



ARTICLE



<https://doi.org/10.1057/s41599-023-02393-4>

OPEN

The Italian colour lexicon in Tuscany: elicited lists, cognitive salience, and semantic maps of colour terms

Maria Michela Del Viva^{1✉}, Serena Castellotti¹ & Galina V. Paramei²

We investigated the Tuscan Italian colour inventory, with the aim of establishing the cognitive salience of the basic colour terms (BCTs) and most frequent non-BCTs. Native speakers from Tuscany ($N = 89$) completed a colour-term elicitation task lasting for 5 min. In total, 337 unique terms were elicited, with an average list length of 30.06. The frequency of each term, its mean list position and cognitive salience index (S) were calculated. The CTs with the highest S (ranked 1–13) included 10 counterparts of the Berlin and Kay BCTs listed in their 1969 seminal work and three basic ‘blue’ terms, *blu*, *azzurro*, *celesti*, estimated for Tuscan respondents by Del Viva et al. in 2022. S -index and Zipf-function (the terms’ “popularity”) indicated that *fucsia* (rank 14) is conceivably an emerging BCT in (Tuscan) Italian. Other cognitively salient non-BCTs are *lilla*, *magenta*, *ocra* and *beige*. The terms’ 3D semantic map (conceptual closeness), assessed using multidimensional scaling and cluster analysis, revealed that in the lists, closely associated CTs were arranged along three competing criteria: the term’s salience gradient; word length; and clustering of fully chromatic concepts with those defined primarily by lightness or desaturation. We also consider salient Italian non-BCTs as indicators of the ongoing process of lexical refinement in certain areas of the colour space. In conclusion, measures of elicitation productivity, as well as the augmented BCT inventory, including the Tuscan ‘triple blues’, and abundant hyponyms and derived forms all indicate (Tuscan) Italian speakers’ “cultural competence” in the colour domain and the need to communicate nuanced information about colour shades.

¹Department of NEUROFARBA, Università degli Studi di Firenze, Via di San Salvi 12, 50135 Florence, Italy. ²Department of Psychology, Liverpool Hope University, Hope Park L16 9JD, Liverpool, UK. ✉email: maria.delviva@unifi.it

Introduction

The universal colour term inventory. Colour nomenclature has been studied in a wide range of languages; to a great degree this was motivated by the question of how the conceptual system in the colour domain is shaped, with three potential solutions offered. According to Berlin and Kay (1969/1991), there are universal, pan-cultural basic colour categories (BCCs) that partition the colour space (the perceived colour gamut) into conceptual regions. BCCs evolve in a certain order, reach a maximum number of 11 and are named by basic colour terms (BCTs), the “core vocabulary” of the colour language one uses for an initial chromatic description of any object in the visual environment. An alternative, relativist view, or linguistic-convention account is posited by Saunders and van Brakel (1997) who conjecture that colour categories/terms are language-specific and shaped by non-trivial cultural constraints.

Currently broadly accepted is the weak relativist account (for a review see Lindsey and Brown, 2021). According to it, across languages universal perceptual constraints govern the position of BCCs in an irregularly shaped colour solid; it admits, however, cross-language variations of the category boundaries and the order of emergence of BCTs (Regier et al., 2007). As pointed out by Conklin (1973), the semantic structure of colour terms (CTs) is affected by various factors, such as the kind of colour (surface or luminous), colourant technology, natural or artefactual object modality, etc. Ronga and Bazzanella (2015) echo this view remarking that colour categorisation and lexicalisation are shaped by sociocultural and historical parameters, in particular, by choices of typical BCC exemplars, as well as the variety of the terms that allow lexical distinction of perceived colour nuances.

Furthermore, Kay and McDaniel (1978, p. 640) acknowledged the possibility that several non-BCTs, used to name intersections of BCCs, would become basic: “there is no apparent reason to believe that the process will not continue, extending basic colour-term lexicons beyond their present eleven terms”. Indeed, there has been accumulating evidence that in some languages the basic nomenclature can exceed the ceiling of 11, whereby an incipient BCC may develop as a cross-language universal, or emerge as a “culturally basic” language-specific, as argued by Paramei (2005). This empirical observation is buttressed by a computational model of Baronchelli et al. (2015), who demonstrated that the universal properties of colour naming patterns emerge independently in every language in response to universal communicative, cognitive or perceptual biases. Along with these, at play is the cultural transmission bias: the specific cultural history of a language is effective in shaping non-universal properties of colour naming due to the need to communicate the prominent colour shades of speakers’ natural and artefactual environment, as well as cross-language contacts.

