		Polish	
	Type	**Token**	**Type**	**Token**	**Type**	**Token**
Adj (simple)	19 (11.7%)	641 (38.2%)	10 (3.8%)	504 (30%)	15 (7.7%)	668 (37.9%)
Adj (derived)	9 (5.5%)	331 (19.7%)	27 (10.3%)	399 (23.8%)	34 (17.4%)	585 (33.2%)
N	–	–	2 (0.8%)	4 (0.2%)	19 (9.7%)	107 (6.1%)
Adj [o] Adj	43 (26.4%)	122 (7.3%)	55 (20.9%)	161 (9.6%)	25 (12.8%)	128 (7.3%)
Adv + Adj	50 (30.7%)	507 (30.2)	67 (25.5%)	483 (28.8%)	–	–
Adj + Adj	–	–	52 (19.8%)	68 (4.1%)	39 (20%)	144 (8.2%)
Adj + N	–	–	2 (0.8%)	2 (0.1%)	27 (13.9%)	96 (5.4%)
Adv + Adj [o] Adj	1 (0.6%)	1 (0.06%)	11 (4.2%)	19 (1.1%)	2 (1%)	2 (0.1%)
N + Ngen	4 (2.5%)	10 (0.6%)	–	–	–	–
Others	37 (22.7%)	68 (4.05%)	37 (14.1%)	40 (2.4%)	34 (17.4%)	34 (2%)
Total	163 (100%)	1680 (100%)	263 (100%)	1680 (100%)	195 (100%)	1764 (100%)

idiosyncratic uses. Among the different types, the use of simple adjectives is the most pervasive in all three languages: it accounts for 38.2% in Croatian, 30% in Czech, and 37.9% in Polish. As for the use of other, structurally more complex constructions, there are several differences among the three languages, with Czech sharing some similarities with Croatian and others, much fewer, with Polish: first, while common in the three languages, the use of derived adjectives is much more frequent in Polish (33.2%) than in Croatian (20.2%) and Czech (23.8%); second, speakers of Croatian and Czech make frequent use of the [*Adv Adj*] construction, which amounts to 30.2% and 28.8% respectively, whereas this construction type was not found in the Polish data. Finally, while Polish speakers resort to the [*N*], [*Adj Adj*], and [*Adj N*] constructions with 6.1%, 8.2%, and 5.4% respectively, these constructions are rare in the Czech data and they are not present in Croatian. It can be observed that among recurrent constructions found in the three languages, there are compound adjectives [*Adj [o] Adj*], the use of which accounts for 7.3% in Croatian, 9.6% in Czech and 7.3% in Polish.

In this chapter, we focus more specifically on the use of structurally complex constructions found in Croatian, Czech, and Polish, including derivation and compounding, and investigate the types of meanings they lexicalize.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4. Types of lexicalization patterns used in color naming in Croatian, Czech, and Polish

The analysis will focus on morphologically complex terms, i.e., terms that are formed via derivation, compounding, or one of the other word-formation processes found in the data (cf. Table 1). Thus, simple terms, either Adj such as Cro. *plava* 'blue', Cze. *žlutá* 'yellow', or Pol. *biała* 'white', or loan words like Cro. *bež* 'beige' will not be taken into consideration. Only complex terms like Cro. *ljubičasta* 'violet-like',[11] Cze. *lososová* 'salmon-like', Pol. *oliwkowy* 'olive-like'; Cro. *bijelo-plava* 'white-blue', Cze. *žlutozelená* 'yellow-green', Pol. *jasnozielony* 'light green'; Cro. *tamno plava* 'dark blue', Cze. *bledě modrá* 'pale blue' will be analyzed. It is important to note that some terms, which are loan words from other non-Slavic languages may have undergone derivation in some of the three languages but not in all three. For example, Czech has a derived term *béžová* 'beige', but in Croatian there is a simple term *bež*. In Polish, the term *bordowy* 'bordeaux' is used, but in Croatian and Czech, this loan word has not undergone derivation (there is *bordo* in Croatian and *bordó* in Czech).[12] Hence, depending on their morphological structure, they will be included in or excluded from the analysis. All these latter terms are instantiations of diverse lexicalization patterns and thus illustrate the most prominent strategies speakers use to name a certain color. However, in the domain of the color spectrum, they are not equally conventionalized.

4.1 Suffixation as a lexicalization pattern in the three languages

In general, the participants often used derived terms in all the three languages (i.e., in 20.2% (340) of cases in Croatian, in 23.8% (399) of cases in Czech, and in 33.2% (585) of cases in Polish). However, the type frequency of derived terms differed widely in the three languages. It seems that Croatian has a fairly limited number of types of derived terms (only nine found in the data), whereas Czech and Polish employ derivation more frequently in forming color terms (the derived terms constituted 10.3% (27) of types in Czech and 17.4% (34) of types in Polish). Thus, suffixation as a process of forming color terms is productive to a different extent in the three languages, with low productivity in Croatian and relatively high productivity in Czech and Polish. The ten most frequent derived terms in all the three languages are presented in Table 2.

---

11. *Ljubičasta* 'purple' or 'violet-like' is a derived color term formed from the stem *ljubič* 'violet' and the derivational morpheme *-asta* 'like'.

12. The non-derived term *bordo* also exists in Polish, but it was used only once in the EoSS task.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. The ten most frequent derived color terms in the three languages

Croatian	N	Czech	N	Polish	N
ljubičasta 'purple' (violet-like)	208	*fialová* 'purple' (violet-like)	148	*fioletowy* 'purple' (violet-like)	157
narančasta 'orange-like'	68	*růžová* 'pink' (rose-like)	103	*różowy* 'pink' (rose-like)	125
ružičasta 'pink' (rose-like)	33	*oranžová* 'orange'	59	*brązowy* 'brown'	83
maslinasta 'olive-like'	11	*okrová* 'ocre'	13	*pomarańczowy* 'orange'	56
plavkasta 'blueish'	3	*lososová* 'salmon-like'	12	*granatowy* 'dark blue'	47
zelenkasta 'greenish'	3	*tyrkysová* 'turqouise'	11	*bordowy* 'bordeaux'	16
breskvasta 'peach-like'	2	*béžová* 'beige'	7	*seledynowy* 'aquamarine'	14
rozikasta 'pinkish'	2	*vínová* 'wine-like'	7	*purpurowy* 'purple'	13
žućkasta 'yellowish'	1	*purporová* 'magenta'	5	*turkusowy* 'turquoise'	13
–	–	*cihlová* 'brick-like'	4	*oliwkowy* 'olive-like'	8
OTHER	0	OTHER	30	OTHER	53
TOTAL	331	TOTAL	399	TOTAL	585

First, regarding Croatian, since suffixation is common in the formation of adjectives in this language, one might expect this process to be productively used in color naming. However, as the data in Table 2 show, this is not the case. We found only nine color terms formed via suffixation in the EoSS Croatian data. These terms are instantiations of two lexicalization patterns: (1) [stem + *-ast(a)*] as in *narančasta* lit. 'orange-like' and (2) [stem + *-kast(a)*] as in *rozikasta* 'pinkish'. The [stem + *-ast(a)*] pattern conveys the meaning of 'N-like' and allows color terms to be derived from nouns referring to flowers and fruits.[13] The second lexicalization pattern, [stem + *-kast(a)*], conveys the meaning of 'approximation' as in *žućkasta* 'yellowish', *breskvasta* 'peach-like' or *zelenkasta* 'greenish'. These are the only four color terms in the Croatian EoSS data formed with the suffix *–kast*. As shown by the number of occurrences, this pattern is used extremely rarely to name colors.

In Czech, there are 27 derived types (399 tokens) and the typical lexicalization pattern is [stem + *-ov(á)*] and it is very similar to the Croatian pattern [stem + *-ast(a)*], while the derived terms are typically formed from nouns referring to flowers or fruits. However, there are also exceptions like *cihlová* ('brick-like'), *lososová* ('salmon-like'), or *petrolejová* ('petrol-like') in the data. The other lexicalization patterns corresponding to the main suffixation processes in the Czech data are [stem + *-ck(á)*] and [stem + *-sk(á)*], which are common in forming general

13. Results of previous research (Conklin, 1973; Wierzbicka, 2005; Malt & Majid, 2013) show that languages regularly and frequently use terms for familiar and well-known entities to name colors.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

relational adjectives in Czech but are rather rarely used in forming color terms. Each pattern was represented only with one token in the data: *námořnická* ('navy') and *svinibrodská* ('color from Svinibrod (Schweinfurt)').

In Polish, we found 34 types of derived color terms (585 tokens). Like in Czech, the most common pattern is [stem + *-ow(y)*], which is typically used to derive color terms from names of fruits (e.g., *pomarańczowy* 'orange'), flowers (e.g., *różowy* 'rose-like'), and other entities such as, for example, minerals (e.g., *turkusowy* 'turquoise'). While many of the derived terms with the suffix *-ow(y)* are highly conventionalized and part of the color lexicon, some of them (e.g., *marchewkowy* 'carrot-like', *pistacjowy* 'pistachio-like') show idiosyncratic uses and little agreement across speakers. The suffix *-ow(y)* is the most frequent in the Polish data and covers 99.8% of the derived terms. Yet, we found another derivational suffix in the dataset which is *-aw(y)* as in *zielonkawy* 'greenish'. Similar in its meaning to the English suffix *-ish* and the Croatian suffix *-kast(a)* we discussed earlier, the suffix *-aw(y)* conveys the meaning of attenuation or approximation.

4.2 Compounding as a lexicalization pattern [Adj [o] Adj] in the three languages

Compounding is the second lexicalization pattern to be described in this section. This lexicalization pattern, structured as [Adj [o] Adj], enables the formation of compound color terms. Terms formed on the basis of this pattern contain two adjectives (typically, simple color terms, though other terms – such as adjectives of brightness (light, dark) – can be used as well) and the linking morpheme -o- that couples the two items. Such terms are, for example, Cro. *zeleno-plava* 'blue-green', Cze. *žlutozelená* 'yellow-green', or Pol. *pomarańczoworóżowy* 'orange-pink'.

In all the three languages, the token frequency of compound terms was generally lower than the token frequency of derivations; it presents 7.3% (122) of the tokens in Croatian, 9.6% (161) in Czech, and 7.3% (128) in Polish. However, the three languages differ in terms of the diversity of types found among the compound terms. In the Polish data, the type frequency of compounds (12.8%, i.e. 25 types) was lower than in Croatian and Czech. For Croatian and Czech, the data show a substantially higher type frequency of this pattern: 26.4% (i.e. 43 types) in Croatian and 20.9% (i.e. 55 types) in Czech. According to the data, the lexicalization pattern [Adj [o] Adj] is more conventionalized and therefore more entrenched in Croatian and Czech speakers' knowledge about the language strategies in color naming than the pattern [stem + *-ast(a)*] or [stem + *-ov(á)*], particularly since it allows for more creativity and productivity and is instantiated with more diverse terms used in color naming. Although the absolute frequency of the occurrences

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

of terms formed via the patterns [stem + *-ast(a)*] or [stem + *-ov(á)*] is higher than that of the terms formed by means of the [Adj [o] Adj] pattern, because of the larger creativity and productivity, we consider the second one more conventionalized, i.e., more entrenched. In Polish, the situation is different and it seems that the pattern [Adj [o] Adj] is only mildly productive, perhaps, even less than derivation. The ten most frequent terms formed according to this pattern in all the three languages are presented in Table 3.

Table 3. The ten most frequent terms formed via the pattern [Adj [o] Adj] in the three languages

Croatian	N	Czech	N	Polish	N
maslinasto-zelena 'olive-like-green'	16	*modrozelená* 'blue-green'	19	*ciemnozielony* 'dark-green'	44
žuto-zelena 'yellow-green'	11	*zelenomodrá* 'green-blue'	15	*jasnozielony* 'light-green'	14
plavo-zelena 'blue-green'	9	*žlutozelená* 'yellow-green'	11	*jasnoniebieski* 'light-blue'	13
zeleno-plava 'green-blue'	9	*tmavozelená* 'dark green'	9	*jasnobrązowy* 'light-brown'	7
zeleno-smeđa 'green-brown'	7	*zelenohnědá* 'green-brown'	9	*ciemnożółty* 'dark-brown'	6
crveno-narančasta 'red-orange'	6	*zelenožlutá* 'green-yellow'	9	*jasnofioletowy* 'light-purple'	6
sivo-zelena 'gray-green'	5	*tmavomodrá* 'dark blue'	8	*ciemnobrązowy* 'dark-brown'	5
smeđe-crvena 'brown-red'	5	*fialovorůžová* 'purple-pink'	6	*bladoróżowy* 'pale-pink'	4
smeđe-zelena 'brown-green'	5	*bílomodrá* 'white-blue'	4	*ciemnoróżowy* 'dark-pink'	4
sivo-maslinasta 'gray-olive-like'	5	*hnědozelená* 'brown-green'	4	*ciemnofioletowy* 'dark-purple'	3
OTHER	44	OTHER	67	OTHER	22
TOTAL	122	TOTAL	161	TOTAL	128

The results in Table 3 demonstrate that the lexicalization pattern [Adj [o] Adj] is more productive than derivation (i.e., suffixation) in Croatian, thus enabling a diversity of terms. However, the results also lead to the observation that speakers more frequently combine some terms than others. Therefore, *zelena* 'green' appears as the first or the second item in 7 out of 10 most frequently used compounds, whereas *crvena* 'red' appears in only 2 compounds. The figures demonstrate that the token frequency of the terms used is lower than the frequency of derived basic color terms *ljubičasta* 'purple', *narančasta* 'orange', and *ružičasta* 'pink'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

In Czech, the situation is roughly similar to Croatian. The only difference is a relatively smaller difference in type frequency of [Adj [o] Adj] pattern and derivation using *-ov(á)*. Similarly to Croatian, two colors are typically used in these compounds (e.g., *zelenomodrá* 'green-blue'; *hnědozelená* 'brown-green', etc.). In addition, we have also recorded compound terms where an adjective is used to describe brightness such as *tmavozelená* ('dark green') or *tmavomodrá* ('dark blue'). This type of compounds is typical of Polish (see below), whereas in Czech, the preferred pattern to specify the brightness of a colour is [Adv + Adj].

In Polish, the token frequency of [Adj [o] Adj] is comparable to what has been observed in Croatian and Czech. Yet, beyond the difference in terms of the types of compounds we mentioned earlier (i.e. 12.8% (25) types in Polish against 26.4% (43) in Croatian and 20.9% (55) in Czech, cf. Table 1, Section 3), there is a difference between Polish, on the one hand, and Croatian and Czech, on the other hand, regarding the type of terms which compose [Adj [o] Adj] compounds. That is, while in Croatian and Czech such compounds typically involve two color adjectives, as shown in Table 3, in Polish this pattern involves an adjective of brightness in most cases (i.e. light or dark) and an adjective of color, e.g., *ciemnozielony* 'dark green', *jasnozielony* 'light-green'. Examples with two color adjectives are scarce, at least in this dataset, accounting for four occurrences in total (e.g., *brązowozielony* 'brown-green', *fioletowoczerwony* 'purple-red', *pomarańczoworóżowy* 'orange-pink, *żółtozielony* 'yellow-green').

4.3 The lexicalization patterns [Adv + Adj] and [Adj + Adj]

The [Adv + Adj] pattern occurs very often in the Croatian and Czech data, but not in Polish. It has both a high type frequency (30.7% (50) in Croatian and 25.5% (67) in Czech) and a high token frequency (30.2% (507) in Croatian and 28.8% (483) in Czech). It thus seems that the [Adv + Adj] pattern is relatively productive in Croatian and Czech. The ten most frequent terms formed via the pattern [Adv + Adj] in Croatian and Czech are presented in Table 4. The Polish participants did not use the pattern [Adv + Adj] at all, showing a preference for a functionally similar [Adj + Adj] pattern instead (20% (39) types and 8.2% (144) tokens). In the Croatian data, the [Adj + Adj] pattern is not present, but it is found in the Czech data with a relatively high type frequency (19.8% (52)), with a contrastively low token frequency (4.1% (68)). Thus, the [Adj + Adj] pattern is also productive in Czech, but it seems there are no conventionalized terms formed according to this pattern. The ten most frequent terms formed via the pattern [Adj + Adj] in Czech and Polish may be found in Table 5.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. The ten most frequent terms formed via the pattern [Adv + Adj] in Croatian and Czech

Croatian	N	Czech	N
svijetlo plava 'light blue'	97	*tmavě zelená* 'dark green'	70
tamno zelena 'dark blue'	80	*světle modrá* 'light blue'	63
svijetlo zelena 'light green'	58	*světle zelená* 'light green'	57
tamno plava 'dark blue'	47	*tmavě modrá* 'dark blue'	45
svijetlo roza 'light pink'	28	*světle fialová* 'light purple'	37
svijetlo ljubičista 'light purple'	27	*tmavě fialová* 'dark purple'	23
tamno ljubičasta 'dark purple'	26	*světle růžová* 'light pink'	21
tamno roza 'dark pink'	12	*tmavě hnědá* 'dark brown'	18
svijetlo smeđa 'light brown'	10	*bledě modrá* 'pale blue'	14
tamno siva 'dark gray'	9	*tmavě růžová* 'dark pink'	14
OTHER	104	OTHER	121
TOTAL	507	TOTAL	483

Table 5. The ten most frequent terms formed via the pattern [Adj + Adj] in Czech and Polish

Czech	N	Polish	N
pruská modrá 'prussian blue'	3	*ciemny zielony* 'dark green'	32
světlejší modrá 'lighter blue'	3	*jasny zielony* 'light green'	21
světlejší zelená 'lighter green'	3	*jasny niebieski* 'light blue'	17
světlá umělá zelená 'light artificial green'	2	*jasny fioletowy* 'light purple'	11
námořnická modrá 'navy'	2	*ciemny różowy* 'dark pink'	6
světlejší hnědá 'lighter brown'	2	*ciemny niebieski* 'dark blue'	5
světlejší růžová 'lighter pink'	2	*jasny błękitny* 'light light.blue'	4
sytější fialová 'more saturated purple'	2	*ciemny brązowy* 'dark brown'	3
tmavější zelená 'darker green'	2	*ciemny fioletowy* 'dark brown'	3
Trávová zelená 'grassy green'	2	*jasny brązowy* 'light brown'	3
OTHER	45	OTHER	39
TOTAL	68	TOTAL	144

In Croatian and Czech, there are some specificities of the [Adv + Adj] pattern that must be pointed out. First, in the first position, there is always an adverb used as a modifier. Adverbs used in this type of terms refer mostly to brightness (e.g., Cro. *svijetlo* 'light', *tamno* 'dark', *jarko* 'bright', *žarko* 'bright', etc., or Cze. *světle* 'light', *tmavě* 'dark'), but they can also refer to saturation (e.g., Cro. *blijedo* 'pale', *zagasito* 'dull', or Cze. *bledě* 'pale'). Second, the adverbs 'light' (*svijetlo* in Croatian and *světle* in Czech) and 'dark' (*tamno* in Croatian and *tmavě* in Czech) are the most frequently used terms that can be combined with all the basic color terms, except the

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

term *bijela* 'white'. In both languages, the type 'light black' is documented in the data (*svijetlo crna* in Croatian and *světle černá* in Czech).

The second pattern, [Adj + Adj], is used slightly differently in Polish and Czech. In Polish, this pattern consists of an adjective of brightness (i.e., light, dark) and an adjective of color (e.g., *ciemny zielony* 'dark green'). Hence, this pattern is functionally similar to the [Adj [o] Adj] compound found in the Polish dataset presented in Section 4.2 (e.g., *ciemnozielony* 'dark-green'). In Czech, the first adjective was often also an adjective of brightness but in many cases it was used in a comparative form, such as *světlejší* 'lighter' or *tmavší/tmavější* 'darker'. Altogether, there were 24 cases of a comparative form and 17 cases of a positive form of an adjective of brightness. The 27 remaining occurrences were other adjectives mostly specifying the hue of the color (e.g., *trávová zelená* 'grass-like green', *námořnická modrá* 'marine blue').

According to the results presented in this section, it can be claimed that the lexicalization pattern [Adv + Adj] is very frequently used in color naming in Czech and Croatian, thus having a more significant role in the partition of the color spectrum than the [Adj [o] Adj] pattern in these two languages. In both languages (Croatian and Czech), this pattern allows for more creativity and productivity since all of the basic color terms are composite elements of this construction. Moreover, both languages share the feature that the colors *plava/modrá* 'blue' and *zelena/zelená* 'green' occur most frequently as composite elements of this lexicalization pattern. The pattern [Adj + Adj] is used typically with adjectives of brightness in Czech and Polish, but in Czech, it is often used in a comparative form. The productivity of this pattern is rather low in Czech and similar to that of the [Adj [o] Adj] pattern in Polish.

4.4 Nouns in color terms

Among different patterns displayed by the speakers for naming colors are nominal patterns, namely [Adj + N] and [N + Ngen]. Nevertheless, nouns were used only to a small extent. Therefore, we find the pattern and its instantiations as borderline examples of conventionalization because of their very low frequency and productivity. Only in Polish we found a more substantial amount of nouns following the [Adj + N] pattern (9.7% (19) of the types and 6.1% (107) of the tokens). In the Croatian and Czech data, the use of nominal patterns is rather scarce. In particular, in Croatian, we recorded only 0.6% (10) tokens of a specific type [N + Ngen] and no tokens of discrete nouns. In Czech, there were only four tokens of discrete nouns. It thus appears that, in contrast to Polish where the use of nouns is a common way to name colors (cf. Kopecka & Janic, 2016), in Croatian and Czech, their use in color naming is rather unproductive and infrequent. The ten most frequent terms formed using nouns in the three languages are shown in the Table 6.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. The ten most frequent terms formed using nouns in the three languages

Croatian	N	Czech	N	Polish	N
boja breskve 'peach color'	4	*tyrkys* '(the) turquoise'	3	*jasny fiolet* '(the) light purple'	20
boja višnje 'sour-cherry color'	1	*akvamarín* '(the) aquamarine'	1	*ciemna zieleń* '(the) dark green'	19
boja cigle 'brick color'	1	*tmavší akvamarín* '(the) darker aquamarine'	1	*ciemny fiolet* '(the) dark purple'	10
boja vina 'vine color'	1	*tmavější, sytější tyrkys* '(the) darker, deeper turquoise'	1	*jasna zieleń* '(the) light green'	6
–	–	–	–	*ciemny róż* '(the) dark pink'	5
–	–	–	–	*zgniła zieleń* '(the) rotten green'	5
–	–	–	–	*blady róż* '(the) pale pink'	3
–	–	–	–	*ciemny pomarańcz* '(the) dark orange'	3
–	–	–	–	*jasny róż* '(the) light pink'	3
–	–	–	–	*intensywny róż* '(the) intense pink'	2
OTHER	0	OTHER	0	OTHER	20
TOTAL	7	TOTAL	6	TOTAL	96

As Table 6 shows, in Polish, the [Adj + N] pattern consists of a nominal form of the color terms (e.g., *fiolet* '(the) purple', *zieleń* '(the) green') and an adjective modifying that term. The most common adjectives are *jasny* 'light' and *ciemny* 'dark', but we also found other types of modifying adjectives such as *zgniła* 'rotten' (especially for green) and *intensywny* 'intense'.

The lexicalization pattern [N + Ngen] and its instantiations are considered to be liminal examples of the lexicalization processes. This pattern is used in Croatian to lexicalize color terms that are in relation to a certain object, i.e. the first term in this construction is always *boja* 'color' and the second term refers to a specific object as in *boja breskve* 'color of peach'. The speakers of Croatian are aware of this pattern and use it creatively in order to form new expressions in color naming. *Boja marelice* 'apricot color', a unit often used in color naming, is one such expression, although it has not been registered in the EoSS data.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

5. Some concluding remarks

On the basis of the EoSS data analyzed in this study, several lexicalization patterns that are productive in the formation of Croatian, Czech and Polish color terms were identified. They are results of different grammatical processes used in the lexicalization process. However, attention should be paid to two facts: first, these patterns are not found in each language; second, they differ with respect to the degree of conventionalization. Table 7 shows the comparison of the most prominent lexicalization patterns for each language based on their token and type frequency.

Table 7. The most prominent lexicalization patterns in Croatian, Czech, and Polish based on their token and type frequency (at least 3%)

Order	Croatian		Czech		Polish	
	Type freq.	Token freq.	Type freq.	Token freq.	Type freq.	Token freq.
1	Adv + Adj	Adj (simple)	Adv + Adj	Adj (simple)	Adj + Adj	Adj (simple)
2	Adj [o] Adj	Adv + Adj	Adj [o] Adj	Adv + Adj	Adj (derived)	Adj (derived)
3	Adj (simple)	Adj (derived)	Adj + Adj	Adj (derived)	Adj + N	Adj + Adj
4	Adj (derived)	Adj [o] Adj	Adj (derived)	Adj [o] Adj	Adj [o] Adj	Adj [o] Adj
5	–	–	Adj (simple)	Adj + Adj	N	N
6	–	–		–	Adj (simple)	Adj + N

The lexicalization patterns were defined in line with the features of the color terms, i.e. their instantiations. According to Langacker's approach to conventionalization as a matter of degree, lexicalization patterns and their instantiations used in color naming can be considered as notable examples of the conventionalization scale created within a closed domain, i.e. the color spectrum. Based on the analysis of the data, it is evident that the defined lexicalization patterns are arranged on a scale from the most conventionalized to less conventionalized or liminal examples.

According to the data presented in Table 1, it can be concluded that:

1. The lexicalization pattern [stem + suffix] is less frequently used (and thus less productive in the formation of color terms) in Croatian in comparison with Czech and Polish;
2. The lexicalization pattern [Adj [o] Adj] is frequently used in all three languages. However, in Polish, it is less productive in comparison with Croatian and Czech;
3. The lexicalization pattern [Adv + Adj] is frequent in Croatian and Czech, and not found in Polish, whereas the [Adj + Adj] pattern is frequent in Polish (and less frequent in Czech), and not found in Croatian;

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4. Nouns are almost not used for color naming in Croatian (only several instantiations of the type [N + Ngen] were found) and Czech, whereas the Polish speakers tend to use nouns in color naming.

Based on the contrastive analysis of the three genetically related and typologically close languages, we can conclude that there are many similarities as well as differences with respect to grammatical mechanisms used in the formation of color terms. Tables 1 and 7 clearly show variations in the degree of conventionalization of certain lexicalization patterns in the three languages as well as the diversity of certain lexicalization patterns.

To sum up, it can be said that the model of lexicalization patterns enables capturing typological regularities as well as specificities with respect to the grammatical devices languages use to encode certain meanings.

References

Berlin, B., & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley, CA: University of California Press.

Comrie, B., & Corbett, G. G. (Eds.) (1993). *The Slavonic Languages*. London and New York: Routledge.

Conklin, H. C. (1973). Color Categorization: Basic Color Terms: Their Universality and Evolution. Brent Berlin, Paul Kay. *American Anthropologist*, *75*(4), 931–942. https://doi.org/10.1525/aa.1973.75.4.02a00010

Heider, E. R., & Olivier, D. C. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology*, *3*(2), 337–354. https://doi.org/10.1016/0010-0285(72)90011-4

Kay, P., Berlin, B., & Merrifield, W. (1991). Biocultural Implications of Systems of Color Naming. *Journal of Linguistic Anthropology*, *1*(1), 12–25. https://doi.org/10.1525/jlin.1991.1.1.12

Kay, P., Berlin, B., Maffi, L., Merrifield, W. R., & Cook, R. (2009). *The World Color Survey*. Stanford: CSLI Publications.

Kopecka, A., & Janic, K. (2016). Dénomination des couleurs en polonais : pratiques et usages. *Neophilologica*, *28*, 122–146.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar. Theoretical Prerequisites I*. Stanford: Stanford University Press.

Lucy, J. A. (1997). The linguistics of "color." In C. L. Hardin, & M. Luisa (Eds.), *Color categories in thought and language* (pp. 320–346). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519819.015

MacLaury, R. E. (1987). *Color and cognition in Mesoamerica: Construing categories as vantages*. Austin University of Texas Press.

Majid, A., Jordan, F., & Dunn M. (2015). Semantic systems in closely related languages. *Language Sciences*, *49*(1), 1–18. https://doi.org/10.1016/j.langsci.2014.11.002

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Majid, A., & Levinson, S. C. (2007). The language of vision I: Color. In A. Maid (Ed.), *Field manual, vol. 10* (pp. 22–25). Max Planck Institute for Psycholinguistics, Nijmegen. http://fieldmanuals.mpi.nl/volumes/2007/language-of-vision-color/.

Malt, B. C., & Majid, A. (2013). How thought is mapped into words. *WIREs Cognitive Science*, *4*(6), 583–597. https://doi.org/10.1002/wcs.1251

Raffaelli, I. (2009). *Značenje kroz vrijeme: poglavlja iz dijakronijske semantike*. Disput: Zagreb

Raffaelli I. (2013). The model of morphosemantic patterns in the description of lexical architecture. *Linuge e linguaggio*, *1*, 47–72.

Raffaelli, I. (2017). Conventionalized patterns of colour naming in Croatian. In K. Cergol Kovačević, & S. L. Udier (Eds.), *Applied Linguistics Research and Methodology: Proceedings from the 2015 CALS conference* (pp. 171–186). Bern, Bruxelles, Frankfurt am Main: Peter Lang Verlag.

Raffaelli, I., & Kerovec, B. (2017). The concept of 'taste' in formation of Croatian and Turkish lexicon: A contrastive analysis. *Suvremena lingvistika*, *1*, 21–48.

Regier, T., Paul, K., & Ketharpal, N. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(4), 1436–1441. https://doi.org/10.1073/pnas.0610341104

Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, *50*(4), 378–411. https://doi.org/10.1016/j.cogpsych.2004.10.001

Roberson, D., Davies, I. R., & Davidoff, J. (2000). Colour categories are not universal: Replications and new evidence from a Stone-age culture. *Journal of Experimental Psychology: General*, *129*(3), 369–398. https://doi.org/10.1037/0096-3445.129.3.369

Wierzbicka, A. (2005). There are no "color universals" but there are universals of visual semantics. *Anthropological Linguistics*, *47*(2), 217–244.

Žic Fuchs, M., & Tuđman Vuković, N. (2000). Pričati: novi lik u priči o glagolima govorenja. *Filologija*, *35*, 141–150.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Color naming in Africa

Guillaume Segerer and Martine Vanhove
Llacan (UMR 8135, CNRS, INALCO, Université Sorbonne Paris-Cité)

This chapter is the first large-scale typological survey of the lexical means used in African languages to express color-related meanings. It is based on a very large sample, with data from 350 languages, most of which come from the *RefLex* online lexical database. It focuses on language-internal semantic sources, morphosyntactic strategies, and contact-induced terminology used for color naming. After a brief discussion of the issues raised by "basic" color terms, and "polychromatic" color terms, the chapter provides a review of the semantic sources of color terms, the origin of borrowings, colexifications and metaphorical uses of color terms, main patterns of lexicalization, and, briefly, color-related ideophones.

Keywords: color, African languages, lexicalization, semantics, borrowing

1. Introduction

Some of the main contributions of African languages to the longstanding debate concerning "basic" color terms which followed Berlin and Kay's (1969) universal, typological, and evolutionary approach based on elicitation with the Munsell color chart,[1] were (a) to test the method and hierarchy on individual languages or language groups (e.g., Wescott, 1970; Tornay, 1973; van Beek, 1977; Blommaert, 1985; Davies & Corbett, 1997; Tosco, 1999; Zelealem, 2016); (b) question the methods and results, and highlight their limits, eventually challenging the universality of "basic" color terms (e.g., Dimmendaal, 1995; Moñino, 2004). An overview of the critical arguments (to which we fully adhere) against the Berlin & Kay approach is found in Dimmendaal (1995, 2015) and Foley (1997).

The aim of this chapter is both broad and narrow. It is narrow because we limit our investigation to bilingual dictionaries which happen to be easily accessible and

1. For data, works and references developed in this approach, including over ten African languages, see the *World Color Survey* project.

https://doi.org/10.1075/sfsl.78.13seg
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

searchable, mainly thanks to the *RefLex* database (Segerer & Flavier, 2011–2018), a lexical database developed at Llacan and DDL French research units, which contains over 1,140,000 lexical items covering 785 languages from 1,289 different sources, but also broad because we cover a large sample of 350 languages of Africa (see map in Figure 1).[2]

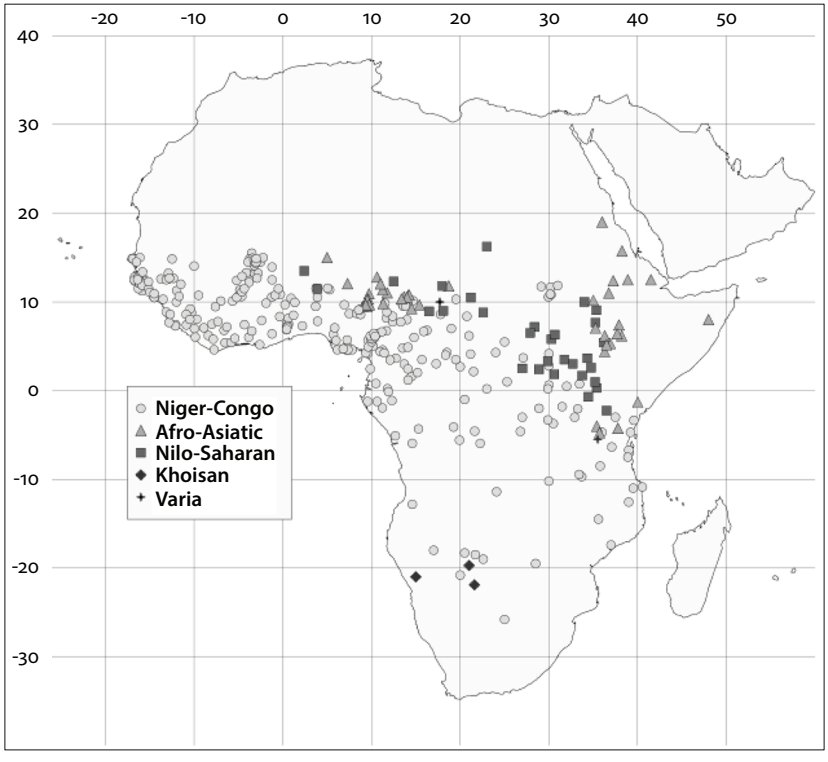


Figure 1. Map of the language sample

In compliance with the objectives of this volume, we focus on semantic sources and morphosyntactic strategies used for color naming and do not address the question of color name inventories (see Section 2). The primary reason is that we are dealing with 350 languages, and although we tried to select the best existing sources, great differences remain in terms of lexicon size and reliability. For instance, in the Joola cluster (a group of closely related Niger-Congo languages spoken in southern Senegal), various sources give from three to ten color names and, remarkably, the

2. Other comparative studies we know of are more limited in scope: Blommaert (1985) deals with 17 Bantu languages (based on dictionaries), and Zelealem (2016), based on a variety of methods, with 16 languages of Ethiopia (Cushitic, Omotic, Semitic and Nilo-Saharan).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

poorest source is one of the biggest dictionaries available for these languages (Kasa, 4,130 entries, Weiss, 1939). Secondly, color name inventories may be misleading in that they do not really reflect the lexical richness of a language. Such inventories tend to be confined to a number of lexically distinguished "primary" colors (in the sense of Berlin & Kay, 1969) in a language. Yet, a language may have few underived terms for "primary" colors but a lot of expressions that distinguish between various shades of these colors: in the biggest Kikongo dictionary (Laman, 1936, approx. 25,000 entries), there are 123 different entries with a definition involving BLACK, 60 with WHITE and 159 with RED. 43 more entries pertain to YELLOW, GREEN, BLUE, GRAY and BROWN. At the other extreme, the Kol lexicon, given in Henson (2007), has 1841 entries, but not a single color term.

In this chapter we first explain our methodology for data collection (Section 2), briefly discuss the issues of the so-called "basic" color terms, and "polychromatic" color terms (also referred to as a "fuzzy set" – Kay et al., 1991 – or a "composite category" in the literature – Foley, 1997) (Section 3). We then provide a review of the various semantic sources of color terms (Section 4), the origin of borrowings (Section 5), the colexification of color terms (Section 6), the different patterns of lexicalization (Section 7). Section 8 briefly discusses the particular case of ideophones and intensifiers used to refer to colors.

2. Data and methodology

The *RefLex* database provided the first set of 197 languages. To get a better coverage of the genetic diversity, we added a second set of 153 languages for which dictionaries, large wordlists in grammars or color-dedicated articles were available to us, or as unpublished data from our own field notes, and data that colleagues in our research unit and other European institutions were kind enough to share with us.[3]

The languages were selected on the basis of three criteria: (a) small word lists were excluded (many African languages are only documented through lists of some 30 items up to a few hundred); (b) sources lacking either BLACK, WHITE or RED were excluded after careful crosschecking with other sources for the same languages, and in all cases it turned out that it was an omission in the data, not a gap in the language; (c) for some languages, various sources were available and comparably reliable, but showed important divergence in color naming. These sources were all

3. Our special thanks go to Odette Ambouroué, Pascal Boyeldieu, Mark Dingemanse, Dmitry Idiatov, Rozenn Guérois, Konstantin Pozdniakov, Nicolas Quint, Serge Sagna, Yvonne Treis, Mark Van de Velde and Alexandra Vydrina. We are also grateful to the two anonymous reviewers for fruitful comments.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

included as we had no way to decide which one to keep[4] (unlike Maffi, 1990, for Somali who could postulate diachronic evolutions and dialectal variation).

The 350-language sample[5] is relatively balanced both geographically and genetically:

- 271 Niger-Congo (NC), approx. 15% of NC languages
- 44 Afro-Asiatic (AA), approx. 13% of AA languages
- 28 Nilo-Saharan (NS), approx. 12% of NS languages
- 3 Khoisan (K), approx. 9% of K languages
- 4 "varia" (2 Songhay, Sandawe, Laal).

We gathered the color tokens in two steps. We first searched for the color terms with the following meanings: BLACK, WHITE, RED, YELLOW, GREEN, BLUE, GRAY, BROWN, ORANGE, PURPLE, and PINK. These correspond to the Berlin and Kay's (1969) eleven "basic" color terms. This choice is purely a matter of convenience for the purposes of comparison, not because we agree with their view. Our aim was not to dwell upon the longstanding debate concerning primary *vs.* secondary "basic" color terms, the associated evolutionary hierarchy, or the "Munsell chart" methodology (e.g., Kay & McDaniel, 1978; Kay et al., 1991; Dimmendaal, 1995), or any of the other issues raised by the definition of "basic" colors as mono-lexemic, non-composite, not restricted to specific classes of objects, or psychologically salient (Berlin & Kay, 1969, p. 6; Hardin & Maffi, 1997; see Dedrick, 1998 for a discussion). To avoid confusion, we refer to the entries which provide translations with these eleven color terms as "BK11 color terms",[6] and it will become clear in the sections below that a good proportion of these terms are not "basic" in the traditional sense.

We tried, whenever possible, to avoid color terms restricted to specific classes of objects, in particular, those specific to cattle. We are of course aware that they play a central role in color denominations in (at least some) African languages (e.g., Dimmendaal, 1995; Last, 1995) and challenge the Berlin and Kay paradigm, but only the biggest dictionaries mention such information, and using it would have introduced too strong a bias. With those sources that make an explicit distinction between "generic" color terms and cattle-specific color terms, we felt free to keep only the former. Of course, when no such distinction was recorded (e.g., Turton

4. There are 21 languages for which we have two or even three sources.

5. All sources and data gathered so far have been made available online at http://guillaumesegerer.fr/Couleurs/. The appendix provides only the references to the languages with their genetic affiliation and their sources cited in this chapter.

6. As suggested by one of the anonymous reviewers.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

et al.'s 2008 Mursi dictionary, contrary to Turton, 1980, cited in Dimmendaal, 1995), we kept all the terms referring to colors. Specific color terms, and especially cattle-specific color terms, would need a dedicated comparative study, far beyond the scope and space limits of this chapter.

We are also aware of the existence of additional terms denoting other patterns of appearance such as stripes, spots, multi-colored aspects, brightness, saturation, etc., but again, for lack of space, we had to leave these terms out of the scope of our study.

In a second step, we searched for color terms that are expressed or intensified with ideophones, an important part of color-related terms in a large number of African languages (e.g., Wescott, 1970; Dingemanse, forthcoming).

Our corpus is biased in some respects. Firstly, for the vast majority of our sources, we have no information about the methodology the linguist used to survey the color terms with the language speakers. It is thus possible that some translations are influenced by the method employed, by the native language of the linguist, or result from typos or omissions. For instance, in the Bété-French dictionary (Zogbo, 2005, approx. 4,500 entries), the entry *pʋpʋ* is translated as 'white', but in the French-Bété index, it is the entry 'yellow' which is *pʋpʋ*. In Nzema (Aboagye, 1992, approx. 2,000 entries), *bɔnyema* 'green' is absent from the dictionary, but it is mentioned in the appendix on color names. In the closely related South Cushitic languages Burunge, Alagwa, and Iraqw, obvious cognate forms are glossed differently: *qancar* is 'green; yellow; blue' in Burunge (Kiessling p.c.), *qantsar* is 'blue, green' in Alagwa and there is no word for YELLOW (Mous p.c.), and *qantsar* is only 'green' in Iraqw which does not have words for YELLOW and BLUE (Mous et al., 2002). Is this due to a real difference between the languages or to the method employed to gather the color vocabulary? Since such cases are marginal, as a rule, we adhered to what we found in the sources.

Another problem is that definitions of color terms tend to be vague: a word for GREEN may have UNRIPE as its primary meaning without the dictionary mentioning it, and thus may not be applicable to any object. As our concern were not "basic" color terms in the traditional sense, we decided to keep these items.

Nevertheless, as will be seen in the following sections, this method proved effective in discovering patterns of lexicalizations for our comparative purposes.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. Monochromatic *vs* polychromatic color terms

3.1 Monochromatic color terms

By monochromatic color terms we refer to the tokens that are translated by only one of the BK11 color terms. Since this is the first large-scale study for African languages, a few statistical facts may be worth presenting here.

Table 1 provides, for each of the BK11 color terms, the number of sources that have at least one expression for this color (ideophones are excluded at this stage). Note that polychromatic terms (see Section 3.2) are counted once for each color, e.g., if one term corresponds to both BLUE and GREEN, it is counted twice.

Table 1. Number of sources per color term in the sample

black, white, red	374	100%
yellow	237	63%
green	259	69%
blue	220	59%
gray	123	33%
brown	108	29%
purple	43	11%
orange	29	8%
pink	29	8%

BLACK, WHITE, and RED are attested in all the sources. The other color terms can be divided into three groups: YELLOW, GREEN and, to a lesser extent, BLUE. These are present in more than half of the sources. GRAY and BROWN come next (in over 25% of the sources). Then come the "marginal" colors, namely PURPLE, ORANGE and PINK.

Larger dictionaries tend to have richer color term inventories. However, the correlation is far from absolute. For example, the small Dhaasanac lexicon (Tosco, 2001; 1,100 entries) distinguishes ten of the BK11 color terms whereas the voluminous Fula dictionary (Seydou, unpublished; 23,200 entries) has only six.

It must be emphasized that Table 1 gives only a very rough account of the situation. First, some languages have rich inventories for various shades of focal colors, which cannot be captured by the above table. In addition, three kinds of terms were not included: those for which the meaning was too vague (such as 'dark' and 'light', without any mention of focal color terms), those with a very specialized meaning, and those with a meaning of visual appearance in which color plays a secondary role (e.g., 'white with black spots'), and which therefore could not be assigned to a focal color (see Section 2). Despite, or because of these restrictions, Table 1 roughly

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

conforms to Berlin and Kay's hierarchy,[7] except that GRAY is far more frequent than it would have been expected (recall that in Berlin & Kay, GRAY goes with the "marginal" terms such as PINK, ORANGE or PURPLE).

3.2 Polychromatic color terms

Table 2 provides an overview of polychromatic color terms, i.e., the color terms translated by several "basic" color terms.

Table 2. Polychromatic color terms in the sample

	Black	White	Red	Yellow	Green	Blue	Gray	Brown	Orange	Purple	Pink	%
Black	1094		3		9	28	1	11		4		5%
White		1044	1	3		1	22	2				3%
Red	3	1	1108	50	1			40	19	3	13	12%
Yellow		3	50	396	12	4	6	23	10		5	29%
Green	9		1	12	399	61	5	4		2		24%
Blue	28	1		4	61	331	12	5		6		35%
Gray	1	22		6	5	12	246	15	3	5	2	31%
Brown	11	2	40	23	4	5	15	248	13	6	5	51%
Orange			19	10			3	13	63	2	12	95%
Purple	4		3		2	6	5	6	2	58	1	50%
Pink			13	5			2	5	12	1	51	75%
Total	56	29	130	113	94	117	71	124	59	29	38	

The figures in the gray-shaded diagonal represent the total number of a color term referring to one or more colors occurring in all sources. For instance, the translation 'black' occurs in 1,094 tokens, representing an average of 3.12 'black' tokens per source. The row "total" corresponds to the addition of all polychromatic terms, whatever the colors involved. The right-hand column gives the percentage of polychromatic terms for each color term as compared to its total number (in the diagonal) of tokens.

The table shows three clear tendencies: (a) the three focal colors, WHITE, BLACK, and RED, in this order, are more likely to be monochromatic in African languages, but RED less so; (b) a middle group of color terms which include GREEN, YELLOW, GRAY, and BLUE, in this order, tend to be significantly polychromatic (on average 25%); (c) the "marginal" color terms, ORANGE, PINK, PURPLE, and also BROWN in

7. In its statistical rather than implicational dimension. It would be easy to give examples of individual languages not conforming to Berlin & Kay's implicational hierarchy, but this is beyond the scope of this paper.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

this case, in this order, are the most likely to be polychromatic (over half to almost all tokens for ORANGE).

Table 2 also shows that color terms are polychromatic in various proportions with other color terms regardless of the total number of tokens per color, and in a different way than the preceding grouping: BROWN can be associated with all the other ten terms, followed by GRAY, YELLOW, RED, PURPLE, GREEN, BLUE, ORANGE, WHITE, BLACK, and PINK.

If one takes into account only the highest number(s) of polychromatic terms within each line in Table 1, another grouping emerges. This grouping is based on whether a color can be classified as "warm" or "cold", or neutral vis-à-vis this classification. In Table 3 the high figures are shaded in gray.

Table 3. Warm and cold colors in the sample

	White	Gray	Red	Yellow	Brown	Orange	Pink	Black	Green	Blue	Purple	
White		22	1	3	2					1		neutral
Gray	22			6	15	3	2	1	5	12	5	neutral
Red	1			50	40	19	13	3	1	1	3	warm
Yellow	3	6	50		23	10	5		12	4		warm
Brown	2	15	40	23		13	5	11	4	5	6	warm
Orange		3	19	10	13		12				2	warm
Pink		2	13	5	5	12					1	warm
Black		1	3		11				9	28	4	cold
Green		5	1	12	4			9		62	2	cold
Blue	1	12	1	4	5			28	62		6	cold
Purple		5	3		6	2	1	4	2	6		neutral

The pairings of polychromatic meanings by category are represented in Figure 2 where the double pointed arrows indicate that for both color terms the polychromy constitutes a majority. A single pointed arrow indicates an asymmetric polychromy: for instance, BROWN has RED as its major polychromy whereas the reverse is not true (RED has YELLOW as its major polychromy). The circles show the two centers of convergence for warm and cold colors.

