



**HAL**  
open science

# Lexicalization patterns in color naming in Croatian, Czech, and Polish

Ida Raffaelli, Jan Chromý, Anetta Kopecka

► **To cite this version:**

Ida Raffaelli, Jan Chromý, Anetta Kopecka. Lexicalization patterns in color naming in Croatian, Czech, and Polish. Lexicalization patterns in color naming, 78, John Benjamins Publishing Company, pp.269-285, 2019, Studies in Functional and Structural Linguistics, 10.1075/sfsl.78.12raf . hal-04653659

**HAL Id: hal-04653659**

**<https://hal.science/hal-04653659v1>**

Submitted on 22 Sep 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **Lexicalization patterns in color naming in Croatian, Czech, and Polish**

**Ida Raffaelli**  
*University of Zagreb*

**Jan Chromý**  
*Charles University, Prague*

**Anetta Kopecka**  
*CNRS & Université Lumière Lyon 2*

**Running head: Lexicalization patterns in Croatian, Czech, Polish**

The goal of this chapter is to 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 project Evolution of Semantic Systems (Max Planck Institute for Psycholinguistics, Nijmegen). Croatian, Czech, and Polish were among more than 50 Indo-European languages included in the project. The sample consisted of 20 native speakers of Croatian, 20 native speakers of Czech and 21 native speakers of Polish. 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.

*Key words:* color terms, lexicalization patterns, word-formation, Croatian, Polish, Czech

## **1. Introduction<sup>1</sup>**

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 study (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 & Greville, 1993; Lucy, 1997) and further developed since its publication. As Majid, Jordan, and Dunn (2015) 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.<sup>2</sup> 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,

---

<sup>1</sup> 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.

<sup>2</sup> Majid, Jordan, & Dunn (2015).

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, 1995; 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 similarities and differences between lexicalization patterns<sup>3</sup> 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

---

<sup>3</sup> 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.

with a color elicitation task that consisted of 84 Munsell color chips (materials developed by Majid & Levinson, 2007).<sup>4</sup> 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

---

<sup>4</sup> Details on the methods employed are presented in the third section.

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 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.<sup>5</sup>

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

---

<sup>5</sup> Langacker (1987, pp. 56–69) elaborates the genesis of linguistic units as conventionalized expressions.



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 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.<sup>6</sup>

The term *lexicalization pattern*, as presented in this chapter, is akin to Talmy's eponymous concept.<sup>7</sup> 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

---

<sup>6</sup> The model is presented in more details also in Raffaelli (2013; 2017) and Raffaelli & Kerovec (2017).

<sup>7</sup> 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.

processes, either morphological or syntactic, regarding them as equally important and enabling lexicalization of meanings.<sup>8</sup>

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.<sup>9</sup> 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

---

<sup>8</sup> See also Raffaelli & Kerovec (2017).

<sup>9</sup> In morphologically rich languages word-formation processes are integrated in lexicalization of new lexemes. Therefore, word-formation and lexicalization are regarded as complementary and unseparable.

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)<sup>10</sup> 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 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,

---

<sup>10</sup> 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.

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.

**Table 1.** Lexicalization patterns used by Croatian, Czech, and Polish speakers in color naming

	CROATIAN		CZECH		POLISH	
	Type	Token	Type	Token	Type	Token
<i>Adj (simple)</i>	19 (11.7%)	641 (38.2%)	10 (3.8%)	504 (30%)	15 (7.7%)	668 (37.9%)
<i>Adj (derived)</i>	9 (5.5%)	331 (19.7%)	27 (10.3%)	399 (23.8%)	34 (17.4%)	585 (33.2%)
<i>N</i>	–	–	2 (0.8%)	4 (0.2%)	19 (9.7%)	107 (6.1%)
<i>Adj [o] Adj</i>	43	122 (7.3%)	55	161	25	128 (7.3%)

	(26.4%)		(20.9%)	(9.6%)	(12.8%)	
<i>Adv + Adj</i>	50 (30.7%)	507 (30.2)	67 (25.5%)	483 (28.8%)	–	–
<i>Adj + Adj</i>	–	–	52 (19.8%)	68 (4.1%)	39 (20%)	144 (8.2%)
<i>Adj + N</i>	–	–	2 (0.8%)	2 (0.1%)	27 (13.9%)	96 (5.4%)
<i>Adv + Adj [o]</i> <i>Adj</i>	1 (0.6%)	1 (0.06%)	11 (4.2%)	19 (1.1%)	2 (1%)	2 (0.1%)
<i>N + Ngen</i>	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%)

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.

#### **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’,<sup>11</sup> 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 *běžová* ‘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).<sup>12</sup> 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.

---

<sup>11</sup> *Ljubičasta* ‘purple’ or ‘violet-like’ is a derived color term formed from the stem *ljubič* ‘violet’ and the derivational morpheme *-asta* ‘like’.