The basic status of the frequent non-BCTs is ascertained using linguistic and psycholinguistic methods against the criteria delineated by Berlin and Kay (1969/1991, p. 6): a colour term is defined as basic if it is (i) monolexemic, (ii) not subsumed under the meaning of any other BCT, (iii) not applied only to a limited class of objects, and (iv) psychologically salient as evidenced by frequent use across occasions, and “occurrence in the ideolects of all informants”. As remarked by Hays et al. (1972, p. 1107), the criteria “are plainly delicate in application”; not infrequently the estimation of CT basic status requires combining various measures of a term’s salience.

The elicitation task. For the exploration of a particular semantic domain, one of the important methods used is free-listing, or the elicitation technique, developed by Weller and Romney (1988).

Free-listing is relatively easy to administer and allows for rapid data collection, whether respondents write down the items themselves or give oral responses. This “deceptively simple, but powerful technique” (Bernard, 2006, p. 301) provides rich and relevant data that enable the researcher to define, for the tested population, the “conceptual sphere”, a mental inventory of items, that constitute the domain in question, and to ascertain the salience of frequently mentioned items (Borgatti, 1999; Quinlan, 2019; Stausberg, 2022; Weller, 2014).

The elicitation of CTs, in the absence of any colour stimuli, was shown to be a convenient method for use in the colour domain. Participants are invited to list all the colour names they can think of during a short period (usually 5 min) (Borgatti, 1998; Corbett and Davies, 1995; Corbett and Morgan, 1988; Jakovljević and Zdravković, 2018; Sutrop, 2001; Uusküla, 2007; Xu et al., 2023, to name just a few). Across the obtained informants’ lists, one estimates the number of times each CT is elicited and the decreasing order of the terms’ frequency. A few CTs are mentioned by many respondents but many other terms, less popular, are listed by only a few. The second measure of a term’s salience is its “height” on the informant’s list: the closer the CT is to the top of the list, the more salient it is.

The term’s frequency and position in the elicited list constitute our equivalent of Berlin and Kay’s criterion (iv) for ascertaining the term’s psychological salience. In their analysis of CT frequencies in four languages (Russian, American English, British English and French), Corbett and Davies (1997) demonstrated that the frequency measure (its rank), obtained from the 5-min listing task, reliably distinguishes BCTs from non-BCTs and can differentiate between primary and secondary BCTs; this measure was found not to correlate with the Berlin and Kay hierarchy of BCTs.

A list of the elicited CTs, sorted according to their frequency (“scree plot”, a downward curve, ordering the measures from largest to smallest), is a function that reflects the term’s salience gradient and follows a roughly exponential decline. It often shows an “elbow”, a natural break that allows identification of the most salient colour terms as having basic status; these are separated from stretches of slower decline for non-BCTs, hyponyms, complex and idiosyncratic terms. However, frequency functions do not always reveal a clear-cut break between BCTs and non-BCTs, as established by CT frequency in corpus analysis or psycholinguistic experiments: that is, a function derived from CT frequency may show either a gradual decline or a few less punctuated drops (Bimler and Uusküla, 2014b, 2018; Jakovljević and Zdravković, 2018). Moreover, as remarked by Uusküla and Bimler (2016, p. 72), “salience on its own is sometimes misleading as an indicator of basicness” since some frequent non-BCTs may fall on the BCT side of the step, ahead of some low(er)-salience secondary BCTs. There may be several factors affecting why non-BCTs come to respondents’ minds first (“leap out”), thus distorting the expected salience order of BCTs, followed by non-BCTs: e.g. a CT may be culturally salient because it describes a significant object in the environment and lives of respondents; there may be semantic links between BCTs and specific objects; or competing CTs may exist for naming the same BCC (Bolton, 1978).

To solve this “riddle of colour term salience”, apart from the term’s frequency, Bolton (1978) suggested using an additional indicator of salience—the tendency of a term to appear at the head of an elicited list; across respondents, “earliness” of the term’s occurrence is reflected by its mean position rank. His idea was operationalised by Sutrop (2001), who introduced the cognitive salience index (S), which combines both the term’s frequency and mean position ranking.

Along with the CT cognitive salience index (S), Bimler and Uusküla (2014b, 2018; Uusküla and Bimler, 2016) introduced an analysis of a *conceptual map* based on the measure of CT adjacencies in the elicited lists. Processed by cluster analysis and multidimensional scaling analysis, this measure allowed estimation of two further aspects of mental representation of CTs: (i) a term's "priority", or the tendency to appear at the start of the list; the *gradient of salience* varies from primary BCTs, through secondary BCTs, to frequent non-BCTs; (ii) the identification of CT "chunks" that reveal respondents' *elicitation strategies*, among these, contrastive associations of terms for opponent primary colours; recollection of CTs according to their perceptual similarity or contiguity; and identification of clusters of hue terms coalesced with their complex forms denoting finer chromatic distinctions. Uusküla and Bimler (2016) found that this pattern of listing associations was similar cross-lingually; they conjecture that it reflects conceptual themes at a higher level of abstraction and the ways in which activation in the nodes of the CT semantic network is spread.