Figure 2 shows that (a) PURPLE is neutral vis-à-vis the warm/cold categories, and has no favorite pattern of polychromy; (b) WHITE and GRAY are also neutral, and regularly associated; (c) RED is the center of many polychromies for warm colors; (d) BLUE, GREEN and BLACK are more frequently associated with each other, but

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

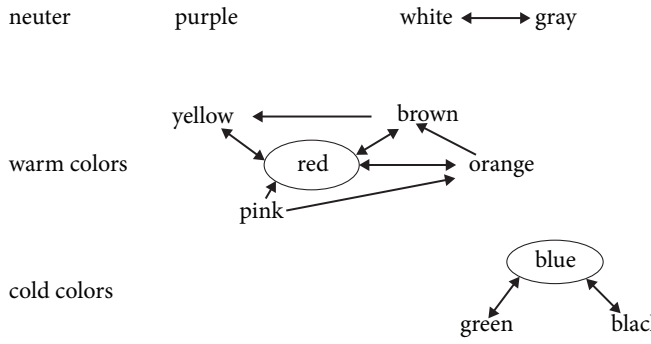


Figure 2. Polychromatic color terms in African languages

not GREEN and BLACK; (e) BLUE is the center for cold colors. Thus, in the majority of cases, polychromatic terms occur within the three categories, not across categories.[8] Some of the polychromatic terms are expected from previous research on the typology of color terms. Such is the case for BLUE and GREEN which are known to be named by the same term in many languages of the world (GRUE was coined to name this conflation). In the African languages of our sample, this represents 12% and 13% of the tokens for GREEN and BLUE, respectively.

Although in a far smaller proportion, polychromatic terms are also well attested for (DARK) BLUE and BLACK, GRAY and WHITE, BROWN and YELLOW, BROWN and GRAY, YELLOW and RED. RED can also cover the spectrum of BROWN, ORANGE, and PINK (see also Zelealem (2016) for Ethiopia).

Some polychromatic terms are surprising from a cross-linguistic viewpoint. Most of them are artifacts of the definitions in the dictionaries, and/or the way the table was established. For instance, the single occurrences of RED and GREEN and RED and BLUE come from Swahili *eusi*, fully translated in the dictionary as 'black, dark, of dark hue, including dark blue, dark green and dark red'. Similarly, GREEN and PURPLE occur in Tswana *tàlá*, which is actually translated as 'green, blue, purple', and in Koegu *mòòlá* 'purple, gray, green'.

The polychromatic terms for WHITE, BROWN (and GRAY), is also an artifact of our database. Thus, in the seven Dogon languages of our sample, the relevant color term is translated as '1. dirty white, off-white, light grey; 2. tan, light brown', while in the eighth language, Landuma, the exact translation of the color term *ferər* is 'whitish, light gray, light brown'.

8. Concerning languages which associate warm and cold colors under one term, namely YELLOW and BLUE, see Foley (1997, p. 157) and Dimmendaal (2015, pp. 130–134)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4. Semantic sources of color terms

Many color terms (of course not only in Africa, see e.g., English *orange*) are identical to or derived from terms that denote real-world objects. This section explores the most common of these metaphorical uses. While we have assumed that color terms, being more abstract, generally originate from nouns referring to concrete objects, it is very likely that in some cases the relationship is reversed, that is, concrete objects are named after their color. Such is the case for Swahili *ahaẓari, ayaẓari* 'green; green leather', a borrowing from Arabic *axḍar* 'green'. Unfortunately, the information provided in the wordlists and dictionaries is not always sufficient to tear apart the source from the target meaning.

Table 4 provides an overview of the various semantic domains of (sometimes putative) sources of the 487 color terms concerned in the languages of our sample, organized by broad semantic domains in Column 1. Column 2 subdivides them further in a number of lower level semantic domains and provides the number of terms concerned. The last line provides the total number of colors. The latter is more numerous than the number of color terms since polychromatic terms are counted for each color they denote.

Table 4. Semantic sources of color terms in the sample

		Black	White	Red	Yellow	Green	Blue	Gray	Brown	Orange	Purple	Pink
Animal	bird (22)	3	1	5	2	5	3		2	1		
	fish (8)	3	1	1			2		1			
	mammal (19)		1	1	1		2	10	4	1		1
	reptiles (5)			1	1	3						
Body	blood (11)			10	1				1	1		1
	other body fluids (6)				5	1			1			
	body parts (5)	1		1	2				1			
	milk (5)		1		4							
Food	egg yolk (3)				3							
	fat (12)				10					2		
	food (11)			1	3	2			4	1		1
Mineral kingdom	clay (29)		7	5	9				7	1	2	3
	dust (6)		1					3	1	1		
	ashes (26)		1					24	1			
	charcoal (2)	2										
	mineral, metal (9)			1	5			2				1

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. (*continued*)

		Black	White	Red	Yellow	Green	Blue	Gray	Brown	Orange	Purple	Pink
Sky	cloud, fog (5)							4	1			
	sky (10)		1				9					
	rainbow (1)						1					
Vegetable kingdom	algae (5)					5	2					
	corn (6)				5	1						
	flower (3)		1		2							
	fruit (20)	2		1	3	1	2		8	2	3	
	grass (29)					28			1			
	kola (7)			3	2				2	2		
	leaf (41)					39	1		2			
	leaf; cassava (9)					9						
	leaf; grass (3)					3						
	locust bean (76)				74				1	1		
	mold (2)					2						
	plant (32)	1	1	6	9	5	2	1	4	1	2	
	saffron (13)				13							
	tree (30)	1		5	11	1	2	1	6	2	3	
Cosmetics	(2)			1	1							
Disease	(4)				4							
Artifacts	(10)		2	2			5				1	
	Total	13	18	44	170	105	31	45	48	16	11	7

In all the languages, the metaphors for BLACK (except in Nzema) and WHITE are the only ones among the lexical means used to express these two focal colors. For the other colors (except RED), metaphors are most often the only lexical means.

The vegetable kingdom is the most frequent source of color terms (258), mostly terrestrial terms (253), more rarely maritime terms ('algae', 5). Next, but far behind, comes the mineral kingdom (72), then the animal kingdom (54), food- and body-related terms on par (26) and (27), and atmospheric elements ('sky', 'cloud', 'fog', 'rainbow'; 16). Names of cosmetics, diseases, and artifacts (fabrics and jewels) are marginal.

Looking from the color terms perspective, YELLOW is clearly the favorite target for the development of metaphorical terms in the languages of Africa, followed by GREEN, then BROWN, GRAY and RED almost on par.

In the following subsections, we review the relationship between semantic sources and color terms in more detail.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.1 Vegetable kingdom

The vegetable kingdom represents the most common source of YELLOW terms with 119 tokens. The most widespread plant name is the 'locust tree' (*Parkia biglobosa*), or rather parts of it, which is an important tree for nutrition and traditional medicine in Western and Central Africa. The terms used for the color denomination designate either the fruit, the yellow pulp, the dry pod, or the flour extracted from it, or in a few instances its leaves. Here are a few examples: Ganja *-dṍhàanθɛ* (< *dɔ̀há* 'powder', *hàanθɛ́* 'locust bean'); Bwamu *dūìnú*; Bambara *nèré*; Tommo So *yùlɔ̀ púnɔ́*; Gbagyi *ōló*.

'Saffron' (or 'turmeric') comes next, but far behind 'locust bean', as e.g., in Afar *húrud*, Sandawe *mādzano*, or Lusoga *kyênvú*. Other plant, fruit and tree names are rarer. 'Corn' is the source of e.g., *makkə* and *mbɔq* in Wolof, and *mbɔx* in Laala; 'ripe banana or plantain' of *nsule* in Akoose; 'country-almond' of *wɔ́lɔ́* in Bambara; '*Combretum glutinosum*' (a bush distributed throughout the Sahel, which produces a yellow dye) of *yeey n / yeey ge* in Sereer; 'lime' of *limɔŋ* in Wolof; '*Lannea acida*' of *sɔɔn* in Wolof. In a few instances, the plant or tree species is not identified, but the lexicographer mentions that the fruit or the bark are yellow, e.g., Yulu *máac(ə̀)*, Krahn *sóèh*.

The vegetable kingdom is also the major source of color terms for GREEN. The majority is related to 'leaf' (39), sometimes more specifically 'cassava leaf' (9), 'grass' (28), or are polysemous between 'leaf' and 'grass' (3): Fonyi *matɔjɛt* 'thing like leaf'; Bambara *bí̱má* 'that contains grass'; Koromfe *fɔ̃frɛ jagalı* 'fresh leaf'; Lega-Beya *lò-bɔ̀bɛ̀là* 'dry leaf of banana tree'; Fulfulde *ɓokko-ɓokko* 'baobab leaf'; Sango *ngùnzā* 'cassava leaf'. A few other plant names have been recruited sporadically, e.g., 'cabbage' in Kambaata (and other Ethiopian languages; see Zelealem, 2016), 'mold' in Gurene and Mankanya, or 'moss' in Jogo. 'Algae' is the source for GREEN in Bedik, Gbagyi, Keeraak, Karon, C'lela (GRUE), and Gurene.

The seven occurrences of the 'kola nut' as the source of color terms, concern only the four warm colors, RED, YELLOW, ORANGE, and BROWN. In two instances, the relevant color terms are polychromatic. It is 'red' in Mofu-gudur, Jula, and Kikongo, while in the latter language another variety of kola nut is used to refer to 'yellow'. 'Kola nut' is used to refer to 'orange' in Maninka, and 'brown' and 'red' in Dii.

Fruit names are also recruited from various species and extended to eight colors, as shown in Table 5.

Other plant, mushroom and tree names are sporadically extended (in one language, sometimes two) to all colors but PINK, such as various types of beans, peas, sorgho, indigo, ficus, cassava, shea tree, henne, and moss. For example, 'indigo tree' gives either 'blue' (Gouro *gālá*, Zarma *sììnì*), or 'purple' (Wolof *nganj*); 'moss' *nìgìnágà* is the source of 'green' in Jogo; '*Sorghum guineense*' *déré* is the source for

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Color terms from fruit names in the sample

Color	Fruit
orange	orange; pumpkin
black	monk's pepper (Vitex sp.); palm nut almond
yellow	lemon; Sodom apple
brown	shea kernel; cowpea; num-num; coffee; calabash fruit
purple	plum; African pear; acacia fruit

one of the terms for 'red' in Bambara; the 'root of *Cochlospermum tinctorium*' (false cotton plant) *kpata'ĭ* is the source of 'orange' in Kyanga.

4.2 Animals

Bird and mammal names are the most common animal sources for color terms. The colors they denote obviously depend on the animal's characteristic color, be it its fur, feathers, or beak.

Various color terms are derived from bird names, mostly warm ones. Ngbaka-Ma'bo alone provides seven of the 22 terms in this semantic category. It is actually the sole language with BLACK, WHITE, and ORANGE terms whose origins are bird names, among other terms for the first two colors. In other languages, bird names are less frequent. For instance, Wolof *cɔy* and Fonyi *cɛɛkʊr* 'green' derive from a type of 'parrot', Hausa *kìliilì* 'green' from the 'greenshank', Dii *bèd yę̀ę̀gná* 'green' from the 'aisle of the roller', Acooli *òcwä̋ä̀k* 'yellow' from a 'yellow weaver bird', Fula (Massina) *wojja* 'be red' from a 'small calao with a red beak', Fang *kula* 'blue turaco; blue', Lomongo *lonkómbé* 'kite; brown', etc.

As for mammals, gazelle, antelope, oryx and duiker names are the most common sources of color terms, often for various shades of GRAY, in Limba, Acooli, Rumanyo, and Maa, and also BROWN and LIGHT RED in Otetela. In Ik, *kodowa*, a type of gazelle, gives a 'yellowish color', and in Ngbaka, *ngbɔ́mò* 'red-flanked duiker' is the source of 'bright red'. In Maa a type of GRAY derives from the 'vervet monkey' *dédɛ̀*, and 'light brown' from the 'serval cat' *e-seperuz*.

Fish names mostly depict various cold or neutral colors: Gubaher *fʊʊn* 'be blue' and Viya *o-ngono* 'black, dark color' derive from the name of the 'wels catfish', Nzema *bile* 'black' from the 'mudfish', Maninka *kàndanfadiláma* 'light brown' from the name of the '*kàndán* fish', and Myene *òsɔ́lɔ́ mɛ́↓ ŋgɛ́lɛ̀* 'blue' derives from a type of shellfish.

Reptile names give either GREEN (in Dii *bɨɨg ha'adná*, a type of 'green snake'), YELLOW (in Bambara *básá* type of 'lizard; crocodile') or BRIGHT RED (in Wolof *xɔnq*, a type of 'garter snake').

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.3 Mineral kingdom

Under this heading, we grouped together terms designating various types of clay, mineral, metal, and dust. As Table 4 shows, 'clay' is the most important one with 29 tokens. 'Clay' is actually a cover term for 'clay' and various types of soil, chalk, ochre, kaolin, and mud.

The color terms for WHITE derived from 'kaolin', 'chalk' or 'white mud' mainly concern Bantu languages (6/7 tokens), Fang, Kikongo (3 terms), Enya and Viya, and one Ubangi language, Ngbaka-Ma'bo.

'Ochre' gives 'red' in Sar, Bambara (as one of the 'red' terms) and Kikongo (3 tokens). In the latter language, one of the tokens is polychromatic with 'yellow' because the term designates, among other things, red soil, reddish clay, and a lead chromate (chrome yellow).

Mineral terms are also the source for YELLOW in one Adamawa, one SBB, two Bantu and three Chadic languages. The color terms are derived from terms meaning either 'soil', 'earth' or 'ochre'. The two 'yellow' terms used in Dii also mean 'brown', and in Otetela the 'yellow' term also means 'pink'.

BROWN is a monochromatic term in Maninka, Wawa, Kenyang and Digo, all NC languages.

PINK is found in Jula, where the 'clay' term means precisely 'pink mud extracted when digging a well', and in Otetela, where it is polychromatic with 'orange'.

The remaining two terms mean 'purple' in two Bantu languages, Enya and Otetela, and they derive respectively from 'mud' and 'loamy soil'.

The remains of burnt vegetation are also important sources of neutral color terms, in particular 'ashes' for GRAY (24, approx. half of the sources for GRAY), and very marginally for WHITE (1) and BROWN (1), and 'charcoal' for BLACK is also very marginal (2).

As for other metals and minerals, 'copper' is recruited twice, in Wolof for 'yellow', in Duala for 'red'; 'rust' once for 'yellow' in Sereer; 'alum' once for 'yellowish' in Bambara; 'stone' twice in Chakali for 'pinkish', in Swahili for 'gray'; and 'rock' once in Kapsiki for 'gray'.

'Dust' is used three times for 'gray' in Digo, Somali, and Ngas ('white-gray'), once for 'white' in Viya (which has other color terms for 'white'), and 'orange' in Bisa.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.4 Body-related terms

'Blood' is unsurprisingly associated with the RED color. This is the case for various languages, from different families of the Niger-Congo phylum[9]: Mande (Maninka, Bambara), Bantu (Rombo, Kikongo, Duala), Igboid (Igbo) and Bak (Karon). In the latter language, it designates more precisely the 'blood of dog', and it is polychromatic with 'orange'. In Kwanyama, 'clot of blood' is not 'red' but 'pink'. In Chumburung, a specific hue of red is designated as 'frog's blood'.

'Milk', or more precisely 'colostrum', also means 'yellow' (Dii; Rumanyo; Acooli, Maa). For the sole instance of 'milk' proper, the color term meaning is 'white' (Maninka).

Other color terms that originate from body fluids are related to 'vomit' (Otetela, Fang 'vomit of bile'; Kapsiki 'vomit of jaundice'), 'children feces' (Mofu-gudur; Karon), or 'children urine' (Mbuko). In all these languages, the referred color is 'yellow'; in Otetela it also means 'brown'. In Iraqw, *qantsar* 'green' may be related to *qantsa* 'green contents of small intestines'.

Five languages derive their RED, YELLOW, BROWN or BLACK color terms from a body part: in Bambara *jàbá* 'female sex' is one of the color terms for 'red'; in Meta' 'breast of dog' gives 'yellow', in Digo 'liver' gives 'brown', in Bongo 'intestine' (associated with fish) gives 'yellow', and in Tene Kan *gɛ́* 'gallbladder' gives 'black'.

4.5 Food

Here, we have grouped terms that designate different types of dishes, egg yolk, and oil. The last category groups together 'oil' and 'fat' extracted from plants, animals, and other unknown sources. The color terms in question are only YELLOW (9) and ORANGE (2). Palm oil is attested in five languages (Sango, Kikongo, Otetela 'yellow'; Ciluba, Idzing 'orange'). In Otetela, *emambamamba* is more precisely 'dregs in palm oil extracted without the fibers and hardly any water'. In three languages, the origin of the oil is not specified: Longando *bosáká* 'oil from which water hasn't been separated as yet', Bassa 'fried oil', and Ijo (Nembe) 'oily'. Three languages have YELLOW terms that come from animal fat: Uduk 'oil of cow', Nzema 'fowl fat', and Buem 'chicken fat'.

In Ngbaka-Ma'bo and Otetela, palm nuts cooked as soups or sauces may be recruited to express color terms, e.g., 'soup of charred palms' > 'red'; 'soup of palm nuts' > 'yellow, orange' (Ngbaka-Ma'bo); 'sauce of palm oil' > 'brown' (Otetela). 'Pea

9. Also in some AA and NS languages of Ethiopia (Zelealem, 2016).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

crushed with palm oil' is recruited for 'yellow' in Duala; 'malt flour' for 'brown' in Buli; a kind of 'leaf sauce' for 'green' in Bedik. 'Chocolate' gives 'brown' and the 'honey of a type of ant[s]' gives 'light brown' in Maninka.

4.6 Atmospheric elements

'Sky' is the most common term in this category that enters in the composition of color terms. It gives, as expected, a BLUE color term in nine languages (e.g., in Swahili, Kambataa, Kapsiki), but in Degema it means 'white'.

'Cloud' and 'fog' usually give 'gray' (Somali, Bambara, Lunda). In Bisa, 'light brown' may derive from 'fog'.

In one instance the BLUE color is related to the word for 'rainbow' (Akoose).

4.7 Other sources

The two tokens labelled 'cosmetics' actually refer to types of paste whose natural origin is not specified, and could well be classified in the category of 'mineral kingdom': 'red' is 'a red paste rubbed on the skin by women in a fattening-room' in Degema, and 'yellow' is the 'name of a yellow substance used in ceremonies' in Bakwe.

In the category 'disease', one finds only YELLOW and the source term is either 'jaundice' (Bambara; Bwamu), 'malaria' (Degema), or 'yellow fever' (Moore).

Material or cloth dyed in either 'blue' (sometimes 'indigo blue') or 'white' are polysemous with the color terms themselves.

In the jewel category, one finds only RED in Ngbaka-Ma'bo (< 'red bead') and in Longando (< 'leg-ring').

Not mentioned in Table 4 is Mbuko *jíyjáy* which is recorded as polysemous with 'brown' and 'light'.

5. Borrowings

145 color terms from 99 languages were identified as (potential) borrowings from other languages, either by the lexicographers or by ourselves.

Table 6 provides the number of borrowed color terms per color, ordered by frequencies.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Borrowings of color terms in the sample

Blue	Yellow	Green	Orange	Black	White	Red	Brown	Purple	Gray	Pink
98	17	14	6	4	4	4	2	2	1	1

Some of the color terms are polychromatic and have been counted once for each color (ORANGE is also YELLOW or GRAY, while GREEN is three times also BLUE). Most of them are loans from French or English (88), followed by Hausa (12), Arabic (9; 2 ultimately from Hindi), Ethio-Semitic (6), Pular (5), Dendi (4), Moore (3), and Fula (2). Lingala, Omotic, Portuguese, Datooga, Jula, Bambara, Swahili, Bantu, and Karimojong provided only one borrowing each, and five terms are given in the sources as probable borrowings from unknown languages.

5.1 Blue

BLUE is the most frequently borrowed color term (98). Most of the borrowings (79) come from either English *blue* or French *bleu*, e.g., Bakwé *plù*, Krim *bùlù*, Ma'di *bùlúù*, Nalu *buluu*, Samo-Maya *bʊ̀la*, Sereer *bule*, Toro Tegu *bulà*, or Wolof *bulɔ*. Sometimes the lexicographer specifies that the loan came via a blue washing powder or a blue dye, e.g., Koromfe *bʊ̀la*. Some of the 'blue' loans actually go back to European trademarks of clothing whiteners: Mankanya *uniilu*, Ma'di *zákārì*. BLUE is borrowed from Hausa in two other Chadic languages (Guus, Zodi) and two Adamawa languages (Peere, Mbum), from Dendi (Songhay) in three Mande languages (Bokobaru, Busa, Kyenga). There are two attestations of borrowing from Arabic (one ultimately from Hindi) in Swahili, and one from Bambara, Karimojong, and Pular in Kagoro, Soo, and Kakabe, respectively.

5.2 Green

GREEN is represented by fourteen borrowings (and three of them actually mean GRUE). Unlike BLUE, the source languages are very diverse (compared to the number of tokens), and European languages form a minority: (a) four come from Hausa into other Chadic or NC languages (Bole; Fulfulde (GRUE); Mbum; possibly Gade); (b) three come from English (Myene, Herero; Sherbro); (c) two from Arabic (Somali; Swahili); (d) two are intra-Gur borrowings from Moore to Koromfe (one is a case of GRUE and appears in a simple and reduplicated form, the other actually comes from a noun meaning 'crushed leaves of baobab'); (e) one is an intra-Mande loan from Jula to Seeku; (f) the last one is a borrowing from French *bleu* with a meaning change to 'green' in Kagoro (BLUE in this language is an intra-Mande borrowing from a Manding variety).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

5.3 Yellow

There are fifteen loans for YELLOW: Central Cushitic languages Bilin, Kemant, and Xamtanga have four borrowings from the neighboring Ethio-Semitic Ge'ez; Somali and Aiki have a borrowing from Arabic *aṣfar*. English, Pular, and Lingala are also sources for this color term; the Lingala term actually means 'refined palm oil' in Sango.

5.4 Other colors

The other borrowed color terms are more sporadically attested in the language sample. There are six for BROWN: three from French *kaki*, the color of military uniforms, one from French *chocolat*, one from English *brown* in Sherbro, one from Amharic *bun* 'coffee; brown' in Gedeo.

There are a few borrowings for WHITE: a putative Omotic borrowing in Awngi, one from Ethio-Semitic in Bilin, one from Arabic in Swahili, a putative one from a Bantu language in Ngbaka Ma'bo; two for RED from Arabic and Hausa in Swahili and C'lela. There are also two borrowings for PURPLE from Arabic and Dendi, and one for PINK from Hausa.

6. Colexification of color terms

In this section we group color terms which are polysemous, but for which we generally had no way to decide which semantic domain was the source, and which was the target of a metaphorical extension, even if in some instances, we had a strong impression that it is the color term that has developed other meanings. For instance, it is probable that the direction of the semantic extension is from the color term to another meaning when the adjectival or noun form denotes only a color, while the derived verb denotes an additional meaning. Even if, in some cases, in the following subsections we hypothesize a direction of the meaning extension, in order to avoid inducing false interpretations, we use a more neutral term, *colexification* coined by François (2008), who distinguishes between *strict colexification* understood as "the capacity, for two senses, to be lexified by the same lexeme in synchrony", and *loose colexification* which also includes diachronic semantic change, derivation, composition, and etymologically related terms (p. 171).

Because of the way the *RefLex* database is organized, we may have overlooked some of the colexifications. Consequently, the figures provided in the following

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

subsections for the less frequent colexifications may have been underestimated, but the overall proportion is probably close to reality.

In what follows, we present the most common patterns of colexification, leaving aside those that occur once or twice, as they could be just a result of coincidental homonymy.

6.1 Colexifications with 'ripe', 'raw', 'wet'

The colexification of color terms with either RIPE or UNRIPE (or RAW or UNCOOKED), is the one most frequently mentioned in our sources, but to different extents for the two semantic domains. There are 83 attestations for RIPE with either RED (62), YELLOW (8), BLACK (8), or WHITE (5), and not even half that many (32 times) for UNRIPE, with GREEN (and once with GRUE). Moreover, in most sources, it is unclear whether RAW really means GREEN, or whether it is an artifact of the French or English translations. There is no doubt when a lexicographer explicitly mentions that the polysemous term refers to the color green or provides an example, but this represents only a minority of cases, e.g., in Dii, Kikongo, Kenga, and Ngbaka. Therefore, the extent of colexification of GREEN and UNRIPE has to be carefully approached. In the translations, such colexification is mentioned for 32 languages in three phyla: 23 languages in ten families of the NC phylum (e.g., Nzema, Mambay, Akan, Swahili), four languages in two families of the AA phylum (Ngizim, Bade, Bole; Bilin), and five languages of the NS phylum (e.g., Yulu, Logo, Soo).

The terms for GREEN are also sometimes translated by 'fresh' (Ganda, Liberian Kpelle, Lunda, Laal), most often in addition to 'unripe' (Sar, Lunda, Akan, Ngizim, Dii) and/or 'wet' (Yao; Ngizim, Dii). Similar doubts as before arise regarding the real semantic status of GREEN. We can be sure for only two languages that it is indeed the color that is denoted by the lexical item. In Dii, *vę́'ę̀d* is translated as: (1) 'uncooked' (peanuts, yam, couscous); (2) 'unripe' (fruit); (3) 'fresh' (grass) = 'green color'. In Giryama, the noun *chanikitsi* is translated as 'greenness', and the lexicographer mentions that the last two syllables *-itsi* mean 'unripe, fresh'. For this particular language, we can be reasonably sure that 'unripe' is probably at the origin of the extension to the green color. But for Ganda, Liberian Kpelle, Lunda, Laal, Yao, Sar, Lunda, Akan and Ngizim, that GREEN really denotes a color remains to be proved. The problem is similar for the other eight putative colexifications of GREEN and WET. Only in Koalib can we be sure that *íklì* 'wet' indeed means 'grue' (Quint, p.c.).

The antonym of WET, DRY, colexifies with RED in only one language, Kikongo: *lōoka* 'be(come) red' is 'be(come) dry' for red leaves, meaning that DRY is a secondary meaning with a very limited combinatorial potential.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

As for RIPE, the rare colexifications with a color term other than RED may depend on the color of the fruit when ripe. The information is not always provided in the sources. It is mentioned only for Kikongo, Cebaara and Day for 'ripe' and 'black', Longando for 'ripe' and 'white', Otetela for 'ripe' and 'yellow'. We may assume, without certainty, that it is also the case for the other colexifications of RIPE with these colors.

For RIPE and RED the authors sometimes mention that the colexification only applies to certain types of fruits (e.g., Banjal, Sar, Wawa, Toro Tegu), or that it is primarily used for RIPE (Siwu), or even that RIPE covers a large spectrum of colors from YELLOW to RED, BROWN and BLACK depending on the fruit (Ngbaka-Ma'bo), so that one may wonder if the term *mōkō* really refers to a color, especially since the language has also other terms for RED.

On the other hand, a few languages have different terms for different colors associated to ripeness: in Fang, there are two terms, *vele* "if the fruits become red or white", and *vine* "if they become black"; in Liberian Kpelle *kpɔlu* (or *kpɔlɔ*) is 'to be red, glowing, yellow, ripe', and *mɔ* (or *mɔŋ*) is 'to be red, ripe, done; to roast'. The sources generally do not mention whether these 'reds' apply to anything else but fruits.

RIPE and RED mainly occur in NC languages (60) from various families, but only in two languages in AA (Bade and Ngas) and one in NS (Jur Modo).

6.2 Colexifications with 'pure, clean' and 'dirty'

There are 54 occurrences of colexifications of WHITE, PURE and/or CLEAN in 39 languages of all phyla: AA (1, Mofu-gudur), K (1, Jul'hoan), NC (31) and NS (5). Almost all languages, except Basaa (which has two terms), Jula (Odienne), Igbo, Nyamwezi, and Susu, have at least one other monosemous lexical entry for WHITE. It is thus possible that in most cases PURE or CLEAN are the origin of secondary terms for WHITE and that those terms could be added to the discussion in Section 4.

The meaning DIRTY is colexified with BLACK 33 times in 28 languages (3 AA, 21 NC, 3 NS), twice with GRAY in two NC languages (Bambara, Jaad), once with DARK BROWN in Kapsiki (Chadic, AA). In one third of these languages, only the verb form is translated as 'be black, be dirty', but the adjective or noun form is only mentioned as meaning 'black' (Attié, Jur Modo, Lumun, Mada, Mundu, Noni, Sango, Southern Samo, Zaar). For most of them, there is another term (verb, adjective or noun) for '(be) dirt(y)' except in Liberian Kpelle and Mmen. Therefore, it may be the case that BLACK is the source of DIRT(Y) in a good number of languages and not vice-versa. Lelemi has only one polysemous verb form 'be black, dirty'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

6.3 Colexifications with terms for personality traits

The colexification of color terms with terms for personality traits occurs in 20 languages, 19 NC and one AA (Afar). The vast majority of cases involve the color WHITE (14) being mainly associated with positive traits, such as kindness, honesty, sincerity, innocence, chastity and sanctity (Afar, Bambara, Buli, Cebaara, Kikongo, Mambay, Minyanka, Nugunu, Nyakyusa, Swahili, Tupuri). In three languages, WHITE is associated with negative traits, such as jealousy in Ngbaka-Ma'bo, and contempt and humiliation in Gbaya-Kara and Nugunu.

Three occurrences of BLACK are colexified with negative traits, such as nastiness (Fonyi), miserliness (Viya) and unsophistication (Bisa). However, in Zimbabwe Ndebele unsophistication is associated with GRUE. Two occurrences of RED are colexified with either positive or negative traits, such as 'being welcoming' (Nalu), and 'being selfish, mean' (Bambara) respectively. The sole occurrence of GRAY is colexified with a negative trait, i.e., 'unlucky' (Fulfulde).

6.4 Colexifications with terms for emotions and affects

The few emotions colexified with color terms are all negative ones. RED colexifies with ANGER in Joola Fogny and six Dogon languages. In Dogon, RED is clearly the source of the semantic extension, since the emotional value is formally derived from the color term by affixes, compounding, or particular syntactic constructions, whereas in the color term no such additional formal means are involved. The formal means are also not involved in the colexification of 'black' with 'pain, affliction' in Beja, with 'discontent' in Jul'hoan and Lunda, with 'sadness' in Somali, and of 'yellow' with 'fear' in Gbaya-Kara.

6.5 Colexifications with terms for evaluative judgments

Evaluative judgments occur in eight languages (3 NC, 5 AA) with BLACK, WHITE, and RED, and may be positive or negative for the same color, depending on the language. In Alagwa and Iraqw, 'good-looking, nice' is associated with 'black' in the former and with 'white' in the latter. But in Idzing, 'black' is 'ugly'. In Beja 'white' is positively associated with 'excellent', while in Hausa, 'black' colexifies with 'bad'. 'Important' or 'abundant' are associated with 'red' in Sura, Bwamu and Bambara (Mande), but with 'black' in another Mande language, Odienne Jula.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

6.6 Colexifications with terms for truth value judgments

Three color terms, WHITE, RED and GREEN are colexified with terms for truth value judgments in seven languages (1 AA, 1 NS, 6 NC). The first two colors are associated with positive judgments: 'certain; white' in Cebaara, Bambara, Mandinka, and Maninka, 'obvious; white' in Afar and Soninke. 'White' is also associated with 'true' in Cebaara, but in Samo-Maya it is the color 'red'. Kanuri is the sole language where a color term is colexified with a term for a negative truth value judgment: 'green' is also 'false'.

6.7 Colexifications with terms pertaining to sensory modalities

Color terms belong to the sense of sight, but they are sporadically polysemous with two other sensory modalities, mostly with the tactile sense and rarely with the gustatory sense. In Bekwel, *zwònzwòn* is 'white; smooth', in Kikongo, *bòbakana* is 'red; really soft', and in Ciluba, *cisengasenga* is 'gray' and 'rough'. In Lomongo, one of the terms for 'green', *bombamba*, is also 'hard' (especially for trees). In other languages, the estimation of temperature mostly colexifies with 'red' for 'warm, hot': Yoruba *kpʠ*, Rundi *gu-tukura*, Eton *və̀*, Fyem *ʃoʃá*. 'Green' is 'cold' only in Akposo (*wlá*). As for the gustatory sense, the sole case is in Bambara where *básí* is 'red; lumpy, bitter, astringent'.[10]

6.8 Colexifications with supernatural, dangerous or remarkable beings

WHITE designates an 'albino' in Bushoong (*ncwɛ̀ɛ̀ky*) and Lomongo (*ɛma*), a 'mythical being' in Viya (*mo-ndɔngɔ*), and a 'character wearing a mask and a raffia loin cloth' in Viya (*mo-nɔmbɛ*). In Bambara, *nsóróbílẹ́* 'brown orange' is a 'dangerous man' or a 'ginger-haired person'.

10. This is an emblematic example of the limitations of dictionary definitions. As an anonymous reviewer pointed out, "the color term is not (or at least not directly) related to the quality verb 'be bitter, be astringent'. Its source is the word *básí* 'blood' (…). Furthermore, as a color term the noun *básí* denotes a 'dirty yellow, brownish color', rather than red."

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

7. Patterns of lexicalization

Dictionaries and wordlists are often far from being specific about word morphology and word categories. For this reason, in what follows we only discuss the cases where the authors provided the necessary grammatical information themselves, or where we could easily deduce it from the sources or from our own knowledge of a language or a language group.

7.1 Noun class alternation

Color terms that come from another semantic domain may be integrated into the color denomination system by a noun class marker shift in languages which have a morphological device, like Atlantic and Bantu languages. For instance, in Nyamwezi *lò-dòótŏ* 'green' belongs to class 11, but *ì-dòótŏ* 'leaf' to class 5. Similarly, in Viya *o-mbulu* 'purple' (noun) is class 11, but *Ø-mbulu* (noun) 'blue turaco' is class 9. In Fonyi, *ɛ-cɛɛkʊr* 'parrot' (class E), shifts to class MA (that of abstract nouns) to denote the 'green' color, *ma-cɛɛkʊr*. In Banjal, the noun *bu-xomboŋ* 'locust bean' produced a color term *e-xomboŋ* 'be yellow' (class E of infinitives). In Rumanyo, *lihengá* 'colostrum', class 5, shifts to class 7 for 'yellow color' *shi-nalihengá*.[11] For another color term in Rumanyo, there is no class shift per se but an extra noun class marker is added to the class marker of the source term: *n-tjéfu* 'eland' (class 9) > *shi-n-tjéfu* 'gray color' (*shi-* = class 7). In some cases, the process of lexicalization to a color term has had an intermediary stage. In Fonyi, the noun *ka-tɔj* 'leaf' yields the verb *-tɔj-ɛt* 'be (like) a leaf = be green', which in turn yields the noun *ma-tɔjɛt* 'green color'.

7.2 Reduplication

In the languages in our sample, reduplication may have several functions. It can be used to derive a color term from a word with a different meaning (7.2.1). The color terms themselves may be reduplicated to express higher (7.2.2) or lower (7.2.3) intensity of the color in question.

11. *na* is of unknown origin. One of the anonymous reviewers suggested that it may be related to 'be with' or 'with' which often has the form *na* in Bantu languages.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

7.2.1 *Derivation of color terms by reduplication*

One of the functions of reduplication, full or partial, with the possible insertion of extra phonemic material, is to mark the meaning change to a color term in a number of languages, e.g., in Bushoong *ʃɛ̀j* 'oil' > *ʃɛ̀j á ʃɛ̀j* 'yellow'; Akoose *mŏl* 'oil' > *mŏlmólɛ* 'yellow'; *vvu* 'ash' > *ki-vuvvu* 'gray'; Bole *bùto* 'ashes' > *butbùto* 'gray, ash-colored', *mocci* 'locust bean' > *momocci* 'yellow'; Kom *mɨghi* 'yellowish type of liquid sodium bicarbonate, drained out of wood ash' > *mɨghi-mɨghi* 'be yellowish'. In Sandawe *!'úpha!'úpha* 'gray' is related to *!'úpha* 'ash', but actually derives from its plural form which is also reduplicated *!'úpha!'úpha* 'ashes'. In Bokobaru *kpà* 'néré, locust bean' > *kpà'ikpa'i* 'bright yellow', an additional meaning is marked by this pattern, i.e., intensification.

7.2.2 *Intensification*

Intensification may be the function of reduplication of color terms. Such is the case in Bade *hēta* 'white' > *hethēta* 'very white'; Akan *tũ̂mm, tũntũm* 'black, dark, dark shades of red, brown, blue, green' > *tuntùuntu(m)* 'very black'; Krahn *ylè* 'black' > *yle yle* 'very black, very dark'; Ngbaka-Ma'bo *pí* 'black' > *pípípí* 'pitch black'. In Toura, only the color terms 'white', 'black' and 'red' can be reduplicated, once or twice, to provide an intensive meaning: *tíí* 'black; any dark color' > *títíí*, *títítíí* 'very black'; *pűű* 'white' > *pűpűű, pűpűpűű* 'very black'; *t̋ɛ̋ɛ̋* 'red' > *t̋ɛ̋t̋ɛ̋ɛ̋, t̋ɛ̋t̋ɛ̋t̋ɛ̋ɛ̋* 'very red'.

7.2.3 *Attenuation*

The most frequent function of reduplication in our language sample is to attenuate the intensity of the color denoted by a color term. From the information provided in the sources, it is not possible to draw a decisive conclusion whether, in a given language, this device is systematic or limited to some color terms, but we have not found a single instance of the former possibility. Some examples are Ma'di *wàrà* 'brown' > *wàrà wàrà* 'brownish'; Ibibio *ŋwèèn* 'become black' > *ŋ̀nwèèn ŋ̀nwèèn* 'blackish, blackened'; Kikongo *lu-mpèmbe* 'whiteness; whitish', *lu-mpèmbe-mpēmbe* 'a little bit of whiteness'; in Bole *ɓûl* 'yellow' can be fully reduplicated once and have the same meaning, *ɓulɓùl* 'yellow', but further partial reduplication results in an attenuative meaning: *ɓuɓulɓùl* 'yellowish'. In Ngwo, two colors are reduplicated: *blā* 'white' > *blāblā* 'whitish', and *fān* 'red' > *fārāfān* 'reddish' (with a change *n* > *r* and insertion of *ā*). In Hausa, more colors can be attenuated by reduplication: *bákíí* 'black; very dark green; very dark blue' > *bákí-bákí* 'blackish'; *shúúdî* 'work of dyeing blue; blue' > *shúúdî-shuudî* 'lightish blue'; *dòòrawà* 'locus bean tree' > *dòòrawà-dòòrawà* 'light-yellow'; *háyáák'í* 'smoke' > *háyáák'í-háyáák'íí* 'brownish, khaki'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

7.3 Compounding and genitive phrases

Compounding and genitive phrases are grouped together as accurate information about these two constructions is not always provided in the sources. For our purposes, it suffices that they both involve at least two elements from two different semantic domains. Some semantic domains are recurrent and could easily be split into subheadings, while others are particularly remarkable and deserve a subheading of their own even though they may not be widespread. The "varia" subsection gathers idiosyncratic sources of color terms.

7.3.1 *Color of*

Genitive phrases or compounds involving the generic term 'color' to create color terms are attested in several languages from different branches of the Niger-Congo phylum. Associated with either 'grass', 'vegetable', 'leaf', or 'manioc leaf', they denote the color GREEN: Cebaara *nyàāpírigē cére* lit. 'fresh grass color', Mba Ndunga-le *ngùnjá ngòmɛ̀* lit. 'color of vegetable', Mba-ne *áɓōkīndá ngòmé* lit. 'color of manioc leaf', Eton *ɲǒl í kpàm* 'color of fresh manioc leaves', Ngwo *ēzɔ́ɔ́rɔ́ ēgɛ́* lit. 'color pasture'. In Bakwé *nápōjījārɛ̄ɲū̄* 'green' is based on *nápō* 'leaf' + *ɲū̄* 'color' and an unidentified element. In Degema *adʊ́ ábɪtaŋ wɔwɔ* 'green' is tentatively proposed by the author as a contraction of *adʊ́ ʊbí ótaŋ wɔwɔ,* lit. 'color leaf tree fresh'.

Other semantic domains as the source of color terms of this type are more sporadically attested: 'color of sky' gives 'blue' in Ngwo *ezɔ́ɔ́rɔ́dí ezob* (classified as an adjective), lit. 'color sky'; in Eton *ɲǒl í dǒb* is lit. 'color of sky'. 'Color of charcoal' *dùgúwɔ̄ʔɔ̀ cére*, categorized as a compound noun, is the origin of one of the two terms for 'black' in Cebaara; and Cebaara is the sole language in our sample which explicitly mentions 'color' as an element of a compound with the widespread 'locust bean flour' for 'yellow', *nánmiigē cére*. 'Yolk color' occurs once for 'yellow' in Ngwo *ēzɔ́ɔ́rɔ́ étáá ākɔ̂n kʷāʔ* (*ēzɔ́ɔ́rɔ́* 'color', *étáá* 'heart', *ākɔ̂n* 'egg', *kʷāʔ* 'chicken').[12]

In some languages, color terms contain the generic term for 'color', but there is no independent entry for the remaining element; it is thus impossible to decide whether they are real compounds with another existing element, or not: Mba Ndunga-Le *mólàngà ngòmɛ̀* 'yellow'; Limbum *nʃòö kàkár* 'green' (*nʃòö* 'color'), Obolo *ùŋʷúnùŋʷán* 'yellow' (*ùŋʷán* 'color'), Ngwo *ēzɔ́ɔ́rɔ́ mbārāŋʷɔ̀* 'green', and *ēzɔ́ɔ́rɔ́ cʷâfàn* 'brown'.

12. Zelealem (2016, p. 66) mentions that *kekino* 'yellow' is derived from *keko* 'fried yolk of egg'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

7.3.2 *Water of*

Compounding with a noun phrase meaning 'water (of)' is a frequent device in Chadic languages. Zaar has *ʒà-k rāās* 'yellow' (lit. 'water of locust bean tree'), *ʒà-k pààtsə́* 'green' (lit. 'water of leaf'), *ʒà-k gàlúúrā* 'blue' (lit. 'water of a plant with a tuberous rhizome'), *ʒà-k nāállē* 'dark brown' (lit. 'water of henna'). In other Chadic languages, only one color term of this sort is mentioned: Zodi *ʃà bàkáŋ* 'green', Guus *ʒà kə̀ wàldə́* 'green' (lit. 'water of leaves'); Mofu-gudur *yám ŋgá gwáraw* 'orange' (lit. 'kola water/juice').

This strategy is sporadically attested in NC languages. Lingala has two color terms formed that way: *mái ma mpɔndú* 'green' (lit. 'water of manioc leaves'), *mái ma lilála* 'orange' (lit. 'water of orange'). In Fon, *fé* is a blue powder used to dye loincloth and to brighten washing and *fé sìn* 'blue' is lit. 'blue powder water'. In Dendi, *sùbù háɽì* 'green' comes from *sùbù* 'bush; grass' and *hàɽí* 'water', and in Adele, *èkɔ̀nlé ù bùtù* 'yellow' is lit. 'tree sp. his water'.

7.3.3 *Plants*

Color terms coming from the vegetable kingdom are often associated with another term to denote the color meaning.

A common association is 'fresh grass' for 'green' in NC languages (see Section 5.1): Nanga-dama *sàwà ɔ̀rí* lit. 'fresh grass'; Mbum (Adamawa) *hòì fókó* lit. 'grass fresh'. In Mokpwe, the verb for 'be green' is actually a copulative construction where 'grass' is the subject and 'fresh' the attributive predicate: *βèwùlé βè ɸéɲá* lit. 'grass is fresh'. 'Leaf' can also be associated with 'fresh' as in Meta' *i̋fù zwî* 'green', lit. 'fresh leaf'.

In Mandinka, *jàmbàkérè* '(be) green' is the result of the compounding of *jàmbà* 'leaf' with *kérè* 'raw', and in Rundi, *akătsi ga-tŏto* 'green' of 'pulp of stalk' with 'leaf'.

For YELLOW we already mentioned 'locust bean flour' in Section 5.1. Here is another example from Mandinka where *nètèmúnkù* '(be) yellow' is made of *nètè* 'locust bean' and *múnkù* 'powder, flour'.

In one language, Fang, the compound for 'chocolate colored, garnet colored', *ésʋ̃isʋ̃i-akona*, associates two plant-related nouns: *ésʋ̃isʋ̃i* 'heart of the raffia bamboo', and *akona* 'fruit of *ngon* (a type of gourd) which has a chocolate color when ripe'.

7.3.4 *Body of*

Compounding with a term meaning 'body' is sporadically attested in several branches of NC. In Jamsay, *yàrù-mànà jèjú* 'blue' is lit. 'sky body'; in Bakwé, *bàbē-fūgbā* 'brown, bronze color' is lit. 'bronze-body'; in Mandinka, *séé-bálà* 'gray' is lit. 'shea-body' and *pòtò-bálà* 'grayish' is lit. 'mud/clay-body'. It should be noted that in none of the sources for these languages a polysemous term 'body; color' exists, a polysemy which is attested in other languages, e.g., Cebaara, Nzadi, or San.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

7.3.5 *Body fluids*

In some Chadic and Bak languages, the terms for 'feces' are associated with terms referring to animals or infants. 'Yellow' in Mofu-gudur *záy-bəzey* is lit. 'feces of a child', and in Joola Karon, *wat añi* is 'feces-child'. Zelealem (2016) also mentions that the same holds true for the languages of Ethiopia.

7.3.6 *Body parts*

Few languages have body parts of animals denoting a color. In two cases the related color term is YELLOW: Meta' *ì̀bən ɨ bu* is lit. 'breast of dog' and Bongo *tì̀kɨ́ gòrò* is an intestine of *gòrò* 'a type of fish with a yellow intestine'. In Mofu-gudur, the related color term is LIGHT BLUE *lakwáy-duwak* 'testicles-monkey' (also in Cushitic Kulisi; see Zelealem, 2016, p. 64).

7.3.7 *Oil, fat*

OIL or FAT can be used in compounds to refer to the color YELLOW: Bassa *ɖāā-cànà* lit. 'fried oil', Nzema *akɔlɛ ɛnlome* lit. 'fowl fat'.

7.3.8 *Combination of two color terms*

Beja is the sole language in our sample that associates two color terms to give the third one. *Hadhamiʃ* 'blackish-grey', made of *had* < *hadal* 'black' and *hamiʃ* 'ashy, dirty-white (of camels)' comes as no surprise, but the combination of *ádar* 'red' and *sóota* 'green' to give *ádarsóotay* 'yellow' is more remarkable as it corresponds to the physiological process of color perception in the eye (called "additive synthesis").

7.3.9 *Varia*

A few other languages have very idiosyncratic compounds. In Akoose, BLACK occurs in genitive phrases to express other colors: *ehínde é byaá* 'green' is lit. 'black, dark of leaf', *ehínde é ngubélɛn* 'blue' is lit. 'black of rainbow'. In Bassa, *hʷúē-jèɖè* 'green' is 'mildew-monkey'. In Mbum, *blú násárá* 'blue' is lit. 'blue European'. In Sar, *ɔ̀gā-dày-déẁ* 'red color' is lit. 'prevent the Day people to hide' which is a reference to a 'war between the Sar-s and the Day-s who were killed in a marigot where they were trying to hide'.

7.4 Similative adpositions

Chadic and Ubangi languages make use of constructions with an explicit similative adposition. In Mofu-gudur, *yám ŋgá gwáraw* 'red', categorized as a compound noun, is lit. 'like kola water/juice'. In Mbuko, *kawa azay sa wan* 'yellow' is lit. 'like the feces of a child'; in Mofu-gudur, *ma məkaday sa wan* 'yellow' is lit. 'like the urine of a child', and *mba mburom* 'blue' is lit. 'like the sky'. In Ngbaka-Ma'bo, *ɓàá.kpáā.*

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

náā 'green' is 'like a tree leaf', *ɓàá.kpáā.bɔ́* 'yellowish green' is 'like a banana leaf', and *ɓàá.kpákpáā* 'green' is 'like foliage'.