<sup>12</sup> The non-derived term *bordo* also exists in Polish, but it was used only once in the EoSS task.

#### 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.

**Table 2.** The ten most frequent derived color terms in the three languages

<b>Croatian</b>	<b>N</b>	<b>Czech</b>	<b>N</b>	<b>Polish</b>	<b>N</b>
<i>ljubičasta</i> 'purple' (violet-like)	208	<i>fialová</i> 'purple' (violet-like)	148	<i>fioletowy</i> 'purple' (violet-like)	157
<i>narančasta</i> 'orange-like'	68	<i>růžová</i> 'pink' (rose-like)	103	<i>różowy</i> 'pink' (rose-like)	125
<i>ružičasta</i> 'pink' (rose-like)	33	<i>oranžová</i> 'orange'	59	<i>brązowy</i> 'brown'	83



<i>maslinasta</i> ‘olive-like’	11	<i>okrová</i> ‘ocre’	13	<i>pomarańczowy</i> ‘orange’	56
<i>plavkasta</i> ‘blueish’	3	<i>lososová</i> ‘salmon-like’	12	<i>granatowy</i> ‘dark blue’	47
<i>zelenkasta</i> ‘greenish’	3	<i>tyrkysová</i> ‘turquoise’	11	<i>bordowy</i> ‘bordeaux’	16
<i>breskvasta</i> ‘peach-like’	2	<i>běžová</i> ‘beige’	7	<i>seledynowy</i> ‘aquamarine’	14
<i>rozikasta</i> ‘pinkish’	2	<i>vínová</i> ‘wine-like’	7	<i>purpurowy</i> ‘purple’	13
<i>žučkasta</i> ‘yellowish’	1	<i>purporová</i> ‘magenta’	5	<i>turkusowy</i> ‘turquoise’	13
–	–	<i>cihlová</i> ‘brick-like’	4	<i>oliwkowy</i> ‘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.<sup>13</sup> The second lexicalization pattern, [stem + *-kast(a)*], conveys the meaning of ‘approximation’ as in *žučkasta* ‘yellowish’ or *zelenkasta* ‘greenish’. These are the only four color terms in the Croatian

<sup>13</sup> 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.

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 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 *-owy* 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 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

<b>Croatian</b>	<b>N</b>	<b>Czech</b>	<b>N</b>	<b>Polish</b>	<b>N</b>
<i>maslinasto-zelena</i> 'olive-like-green'	16	<i>modrozelená</i> 'blue-green'	19	<i>ciemnozielony</i> 'dark-green'	44
<i>žuto-zelena</i>	11	<i>zelenomodrá</i>	15	<i>jasnozielony</i>	14

'yellow-green'		'green-blue'		'light-green'	
<i>plavo-zelena</i>	9	<i>žlutozelená</i>	11	<i>jasnoniebieski</i>	13
'blue-green'		'yellow-green'		'light-blue'	
<i>zeleno-plava</i> 'green-blue'	9	<i>tmavozelená</i> 'dark green'	9	<i>jasnobrązowy</i> 'light-brown'	7
<i>zeleno-smeđa</i>	7	<i>zelenohnědá</i>	9	<i>ciemnożółty</i>	6
'green-brown'		'green-brown'		'dark-brown'	
<i>crveno-narančasta</i> 'red-orange'	6	<i>zelenožlutá</i>	9	<i>jasnofioletowy</i>	6
		'green-yellow'		'light-purple'	
<i>sivo-zelena</i>	5	<i>tmavomodrá</i>	8	<i>ciemnobrązowy</i>	5
'gray-green'		'dark blue'		'dark-brown'	
<i>smeđe-crvena</i>	5	<i>fialovorůžová</i>	6	<i>bladorůžowy</i>	4
'brown-red'		'purple-pink'		'pale-pink'	
<i>smeđe-zelena</i>	5	<i>bílomodrá</i>	4	<i>ciemnorůžowy</i>	4
'brown-green'		'white-blue'		'dark-pink'	
<i>sivo-maslinasta</i>	5	<i>hnědozelená</i>	4	<i>ciemnofioletowy</i>	3
'gray-olive-like'		'brown-green'		'dark-purple'	
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’.

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 both 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.