Italian colour terms. To preface the discussion of Italian colour terms and the report of the present findings, we would like to remark that Modern Standard Italian is quite a recent language: before Italian unification (*Risorgimento*, 1815–1870) the use of dialects was widespread throughout the peninsula (De Mauro, 1983). The spread of the Florentine dialect as the common language has not eliminated the traditional multilingual aspect of Italy (De Renzo, 2009), with the La Spezia-Rimini Line marking a series of isoglosses distinguishing Northern Italian speech from that of Tuscany (Società Linguistica Italiana, 1995). Apparently, conceptualisation, including in the colour domain, depends on the region of Italy, i.e. exposure to the corresponding dialect.

In contemporary Italian, counterparts of the ten Berlin and Kay BCTs were established in a psycholinguistic experiment (Paggetti et al., 2016): *bianco* 'white', *nero* 'black', *rosso* 'red', *verde* 'green', *giallo* 'yellow' (primary BCTs), and *marrone* 'brown', *arancione* 'orange', *viola* 'purple', *rosa* 'pink', and *grigio* 'grey' (secondary BCTs).

Compared to English, German or French which all possess one basic 'blue' term (*blue*, *blau*, *bleu* respectively), in Italian the blue area is denoted by more than one basic colour term, as demonstrated in linguistic studies (Giacalone Ramat, 1967; Grossmann, 1988; Sandford, 2015; Vincent, 1983). It is noteworthy that this observation does not imply a lack of linguistic richness in the denotation of the blue area in all these languages by means of multiple non-BCTs, compounds or modified BCTs denoting various blue shades (e.g. *periwinkle* or *baby blue* in English; *turkiz* or *königsblau* in German; *céruléen* or *bleu pétrole* in French).

Recent psycholinguistic studies carried out in different regions of Italy, adduce considerable evidence of more than one basic 'blue' term and, in addition, diatopic variation of Italian basic 'blue' terms. Specifically, *azzurro* 'light-and-medium blue' and *blu* 'dark blue' are basic for speakers of the Veneto region (Paggetti and Menegaz, 2012; Paggetti et al., 2016; Paramei et al., 2014) and Trento (Albertazzi and Da Pos, 2017). For speakers of the Algherese-Catalan dialect (Sardinia), the 'light-and-medium blue' basic term is lexicalised as *celeste*, complemented by *blu* 'dark blue' (Paramei et al., 2018a). Tuscan speakers, in comparison, appear to require triple basic "blues": *celeste* 'light/sky blue', *azzurro* 'medium blue' and *blu* 'dark blue' (Bimler and Uusküla, 2014a; Del Viva et al., 2022).

This last finding is relevant for the present study since the reported data were collected from speakers in Florence and the areas immediately surrounding it, i.e. from the Tuscan dialect,

which presents some lexical differences with respect to standard Italian (Wieling et al., 2014). The choice of the Tuscan area was motivated by practical reasons—by its proximity to the authors' institution. In an ideal world, one would wish to collect data in various regions of Italy; this would require, however, a large-scale collaborative endeavour across Italian institutions.

Aims of the study. The aims of the present study are the following: (i) to assess the (Tuscan) Italian colour term inventory using analysis of the elicited lists and, based on the terms' frequency and cognitive salience, to estimate the BCT inventory; (ii) to further explore the (basic) status of the three "Tuscan blues"; and (iii) to estimate the cognitive salience of Italian high-frequency non-BCTs that may be emerging as (culturally) basic terms.

The novelty of the present study is the estimation of the cognitive salience of (Tuscan) Italian BCTs ($N = 13$) and high-frequency non-BCTs (about 20), whereby the (basic) status of the terms was ascertained using linguistic and psycholinguistic indices against the criteria delineated by Berlin and Kay (1969/1991) outlined above. In addition, we found that the Modern Italian BCT inventory undergoes augmentation, with *fucsia* 'fuchsia' as an incipient basic term. Furthermore, we ascertained linguistic patterns of non-monolexic expressions that apparently enable greater communication efficiency of a perceived colour. Finally, we identified areas of the colour space that undergo lexical refinement in contemporary Italian.

Methods

Participants. Participants ($N = 89$; 68 females) were students or graduates from the University of Florence or high school graduates in Tuscany, with a mean age of 23.2 ± 3.0 and an age range of 20–31 years. All participants were native Italian speakers born and residing in Tuscany. For detailed information on demographic characteristics (see Table S1).