7.5 Affixation

Affixation is not a common device for color naming in our language sample. It is only attested in three languages. Interestingly, in Kambaata it is the highly productive similative suffix *-gud* (not an adposition as in the languages mentioned in Section 7.4) which is used to derive color terms from other semantic domains: *hamiila-gud-* 'green' (lit. cabbage-like), *buna-gud-* '(dark) brown' (lit. coffee-like), *bonxa-gud-* 'green' (lit. leaf-like), *samaa-gud-* 'blue' (lit. sky-like) (the process is also known in other Ethiopian languages; see Zelealem, 2016).

Kakabe (as well as other Greater Manding languages) has a special suffix *-lama* (*ma* is originally a similative device)[13] which occurs only with three color adjectives derived from other semantic domains. These three color terms are *yànba-lama* 'green' < *yànba* 'leaf', *fíta-kɛndɛ-lama* 'green' < *fíta* 'leaf' + *kɛ́ndɛ* 'unripe', and *nétɛ-lama* 'yellow' < *nètɛ* 'locust bean'.

7.6 Diminutives

In our language sample, four languages provide evidence that diminutives may be used as attenuatives of focal color terms. In Beja, there are two color terms: *eera* 'white' > *eela* 'a little white', and *adar* 'red' > *adal* 'pink, light red, light brown'; in Tswana, there is only one: *-híbídú* 'red' > *híbítsʷánà* 'pink'. Kikongo has more. Excluding the dialectal phonetic variants, Laman's (1936) dictionary provides four color terms with the diminutive suffix *-(l)uluka*: *lòmba* 'black' > *lòmbuluka* 'be a bit dark, black; be blackish', *bènga* 'ginger, light' > *bènguluka* 'be reddish, yellow, etc.', *bwàka* 'be(come) red, yellow' > *bwàkuluka* 'be(come) reddish, brown', *nánzò* 'red' > *nanzóluluka* 'light red'. In Kom, a diminutive suffix is attested with one color verb, *bàŋ* 'be red', as a way to attenuate the intensity of the color: *bàŋlɨ̀* 'be reddish'.

13. We thank one of the anonymous reviewers for this information.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

8. Ideophones

A working definition of ideophones can be that of Ameka (2001): "[I]deophones are a phonosemantic class of words with expressive and imaginistic semantics" (p. 26). This definition makes no assumptions about morphology or syntax. We could add that, in general, ideophones have a limited combinability, with many of them attached to a specific verb. The linguistic features that define ideophones in various languages may vary considerably. However, similar semantic properties of ideophones can be found in a number of languages. This is especially true in the domain of colors.

One problem we had to face when sorting the data, was that this word type may not be labeled as "ideophone" in our sources, and we had to search for other equivalents such as "intensifiers" or "attenuatives", or decide by ourselves that a color term was indeed a member of this category when no word category was indicated. Furthermore, in some instances, other language specialists may disagree with our analysis or terminology, and consider that the terms belong yet to another specific class, different from the more common word classes such as nouns, verbs, adjectives, or adverbs (see e.g., Roulon-Doko, this volume, for Gbaya-Bodoe and her use of 'adjective-adverbs'). Within the scope of this chapter, we counted as ideophones (a) those labeled as ideophones (or an equivalent term such as 'adverbe expressif' in Bambara, for instance) by the authors; (b) any modifier which combines only with a specific color term, generally with an intensifying or attenuative meaning. In some sources, however, the information provided does not allow the use of any of the above criteria. In those cases, we arbitrarily counted as ideophones lexical items that looked phonologically atypical (for instance, reduplicated forms as Bini *ɾèɣèɾèɣè* 'pink, light red') or semantically specific (for instance, Longando *tolí* 'black, very black').

Among the 350 languages in our sample, at least 106 have dedicated ideophones in the domain of color, which makes a corpus of nearly 1,000 ideophones (counting only ideophones for focal colors, i.e., not 'multicolored', 'spotted', and the like).

As Table 7 shows, the color ideophones are not evenly distributed color-wise.

Table 7. Proportion of ideophones per color in the sample

Color	Black	White	Red	Green	Yellow	Blue	Gray	Other
Nb of id.	314	283	324	26	11	7	15	6

The three "basic" colors, BLACK, WHITE and RED, concentrate nearly all the ideophones. This might be due to the fact that these color terms are typically verbs, and that ideophones typically modify verbs. Interestingly, each one of the 106 languages

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

which have color ideophones has, on average, 3 different ones for each of the three BLACK, WHITE, and RED colors. In very large dictionaries, however, we can find many more ideophones than in smaller ones. For example, the Bambara dictionary by Dumestre (2011) has nine different ideophones for BLACK, sixteen for WHITE and fourteen for RED, but none for the other colors.

The vast majority of color ideophones are intensifiers, i.e., they are used to emphasize the hue of a color. This can be achieved in various ways:

– The ideophone can be used with only one color term adding emphasis:

  Keeraak (Segerer, pers. fieldnotes)

  (1) *kajılɔɔnak kɔɔkʊ kɔɔmɔ kʊhıtʊm far*
  rooster this COPULA white IDEO
  'This rooster is really white!'

  Maninka (Creissels, 2013, p. 111)

  (2) *sáŋ-o be fín-diŋ kílő*
  Sky COPULA be black IDEO
  'The sky is very dark.'

– The ideophone may be a predicate:

  Dii (Bohnhoff, 2014, p. 264)
  zǫ́ḿ ạn mbàà vìḿ
  empty land burn COPULA IDEO
  'Burnt grass is black'

– The ideophone may also modify a stative verb whose primary meaning is one of low intensity only, not including color, as in Tima (Schneider-Blum, 2013)

  -héh (classified as an adjective by the author) 'light':

  àhéh kɨ̀rɨ̀ndì 'light yellow (like millet)'
  àhéh wèlwèl 'light blue'
  àhéh kʌ́yì 'grass green'

Much more could be said about ideophones in African languages, but this would go far beyond the scope and limits of this chapter (see Segerer and Vanhove, under revision, for further details).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

9. Conclusion

Our study is analogous to big data studies where "noise" (biases) is compensated by quantity. We have surveyed 374 lexical sources corresponding to 350 African languages, looking for the strategies used for color naming. For a number of reasons pertaining to the reliability of the data (the exhaustiveness of inventories, the disparity between the number of entries in sources, possible omissions, the absence of an indication of methodology, etc.), the question of color inventories has not been really addressed here. Instead, our research focused on the strategies African languages use to name colors. Metaphorical uses of nouns, borrowings and colexifications were given special attention, as well as patterns of lexicalization. Derivatives and ideophones were only briefly treated and deserve a much more detailed investigation in future research. One of the main conclusion of this very large survey is that there is not much that can be considered specifically African as far as strategies for color naming are concerned. Still, a few local patterns emerge, such as the naming of YELLOW after the tree *Parkia biglobosa* (locust tree) attested throughout West Africa, the massive borrowing of BLUE from European languages, the (not as frequent) cases of borrowings from neighboring languages of GREEN, and the tendency to get polychromatic terms within but not across the warm/cold distribution of colors. Apart from a few idiosyncratic cases, patterns of lexicalization, metaphors, and colexifications of color terms do not deviate from general cross-linguistic tendencies and the borrowing of color terms is not particularly frequent (except for BLUE). Moreover, although some patterns of lexicalization may be more specific to certain language groups (e.g., 'water of' in Chadic or the colexification of 'ripe' and 'red' in Niger-Congo), none is particular to any of them and they are attested (sporadically) in other language groups for which areal patterns, due to language contact, could not easily be deduced and should be further researched. A textual analysis would also be an interesting approach for future research, but this would not be an easy task for a large-scale study. One solution would be to examine texts in languages for which big corpora are available (e.g., Bambara) and compare the results with those extracted from the dictionaries.

References

(except data sources, see Appendix)

Ameka, F. K. (2001). Ideophones and the nature of the adjective word class in Ewe. In E. F. K. Voeltz, & C. Kilian-Hatz (Eds.), *Ideophones* (pp. 25–48). Amsterdam/Philadelphia: Benjamins. https://doi.org/10.1075/tsl.44.04ame

Berlin, B., & Kay, P. (1969). *Basic color terms: their universality and evolution*. Berkeley and Los Angeles: University of California Press. (First Paperback Edition: 1991).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Blommaert, Jan M. E. (1985). The semantics of Bantu colour terms. *Grazer Linguistische Studien*, 24, 63–75.

Davies, I., & Corbett, G. (1997). Colour categories in African languages: A test of the Berlin & Kay's theory of colour universals. In R. K. Herbert (Ed.), *African Linguistics at the Crossroads: Papers from Kwaluseni* (pp. 581–598). Köln: Rüdiger Köppe.

Dedrick, D. (1998). *Naming the rainbow*. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-017-2382-4

Dimmendaal, G. J. (1995). Studying lexical-semantic fields in languages: nature versus nurture, or where does culture come in these days. *Frankfurter Afrikanistische Blätter*, 7, 1–28.

Dimmendaal, G. J. (2015). *The Leopards Spots. Essays on Language, Cognition and Culture*. Leiden/Boston: Bril.

Foley, W. A. (1997). *Anthropological Linguistics: An Introduction*. Oxford: Blackwell.

François, A. (2008). Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. In M. Vanhove (Ed.), *From polysemy to semantic change* (pp. 163–215). Amsterdam/Philadelphia: Benjamins. https://doi.org/10.1075/slcs.106.09fra

Hardin, C. L., & Maffi, L. (1997). Introduction. In C. L. Hardin, & L. Maffi (Eds.), *Color categories in thought and language* (pp. 1–20). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519819.001

Kay, P., Berlin B., & Merrifield, W. R. (1991). Biocultural implications of systems of color naming. *Journal of Linguistic Anthropology*, 1, 12–25. https://doi.org/10.1525/jlin.1991.1.1.12

Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meanings of basic color terms. *Language*, 54, 610–646. https://doi.org/10.1353/lan.1978.0035

Last, M. (1995). *A Sketch of Chai*. M.A. thesis, *Leiden University*.

Maffi, L. (1990). Somali color term evolution: Grammatical and semantic evidence. *Anthropological Linguistics*, 32 (3/4), 316–334.

Moñino, Y. (2004). Une autre conception des lumières: Sur les noms des couleurs en gbaya. In E. Motte-Florac, & G. Guarisma (Eds.), *Du terrain au cognitif: Linguistique, ethnolinguistique, ethnosciences. A Jacqueline M. C. Thomas*, (pp. 241–265). Leuven/Paris: Peeters.

Segerer, G., & Flavier, S. (2011–2018). *RefLex: Reference Lexicon of Africa*, Version 1.1. Paris, Lyon. http://reflex.cnrs.fr/.

Segerer, G., & Vanhove, M. (under revision). Areal patterns of color naming in the languages of Africa. Linguistic Typology.

Tornay, S. (1973). Langage et perception: La dénomination des couleurs chez les Nyangatom du sud-ouest éthiopien. *L'Homme*, 13(4), 66–94.

Tosco, M. (1999). The Color Terms of Dhaasanac. In M. Lamberti, & L. Tonelli (Eds.), *Afroasiatica Tergestina. Papers from the 9th Italian Meeting of Afroasiatic (Hamito – Semitic) Linguistics* (pp. 381–392). Padova: Unipress.

Turton, D. (1980). There's no such beast: cattle and colour naming among the Mursi. *Man*, 152, 320–338. https://doi.org/10.2307/2801674

van Beek, W. E. A. (1977). Color terms in Kapsiki. In P. Newman, & R. M. Newman (Eds.), *Papers in Chadic Linguistics* (pp. 13–20). Leiden: Afrika-Studie centrum.

Wescott, R. W. (1970). Bini Color Terms. *Anthropological Linguistics*, 12(9), 349–360.

World Color Survey. http://www1.icsi.berkeley.edu/wcs/. Accessed 02 April 2018.

Zelealem, L. (2016). The Description of Colour Terms in Ethiopian Languages. *Eastern Africa Social Science Research Review,* 32 (2), 53–86. https://doi.org/10.1353/eas.2016.0008

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Appendix. Language list and source references

AA: Afroasiatic; KS: Khoi-San; NC: Niger-Congo; NS: Nilo-Saharan, VARIA: other.

Acooli (NS, Eastern Sudanic: Nilotic)
Crazzolara, J. P. (1955). *A study of the Acooli language: grammar and vocabulary*. London, New York: Oxford University Press for the International African Institute (IAI).
Adele (NC, Kwa: Nyo)
Rongier, J. (1998). *Dictionnaire français–adele*. Abidjan: Institut de Linguistique Appliquée.
Afar (AA, Cushitic: Eastern)
Parker, E. M., & Hayward, R. J. (1985). *An Afar-English-French dictionary (with grammatical notes in English)*. London: SOAS. In RefLex.
Aiki (NS, Maban: Mabang)
Nougayrol, P. (1989). *La langue des Aiki dits Rounga (Tchad et République Centrafricaine): Esquisse descriptive et lexique*. Paris: Geuthner.
Akan (Asante) (NC, Kwa: Nyo)
Christaller, J. G. (1933). *Dictionary of the Asante and Fante language called Tshi (Twi)*. Basel: Evangelical Missionary Society.
Akoose (NC, Benue-Congo: Bantu A)
Hedinger, R. (Comp.). (2012). *Akoose-English Dictionary*. Yaounde: Akoose Language Committee and SIL Cameroon.
Akposo (NC, Kwa: Kposo-Ahlo-Bowili)
Soubrier, A. (2013). *Description de l'ikposso uwi*. Lyon: Université Lumière Lyon 2. In RefLex.
Angas (Ngas) (AA, Chadic: Western)
Jungraithmayr, H., & Holubová, M. (in collaboration with Luka J. Jiwul and Sonja Ermisch, 2016). *The Ngas language - Shik Ngas (Northern Nigeria): Fundamentals of Grammar – Texts – Dictionary*. Berlin: Dietrich Reimer.
Attié (NC, Kwa: Nyo)
N'Guessan, J. K. (1996). *Description systématique de l'attie de Memni (langue Kwa de Côte d'Ivoire)*. Grenoble: Université Stendhal (Grenoble 3). In RefLex.
Awing (NC, Benue-Congo: Bantoid)
Alomofor, C. (Comp.) (2007). *Awing–English dictionary and English–Awing index*. Yaoundé: CABTAL (Cameroon Association for Bible Translation and Literacy).
Awngi (AA, Cushitic: Central)
Appleyard, D. L. (2006). *A Comparative Dictionary of the Agaw Languages*. Köln: Rüdiger Köppe.
Bade (AA, Chadic: Western)
Dagona, B. W. (2009). *Bade-English–Hausa Dictionary (Western Dialect)*. Potiskum, Yobe State: Ajami Press.
Bakwé (NC, Kru: Eastern)
Leidenfrost, C. T., & CTAB. (2008). *Bakvɛwaklüüa–Pɔɔkv (Dictionnaire Bakwé)*. Méagui (Côte d'Ivoire). In RefLex.
Bambara (NC, Mande: Western)
Dumestre, G. (2011). *Dictionnaire bambara–français – suivi d'un index abrégé français-bambara*. Paris: Karthala. In RefLex.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Banjal (NC, Atlantic-Bak: Joola)
Bassène, A.-C. (2006). *Description du jóola Banjal (Sénégal)*. Lyon: Université Lumière (Lyon 2). In RefLex.
Basaa (NC, Benue-Congo: Bantu A)
Lemb, P., & Gastines, F. D. (1973). *Dictionnaire basaá–français*. Douala: Collège Libermann.
Bassa (NC, Kru: Western)
Slager, D., Payne, S., Glaygbo, R., Meece, T., & Boen, W. (n.d.). *Bassa Dictionary*. In RefLex. Also online at http://cefliberia.org/?q=bassa.htm (Accessed December, 2017).
Bedik (NC, Atlantic-North: Tenda)
Ferry, M.-P. (1991). *Thesaurus tenda: dictionnaire ethnolinguistique de langues sénégalo-guinnéennes (bassari, bedik, konyagi)* (3 vols). Paris: SELAF. In RefLex.
Beja (AA, Cushitic: Northern)
Roper, E. M. (1928). *Tu Beḍawiɛ: an elementary handbook for the use of Sudan government officials*. Hertford UK: Stephen Austin. In RefLex.
Bekwel (NC, Benue-Congo: Bantu A)
Cheucle, M. (2012). *Bekwel data (personal fieldwork)*. In RefLex.
Bété-Daloa (NC, Kru: Eastern)
Zogbo, G. R. (2005). *Dictionnaire bété–français*. Abidjan: CERAP
Bilin (AA, Cushitic: Central)
See Awngi
Bisa (NC, Mande: Western)
Vanhoudt, B. (1999). Lexique bisa–français, suivi d'un index français–bisa. *Mandenkan* 34: 1–113.
Bokobaru (NC, Mande: Eastern)
Jones, R. (2004). *Bokobaru dictionary*. München: Lincom Europa. In RefLex.
Bole (AA, Chadic: Western)
Gimba, A. M., Baba Ali, M., & Bah, M. (2009). *Bole–English–Hausa Dictionary* (Second Edition). Potiskum, Yobe State: Ajami Press.
Bongo (NS, Central Sudanic: SBB)
Nougayrol, P. (2010). *Bongo data (personal fieldwork)*. In RefLex.
Buem (NC, Kwa: Nyo)
Allan, E. J. (1973). *A grammar of Buem, the Lelemi language*. London: SOAS (PhD thesis).
Buli (NC, Gur: Oti-Volta)
Kröger, F. (1992). *Buli–English dictionary*. Münster, Hamburg: Lit.
Busa (NC, Mande: Eastern)
McCallum Jones, R. (2017a). *Illo–Busa Dictionary*. München: Lincom Europa.
Bushoong (NC, Benue-Congo: Bantu C)
Vansina, J. (1959). *Esquisse de grammaire bushong*. Tervuren: Musée Royal du Congo Belge (MRCB). In RefLex.
Bwamu (NC, Gur: Bwamu)
Bloemarts, M., & de Rasilly, P. (2012). *Dictionnaire boore*. Bamako. In RefLex.
C'lela (NC, Benue-Congo: Kainji)
SIL Nigeria (D. J. Rowbory). (2001). *C'Lela dictionary*. http://blog.rowbory.co.uk/linguistics/ (Accessed December, 2017).
Cebaara (NC, Gur: Senufo)
Mills, R. (2003). *Dictionnaire sénoufo–français (sénanri, parler Tyébara, Côte d'Ivoire)* (2 vols). Köln: Rüdiger Köppe.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chakali (NC, Gur: Gurunsi)
Brindle, J. (2017). *A dictionary and grammatical outline of Chakali*. Berlin: Language Science Press. http://langsci-press.org/catalog/book/74 (Accessed December, 2017).
Chumburung (NC, Kwa: Nyo)
Hansford, G. F. (2010). Red is a verb: the grammar of colour in Chumburung. *Journal of West African Languages* 37(2): 109–137.
Ciluba (NC, Benue-Congo: Bantu L)
Kabuta, N. S. et al. (n.d.). *Dictionnaire Cilubà – Français*. www.ciyem.ugent.be/ (Accessed December, 2017).
Day (NC, Adamawa: Day)
Nougayrol, P. (1980). *Le day de Bouna (Tchad), 2: Lexique day–français, index français–day*. Paris: SELAF. In RefLex.
Degema (NC, Benue-Congo: Edoid)
Kari, E. E. (2008). *Degema–English dictionary with English index*. Tokyo: ILCAA.
Dendi (VARIA, Songhai: Southern)
Zima, P. (1994). *Lexique dendi (songhay) (Djougou, Bénin), avec un index français–dendi*. Köln: Rüdiger Köppe Verlag. In RefLex.
Dhaasanach (AA, Cushitic: Eastern)
Tosco, M. (2001). *The Dhaasanac language: Grammar, texts and vocabulary of a Cushitic language of Ethiopia*. Köln: Rüdiger Köppe.
Digo (NC, Benue-Congo: Bantu E)
Mwalonya, J., Nicolle, A., Nicolle, S., & Zimbu, J. (2004). *Mgombato: Digo–English–Swahili Dictionary*. Kwale, Kenya: Bible Translation and Literacy.
Dii (Yag Dii) (NC, Adamawa: Duru)
Bohnhoff, L. E. (2014). *Dictionnaire de la langue dii (duru)*. Mbé (Cameroun): Equipe de littérature Dii.
Ding (NC, Benue-Congo: Bantu B)
Mertens, J. (1939). *Les badzing de la Kamtsha, III: Dictionnaire idzing–français, suivi d'un aide-mémoire français–idzing*. Bruxelles: Librairie Falk fils.
Dott (=Zoɗì) (AA, Chadic: Western)
Caron, B. (2002). Dott, aka Zoɗì: Grammatical notes, vocabulary, text. *Afrika und Übersee* 85: 161–248. In RefLex.
Duala (NC, Benue-Congo: Bantu A)
Helmlinger, P. (1972). *Dictionnaire duala–français, suivi d'un lexique français–duala*. Paris: Klincksieck.
Edo (NC, Benue-Congo: Edoid)
Wescott, R. W. (1970). Bini color terms. *Anthropological linguistics* 13: 251–252.
Enya (NC, Benue-Congo: Bantu D)
Spa, J. J. (1975). *Vocabulaire Enya*. Tervuren: Musée Royal de l'Afrique Centrale.
Eton (NC, Benue-Congo: Bantu A)
Van de Velde, M. (2017). *Eton data (personal fieldwork)*.
Fang (NC, Benue-Congo: Bantu A)
Galley, S. (1964). *Dictionnaire fang–français et français–fang, suivi d'une grammaire fang*. Neuchâtel: Henri Messeiller.
Fon-Gbe (NC, Kwa: Gbe)
Höftmann, H. (2003). *Dictionnaire fon–français, avec une esquisse grammaticale*. Köln: Rüdiger Köppe.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Fonyi (NC, Atlantic-Bak: Joola)
Sapir, J. D. (1993 [1970]). *Dictionnaire Jóola Kujamutay*. Ms. In RefLex. Also online at http://people.virginia.edu/~ds8s/Kujamaat-Joola/DIC/Joola-Dic.html (Accessed December, 2017).
Weiss, [P.] H. (1939). Grammaire et lexique diola du Fogny (Casamance). *Bulletin de l'IFAN* 1–2: 412–578. In RefLex.
Fula (Maasina) (NC, Atlantic-North: Fula-Sereer)
Seydou, C. (2014). *Dictionnaire peul–français (variété du Massina)*. In RefLex.
Fulfulde (Adamawa) (NC, Atlantic-North: Fula-Sereer)
Noye, D. (1989). *Dictionnaire foulfouldé–français, dialecte peul du Diamaré du Nord-Cameroun*. Paris, Garoua (Cameroon): Geuthner; Procure des Missions.
Fyem (NC, Benue-Congo: Plateau)
Nettle, D. (1998). *The Fyem language of Northern Nigeria*. München: Lincom Europa.
Gade (NC, Benue-Congo: Nupoid)
Sterk, J. P. (1994). *Gade–English dictionary, including English–Gade reference dictionary and summary of Gade grammar*. Berlin: Dietrich Reimer.
Ganda (NC, Benue-Congo: Bantu J)
Snoxall, R. A. (1967). *Luganda–English dictionary*. Oxford: Clarendon Press.
Ganja (NC, Atlantic-Bak: Balanta)
Creissels, D., & Biaye S. (2015). *Le balant ganja: Phonologie, morphosyntaxe, liste lexicale, textes*. Dakar: IFAN. In RefLex.
Gbagyi (NC, Benue-Congo: Nupoid)
Hyman, L. M., & Magaji, D. J. (1970). *Essentials of Gwari grammar*. Ibadan: University of Ibadan. In RefLex.
Gbaya-Kara (NC, Ubangi: Gbaya-Manza-Ngbaka)
Blanchard, Y., & Noss, P. A. (1982). *Dictionnaire gbaya–français, dialecte yaayuwee*. Meiganga (Cameroun): Centre de Traduction Gbaya. In RefLex.
Gciriku (NC, Benue-Congo: Bantu K)
Möhlig, W. J. G., & Shiyaka-Mberema, K. P. (2005). *A dictionary of the Rumanyo language: Rumanyo–English/English–Rumanyo, including a grammatical sketch*. Köln: Rüdiger Köppe.
Gedeo (AA, Cushitic: Eastern)
Wedekind, K. (2008). *Gedeo Dictionary, Revision of 1978*. In RefLex. Also online at http://kwedekind.de/Eingang1/Biblio1.htm (Accessed December, 2017).
Giryama (NC, Benue-Congo: Bantu E)
Deed, F. I. (1964). *Giryama–English dictionary*. Nairobi: East African Literature Bureau.
Gorowa (AA, Cushitic: Southern)
Kiessling, R. (2016). *Burunge worldist*. Ms.
Gouro (NC, Mande: Eastern)
Grégoire, H. C. (1975). *Étude de la langue gouro (région de Zuénoula): Lexique*. Abidjan: Institut de Linguistique Appliquée, Université d'Abidjan. In RefLex.
Guus (Sigidi) (AA, Chadic: Western)
Caron, B. (2001). Guus, aka Sigidi (Chadic, West-B, South-Bauchi): Grammatical notes and vocabulary. *Afrika und Übersee* 84: 1–60. In RefLex.
Hausa (AA, Chadic: Western)
Bargery, G. P. (1934). *A Hausa–English dictionary, and English–Hausa vocabulary*. London: Humphrey Milford for Oxford University Press

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ibibio (NC, Benue-Congo: Cross River)
Urua, E.-A., Ekpenyong, M., & Gibbon, D. (2004). *Uyo Ibibio Dictionary*. ABUILD Language Documentation Curriculum Materials #1. University of Uyo and Universität Bielefeld. In RefLex.
Igbo (NC, Benue-Congo: Igboid)
Echeruo, M. J. C. (1998). *Igbo–English dictionary: A comprehensive dictionary of the Igbo language, with an English–Igbo index*. New Haven CN & London: Yale University Press.
Ik (NS, Eastern Sudanic: Kuliak)
Schrock, T. (2017). *The Ik language: Dictionary and grammar sketch*. Berlin: Language Science Press. In RefLex. Also online at http://langsci-press.org/catalog/book/98 (Accessed December, 2017).
Iraqw (AA, Cushitic: Southern)
Mous, M., Qorro, M. P., & Kiessling, R. (2002). *Iraqw–English dictionary, with an English and a thesaurus index*. Köln: Rüdiger Köppe.
Jaad (Badiaranke) (NC, Atlantic-North: Jaad-Biafada)
Meyer, [P.] G. (2001). *Lexique badiaranke–français: Parler de la région de Koundara, Guinée*. Dakar. In RefLex.
Jamsay (NC, Dogon: South-Eastern)
Heath, J. (2013a). *Jamsay lexicon (Douentza)*. In RefLex. Also online at http://dogonlanguages.org/values?parameter=06853 (Accessed December, 2017).
Jul'hoan (Khoisan, Non-Khoe: Ju)
Dickens, P. J. (1994). *English–Ju|'hoan, Ju|'hoan–English dictionary*. Köln: Rüdiger Köppe. In RefLex.
Jula-Odienné (NC, Mande: Western)
Braconnier, C. (1999). *Dictionnaire du dioula d'Odienné* (2 volumes). Paris: Université Paris 7.
Jur Modo (NS, Central Sudanic: SBB)
Persson, A. M., & Persson, J. (1991). *Modo–English dictionary with grammar*. Nairobi: Summer Institute of Linguistics (SIL). In RefLex.
Kagoro (NC, Mande: Western)
Vydrin, V. (2001). *Esquisse contrastive du Kagoro (Manding)*. Köln: Rüdiger Köppe.
Kakabe (NC, Mande: Western)
Vydrina, A. (2017). *Kakabe data (personal fieldwork)*.
Kambaata (AA, Cushitic: Eastern)
Treis, Y. (2017). *Kambaata data (personal fieldwork)*.
Kanuri (NS, Saharan: Western)
Cyffer, N., & Hutchison, J. P. (Eds.). (1990). *Dictionary of the Kanuri language*. Dordrecht: Mouton de Gruyter; Foris Publications. In RefLex.
Kapsiki (AA, Chadic: Central)
van Beek, W. E. A. (1977). Color terms in Kapsiki. In P. Newman, & R. M. Newman (Eds), *Papers in Chadic linguistics: Papers from the Leiden colloquium on the Chadic language family* (pp. 13–20). Leiden: Afrika-Studiecentrum.
Karon (NC, Atlantic-Bak: Joola)
Wilkinson, S., & Berndt, W. (2011). *Esquisse de grammaire de la langue karone*. SIL Senegal.
Keeraak (NC, Atlantic-Bak: Joola)
Segerer, G. (2017). *Joola Keeraak data (personal fieldwork)*.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kemant (AA, Cushitic: Central)
See Awngi
Kenga (NS, Central Sudanic: SBB)
Palayer, P. (2004). *Dictionnaire kenga (Tchad)*. Louvain: Peeters.
Kenyang (NC, Benue-Congo: Bantoid)
Mbuagbaw, T. E. (1998). *Kenyang lexicon*. Yaoundé: Society for Kenyang Literature (SKL); Cameroon Association for Bible Translation and Literacy (CABTAL).
Kikongo (NC, Benue-Congo: Bantu H)
Laman, K. E. (1936). *Dictionnaire kikongo–français, avec une étude phonétique décrivant les dialectes les plus importants de la langue dite kikongo*. Bruxelles: Librairie Falk fils.
Koalib (NC, Kordofanian: Heiban)
Quint, N. (2017). *Koalib data (personal fieldwork)*.
Kol (NC, Benue-Congo: Bantu A)
Begne, L. P. (1980). *The phonology of Bikele, a Cameroonian language*. Illinois Institute of Technology (PhD thesis).
Henson, B. J. (2007). *The Phonology and Morphosyntax of Kol*. Berkeley: University of California.
Kom (NC, Benue-Congo: Bantoid)
Jones, R. (2001). *Provisional Kom–English lexicon*. Yaoundé: SIL International.
Koromfe (NC, Gur: Kurumfe)
Rennison, J. R. (1997). *Koromfe*. London, New York: Routledge
Kpelle (Liberia) (NC, Mande: Western)
Westermann, D., & Melzian, H. J. (1930). *The Kpelle language in Liberia: Grammatical outline, coloquial sentences and vocabulary*. Berlin: Dietrich Reimer. In RefLex.
Krahn (Eastern) (NC, Kru: Western)
Sauder, D., & Wright, P. (2000). *Krahn–English Dictionary, English–Krahn (Tchien Dialect)*. Mississauga (Canada).
Krim (NC, Mel: Southern)
Childs, G. T. (2012). *Kim data (personal fieldwork)*. In RefLex.
Kwanyama (NC, Benue-Congo: Bantu R)
Halme, R. (2004). *A tonal grammar of Kwanyama*. Köln: Rüdiger Köppe.
Kyanga (NC, Mande: Eastern)
McCallum Jones, R. (2017b). *Kyanga Dictionary*. München: Lincom Europa.
Laal (VARIA, Unclassified)
Boyeldieu, P. (2012). *Laal wordlist (personal fieldwork)*.
Laala (NC, Atlantic-North: Cangin)
Pichl, W. J. (1981). *Vocabulaire laala*. Ms. In RefLex.
Landuma (NC, Mel: Northern)
Rogers, K., & Bryant, D. (2012). *Diksiyɔnɛr kəlʌndma – kətabu – Dictionnaire landouma – français*. Boké: Mission Evangélique de Boké. In RefLex.
Lega-Beya (NC, Benue-Congo: Bantu D)
Botne, R. D., & Salama-Gray, K. (1994). *A Lega and English dictionary, with an index to Proto-Bantu roots*. Köln: Rüdiger Köppe. In RefLex.
Lelemi (Lefana) (NC, Kwa: Nyo)
Kropp Dakubu, M. E. (1967). *Lefana, Akpafu and Avatime, with English gloss*. Legon: Institute of African Studies, University of Ghana.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Limba (NC, Limba: Nyo)
Clarke, M. L. (1922). *A Limba–English dictionary, or: Tampen ta ka talun ta ka hulimba in huinkilisi ha*. New York: Houghton. In RefLex.
Limbum (NC, Benue-Congo: Bantoid)
Fransen, M. A. E. (1995). *A grammar of Limbum, a Grassfields Bantu language spoken in the North-West Province of Cameroon*. Amsterdam: Free University of Amsterdam. In RefLex.
Lingala (NC, Benue-Congo: Bantu C)
Kawata, A. T. (2004). *Bagó – dictionnaire Lingala/Falansé – Français/Lingala*. Paris: Karthala.
Logo (NS, Central Sudanic: Moru-Madi)
Vallaeys, A. (1986). Dictionnaire logo–français suivi d'un index français–logo. *Tervuren: Musée Royal de l'Afrique Centrale (MRAC)*.
Longando (NC, Benue-Congo: Bantu C)
Tuerlings, W. (2008). *Longandó dictionary – Longandó–English, English–Longandó*. Waitakere (New Zealand).
Lumun (NC, Kordofanian: Talodi)
Smits, H. (2017). *A grammar of Lumun: A Kordofanian language of Sudan*. Leiden: Leiden University (PhD Thesis). https://www.lotpublications.nl/Documents/465_fulltext.pdf (Accessed March, 2018).
Lunda (NC, Benue-Congo: Bantu K)
White, C. M. N. (1957). *A Lunda–English Vocabulary*. London: University of London Press.
Maasai (NS, Eastern Sudanic: Nilotic)
Payne, D. L. (2003). Maa color terms and their use as human descriptors. *Anthropological Linguistics* 45–2: 169–200.
Mada (NC, Benue-Congo: Plateau)
Blench, R., & Kato, B. (2015). *A dictionary of Mada, a Plateau language of Central Nigeria*. http://www.rogerblench.info/Language/Niger-Congo/BC/Plateau/Ninzic/Mada/ (Accessed March, 2018).
Mambai (NC, Adamawa: Mbum)
Anonby, E. J., & Oussoumanou, K. B. (2014). *Dictionnaire mambay–français*. Garoua: SIL Cameroun – Comité de Langue Mambay.
Mandinka (NC, Mande: Western)
Creissels, D. (2012). *Lexique mandinka–français*. Ms. In RefLex.
Maninka-Kankan (NC, Mande: Western)
Diané, M. (2012). L'expression des couleurs en maninka de Kankan. *Mandenkan* 48: 21–38
Maninka-Niokolo (NC, Mande: Western)
Creissels, D. (2013). Le maninka du Niokolo (Sénégal oriental), esquisse phonologique et morphosyntaxique, liste lexicale, textes glosés. *Mandenkan* 49: 1–218.
Mankanya (NC, Atlantic-Bak: Manjaku)
Gaved, T., & Stammers, J. (2004). *Petit lexique mancagne–français, suivi d'un index français–mancagne*. Dakar: SIL Sénégal. In RefLex.
Mano (NC, Mande: Eastern)
Khachaturyan, M. (2012). *Mano data (personal fieldwork)*. In RefLex.
Mba-ndunga-le (NC, Ubangi: Sere-Ngbaka-Mba)
Pasch, H. (1987). *Die Mba-Sprachen: die Nominalklassensysteme und die genetische Gliederung einer Gruppe von Ubangi-Sprachen*. Hamburg: Helmut Buske. In RefLex.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mba-ne (NC, Ubangi: Sere-Ngbaka-Mba)
See Mba-ndunga-le.
Mbuko (AA, Chadic: Central)
Gravina, R., Nelezek, R., & Tchalalao, R. (1999). *Mbuko Lexicon*. Yaoundé: SIL International. In RefLex. Also online at www.silcam.org (Accessed December, 2017).
Mbum (NC, Adamawa: Mbum)
Hino, S. (1978). *The classified vocabulary of the Mbum language in Mbang Mboum, with ethnographical descriptions*. Tokyo: ILCAA.
Minyanka (Mamara) (NC, Gur: Senufo)
Association pour la promotion de la langue Mamara. (2009). *Dictionnaire Mamara – Mamaara Kafilakaya Kpuun*. Koutiala, Mali: Association pour la promotion de la langue Mamara. In RefLex.
Mmen (NC, Benue-Congo: Bantoid)
Mua, B. (Comp.). (2013). *Mmen–English Lexicon, with English–Mmen index*. Yaoundé: Cameroon Association for Bible Translation and Literacy (CABTAL).
Mofu-Gudur (AA, Chadic: Central)
Barreteau, D. (1983). *Description du mofu–gudur (langue de la famille tchadique parlée au Cameroun)* (2 volumes). Paris: Université Paris 3.
Mokpwe (NC, Benue-Congo: Bantu A)
Connell, B. A. (Ed.). (1997). *Mòkpè (Bakweri)–English dictionary*. Köln: Rüdiger Köppe. In RefLex.
Mongo-Nkundu (NC, Benue-Congo: Bantu C)
Hulstaert, G. (1957). *Dictionnaire lomongo–français* (2 volumes). Tervuren: Musée Royal du Congo Belge.
Mòoré (NC, Gur: Oti-Volta)
Niggli, U. (Ed.). (2016). *Moore–Français–English Dictionnaire*. SIL International. https://moore.webonary.org/ (Accessed December, 2017).
Mundu (NC, Ubangi: Sere-Ngbaka-Mba)
Vallaeys, A. (1991). *La langue mondo: Esquisse grammaticale, textes et dictionnaire*. Tervuren: Musée Royal de l'Afrique Centrale. In *RefLex*.
Mursi (NS, Eastern Sudanic: Surmic)
Turton, D., Yigezu, M., & Olibuli, O. (2008). *Mursi–English–Amharic Dictionary*. Addis Ababa. http://www.mursi.org/pdf/dictionary.pdf (Accessed December, 2017).
Nalu (NC, Atlantic-North: Nalu)
Seidel, F. (2013). *Nalu data (personal fieldwork)*. In RefLex.
Nanga-Dama (NC, Dogon: Northern)
Heath, J. (2013b). *Nanga lexicon (Anda)*. In RefLex. Also online at http://dogonlanguages.org/values?parameter=06853 (Accessed December, 2017).
Ndebele (Zimbabwe) (NC, Benue-Congo: Bantu S)
Pelling, J. N. (1971). *A practical Ndebele dictionary*. Salisbury: Longman Rhodesia. In RefLex.
Nembe (NC, Ijoid: Western Ijo)
Kaliai, M. H. I. (2008). *A Nembe–English dictionary*. Cambridge: Mallam Dendo. In RefLex. Also online at http://www.rogerblench.info/Language%20data/NC/Ijoid/ (Accessed December, 2017).
Ngbaka (NC, Ubangi: Gbaya-Manza-Ngbaka)
Henrix, M. (2000). *Dictionnaire ngbaka–français*. Ghent: Research Centre of African Languages and Literatures, Ghent University.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ngbaka Ma'bo (NC, Ubangi: Sere-Ngbaka-Mba)
Thomas, J. M. C. (1989). Des noms et des couleurs. In *Graines de Paroles. Puissance du verbe et traditions orales. Textes offerts à Geneviève Calame-Griaule*. Paris: CNRS.
Ngizim (AA, Chadic: Western)
Adamu, M. A., & Potiskum, U. B. G. . (2009). *Ngizim–English–Hausa dictionary* (edited by Russell G. Schuh). Potiskum, Yobe State: Ajami Press.
Ngwo (NC, Benue-Congo: Bantoid)
Eyoh, J. A. (2011). *Engwo Lexicon – Engwo–English, English–Engwo*. Ngwo (Cameroon): Ngwo Language and Cultural Committee; Ngwo Cultural and Development Association. In RefLex.
Noone (NC, Benue-Congo: Bantoid)
Lux, D. (2003). *Noni Provisional Lexicon*. Yaoundé: SIL Cameroon. In RefLex. Also online at www.silcam.org. (Accessed December, 2017).
Nugunu (NC, Benue-Congo: Bantu A)
Robinson, C. D. W. (2003). *Nugunu provisional lexicon*. Yaoundé: SIL International / Nugunu Language Committee (GULICO). In RefLex. Also online at www.silcam.org (Accessed December, 2017).
Nyakyusa-Ngonde (NC, Benue-Congo: Bantu M)
Felberg, K. (1996). *Nyakyusa–English–Swahili & English–Nyakyusa dictionary*. Dar es Salaam: Mkuki na Nyota Publishers.
Nyamwezi (NC, Benue-Congo: Bantu F)
Maganga, C., & Schadeberg, T. C. (1992). *Kinyamwezi: Grammar, texts, vocabulary*. Köln: Rüdiger Köppe. In RefLex.
Nzadi (NC, Benue-Congo: Bantu B)
Crane, T. M., Hyman, L. M., & Nsielanga Tukumu, S. (2011). *A Grammar of Nzadi [B865]: A Bantu Language of Democratic Republic of Congo*. Berkeley: University of California Press. In RefLex. Also online at http://escholarship.org/uc/item/846308w2 (Accessed December, 2017).
Nzema (NC, Kwa: Nyo)
Aboagye, P. A. K. (1992). *Nzema–English, English–Nzema dictionary*. 3rd Edition. Accra: Ghana Publishing Corporation.
Obolo (NC, Benue-Congo: Cross River)
Rowland Oke, M. (2003). *Description systématique de la langue obolo-andoni*. Paris: L'Harmattan. In RefLex.
Otetela (NC, Benue-Congo: Bantu C)
Hagendorens, J. (1975). *Dictionnaire otetela–français*. Bandundu: Centre d'Etudes Ethnologiques de Bandundu (CEEBA).
Peere (NC, Adamawa: Duru)
Raen, K. (1985). *Dictionnaire pere–français*. Ngaoundéré (Cameroun): Eglise Evangélique Luthérienne du Cameroun. In RefLex.
Pular (NC, Atlantic-North: Fula-Sereer)
Bah, O. (2009). *Saggitorde Oumar Bah [Dictionnaire pular]*. In RefLex. Also online at https://pular.webonary.org/ (Accessed December, 2017).
Rombo (Northern) (NC, Benue-Congo: Bantu E)
Shinagawa, D. (2015). *A basic vocabulary of Chaga-Rombo (Bantu E623)*. Tokyo: ILCAA.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rundi (NC, Benue-Congo: Bantu J)
Rodegem, F. M. (1970). Dictionnaire rundi–français. Tervuren: Musée Royal de l'Afrique Centrale (MRAC).
Samo-Maya (NC, Mande: Eastern)
Morris, P., Koussoubé, M., & Seme, P. (2011). *Lexique San Mayaa avec guide d'orthographe*. Tougan, Burkina Faso: Association nationale pour la traduction de la Bible et alphabétisation (ANTBA). In RefLex.
Sandawe (VARIA, Hatsa-Sandawe)
Ehret, C., & Ehret, P. (Eds.). (2012). *A Dictionary of Sandawe – compiled from the field collections of Eric Ten Raa*. Köln: Rüdiger Köppe.
Sango (NC, Ubangi: Ngbandi)
Bouquiaux, L., Kobozo, J.-M., Diki-Kidiri, M., Vallet, J., & Behaghel, A. (Eds.). (1978). *Dictionnaire sango–français et lexique français–sango*. Paris: SELAF.
Sar (NS, Central Sudanic: SBB)
Gotengaye, C., & Keegan, J. M. (2016). *Dictionnaire Sar*. Cuenca: Morkeg Books. http://morkegbooks.com/Services/World/Languages/SaraBagirmi/#title (Accessed December, 2017).
Seenku (NC, Mande: Western)
McPherson, L. (n.d.). *Seenku worldist*. Ms. In RefLex.
Sereer (NC, Atlantic-North: Fula-Sereer)
Crétois, L. (1973/77). *Dictionnaire sereer–français* (6 volumes). Dakar: Centre de Linguistique Appliquée de Dakar (CLAD). In RefLex.
Sherbro (NC, Mel: Southern)
Pichl, W. J. (1967). *A Sherbro–English dictionary*. Pittsburgh: Duquesne University Press. In RefLex.
Siwu (Akpafu) (NC, Kwa: Nyo)
Dingemanse, M. (in press). The Language of Perception in Siwu. In A. Majid, & S. C. Levinson (Eds.), *The Oxford Handbook of the Language of Perception*. Oxford: Oxford University Press.
Soga (NC, Benue-Congo: Bantu J)
Gonza, R. K. (2007). *Lusoga–English dictionary and English–Lusoga dictionary*. Kampala: MK Publishers Ltd.
Somali (AA, Cushitic: Eastern)
Zorc, R. D. P., & Osman, M. C. (Eds.). (1993). *Somali–English dictionary, with English index*. Wheaton MD: Dunwoody Press.
Soninke (NC, Mande: Western)
Creissels, D. (2016). *Phonologie segmentale et tonale du soninké (parler du Kingi)*. *Mandenkan* 55. 3–174. In RefLex.
Soo (NS, Eastern Sudanic: Kuliak)
Carlin, E. B. (1993). *The So language*. Köln: Institut für Afrikanistik, Universität zu Köln.
Southern Samo (San) (NC, Mande: Eastern)
SIL Burkina. (2003). *Boo nɛn sɛ́wɛ san–fransi, fransi–san [Lexique san–français, français–san]*. SIL Burkina Faso. In RefLex.
Sura (AA, Chadic: Western)
Jungraithmayr, H. (1963/64). Die Sprache der Sura (Maghavul) in Nordnigerien. *Afrika und Übersee* 47: 8–89, 204–220. In RefLex.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Susu (NC, Mande: Western)
Engelbrecht, R. (Comp.). (2004). *Susu Dictionary and English–Susu Word List – English–Susu Bible Names*. New Tribes Missions. In RefLex.
Swahili (NC, Benue-Congo: Bantu G)
Lenselaer, A. (1983). *Dictionnaire swahili–français*. Paris: Karthala.
Tima (NC, Kordofanian: Katla)
Schneider-Blum, G. (2013). *A Tima–English Dictionary: An Illustrated Lexicon of a Niger-Congo Language Spoken in the Nuba Mountains (Sudan)*. Köln: Rüdiger Köppe.
Togo-Kan (NC, Dogon: South-Eastern)
Heath, J. (2013c). *Togo-Kan lexicon (Koporo-pen)*. In RefLex. Also online at http://dogon-languages.org/values?parameter=06853 (Accessed December, 2017).
Tommo So (NC, Dogon: Central)
McPherson, L. (2013). *Tommo-So lexicon (Tongo Tongo)*. In RefLex. Also online at http://dogonlanguages.org/values?parameter=06853 (Accessed December, 2017).
Toro Tegu (NC, Dogon: South-Eastern)
Heath, J. (2013d). *Toro Tegu lexicon (Tabi)*. In RefLex. Also online at http://dogonlanguages.org/values?parameter=06853 (Accessed December, 2017).
Toura (NC, Mande: Eastern)
Idiatov, D. (2017). *Toura data (personal fieldwork)*.
Tswana (NC, Benue-Congo: Bantu S)
Creissels, D., & Chebanne, A. M.. (2000). *Dictionnaire français–setswana, setswana–français*. Mogoditshane (Botswana) & Gaborone: Tasalls Publishing & Books. In RefLex.
Tupuri (NC, Adamawa: Mbum)
Ruelland, S. (1988). *Dictionnaire tupuri–français–anglais (Région de Mindaore, Tchad)*. Paris: Peters; SELAF; CNRS. In RefLex.
Uduk (NS, Komuz: Koman)
Killian, D. (2015). *Topics in Uduk Phonology and Morphosyntax*. Helsinki: University of Helsinki.
Viya (NC, Benue-Congo: Bantu B)
van der Veen, L. J., & Bodinga-Bwa-Bodinga, S. (2002). *Gedandedi sa geviya/Dictionnaire geviya–français*. Leuven, Paris: Peeters. In RefLex.
Wasi (AA, Cushitic: Southern)
Mous, M. (2016). *Alagwa – A South Cushitic Language of Tanzania – Grammar, Texts and Lexicon*. Köln: Rüdiger Köppe.
Wawa (NC, Benue-Congo: Bantoid)
Martin, M. (2012). *A grammar of Wawa: An endangered language of Cameroon*. Phd Thesis, University of Kent.
Wolof (NC, Atlantic-North: Wolof)
Diouf, J.-L. (2003). *Dictionnaire wolof–français et français–wolof*. Paris: Karthala. In RefLex.
Fal, A., Santos, R., & Doneux, J. L. (1990). *Dictionnaire wolof–français suivi d'un index français–wolof*. Paris: Karthala. In RefLex.
Xamtanga (AA, Cushitic: Central)
cf. Awngi.
Xamtanga (AA, Cushitic: Central)
Teshome, B. (2015). *A Grammar of Khimt'anga*. Addis Ababa: Addis Ababa University.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Yao (NC, Benue-Congo: Bantu P)
Ngunga, A. S. A. (2001). *Yao lexicon*. CBOLD. http://www.cbold.ish-lyon.cnrs.fr/Dico.asp?Langue=Yao (Accessed December, 2017).
Yoruba (NC, Benue-Congo: Defoid)
Sachnine, M. (1997). *Dictionnaire usuel yoruba–français suivi d'un index français–yoruba*. Paris: Karthala. In RefLex.
Yulu (NS, Central Sudanic: SBB)
Boyeldieu, P. (1990). *Yulu data (personal fieldwork)*.
Zarma (VARIA, Songhai: Southern)
Ducroz, J. M., & Charles, M. C. (1978). *Lexique soŋej (songhay) français: Parler kaado du Gorouol*. Paris: L'Harmattan. In RefLex.
Zaar (AA, Chadic: Western)
Caron, B. (2005). *Za:r (Dictionary, grammar, texts)*. Ibadan (Nigeria): IFRA. In RefLex.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

PART III

Languages in culture and languages in contact

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Universals and variability of color naming in Icelandic, Icelandic Sign Language, and North American Icelandic

Þórhalla Guðmundsdóttir Beck and Matthew Whelpton
University of Iceland

Color naming remains a lively area of debate pitting supporters of universal neuro-psychological constraints against cultural relativists. The Evolution of Semantic Systems project offered an extensional semantic approach, using statistical methods to measure relative variability. This chapter uses a similar methodology to compare color naming in Icelandic with a typologically unrelated sign language (Icelandic Sign Language) in the same community, and with a heritage language developing in an English-speaking environment (North American Icelandic). Data from North American English and British English are provided for context. The statistical analysis shows very little variation in naming patterns. However, a more detailed analysis of formal mechanisms of color naming (compounding, object-naming, opaque lexemes) shows rich variation obscured by a methodological emphasis on transparent semantic heads and basic color terms.