**Table 4.** The ten most frequent terms formed via the pattern [Adv + Adj] in Croatian and Czech

Croatian	N	Czech	N
<i>svijetlo plava</i> ‘light blue’	97	<i>tmavě zelená</i> ‘dark green’	70
<i>tamno zelena</i> ‘dark blue’	80	<i>světle modrá</i> ‘light blue’	63
<i>svijetlo zelena</i> ‘light green’	58	<i>světle zelená</i> ‘light green’	57
<i>tamno plava</i> ‘dark blue’	47	<i>tmavě modrá</i> ‘dark blue’	45
<i>svijetlo roza</i> ‘light pink’	28	<i>světle fialová</i> ‘light purple’	37
<i>svijetlo ljubičista</i> ‘light purple’	27	<i>tmavě fialová</i> ‘dark purple’	23
<i>tamno ljubičasta</i> ‘dark purple’	26	<i>světle růžová</i> ‘light pink’	21
<i>tamno roza</i> ‘dark pink’	12	<i>tmavě hnědá</i> ‘dark brown’	18
<i>svijetlo smeđa</i> ‘light brown’	10	<i>bledě modrá</i> ‘pale blue’	14
<i>tamno siva</i> ‘dark gray’	9	<i>tmavě růžová</i> ‘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
<i>pruská modrá</i> ‘prussian blue’	3	<i>ciemny zielony</i> ‘dark green’	32
<i>světější modrá</i> ‘lighter blue’	3	<i>jasny zielony</i> ‘light green’	21
<i>světější zelená</i> ‘lighter green’	3	<i>jasny niebieski</i> ‘light blue’	17
<i>světlá umělá zelená</i> ‘light artificial green’	2	<i>jasny fioletowy</i> ‘light purple’	11
<i>námořnická modrá</i> ‘navy’	2	<i>ciemny różowy</i> ‘dark pink’	6
<i>světější hnědá</i> ‘lighter brown’	2	<i>ciemny niebieski</i> ‘dark blue’	5
<i>světější růžová</i> ‘lighter pink’	2	<i>jasny błękitny</i> ‘light light blue’	4
<i>sytější fialová</i> ‘more saturated purple’	2	<i>ciemny brązowy</i> ‘dark brown’	3



<i>tmavější zelená</i> ‘darker green’	2	<i>ciemny fioletowy</i> ‘dark brown’	3
<i>Trávová zelená</i> ‘grassy green’	2	<i>jasny brązowy</i> ‘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 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 [Adv [o] Adv] 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 and 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.

**Table 6.** The ten most frequent terms formed using nouns in the three languages

<b>Croatian</b>	<b>N</b>	<b>Czech</b>	<b>N</b>	<b>Polish</b>	<b>N</b>
<i>boja breskve</i> ‘peach color’	4	<i>tyrkys</i> ‘(the) turquoise’	3	<i>jasny fiolet</i> ‘(the) light purple’	20
<i>boja višnje</i> ‘sour-cherry color’	1	<i>akvamarin</i> ‘(the) aquamarine’	1	<i>ciemna zieleń</i> ‘(the) dark green’	19
<i>boja cigle</i> ‘brick color’	1	<i>tmavši akvamarin</i> ‘(the) darker aquamarine’	1	<i>ciemny fiolet</i> ‘(the) dark purple’	10
<i>boja vina</i> ‘vine color’	1	<i>tmavější, sytější tyrkys</i>	1	<i>jasna zieleń</i> ‘(the) light	6

		'(the) darker, deeper turquoise'		green'	
–	–	–	–	<i>ciemny róż</i> '(the) dark pink'	5
–	–	–	–	<i>zgnila zieleń</i> '(the) rotten green'	5
–	–	–	–	<i>blady róż</i> '(the) pale pink'	3
–	–	–	–	<i>ciemny pomarańcz</i> '(the) dark orange'	3
–	–	–	–	<i>jasny róż</i> '(the) light pink'	3
–	–	–	–	<i>intensywny róż</i> '(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 *zgnila* '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.

## 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%).

	CROATIAN		CZECH		POLISH	
ORDER	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;
- 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: Routledge.
- Conklin, H. C. (1973). Color Categorization: Basic Color Terms: Their Universality and Evolution. Brent Berlin, Paul Kay. *American Anthropologist*, 75(4), 931–942.
- Fábregas, A., & Scalise, S. (2012). *Morphology. From data to theories*. Edinburgh: Edinburgh University Press.
- 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.
- Kay, P., Berlin, B., & Merrifield, W. (1991). Biocultural Implications of Systems of Color Naming. *Journal of Linguistic Anthropology*, 1(1), 12–25.
- 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.



- 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.
- 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.
- Raffaelli, I. (2009). *Značenje kroz vrijeme: poglavlja iz dijakronijske semantike*. Disput: Zagreb
- Raffaelli, I. (2012). The conceptual category of ‘light’ in Croatian: A diachronic perspective. In M. Brdar, I. Raffaelli, & M. Žic Fuchs (Eds.), *Cognitive linguistics between universality and variation* (pp. 383–410). Newcastle upon Tyne: Cambridge Scholars Publishers.
- Raffaelli I. (2013). The model of morphosemantic patterns in the description of lexical architecture. *Lingua 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.
- 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.
- 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.
- Silić, J., & Pranjković, I. (2007). *Gramatika Hrvatskoga jezika*. Zagreb: Školska knjiga.
- 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.