Keywords: color, semantic categories, universals, relativism, extensional semantics, Icelandic, Icelandic Sign Language, North American Icelandic, compounding, heritage language

1. Introduction

Color is one of the most salient aspects of our environment, apparently leaping out for our attention, like red holly berries against a dark green background or the shock of blood on a white sheet. It is easy to assume that the word *red* is simply a label for an aspect of our environment which must "cry out to be named" (Berlin, 1992, p. 290). It may, therefore, surprise us to learn that some languages do not have a term that can be considered as equivalent to red, even though, along with white and black, red is indeed one of the most common terms included in the color lexicon. Less surprising is the fact that languages do not always distinguish red from

https://doi.org/10.1075/sfsl.78.14bec
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

orange or orange from yellow, while some distinguish as basic colors what English speakers would call light blue and dark blue.

This diversity of naming made the lexical field of color a flagship example of linguistic relativity amongst the early structuralists, who argued that color is a constructed category, by which our linguistic system helps us to make sense of the world (Conklin, 1955). There is no better example of what Whorf (1956) calls the "kaleidoscopic flux of impressions which has to be organized by our minds" than the perceptual input of color. Nevertheless, the classic study of Berlin & Kay (1969) overturned this relativist consensus arguing for robust universal perceptual constraints on color naming. Contemporary socio-cultural approaches have returned the focus of interest to the microvariation in naming and its cultural underpinnings.

This chapter reports on a comparison of color naming strategies in Icelandic (ICE), Icelandic Sign Language (ISL) and North American Icelandic (NAI), using standardized methodologies developed in the Evolution of Semantic Systems (EoSS) project (Majid, Jordan, & Dunn, 2011). To provide linguistic and cultural context, data is also provided from the North American English (NAE) spoken in the same heritage communities as North American Icelandic, as well as British English (BRE; European context). North American Icelandic is a heritage language in Canada and the northern United States, spoken by the descendants of nineteenth-century immigrants from Iceland (Arnbjörnsdóttir, 2006). We might, therefore, expect a contrast in color naming strategies between (European) Icelandic, which developed in relative isolation and with a strong nationalist and purist tradition (Rögnvaldsson, Jóhannsdóttir, Steingrímsson, Loftsson, & Helgadóttir, 2011), and North American Icelandic, which developed in a predominantly English-speaking environment. The difference here is cultural rather than typological. In contrast, both Icelandic and Icelandic Sign Language are spoken in Iceland by Icelanders (Sverrisdóttir & Thorvaldsdóttir, 2016), but Icelandic Sign Language is quite different in formal and typological characteristics from Icelandic (although, as there is no written standard for Icelandic Sign Language, many in the deaf community can read and write Icelandic). Any differences here would be likely related to typological differences in the linguistic systems.

It should be noted that we use *cultural* here in a broad sense to encompass social, political and shared linguistic aspects of community life, as opposed to formal typological aspects of the linguistic systems and inherited conventions of linguistic usage. As one reviewer points out, we are therefore not distinguishing between the effects of language contact and influence from non-linguistic aspects of culture, as Vos (2011) does. Cultural influences on North American Icelandic will include the usage conventions of North American English, such as the names of paint or yarn colors. Similarly, the cultural influences on Icelandic Sign Language will include the written conventions of Icelandic, which is used in deaf education in Iceland.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The statistical analysis of color naming patterns in the five languages supports the results reported for the EoSS Germanic languages (Majid, Jordan, & Dunn, 2015), which show strong universal tendencies in color naming. However, within this broadly similar system, there is ample evidence of the sort of microvariation that the socio-cultural theorists have emphasized. There are differences in the number of dominant color terms and some interesting contrasts concerning the areas of the color space emerging for lexicalization. There is an enormous amount of variation in the full range of color terms used. This in part reflects significant differences in the formal strategies exploited by each language to elaborate their color lexicon: for instance, both varieties of English make extensive use of derivational morphology, but Icelandic relies almost exclusively on compounding and modification.

In Section 2, we set our study in the context of the broader research literature. In Section 3, we outline the EoSS methods employed to collect our data. In Section 4, we report our results: a statistical analysis of correlations in naming (4.1), an overview of the main terms used to categorize the stimuli (4.2), and a review of the formal methods used to elaborate these main terms, including an analysis of the semantic features most commonly coded in these elaborations (4.3). In Section 5, we draw our conclusions.

2. Background

Color naming has proved to be an enduringly controversial topic of research, pitting proponents of cultural relativity against proponents of cognitive universals. In the 1950s, anthropologists took color naming to be the paradigmatic case of free variation across languages, where the divisions of particular hues into categories was driven by the vocabulary system itself and not by any inherent divisions in the physical color space (Nida, 1959), reflecting cultural concerns, e.g., with freshness, which had nothing to do with the spectral properties of hue (Conklin, 1955).

The physical phenomenon perceived as color is simply light waves striking the retina. These waves vary along three continua: the wavelength of light, which determines hue; the intensity of the wave, which determines luminosity or brightness; and the degree to which the wave is free from white, which determines saturation. As these are smooth continua, there are no disruptions in the physical space which would favor boundaries. The fact that English distinguishes between *red*, *orange*, and *yellow* has therefore nothing to do with the measurable qualities of the physical waves themselves. Languages do indeed vary widely in the number of color terms they make available (Kay, Berlin, Maffi, Merrifield, & Cook, 2011) and in the placement of the boundaries between those colors. There is no obvious source for such variation other than the internal organization of the color vocabulary system and

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

the cultural conventions for their use. The linguistic and cultural nature of color naming is further reinforced by the fact that the color vocabulary appears to encode more than hue-based distinctions, for instance relating to shininess (Lucy, 1997).

Despite the strength of the structuralist relativist position, this consensus was overturned by a landmark work (Berlin & Kay, 1969) arguing that the color space is indeed structured by universals of perception. They showed that, although the number of color terms and the boundaries between them might vary widely, the focal colors with which a color name is associated tend to be stable across languages: a "good red" in one language is usually very close to a "good red" in another language. They argued that this is not surprising given that the perceived light does not in fact represent a uniform continuum. The human visual system perceives certain distinctions in the physical properties of light more easily than others, due to the functioning of the retina and the neurological structure which supports it. At the most obvious level, we see red and violet but not infrared and ultraviolet, even though all four "colors" are parts of the same continuous spectrum. Similarly, colors such as red are perceptually far more salient than colors such as brown. Peaks in perceptual salience in the color space attract the linguistic terms provided for naming the space. Further, although languages vary in the number of terms, the order in which items are added to the system is not random, but is again structured by the relative salience of the focal colors in the perceived color space. Berlin and Kay therefore argue that there are semantic universals of color naming which are underpinned by innate properties of the visual-perceptual system in humans.

Berlin and Kay (1969) also shifted attention away from the full range of color terms and color naming strategies in a language to focus on basic color terms, the core lexical set which provides the major categorial divisions in the extensional space. Basic color terms are generally monolexemic (*brown* but not *light brown*), frequent (*purple* but not *indigo*), applied to many objects rather than just certain objects (*yellow* but not *blond*), and judged by speakers not to be a hyponym or subcase of another color (*red* but not *scarlet*). Their central claim is that basic color terms are added to a language in a fixed order. Although languages may differ in the number of color terms overall and may (in more recent versions of the proposal) have alternate paths for adding color terms, there is a predictability to the order in which terms are added. For instance, no language will have a basic color term for yellow without also having a color term for red. Further, whatever the extensional space that the new term covers in the language (which will depend on how many competing terms have been added), the focal point defining a good example of the color will tend to be stable cross-linguistically. Whether the equivalent term for *red* extends to include *orange* and warm colors generally or has separate terms dividing up this larger space, the focal example of a good "red" will be roughly the same.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The debate between these two broad approaches, relativist and universalist, continues unabated and both approaches have faced challenges. Work in the universalist tradition (Regier, Kay, & Cook, 2005; Webster & Kay, 2007) has shown that color naming cannot be completely free in the way suggested by relativists. However, the universalist tradition has itself faced problems, with languages showing a diversity of semantic encoding which goes beyond a narrow number of universal foci (Lucy, 1997). A third approach (Jameson & D'Andrade, 1997; Regier, Kay, & Khetarpal, 2007, 2009) has resolved some of these tensions, essentially extending the universalist system but in a way that allows for shifting foci and varying boundaries. Jameson and D'Andrade argue that color naming represents optimal divisions in a bumpy color space, where any color name should maximize perceptual similarity within the category and minimize similarity outside the category. The color space division will be sensitive both to the structure of perceived color space, which is even bumpier when factors such as saturation are considered, and to the number of terms being provided for dividing up that space. Their system is essentially a well-formedness condition on color naming. Nevertheless, all these approaches assume that it is meaningful to talk about color naming and color practices as a linguistic phenomenon. This assumption has been strongly challenged on the grounds that color naming, in the sense of labeling particular clusters of hues, does not in fact reflect a coherent linguistic system at all, whatever its perceptual or cultural status (Lucy, 1997; Wierzbicka, 1996).

More recently, there has been increased interest in the diversity of naming, driven by socio-cultural and cognitive factors, moving away from the strict focus on basic color terms. When the full range of color naming practices is investigated, more diverse patterns emerge than are implied by the Berlin and Kay approach. For instance, languages often name colors with a term derived from the name of an object (Conklin, 1973; Malt & Majid, 2013; Wierzbicka, 2008). This is particularly true of sign languages, where iconicity is a prominent characteristic of sign formation. As Sverrisdóttir and Thorvaldsdóttir (2016) discuss in detail, even the basic color terms in Icelandic Sign Language display iconicity, either by direct association with an object associated with a color or by an association between the color sign and the sign for an object associated with that color. As shown in Figure 1, the sign for red is associated with the lips, the sign for pink with the cheeks, and the sign for purple with a black eye. As shown in Figure 2, the sign for green is derived from the sign for cutting grass and the sign for yellow from the sign for cheese.

Both English and Icelandic also make extensive use of object-oriented color terms but they differ in the formal mechanisms used: in English, object names are simply used with zero derivation (*salmon, lemon, orange*); in Icelandic, the object name must be compounded either with a basic color term (*sítrónugulur* 'lemon yellow') or the word for *color/colored* (*húðlitaður* 'skin-colored'). As we discuss in

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

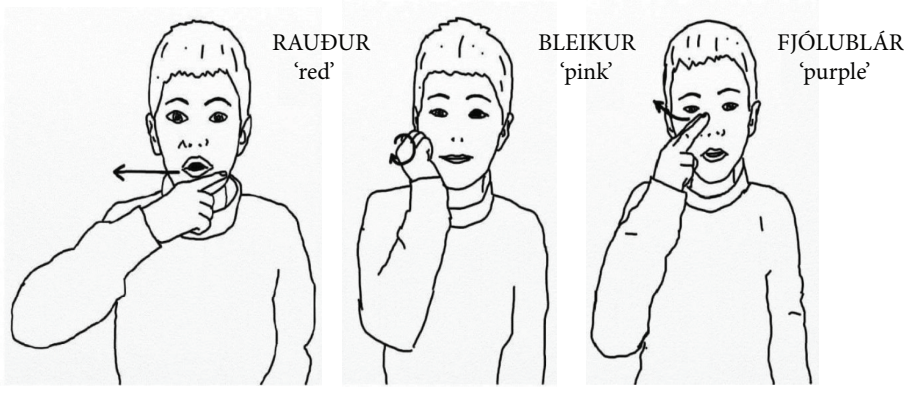


Figure 1. Three iconic color signs in ISL associated with the face

Figure 2. Two iconic color signs in ISL and the signs from which they derive

Section 4, both the formal mechanisms exploited to create color terms (e.g., compounding, affixation) and the semantic fields exploited to provide these terms can be of great interest in cultural and cognitive analyses of the color space.

The work reported here originated in the Evolution of Semantic Systems (EoSS) project (Majid et al., 2011), which aimed to assess the degree and nature of variation in semantic categories over space and time. Naming data was collected for 50 Indo-European languages for four domains: colors (attributes), body parts (parts), containers (objects), and spatial relations (relations). Use of fixed stimulus arrays allowed for comparisons across a uniform extensional space. The EoSS results for the Germanic languages are reported in a special volume of Language Sciences (Berthele, Whelpton, Næss, & Duijff, 2015; Majid et al., 2015; Vejdemo et al., 2015; Whelpton, Guðmundsdóttir Beck, & Jordan, 2015; Zimmermann, Levisen, Guðmundsdóttir Beck, & van Scherpenberg, 2015). The statistical analysis of

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

naming patterns in Majid et al. (2015, p. 15) shows that color terms, along with body parts, show the least variation in category extensions. Nevertheless, as Vejdemo et al. (2015) point out, the EoSS data also provides ample support for the diversity of naming at the micro-level. Our own analysis shows exactly this diversity of naming strategies within a broadly uniform pattern.

3. Method

The data was collected using EoSS standardized experimental protocols (Majid et al., 2011) and we include here the EoSS color data from Icelandic (ICE) and British English (BrE). This data was then supplemented with data from two other projects, elicited using the same protocols. As part of the *Heritage Language, Linguistic Change and Cultural Identity* project (PIs: Höskuldur Þráinsson and Birna Arnbjörnsdóttir), funded by the Icelandic Research Fund, data was collected from speakers of North American Icelandic (NAI) in Manitoba, Saskatchewan and North Dakota, as well as from speakers of North American English (NAE) in the same communities. As part of the *Colour in Context* project (PI: Matthew Whelpton), funded by the University of Iceland Research Fund, data was collected from Icelandic Sign Language (ISL) users.

3.1 Participants

The EoSS protocols require that all participants be native speakers of their respective languages and that the experimenter be a native speaker of the target language. Most participants in the main EoSS study were undergraduates with a mean age between 20 and 30. These protocols were adhered to as closely as possible but a number of issues must be mentioned, largely relating to the moribund status of North American Icelandic and the very small community of Icelandic Sign Language users.

North American Icelandic (NAI) is a heritage language which is dying out. The data was therefore collected from much older speakers (mean 78). However, all NAI participants had learned Icelandic at home as children, often as their only language before schooling. The aim was to collect North American English (NAE) from the same communities, in order to provide a cultural-linguistic context for the heritage language: the NAE speakers were therefore also older speakers (often caretakers of the NAI speakers; mean 67). Due to funding and practical constraints, the experimenter for NAI was a native speaker of Icelandic and the experimenter in NAE was a native speaker of British English.

The Icelandic Sign Language community is very small, comprising 250–300 signers (Thorvaldsdóttir & Stefánsdóttir, 2015; Sverrisdóttir & Thorvaldsdóttir,

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2016). As is common for deaf signers (who usually do not have deaf parents), our participants began acquisition later than hearing participants: of the 21 participants, 6 began by the age of 2, 10 by the age of 5, and 5 by the age of 12. The mean age of the group was 48. There is no evidence of regional variation in ISL but there is considerable generational variation (Sverrisdóttir & Thorvaldsdóttir, 2016). Data collection was contracted to the Communication Centre for the Deaf and Hard of Hearing; the experimenter was a fluent hearing signer. The experiment was conducted exclusively in sign language, but the participant information forms were filled out in Icelandic (which all participants knew).

Table 1 shows information on the participants. Given the variation in age, the overall uniformity of naming practices reported in Section 4 is particularly striking.

Table 1. Participant information

Language (Project)	Participants (F)	Mean Age	Researchers
Icelandic (EoSS)	21 (10)	29	Matthew Whelpton; Þórhalla Guðmundsdóttir Beck
British English (EoSS)	20 (9)	22	Linnaea Stockall; Euphemia Snell
North American English (Heritage)	51 (36)	67	Matthew Whelpton; Þórhalla Guðmundsdóttir Beck
North American Icelandic (Heritage)	30 (19)	78	Íris Edda Nowenstein; Matthew Whelpton; Þórhalla Guðmundsdóttir Beck
Icelandic Sign Language (Color in Context)	21 (14)	48	Kristín Lena Þorvaldsdóttir; Matthew Whelpton; Þórhalla Guðmundsdóttir Beck

3.2 Data elicitation

Following the EoSS protocol, the color stimuli comprise 84 Munsell colors. Four colors are achromatic (black-gray-white). The remaining 80 vary in hue, brightness, and saturation. There are 20 equally spaced hues at 4 degrees of brightness. Saturation varies so that colors are generally at the maximal possible chroma for that point in the color space. Only Munsell certified color sheets were used. As light conditions critically affect perception, colors were presented under a daylight bulb simulating the spectral range of full daylight.

There were two elicitation tasks: a color naming task and a focal color task. In the color naming task, participants were presented with the 84 color chips in a single randomized order and asked to name the color (free naming). In the focal color task, participants were presented with an array displaying the 84 colors and then asked to point to the best example of colors from a list of basic color terms. The same

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

standard list of basic color terms was used for both varieties of English; the list for Icelandic Sign Language was taken from Rannveig Sverrisdóttir and Kristín Lena Thorvaldsdóttir´s (2016) detailed study. No standard list was available for Icelandic, so a rapid listing test was run to determine a salient set of color terms (cf. Berlin & Kay, 1969; Corbett & Davies, 1997). Given constraints on data collection with the heritage community of North American Icelandic speakers, it was not possible to run such an elicitation test; the list for Icelandic was used. The consequences of this decision are discussed in Section 4.1.2. All five languages had the eleven basic color terms listed by Berlin and Kay (Berlin & Kay, 1969; Kay et al., 2011).

In addition to the two elicitation tasks, all participants took a color blindness test to allow the exclusion of results from color-blind participants. The British English and Icelandic participants took the Waggoner (2002) color blindness test. The North American and Sign Language participants took the 10-plate Ishihara test for color blindness (The Isshinkai Foundation, 2005). The experiment was recorded (video recorded for the sign language participants; audio recorded for the others) and the responses coded afterward.

3.3 Coding

Each participant's complete response to each stimulus was transcribed into a spreadsheet as the "full response" and, in each language, the researcher then coded these full responses for the "main response". In most cases, the coded main response is the transparent morphological and semantic head of the noun used to name the stimulus. If a participant responded "whisper pink, very light pink", then *pink* was the coded main response because *pink* is the head of the compound *whisper pink* as well as the head of the phrase *very light pink* and the semantic relation in both cases is transparent: whisper pink is a kind of pink and very light pink is a kind of pink (hyponym relation).

In some cases, as we will see in more detail in Section 4, the choice of a coded main response was problematic, usually because the relationship between the nominal in the full response and the morphological head noun was not (entirely) semantically transparent. For instance, in Icelandic, the term for the conventional translation for *orange* is *appelsínugulur* which is literally *orange yellow*. However, most speakers no longer accept that *appelsínugulur* is a kind of *gulur* 'yellow', treating the compound as a distinct color term. In this case, *appelsínugulur* was coded as a main term in its own right and special attention was paid in the analysis to see whether it behaved as an independent color term or not.

Following the standard conventions, ISL signs were coded using Icelandic labels in capital letters (to indicate that it is a sign label not an Icelandic lexeme), e.g.,

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

RAUÐUR 'red'. Variant forms of signs were distinguished by a number index, e.g., GULUR2 'yellow2'. Terms that were finger-spelled words were coded with punctuation, e.g., L.I.M.E. GRÆNN 'lime green', where L.I.M.E. is a finger-spelled English word and GRÆNN is a sign. Finger-spelling was generally used only for the first part of a sequential compound, especially with an English or loanword, and was not frequent. Only one sign involved initialization, where the form of the sign is based on the first letter of the written word: TÚRKIS 'turquoise'.

4. Results

4.1 Statistical analysis

4.1.1 *Color naming task*

Following Majid et al. (2015) and Malt et al. (1999), we conducted Pearson correlations on similarity matrices for the languages, as well as a principal component analysis.[1]

A similarity matrix represents groupings of color stimuli by naming. For each individual participant, the response for a pair of stimuli was compared, e.g., the response to stimulus 5 and the response to stimulus 62. If the same main response was given to name both stimuli, the value of 1 was assigned; otherwise, a value of 0 was assigned. Therefore, for each individual participant there was an 84 × 84 matrix representing which stimuli were grouped under the same name and which were not. The matrices for participants in a language were then averaged to give a value between 0 and 1, representing the proportion of participants in that language who assigned each pair of stimuli the same name. Simple Pearson correlations were then run on each pair of languages within a particular semantic category to see how similarly two languages divided up the extensional space lexically. All the languages are extremely highly correlated (mean 0.92, at $p = .01$), as shown in Table 2.

1. A preliminary statistical analysis was conducted on a snapshot of the data from May 2016, by the Social Sciences Institute of the University of Iceland (Guðný Bergþóra Tryggvadóttir & Guðbjörg Andrea Jónsdóttir, 2016); we would like to thank the authors for their invaluable analytical work and discussions. The analysis presented here is based on a snapshot of the data from September 2017. The analysis is conducted in R version 3.3.2 (2016-10-31). We would like to thank Michael Dunn, Joe Jalbert and Helgi Guðmundsson for their help with the R analysis. All errors and misunderstandings remain solely ours.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Pearson correlations

	BRE	ICE	ISL	NAE	NAI
BRE	1.00	0.90	0.90	0.94	0.89
ICE	0.90	1.00	0.94	0.91	0.91
ISL	0.90	0.94	1.00	0.93	0.92
NAE	0.94	0.91	0.93	1.00	0.94
NAI	0.89	0.91	0.92	0.94	1.00

This is particularly striking given the wide variation in age of participants for the five languages and in the typological diversity introduced by the inclusion of Icelandic Sign Language (ISL). To the extent that the small variation between language pairs suggests anything, it is interesting to note that the top end of the ranking includes languages sharing a cultural context: Icelandic and Icelandic Sign Language (0.94), typologically diverse but sharing a common cultural environment; and North American Icelandic and North American English (0.94), different languages spoken by members of the same communities. Language contact in a shared cultural environment would appear to be important here. Nevertheless, British and North American English also rank highly (0.94).

The principal component analysis identifies a small number of components and quantifies how much of the variation in the data can be attributed to each component. It shows that 93.6% of the variance was accounted for by a single component and only the first component received an eigenvalue score greater than 1. Statistically, there is relatively little variation in naming patterns, despite cultural and typological differences, and despite wide variations in age.

4.1.2 *Focal color task*

As summarized in Section 3.2, participants also performed a focal color task in which they were presented with all 84 colors in a fixed 4-by-21 array and asked to point to the best example of a series of basic color terms. Each of the 21 columns represented a separate hue and each of the rows a different degree of brightness, with the first column being achromatic (white – gray – black). A participant's response is therefore a vector: for instance, one might expect white to be (1,1) and black to be (4,1), i.e., the first and the fourth row of the first column, respectively. The vector for all participants in a language can then be averaged to find the center of gravity for that color in the color space. Figure 3 shows the centers of gravity for eleven basic color terms in the five languages.

Overall, there is a remarkable consistency in the distribution of color terms. Most of the potentially anomalous results relate to methodological issues. The first and most obvious is the placement of *bleikur* 'pink' for North American Icelandic

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

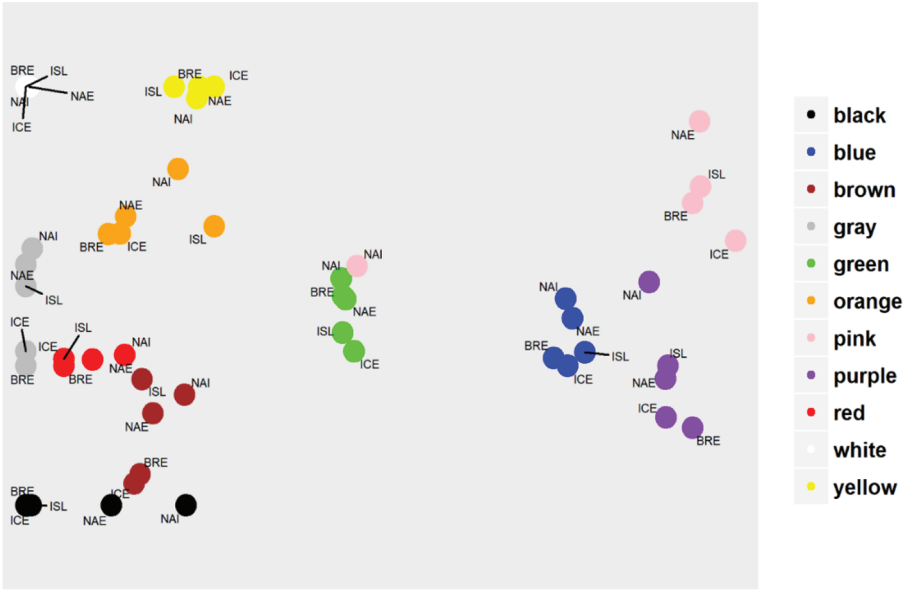


Figure 3. Center of gravity for focal color selections

(NAI) in the center of the array, far from the centers in the other four languages. This relates to the methodological issue mentioned in Section 3.2: due to practical constraints, speakers of North American Icelandic were presented with the basic color list for Icelandic, which included the terms *bleikur* 'pink', *appelsínugulur* 'orange', and *fjólublár* 'purple'. As the results of the naming task make clear (see Section 4.2.2 and the discussion in Guðmundsdóttir Beck & Whelpton (2018)), these relatively new standard color terms in Icelandic did not establish themselves in North American Icelandic, where the equivalent English terms have been taken in as loan words. For all three of these terms, the most common response was a non-response (error). For those who did answer, the compound structure of *appelsínugulur* 'orange' (literally, yellow like the orange fruit) and *fjólublár* 'purple' (literally, blue like the violet flower) provided a sufficient clue to guess the right area (notice though that the NAI response for orange tends towards the yellow area). However, the term *bleikur* is an Old Norse term meaning pale or faded, which has shifted semantically to become the standard term for pink. It is entirely opaque. Some of those who responded knew the term and pointed to the top right quadrant; some of those who did not guess on the basis of the form that it meant black and pointed to the bottom left quadrant: the center of gravity is therefore in the center of the space.

Another interesting anomaly is the difference between *white* and *black*. The term *white* provides the tightest cluster as the top left cell was the only response.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

However, for *black*, speakers of North American English did not choose the darkest achromatic cell in the bottom left, drifting rather into the darkest brown/red area; speakers of North American Icelandic drifted even further.

Notwithstanding the aforementioned, there is little to suggest significant variation in focal color placement. Statistical analysis of both free color naming and focal color selection therefore suggests a largely uniform pattern. To see the kinds of microvariations in which we are interested – and in particular, to understand the crucial methodological issue related to analysis based on transparent semantic heads and basic color terms – one must consider the main responses and then the full responses from which they are derived.

4.2 Main terms and dominant terms

4.2.1 *Main coded terms*

As discussed in Section 3.3, each full response was coded for the main term, which was the transparent semantic head term describing the color. If there was no response or the response could not be understood as describing the color, then an error was coded. For each color chip, the dominant term was the main term used more often than the others to describe it (if there was a tie between more than one term, then both are considered as dominant terms for that color). Table 3 shows the number of main terms, dominant terms, and errors for each language, as well as object-oriented terms, to which we turn shortly.

Table 3. Main terms (± object-oriented), dominant terms and errors for each language

	Main	Dominant	Error	Main (Not Object-Oriented)	Main (Object-Oriented)
BRE	54	12	0	17	37
ICE	26	11	2	17	9
ISL	52	11	0	23	29
NAE	93	13	1	20	73
NAI	63	11	16	26	37

The number of main terms varied significantly from language to language. The Icelandic speakers were by far the most conservative, using only 26 main terms, and the North American English speakers were by far the most elaborate, using 93 main terms. This gives a sense of the degree of variation that is hidden beneath the broad uniformity shown by the statistical analysis. However, it is not clear that this variation reflects diversity in color categories per se, as it relates to an interaction of the coding strategy with a typological difference between English and Icelandic. English speakers freely use object terms such as *mint* to name colors. Icelandic

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

does not allow this and, instead, the object must be compounded with a traditional color term, e.g., *myntugrænn* 'mint-green' or with the general term for color, e.g., *ferskjulitur* 'peach-color'. The English object name (*mint*) is coded as the main term in its own right but the Icelandic compounded term is coded by its transparent semantic head morpheme as *grænn* 'green'; only the compounds based on *litur* 'color' or *litaður* 'colored' are retained. When all these object-oriented terms are removed from the list of main terms in all the languages, then the languages are more similar (see Table 3, Main (Not Object-Oriented)).

The range and diversity of the object-oriented terms, as opposed to the purely hue-based terms, underlines the dangers of characterizing the color vocabulary system purely in terms of specialized hue vocabulary, as encouraged by Berlin and Kay's methodology. This is especially important for sign languages, where iconicity is so prevalent; but it is also true for languages like English, as the NAE data shows. We will see in Section 4.3.3 that this difference in object-orientation also has an effect on the saliency of different semantic features in compounding and modification.

4.2.2 *Dominant terms*

Dominant terms (the main term used more often than others to describe a particular stimulus) also provide an interesting perspective on variation between the languages. Table 3 showed that there was considerably less variation in the number of dominant terms than main terms. Table 4 shows these dominant terms, listed by rough translation equivalent.[2] Where two terms are used by an equal number of participants for the cell, both terms are treated as dominant.

Ten dominant terms are shared by all five languages. Black is not one of them (and many participants complained in the focal color task that there was no good example of black). Speakers of both British English and (European) Icelandic (younger) preferred to call the darkest color chip *gray*/*grár* rather than *black*/*svartur*, but the North American speakers and sign language speakers (older) used the equivalent of *black* as expected for this color chip. In addition, the speakers of both North American and British English used the term *peach* twice as a dominant term (and for the same color chips). For one of these, (European) Icelandic speakers used *húðlitur* 'skin-color'. Zimmerman et al. (2015, p. 42) discuss the displacement of *skin color* or *flesh color* by *peach* in English as the result of pressure from the civil rights

2. Obviously, the rough translation equivalence does not mean semantic coextension – exactly the factor that it is at stake here. However, this is a useful way for the readers to orient themselves to the data and the statistical analysis provided in the last section suggests that this is not an egregious simplification.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Dominant color terms, listed by rough translation equivalent (number of color stimuli in extension, shown in brackets)

ICE (11 of 26)	NAI (11 of 63)	NAE (13 of 93)	BRE (12 of 54)	ISL (11 of 52)
grænn (23)	grænn (23)	green (26)	green (26)	GRÆNN (26)
blár (17)	blár (16)	blue (15)	blue (16)	BLÁR (15)
fjólublár (12)	purple (9)	purple (10)	purple (12)	FJÓLUBLÁR (10)
bleikur (9)	pink (7)	pink (11)	pink (11)	BLEIKUR (13)
brúnn (7)	brúnn (4)	brown (4)	brown (5)	BRÚNN (5)
gulur (6)	gulur (5)	yellow (4)	yellow (5)	GULUR (4)
appelsínugulur (4)	orange (4)	orange (4)	orange (3)	APPELSÍNUGULUR (2)
rauður (3)	rauður (8)	red (3)	red (2)	RAUÐUR (4)
grár (3)	grár (2)	gray (2)	gray (3)	GRÁR (3)
hvítur (1)	hvítur (3)	white (2)	white (1)	HVÍTUR (2)
	svartur (3)	black (1)		SVARTUR (1)
		peach (2)	peach (2)	
húðlitur (1)				
		turquoise (1)		
			maroon (1)	

movement and the general trend towards multiculturalism. Icelandic retains the original skin terminology as the dominant term. In addition, the two Englishes each introduced one additional term but in a different area of the color palette. North American English introduced *turquoise* as the dominant term in the blue-green area, which is a well-known composite color (Kay et al., 2011; Kay, Berlin, & Merrifield, 1991; Kay & McDaniel, 1978). British English introduced *maroon* for one color in the dark brown-red area. As noted in Section 4.1.2, the North American Icelandic list of dominant terms includes both Icelandic and English terms. The English terms (*pink*, *purple*, *orange*) are used instead of the European Icelandic terms (*bleikur*, *fjólublár*, *appelsínugulur*), which had not established themselves as standard color terms in Icelandic at the time of the emigration to North America. Overall, a broadly uniform pattern shows interesting signs of variation in salient areas for lexicalization.

We now turn to a broader analysis of the full responses. First, we consider differences in the formal mechanisms exploited to enrich color descriptions in the five languages, in particular, derivational morphology and compounding. Then we offer a broad analysis of the semantic features encoded in the enriched color descriptions. We see that the formal mechanisms exploited are fairly diverse, while the overall range of semantic features is remarkably stable, with one interesting exception that shows the interaction of formal mechanisms and semantic range.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.3 Morphology and semantics of full responses

In this section, we consider the original full responses (pre-coding), in particular, the use of derivational morphology, compounding and modification to elaborate the description of the main color term. This will bring us back to the issue of object-oriented color naming and the distorting effect that removing such terms can have on a discussion of the domain.

4.3.1 *Derivation*

There is considerable variation in the use of derivational morphology, partly for typological reasons but also, simply, in the speakers' practice. Speakers of both British and North American English made productive use of the suffixes *-y* and *-ish* to mark peripheral and intermediate colors, e.g., *bluey* and *bluish*. Although both suffixes mark marginality, we assume that their core function is slightly different.

The suffix *-y* is an implicit comparative, so *bluey* essentially means *blue-like* and marks marginality because the color is not actually blue but is qualitatively like blue. It is used in our data not only with color terms like *blue*, but also with nouns associated with a visual quality. For instance, *nighty* is used presumably to indicate a quality of luminous darkness in the phrase "darky bluey nighty sky type thing, blue" by one speaker of British English.

The suffix *-ish*, on the other hand, is a hedge, indicating that the object is only a marginal member of the relevant category or has the property only to a slight extent. So *bluish* would mean *slightly blue* or *only marginally blue*. The two suffixes can occur together but always with *-ish* further from the root, as one might expect with a hedge suffix, e.g., "pinkish purplish" in our British English data (see Hay, 2002, for further discussion).

In principle, Icelandic also has the option of using suffixes such as *-leitur* '-like' for a similar purpose. However, suffixation never occurs in our data as a way of describing marginal colors or specifying comparative qualities. That function is served exclusively by compounding. This is true both in (European) Icelandic and North American Icelandic. The overriding role of compounding in elaborating semantic information is typical of the Icelandic data in EoSS more generally (see, for instance, Whelpton et al., 2015). We return to compounding in Section 4.3.2.

Icelandic Sign Language, like other signed languages, does not use derivational affixation in this way but rather simultaneous modulation of the performance of the sign, including changes to the manner of the gesture and accompanying facial expressions (cf. Sandler & Lillo-Martin, 2006). In this data, we found the use of such simultaneous modulations to mark the degree of commitment or prototypicality. Figure 4 shows the contrast between normal, sharp, and drawn modes.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

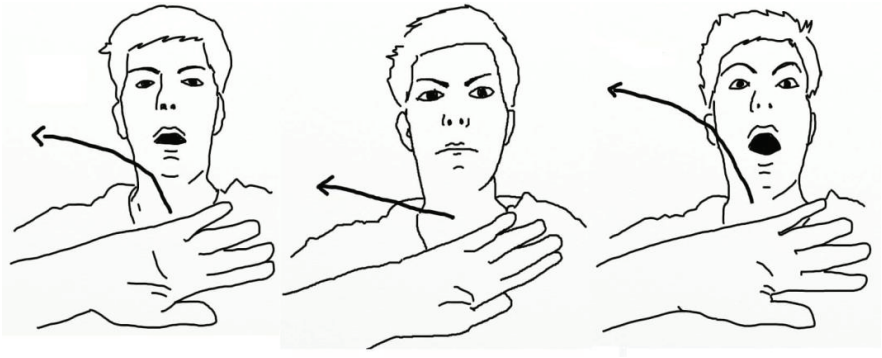


Figure 4. BLÁR[blá] 'blue': normal mode (unmarked) / sharp mode (commitment) / drawn mode (prototypicality)

The sharp mode involves a fast, short, direct movement, accompanied by lowered eyebrows and slightly squinted eyes. It implies certainty and commitment: e.g., "it's blue – period!". The drawn mode involves a slow, exaggerated movement, accompanied by raised eyebrows and wide eyes. It implies extreme prototypicality: e.g., "it's reeeeeeeeally blue!".

There is, therefore, considerable variation in whether and how the languages use derivational morphology to elaborate the description of color. Both varieties of English use it to mark intermediate and marginal colors. Icelandic Sign Language uses simultaneous modulation of the sign to indicate certainty and prototypicality. Neither variety of Icelandic uses it at all, favoring instead the use of compounding to which we now turn.

4.3.2 *Compounding*

Speakers of all languages used compounding productively, though it was particularly productive in the Icelandic data. The prominence of compounding in the Icelandic data is partly for typological reasons, as already discussed in Section 4.2.1. While English speakers freely use object terms such as *mint, mustard, navy, salmon, coral, jade, aubergine, charcoal,* and *grape* to name colors, this is generally not allowed in Icelandic. The only instance in our data was the use of *sítróna* 'lemon' on its own to describe one color. In all other instances, the object term was compounded with a traditional color term, or with the term color or colored. Examples with *litur* 'color' include *blóðappelsínulitur* 'blood orange color' and with *litaður* 'colored', *húðlitaður* 'skin-colored'. Examples with compounded objects include *vínrauður* 'wine-red', *mosagrænn* 'moss-green', *myntugrænn* 'mint-green', *himinblár* 'sky-blue', and *karrígulur* 'curry-yellow'. This obviously imposes a requirement on the speaker

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

to classify the color according to a standard color term, a requirement not imposed by the English naming strategy. In two cases, the participant was unable to think of an appropriate term but still formed a compound using the pronominal *something* as a head instead of a particular color term: *apríkósueitthvað* 'apricot-something', *sinnepseitthvað* 'mustard-something'.

North American Icelandic largely employs the English strategy, naming colors directly with objects, rather than compounding. The object names are primarily in English (e.g., *lime*, *gold*, *peach*), but also in Icelandic (*rúsína* 'raisin', *fjóla* 'violet (flower)'). There are examples of Icelandic compounds such as *rjómalitur* 'cream-color' and *gulllitaður* 'gold-colored'. And like English, compounded object terms also occur, e.g., *lime green, sky blue, coral pink*.

Icelandic Sign Language used both the English strategy (e.g., HÚÐ 'skin'; RAUÐVÍN 'red wine') and the Icelandic strategies, both with color headwords (VÍNRAUÐUR 'wine red') and much more commonly compounding with the term for color (e.g., VÍNLITUR 'wine color').

Another factor concerns color-oriented adjectives such as *light*, *dark*, and *pale*. In English, these are added as adjectival modifiers. In Icelandic they are usually compounded in their root form, rather than appearing as independently inflected attributive adjectives, e.g., *dökkgulur* 'dark-yellow' instead of *dökkur gulur* 'dark.masc.sg.nom yellow'. The compounds in Icelandic can therefore be quite elaborate, e.g., *földökkgrænblár* 'pale dark green blue' and *ljóshvítneongrænn* 'light white neon green'. Icelandic Sign Language can also use a sequence of signs for this information, in which case the response is coded as a sequential compound, following the conventions of Icelandic orthography.

As Guðmundsdóttir Beck and Whelpton (2018) discuss, the Icelandic strategy of compounding with a general color term head has a significant effect on the extensional range of terms such as *grænn* 'green' and *blár* 'blue' in Icelandic as opposed to the two Englishes. If the extensional range of the green and blue terms is considered, where there is 80% consensus between participants in the naming, then the terms in Icelandic extend right up to each other, whereas in the two Englishes there is a gap between the two domains. This is because the contact area in the Englishes has competing terms such as *teal* and *turquoise*, whereas in Icelandic we find compounds such as *blágrænn* 'blue-green' or *sjávarblár* 'sea-blue', which are then coded as <blár; grænn> and <blár>, respectively. The methodological choice to code for transparent semantic heads therefore has a distorting effect when the extensional divisions in the Icelandic and English are compared.

In Section 4.3.3, we look more closely at the kinds of semantic information used to enrich the color description by modifying the main term. Once again, we see a broadly similar pattern but with one significant exception that relates to our discussion of compounding.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.3.3 *Semantic feature analysis*

In this section, we consider the semantic contribution of terms that modify the main color term. The data were recoded to tag all modifiers including compounded items (*woodland green*), attributive modifiers (*light green*), and phrasal modifiers (*blue like a midnight in Saskatchewan*). The modifiers were then reviewed by the researchers and classified by intuition according to the kind of semantic information they contributed. The semantic features used in the analysis are shown in Table 5.

Table 5. Semantic features

Semantic feature	Definition	Example
Hue	another color term is used to modify the head term	green blue
Non-Hue	this is used rather broadly to cover non-hue-related properties, such as brightness (*light-dark*) and reflectivity (*shiny*)	light-dark, bright, neon, shiny
Property of Object	the color of an object is used to specify the hue	sky blue
Related to Object	a color associated with an object is used to specify the hue, even though the thing itself does not have that color	party pink
Typicality	these terms relate to the prototypicality (*true*, *standard*) or marginality (*funny, off*) of the color	true, standard funny, off
Evaluative	these terms express subjective value judgments concerning the color	beautiful, ugly

With all but one rather interesting exception, all the languages use the vast majority of modifiers to add information concerning non-hue related aspects of the color and to comment on the degree to which the color is a typical or non-typical example of that color. Table 6 shows proportions of modifiers expressing different semantic features in each of the languages.

Table 6. Proportions of modifiers coding semantic features

Semantic type	BRE	ICE	ISL	NAE	NAI
Non-Hue	30.6%	44.7%	57.4%	38.2%	34.1%
Typicality	39.1%	6.7%	20.2%	31.4%	45.5%
Hue	14.1%	19.2%	12.4%	12.7%	10.9%
Property of Object	12.2%	23.6%	6.9%	12.2%	7.7%
Related to Object	3.3%	2.9%	0.9%	3.4%	0.9%
Evaluative	0.7%	2.9%	2.2%	2.0%	0.9%

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Icelandic, Icelandic Sign Language, and North American English favor non-hue information, while British English and North American Icelandic favor typicality information. Icelandic is the only one of the languages which does not have both of these semantic features in the top two positions: indeed, typicality is only the fourth of the six features. Instead, Icelandic favors Property of Object and Hue information in the second and the third place. This is significant given our discussion on the role of compounding in Section 4.3.2.

Recall that instead of using separate lexical terms such as *teal* and *turquoise* for intermediate colors, Icelandic uses compounds such as *blágrænn* 'blue-green' or *sjávarblár* 'sea-blue'. Modifiers naming another hue and modifiers naming objects that bear that hue are therefore much more prominent in the Icelandic data than in the other languages, which make much freer use of bare object names (*salmon, peach, leather*) and distinct lexemes (*teal, maroon, mauve, turquoise*) to characterize such intermediate colors. This emphasizes the long-standing warning from relativists that analysts must be cautious of reductive techniques which obscure diversity in the domain, whether that be coding for transparent semantic heads or a focus on basic color terms.

5. Conclusion

The initial statistical analysis presented a picture of nearly monolithic uniformity. This, in itself, is of significant interest when one considers color naming in a broader context, comparing it to variations in other semantic categories (Majid et al., 2015; Whelpton, 2018). Nevertheless, within this broadly similar system, there is ample evidence of the sort of microvariation that relativists have emphasized. Even in the dominant terms which carve up the main color space, there are suggestive differences. It is clear that *peach* has introduced itself into the core naming vocabulary of English in a way that has not happened in either Icelandic or Icelandic Sign Language, though Icelandic does pick out part of the same area as *húðlitur* 'skin-color'. The two Englishes both introduce an extra term in the basic division of the color space but diverge in the area of the palette they focus on for differentiation: North American English privileges the blue-green area with the term *turquoise*; British English privileges the dark red-brown area with *maroon*.

There are also dramatic differences in naming strategies. Both Englishes use object names as color terms, a strategy excluded in Icelandic, which must compound either with a general color term or with the term *color* itself; Icelandic Sign Language uses a mix of the two strategies. In line with contemporary criticisms of Berlin and Kay's methodology, the removal of such object-oriented terms has a clear distorting effect on the comparison of these color systems. Furthermore,

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

both Englishes make extensive use of derivational morphology, but Icelandic relies exclusively on compounding and modification, even though derivational morphology is in principle available; Icelandic Sign Language does not have derivational morphology of this kind but does make use of sign modulation.

The overall picture is one of highly constrained macrovariation: within the bounds set by universal cognitive-perceptual constraints, there is an enormous amount of microvariation, suggesting sensitivity to cultural and linguistic factors.

Acknowledgment

This work stems from the Evolution of Semantic Systems (EoSS) project which received financial support from the Max Planck Gesellschaft. The EoSS data for Icelandic was collected by the authors. We would like to thank Linnaea Stockall for the use of the British English data. Data for North American English and North American Icelandic was collected as part of the *Heritage Language, Linguistic Change and Cultural Identity* project (PIs: Höskuldur Þráinsson and Birna Arnbjörnsdóttir) funded by the Icelandic Research Fund. Data for Icelandic Sign Language was collected as part of the *Colour in Context* project (PI: Matthew Whelpton) funded by the University of Iceland Research Fund. We would like to thank Rannveig Sverrisdóttir for her input on Icelandic Sign Language and the two anonymous reviewers for their invaluable comments. All views, errors and omissions are ours solely.

References

Arnbjörnsdóttir, B. (2006). *North American Icelandic – the Life of a Language*. Winnipeg: University of Manitoba Press.

Berlin, B. (1992). *Ethnobiological Classification: Principles of Categorization of Plants and Animals in Traditional Societies*. Princeton, NJ: Princeton University Press. https://doi.org/10.1515/9781400862597

Berlin, B., & Kay, P. (1969). *Basic Color Terms, Their Universality and Evolution*. Berkeley and Los Angeles: University of California.

Berthele, R., Whelpton, M., Næss, Å., & Duijff, P. (2015). Static spatial descriptions in five Germanic languages. *Language Sciences*, 49, 82–101. https://doi.org/10.1016/j.langsci.2014.07.006

Conklin, H. C. (1955). Hanunóo Color Categories. *Southwestern Journal of Anthropology*, 11(4), 339–344. https://doi.org/10.1086/soutjanth.11.4.3628909

Conklin, H. C. (1973). Color Categorization. *American Anthropologist*, 75(4), 931–942. https://doi.org/10.1525/aa.1973.75.4.02a00010

Corbett, G. G., & Davies, I. R. L. (1997). Establishing basic color terms: measures and techniques. In C. L. Hardin & L. Maffi (Eds.), *Color Categories in Thought and Language* (pp. 197–223). Cambridge University Press. https://doi.org/10.1017/CBO9780511519819.009

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Guðmdundsdóttir Beck, Þ., & Whelpton, M. (2018). Það besta úr báðum heimum – um litaheiti í vesturíslensku. In B. Arnbjörnsdóttir, H. Þráinsson, & Ú. Bragason (Eds.), *Sigurtunga – Vesturíslenskt mál og menning* (375–400). Reykjavík: Háskólaútgáfan.

Hay, J. (2002). From Speech Perception to Morphology: Affix Ordering Revisited. *Language*, 78(3), 527–555. https://doi.org/10.1353/lan.2002.0159

Jameson, K., & D'Andrade, R. (1997). It's not really red, green, yellow, blue: an enquiry into perceptual color space. In Clyde L Hardin & L. Maffi (Eds.), *Color categories in thought and language* (pp. 295–319). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519819.014

Kay, P., Berlin, B., Maffi, L., Merrifield, W. R., & Cook, R. (2011). *The World Color Survey* (1st edition). *Stanford*, California: Center for the Study of Language and Information.

Kay, P., Berlin, B., & Merrifield, W. (1991). Biocultural Implications of Systems of Color Naming. *Journal of Linguistic Anthropology*, 1(1), 12–25. https://doi.org/10.1525/jlin.1991.1.1.12

Kay, P., & McDaniel, C. K. (1978). The Linguistic Significance of the Meanings of Basic Color Terms. *Language*, 54(3), 610–646. https://doi.org/10.2307/412789

Lucy, J. A. (1997). The linguistics of "color." In C. L Hardin & L. Maffi (Eds.), *Color categories in thought and language* (pp. 320–346). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511519819.015

Majid, A., Jordan, F., & Dunn, M. (2011). *Evolution of Semantic Systems Procedures Manual*. Nijmegen: Max Planck Institute for Psycholinguistics.

Majid, A., Jordan, F., & Dunn, M. (2015). Semantic systems in closely related languages. *Language Sciences*, 49, 1–18. https://doi.org/10.1016/j.langsci.2014.11.002

Malt, B. C., & Majid, A. (2013). How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(6), 583–597. https://doi.org/10.1002/wcs.1251

Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus Naming: Similarity and the Linguistic Categorization of Artifacts. *Journal of Memory and Language*, 40(2), 230–262. https://doi.org/10.1006/jmla.1998.2593

Nida, E. A. (1959). Principles of translation as exemplified by Bible translating. In R. A. Brower (Ed.), *On translation* (pp. 11–31). Cambridge MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674731615.c3

Regier, T., Kay, P., & Cook, R. S. (2005). Focal Colors Are Universal after All. *Proceedings of the National Academy of Sciences of the United States of America*, 102(23), 8386–8391. https://doi.org/10.1073/pnas.0503281102

Regier, T., Kay, P., & Khetarpal, N. (2007). Color Naming Reflects Optimal Partitions of Color Space. *Proceedings of the National Academy of Sciences of the United States of America*, 104(4), 1436–1441. https://doi.org/10.1073/pnas.0610341104

Regier, T., Kay, P., & Khetarpal, N. (2009). Color naming and the shape of color space. *Language*, 85(4), 884–892. https://doi.org/10.1353/lan.0.0177

Rögnvaldsson, E., Jóhannsdóttir, K. M., Steingrímsson, S., Loftsson, H., & Helgadóttir, S. (2011). *Languages in the European Information Society – Icelandic*. Berlin: META-NET DFKI Projektbüro.

Sandler, W., & Lillo-Martin, D. (2006). *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139163910

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sverrisdóttir, R., & Thorvaldsdóttir, K. L. (2016). Why is the SKY BLUE? On colour signs in Icelandic Sign Language. In U. Zeshan & K. Sagara (Eds.), *Semantic Fields in Sign Languages, Colour, Kinship and Quantification* (pp. 209–250). Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9781501503429-006

The Isshinkai Foundation. (2005). *Ishihara's Design Charts for Colour Deficiency of Unlettered Persons*. Tokyo, Japan: Kanehara Trading.

Thorvaldsdóttir, K. L., & Stefánsdóttir, V. (2015). Icelandic Sign Language. In J. B. Jepsen, G. De Clerck, S. Lutalo-Kiingi, & W. B. McGregor (Eds.), *Sign Languages of the World: A Comparative Handbook* (pp. 409–430). Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9781614518174-022

Vejdemo, S., Levisen, C., van Scherpenberg, C., Guðmundsdóttir Beck, Þ., Næss, Å., Zimmermann, M., Stockall, L., & Whelpton, M. (2015). Two kinds of pink: development and difference in Germanic colour semantics. *Language Sciences*, 49, 19–34. https://doi.org/10.1016/j.langsci.2014.07.007

de. Vos, C. (2011). Kata Kolok Color Terms and the Emergence of Lexical Signs in Rural Signing Communities. *The Senses and Society*, 6(1), 68–76. https://doi.org/10.2752/174589311X12893982233795

Waggoner, T. L. (2002). Quick six colour vision test pseudoisochromatic plates. *Colour vision testing made easy*. Good-Lite company.

Webster, M., & Kay, P. (2007). Individual and population differences in focal colors. In R. E. MacLaury, G. V. Paramei, & D. Dedrick (Eds.), *Anthropology of Color: Interdisciplinary Multilevel Modeling* (pp. 29–54). John Benjamins Publishing. https://doi.org/10.1075/z.137.05web

Whelpton, M. (2018). The variability of semantic categories: An experiment in extensional semantics. *Nordiske Studier I Leksikografi*, 14, 29–44.

Whelpton, M., Guðmundsdóttir Beck, Þ., & Jordan, F. M. (2015). The semantics and morphology of household container names in Icelandic and Dutch. *Language Sciences*, 49, 67–81. https://doi.org/10.1016/j.langsci.2014.07.014

Whorf, B. L. (1956). Science and linguistics. In S. Chase & J. B. Carroll (Eds.), *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf* (pp. 207–219). New York: MIT Press.

Wierzbicka, A. (1996). The Meaning of Colour Terms and the Universals of Seeing. In *Semantics, Primes and Universals* (pp. 287–330). Oxford University Press.

Wierzbicka, A. (2008). Why there are no "colour universals" in language and thought. *Pourquoi Il N'y a Pas D'universaux de La Couleur Dans Le Langage et La Pensée*, 14(2), 407–425. https://doi.org/10.1111/j.1467-9655.2008.00509.x

Zimmermann, M., Levisen, C., Guðmundsdóttir Beck, Þ., & van Scherpenberg, C. (2015). Please pass me the skin coloured crayon! Semantics, socialisation, and folk models of race in contemporary Europe. *Language Sciences*, 49, 35–50. https://doi.org/10.1016/j.langsci.2014.07.011

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Symbolic and cultural meaning of colors in phraseology

A cross-linguistic and cross-cultural study of Russian and German phraseological units

Branka Barčot and Anita Hrnjak
University of Zagreb

In the present study, we assume that color terms, as components of phraseological units, contribute to the meaning of the whole figurative unit in two ways: (1) through their denotative lexical meaning; (2) through their symbolic and/or cultural potential. We examine the phraseological subsystem of Russian and German, and then perform a comparative phraseological analysis. The main aim of this research is to systematically depict the main features of Russian and German color phraseology, to determine how much of the basic lexical meaning color components bring into the formation of the phraseological meaning, and to which extent the symbolism of colors has influenced the formation of that meaning.

Keywords: phraseological units, color components, denotative meaning, symbolic and cultural meaning, phraseological meaning, Russian, German

1. Introduction

Considering the ability of the human eye to distinguish colors as well as the psychological effects ascribed to colors, on the one hand, and considering the well-known fact that the way of conceptualizing reality inherent in any language is partly universal and partly nation-specific, on the other hand, in this study[1] we focus on figurative units (i.e., phraseological units (PUs)) that are probably the most language-specific units and thus commonly seen as problematic and unmotivated cross-culturally. Aware of the fact that color is not a universal human concept but a

1. This paper has been written as a part of the project *Application of phraseological theory in phraseography (No. 4054)*, supported by the Croatian Science Foundation.

https://doi.org/10.1075/sfsl.78.15bar
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

part of the visual semantics as a wider and more fundamental field, and cognizant that color is an important semantic molecule in many languages as it underlies (and is a part of) the meaning of words like *blue*, *red*, *yellow* and *pink* (Wierzbicka, 2005, p. 223), in this study we focus on the phraseological subsystems of one Slavic and one Germanic language, i.e., Russian and German. Therefore, our approach is contrastive in order to achieve one of our main aims: to show to which extent the speakers of these two languages share visual descriptions when using phraseological units with color components.

In this study, we assume that color terms, as components of PUs, contribute to the meaning of the whole figurative unit in two ways: (1) through their denotative lexical meaning; (2) through their symbolic and/or cultural potential (whereas examining contemporary color PUs can actually be seen as key to color symbolism in the two cultures). We thus examine the phraseological subsystems of Russian and German. First, we access them separately and then perform a comparative phraseological analysis. We distinguish between PUs that rest on natural color features (e.g., *чёрный как ночь* lit. black as night – *schwarz wie die Nacht* lit. black as the night; *красный как кровь* lit. red as blood – *rot wie Blut* lit. red as blood); PUs which contain a literal color designation but require a semantic interpretation (e.g., *как бык на красную тряпку* [реагировать, действовать и т. п.] lit. like bull on red rag [to react, to act etc.] – *ein rotes Tuch* für jmdn. *sein*[2] lit. to be a red rag to sb.), and PUs in which the color is used in a transferred sense (e.g., *молодо-зелено* lit. young-green – *grün hinter den Ohren sein* lit. to be green behind the ears; *голубая кровь* lit. blue blood – *blaues Blut in den Adern haben* lit. to have blue blood in the veins). In the analytical section, we merge these last two groups into one since in both cases color components contribute to the fact that those PUs can be seen as cultural signs. We take into consideration eleven color terms interpreted as basic according to the seven stages in an evolutionary sequence identified by Berlin and Kay in their theory of universals in color naming, one of the most influential theories in anthropological linguistics, originally formulated in 1969: *black*, *white*, *red*, *green*, *blue*, *yellow*, *grey*, *pink*, *purple*, *orange*, *brown*.

The data for the present study were extracted from the following Russian and German phraseological dictionaries: Mokienko, V. M. & Nikitina, T. G. (2007), *Bol'šoj slovar' russkih pogovorok*; Mokienko, V. M. & Nikitina, T. G. (2008), *Bol'šoj slovar' russkih narodnyh sravnenij*; Birih, A. K., Mokienko, V. M. & Stepanova, L. I. (1998), *Slovar' russkoj frazeologii. Istoriko-etimologičeskij spravočnik*; Duden (2002), *Redewendungen. Wörterbuch der deutschen Idiomatik*; Duden (2012), *Redensarten:*

2. The word order in all literal translations of German PUs differs (or might differ) from the original word order in German PUs due to different canonical word order in English and in German.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Woher sie kommen, was sie bedeuten; Walter, H. (2008), *Wörterbuch deutscher sprichwörtlicher und phraseologischer Vergleiche, Teil 1*. The gathered PU collection contains 128 Russian PUs and 116 German PUs.

The gathered data allow closer observation of structural types of PUs. Čermák (2007) distinguishes between three structural types of PUs, i.e., idioms: (1) *collocational idioms* (*смотреть сквозь розовые очки* на кого, на что lit. to look through pink glasses at sth., sb. – *durch die rosa<-rote> Brille sehen* lit. to look through the pink<-red> glasses; *проходить красной нитью* через что lit. to go like red thread through sth. – *roter Faden* lit. red thread); (2) *propositional or sentence idioms* (*чёрный день наступил* для кого lit. black day has come for sb.; *bei Nacht (in der Nacht) sind alle Katzen grau* lit. at night (in the night) are all cats grey); (3) *lexical idioms (cranberry idioms)* (*снежнобелый* lit. snowwhite – *schneeweiß* lit. snow-white; *rabenschwarz* lit. ravenblack). We add a fourth structural type: *similes*, since Russian and German phraseologies recognize them as PUs (Mokienko & Nikitina, 2008; Walter, 2008). In most structural types of PUs we cannot detect potential lexicalization patterns in color naming since color words can be viewed primarily as phraseological components, i.e., as part of the phraseological pattern forming the meaning of the whole PU in the two ways mentioned above. Nevertheless, lexicalization patterns in color naming can be observed on a phraseological level only through similes as a specific and very frequent structural type of PUs found both in German and Russian (*белый как снег* lit. white as snow > *снежнобелый* lit. snow-white, *белоснежный* lit. whitesnowy, *снежно-белый* lit. snowy-white, *снежный* lit. snowy – *weiß wie Schnee* lit. white as snow > *schneeweiß* lit. snowwhite).

Since the total number of PUs is too big (cf. Figure 2) for the analysis to be performed to its full extent in this chapter, the gathered data of Russian and German color phraseology will be represented in the analytical section only by several PUs per language for each color component. The descriptions of the PUs chosen to represent the Russian and German phraseological subsystems are supposed to show similarities, but also highlight differences between the two language subsystems. Thus, we look into the color-based phraseology that rests either on universal conceptual metaphors or on common symbolic meaning, e.g., on similar cultural connotations connected to color terms. On the other hand, a part of the analyzed data can be considered as primarily culture-bound and thus culturally motivated. The most representative PUs are divided and observed in the two groups mentioned above: (1) PUs based on a denotative lexical meaning of a color component; (2) PUs based on a symbolic and/or cultural potential of a color component.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. Basic color terms in Russian and German phraseology

2.1 Black

2.1.1 *PUs based on denotative lexical meaning*

The denotative lexical meaning of the adjective *black*, i.e., "the achromatic color of least lightness characteristically perceived to belong to objects that neither reflect nor transmit light" (Merriam-Webster Dictionary), is preserved in numerous Russian and German similes. In such constructions, the black color of one thing is explicitly compared with the color of another thing that is supposed to be recognized and visualized as something generally common in that specific linguo-cultural community. Some of the most frequent similes with the component *чёрный*/*schwarz* in Russian and German are: *чёрный как ночь* (lit. black as night) – *schwarz wie die Nacht* (lit. black as the night); *чёрный как вороново крыло* (lit. black as raven's wing), *чёрный как ворон* (lit. black as raven) – *rabenschwarz*/*kohlrabenschwarz* (lit. ravenblack/coalravenblack). Several Russian PUs in which the color black appears as a component contributing to the phraseological meaning through its denotative lexical meaning can be seen as culturally specific. For example, in similes *чёрные как смородины* (lit. black as blackcurrants), *чёрные как вишни* (lit. black as cherries), and *чёрные как сливы* (lit. black as plums) a comparison is drawn between the dark color of human eyes and fruits that are actually just dark and not of a prototypically black color. All of them play a significant role in Russian folklore and cuisine.

The German PU *das schwarze Brett* (lit. the black board) is one of the few examples of figurative units in our collection of PUs that rests on natural color features and does not have the structure of a simile. Nowadays, it denotes a bulletin board, but *das schwarze Brett* was originally a writing surface made of black or dark grey slate stone on which one wrote with a chalk. Blackboards as objects are hard to find nowadays in the German linguo-cultural community, but this PU still denotes different bulletin boards that are not black.

2.1.2 *PUs based on symbolic and/or cultural potential*

The symbolism of black varies according to time and place. But, as the color psychologist Jill Morton (1997, p. 40) points out, idiomatic expressions with color components demonstrate significant associations, many of which reinforce the general psychological symbolism of specific colors. Symbolically the color black is the color of the unknown, the end of a cycle, emptiness, death, corruption, ominous forces and depression. On the other hand, black can symbolize power, sexuality, and sophistication. In religion, black symbolizes evil and it is the color of mourning in most Western cultures (Morton, 1997, pp. 35–36).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The following figurative units reflect the negative symbolism of the color black: *изобразить / изображать в чёрных красках* кого, что (lit. to paint sb., sth. in black colors) – etw. *schwarz in schwarz (in den schwärzesten Farben) malen* (lit. to paint sth. black in black (in the blackest colors)). The meaning of both PUs is 'to depict sb., sth. in a pessimistic manner'. The mentioned negative associations of black are hidden behind the next two German PUs: *alles durch die schwarze Brille sehen* (lit. to see everything through black glasses); *schwarz sehen* (lit. to see black) 'to be very pessimistic'. If one wants to lay stress on the black color of something in German, the following PU, which rests on the mysterious symbolism of the color black, can be used: *pechschwarz/so schwarz wie Pech* (lit. misfortune black/black as misfortune). It is the only simile in the gathered PU collection of the two analyzed languages that does not rest on the denotative lexical meaning of a color.

The Russian PU *чёрный год* (lit. black year) meaning 'a bad year full of problems and misfortunes', *чёрный день* (lit. black day) meaning 'hard times, misfortune', and *чёрный день наступил* <для кого> (lit. black day has come for sb.) meaning 'hard times have come for sb.' are all based on a similar mental image. Their semantics rests on a negative symbolic potential of the color black that is perceived as a symbol of bad luck and misfortune.

The illegal trade of goods or money is also associated with the color black in the phraseological subsystems of Russian and German. The PUs that demonstrate this are: *с чёрного хода* (paraphrase: through the backdoor entry), *чёрное дело* (lit. black business); *schwarzarbeiten* (lit. to work black), *der Schwarzmarkt* (lit. the black market), *Schwarzhandel treiben* (paraphrase: to operate on the black market) 'to be engaged in illicit business'. The association with the illegal action is present in the German PU *schwarzfahren* (lit. to travel/drive black) that has two meanings: (1) 'to travel without a valid ticket'; (2) 'to drive without a driving license'.

C. G. Jung considered black to be the dark side of personality and one of the first stages to be passed, while white is at the end stage, when perfection is reached. In this, he was in accord with the theories of the alchemists, for whom the *Great Work* originated from *blackness* (Chevalier & Gheerbrant, 1996, p. 96). From a religious and philosophical point of view, black represents self-denial and symbolizes evil forces (Morton, 1997, p. 36). Since the soul is considered to be the essence of a person and his nature, the connection between a soul and the black color as a symbol of evil clearly underlies the Russian PU *чёрная душа* <у кого> (lit. <sb. has> a black soul) and German *eine schwarze Seele haben* (lit. to have a black soul) 'to have a bad character, to be a bad person'.

The German PU *schwarze Zahlen schreiben* (lit. to write black numbers) refers to having money or profiting and doing well in business. The etymology of this PU rests on the bookkeeping practice where black numbers were used to denote positive income and red numbers financial loss (Duden, 2002, p. 889).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

There are several culturally specific PUs with the color component *чёрный* / *schwarz* that are language-specific and thus considered to be culturemes. We will illustrate this by giving one German and one Russian example. If one wants to point the finger of blame at someone in German, one can use the PU jmdm. *den Schwarzen Peter zuschieben (zuspielen)* (lit. to slip (to play) the Black Peter to sb.). *Schwarzer Peter* is one of the most popular card games for children; it is known in different countries, and the names differ slightly (e.g., *Old maid* in English). The German PU jmdm. *den Schwarzen Peter zuschieben (zuspielen)* comes from this card game. Whoever holds the card "schwarzer Peter" at the end of the game is the loser (Duden, 2002, p. 691).

A negative symbolism connected with a specific animal is crucial for understanding the Russian PU *чёрная кошка пробежала (проскочила)* между кем (lit. a black cat ran (jumped) between two people) meaning 'there has been an argument between two people'. The symbolic meanings and cultural connotations linked to black cats vary from positive to extremely negative. In Slavic folklore, black cats are symbolically associated with evil forces, witches, and demons. It is believed that all of them can assume the shape of an animal, most commonly that of a black cat. Therefore, in Slavic folklore, witches are always accompanied by a black cat or a black rooster (Afanas'ev, 2008, p. 1373). The expression *чёрная кошка пробежала (проскочила)* между кем points to an evil demon running or jumping between two people, causing them to become hostile towards one another (Afanas'ev, 2016, p. 233). Exactly this kind of an irrational belief about black cats plays a key role in understanding the mentioned PU.

2.2 White

2.2.1 *PUs based on denotative lexical meaning*

Adjectives *белый* and *weiß* as components of PUs contribute to the meaning of the whole figurative unit through their denotative lexical meaning by resting on natural color features of specific objects. Some of the comparisons on which they are based are common to both Russian and German phraseology. For example, the white color of an object, e.g., human skin or hair, is compared to the color of snow in the Russian PU *белый как снег* and its German equivalent *weiß wie Schnee* (lit. white as snow). In both analyzed languages, we find the comparison of a pale human face to the color of a wall: *белый как стена* (lit. white as wall) – *weiß wie die <gekalkte> Wand* (lit. white as the <limed> wall). Similes with the color white as a component are more frequent in Russian phraseology and some of them can be considered as culturemes. One of them is the simile *белый как репа* (lit. white as a turnip), in which white color (describing primarily human teeth) gets to be compared to the

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

color of a turnip, a vegetable that is very common and highly appreciated both in ancient and contemporary Russian cuisine.

2.2.2 *PUs based on symbolic and/or cultural potential*

Белый and *weiß* as components of PUs contribute to the meaning of the whole figurative unit through their symbolic and/or cultural potential as well. The symbolism of the color white associated with purity, virginity and innocence is present in the German PU e*ine (keine) weiße (saubere, reine) Weste haben* (lit. to have/not have a white (clean) vest) and jmdn. *weißwaschen* (lit. to wash sb. white). The meaning of the former is '(not) to be innocent', whereas the latter relates to the action of releasing someone from suspicion. In this case, a white vest represents an external sign of someone's good and uncorrupted character. It is interesting to note that this element of phraseological meaning might be associated with the fact that white was the color of candidates' robes when they ran for public office in Ancient Rome (*candidus* = 'white' in Latin) (Chevalier & Gheerbrant, 1996, p. 1108).

The inherently positive symbolism of the color white obviously plays a considerable role in PUs such as the Russian *белая ворона* (lit. a white crow) and its German equivalent *ein weißer Rabe* (lit. a white raven). They both refer to someone who is rare and exceptional, someone who stands out because of his/her exceptional characteristics. They are based on a mental image of a white crow/raven that stands out in a flock of black birds of the same species. An additional semantic interpretation is necessary because we must take into consideration the fact that white crows/ravens are extremely rare since they are affected by albinism as the result of a genetic mutation. The inherently positive symbolism of the color white is also important because it brings us to the conclusion that, in this case, being different is considered to be better.[3]

In some PUs, white color occurs primarily as a visual part of a mental image behind a PU containing a literal color designation, but this also requires a certain semantic interpretation. As an illustration, we can mention Russian and German PUs that are considered to be euphemisms because the white color component is associated with drunkenness and visual hallucinations that occur as a result of excessive drinking. Russian idioms *увидеть белого слона* (lit. to see white elephant) meaning 'get extremely drunk' and *до белых чертей* [пить, напиться, допиться etc.] (lit. till white devils [to drink, get drunk, etc.]) meaning 'to drink heavily, to get extremely drunk' are based on the perception that extremely drunk people

3. There is a second meaning of the Russian PU *белая ворона* that is opposite to the first one, i.e., it can be used to denote a disfavored or disreputable member of a group. In this case, the German equivalent is *schwarzes Schaf* (lit. black sheep).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

suffer from hallucinations during which they see white elephants or white devils. In German phraseology, hallucinations are associated with seeing white mice: *weiße Mäuse sehen* (lit. to see white mice) meaning 'to be extremely drunk'. In all three cases, the semantic interpretation includes the speaker's perception of white as a very rare natural color of elephants and mice, as well as the common concept of devils not being white but black. In fact, white is an inherently positive color associated with purity, goodness, and heaven, while black, on the contrary, is considered to be a color symbolically linked to death and evil. Understanding the Russian idiom *до белой горячки* [пить, напиться, допиться etc.] (lit. till white fever [to drink, get drunk, etc.]) used in the meaning of 'to drink heavily, to get extremely drunk', implies that the speaker is aware of the psychotic condition typical of withdrawal in chronic alcoholics, involving tremors, hallucinations, anxiety, and disorientation called *delirium tremens* in English and *белая горячка* (lit. white fever) in Russian.

The Russian PU *сказка про белого бычка* (lit. fairy tale about a little white ox) denotes a constant repetition, but this meaning cannot be linked to a mental image behind the PU unless we are aware of its etymology. The PU refers to a story that grown-ups tell to tease children when they annoy them with their constant nagging to tell them a fairy tale. Actually, it is not a real story but some kind of a word game. One must also take into consideration the fact that *сказка про белого бычка* can literally be translated as 'a fairy tale for a little invisible ox' and interpreted as a story told to child where a little creature makes annoying noises that makes you tell him a fairy tale (Birih, Mokienko, & Stepanova, 1998, p. 525).

A considerable number of Russian and German PUs are based on a contrast between the colors white and black. Some of them are based on a symbolic contrast, while others include a visual contrast between these two colors.

The contrast between the inherently positive symbolism of white and the inherently negative symbolism of black can be tracked in PUs such as: *называть чёрное белым* (lit. to name black as white; to present black as white) meaning 'to portray something negative as positive', *принимать белое за чёрное* (lit. to take black for white) meaning 'to make a mistake by considering something good to be bad', and *разделять (делить) на белое и чёрное* (lit. to divide into black and white) meaning 'to distinguish clearly between good and bad'.

The Russian PU *чёрным по белому* (lit. with black on white) and its German equivalent *schwarz auf weiß* (lit. black on white), meaning 'in written form, so that one can rely on it as proof', can be seen as a good example of the role of visual contrast between white and black, i.e., between the two opposite sides of a color spectrum in the semantics of a PU.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.3 Red

2.3.1 *PUs based on denotative lexical meaning*

Красный and *rot* as components of PUs contribute to the meaning of the whole figurative unit through their denotative lexical meaning. The red color of an object is in both languages compared to the color of blood (*красный как кровь* lit. red as blood – *rot wie Blut* lit. red as blood), while the red color of a human face is compared to the color of a crab or a boiled crab (*красный как <вареный> рак*, (lit. red as <cooked> crab), *rot wie ein Krebs* (lit. red as a crab)). An interesting cultural element can be observed in Russian similes *красный как свёкла* (lit. red as beetroot) and *красный как бурак* (lit. red as beetroot). Once again, Russian phraseology reflects the typical Russian everyday life by using a vegetable that is very common and highly valued in both ancient and contemporary Russian cuisine.

2.3.2 *PUs based on symbolic and/or cultural potential*

In the analyzed collection of Russian and German PUs, there are not as many PUs with the color component *red* based exclusively on color symbolism as one would expect, taking into account the particularly rich symbolism of the color red.

Красный and *rot*, as components of PUs, contribute to the meaning through its symbolic and/or cultural potential in such PUs as those found in a series of Russian and German figurative units that refer to fire. The Russian PU *красный петух* (lit. red rooster) denotes fire, while *пустить/пускать красного петуха* (lit. to release a red rooster), *красного петуха на крышу посадить/сажать* and its German equivalent *den roten Hahn aufs Dach setzen* (lit. to put the red rooster on the roof) carry the meaning 'to set a fire'. Red is symbolically the color of fire, but in this case, one must take into account a mental image behind all of them because it involves a red rooster, an animal with an abundantly rich symbolism. Most cultures embrace a rooster as a solar symbol, while a red rooster is associated with the symbolism of fire and the ancient cult of the god of thunder. As such, it finds its place in Slavic, Baltic, and German folklore and languages. In these cultures, red roosters have been sacrificed in ritual killings as part of a religious ritual to appease or maintain favor with the gods of fire. The same cultures cultivated a belief according to which a red rooster descends from the heavens during a storm, accompanied by lightning, and sets fire to rooftops and houses. An evidence of the mentioned cult may be observed in Slavic material culture where it is customary to put weather vanes or lightning conductors in the shape of a metal rooster on top of a roof (Birih, Mokienko, & Stepanova, 1998, pp. 443–444).

In most of the PUs with *красный* and *rot* as phraseological components, color contributes to the meaning of the whole figurative unit primarily through its denotative lexical meaning, but some sort of a semantic interpretation is also required

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

because PUs are not completely deprived of symbolic and/or cultural potential of the color red. Cultural information about bullfighting as a traditional spectacle of many countries including Spain, Portugal, and some Latin American countries is crucial for understanding the Russian PU *как бык на красную тряпку* [реагировать, действовать и т. п.] (lit. like bull on red rag [to react, to act, etc.]) and the German PU *ein rotes Tuch* für jmdn. *sein / wie ein rotes Tuch* auf jmdn. *wirken* (lit. to be a red rag to sb. / to act like a red rag on sb.') meaning 'to make sb. furious, to be certain to produce an angry or violent reaction'. A mental image behind the German PU actually lacks the figure of a bull itself, but a red rag as an integral part of bullfighting suggests that the phraseological meaning rests on the fact that a matador practices waving a red rag at a bull during a bullfight to enrage the animal.

Sometimes the color red as a component contributes to the meaning of a PU through its specific coloristic characteristics. As red is a strong color, it is used in fashion as attention-getting and in optics as a color that advances, thereby creating the impression that red objects are closer than they are (Morton, 1997, p. 24). Apart from that fact, in the German PU *roter Faden* (lit. red thread), used in the meaning of 'a common theme', there is an additional explanation of its etymology: the PU goes back to the *Ellective Affinities*, the novel by Johann Wolfgang von Goethe. Goethe's character Ottilie keeps a diary and the main idea that goes through the whole diary is compared to the red thread in the lashing of the English navy (Duden, 2002, p. 204). There is a similar Russian PU *проходить красной нитью* через что (lit. to go like red thread through sth.) used in the meaning of 'to be the most important part of sth., to be the main idea of sth.'.

The German PU *rote Zahlen schreiben* (lit. to write black numbers) is an antonym to the PU *schwarze Zahlen schreiben*. As mentioned above, the etymology of this PU rests on the bookkeeping practice where red numbers mean financial loss (Duden, 2002, p. 889). There are several more PUs with the same phraseological image: *in die roten Zahlen kommen (geraten)*, *aus den roten Zahlen <heraus>kommen*; *in den roten Zahlen sein* (lit. to get to the red numbers, to get out from the red numbers, to be in the red numbers).

Another figurative unit which is widespread in different languages is *den roten Teppich ausrollen* (lit. to roll out the red carpet), with its Russian equivalent *расстелить красную <ковровую> дорожку* (lit. to spread out red <carpet> path). It is used when we want to welcome a guest with high honors. The etymology of the PU is based on the literal meaning of red carpet, i.e., a thick red covering for a floor or other surface that is put down for important guests to walk on, but it requires an additional semantic interpretation.

There are a few gathered PUs which are culturally rather interesting. One of them is the German *roter Hering* (lit. red herring). In the book *64 Fehlschlüsse*

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

in Argumenten: Logische und rhetorische Irrwege erkennen und vermeiden (2014), the author Albert Mößmer discusses argumentation fallacies and in the chapter *Der rote Hering (von der Fährte abbringen)*, he explains that *roter Hering* is a loan translation from the English PU *red herring* denoting something that misleads or distracts from a relevant or important issue. Mößler mentions the English journalist William Cobbett (1763–1835) as being the first one to use it in this sense.

2.4 Green

2.4.1 *PUs based on denotative lexical meaning*

Only a few gathered PUs rest on the denotative lexical meaning of the adjective *green*, i.e., "of the color between blue and yellow in the spectrum; colored like grass or emeralds" (Oxford Online Dictionary). The color green appears only in one Russian PU in which color contributes to the phraseological meaning through its denotative meaning. *Зелёный как <молодая> трава* (lit. green as <young> grass) rests on a comparison between the green color of an object or unhealthy skin color and the color of grass. The color green, in its denotative meaning, also appears in one German PU: *bei Mutter Grün <schlafen>* (lit. <to sleep> at Mother's Green). The meaning of this PU is 'to sleep out in the open, to sleep in nature'. It derives from the allegorical perception of 'Mother Nature'.

2.4.2 *PUs based on symbolic and/or cultural potential*

There are a few PUs with the color component *зелёный/grün*, which contain a literal color designation but require a semantic interpretation. For instance, *зелёная улица* (lit. green street) – *grüne Welle* (lit. green wave); *открыть/открывать зелёную улицу* кому, чему, за что (к чему) (lit. to open green street for sb., for sth.) – jmdm. *grünes Licht geben* (lit. to give green light to sb.). The former PU is associated with a green wave, a coordinated mode of operation of traffic lights that allows continuous travel of vehicles moving at a certain speed along a road with a series of intersections controlled by traffic lights.[4] The latter denotes giving permission for someone to do something or for something to happen. The semantic interpretation of the PU is connected with the universal color code of colors red, amber (yellow), and green as traffic lights (signals) that alternate the right of way accorded to vehicles and pedestrians.

Green is said to be the most restful to the eye. It is the color of nature and growth, fruitfulness, renewal, freshness, tranquility, hope, health, youth and

4. Hrvatska enciklopedija (Croatian encyclopedia online) Retrieved from: http://www.enciklopedija.hr/natuknica.aspx?id=67083

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

immaturity. In contemporary culture, the color green is associated with ecology and conservation (Morton, 1997, pp. 29–30). There are PUs that can be seen as a reflection of that connection with growing, nature, and the environment, such as *einen grünen Daumen haben* (lit. to have a green thumb). If someone has a green thumb, s/*he* has the ability to grow plants well.

As mentioned above, green stands for renewal, growth, and hope as well as for immaturity (Morton, 1997, p. 29). The Russian and German phraseological subsystems capture that in the following PUs: *молодо-зелено* (lit. young-green) – *ein grüner Junge* (lit. a green boy), *grün hinter den Ohren sein* (lit. to be green behind the ears). By using these PUs, we denote a green and inexperienced person.

The color green is considered to be relaxing, soothing and peaceful (Morton, 1997, p. 63), but it is also a color of anger, jealousy, and envy. This symbolic meaning of the color green, in combination with yellow, has been captured and preserved in the German phraseological subsystem in the following PUs: *grün und gelb vor Ärger werden* (lit. to get green and yellow from anger), *sich gelb und grün ärgern* ('to be yellow and green with anger'), *grün vor Neid sein (werden)* ('to be (turn) green with envy'). This has been explained by Wanzeck (2003), who took into account the theory of the four humors of Hippocratic medicine (black bile, yellow bile, phlegm, and blood), with each humor corresponding to one of the traditional four temperaments (pp. 78–80). Thus, the yellow-green color of the yellow bile is connected with anger.

There are several culturally specific PUs with the color component *зелёный* that are unique, i.e., language-specific, and therefore seen as culturemes. The Russian PU *зелёный змий* (lit. a green dragon) which refers to alcohol in general and *напиться / напиваться (допиться / допиваться, нарезаться / нарезываться) до зелёного змия* (lit. till a green dragon (drink, get drunk)) meaning 'to get extremely drunk', are based on a mental image of a green dragon seen by somebody during alcohol-induced hallucinations. However, both of them can also be associated with a poetic syntagma *зелёное вино* (lit. green wine) used in the Russian language to refer to the green color of grapes, a fruit from which wine is made.

2.5 Blue

The Russian language has more than one color term that refers to the color blue. Two different shades of blue are denoted by *голубой* and *синий*. For Russian language speakers, they are considered to be two different basic color terms: *голубой* for light blue and *синий* for dark blue.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.5.1 *PUs based on denotative lexical meaning*

There are a few gathered PUs that rest on the denotative lexical meaning of the adjective *blue*, i.e., "of a color intermediate between green and violet, as of the sky or sea on a sunny day" (Oxford Online Dictionary). Since blue represents the sky, it has been captured in a form of phraseological units in Russian and German, which are mainly used to describe the color of eyes: *глаза* у кого *голубые как небо* (lit. sb. has eyes blue as sky) – *blau wie der Himmel* (lit. blue as the sky). There is another German lexical idiom with the color blue and it rests on a natural color of azurite, a soft, deep blue copper mineral: *azurblau* (lit. azureblue).

2.5.2 *PUs based on symbolic and/or cultural potential*

Голубой as a component of the Russian PU *голубая мечта* (lit. blue dream), denoting a passionate wish, a longing for something imaginary, contributes to its meaning through the symbolism of the color blue. Despite the fact that blue is the color of the sky, oceans, and lakes, it is not commonly found in natural objects and it can be described as the color of transparency and immateriality (Morton, 1997, p. 27). Therefore, it evokes the idea of the imaginary, introspection, eternity and calmness (Chevalier & Gheerbrant, 1996, pp. 102–103).

Although the PU *синяя птица* (lit. blue bird) used in the meaning of happiness and *гоняться за синей птицей* (lit. to chase after a blue bird) 'to look for happiness' can be associated with the name of a play *The Blue Bird* (French: *L'Oiseau bleu*) written by Maurice Maeterlinck (1908) in which a blue bird figures as a symbol of elusive happiness, in both of them the aforementioned symbolism also plays an important role.

In the gathered phraseological collection, there are several PUs that are interesting from a symbolic and cultural point of view. One of them is *голубая кровь* (lit. blue blood) – *blaues Blut <in den Adern haben>* (lit. <to have> blue blood <in the veins>). 'Blueblood' refers to a person of royal, noble, or superior birth, aristocracy. From a synchronic perspective, the color blue in this PU can be considered as used in a transferred sense, but if we look at its etymology, we find that the term probably originates from the notion of veins being more visible in people with a fair complexion. It is, in origin, a translation of the Spanish *sangre azul*, an expression that can be traced to the period after the Christian conquest of Muslim Spain when many Muslims and Jews were compelled to accept Christianity, and these converts were often viewed with suspicion and contempt by Christian authority (SWHM, 2007, p. 29).

> Since many Spanish Muslims and Jews were ultimately of North African origin, they often had quite dark skin in which the blue lines of veins were not visible. Fair skin, on the other hand, was thought to be typical of the Christian nobility. *Sangre azul* thus came to be considered to be proof of Christian ancestry – without any admixture of Muslim or Jewish elements. (SWHM, 2007, p. 29)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The color blue is, on the one hand, psychologically associated with spirituality and, on the other hand, it denotes the sky and oceans in nature (Morton, 1997, p. 27). If we would try to reduce those associations to a common denominator, it could be open space and freedom. Both of them are hidden behind two German PUs: *ins Blaue (hinein)* (lit. into the blue (inside)), meaning 'into the unknown, without aim', and *eine Fahrt ins Blaue* (lit. a ride into the blue), meaning 'a journey into the blue, into the unknown, without aim'.

The most culturally interesting Russian PU in the gathered PUs' collection is *синий чулок* (lit. blue stocking). It is usually used derogatively to refer to an educated, intellectual woman who has no feminine charm or seductiveness. It originates from the English term *bluestocking,* which is associated with the name of a society founded in the late 18th century exclusively for the purpose of discussing literary, scientific and other intellectual topics. It is believed that the name *Blue Stocking Society* was given to the group by General Boscawen in reference to Benjamin Stillingfleet, a scholar and a regular guest of the salon who used to wear blue stockings (Birih, Mokienko, & Stepanova, 1998, p. 630). Although the term *bluestocking* was originally used to describe both men and women, in time, it started to refer strictly to women dedicated to intellectual pursuits and it became even pejorative, "used within a broad reaction against women's pursuit of intellectual life and social usefulness beyond mere domesticity" (Kelly, 1999, p. xi).

2.6 Yellow

2.6.1 *PUs based on denotative lexical meaning*

Yellow is "the color between green and orange in the spectrum, a primary subtractive color complementary to blue; colored like ripe lemons or egg yolks" (Oxford Online Dictionary). In the consciousness of Russians and Germans, a lemon is a prototype for this color, which is shown in these PUs: *жёлтый как лимон* (lit. yellow as lemon); *жёлтый как лимонная корка* (lit. yellow as lemon peel) – *gelb wie eine Zitrone* (lit. yellow as a lemon). There is another German simile, less frequent in use than the one just mentioned: *gelb wie die Sonne* (lit. yellow as the Sun). It rests on the color feature of yellow as the color of sunshine.

The yellow color of an object in Russian phraseology is compared to amber, a fossilized tree resin appreciated for its color, valued as a gemstone, and used in jewelry (*жёлтый как янтарь* (lit. yellow as amber). This kind of a comparison is specific for the Russian language because this gemstone is particularly valued in Russia and most of the world's extractable amber is located in the Russian area of Kaliningrad.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.6.2 *PUs based on symbolic and/or cultural potential*

Culturally interesting are the PUs *жёлторотый птенец* (lit. yellow-beaked nestling) – *Gelbschnabel* (lit. yellow beak). The color of a nestling's beak is usually yellow and it metonymically stands for a nestling, which has been metaphorically transferred to denote a young, inexperienced person. Taking into account the symbolic meaning of green, it is possible to say *Grünschnabel* as well.

The Russian PU *жёлтый дом* (lit. yellow house), denoting a mental hospital, originated through the association with the yellow color of the facade of the well-known mental hospital Oluhovski in Saint-Petersburg (Birih, Mokienko, Stepanova, 1998, p. 215).

2.7 Grey

2.7.1 *PUs based on denotative lexical meaning*

The denotative lexical meaning of the adjective *grey* is "of a color intermediate between black and white, as of ashes or lead" (Oxford Online Dictionary). It has been reflected in only one PU in the gathered PU collection, namely *grau wie Asche* (lit. grey as ash).

2.7.2 *PUs based on symbolic and/or cultural potential*

The color gray is associated with sadness (Morton, 1997, p. 39). Thus, the German PU *alles grau in grau sehen (malen)* (lit. to see (paint) everything grey), meaning 'to depict everything in a pessimistic way', reflects that symbolism.

Culturally interesting are the PUs *серый кардинал* (lit. grey cardinal) and *graue Eminenz* (lit. grey eminence), both denoting *grey eminence*, a person who wields power and influence unofficially or behind the scenes (Collins English Dictionary). The etymology of that expression goes back to the nickname of François Leclerc du Tremblay, the Capuchin monk considered to be the right-hand of Cardinal Richelieu because of his immense, but covert influence over him. The nickname people addressed him by is directly associated with the mentioned influence as well as with the color of his religious robe (cf. Birih, Mokienko, & Stepanova, 1998, p. 523).

2.8 Pink

All Russian and German PUs with *розовый*/*rosa* as phraseological components contribute to the phraseological meaning only through the inherently positive symbolic meaning of the color pink. It is a color psychologically perceived as gentle, comforting, sweet, festive, warm, vibrant, and happy (Morton, 1997, pp. 47–49). Thus, *розовый момент* (lit. a pink moment), for example, denotes a moment of

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

excessive joy due to which someone does not see things rationally and clearly. The mental image behind this PU suggests that everything looks much better in pink and the color pink prevents us from seeing the negative side of things. A similar background mental image is found in several Russian and German PUs which refer to the impossibility of seeing the world around us clearly and objectively. The same idea shapes the Russian PU *видеть (представлять) в розовом свете* кого, что (lit. to see (to imagine) sth., sb. in pink light) and its German equivalent etw., jmdn. *in rosa <rotem> Licht sehen* (lit. to see sth., sb. in pink <red> light) meaning 'to be very positive in giving an opinion on sth., sb.', as well as the Russian PU *представить/представлять в розовом свете* кого, что and its German equivalent etw., jmdn. *in rosigem Licht darstellen (schildern)* (lit. to depict sth., sb. in pink light) meaning 'to be very positive in depicting sth., sb.'. A very similar meaning and structure can be observed in the Russian PU *смотреть сквозь розовые очки* на кого, на что (lit. to look through pink glasses at sth, sb.) meaning 'to see only good sides of sth., sb., to be very positive in giving an opinion on sth., sb.' and the German PU *durch die rosa <-rote> Brille sehen (betrachten)* (lit. to look through the pink<-red> glasses) meaning 'to be very optimistic'. They rely on the idea that wearing colored glasses affects one's view of the word in a subjective and biased way, as if the color pink (or pink-red in case of the German PU) makes the world look better than without wearing the glasses. All the PUs mentioned above, except the Russian *розовый момент*, are confirmed to be widespread in Indo-European languages in Europe. Consequently, this aspect of the color pink may be considered to be a cultural symbol of the figurative lexicon of various European languages (Piirainen, 2016, p. 279).

2.9 Purple

The color purple appears as a phraseological component only in the Russian language and only in two PUs: *все фиолетово* кому (lit. everything is purple to sb.) denoting someone who is completely nonchalant about everything; and *сугубо (глубоко) фиолетовый* кто <к кому, к чему> (lit. sb. is particularly purple towards sb./sth.) meaning 'sb. doesn't care at all about sb./sth.'. Both of them are quite new and used primarily in youth slang. The etymology is vague and one can only presume that the symbolic potential of the color purple plays a role in structuring the phraseological meaning. The Russian adjective *фиолетовый* is lexically close to the English adjective *violet*: "[v]iolet is the color of temperance, clarity of mind, deliberate action, of balance between Heaven and Earth, senses and spirit, passion and reason, love and wisdom" (Chevalier & Gheerbrant, 1996, pp. 1068–1069). Taking this into consideration, we can connect this symbolic potential with the fact that somebody who is oriented on spirituality does not care much about everyday life and things that surround him/her.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.10 Orange

According to Morton (1997), orange symbolizes warmth, energy, and excitement (p. 33). Although rich in symbolism, the color orange is extremely rare as a phraseological component. However, the above-mentioned symbolism influences the Russian PUs *оранжевое настроение* (lit. orange mood) and *быть (пребывать) в оранжевом настроении* (lit. to be in orange mood), denoting a particularly optimistic and carefree mood.

2.11 Brown

In the phraseological subsystems of Russian and German, the color brown is hardly productive. There is only one PU with the component brown: *aussehen wie Braunbier mit Spucke* (lit. to look like brown beer with spit). It is used to denote a person who looks unwell, ill. The meaning of the figurative unit rests on the perception of brown beer as something unappealing (which is surprising if we consider the status that beer has in German culture), additionally worsened with spit.

3. Lexicalization patterns of color terms based on Russian and German phraseology

In most structural types of the analyzed PUs we cannot detect potential lexicalization patterns in color naming since color words can be viewed primarily as phraseological components that contribute to the phraseological meaning through their symbolic and/or cultural potential. Nevertheless, lexicalization patterns in color naming can be observed even on the phraseological level through similes as a specific and very frequent structural type of PUs found both in Russian and German phraseology. Similes are specific not only because they represent a specific structural type of PUs, but because they rest on a denotative lexical meaning of a color described through the comparison with objects, materials, or natural phenomena. Some similes in the PU collection are considered to the basis for the lexicalization of certain color terms resulting from the usage frequency of the comparison underlying them since their phraseological status is based and confirmed due to their frequent usage in language.

Similes of this type are, in general, more frequent in the analyzed Russian phraseological subsystem than in German. Subsequently, the lexicalization of color terms based on PUs occurs more often in the Russian language. At the same time, the Russian language shows two basic lexicalization patterns. Sometimes a color term is derived from a lexeme that plays the role of *secundum comparatum* in the

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

similes (*белый как снег* lit. white as snow > *снежный* lit. snowy), and sometimes a color term is lexicalized as a compound adjective composed of both *primum comparandum* and *secundum comparatum* of similes (*белый как снег* lit. white as snow > *снежнобелый* lit. snowwhite, *белоснежный* lit. whitesnowy, *снежно-белый* lit. snowy-white). In the gathered German corpus, we detected only the color terms lexicalized in a form of compound adjectives. Black, white, and red as phraseological components are most productive when it comes to lexicalization, while pink, purple, orange, and brown are not productive, neither in Russian nor in German.

(1) Black: *чёрный как уголь* lit. black as coal > *угольный* lit. coal-like, *чёрный как смоль* lit. black as pitch > *смоляной* lit. pitch-like, *чёрный как ворон* lit. black as raven > *вороной* lit. raven-like, *чёрный как смородины* lit. black as blackcurrants > *смородинный* lit. blackcurrant-like, *чёрный как угли* lit. black as coal > *угольно-чёрный* lit. coal-black, *чёрный как агат* lit. black as agate > *агатовый* lit. agate-like, *чёрный как антрацит* lit. black as anthracite > *антрацитный* lit. anthracite-like, *чёрный как чернила* lit. black as ink > *чернильно-чёрный* lit. ink-black; *schwarz wie die Nacht* lit. black as the night > *nachtschwarz* lit. nightblack, *rabenschwarz* lit. ravenblack / *kohlrabenschwarz* lit. coalravenblack.

(2) White: *белый как снег* lit. white as snow > *снежнобелый* lit. snowwhite, *белоснежный* lit. whitesnowy, *снежно-белый* lit. snowy-white, *снежный* lit. snowy – *белый как известь* lit. white as lime > *известково-белый* lit. lime-white, *известковый* lit. lime-like; *белый как алебастр* lit. white as alabaster > *алебастровый* lit. alabaster-like, *белый как кипень* lit. white as foam> *кипенный* lit. foam-like, *белокипенный* lit. whitefoam-like, *белый как молоко* lit. white as milk > *молочный* lit. milk-like, *молочно-белый* lit. milk-white, *молочнобелый* lit. milkwhite, *белый как мука* lit. white as flour > *мучнисто-белый* lit. flour-white, *белый как мрамор* lit. white as marble > *мраморно-белый* lit. marble-white; *weiß wie Schnee* lit. white as snow > *schneeweiß* lit. snowwhite.

(3) Red: *красный как кровь* lit. red as blood > *кровавый* lit. blood-like, *кровяной* lit. blood-like, *кроваво-красный* lit. blood-red, *красный как рубин* lit. red as ruby > *рубиновый* lit. ruby-like, *рубинный* lit. ruby-like, *красный как огонь* lit. red as fire > *огневой* lit. fire-like, *огненный* lit. fire-like, *огненно-красный* lit. fire-red, *красный как кирпич* lit. red as brick > *кирпичный* lit. brick-like, *кирпично-красный* lit. brick-red, *красный как клюква* lit. red as cranberry > *клюквенно-красный* lit. cranberry-red, *красный как кумач* lit. red as red calico > *кумачовый* lit. red calico-like, *кумачный* lit. red calico-like, *кумачевый* lit. red calico-like; *rot wie Blut* lit. red as blood > *blutrot* lit. bloodred.

(4) Green: *зеленый как <молодая> трава* lit. green as <young> grass > *травяной* lit. grassy, grass-like.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(5) Blue: *голубой как небо* lit. blue as sky > *небесный* lit. sky-like, *небесноголубой* lit. skyblue, *небесно-голубой* lit. sky-blue, *голубой как бирюза* lit. blue as turquoise > *бирюзовый* lit. turquoise; *blau wie der Himmel* lit. blue as the sky > *himmelblau* lit. skyblue, *azurblau* lit. azureblue.

(6) Yellow: *жёлтый как лимон* lit. yellow as lemon > *лимонный* lit. lemon-like, *жёлто-лимонный* lit. yellow-lemon-like, *лимонно-жёлтый* lit. lemon-yellow, *лимонножёлтый* lit. lemonyellow, *жёлтый как шафран* lit. yellow as saffron > *шафрановый* lit. saffron-like, *шафранный* lit. saffron-like, *шафранножёлтый* lit. saffronyellow, *жёлтый как янтарь* lit. yellow as amber > *янтарный* lit. amber-like, *жёлтый как мёд* lit. yellow as honey > *медовый* lit. honey-like, *медно-жёлтый* lit. honey-yellow; *gelb wie eine Zitrone* lit. yellow as a lemon > *zitronengelb* lit. lemonyellow.

(7) Gray: *grau wie Asche* lit. grey as ash > *aschgrau* lit. ashgrey.

4. Conclusion

One of the main aims of this chapter was to systematically depict the main features of Russian and German color phraseology. Thanks to the insight into the phraseological image of the world of Russians and Germans, we can conclude that the most productive color in both languages, from the phraseological point of view, is the color black (see Figure 1). This correlates with the general tendency of phraseology to show primarily negative concepts. As Figure 1 shows, the most productive colors in phraseology of the two analyzed languages are: *black*, *white*, *red*, *green*, and *blue*. It has to be pointed out that the reason why Russian PUs stand out regarding the total number is the maximalist approach that Russian lexicographers have when compiling phraseological dictionaries.

As it is common in the phraseological practice, background images, motivational elements, etymology, cultural, and symbolic features were taken into consideration (cf. Barčot, 2014; Hrnjak, 2005, 2015) and the semantic analysis of the collected data was performed.

Finally, it can be stated that the analysis of the collected phraseological material with color components in Russian and German confirmed the premise about similarities between one part of the phraseological linguistic worldview of Russians and Germans. On the other hand, there is a great number of nation-specific PUs, the so-called culturemes that require additional explanation.

There is a larger number of German PUs with symbolic and/or cultural meaning, whereas there are more Russian similes with color components, i.e., more PUs with color components that rest on natural color features (see Figure 2).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

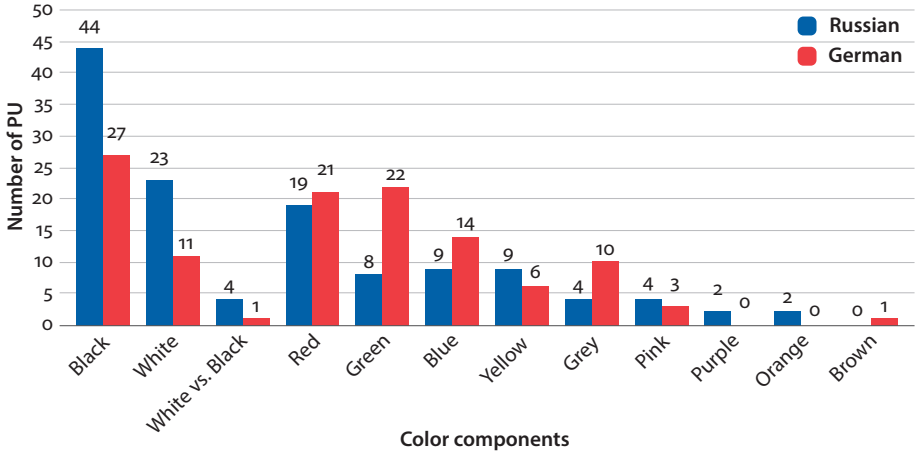


Figure 1. Distribution of PUs regarding the color component in Russian and German

Similes are the only structural type of PUs in the analyzed Russian and German subsystems that can be considered as the basis for the lexicalization of certain color terms resulting from the usage frequency of the comparison underlying similes. This type of lexicalization is most frequent in Russian similes and mainly with the color components *black*, *white*, and *red*.

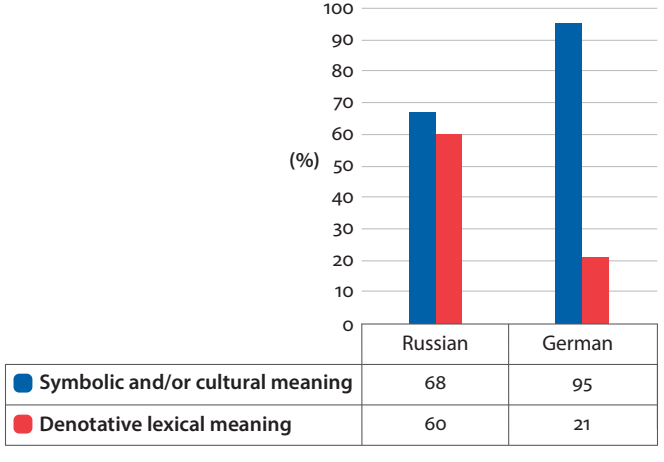


Figure 2. Distribution of PUs regarding symbolic and/or cultural meaning and denotative lexical meaning

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

References

Afanas'ev, A. N. (2008). *Slavjanskaja mifologija*. Moskva: Eksmo, Sankt-Peterburg: Midgard.

Afanas'ev, A. N. (2016). *Mify drevnih slavjan*. Moskva/Berlin: Direct Media.

Barčot, B. (2014). *Divlja životinja kao sastavnica u hrvatskoj, ruskoj i njemačkoj frazeologiji*. Zagreb: Filozofski fakultet. (doctoral dissertation)

Birih, A. K., Mokienko, V. M., & Stepanova, L. I. (1998). *Slovar' russkoj frazeologii. Istoriko-etimologičeskij spravočnik*. Sankt-Peterburg: Folio-Press.

Chevalier, J., & Gheerbrant, A. (1996). *Dictionary of Symbols*. London: Penguin Books.

Čermák, F. (2007). Idioms and morphology. In H. Burger, D. Dobrovol'skij, P. Kühn, & N. R. Norrick (Eds.), *Phraseologie: ein internationales Handbuch der zeitgenössischen Forschung. 1. Halbband* (pp. 20–27). Berlin/New York: Walter de Gruyter. https://doi.org/10.1515/9783110171013.20

Duden (2002). *Redewendungen. Wörterbuch der deutschen Idiomatik*. Band 11, hrsg. von der Dudenredaktion. Mannheim/Leipzig/Wien/Zürich: Dudenverlag.

Duden (2012). Redensarten: Woher sie kommen, was sie bedeuten. Mannheim–Zürich: Dudenverlag.

Hrnjak, A. (2005). Frazemi s bojom kao komponentom (na primjeru hrvatskih i ruskih frazema). Zagreb: Filozofski fakultet. (master thesis)

Hrnjak, A. (2015). *Rodni elementi u hrvatskoj i ruskoj frazeologiji*. Zagreb: Filozofski fakultet. (doctoral dissertation)

Kelly, G. (1999). General Introduction. In E. Eger (Ed.), *Elizabeth Montagu*, Vol. 1 of G. Kelly (Gen. Ed.), Bluestocking Feminism: Writings of the Bluestocking Circle, 1738–1785, 6 vols. (pp. ix–liv). London: Pickering & Chatto. Retrieved from: https://sites.google.com/a/georgiasouthern.edu/nonconformist-women-writers-1650-1850/kelly-introduction-bluestockings-1999.

Mokienko, V. M., & Nikitina, T. G. (2007). *Bol'šoj slovar' russkih pogovorok*. Moskva: ZAO "OLMA Media Grupp".

Mokienko, V. M., & Nikitina, T. G. (2008). *Bol'šoj slovar' russkih narodnyh sravnenij*. Moskva: ZAO "OLMA Media Grupp".

Morton, J. (1997). *A Guide to Color Symbolism*. New York: Colorcom. Retrieved from: https://www.scribd.com/document/271564820/A-Guide-to-Color-Symbolism

Mößmer, A. (2014). *64 Fehlschlüsse in Argumenten: Logische und rhetorische Irrwege erkennen und vermeiden* [Kindle version]. Retrieved from: www.amazon.com

Piirainen, E. (2012). *Widespread Idioms in Europe and Beyond. Toward a Lexicon of Common Figurative Units*. New York: Peter Lang Publishing Inc. https://doi.org/10.3726/978-1-4539-0845-7

Piirainen, E. (2016). Lexicon of Common Figurative Units. Widespread Idioms in Europe and Beyond. New York: Peter Lang Publishing Inc. https://doi.org/10.3726/978-1-4539-0845-7

SWHM = Spanish word histories and mysteries: English words that come from Spanish. (2007). From the editors of the *American Heritage Dictionaries*. Boston: Houghton Mifflin Co.

Walter, H. (2008). *Wörterbuch deutscher sprichwörtlicher und phraseologischer Vergleiche, Teil 1*. Hamburg: Verlag Dr. Kovač.

Wanzeck, Ch. (2003). *Zur Etymologie lexikalisierter Farbwortverbindungen. Untersuchungen anhand der Farben Rot, Gelb, Grün und Blau*. Amsterdam/New York: Rodopi.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wierzbicka, A. (2005). There Are No "Color Universals" But There Are Universals of Visual Semantics. *Anthropological Linguistics*, 47(2), 217–244.
Collins English Dictionary: https://www.collinsdictionary.com/dictionary/english
Merriam-Webster Dictionary: https://www.merriam-webster.com/
Oxford Online Dictionary: https://en.oxforddictionaries.com/

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

From object to color and back

Seeing the world in color in Croatian, Turkish, and Arabic

Daniela Katunar, Barbara Kerovec and Nawar Ghanim Murad
University of Zagreb

This chapter deals with lexicalization patterns of color terms and lexical units derived from color terms in Croatian, Turkish, and Arabic, three typologically and socioculturally different languages. The focus of the analysis is the relationship between object and color in: (a) the way the languages exploit the OBJECT FOR COLOR conceptualization strategy in order to name colors and (b) the way languages use the COLOR FOR OBJECT conceptualization strategy in order to lexicalize different phenomena of experience. In addition, one of the main aims of the analysis is to see which lexicalization patterns in the three languages are used productively in both types of conceptualization strategies.

Keywords: color terms, lexicalization patterns, conceptualization strategies, Croatian, Turkish, Arabic

1. Introduction

Literature on color terms has been, for the most part, dealing with the way color terms are used to partition the color spectrum in languages of the world. Such research began with the pioneering work of Berlin and Kay (1969) and has been extended to psycholinguistic research by E. Rosch (1973). Such research shows that color is one of the basic semantic domains which demonstrates the intricate relation between perception and conception, as well as the principles of prototype categorization. It also demonstrates some universal tendencies in color naming (Kay & Regier, 2003; Raffaelli, 2017). On the other hand, some studies have focused on what languages do with their existing color terms, especially in terms of their semantic extensions into other domains of experience – these often deal with phraseological units (e.g., Hrnjak, 2009). Ronga and Bazzanella (2015) thus point out that the relation of universalist and relativist tendencies in color naming is

https://doi.org/10.1075/sfsl.78.16kat
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

highly complex, especially because of (in our own words) a 'feedback loop' between sociocultural and perceptual factors that influence the way we conceptualize particular colors in a given sociocultural context. Another 'feedback loop' is discussed by Wierzbicka (1990, 2005) and it focuses on the relationship between colors and visual prototypes that instantiate them in the world around us. One factor these authors emphasize is the cultural specification of particular colors via idioms, as well as historical factors which can give a broader insight into cross-linguistic variations and similarities in naming and saliency of certain colors in the lexicon of a language. Sociocultural context is not only an extremely relevant factor in our closer understanding of the universal and (linguistically) relative properties of colors, but also in our understanding of the limits to which sociocultural context can influence the universal tendencies in color perception and categorization.

This study compares lexicalization patterns of color terms in Croatian, Turkish, and Arabic, three languages which are typologically, genetically, and socioculturally quite different. We focus particularly on the OBJECT FOR COLOR relationship as it is expressed structurally in these three languages. We also focus on the 'feedback loop' between objects and colors by examining the COLOR FOR OBJECT[1] conceptualization strategy, i.e., by examining word-formation patterns (derivation, compounding, and multi-word expressions) in which color terms are used to name phenomena from domains of experience other than color itself. As to our knowledge, such an investigation has not been undertaken in the context of examining lexicalization patterns related to color terms (as well as their cultural saliency) in these three languages contrastively. By comparing three typologically and culturally different languages we aim to take into account the structural and sociocultural variables that allow for an account of the diversity of language expressions formed by different color terms and thus examine the question of their saliency in a novel way. For instance, in Croatian, the lexemes *zelen**ilo*** 'greenery.N' and *zelen**java*** 'green legumes.N', derived by suffixation (*zelen* 'green.Adj' + *ilo/* + *java*), have specialized to mean green vegetation and green legumes, respectively. This means that both terms have developed new and specific meanings and denote particular entities related to (but not themselves part of) the color domain. Similarly, the Arabic root system motivates the formation of morphosemantically related lexemes using systematic

1. We must note that in this chapter we will not focus on defining the details of cognitive mechanisms of meaning extension which play a role in the OBJECT FOR COLOR and COLOR FOR OBJECT relationship, most notably, metonymy and metaphor. For this reason, we name both relations as conceptualization strategies, which we believe is a cover term for both metonymy and metaphor, especially because from a morphological and syntagmatic point of view one stem can instantiate both metonymic and metaphoric processes. For more on the theoretical aspects of metonymy and metaphor in color term use and formation see, e.g., Sanford (2014).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

patterns; this is also true for words derived from basic color terms, e.g., أخضَر / *akhḍar*/ 'green' > خُضار */khuḍār/* 'vegetables', خُضَرِي */khuḍariyy/* '(green)grocer', خُضرَة or اِخضِرار */khuḍra/* or */ikhḍirār/* 'greenness', يَخضور */yakhḍūr/* 'chlorophyll', مُخضَر */mukhḍar/* 'greenish; virescent'. In Turkish, on the other hand, a lexeme which refers primarily to color can, without any additional affixes, refer to an entity which is not a part of the color domain, e.g., *siyah* 'black' > 'garments worn during mourning'; 'bold letter'. A similar situation can be observed when the strategy OBJECT FOR COLOR is used – the noun *gök* denoting 'sky' can, with no additional affixes, denotes blue color as well.

This chapter thus has several goals. Firstly, the goal will be to investigate the lexicalization patterns of color terms in Croatian, Turkish, and Arabic. Such a comparison of these three languages, to our knowledge, has not been presented in extant literature. It thus serves to establish the general structural (lexicalization) processes at work in these languages in the color domain. We will especially focus on the way the three languages exploit the OBJECT FOR COLOR strategy. Secondly, the aim is to sketch out the main tendencies of morphosemantic extensions for color terms in the three languages (based on the COLOR FOR OBJECT strategy), especially with regards to (1) saliency of particular colors and their importance in the formation of the lexicon, (2) (ir)regularities in naming strategies, and (3) similarities and differences in morphological devices (derivation, compounding, multi-word formation) these three languages use to name 'objects' with colors terms.

2. Lexicalization patterns of color term formation in Croatian, Turkish, and Arabic

Several mutually unrelated studies have investigated inventories for basic and non-basic color terms in Croatian, Turkish, and Arabic.[2] Based on the results reported by these studies and corpus and lexicographic data investigated for the purposes of this chapter, we provide a list of the most common lexicalization patterns in these three languages. Firstly, it must be noted that there are comparable lexicalization patterns in Croatian, Turkish, and Arabic, especially in terms of the devices used to lexicalize morphologically complex color terms (derivation and compounding). There are also some major differences that stem from the structural and typological types these languages belong to, for example, word order in compounds and multi-word units, the use of infixes (for Arabic) or the lack

2. For Croatian this was most notably investigated in the project *Evolution of Semantic Systems* (EoSS), Max Planck Institute of Psycholinguistics in Nijmegen (see Raffaelli, 2017). For Turkish see Özgen and Davies (1998), and for Arabic see Hasan, Al-Sammerai, and Kadir (2011).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

of prefixation (for Turkish). All three languages have a similar inventory of basic color terms (11 for Turkish and Arabic, and 13 for Croatian[3]), but they differ in the number of simplex vs. complex basic color terms and the productivity of particular lexicalization patterns used to form secondary color terms. As for lexicalization patterns, we determined three productive morphological strategies in these languages, which we can divide into three categories: suffixation, genitive construction and compounding. The conceptualization strategy OBJECT FOR COLOR is used predominantly in two of the most productive lexicalization patterns for all three languages, as we will show below. However, there are differences in the tendencies of choosing socioculturally salient objects to form color terms, as we will also discuss in the following sections.

2.1 OBJECT FOR COLOR lexicalization patterns

2.1.1 *Basic color terms in Croatian, Turkish, and Arabic – simplex forms and suffixation*

2.1.1.1 *Croatian*

Based on the data collected via the 84 Munsell color chips in the project *Evolution of Semantic Systems* (EoSS),[4] the following 13 basic color terms were established for Croatian: *bijela* 'white', *crna* 'black', *crvena* 'red', *plava* 'blue', *modra* 'dark blue', *zelena* 'green', *žuta* 'yellow', *narančasta* 'orange', *ljubičasta* 'purple', *ružičasta* 'pink', *smeđa* 'brown', *siva* 'grey' and *tirkizna* 'turquoise'.[5] Out of these 13 color terms, 9 are morphologically simplex and 4 are complex and derived via the pattern [OBJECT + suffix *-ast(a) / -an(a)*] (e.g., *ljubica* 'violet (flower)' + *-asta* > *ljubičasta* 'purple', *tirkiz* 'turquoise (mineral)' + *-na* > *tirkizna*).

3. As reported by the EoSS research.

4. Max Planck Institute of Psycholinguistics in Nijmegen. For more on the project, see Raffaelli (2017) and references therein.

5. For Croatian we list color terms with the inflection ending *-a* which is the singular feminine nominative ending for adjectives, e.g. *crn-a* 'black', *ljubičast-a* 'violet'. The term for color, *boja*, is also feminine. These are the forms elicited by the EoSS research. Indefinite adjectives have the zero morph for masculine gender and are the default lexical entry forms in dictionaries, e.g. *crnø* 'black', *ljubičastø* 'violet', etc. Indefinite adjective forms are also isomorphic to color stems used to form derivatives and compounds, as can be seen in the sections below.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.1.1.2 *Arabic*

Arabic recognizes 11 main colors as basic color terms based on the criteria by Berlin and Key (1969). These 11 basic color terms can be divided into two categories according to their lexical organization: primary (6 colors) and secondary (5 colors) (Hasan et al., 2011). The first group can be easily recognized due to its four-letter lexemes which are considered to be an essential feature of this group. Additionally, color terms in this group start with the letter الف */alif/* Eng. A, where the short vowel of فتحة */fatḥa/*[6] /a/ settles on the third consonant, e.g., أحمر */aḥmar/* 'red', أزرق */azraq/* 'blue', أسود */aswad/* 'black'.[7] The remaining 5 colors belong to the secondary group, e.g., بني */bunniyy/* 'brown', بنفسجي */banafsajiyy/* 'purple', وردي */wardiyy/* 'pink', برتقالي */burtuqāliyy/* 'orange', and رصاصي */raṣāṣiyy/* 'grey'. They are morphologically complex and do not correspond to the first primary group pattern. The pattern of the secondary group utilizes the OBJECT FOR COLOR strategy and it is one of three fundamental patterns for forming non-basic color terms in Arabic, e.g., بنفسج */banafsaj/* 'violets' + ـي '*-iyy*' > بنفسجي */banafsajiyy/* 'purple', برتقال */burtuqāl/* 'oranges' + ـي '*-iyy*' > برتقالي */burtuqāliyy/* 'orange-color', نيلة */nīla/* 'indigo' + ـي '*-iyy*' > نيلي */nīliyy/* 'dark blue'. The stem here stands for the object, while the insertion of the suffix ي */iyy/* leads to the creation of the color term. Several domains can provide objects (i.e., stems) for naming colors: flowers, plants, minerals, animals, and others, e.g., عناب */'innāb/* 'jujube' + ـي '*-iyy*' > عنابي */'innābiyy/* 'russet, burgundy', وردة */warda/* 'rose' + ي '*iyy*' > وردي */wardy/* 'pink', حشيش */ḥashīsh/* 'grass' + ـي '*-iyy*' > حشيشي */ḥashīshiyy/* 'sap-green, reseda-colored', رصاص */raṣāṣ/* 'lead' + ـي '*-iyy*' > رصاصي */raṣāṣiyy/* 'grey', فيل */fīl/* 'elephant' + ي '*iyy*' > رمادي */ramād/* 'ash(es)' + ي '*iyy*' > فيلي /fīliyy/ 'elephant-like color', */ramādiyy/* 'dark-grey', lit. 'ash-colored'. The grammatical process of the insertion of the suffix ي */iyy/* is known as '*Nisba* النسبة'[8] 'relative adjectives', lit. 'kinship, relation'.[9] This is the most productive pattern in Arabic for creating novel color terms.

2.1.1.3 *Turkish*

Turkish has 11 basic color terms corresponding to Berlin and Kay's 11 color universals: *yeşil* 'green', *sarı* 'yellow', *siyah* 'black', *kırmızı* 'red', *mavi* 'blue', *beyaz* 'white',

---

6. See Ryding (2005, pp. 25–34).

7. We can add to this group أسمر */asmar/* 'brown-tan' and أشقر */ashqar/* 'blond' that fit this lexical criterion, and are used to describe humans (skin and/or hair). أسمر */asmar/*, however, can be used also with some other nous like sugar, bread, etc.

8. See Ryding (2005, pp. 261–275).

9. Using the very same pattern we can form, for instance, the noun عراقي /'irāqiyy/ 'Iraqi' out of the lexeme العراق /al-'irāq/ 'Iraq'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

mor 'purple', *kahverengi* 'brown', *pembe* 'pink', *turuncu* 'orange', and *gri* 'grey' (Özgen & Davies, 1998). It must be noted that besides *siyah, beyaz, kırmızı, mavi*, and *gri* Turkish has additional terms for 'black', 'white', 'red', 'blue', and 'grey', respectively *kara*, *ak, kızıl/al, gök,* and *boz*. These are all originally Turkish words while the terms recognized as basic by Özgen and Davis (1998) are of Arabic, Persian or French origin.[10] Out of these eleven basic color terms, nine are morphologically simplex and two, *turuncu* 'orange' and *kahverengi* 'brown', are complex and derived via two different OBJECT FOR COLOR patterns – a) genitive/possessive construction [OBJECT-color$_{POSS.3SG}$], e.g., *kahve* 'coffee' + *renk* 'color' + -*i*$_{POSS.3SG}$ and (b) [OBJECT -*i/u*], e.g., *turunç* 'bitter orange' + -*u* > *turuncu* 'orange (color)'. The suffix in (b) ([N–*ī*]) is borrowed from Arabic ([stem + ـي */iyy/*]).[11] This correspondence is due to long historical and civilizational contacts between Turkish and Arabic and points to the importance of such factors in color term formation and the potential for borrowing lexicalization patterns.

Thus, in all three languages the basic color term inventory can be divided into two groups – morphologically simplex color terms and morphologically complex color terms which are formed via suffixation (Croatian [stem + *ast(a)*], Turkish [N–*ī*], and Arabic [stem + ـي */iyy/*]) and the OBJECT FOR COLOR strategy. Turkish, in addition, uses the genitive/possessive construction to lexicalize one of the basic color terms – the term for 'brown' – *kahverengi.*

Furthermore, these suffixation patterns are not reserved for basic colors only, but are used for lexicalizing various hyponyms of basic color terms as well, and to various degrees of productivity. In Croatian, the [stem-*ast(a)*] which realizes the *N-like* meaning is usually formed from stems referring to flowers of fruits. Apart from basic color terms *ljubičasta* 'purple, violet', *ružičasta* 'pink' ('rose-like'), and *narančasta* 'orange' ('orange-like'), this pattern is used for two other color terms in the EoSS data as well, *maslinasta* 'olive green', and *breskvasta* 'peach-like orange' (Raffaelli, 2017). Only a few more terms found in dictionaries can be added to these – *kestenjasta* 'chestnut brown', *pepeljasta* 'ash-like grey', and *riđasta* 'brown-red'. Raffaelli (2017) points out that this pattern is not productive across the color spectrum as one would expect due to its presence in forming (secondary)

10. Bayraktar (2004, 2005, 2006) has done extensive work comparing the meanings and the use of borrowed terms with the meanings and the use of originally Turkish terms. The results of her analysis show that borrowed terms carry rather neutral meanings related to a color itself, while originally Turkish terms can, besides colors, relate to abstract concepts as well and are much more productive in the formation of compounds, multi-word expressions (collocations) and proverbs.

11. The Arabic suffix -*ī* is called *nisbet-ī* in Turkish. The form -*u* of this suffix in the noun *turuncu* is due to the assimilation according to the principles of vowel harmony.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

basic color terms (*ljubičasta* 'violet' and *narančasta* 'orange'). Furthermore, this goes against expectations as the suffix *-ast* is used quite productively to form other *N-like* meanings not related to the color domain (e.g., *budal-ast* 'foolish', *krem-ast* 'creamy', *sedl-ast* 'saddle-like (shape)', *četveroznamenk-ast* '4-digit number'). The Croatian suffix *-na* used to form colors such as *tirkizna* 'turquoise', *zlatna* 'golden', *srebrna* 'silver', and *grimizna* 'crimson' is also more productive in other domains.[12] Among basic color terms in Turkish, only *turuncu* 'orange' is derived by the pattern [N–*ī*], but this pattern can be observed in naming some non-basic color terms as well, such as *altunî* 'golden', *atesî* 'flame-like', *erguvanî* 'redbud-like', *limonî* 'lemon-like', *mercanî* 'coral-like', *samanî* 'hay-like', *sincabî* 'squirrel-like', *şarabî* 'wine-like', *zeytunî* 'olive-like', *kuzgunî* 'raven-like', etc.[13] As the examples show, this pattern in Turkish is somewhat more productive than the Croatian pattern with the suffixes *-ast* and *-an* but it is not, as opposed to Arabic, the most productive Turkish pattern for color term formation.

2.1.2 *The genitive/possessive construction in Croatian, Turkish, and Arabic*

Another comparable lexicalization pattern is the genitive/possessive construction. In Croatian it takes the form of [N + Ngen > boja 'color' + Object], e.g., *boja vina* 'color of wine', *boja cigle* 'color of brick', *boja višnje* 'color of cherry'.[14] This pattern is considered to be borderline conventionalized by Raffaelli (2017) based on its low frequency in the EoSS data. However, the pattern is also attested by some relatively frequent uses in the Croatian Web Corpus (hrWaC), e.g., *boja lavande* 'lavender', *boja šampanjca* 'champaigne color', etc., and this leads Raffaelli (2017) to conclude that speakers have a knowledge of this pattern and use it creatively to make novel color terms. Because the pattern can be considered more schematized than other patterns, it can give speakers an opportunity for *ad hoc* formation of particular color terms depending on the context.[15]

---

12. These are denominal adjective forms derived by the suffix *-an* and appear with the feminine singular ending *-na*, e.g. *zlato* 'gold.N' > *zlat-an* 'golden.Adj' > *zlatna* 'golden color'. This suffix is also productively used for other adjectives, e.g. *pamet-an* 'smart', *mek-an* 'soft', etc.

13. For more examples see Küçük (2010). This is different from Croatian, as in Croatian some of the corresponding examples of object based color terms are formed via the genitive construction, i.e., a multi-word expression such as *boja vina* 'color of wine', *boja limuna* 'lemon color' (see 2.2.2).

14. It is worth noting that the noun *boja* 'color' is in fact a Turkish loanword itself, but Turkish uses the noun *renk* (a Persian loanword) in a similar type of construction, as we will show in the next section.

15. Some expressions found in hrWaC, for instance, are related to specialized contexts and objects, e.g., *oči boje badema* 'eyes the color of almonds'. It is interesting that, in this example, the expression *bademaste oči* 'almond eyes' would refer to the almond shape of the eyes, and not color.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

A similar pattern [N + Ngen] can be found in Arabic as well. The first noun is the lexeme لون */lawn/* 'color', and the second one is a noun in its genitive case. This pattern produces phrases like لون السماء */lawnu as-samā'-i/* 'light blue or sky-like color', lit. 'color of the sky', لون الصحراء */lawnu aṣ-ṣaḥrā'-i/* 'desert-like color', lit. 'color of the desert', لون التفاح */lawnu at-tuffāḥ-i/* 'apple-green-like color', lit. 'color of the apple'. This pattern is interchangeable with the more productive [*stem* + ي */iyy/*], with the same meanings, e.g., صحراء */ṣaḥrā'/* 'desert', تفاح */tuffāḥ/* 'apples' or سماء */samā'/* 'sky' would be سماوي */samāwiyy/*[16] 'light blue', صحراوي */ṣaḥrāwiyy/* 'desert-like color', تفاحي */tuffāḥiyy/* 'apple-like color'. To some degree, this is possible in Croatian as well, most notably in the EoSS example *boja breskve* 'color of peach' and *breskvasta* 'peach-like'. A few other examples can be found in hrWaC, e.g., *boja zlata* 'color of gold' vs. *zlatna* 'golden color'. However, the genitive construction in Croatian is not, in most cases, interchangeable with the suffix *-ast* / *-na.*

Unlike Croatian and Arabic, in Turkish the genitive/possessive construction is ubiquitous in forming color terms, has the opposite order of constituents (OBJECT first) and comes in more than one variety. There is one basic color term formed by the genitive/possessive construction – *kahverengi* 'brown'. It follows the basic pattern of the genitive/possessive construction and consists of two components: the first component is the term for 'coffee' (*kahve*), and the second component the term for 'color' (*renk*) (lit. 'coffee-color'). The two components form a so-called genitive construction,[17] a sort of Turkish N-N construction (or noun compound) which consists minimally of two juxtaposed nouns, the first of which has no suffixes while the second is marked with the 3rd person possessive suffix *-(s)I.* In Turkish, this [$N\text{-}N_{POSS.3SG}$] construction is generally used in denoting category membership (e.g., *el çantası* 'handbag', *diş hekimi* 'dentist, dental surgeon', *alan korkusu* 'agoraphobia') and it is used as the most productive pattern in non-basic color term formation. Many Turkish non-basic color terms are formed via this pattern but they differ in whether their second component is (1) the term meaning 'color', like in *kahverengi*, lit. 'coffee color', (2) a term for a specific color, like in *gök mavisi* 'sky-blue', or (3) some other noun, like in *kavuniçi* (lit. 'melon inside') 'pale orange'. The most frequently used are the subtypes (1) and (2), while the third subtype is less frequent.

---

16. With words that end with الف /alif/ or همزة-الف /alif-hamza/, النسبة /an-Nisba/ requires additional واو /wāw/ letter (sometimes نون /nūn/ letter) to the suffix ي /iyy/. Thus, we have صحراوي */ṣaḥrāwiyy/* and not صحرائي /ṣaḥrā'iyy/.

17. For a detailed explanation of the genitive construction see, e.g., Čaušević (1996, pp. 67–74, 111–113), and Göksel and Kerslake (2005, pp. 95–99, 161–164). As can be seen from the examples, the final *k* of the noun *renk* is replaced by his voiced pair *g* when followed by a suffix which begins with a vowel.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.1.3 *Differing sociocultural tendencies in the* OBJECT FOR COLOR *lexicalization patterns*

As described in previous sections, color terms in Croatian, Arabic, and Turkish can be either morphologically simplex or, more frequently, morphologically complex and formed by various morphosemantic processes (especially when it comes to non-basic color terms). Some of these processes demonstrate typological specificities and similarities of the three languages under investigation (e.g., types of affixation and word order). Lexicalization patterns, viewed from this perspective, have to do with the way languages use available means to demarcate and name a particular part of the color spectrum. One strong tendency in all three languages is to use the OBJECT FOR COLOR strategy. All three languages have in common the use of plants (particularly flowers), fruits, vegetables, and different parts of the environment (sky, sea, vegetation). However, some sociocultural differences can be observed, as well as varying degrees of productivity of particular constructions among the three languages.

A clear example of this is the [N + N_{GEN}] / [N-$N_{POSS.3SG}$] construction, which is a borderline lexicalization pattern in Croatian and Arabic, but fully productive in Turkish. The productivity of this pattern in Turkish allows for constructions which include neither the term meaning 'color', nor a term for a specific color, as shown by the third subtype of the pattern N-$N_{POSS.3SG}$ (*kavuniçi* 'pale orange', lit. 'melon inside'.). This pattern is not used in Croatian and Arabic. It can be observed that the pattern [N + N_{GEN}] / [N-$N_{POSS.3SG}$] depends in some degree on specific sociocultural and environmental facts which are used in color conceptualization. For example, in Turkish, the terms for sociocultural artefacts can appear as a component of this pattern (*fes rengi* 'crimson', lit. 'fez color', *bayrak kırmızısı* 'flag red', *türbe yeşili* 'turbe green'), which was not found in Croatian and Arabic. Additionally, the terms related to animals appear more frequently as the component of the pattern [N + N_{GEN}] / [N-$N_{POSS.3SG}$] in Turkish than in Croatian and Arabic. Many such examples can be found particularly in the third subtype of the pattern N-$N_{POSS.3SG}$ where the name of an animal appears as the first component[18] while the second component is in a WHOLE-PART relation to the first, e.g., *yavruağzı* (lit. 'baby mouth') 'light pink', *güvercinboynu* (lit. 'pigeon neck') 'green-blue-pink gradiation', *horoz ibiği* (lit. 'rooster crest) 'amaranth pink', *kazayağı* (lit. 'goose foot') 'light orange', *kumru göğsü* (lit. 'dove chest') 'dark grey', *ördekbaşı* (lit. 'duck head') 'dark blue green', *ördekgagası* (lit. 'duck beak') 'light orange', *turnagözü* (lit. 'crane eye') 'bright yellow', *deve tüyü* (lit. 'camel hair') 'buff', *tavşan kanı* (lit. 'rabbit blood')

18. Mostly birds.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

'bright red'.[19] In Croatian, there are only a couple of examples of where the object is an animal. In such examples, the animal specifies a particular hue, but not in the pattern [N + N_{GEN}] / [N-$N_{POSS.3SG}$], but rather in the compound pattern [Adj+Adj] (e.g., *mišje siva* 'mouse grey', *golublje plava/siva* 'pigeon blue/grey').[20] Such different naming strategies may have to do with the tendency of some languages for greater levels of genericity versus specificity when it comes to particular lexical domains (e.g., Ullmann, 1966). Such differences also support the relativist account of color naming, not only in word-formation processes present in three typologically different languages but also in conceptualization strategies that are prevalent when it comes to the entire color spectrum. This relativist account falls in line with the point made by Wierzbicka (1990, 2005) that concepts, and not direct sensory data, are what governs color term formation in languages. Concepts are both biological and cultural in nature and thus serve to communicate about our shared environment.

2.2 Compounding and hue overlap

Croatian, Turkish, and Arabic use a similar pattern to denote transitory colors, i.e., colors which overlap in hue. In Croatian, this is the pattern [Adj-[o]-Adj] which is a common pattern for compounds in Croatian with the use of the interfix morpheme *-o-*, e.g., *zeleno-plava* 'green-blue', *sivo-zelena* 'grey-green', etc. This pattern can be used in color naming to denote either a transitory color property such as *zeleno-plava* 'green-blue' or a subtype of color, either with the *N-like* suffix *-ast* e.g., *maslinasto-zelena* 'olive-like green' or the approximative/diminutive suffix *-kast*, e.g., *zelenkastoplava* 'greenish blue'.[21] There are some borderline cases of lexicalized

19. In Turkish, there is one more N-N pattern which differs from N-$N_{POSS.3SG}$ in the fact that the second component is always a term for a specific color, but it is not marked with the 3rd person possessive suffix *-(s)I.* (e.g., *kan kırmızı* (lit. 'blood red') 'crimson', *süt beyaz* 'milk white', *kar beyaz* 'snow white', *tozpembe* (lit. 'powder pink') 'light pink'). This pattern can be regarded as a borderline case of lexicalization since there is a very small number of terms formed via this pattern.

20. These examples have not been attested in the EoSS data, but they were found in the Croatian Web Corpus hrWaC http://nlp.ffzg.hr/resources/corpora/hrwac/

21. A note on the spelling of such compounds in Croatian is in order. Orthographic norms (e.g., Babić, Finka & Moguš, 1995; cf. Raffaelli, 2017) seem to agree that terms with a dual accent and compositionality of meaning are to be written with a hyphen, e.g., *zeleno-plava* 'green-blue'. Other forms, e.g., *maslinasto-zelena* / *maslinastozelena* are not as straightforward. In order to preserve consistency with extant literature on color terms (e.g., Raffaelli, 2017), we follow the argumentation for spelling presented there. When it comes to less frequent examples, e.g., *krvavocrvena* / *krvavo crvena* 'blood red', *mišjesiva* / *mišje siva* 'mouse grey', we present them in the most frequent spelling form found in hrWaC.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

color terms which were not elicited by the EoSS experiment, but can be found in hr-WaC. Most notably, it is the compound pattern OBJECT + (BASIC) COLOR, in which the first component can be either an adjective (e.g., *krvavo crvena* 'blood red', *mišje siva* 'mouse grey'), or a noun in nominative case (e.g., *limun žuta* 'lemon yellow', *grašak zelena* 'pea green'). These are, however, highly restricted to specific contexts, mostly fashion, make-up or home design. In Turkish, a similar pattern [Adj + Adj] consists of two color terms. The first one can be either simplex (e.g., *mavi-gri* 'blue-grey', *sarı-yeşil* 'yellow-green' – transitory color property) or modified with the suffixes *-(I)msI* or *-(I)mtırak*,[22] both denoting approximation (e.g., *mavimsi/ mavimtırak gri* 'bluish-grey', or *sarımsı/sarımtırak yeşil* 'yellowish-green' – a subtype of color). Color terms derived by the originally Arabic suffix *-ī* can also occur as the first component of the pattern [Adj + Adj] (e.g., *zeytunî yeşil* 'olive green', *limonî sarı* 'lemon yellow', *kuzgunî siyah* (lit. 'raven black') 'jet-black'). In Arabic the corresponding pattern is [Adj-/mu/Adj]. It forms its compound color terms by omitting the first letter ألف */alif/* of the second lexeme (i.e., the *alif* /ا/), which must belong to the first group of the basic color terms (in words like أصفر */**a**ṣfar/* 'yellow') and then, inserting a consonant ميم */m/* instead, followed by a short vowel */u/* known as ضمة *ḍamma*,[23] e.g., أبيض مصفر */abyaḍ muṣfar/* 'yellowish-white', برتقالي محمر */burtuqāliyy muḥmar/* 'reddish-orange'. As the examples show, the described process [/mu/Adj] serves to form adjectives denoting approximation. The first stem can be also omitted or substituted, whereas the second adjective[24] [/mu/Adj] can stand unaccompanied describing the color, e.g., لونه مزرق */lawnuhu muzraq/* 'its color is bluish', شكلها محمر */shakluhā muḥmar/* 'it looks like red', lit. 'its shape is reddish'.

2.3 Brightness and intensity modification

The fourth and final corresponding set of lexicalization patterns have to do with modification of color brightness and intensity with respect to hue. They are different in that they do not specify the demarcation lines of colors and their subtypes but refer to the intensity and brightness of particular colors. In Croatian, this is the [Adv + Adj] pattern.[25] The typical modifiers are *svijetlo* 'light' and *tamno* 'dark',

22. The vowels in the suffix *-(I)ms(I)* vary according to the principles of the vowel harmony.

23. See Ryding (2005, pp. 25–34).

24. Which can belong to the basic or non-basic color terms.

25. For this pattern we follow the word class definitions provided in Croatian grammars (e.g., Babić et al., 1995) and the analysis of EoSS data (Raffaelli, 2017). We also follow the spelling for adverbs (e.g., *svijetlo siv* 'light grey') as suggested by Babić et al. (1995, p. 81).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

e.g., *svijetlo plava* 'light blue', *tamno plava* 'dark blue'. Other adverbs in the EoSS data are *jarko* 'bright', *žarko* 'bright ('live coal – like')', *kričavo* 'flashy' ('screaming'), *fluorescentno* 'fluorescent', *zagasito* 'dull', *blijedo* 'pale' (Raffaelli, 2017). In Turkish, brightness, and intensity are expressed by the pattern [Adj + N/Adj] in which different adjectives occur as the first component, while the second component is a specific color. The most frequently used adjectives in this pattern are *açık* and *koyu*, e.g., *açık mavi* 'light blue', *koyu yeşil* 'dark green', while less frequently used are *acı* 'hot (spicy)', *tatlı* 'sweet', *donuk* 'dull', *canlı* 'vivid', *parlak* 'shiny', *menevişli* 'tempered', *ebruli* 'variegated, marbled', *fosforlu* 'fluorescent', *sütlü* 'milky', *uçuk* 'pale', *kirli* 'dirty'.[26] The primary meaning of the term *açık*, which is used to denote light colors, is not related to the concept of 'light', like *svijetlo* in Croatian, but to 'openness' (*açık mavi* lit. 'open blue'). The term *koyu*, which is used to denote dark colors, is etymologically related to a thick viscid liquid (Clauson, 1972), that is, to the concepts of 'density', 'intensity', and also 'depth'.[27] In Arabic, the [Adj + Adj] pattern is a description to the degree of the intensity of the color. The first component can be any color term, while the second component must define the color through adjectives like غامق */ghāmiq/* 'dark', lit. 'deep', فاتح */fātiḥ/* 'light' lit. 'open', قاتم */qātim/* 'dark', lit. 'intense', 'severe', خفيف */khafīf/* 'light', etc., e.g., احمر فاتح */aḥmar fātiḥ/* 'light red', and زيتي غامق */zaytiyy ghāmiq/* 'dark olive-green or dark oil-green color'. Therefore, similar concepts of 'depth' and 'intensity' for dark colors and 'openness' for light colors are found in Turkish and Arabic but are not attested in Croatian.[28] Main similarities and differences of lexicalization patterns in Croatian, Turkish, and Arabic are summarized in Table 1.

26. For more see Küçük (2010).

27. See http://www.nisanyansozluk.com/?k=koyu&lnk. In today's Turkish, in some contexts other than colors, they still have the meanings related to these concepts (e.g., *koyu çorba* 'potage', *koyu şive* 'thick accent', *koyu asit* 'strong acid', *koyu sohbete dalmak* 'move into deep conversation').

28. A very borderline case is found in Croatian dictionaries or orthography manuals, e.g., *otvoreno zelen* 'open green; intense light green' (Babić et al., 1995, p. 81), but a search in hrWaC yielded no results with respect to 'green', 'blue', 'yellow', or 'red'. Therefore, usage data for the pattern do not confirm any degree of conventionalization.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Lexicalization patterns of color terms in Croatian, Turkish, and Arabic

Croatian	Turkish	Arabic
9 simplex basic color terms	9 simplex basic color terms	6 simplex basic color terms
OBJECT FOR COLOR		
[stem + *asta*] *ljubičasta* 'purple', lit. 'violet-like' *maslinasta* 'olive-like'	[N–*ī*] *limoni* 'lemon-like' *turuncu* 'orange', lit. 'orange-like'.	[stem + ي /*iyy*/] وردي /*wardiyy*/ 'pink, lit. rose-like'
[N + N_{GEN}] (borderline) *boja višnje* 'color of cherry'	[N-$N_{POSS.3SG}$] (productive) *kahverengi* 'brown', lit. ' coffee-color'. *gök mavisi* 'sky-blue' *yavruağzı* (lit. 'baby mouth') 'light pink'	[N + N_{GEN}] (borderline) لون الصحراء/*lawnu aṣ-ṣaḥrā'-i*/ 'color of the desert'
HUE OVERLAP		
[Adj-o-Adj] *zeleno-plava* 'green-blue'	[Adj + N/Adj] *mavi-gri* 'blue-grey' *zeytuni yeşil* 'olive green'	[Adj-/mu/Adj] أبيض مصفر /*abyaḍ muṣfar*/ 'yellowish-white color'
BRIGHTNESS AND INTENSITY MODIFIER		
[Adv + Adj] *svijetlo/tamno-plava* 'light/dark blue'	[Adj + N/Adj] *açık/koyu mavi* 'open/thick blue'	[N + Adj] احمر فاتح/غامق /*aḥmar fātiḥ*/ *ghāmiq*/ 'light/dark red'

3. From object to color and back

In examining the relationship between colors and objects we observed two main trends in color naming. Firstly, objects used by all three languages are mainly rooted in the names of plants and flowers, and to some extent the main features of the environment (sky, sea, vegetation, etc.). Secondly, animals and cultural artifacts are rarely used in Croatian and Arabic but are more productive as the source domain in Turkish. In this chapter, we will examine another direction in which objects and colors are related and we will show that those very domains commonly used for color term formation are also enriched by using extant color terms in naming various phenomena. A common tendency in Croatian, Turkish, and Arabic is to form color-based lexemes from simplex color terms (most of them being basic color terms as well), thus completing the feedback loop between object and color at varying levels of morphosemantic productivity. What we mean by the morphosemantic productivity of colors is the way particular color roots/stems are used to form other lexemes via derivation and compounding (as well as multi-word

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

expressions). For instance, Cro. *crven* 'red.Adj' is used as a stem to form words such as *crven-ilo* 'redness of a particular body part', *crven-ica* 'terra rossa; a type of soil', *crven-dać* 'robin (bird)', *crven-o-kos* 'red-haired', among others. These lexemes can be viewed as forming a *morphosemantic field,*[29] i.e., a group of lexemes related both morphologically (based on the root) and semantically (based on the extensions of meaning which are metonymically or metaphorically related to the concept of 'red'). Different color terms have larger or smaller morphosemantic fields, i.e., more or fewer lexemes are formed via a particular color root. For example, a synonym to *crvenokos* 'red-haired' is *riđokos*, from the color term *riđi* '(a type of) red'. However, *riđa* 'red-brown' is an archaic color term which is highly collocationally restricted in use (used only in expressions such as *riđa kosa* 'red hair', *rid-o-kos* 'red-haired' or *riđi konj* 'horse with brownish-red hair') and thus has a significantly smaller morphosemantic field than *crven.* The same can be said of types of red, such as *grimizna* (morphologically complex) or *bordo* (loanword). Complex color terms do not participate in such derivational processes, as they are constrained by the fact that they are already morphologically complex. Thus the simplex, i.e., predominantly basic color terms are the base for further morphosemantic extensions.

3.1 Data collection and analysis

In order to investigate the morphosemantic productivity of particular color terms, we used dictionary data for Croatian, Arabic, and Turkish. Although there are certain shortcomings to lexicographic sources as opposed to, for example, corpus data, dictionaries have the advantage of containing conventionalized lexemes which are readily available for purposes of this research and offer a comparable source of lexemes. Furthermore, many of the lexemes found in dictionaries are underrepresented in corpora but still important when it comes to our analysis.[30]

Croatian data was collected using two monolingual dictionaries, *Rječnik hrvatskog jezika* (Šonje, 2000) and an online dictionary *Hrvatski jezični portal* (Croatian Language Portal).[31] In order to investigate the morphosemantic productivity of basic color terms in Turkish, we used data from two monolingual Turkish dictionaries (Püsküllüoğlu, 2005; Akalın et al., 2011) and one bilingual dictionary (Đinđić, 2014). Online dictionaries used for checking the data from written

29. See Raffaelli and Kerovec (2008) and Raffaelli (2013).

30. For instance, *modrulj* 'blue shark' < *modar* 'dark blue' is present in the hrWaC corpus (194 tokens), but *modruljica* 'titmouse (bird)' has no entries. Yet both lexemes are relevant from the point of view of derivation via *modar* 'dark blue' > animal property (blue color).

31. http://hjp.znanje.hr/.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Turkish dictionaries were the dictionary TDK[32] and the Turkish-English dictionary *Tureng*.[33] The Arabic data was collected using two different bilingual dictionaries: *'Hans Wehr' A Dictionary of Modern Written Arabic* (Milton Cowan, 1980), and *'Almawrid' A Modern Arabic-English Dictionary* (Baalbaki, 1995). The first one is a root-pattern Arabic-English dictionary and the second has an alphabetical order.

We extracted all the lexemes with a simplex color term into a database and morphologically segmented them. Three types of lexical units were investigated: derivatives, compounds, and multi-word units.[34] In comparing Croatian, Turkish, and Arabic we must point out that there are some pronounced structural differences in their formation. For example, the compound expressions for the 'white of the eye' or 'egg white' in Turkish consist of two components out of which neither is morphologically changed except the addition of the possessive suffix *-i* on the second component in order to form a genitive construction, e.g., *göz akı* < *göz* 'eye' + *ak* 'white', *yumurta akı* < *yumurta* 'egg' + *ak* 'white' + *i*. In Croatian, the terms for the same entities are clearly morphologically changed with additional affixes (compounding and derivation in *bjel-o-oč-n-ica* 'white of the eye' < *bijel* 'white' + *oko* 'eye' and derivation in *bjel-anjak* 'egg white'). Additionally, the interfix *-o-* used for forming compounds in Croatian has no equivalent in Turkish. In Turkish, a two-lexeme construction without any morphological change is also a very common pattern for naming animals and plants (e.g., *kızıl gerdan* or *kızılgerdan*[35] 'robin', lit.

---

32. *TDK* is the abbreviation for *Türk Dil Kurumu* 'Turkish Linguistic Society'.

33. http://www.tdk.gov.tr/index.php?option=com_gts&view=gts
http://tureng.com/tr/turkce-ingilizce

34. Multi-words units are sometimes called collocations (see Cruse, 1986) or phrasal names (see Booij, 2009) based on the fact that they share the naming function with words, and often the divide between multi-word units and compounds proper (single-word units) is not clear cut.

35. It has to be noted that in Turkish it is not always clear whether a unit is a compound or a multi-word expression (collocation) since in many cases different Turkish dictionaries are not consistent in the way of noting it. In the case of the name for the bird robin, it is noted as a compound in TDK (*kızılgerdan*) while in the online *Tureng* dictionary both ways of noting appear (*kızıl gerdan* and *kızılgerdan*). Sometimes, even in the same dictionary, the "same" unit is noted differently. For example, in the online dictionary of TDK two variants of the "same" unit meaning 'hydroelectric power' are noted: *beyaz kömür* ve *beyazkömür*, lit. 'white coal'. The difficulty of discerning compounds from multi-word expression comes also from the fact that in Turkish compounds, as noted previously, there are usually no additional morphemes which would link the two components or change their original form (for example, the linking morpheme *o* in Croatian compound *plav-o-kos* 'blond' from *plav* 'blue' + *kosa* 'hair' and the loss of final *-a* in the noun *kosa* 'hair' of the same compound). For the purposes of this chapter, especially in the quantitative analysis of lexical units, we used TDK dictionary since Turkish Linguistic Society is a reference point for standard Turkish language.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

'red neck'), which is not a common pattern in Croatian (compare the derivative *crven-dać* 'robin'). In that sense Arabic functions more like Turkish than Croatian, e.g., the term بياضُ البيضِ 'albumen, white (of egg)' stands for بياض 'white or whiteness' + البيض 'eggs-genitive case'.

These structural preferences in word-formation strategies are supported by the number of different kinds of lexical units (derivatives, compounds, and multi-word units) found in our data. Thus for all examined colors in Turkish, there are 296 derivatives, 208 compounds proper, and 180 multi-word units; in Croatian, there are 141 derivatives, 106 multi-word units, and 40 compounds, and in Arabic, there are 133 derivatives, 89 multi-word units, and no compounds.[36]

As for the productivity of color roots, there are some notable differences among the three languages. The productivity cline from most to least productive is the following:

1. Croatian: 'black' (70), 'white' (59), 'red' (42), 'green' (33), 'yellow' (22), 'blue' (12), 'grey' (16), and 'brown' (8);
2. Turkish: 'black' (221), 'white' (152), 'red' (129), 'yellow' (55), 'green' (33) 'blue' (32), 'grey' (23), 'purple' (23), 'pink' (11), 'orange' (4), and 'brown' (1);
3. Arabic: 'white' (65), 'black' (35), 'yellow' (33), 'green (32), 'red' (28), 'blue' (18), 'pink' (8), 'brown' (2), 'purple' (1).

The productivity cline could provide some interesting insights into the saliency of certain colors over others. For instance, 'black', 'white', and 'red' can carry certain experientially based connotations (emotion, mood, danger, etc.) which can be exploited in word-formation as well. However, such an analysis exceeds the limitations of this chapter as it would require establishing a broader connection with extant phraseological and interdisciplinary studies (from psychology, sociology, and anthropology).

Interestingly, Croatian and Turkish color root productivity follows (with slight variation) the Berlin and Kay (1969) hierarchy of color term evolution, while the same cannot be said for Arabic. However, it has to be noted that the above results for Turkish 'black', 'white', 'red', 'blue', and 'grey' include both basic color terms borrowed from Persian, Arabic, and French (see Section 2.1.1.) and original Turkish terms. These original terms were not included as basic by Özgen and Davis (1998) but are much more productive than their basic-color counterparts in the formation of compounds, multi-word expressions (collocations) and proverbs.

36. These numbers do not present dictionary entries, but different meanings, i.e., a dictionary entry or a lexical form which has, for instance, three different meanings was counted as three, and not one lexical unit.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.2 Domains lexicalized by color terms

In order to get a clear insight into the types of meaning extensions (in terms of what referents are commonly named in regard to color roots), we added a target domain for each of the analyzed lexemes, e.g., Cro. *plav-uša* 'blonde' – PERSON PROPERTY, Tur. *gökgüvercin* 'stock dove', lit. 'sky/blue pigeon' – ANIMAL, Ar. السودان /As-Sūdān/ 'Sudan' – GEOGRAPHY. As we mentioned previously (Chapter 2), some objects such as plants, animals, and environmental entities form the backbone of lexicalization of non-basic color terms in all three languages, to varying degrees. Through examining lexemes formed with a color root, we have noticed that the same is true in the opposite direction, i.e., that different entities from domains of plants, animals, and the environment are commonly lexicalized via the strategy COLOR FOR OBJECT. However, there are particular trends in naming strategies that surface from the investigation of the data collected so far for the three languages. Our observations can be summarized as follows:

a. Color as a visual property of a body part is commonly used to denote the entire entity. This is particularly common in naming plants, mushrooms, and animals, e.g., Cro. *crvendać* 'robin', *bjelouška* 'grass snake', lit. 'white-ear loaf', Tur. *kızıl gerdan* 'robin', lit. 'red neck', *sarıca* 'wasp', *sarıgöz* 'sea-bream', lit. 'yellow eye', Arab. أخضَر /akhḍar/ 'greenfinch', lit. 'green'.
b. Color as a visual symptom is used to denote medical conditions in general, e.g., Cro. *žutica* 'yellow fever', *crveni vjetar* 'erysipelas', lit. 'red wind', Tur. *sarılık* 'hepatitis', *kızamık* 'rubella', Ar. صَفَر /ṣafar/ 'jaundice, yellows, icterus', حميراء /ḥumayrā'/ 'German-measles, rubella'
c. Color of a (body) part is used to denote an entire group or subcategory, e.g., Cro. *bijelac* 'white person', *crnka* 'black-haired woman', *plavuša* 'blonde', *crnokošuljaš* 'a fascist', lit. 'black-shirted', Tur. *beyaz* 'white person', *sarışın* 'blonde', lit. 'yellow-like', Ar. أسود /aswad/ 'a male person of African descent', lit. 'black', صفّار /ṣaffār/ 'brass founder' خُضري /khuḍariyy/ 'greengrocer'.
d. Color of an entity can be synonymous with the entity itself. This is particularly common for environmental phenomena that are usually considered as visual prototypes for basic colors (see Wierzbicka, 1990). Terms such as Cro. *nebo* 'sky' and *more* 'sea' have synonyms such as *plaventilo* 'blueness' (lit.), 'sky', 'sea' and 'sea' or *zelenilo* 'green vegetation', Tur. *gök mavisi* 'sky-blue, sky', *yeşillik* 'vegetation; green vegetables; potherbs'. Ar. الزرقاء /az-zaarqā'/'the blue' (lit.), 'the sky', 'the firmament'.
e. Color of an entity is used to denote some other property or quality of the same entity. This category often has metaphorical meanings attached to color as well, due to connotations ascribed to cultural values of colors and mediated by our

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

knowledge of the world, e.g., Cro. *zelembać* (lit. 'green person') 'inexperienced person',[37] *žutokljunac* (lit. 'yellow-beaked') 'a novice', Tur. *yeşil* (lit. 'green') 'underripe', *kızıl* (lit. 'red') 'exaggerated', 'communist', Ar. سوداوي /sawdāwiyy/ 'melancholic, melancholy; melancholiac'.

f. Color roots used to form verbs, adjectives and some nouns are used to construe processes, attributes, and color as substance, related in a generic way to color itself, e.g., Cro. *crniti* 'to make something black', *crnkast* 'blackish', *crnilo* 'black substance used for coloring', Tur. *karartmak* 'to make something black', *karanlık* 'blackness, darkness', *beyazlatıcı* 'whitener, bleach', Ar. بَيَضَ /bayyaḍa/ 'to whiten, make white or whiter', مُبيض/mubyaḍ/ 'whitish', مُبيّض /mubayyiḍ/ 'bleaching powder'. This last category is significant because it is most regular with respect to affixes used to form such types of lexemes in the three languages. This, perhaps, has to do with generic word-class meanings which are denoted by these lexemes. In contrast, previously mentioned naming strategies use different and varied affixes both intra- and inter-linguistically.

Some differences have to be pointed out as well. In Arabic, for example, some colors are used as synonyms for some countries, e.g., الخضراء /al-khaḍrā'/ 'epithet of Tunis', (lit. 'the verdant'). The synonym for 'paradise' is الخضيراء /al-khuḍayrā'/.[38] Two other interesting examples are الاصفران /al-aṣfarān/ lit. 'the two yellows' and الاحمران /al-aḥmarān/ lit. 'the two reds' which have the meaning 'saffron and gold' and 'wine and meat', respectively.

Turkish is, in comparison to Croatian and Arabic, specific in the fact that it uses originally Turkish color terms more productively than loanwords (which are considered to be basic color terms) in the formation of derivatives, compounds, and multi-word expressions. As noted by Bayraktar (2004, 2005, 2006), who has done an extensive work comparing the meanings and the use of the borrowed and originally Turkish terms, the borrowed terms carry rather neutral meanings related to a color itself, while the originally Turkish terms can, besides colors, relate to abstract concepts as well. Originally Turkish terms are also much more productive in the formation of toponyms and in the naming of plants and animals. Such a low productivity of basic color terms and the coexistence of synonymous terms for cognitively salient colors, which cannot be seen in Croatian and Arabic and is thus specific for Turkish, shows the impact of historical and sociocultural factors on the formation of the vocabulary.

---

37. For example, the knowledge that green fruits are not ready to be consumed is related to age and, metaphorically, to experience.

38. 'Green' is connected to the concept of 'paradise' based on cultural and environmental reasons. More specifically, in a desert-like climate greenery is rare, therefore, it symbolizes abundance.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Finally, some morphological devices, as we mentioned previously, are used to a different extent in the three languages. Notably, Turkish uses conversion to a much greater extent than Croatian and Arabic, while in Croatian and Arabic, a change of meaning is usually accompanied by a morphological change.

Besides differences, some similarities can also be observed. For instance, some artifacts are formed similarly because of their cultural salience in all three languages. A good example is the name for 'mourning garments', Cro. *crnina*, Tur. *siyahlar*, Ar. سواد /sawād/, which has to do with the high salience of some connotations ascribed to particular colors in different languages.

4. Concluding remarks and future work

In this chapter, we focused on lexicalization patterns of color terms and color-based lexemes in three typologically, genetically, and socioculturally different languages. In spite of the structural differences among the three languages, we have established some similarities in the lexicalization patterns used by all three. We also pointed out some notable differences, some of which are structural and some sociocultural. These similarities and differences can be summarized as follows:

1. As for lexicalization patterns, we determined three similar lexicalization strategies in all three languages, which we divided into three categories: suffixation, genitive construction and compounding. Despite similarities of these lexicalization strategies, the three languages differ in the productivity of a particular lexicalization pattern used to form secondary color terms. A good example is the [N + N_{GEN}] / [N-$N_{POSS.3SG}$] construction, which is a borderline lexicalization pattern in Croatian and Arabic, but fully productive in Turkish. The productivity of this pattern in Turkish allows for some constructions which were not found in Croatian and Arabic (e.g., *kavuniçi* 'pale orange', lit. 'melon inside').
2. Although in all three languages there is a strong tendency to use plants and different parts of the environment in the OBJECT FOR COLOR strategy, some differences can be observed in color conceptualization. These differences depend in some degree on specific sociocultural and environmental facts. For instance, in Turkish, the terms for sociocultural artifacts can appear as a component of the pattern [N + N_{GEN}] / [N-$N_{POSS.3SG}$], which is not the case in Croatian and Arabic. Additionally, the terms related to animals appear more frequently as a component of this pattern in Turkish than in Croatian and Arabic.
3. All three languages use a similar pattern to denote color intensity and transitory colors, i.e., colors which overlap in hue. Nevertheless, Turkish and Arabic use similar concepts of 'depth' and 'intensity' for dark colors, and 'openness' for light colors, while Croatian uses the concepts related to 'light'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

In this chapter, we also aimed to establish a basis for a contrastive analysis of color-based lexemes, especially with regard to the morphosemantics of color terms. This includes examining the intricate relationship between colors and objects (i.e., entities related to different concrete and abstract experiences), and 'the feedback loop' that enriches the lexicon in both directions – from object to color and back. By examining the direction 'COLORS → OBJECTS', we have shown that those very domains commonly used for color term formation are also enriched by using extant color terms in naming various phenomena, i.e., that different entities from domains of plants, animals, and the environment are commonly lexicalized via the strategy COLOR FOR OBJECT. Although it is evident that in languages of the world there is a firm relation between experientially basic entities (animals, plants, the environment) and colors, we hope that we added a new perspective of this relation, and showed that they are mutually intertwined in many different ways.

Investigating the strategy COLOR FOR OBJECT we have found that there are some pronounced structural differences among the three languages: Croatian often uses different affixes in the formation of color-term based expressions, while in Turkish and Arabic, a morphological change of words' constituents is less common. These structural preferences in word-formation strategies are reflected in the number of different kinds of lexical units (derivatives, compounds, and multi-word units) in the three languages. Notable differences among the languages are also found in the productivity of different color terms in the word-formation process. Although Croatian and Turkish show some similarities in this regard, Turkish can be considered as specific in the fact that it uses more extensively the color terms of Turkish origin which are not acknowledged as basic color terms.

As our analysis shows, the lexicalization patterns of color terms have certain properties of a contained system of structural relations, which allowed us to examine them in greater detail. On the other hand, we are aware that future studies are needed in order to fully investigate the morphosemantics of color terms as presented in the second part of the study. For this reason, we feel it is necessary to suggest future directions of study, which would require: (a) a more detailed study of the meanings of numerous affixes which appear in our data since they are a part of a larger system of morphological devices in each language, (b) an investigation of the productivity of color roots in relation to other productive roots in these languages, (c) a study of connotations and values of color terms as an interdisciplinary field of research. However, we hope that we presented a novel contrastive analysis of three languages which, as to our knowledge, were not studied contrastively from the point of view of lexicalization processes and the role colors play in the formation of the Croatian, Turkish, and Arabic lexicon.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

References

Al-Adaileh, B. A. (2012). The connotations of Arabic colour terms. *Linguistica online*, 1–18.

Al-Rasheed, A. S., Al-Sharif, H. H., Thabit, M. J., Al-Mohimeed, N. S., & Davies, I. R. L. (2011). Basic colour terms of Arabic. In C. P. Biggam, C. A. Hough, Ch. J. Kay & D. R. Simmons (Eds.), *New directions in colour studies* (pp. 53–58). Amsterdam/Philadelphia: John Benjamins Publishing Company. https://doi.org/10.1075/z.167.10alr

Babić, S., Finka, B., & Moguš, M. (1995). *Hrvatski pravopis*. Zagreb: Školska knjiga.

Bayraktar, N. (2004). Kara ve Siyah renk adlarının Türkçedeki kavram ve anlam boyutu üzerine. *TÖMER Dil Dergisi*, 126, 56–77.

Bayraktar, N. (2005). Kavram ve anlam boyutunda al, kırmızı ve kızıl. *International Journal of Central Asian Studies Mustafa Canpolat Armağanı*, 10 (1), 145–165.

Bayraktar, N. (2006). Kavram ve anlam boyutunda Türkçede ak ve beyaz. In C. Yıldız, & L. Beyreli (Eds.), *Dilbilim, Dil Öğretimi ve Çeviribilim Yazıları I–II* (pp. 35–50). Ankara: Pegem-A Yayıncılık.

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

Booij, G. (2009). Phrasal names: A constructionist analysis. *Word Structure*, 2 (2), 219–240. https://doi.org/10.3366/E1750124509000427

Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.

Čaušević, E. (1996). *Gramatika suvremenoga turskog jezika*. Zagreb: Hrvatska sveučilišna naklada.

Göksel, A., & Kerslake, C. (2005). *Turkish. A Comprehensive Grammar*. Routledge: London & New York.

Hasan, A. A., Al-Sammerai, N. S. M., & Kadir, F. A. B. A. (2011). How Colours Are Semantically Construed in the Arabic and English Culture: A Comparative Study. *English language teaching*, 4 (3), 206–213. https://doi.org/10.5539/elt.v4n3p206

Hrnjak, A. (2009). Secondary colours in Croatian and Russian phraseology. In E. E. Stefanskij (Ed.), *Kartina mira v slavjanskih i romano-germanskih jazykah* (pp. 82–86). Samarskaja gosudarstvennaja akademija.

Kay, P., & Regier, T. (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences*, 100 (15), 9085–9089. https://doi.org/10.1073/pnas.1532837100

Küçük, S. (2010). Türkiye Türkçesinde renk andları ve özellikleri. *Uluslararası Sosyal Arastırmalar Dergisi / The Journal of International Social Research*, 3 (10), 420-427. http://www.sosyalarastirmalar.com/cilt3/sayi10pdf/kucuk_salim.pdf

Masliyah, S. (2004). Colours in Iraqi Arabic. *Zeitschrift Für Arabische Linguistik*, 43, 26–42.

Özgen, E., & Davies, I. R. L. (1998). Turkish color terms: tests of Berlin and Kay's theory of color universals and linguistic relativity. *Linguistics*, 36 (5), 919–956.

Raffaelli, I. (2017). Conventionalized patterns of colour naming in Croatian. In K. Cergol Kovačević & S. L. Udier (Eds.), Proceedings of the XXIX. CALS conference (pp. 171–186). Peter Lang: Frankfurt a/M.

Raffaelli, I. (2013). The model of morphosemantic patterns in the description of lexical architecture. *Lingue e linguaggio*, 1, 47–72.

Raffaelli, I., & Kerovec, B. (2008). Morphosemantic fields in the analysis of Croatian vocabulary. *Jezikoslovlje* 9 (1–2), 141–169.

Ronga, I., & Bazzanella, C. (2015). The emergence of color categories: Variance and invariance. *Proceedings of the EuroAsianPacific Joint Conference on Cognitive Science*, 222–227.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rosch, E. H. (1973). Natural categories. *Cognitive psychology*, 4 (3), 328–350. https://doi.org/10.1016/0010-0285(73)90017-0
Ryding, K. C. (2005). *A Reference Grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.
Ryding, K. C. (2005b). *A reference grammar of modern standard Arabic*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486975
Sandford, J. L. (2014). *Her blue eyes are red: An idealized cognitive model of conceptual color metonymy in English*. In W. Anderson, C. P. Biggam, C. Hough & Ch. Kay (Eds.), Color studies: A broad spectrum (pp. 109–125). Amsterdam/Philadelphia: John Benjamins.
Topçu, N. (2001). Fransızca ve Türkçe renk isimleri içeren deyimlerin karşılaştırmalı incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 20, 131–140.
Ullmann, S. (1966). Semantic universals. In J. H. Greenberg (Ed.), *Universals of language* (pp. 217–262). Cambridge, MA: MIT Press.
Wehr, H., & Cowan, J. M. (1979). *A dictionary of modern written Arabic (Arab.-Engl.)*. Wiesbaden: Otto Harrassowitz Verlag.
Wierzbicka, A. (1990). The meaning of color terms: semantics, culture, and cognition. *Cognitive Linguistics*, 1 (1), 99–150. https://doi.org/10.1515/cogl.1990.1.1.99
Wierzbicka, A. (2005). There are no "color universals" but there are universals of visual semantics. *Anthropological linguistics*, 47 (2), 217–244.

Dictionaries

Akalın, Ş. H. et al. (Eds.). (2011). *Türkçe Sözlük*. Ankara: Türk Dil Kurumu.
Baalbaki, R. (Ed.). (1995). *Al-Mawrid: A Modern Arabic-English Dictionary*. Beirut: Dar El-ilm Lilmalayin.
Canpolat, M. et al. (Eds.). (1983). *Türkçe Sözlük*. Ankara: Türk Dil Kurumu Yayınları.
Clauson, G. (1972). *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*. Oxford: Oxford University Press.
Đinđić, M. (2014). *Yeni Türkçe-Sırpça Sözlük*. Ankara: Türk Dil Kurumu Yayınları.
Milton Cowan, J. (Ed.). (1980). *A Dictionary of Modern Written Arabic, Hans Wehr*. Beirut: Librairie Du Liban.
Püsküllüoğlu, A. (2005). *Arkadaş Türkçe Sözlüğü*. Ankara: Arkadaş Yayınevi.
Šonje, J. (2000). *Rječnik hrvatskoga jezika*. Zagreb: Leksikografski zavod Miroslav Krleža i Školska knjiga.

Online dictionaries and corpora

http://www.nisanyansozluk.com/
Hrvatski jezični portal, http://hjp.znanje.hr/
http://nlp.ffzg.hr/resources/corpora/hrwac/
http://www.tdk.gov.tr/index.php?option=com_gts&view=gts

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ossetic color terms system

Arseniy Vydrin
Institute for Linguistic Studies of the Russian Academy of Sciences,
Saint Petersburg

In this chapter, I determine Iron Ossetic[1]'s basic color terms and the place of the Iron Ossetic[2] system in the evolutionary sequence of the development of basic color terms. I examine the semantics and use of basic and non-basic color terms, morphological and lexical color modifiers, the main patterns used to describe color in modern Ossetic as well as connotations of color terms. I conclude by comparing the peculiarities of the Ossetic color system with the systems of some other Iranian languages.

This study is based on the methodology proposed by A. Majid, F. Jordan, and M. Dunn (2011). The Ossetic material is taken from the field data collected in North Ossetia, Russia, and data from the Ossetic corpora[3].

Keywords: Iranian languages, Caucasian languages, Ossetic, color terms, semantics, morphology

---

1. Ossetic is a modern Eastern Iranian language spoken in the republics of North Ossetia (Russia) and South Ossetia. The total number of native Ossetic speakers in the world is about 500,000 people (Ethnologue 2017). There are two main dialects, Iron and Digor. Iron is considered the main dialect. I will use the term Ossetic for Iron Ossetic hereinafter.

2. I thank my Ossetic native speakers for their patience. I would like to express my deep gratitude to Madina Darchieva, the Ossetic native speaker born in Alagir, North Ossetia, and permanently living there, who keeps helping me with Ossetic data. The examples from Ossetic National Corpus included in this chapter were cross-checked by her. I also thank the anonymous reviewers for their critiques and comments.

3. I used Ossetic National Corpus (ONC, 12 million tokens at the time of submission, http://corpus.ossetic-studies.org/en) and Ossetic Oral Corpus (60,000 tokens, http://ossetic-studies.org/en/texts).

https://doi.org/10.1075/sfsl.78.17vyd
EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use
© 2019 John Benjamins Publishing Company

1. Previous studies

There are a number of existing studies of color words, their etymologies, meanings, and even connotations in Iranian languages. For instance, Khotanese colors (Rossi, 2007), colors in Middle Persian (Čunakova, 1994), color terms in Eastern Iranian languages (Edel'man, 2009), etc. Among the most significant recent studies is an article by Adriano Rossi (2006), which gives a history of the study of color terms in Iranian languages and provides important references on the topic. Ossetic color terms have also attracted scholars; however, most of the studies have been focused on the etymology of Ossetic color terms (see Abaev's etymology dictionary of Ossetic (1958, 1973, 1989)).

There are a few studies of color terms in Iranian languages based on typologically relevant works and appropriate linguistic methods. One of them is a study by Erika Friedl of color terms in Southwest Iran (Friedl, 1979). However, there are no special linguistic studies of the Ossetic color system based on the latest typologically oriented approaches to color.

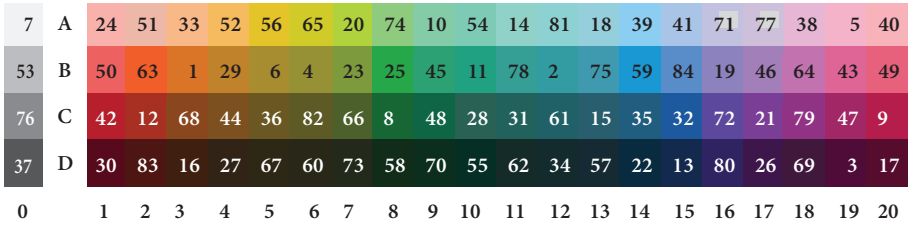
This chapter presents a case study of modern Ossetic (Iron) color terms. My principal aims are to describe the Ossetic color system and to highlight its peculiarities.

2. Methodology

This chapter is mainly based on the Evolution of Semantic Systems (EoSS) project methodology (Majid, Jordan, & Dunn, 2011). The EoSS project was organized by the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands. The principal investigators are Asifa Majid, Fiona Jordan, and Michael Dunn. The project aims to collect data from 50 Indo-European languages in order to assess the degree and nature of variation in semantic categories over space and time. It studies four lexical domains: containers (objects), body parts (parts), colors (attributes), and spatial relations (relations). The data for Ossetic was collected in 2013 by Oleg Belyaev and me. I carried out the analysis of the data for color terms and the collection of missing data for this chapter from native speakers and the Ossetic corpora. Since only the colors are relevant for the current study, I will briefly describe the color data collection procedure. We used the EoSS project material to collect the Ossetic color-naming data (Majid & Levinson, 2007; Majid, Jordan, & Dunn, 2011). The stimuli are standardized Munsell colors. The EoSS project uses the reduced version color kit of the World Color Survey (Kay, Berlin, Maffi, Merrifield, & Cook, 2009). It has 84 chips: 4 chips are achromatic, while 80 vary in hue, brightness and saturation. There are 20 equally spaced hues at 4 degrees of brightness. Saturation varies such that colors are generally at the maximum possible chroma for that point in the color space (see Table 1).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Color chips



Every chip has a number from 1 to 84. Correspondences between the numbers and Munsell codes are shown in Table 2.

Table 2. Stimulus number vs. Munsell code

Stimulus number	Munsell code	Stimulus number	Munsell code	Stimulus number	Munsell code
1	5YR 6/14	29	10YR 6/12	57	5B 2/6
2	10BG 6/8	30	5R 2/8	58	10GY 2/4
3	5RP 2/8	31	5BG 4/8	59	10B 6/10
4	10Y 6/10	32	5PB 4/12	60	10Y 2/2
5	5RP 8/6	33	5YR 8/8	61	10BG 4/6
6	5Y 6/10	34	10BG 2/6	62	5BG 2/6
7	N9.5	35	10B 4/10	63	10R 6/14
8	10GY 4/8	36	5Y 4/6	64	10P 6/10
9	10RP 4/14	37	N2	65	10Y 8/12
10	5G 8/6	38	10P 8/6	66	5GY 4/8
11	10G 6/10	39	10B 8/6	67	5Y 2/2
12	10R 4/12	40	10RP 8/6	68	5YR 4/8
13	5PB 2/8	41	5PB 8/6	69	10P 2/6
14	5BG 8/4	42	5R 4/14	70	5G 2/6
15	5B 4/10	43	5RP 6/12	71	10PB 8/4
16	5YR 2/4	44	10YR 4/8	72	10PB 4/12
17	10RP 2/8	45	5G 6/10	73	5GY 2/2
18	5B 8/4	46	5P 6/8	74	10GY 8/8
19	10PB 6/10	47	5RP 4/12	75	5B 6/10
20	5GY 8/10	48	5G 4/10	76	N4.5
21	5P 4/12	49	10RP 6/12	77	5P 8/4
22	10B 2/6	50	5R 6/12	78	5BG 6/10
23	5GY 6/10	51	10R 8/6	79	10P 4/12
24	5R 8/6	52	10YR 8/14	80	10PB 2/10
25	10GY 6/12	53	N7	81	10BG 8/4
26	5P 2/8	54	10G 8/6	82	10Y 4/6
27	10YR 2/2	55	10G 2/6	83	10R 2/6
28	10G 4/10	56	5Y 8/14	84	5PB 6/10

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Initially, we tested the participants for color-blindness using the EoSS material. We had 9 color plates, consisting of different color patches (the number vs. the background). The participants were asked to name the numbers and we continued only with the participants who had passed the color-blindness test. The color-naming data were elicited from 15 native Ossetic speakers: both men and women from 15 to 60 year of age. All are permanent residents of North Ossetia, Russia, and bilingual in Russian. We mainly interviewed high school students and their teachers.

The main experiment consisted of two tasks, namely, the color-naming task and the focal-color test. In the color-naming experiment, we used 84 color chips, which were organized in a fixed random order. The participants were shown each chip and asked to name the color. In the focal-color task, the same 84 color circles were used. The participants were asked to indicate which circle was the best example of each basic color term. The material for the experiment was developed by Majid (2008; see also Majid, Jordan, & Dunn, 2011). The focal colors are the best examples of each basic color term. The key question about the focal colors (both in a certain language and cross-linguistically) addressed in theoretical literature is whether there is a tight clustering of focal colors or if they are diffused. For more information about focal colors see Berlin and Kay (1969), Davidoff, Davis, and Roberson (1999), Roberson, Davies, and Davidoff (2000), Regier, Kay, and Cook (2005).

3. Basic color terms

Typological studies of color terms distinguish between basic color terms and color terms more generally. This distinction was first made by Brent Berlin and Paul Kay (1969) who suggested the following criteria to establish basic color terms in a particular language (p. 6):

1. The basic color term is monolexemic: its meaning is not predictable from the meaning of its parts (e.g., *salmon-colored* is not a basic color);
2. The basic color term signification is not included in that of any other color (e.g., *crimson* vs. *red*; *blond* vs. *white*);
3. The basic color term application must not be restricted to a narrow class of objects (e.g., *blond*);
4. The basic color term must be psychologically salient for informants. Indices of psychological salience include, among others, (a) a tendency to occur at the beginning of elicited lists of color terms, (b) stability of reference across informants and across occasions of use, and (c) occurrence in the idiolects of all informants (e.g., *the color of the rust on my aunt's old Chevrolet*).
   For difficult cases Berlin and Kay suggest several additional criteria (pp. 6–7):

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

5. The forms under question should have the same distributional potential as the previously established basic terms. For example, English basic colors are used with the suffix *-ish* (*reddish*, *whitish,* etc.), while general colors are not (e.g., *aqua color*, **aquaish*);
6. Color terms that are also the name of an object characteristically having that color are suspect, e.g., *gold*, *silver,* etc.;
7. Recent foreign loan words may be suspect;
8. In cases where lexemic status is difficult to assess (see criterion 1.), morphological complexity is given some weight as a secondary criterion. The English term *blue-green* might be eliminated by this criterion.

Since Berlin and Kay's work, there has been a lot of literature published on basic colors criteria. However, most of the linguists working on color terms are still guided by Berlin and Kay's criteria. In my study, I used the aforementioned criteria to distinguish basic color terms in Ossetic.

According to the data collected from our color naming experiment, there are six basic colors in Ossetic, namely *šaw*[4] 'black', *urš* 'white', *šərχ* 'red', *bur* 'yellow', *cʔɜχ* 'green, blue, grey', and *morɜ* 'brown'. They consist of a single word without any affixes; they convey only color; they are not restricted to a narrow class of objects, and all are psychologically salient for native Ossetic speakers.

The Ossetic basic color terms can function as attributes or as nouns, in the latter case expressing the meaning 'something of this color', e.g., *šərχ* 'something red, redness'. The color terms are not unique in this function. Generally, Ossetic does not have a clear-cut distinction between nouns and adjectives. There is a large class of nouns which can function both as heads of an NP or as attributes, e.g., *rɜšuyd* 'beautiful', 'a beauty', *žən* 'difficult', 'difficulty'. There are also nouns that can be marked by the attributive suffix of the comparative degree *-dɜr*, e.g., *lɜg-dɜr* (man-COMPAR) lit. 'more man'. The basic color terms, which can be used either as an attribute or as a noun, are within the general logic of the Ossetic nouns. In Section 5, I will show that some basic color terms can also be nominalized by a special suffix.

After the color-naming experiment, we elicited 'focal colors' of the basic color terms from 12 participants, from whom we have collected the color-naming data. We got the following data for the Ossetic foci of the basic color terms:[5]

---

4. All the examples are transcribed in accordance with the modern standard (Iron) Ossetic pronunciation (for details see Dzahova, 2009). In general, the most important phonetic differences are connected to the pronunciation of the consonants (in Ossetic script) *c*, *ɜ*, *ð3* and *ц*. In most cases, the letter *c* is pronounced as /š/, *ɜ* as /ž/, *ð3* as /z/, and *ц* as /s/, transcribed into IPA symbols.

5. The Latin letter with the number means the cell in the color Table 1. The figure in the brackets shows the number of the participants who chose this answer.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. *šaw* ‘black’: 0D (9), 4D (1), 5D (1), 6D (1), 7D (1)[6]
2. *urš* ‘white’: 0A (12)
3. *šərχ* ‘red’: 1C (11), 2C (1)
4. *bur* ‘yellow’: 5A (7), 4A (4), 6A (2)
5. *cʔɜχ* ‘green, blue, grey’: 14C (3), 0C (2), 15D (2), 0B (1), 8C (1), 14B (1), 10B (1), 9C (1)
6. *morɜ* ‘brown’: 3D (4), 3C (2), 2D (2), 4C (2), 0D (1), 0C (1)

Most of the basic colors are focal. The best example is *urš* ‘white’, which was described using the same circle by all participants. The most diffused basic color is *cʔɜχ*, with nearly every participant describing it differently, choosing ‘green’, ‘blue’, or zero (‘grey’) hues with varying brightness. According to our color-naming data, the largest number of answers for *cʔɜχ* have the following chips: 10B, 11C, 13C, 9B, 10C, and 12C. Between 10 and 12 participants chose *cʔɜχ* to describe these chips. If we start to broaden the number of answers considered, the number of color circles which can be named by *cʔɜχ* will gradually increase (see Table 3 where the figures mean the number of native speakers who chose *cʔɜχ* to describe this color):

Table 3. Color-naming test for *cʔɜχ*, from 7 to 12 answers (10–12 answers are marked in bold)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A									8	7	8										
B							7	7	8	11	12	9	8								
C							7			9	10	12	10	11	8	8					
D												9	10	9	7	9					

If we broaden the number of answers again (from 1 to 6) we get the picture shown in Table 4.

Table 4. *cʔɜχ* color-naming test from 1 to 6 answers

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A							2	5				5	3	2	3	1					
5 B						2								5	6	5	1				
5 C						4		3	3								2	1	1		
D						1		1	4		6						4	1	1		

The color-naming test shows that it is easier to describe the distribution of the color term *cʔɜχ* by saying where it cannot be used. It is never used in hues 1–4 and 19–20,

6. One of the participant gave two answers.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

and it is very rarely employed in hues 17–18. It is a widely diffused color, spreading from the 5th to the 16th hue and from A to D in brightness, while also serving two achromatic chips of brightness B and C.

It is worth noting that the achromatic use of *cʔɜχ* has only 5 answers, which is less than half its use in 'green'–'blue' hues (see Table 3). However, the frequency of *cʔɜχ* in achromatic space is higher than the frequency of any other Ossetic color term. For example, the chip 0B, besides *cʔɜχ*, is served by *fɜnək-χʷəž* (ash-color) which has only three answers and some other color terms which have only one answer each: *kɜrdɜg-χʷəž* (lit. grass-color), *arv-χʷəž-ɜj tar-dɜr* 'darker than the sky color' (sky-color-ABL dark-COMPAR), *miɣ-χʷəž* (cloud-color), *morɜ-χʷəž* (brown-color), *šaw urš* (black white), *urš šaw-imɜ žmɜšt* 'a turbidity of white with black' (white black-COM turbidity). The chip 0C, other than 5 answers of *cʔɜχ*, has 4 answers of *fɜnək-χwəž* (ash-color), other color terms have only one answer each: *morɜ* 'brown', *morɜ-χʷəž* (brown-color) 'brown color', *miɣ-t-ə χʷəž* (cloud-PL-GEN color) 'the color of clouds', *seryj* 'grey' (a loanword from Russian), *dur-χʷəž* (stone-color) 'stone color', *urš šaw-imɜ* (white black-COM) 'white with black'. Although *cʔɜχ* has only 5 answers in 0B and 0C, it is the most common color term in achromatic space.

The results of the focal-color test are also widely dispersed. As shown above, we got 8 different answers for the best chip for *cʔɜχ* from the 12 participants asked. The only relevant result from the focal-color test is the fact that none of the participants chose a chip with brightness A. The color-naming test shows that most of the answers with *cʔɜχ* are concentrated in B and D brightness (hues 5–18, 5–16). It means that brightness A is less common for *cʔɜχ* than brightness B, C, or D. The native speakers who were asked to explain the meaning of *cʔɜχ* in Russian (all our participants were bilingual in Russian and Ossetic) usually used 'dark green' or 'dark blue' expressions.

Another candidate for a diffused color is *morɜ* 'brown'. The focal-color test shows that the most frequent answer is 3D, but it has only 4 answers from 12 participants; other circles have from 1 to 2 answers (hues 0, 2–4; brightness C and D). However, the color-naming experiment shows that *morɜ* is a focal color with foci in 2D, 3D, and 3C, 4C. The table below shows the number of answers for *morɜ*.

Table 5. Color-naming test *morɜ*, the foci are marked in bold

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	A																				
1	B			1																	
2	C		3	**9**	**9**	5	1											1			
	D	1	**9**	**11**	5	1														2	2
0																					

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

It is worth noting that some native speakers used *mor3* in hues 17 and 19–20 and also in achromatic colors (0B and 0C). It is never used in brightness A or in hues 7–16 and 18. There are only two uses of *mor3* in brightness B (each use has one answer from two different participants). Together with *c'3χ* the color term *mor3* is a dark color.

Other basic colors, namely, *šarχ* 'red', *bur* 'yellow', *urš* 'white', and *šaw* 'black' are focal. The data for these colors collected in the focal-color and color-naming experiments are identical. Table 6 below summarizes foci of all basic colors except *c'3χ*.

Table 6. Foci. š – *šaw*; u – *urš*; šə – *šarχ*; b – *bur*, m – *mor3*

		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
u	A					b	b															
	B																					
	C		šə		m	m																
š	D			m	m																	

Note that the foci of the Ossetic basic colors are concentrated in hues 0–5 and brightness A, C, and D and do not occur in hues 6–20 or brightness B.

Other than the six basic colors considered in this section, Ossetic also had a seventh basic color term *3χšin* 'dark grey', 'blue-grey'. This color term follows all the criteria of basic colors, however, it is no longer in use in modern Ossetic. There are no examples of *3χšin* either in the Ossetic corpora or in our elicited data. *3χšin* is derived from the Old Iranian *axšaina* 'dark grey', 'blue, grey', which initially meant *a-xšaina-* 'not shining' (Abaev, 1958, p. 220).[7] In Standard (Iron) Ossetic, it was replaced by *c'3χ* 'green, blue, grey'. The root *axšaina* was preserved in Ossetic only within the word *3χšin3g* 'wild dove'. Color terms are often used in names of birds,[8] for example, the Russian *golub'* 'dove' and *goluboj* 'blue'; *sinica* 'tomtit' and *sinij* 'dark blue' or the Persian *kabud* 'blue' and *kabutar* 'dove' (Abaev, 1958, p. 221). Though the color term *3χšin* was lost in Ossetic, its cognate remained in some other modern Iranian languages, e.g., Pashto *šin* 'green, blue, grey', and Persian *χäšin* 'dark, grey, bluish'.

Another archaic color term is *χ3r3* 'grey'. It was completely replaced by *c'3χ*. Ossetic preserved only a few fixed phrases and compounds with *χ3r3*: *χ3r3-miy* (grey cloud.or.mist) 'heavy mist', *χ3riš* 'willow', lit. 'grey twig' (*wiš* 'twig'), *χ3r3g* 'donkey' (Abaev, 1989, p. 175) and, possibly, *χ3r3zašt* 'stony, sinister glance' (lit. 'grey glance').

7. I preserved Abaev's transcription hereinafter in all language examples borrowed from Abaev's Etymology dictionary. The only exception are examples from Ossetic where I used IPA transcription.

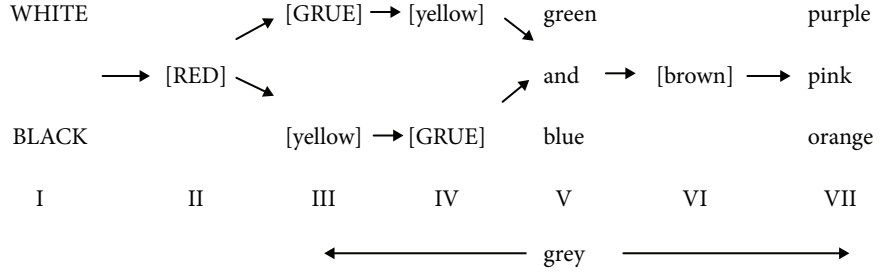
8. There are also a few compounds with a basic color and the word 'bird' in Ossetic, e.g., *šaw-c'iw* (black-bird) 'starling', *bur-c'iw* (yellow-bird) 'oriole'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

4. Ossetic stage of color terms development and the origin of basic color terms

The evolutionary sequence of the development of basic color terms in language was first presented in Berlin and Kay (1969) and has subsequently received certain updates. Schematically, the sequence is shown in Table 7 (after Kay & McDaniel, 1978, p. 615).

Table 7. Evolutionary sequence of the development of basic color terms



The arrow signifies the sequence of development, e.g., the color term RED appeared after the colors WHITE and BLACK, and before other color terms. The term GRUE means the combination of green and blue.

The Ossetic basic color system is located at Stage VI. It has terms for 'black', 'white', 'red', 'yellow', 'green-blue-grey', and 'brown'. Interestingly, Ossetic does not divide 'blue' and 'green' and uses one term for both (*цʔɜχ*). It is believed that the Old Iranian color system was at Stage III ('white', 'black', 'red', and 'green-blue') (Rossi, 2006). The existence of one term for 'green' and 'blue' in Ossetic could be a remnant of the Old Iranian system. It is well known that Ossetic retained some Old and Middle Iranian grammatical features that were completely lost in other modern Iranian languages. The preservation of some peculiarities of extinct Iranian languages in Ossetic (which became atypical of modern Iranian languages), in some cases, can be explained by the influence of the Caucasian languages. Due to historical reasons, Ossetic has been in full isolation from any Iranian language for centuries, as well as from any Indo-European language, if we do not take into account the Russian language which came to the Caucasus only a few centuries ago. Some features common to the Caucasian languages and Alanian (Ossetic's ancestor) are retained in Ossetic. For example, a morphological verbal Optative was typical of the Middle Iranian languages and has been almost completely lost in the modern Iranian languages (the exceptions are Ossetic, some Balochi dialects (Northwestern Iranian) and Northern Talyshi spoken in Azerbaijan (Northwestern Iranian)). The Caucasus is believed to be an Optative language area (Dobrushina, van der Auwera,

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

& Goussev, 2005). The preservation of the verbal Optative in modern Ossetic is explained by the Caucasian influence (Vydrin, 2017).

The Ossetic color term *c'ɜχ* (green.blue.grey) is likely of Caucasian origin (Abaev, 1958, p. 334). The color syncretism of green, blue, and grey is attested in some Northwest Caucasian languages, e.g., Adyghe *šχ$_w$anṭe* 'green, blue, grey'. The color term *morɜ* is also a loanword, which is attested in many languages, both geographically close to Ossetic (Caucasian, Turkic) and more distant (Greek, Czech), e.g., Georgian *mura* 'dark red, dark brown', Kabardian *morä*, Ingush *mōrɜ*, *moažɜ* 'brown' (Abaev, 1973, p. 131). Abaev assumes that Ossetic *morɜ*, which has the same vocalism in the Iron and Digor dialects, should be a recent loanword. Under heavy Caucasian influence, it appears that Ossetic enlarged the Old Iranian color system by adding a 'grey' component to Iranian 'green-blue' syncretism and borrowing a color term for 'brown'.

A cognate to the Ossetic color term *bur* is attested in Scythian personal names and in some modern Iranian languages: Persian *bur* 'brown, chestnut', 'blond-brown', Yaghnobi *vur* (Abaev, 1958, p. 271). Rastorgueva and Edel'man (2003, p. 152) derive it from the Old Iranian **baura*. However, it should be noted that this root denoting a color close to brown is attested in other non-Indo-European languages, e.g., Mongolian *boro* and Turkic *boro*, *bora, boz* 'earth color, grey'. Abaev supposes that it is a Eurasian substrate word (see also Rossi, 2006, pp. 473–474). However, in most of these other languages it means 'brown', while in Ossetic *bur* is brighter and is used for 'yellow'. In some modern Iranian languages, the root *bur* is specific to the chromatic lexicon and is generally translated as 'blond' or 'grey' (see examples in Rossi, 2006, p. 473).

Other Ossetic basic color terms (*šərχ*, *urš*, and *šaw*) have a clear Iranian origin and are attested in other modern and extinct Iranian languages. *šərχ* is derived from Iranian **suxra*, e.g., Avestan *suxra-* 'red'. The root *suxra-* can be traced in many modern Iranian languages, e.g., Persian *sorχ* 'red', Pashto *sūr* 'red' (Abaev, 1979, p. 209). The color term *šaw* goes back to the Iranian **syāva-*, e.g., Persian *siyāh* 'black', Yaghnobi *sāw*, Sogdian *šāw*, etc. (Abaev, 1979, pp. 42–43).

The color term *urš* originates in the Iranian **auruša*, from the root **aruša-* 'flame-colored, bright'. There are some examples from extinct Iranian languages of the words for 'white' derived from this root, e.g., Avestan *auruša-* 'white', Middle Persian *arus* 'white' (Abaev, 1989, pp. 18–19). However, in many Iranian languages **aruša-* is associated with red, e.g., Yazghulami *rust* 'red, foxy', Shughni *rūšt* (feminine *rōšt*) 'red, foxy, ruddy' (Edel'man, 2009, p. 186; Rastorgueva & Edel'man, 2000, pp. 228–229). Interestingly, other modern Iranian languages usually use a color term for white derived from either Iranian **śu̯aita-* / *śu̯ita-* 'light, white' (e.g., Avestan *spaēta*, Persian *sefid* 'white') or **ruk-na* 'shining' (Edel'man, 2009, pp. 185–186). The latter is derived from the root **rauk-* / *rauč-* / *ruk-* 'light, bright',

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

'shine', e.g., Yazghulami *roxn* 'white, bright', Wakhi *ruxn* 'white, bright', 'grey haired' (Edel'man, 2009). Ossetic is a rare (if not the only) modern Iranian language that uses a color term for 'white' which is derived from the Iranian root **aruša-*.

5. Some semantic and morphological peculiarities of basic colors

The primary semantic peculiarity is found in the use of the color term *cʔɜχ* in modern Ossetic. The syncretism of the color term *cʔɜχ* (green.blue.grey) can be useful for certain objects. For example, the color of a body of water may be green, blue, or dark. According to ONC, *cʔɜχ* is frequently used to describe the color of water, e.g., *cʔɜχ dendžəž* 'green-blue-grey sea' or the color of grass and leaves. For more color specific objects the term *cʔɜχ* is ambiguous. Ossetic has some set expressions where *cʔɜχ* is used as either 'grey' or 'blue' or 'green'. For instance, *cʔɜχ sɜštə-tɜ* (green.blue.grey eye-PL.NOM) 'blue eyes' or 'grey eyes' (not 'green eyes'), *ɜnušon-cʔɜχ* (everlasting-green.blue.grey) 'evergreen', e.g., a tree (not everblue or evergrey). Some objects have their own color and the use of the color term is semantically odd. The meaning of *cʔɜχ* in these cases is clear as one knows the color of the object, e.g., *cʔɜχ birɜg* (green.blue.grey wolf) 'grey wolf' (blue or green wolves do not exist in nature).

When *cʔɜχ* is used with an object which can be of any color (e.g., cloth), one can use an additional color term (a non-basic one) to specify the hue of *cʔɜχ*. According to the data from ONC, the most common modifiers for 'grey' are *donχʷəž* (water.color) 'dark grey' and *fɜnəkχʷəž* (ash.color) 'grey'. For the 'blue' hues of *cʔɜχ* Ossetic usually employs the modifier *ɜrvχʷəž* (sky.color) 'blue' or just *ɜrv* (sky). The derivational color term *kɜrdɜgχʷəž* (grass.color) or just the word *kɜrdɜg* (grass) are used to show the 'green' hues of *cʔɜχ*.

The morphological peculiarity of basic color terms is their ability to be nominalized by the suffix *-oj*. In ONC, there are only examples of *uršoj* 'something white', 'the white one' and *šərχoj* 'something red', 'the red one' (see example (1)).

(1) *dəwwɜ fšad-ə kɜrɜzi-jə sɜyd-əns ɜmɜ se χšɜn*
two army-GEN each.other-GEN destroy-PRS.3PL and POSS.3PL between
minɜvar sɜw-ə. urš-oj naj kɜn-ə šərχ-oj
mediator go-PRS.3SG white-SUF milling do-PRS.3SG red-SUF
fɜldaχ-gɜ
overthrow-PART
'Two armies are fighting with each other, there is a mediator in between. The white one is milling, the red one is overthrowing' (a puzzle about teeth and tongue).
(ONC. Makh dug, 1996, no. 12)

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The consulted native Ossetic speaker confirmed the examples and said that the nominalized form of *cʔɜχ* is also possible–*cʔɜχoj* 'something green, blue, or grey'. However, the native speaker expressed doubt that the nominalized forms with the suffix *-oj* from other basic colors exist in Ossetic: **šawoj*, **buroj*, **morɜoj*, **morɜjoj*.

The suffix *-oj* is not active[9] in modern Ossetic (Ahvlediani (red.), 1963, pp. 107, 139). There are a few cases of the derivation of adjectives and nouns by this suffix (sometimes the semantics is caritive): *kʔaχ* 'leg' – *kʔaχoj* 'footless', *song* 'hand'–*songoj* 'handless', *lɜg* 'man'–*lɜgoj* 'manlike', *wiš* 'twig'–*wišoj* 'broom', *žɜng* 'shin'–*žɜngoj* 'gaiter'.

6. Derivational and compound color terms

Aside from its basic color terms, Ossetic possesses a number of derivational and compound colors. In this section, I will consider the most common cases. According to my data, the only active morphological way to derive a new color is by adding the suffix *χʷəž* 'color', 'appearance, look'. Interestingly, Ossetic does not use other derivational suffixes that modify nouns to adjectives, such as the suffix *-on* (*χɜχχ* 'mountain' – *χɜχχon* 'mountainous') or *-džən* (*sɜχ* 'salt' – *sɜχdžin* 'salty', *təχ* 'strength' – *təχdžən* 'strong'). There are no examples of such derivations with color terms in ONC, and the consulted native speaker considered my examples built by this pattern ungrammatical, e.g., **fioleton* intended meaning 'violet', **oranžon* intended meaning 'orange color'.

6.1 *-χʷəž* pattern

The general way to form a derivational color is to use the suffix *-χʷəž* 'color', 'appearance, look' with an object. The suffix *-χʷəž* has been grammaticalized[10] from the word *χʷəž* 'color', 'appearance, look'. Etymologically it is close to Sogdian *kršn* 'appearance', 'look', *kršn'w* 'beautiful', *qršn'wty'* 'beauty' which is in turn connected to the Old Indian *kṛṣna* 'black' (Abaev, 1989, p. 274). The most frequent color terms formed after this pattern are *kɜrdɜgχʷəž* (lit. grass color) 'green, blue' and *ɜrvχʷəž*

9. One of the anonymous reviewers called my attention to Isaev's work on Ossetic (Isaev, 1987, p. 639) where the suffix *-oj* is considered a 'productive' suffix. However, my field data and the data from ONC shows that the suffix is not active in modern Ossetic, i.e., its use is limited to certain lexemes.

10. Technically there are no morphological or syntactic criteria to argue that *-χʷəž* is not the second part of a compound word and purely functions as a suffix in modern Ossetic. I consider it a suffix because of its high productivity and ability to attach to nouns of different semantic classes.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

(lit. sky color) 'blue'. Both color terms are close to the basic color *cʔɜχ*. The semantic differences between *kɜrdɜgχʷəž*, *ɜrvχʷəž,* and *cʔɜχ* will be considered in Section 7.

Color terms with *-χʷəž* were provided by our participants during the color-naming test. I also used data from ONC. At the time of submission, ONC had 1043 contexts of the suffix *-χʷəž* used with nouns.

-χʷəž as a color term derivational suffix can be used with objects of different semantic types. In my data, this pattern is used with the following types of objects: body, *bʷar-χʷəž* (body-color) close to pink and violet (20A, 19A, 17A, 3A, 2A); internal parts of the body, *igɜr-χʷəž* (liver-color) 'dark-red'; body liquids, *tug-χʷəž* (blood-color) 'blood-red'; metals (mostly precious), *ɜvžišt-χʷəž* (silver-color) 'silver', *šəžyɜrin-χʷəž* (gold-color) 'gold'; precious gems, *lal-χʷəž* (true.ruby-color) 'ruby-colored'; articles of domestic utility, *černilɜ-χʷəž* (ink-color) close to violet and blue (19D, 17A, 18C, 16D, 12A); *fɜnək-χʷəž* (ash-color) 'grey', *kopraz-χʷəž* (laundry.blue-color) 'blue'; flowers, *wardi-χʷəž* (rose-color), *rozɜ-χʷəž* (rose-color) 'pink', *siren-χʷəž* (lilac-color) 'lilac'; plants or the material produced from them, *gɜn-χʷəž* (flax-color) 'flaxen' (e.g., hair), *χoraša-χʷəž* (worm.seed-color) 'worm seed color'; food or drink, *məd-χʷəž* (honey-color) 'chestnut brown', *kʔofi-χʷəž* (coffee-color) 'coffee-colored'; trees or their berries, *bal-χʷəž* (cherry-color) 'cherry-colored', *mərtgɜ-χʷəž* (guelder.rose-color) 'guelder rose-colored, rose'; animals (a rare case), *məšt-χʷəž* (mouse-color) 'mousy color' (about horse); birds (a rare case), *bɜlon-χʷəž* (dove-color) 'dove-colored'.

The instances of the occurrence of *-χʷəž* with more abstract objects are also attested in Ossetic: *žɜχχ-χʷəž* (earth-color) close to 'brown' and 'green-blue' (4C, 5C, 10D), *šədžət-χʷəž* (clay-color) 'clay color, grey', *don-χʷəž* (water-color) 'dark grey'. The last type of the *-χʷəž* pattern is non-compositional. The meaning of a color term cannot be guessed from the meaning of the object used since the object does not have a specific color. For example, clay can be white, brown, red, etc. These are cases in which the language chooses a specific color for this type of color terms according to the ethnographic or geographic peculiarities of the nation.

I did not come across a single example in ONC of a color term derived from a noun belonging to the semantic class of fish (cf. English *salmon*). This can be explained geographically. The Ossetians do not have any large bodies of water in their territory and fish has never been a staple of their diet. The color terms derived from the names of animals or birds are rarer and occur less frequently in the corpus than the color terms derived from names of trees, plants, and flowers.

The suffix *-χʷəž* can be attached to colors derived from Russian color loans, e.g., *fiolet* 'violet' (Russian *fioletovyj*) – *fioletχʷəž* 'violet'. It can be used with some basic colors without changing the meaning, (e.g., *morɜ* 'brown' – *morɜχʷəž* 'brown color') or with a basic color marked by a modifier suffix, e.g., *cʔɜχ-bən-χʷəž* (green.blue.grey-DIM-color) 'greenish, bluish, greyish color'. The suffix can be employed with

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

color modifiers, e.g., *ird* 'bright', 'shining' – *irdχ^{w}əž* close to pink, 18A, *lak^{ʔ}on-χ^{w}əž* (turbid-color) 'turbid color'. The color modifiers will be considered in Section 8.2.

The suffix *-χ^{w}əž* can also be attached to abstract nouns to form attributes; the meaning of the suffix, in this case, is 'appearance, look', e.g., *rənčən-χ^{w}əž* (sick-color) '(someone) with an unhealthy appearance'. The derivative sometimes undergoes a metonymical shift: *χɜrž-χ^{w}əž* (good-color) 'fat, well-fed', *χɜd-χ^{w}əž* (self-color) 'indigenous', 'original', *sard-χ^{w}əž* (life-color) 'cheerful' (about a human).

Some of the words derived by the *-χ^{w}əž* pattern have both a color meaning and a non-color meaning. For example, *məd-χ^{w}əž* (honey-color) means either 'chestnut brown' (e.g., hair) or 'honey-like' (*məd-χ^{w}əž χɜlaš* (honey-color voice) 'honeylike voice', compare it to *šədžət-χ^{w}əž* (clay-color) which according to my data from the native speakers consulted has only the color meaning 'clay color, grey' and cannot be used as an attribute 'clayey'.

The *-χ^{w}əž* pattern is an active model of producing adjectives with the meaning 'colored' or 'looking like, similar to'. The exact meaning of the derivative is determined by the context. Compare the examples below where the first example expresses the meaning 'looking like, similar to', while other examples appear to express color:

(2) *šofər-ə faršmɜ bad-ɜn-ɜj wɜžžaw išt š-kod-t-a*
driver-GEN near sit-NMLZ-ABL heavy taken PREF-do-TR-PST.3SG
čumɜdan-χ^{w}əž χɜsɜn-džən χəž-ən
suitcase-color hand-SUF take-INF
'He took (lit. 'heavily lifted') from the seat near the driver the bag which looked like a suitcase with handles'. (ONC)

Other examples of this pattern from the elicited data are: *kivi-χ^{w}əž* (kiwi-color) 'kiwi-colored', *p^{ʔ}amidor-χ^{w}əž* (tomato-color) 'tomato-colored', *č^{ʔ}erami-χ^{w}əž* (apricot-color) 'apricot-colored' (see the example from ONC *wɜžžaw č^{ʔ}erami-χ^{w}əž traktor-tɜ* (heavy apricot-color tractor-PL.NOM) 'heavy apricot-colored tractors' (ONC. Dzabaeva T. T. Siroty i vdovy. Vladikavkaz, 2009)).

The general grammatical difference between the basic colors and the derivatives in *-χ^{w}əž* is the inability of the derivational colors to employ most of the morphological color modifiers used with the basic colors (see Section 8).

The grammaticalization of the meaning 'look, appearance' to the color derivational suffix is typologically trivial. However, it is not clear from the available literature devoted to color terms whether the same pattern is common in other Iranian languages. For example, in Persian (Southwestern Iranian) the pattern with *rang* 'color' as the second part of a color term is possible (*sabz-rang* (green-color) 'green') but not very active. The common way to derive a new color term is to add the suffix *-i* (generally modifies nouns to adjectives) to a noun, e.g., *rang-e piyåz* (color-EZF onion) 'color of onion, red, brown' – *piyåz-i* (onion-SUF) 'onion color, red, brown'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Some Caucasian languages that are geographically close to Ossetic have a color pattern with a word/suffix 'color' which attaches to an object. For example, Adyghe (Northwest Caucasian) uses the word *ŝ$_w$e* to denote 'skin; color' which can derive color terms: *wəcə-ŝ$_w$e* (grass-color) 'grass color'.

The development of the *χ^{w}əž* pattern in Ossetic as a general way to derive new color terms could be either an internal development of the Ossetic color system or a result of the influence of neighboring Caucasian languages. A further study of color derivational patterns in the Caucasian and Iranian languages is needed.

6.2 Compounds

In this section, I will briefly describe Ossetic color term compound patterns, which are found in elicited and corpora data. The list of patterns is far from complete and requires further study.

Color compounds are basically derived by an object as the first component and a basic color as the second component (*ɜrv-c^{ʔ}ɜχ* (sky-green.blue) 'blue', 12B, 9A; *tug-šərχ tərəša* (blood-red flag) 'blood red flag' ONC). The object used in this pattern can function as a color modifier, e.g., *ɜndon-c^{ʔ}ɜχ gagwə-tɜ* (steel-green.blue.grey apple.of.eye-PL.NOM) 'iron grey apples of eyes' (ONC). Rarely, the object can be marked by case, e.g., *tug-ɜj-šərχ səšt* (blood-ABL-red eye) 'blood-colored red eye'. The pattern where the object is marked by the adjectival suffix *–on* is extremely rare; there is only one example of this pattern from ONC: *ɜrv-on-c^{ʔ}ɜχ žɜldag* (sky-SUF-green.blue.grey silk) 'sky-blue silk'. There can be a compound with two objects at the beginning and one basic color term at the end: *ruχšə-tɜ-kɜrdɜg-c^{ʔ}ɜχ* (light-PL-grass-green.blue.grey) 'light green' (ONC).

Another pattern consists of a basic color at the beginning and an object at the end: <basic color + object>, e.g., *šərχ-wadul* (red-cheek) 'red-cheeked, ruddy' (can be used both for animate objects (*red-cheeked girl*) and inanimate objects, e.g., *šərχ-wadul fɜtk^{ʔ}wə* 'ruddy apple', ONC). This pattern is rare in ONC.

The next compound pattern looks like <color modifier + basic color>: *ird-c^{ʔ}ɜχ* (bright-green.blue) 'light blue', 13A, 11B, 12A). Another interesting pattern is the combination of an object and a color modifier: *fɜlurš* 'pale': *mit-fɜlurš* (snow-pale) '(someone) very pale (lit. pale as snow)' ONC.

A combination of two basic color terms (as bare stems or with a suffix) is also possible: *bur-bən-šərχ* (yellow-SUF-red) 'orange' (*burbənšərχ revoluci* 'Orange Revolution', ONC); *šaw-urš* 'grey' (lit. black-white, 0B), *bur-šərχ šɜrəq^{w}ənt*ɜ (yellow-red hair) 'yellow-red hair' (ONC). There can be three basic colors in this pattern: *qulon (bur-c^{ʔ}ɜχ-šərχ aχ^{w}əršt) kɜlmɜržɜn* 'variegated (pained in yellow-green.blue.grey-red) shawl'. This pattern is very rare in the elicited data, however, it occurs

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

with a normal frequency in ONC. There are examples of such patterns used in verbal compounds as a nominal part of the compound: *šɜ sɜštə-tɜ nə-wurš-šərχ əštə* (POSS.3PL eye-PL.NOM PREF-white-red be.PRS.3PL) 'his eyes became white-red'.

In ONC, another pattern is attested: <basic color + object-*χʷəž*+basic color, *bur-kɜrdɜg-χʷəž-cʔɜχ kalm* (yellow-grass-color-green.blue.grey) 'yellow-green snake'. There are examples with two derivative colors in *χʷəž*: *ɜrv-χʷəž-kɜrdɜg-χʷəž-cʔɜχ* (sky-color-grass-color-green.blue.grey) 'blue-green'.

Another rare pattern (cited by only one participant) is <basic color + basic color marked by Comitative + (optionally) *zmɜšt* 'turbidity, darkness'>: *cʔɜχ šaw-imɜ* (green.blue.grey black-COM) 'dark green, dark blue' 8D, 10D, 11D, 13D, 7C, 18D , *morɜ urš-imɜ* (brown white-COM) close to light violet 17A.

7. Synonyms

The most semantically complex Ossetic color terms are *kɜrdɜgχʷəž* (lit. grass.color), *cʔɜχ* and *ɜrvχʷəž* (lit. sky.color). The translation of each of these terms into Russian or English will contain the component 'blue'. The distribution of the terms for a non-native speaker is unclear. In this section, I will discuss the semantic differences between three color terms.

7.1 Distribution of *kɜrdɜgχʷəž* and *cʔɜχ*

Ossetic-Russian dictionaries describe *kɜrdɜgχʷəž* as 'green' and *cʔɜχ* as 'green', 'blue', 'grey' (e.g., Abaev, 1958, pp. 333–334, 582). Our data argue that there are intersections between these two color terms in green and blue hues.

The results of the focal-color test for *kɜrdɜgχʷəž* argue that it is diffused: 8B (3), 8C (3), 7C (2), 9C (1), 10C (1), 10D (1), 11C (1). According to our color-naming data, 10 participants chose *kɜrdɜgχʷəž* to describe 7C, 8C, 9C; 9 participants used *kɜrdɜgχʷəž* for 7B and 8 participants selected this color term for 8B. Other chips received 6 or fewer *kɜrdɜgχʷəž* answers. See Table 8 below.

Table 8. *kɜrdɜgχʷəž*. Color-naming test, the foci are marked in bold

		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	A						4		5	5	3	3										
1	B					3			**9**	**8**	6	3	6			1						
	C					1			**10**	**10**	**10**	6		2								
	D								2	2	5	3	2									

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The foci of *k3rd3gχʷəž* are brightness C (and partly B) and hues 7–9. It can be used in both brightness A and D, and its total distribution spreads from hue 5 ('yellow hue') to hue 12 ('blue hue'). We also have one answer for hue 14 and one achromatic use (0B). The correct translation of *k3rd3gχʷəž* into English is 'green', 'blue', or 'yellow'.

Let us compare *k3rd3gχʷəž* and *cʔ3χ*. Table 9 shows the distribution of *cʔ3χ*.

Table 9. Distribution of *cʔ3χ*

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	A						2	5	8	7	8	5	3	2	3	1					
5	B					2	7	7	8	11	12	9	8	5	6	5	1				
5	C					4	7	3	3	9	10	12	10	11	8	8	2	1	1		
	D					1		1	4		6	9	10	9	7	9	4	1	1		
0																					

It is clear that both color terms cannot function in hues 1–4 and 19–20. Both terms can be used in all types of brightness, however, brightness D is less common for *k3rd3gχʷəž* (2–5 answers) than for *cʔ3χ* (maximum of 10 answers). The distribution of *k3rd3gχʷəž* in hue is less than that of *cʔ3χ*. *k3rd3gχʷəž* is rarely employed in hues 13–20 and with achromatic colors. *k3rd3gχʷəž* has foci, while *cʔ3χ* is diffused.

Comparing Tables 8 and 9, it is clear that *cʔ3χ* covers all the uses of *k3rd3gχʷəž*. However, taking into account the number of answers for *k3rd3gχʷəž* and *cʔ3χ*, it is clear that the main intersections of *k3rd3gχʷəž* and *cʔ3χ* are 9C, 7–8B, 7A and, possibly, 5B and 11B.

7.2 Distribution of *3rvχʷəž, k3rd3gχʷəž* and *cʔ3χ*

Another complex case is the color term *3rvχʷəž* (lit. 'sky color') and its interaction with *k3rd3gχʷəž* (grass color) and *cʔ3χ* (the basic color used for green, blue and grey). Dictionaries translate *3rvχʷəž* as 'blue', 'azure', 'sky-blue'. Our data elicited from the color-naming test is demonstrated in Table 10.

Table 10. *3rvχʷəž*, color-naming test

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	A									1		6	11	7	10	9	1				
	B												9	11	10	10					
	C											1	1	4	7	7					
	D								2			1		1	1	4					
0																					

I did not conduct a focal-color test with *3rvχʷəž*. However, color-naming data shows that the foci of *3rvχʷəž* should be hues 12–15 and brightness A-B (9–11 answers).

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

ɜrvχʷəž can be used in brightness C. The fewest number of answers were received for brightness D. The hue distribution is small, mainly, from 11 to 15 ('blue hue'); we got 2 answers for hue 8 ('green hue'), one each for hues 9 and 16.

Comparing the distribution of *ɜrvχʷəž*, *kɜrdɜgχʷəž* and *cʔɜχ* one notices that *ɜrvχʷəž* virtually does not intersect with *kɜrdɜgχʷəž*, which occupies hues 7–10. It is also brighter than *kɜrdɜgχʷəž*. *ɜrvχʷəž* does intersect with *cʔɜχ* in hues 12–15. However, *cʔɜχ* is darker than *ɜrvχʷəž*. In the aforementioned hues, *cʔɜχ* is mainly employed in brightness C and D, while brightness B has fewer answers (5–8) and brightness A lies on the periphery of the *cʔɜχ* basic use (1–3 answers). There are interesting cases in 11A and 14–15C, all of which can be served equally by either *cʔɜχ* or by *ɜrvχʷəž*.

To summarize, *kɜrdɜgχʷəž* is mainly employed in hues 7–9 and brightness B and C, *ɜrvχʷəž* occupies hues 12–15 and mainly brightness A and B, while *cʔɜχ* is a diffused color with very broad application, which sometimes intersects with *kɜrdɜgχʷəž* and *ɜrvχʷəž*.

8. Color modifiers

8.1 Morphological markers

Ossetic possesses at least five different affixes, which are mainly used to mark different degrees of saturation and brightness. There are two markers for the intensive degree and three diminutive suffixes. Most of the color affixes can only be used with basic colors. I list the affixes and consider their semantics and use below.

Intensive affixes:

a. The prefix[11] *šɜnt-* or *χɜnt-* is used with color terms to convey high saturation of the color or 'absolute' color without any admixture, e.g., *šɜnt-šərχ* 'very red', *šɜnt-urš* 'absolutely white'. According to our corpora data, it can only be used with the basic colors except *morɜ* 'brown'. However, the Ossetic native speaker assessed *šɜnt-morɜ* 'very brown' as natural. The prefix is not a dedicated color marker; it is also employed with other adjectives, e.g., *šɜnt-χuš kʔaliw* 'absolutely dry limb'.

11. The morphological state of *šɜnt-/xɜnt-* is unclear. Standard grammars call *šɜnt-* a prefix, however, its behavior is the same as that of other attributes and, possibly, *šɜnt-/xɜnt-* can be analyzed as lexical color modifiers. In the Ossetic orthography *šɜnt-/xɜnt-* and its head are written either as a solid word or (in rarer cases) as two words.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

b. The suffix *-id* is attached to the second component of the reduplicated basic color term. It expresses the same meaning as the prefix *šɜnt-*, e.g., *šərχ-šərχ-id* 'very red', *šaw-šaw-id* 'very/absolutely black'. It cannot be used with any kind of compound colors. The only basic color term which is not used in this pattern in *morɜ* 'brown'. The suffix *-id* cannot be used with any other attributes aside from color.

Diminutive suffixes:

a. The suffix *-bən* can only be used with basic colors: *urš-bən* 'whitish', 'grey'. It can also be employed with color modifiers: *tar-bən* (dark-SUF) 'darkish'. The color modifiers marked by the suffix can combine with basic colors: *tar-bən-cʔɜχ* (dark-SUF-green.blue.grey) 'dark blue', 'dark grey'. The suffix can be used with other attributes (conveying 'the weaker quality') and nouns (forming an attribute): *wəmɜl* 'wet', 'damp' – *wəmɜl-bən* 'dampish'; *χʷər-bən* (gravel-SUF) 'stony'.
b. The suffix *-gomaw* is the only morphological color modifier that is used with both the basic colors and with the compounds in *-χʷəž*: *šaw-gomaw* 'blackish', *ɜrvχʷəž-gomaw* 'bluish'. This suffix can be employed with other attributes (the suffix diminishes the quality), e.g., *štər-gomaw* (big-DIM) 'biggish', *lɜmɜy-gomaw* (weak-DIM) 'weakish'. An interesting point is that according to the standard grammar (Ahvlediani (red.), 1963, p. 140) *-bən* and *-gomaw* can be used together: *bur-bən-gomaw* 'yellowish'. However, native speakers consider *bur-bən-gomaw* ungrammatical. In ONC, there is only one example of the simultaneous use of the suffixes *-bən* and *-gomaw*: *wɜž-bən-gomaw* (heavy-SUF-SUF) 'a bit lazy'.
c. The suffix *-gond* is used exclusively with the basic color terms: *šərχ-gond* 'reddish' (lit. 'close to red'). The suffix *-gond* is derived from the past stem of the verb *kɜnən* 'to do'. It can be used with nouns to mark the nouns of action (*χʷəm-gɜnd* field-do.PST 'ploughed field') or the similarity with the object (*χəž-gond* net-do.PST 'lace' (something similar to a net).

The distributional and semantic differences of the diminutive suffixes are not clear. Native speakers consider the examples of the basic color terms used with different diminutive suffixes semantically identical (e.g., *uršbən*, *uršgomaw,* and *uršgond* mean 'whitish'). The only distributional difference found consists of the ability of the suffix *-gomaw* to mark color terms in *-χʷəž*. The other two suffixes function only with the basic color terms.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

8.2 Lexical modifiers

There are a significant number of lexical modifiers of color in Ossetic. I will list some modifiers which specify the brightness and saturation of a given color: *təng* 'very', *šat3g* 'pure' (can be used as an intensifier: *šat3g-šaw* 'very black'), *tar* 'dark', *f3lurš* 'pale', *ird* 'bright', *f3lm3n* 'soft', *χʷəžzəd* 'faded', *b3žžən* 'thick', *žm3št* 'turbid', *qulon* 'variegated', etc. Lexical modifiers are always preposed to the color term.

No formal restrictions for lexical color modifiers were noticed in Ossetic. They can be used with basic colors, derivational color terms, and compounds, for example:

tar k3rd3g-χʷəž (dark grass-color) 'dark green' B10,
tar cʔ3χ (dark green,blue,grey) 'dark blue' D15,
tar mor3 (dark brown) 'dark brown' D3,
tar 3rv-χʷəž (dark sky-color) 'dark blue' D14,
tar siren-χʷəž (dark lilac-color) 'dark purple' D17,
tar-3rv-cʔ3χ (dark-sky-green.blue.grey) 'dark blue'.

A lexical modifier can also be marked by a morphological modifier, e.g., *tar-bən-šərχ* (dark-SUF-red) 'a bit bright red'.

There are also lexemes which are used to specify hue in Ossetic: *3mχ3cc3* 'mixed', *3rχ3cc3* 'mixed' (pattern <first color term-second color term-COMITATIVE-*3mχ3cc3*>), *χ3cc3* 'mixture' (pattern <*χ3cc3*-first color term-*3m3* 'and'-second color term>), *3vvaχš* 'close to', *χ3št3g* 'close to' (pattern <color term-ALLATIVE *3vvaχš*/*χ3št3g*>). Compare the examples:

urš cʔ3χ-im3 3mχ3cc3 (white green.blue.grey-COM mixed) 'blue' A11,
urš šərχ 3rχ3cc3 (white red mixed) 'pink' A19, from complex predicate *3mχ3cc3 k3nən* (mixture do) 'to mix',[12]
χ3cc3 šərχ 3m3 urš (mixture red and white) 'pink' B1,
k3rd3g-χʷəž-im3 3vvaχš (grass-color-ALL close) 'green-blue' C10,
šərχ-m3 3vvaχš (red-ALL close) 'red, pink' C19,
k3rd3g-χʷəž-im3 χ3št3g (grass-color-ALL close) 'green-yellow' B5.

12. Note that in the cited examples *urš* functions as a brightness modifier, *urš cʔ3χ* and *urš šərχ* are *cʔ3χ* and *šərχ* with brightness A. In this function *urš* is interchangeable with *f3lurš* 'pale'.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

9. Connotations of the basic colors

The connotations and non-color meanings of the Ossetic basic color terms need a special study. In this chapter, I will mention only some cases of the use of basic colors in non-color meanings.

šaw 'black' has mostly negative connotations in Ossetic, e.g., *šaw bon* (black day) 'bad day, disaster, misfortune'. It is widely used in curses. However, there are some composites where *šaw* has positive connotation: *šaw-čəžg* (black-girl) 'dark-complexioned[13] girl', *šaw-rзšuγd* (black-beautiful) 'beauty', *šaw lзppu* (black boy) 'good boy' (Besolova, Balova & Berbekova, 2009, pp. 158, 163). As in other Indo-European languages, *urš* 'white' mostly has a positive connotation and is opposed to *šaw* 'black'. This opposition becomes clear in Ossetic proverbs, as in:

(3) *šaw χoχ-зj urš dur nз χaw-ə*
black mountain-ABL white stone NEG fall-PRS.3SG
'A white stone does not fall from a black mountain' (meaning: something good does not come from something bad).
(Besolova, Balova, & Berbekova, 2009, p. 161)

The basic color *šərχ* 'red' in Ossetic idioms and fixed phrases has the meanings 'beautiful' and 'warm' (Besolova, Balova, & Berbekova, 2009, pp. 165–166). For example:

šəžγзrin šərχ Alardə (gold red Alardy) 'gold beautiful Alardy (a deity of smallpox)'
šərχ-žəng (red-fire) 'aglow'.

The color term *cʔзχ* can mean 'not yet ready, green, unripe', for example, *cʔзχ nзməg* 'green grain'. It can have negative connotations, e.g., *mзrdon-cʔзχ* 'dead color', 'dead grey, blue', *žзrdз-cʔзχ* (heart-green.blue.grey) 'fatigue', the complex verb *acʔзχ wзvən* (PREF.green.blue.grey be) 'become green, blue' (can also describe humans with the meanings 'pale', 'feel bad', or 'feel cold').

The color term *cʔзχ* can have positive connotations. For example, it is associated with the color of the sky early in the morning and with the healthy color of greenery. Compare the example below, which describes the morning sky:

(4) *šзwwon arv-ə nog-зχšad cʔзχ tзbзγ kзšadž-ə*
morning.ADJ sky-GEN new-washed green.blue.grey plate fish-GEN
sзšt-aw š-зrttəvt-a
eye-EQU PREF-shine-TR.PST.3SG
'The just washed blue-grey plate of the morning sky was shining like a fish's eye'.
(ONC. Makh dug, 1998, no. 9)

13. Considered one of the attributes of beauty in the Caucasus.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The color term *bur* also means 'browned, crisp, roasted'. It has connotations of prosperity, harmony, something good: *Burɜfɜrnəg* – a proper name of the forefather of the Borata family, consists of *bur* 'yellow' and *farn* 'prosperity, harmony, blessing'; *Bur-ɜrtχuron* 'the god of fire' (yellow-fire.sun.SUF). *bur* can also mean illness: *jɜ sɜšgom nə-bbur-iš* (POSS.3SG face PREF-yellow-PRS.3SG) 'his face turned yellow' / 'he turned sick'.

As noted above, the basic color *morɜ* is supposed to be a late loanword (Abaev, 1973, pp. 130–131). Like all other basic color terms, *morɜ* can have color morphological and lexical modifiers (*morɜ-bən* 'brownish' *tar-morɜ* (dark-brown) 'dark brown', *fɜlurš-morɜ* (light-brown) 'light brown'). The frequency of basic colors used separately without any morphological markers in ONC shows the followings results: *šaw*–11088, *urš*–4855, *cʔɜχ*–4226, *šərχ*–2638, *bur*–1125, *morɜ*–136. The highest frequency of *šaw* 'black' is explained by its wide use in idioms and fixed phrases. The least frequently cited basic color term is *morɜ*; it has only 136 uses in ONC. It can be compared to non-basic color terms *kɜrdɜgχʷəž* 'green' and *ɜrvχʷəž* 'blue', which have 125 and 189 uses respectively.

There are no non-color uses of *morɜ* in the corpus. It is employed mostly with cloth (cloth itself and things made from it, e.g., dress, footwear). There are some examples of *morɜ* used to express the color of eyes, hair, leaves and earth. The low frequency in ONC and the absence of non-color uses of *morɜ* suggest that it only recently became a basic color.

Some basic colors are used in compounds with generic nouns to denote a species (the pattern <basic color + generic noun>. For example, *cʔiw* 'bird', *šaw-cʔiw* (black-bird) 'starling', *bur-cʔiw* (yellow-bird) 'oriole', *nɜməg* 'grain', *šaw-nɜməg* (black-grain) 'blueberry', *šərχ nɜməg* (red grain) 'cranberry'. The pattern is neither productive nor predictable, e.g., for 'cowberry' Ossetic uses a loanword *mɜckʔwə*. While berries and birds are used in this model, mushrooms and fruits are not (cf. Russian *belyj grib* (lit. white mushroom) 'edible boletus', *krasnyj grib* (red mushroom) 'orange-cup boletus').

10. Special color terms in Ossetic

There are special color terms for domestic animals in some other modern Iranian languages. For example, in Luri (Southwestern Iranian), colors of domestic animals are object specific (at least for horse, donkey, sheep, goat, and cow) (Friedl, 1979, p. 53).

I studied the usual way of naming color for the following domestic animals: chicken, cow, sheep, goat, donkey, and horse. The data from Ossetic native speakers show that the Ossetians usually use basic color terms to specify the color of domestic animals, see Table 11. For grey, *cʔɜχ* can be replaced by *fɜnəkχʷəž*.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 11. Color of domestic animals in Ossetic

animal	black	white	brown	grey
horse	*šaw*	*urš*	*morз*	*cʔзχ*
donkey	*šaw*	*urš*		*cʔзχ*
sheep	*šaw*	*urš*	*morз*	*cʔзχ*
goat	*šaw*	*urš*	*morз*	*cʔзχ*
cow	*šaw*	*urš*	*morз*	*cʔзχ*
chicken	*šaw*	*urš*	*morз*	*cʔзχ*

There are also a few specific color terms for horse in ONC, formed by the *-χʷəž* pattern: *bзgзnə-χʷəž bзχ* (beer-color horse) 'bay horse', *ruvaš-χʷəž bзx* (fox-color horse) 'chestnut horse'. The color-naming of domestic animals is not lexicalized in Ossetic.

In general, Ossetic has very few special color terms, the use of which is limited to some objects. There are only two examples in my data from ONC: *šзyzašt* 'brown' (used only for eyes) from *šзy* 'goat' and *sзšt* 'eye', *məlažon* 'dark, black' (usually about clouds). Abaev (1973, p. 139) considers *məlažon* a loanword from Greek, which might have been borrowed during the Greek-Scythian language contacts in what is now Southern Russia and Ukraine.

11. Conclusions

I will here summarize the peculiarities of the Ossetic color system. Ossetic has six basic color terms referring to white, black, red, yellow, brown, and one word for green, blue, and grey. The Ossetic color system is located at Stage VI in the evolutionary sequence of basic color terms development. According to the diachronic data, Ossetic extended the Old Iranian three-colored terms system by developing a color term for yellow and by borrowing a special color term for green, blue, and grey, and another color term for brown from the neighboring non-Indo-European languages. The data presented in this chapter argues that the Ossetic color term for brown was borrowed recently: it does not have any non-colored meanings and its frequency in the corpus is very low (compared to the frequency of other basic color terms).

The only diffused basic color term is *cʔзχ*, which covers the spectrum from hue 5 to hue 18 and brightness from A to D. It is also the most frequent color term for the achromatic chips B0 and C0. I have shown how the Ossetians distinguish green, blue, and grey hues using special lexical color modifiers together with the syncretic color term *cʔзχ*.

The most active derivational model is the use of the suffix *-χʷəž* 'color, look, appearance', which can derive a color from practically any object that has a specific

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

color. A further comparative study is needed to determine whether this pattern is typical of other Iranian languages or not. My preliminary study shows that the pattern is common in at least some Caucasian languages that are geographically close to Ossetic, while in some of the modern Iranian languages this pattern is not active.

Another peculiarity of the Ossetic color system is the developed system of morphological intensifiers and diminutives. Ossetic has two intensive and three diminutive affixes. Most of the affixes are used only with the basic color terms. Such a developed system seems uncommon in other Iranian languages. However, the Ossetic morphology (9 cases, 5 moods) is highly developed in general and atypical of other modern Iranian languages.

Another feature that distinguishes the Ossetic color system from other modern Iranian color systems is the rarity of the reduplication pattern in Ossetic. This pattern seems to be common in at least some Iranian (e.g., Persian) and Indo-Aryan languages (e.g., Hindi, see Drocco & Risato, this volume). However, this feature also needs a special comparative study.

Ossetic basic color terms are frequently used in fixed phrases, curses, and blessings, and other phraseological expressions. At the same time, Ossetic has not developed special color terms for specific objects, e.g., domestic animals. There are very few color terms that are used exclusively with specific types of objects in Ossetic.

The further study of the Ossetic color system requires its comparison to color terms in the Caucasian languages that are geographically close to Ossetic and in the modern Iranian languages that are genetically close to Ossetic.

Abbreviations

ABL	ablative
ADJ	adjective
ALL	allative
COM	comitative
COMPAR	comparative
DIM	diminutive
EQU	equative
EZF	ezafe
INF	infinitive
GEN	genitive
NEG	negation
NMLZ	nominalization
NOM	nominative
ONC	Ossetic National Corpus
PART	particle
PL	plural
POSS	possessive
PREF	prefix
PRS	present
PST	past
SG	singular
SUF	suffix
TR	transitive

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

References

Abaev, V. I. (1958). *Istoriko-etimologičeskij slovar' osetinskogo jazyka. T. 1*. Moscow and Leningrad: Izdatel'stvo akademii nauk SSSR [Etymology dictionary of Ossetic language. Volume 1].

Abaev, V. I. (1973). *Istoriko-etimologičeskij slovar' osetinskogo jazyka. T. 2*. Leningrad: Izdatel'stvo 'Nauka' [Etymology dictionary of Ossetic language. Volume 2].

Abaev, V. I. (1989). *Istoriko-etimologičeskij slovar' osetinskogo jazyka. T. 4*. Leningrad: Izdatel'stvo 'Nauka1' [Etymology dictionary of Ossetic language. Volume 4].

Ahvlediani, G. S. (Ed.). (1963). *Grammatika osetinskogo jazyka. T. I. Fonetika i morfologija*. Ordžonikidze: Naučno-issledovatel'skij institut pri sovete ministrov Severo-Osetinskoj ASSR [Ossetic grammar. Volume 1].

Berlin, B., & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.

Besolova, E. B., Balova, I. M., & Berbekova, Kh. M. (2009). Cvetooboznačenija v tradicionnoj kul'ture adygov i osetin [Color naming in the traditional culture of Adyghe and Ossetians]. *Izvestija SOIGSI. Vyp.* 3 (42), 157–170.

Čunakova, O. M. (1994). Cvetovaja simvolika v pexlevijskix tekstax [Color symbolism in Pahlavi texts], *Peterburgskoe vostokovenie 6*, 194–201.

Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories of a stone-age tribe. *Nature*, 398, 203–204. https://doi.org/10.1038/18335

Dobrushina, N., van der Auwera, J., & Goussev, V. (2005). Optative. In M. Dryer, M. Haspelmath, D. Gil, & B. Comrie (Eds.), *World Atlas of Language Structures* (pp. 298–301). Oxford: Oxford University Press.

Dzahova, V. T. (2009). *Foneticheskie harakteristiki fonologicheskoj sistemy sovremennogo osetinskogo (ironskogo) literaturnogo jazyka. Monografija*. Vladikavkaz: Izdatelstvo SOGPI. [Phonetic description of phonological system of modern standard (Iron) Ossetic].

Edel'man, D. I. (2009). *Sravnitel'naja grammatika vostočnoiranskih jazykov. Leksika.* [Comparative grammar of Eastern Iranian languages. Vocabulary]. Moscow: Vostočnaja literatura.

Friedl, E. (1979). Colors and culture change in Southwestern Iran. *Language in Society*, 8, 51–68. https://doi.org/10.1017/S0047404500005947

Isaev M. I. (1987). Osetinskij jazyk. In V. I. Abaev, M. N. Bogoljubov, V. S. Rastorgueva (Eds.), *Osnovy iranskogo jazykoznaniya. Novoiranskie jazyki: vostočnoiranskie jazyki* (pp. 537–643). Moscow: Nauka [Ossetic language].

Kay, P., Berlin, B., Maffi, L., Merrifield, W. R., & Cook, R. (2009). *The World Color Survey*. Stanford: CSLI.

Kay, P., & McDaniel, C. K. (1978). The linguistic significance of the meaning of basic color terms. *Language*, 54, 610–646. https://doi.org/10.1353/lan.1978.0035

Majid, A., Jordan, F., & Dunn, M. (2011). *Evolution of Semantic Systems Procedures Manual*. Nijmegen: Max Planck Institute for Psycholinguistics.

Majid, A., & Levinson, S. C. (2007). The language of vision I: Color. In A. Majid (Ed.), *Field Manual* Volume 10 (pp. 32–35). Nijmegen: Max Planck Institute for Psycholinguistics.

Majid, A. (2008). Focal colours. In A. Majid (Ed.), *Field Manual* Volume 11 (pp. 8–10). Nijmegen: Max Planck Institute for Psycholinguistics.

Rastorgueva, V. S., & Edel'man, D. I. (2000). *Etimologičeskij slovar' iranskih jazykov. Tom 1*. Moscow: Vostočnaja literatura [Etymology dictionary of Iranian languages. Volume 1].

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rastorgueva, V. S., & Edel'man, D. I. (2003). *Etimologičeskij slovar' iranskih jazykov. Tom 2*. Moscow: Vostočnaja literatura [Etymology dictionary of Iranian languages. Volume 2].

Regier, T., Kay, P., & Cook, R. S. (2005). Focal colours are universal after all. *PNAS*, 102, 8386–8391. https://doi.org/10.1073/pnas.0503281102

Roberson, D., Davies, I., & Davidoff, J. (2000) Colour categories are not universal: Replications and new evidence from a Stone-age culture. *Journal of Experimental Psychology: General*, 129, 369–398. https://doi.org/10.1037/0096-3445.129.3.369

Rossi, A. (2006). Colours and lexical taxonomies: Linguistic and cultural categories in Iranian. In A. Panaino & R. Zipoli (Eds.), *Proceedings of the 5th Conference of the Societas Iranologica Europæa*. Vol. II: *Classical and Contemporary Iranian Studies* (pp. 459–480). Milano.

Rossi, A. (2007). More colours in Khotanese. In M. Macuch, M. Maggi, & W. Sundermann (Eds.), *Iranian Languages and Texts from Iran and Turan: Ronald E. Emmerick Memorial Volume* (pp. 343–354). Wiesbaden.

Vydrin A. (2017). Ossetic Optative: Iranian heritage vs. Caucasian influence. G. Authier, H. Gérardin, M. I. Magomedov, & T. A. Maisak (Eds.), *Historical Linguistics of the Caucasus: Book of abstracts, Paris*, 12–14 April, 2017 (pp. 201–206). Makhachkala.

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

Index

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use

The volume presents sixteen chapters focused on lexicalization patterns used in color naming in a variety of languages. Although previous studies have dealt with categorization and perceptual salience of color terms, few studies have been consistently conducted in order to investigate phonological, morphological, syntactic, and semantic devices languages use to form color terms.

The aim of this volume is to approach color data from a relativist and typological perspective and to address some novel viewpoints in the research of color terms, such as:
(a) the focus on language structure per se in the study of lexicalization data; (b) investigation of inter- and intra-language structural variation; (c) culture and language contact as reflected in language structure.

Topics of this book have a broad appeal to researchers working in the fields of linguistics, anthropology, sociology, and psychology.

ISBN 978 90 272 0403 5

John Benjamins Publishing Company

EBSCOhost - printed on 2/10/2023 1:40 AM via . All use subject to https://www.ebsco.com/terms-of-use