Procedure. Each participant undertook a colour-term elicitation task. The procedure followed the one developed by Davies and Corbett (1994, 1995), with the following instruction to participants (provided in Italian by a native speaker experimenter): "Please write as many colours as you know". The elicitation task was carried out online (via Zoom or Meet). Just before data collection, the experimenter shared, via Google Drive, an Excel file with a participant, who was instructed to type into it colour terms using her/his keyboard. The task was limited to 5 min. The experimenter started the clock and informed the participant about the time lapse after each minute.

Data analysis. All morphologically modified terms were counted separately. The obtained lists underwent data cleansing. Specifically, spelling was regularised (i.e. *verdeacqua*, *verdeacqua* → *verde acqua*; *acqua marina* → *acquamarina*); spelling errors and typos were corrected (e.g. *fuxia* → *fucsia*; *bordaux*, *bordeaux*, *bordoux* → *bordeaux*; *bnero* → *nero*). We treated two Italian variants of 'orange', *arancione* (a BCT) and *arancio* (less frequent) as different items; in comparison, two versions of 'straw yellow', *giallo paglierino* and *paglierino giallo*, were treated as synonyms. Also, the entrenched Italian metonymy *carta da zucchero* (denoting the powder blue colour of sugar paper used from ca. 1890) replaced other term variants (*carta zucchero*; *blu carta da zucchero*; *color carta da zucchero*). Finally, in a few cases, where some colour terms were listed twice, the duplicates were eliminated. If any item appeared more than once on a list, only the first occurrence was taken as valid.

Individual lists of colour terms were recorded in the order in which the participants typed them. From the collected data, we calculated the frequency (F) of each colour term occurrence across the lists, as well as the term's mean position (mP), and estimated the corresponding colour-term ranks, R_F and R_{mP} (cf. Morgan and Corbett, 1989).

In addition, we estimated the Zipf-function, or power-law relation, that reflects the term's popularity. As suggested by Lindsey and Brown (2014), we calculated the number of participants (\log_{10}) who listed the term (F) as a function of the value (\log_{10}) of the term's frequency-based rank (R_F). When graphed, this function falls onto a line (for a review, see Mitzenmacher, 2003). The Zipf-function allows the researcher to assess a double-power law behaviour that reflects the difference between the terms used for general but imprecise communication (a kernel lexicon) and the terms for more specific communication, in this case: the difference in the use of BCTs and (frequent) non-BCTs (Ferrer i Cancho and Solé, 2003). Guided by Lindsey and Brown's (2014) and Kuriki et al. (2017) approach, we fitted limbs of the Zipf-function by linear equations using a least-squares criterion.

For the frequency (ranking) analysis and Zipf-function, we used the list of frequent colour terms ($N_{FCTs} = 96$) that were offered by at least five respondents. This cut-off number slightly deviates from the cut-off of four suggested by Sutrop (2001), who argues that, for a sample of over 50 participants, the terms listed by three or fewer respondents are likely to be part of their idiolects. Since in the present dataset the list of colour terms produced by at least four participants was relatively long ($n = 112$), we constrained the list to colour terms elicited by at least five respondents while relying on Borgatti's (1998) guidance that the top number can be arbitrarily chosen to handle in the remaining part of the study.

For further analysis of the terms' cognitive salience, the two measures, F and mP, were combined into the Cognitive Salience Index (S) using the formula suggested by Sutrop (2001):

$$S = F/(N \times mP) \quad (1)$$

where F is the frequency (the number of participants who listed the term), N is the total number of participants and mP is the mean position of the term in the list. The S -index can vary between 1 (when the term is present in all lists, i.e. maximally salient) and 0 (the term is not present in any list), with the salience ranking (R_S) estimated accordingly, in descending order. The S -index is independent of the list length and allows comparison of the present results with analogous indices reported for other languages.

Finally, we explored the semantic map, or the pattern of conceptual closeness among elicited colour names, assuming that in a semantic network, closely associated terms tend to prime each other and, hence, appear in the lists in close succession (cf. Friendly, 1977). We were guided by Uusküla and Bimler (2016), who used colour-term listing data to estimate inter-name link strength. Specifically, following Friendly's (1977) approach, Uusküla and Bimler (2016) defined the measure of 'adjacency', $ADJ(i, j)$, between the i th and the j th terms as the absolute difference in the sequence position, for each colour-term pair. For an individual p th respondent, a matrix of separation, SEP_p , was obtained with entries:

$$SEP_{p,ij} = \ln |s_{pi} - s_{pj}| \quad (2)$$

For two consecutive terms, $SEP_{p,ij} = 0$. The number of participants, for whom $SEP_{p,ij}$ is defined (i.e. both terms i and j are present in their lists), is c_{ij} . $ADJ(i, j)$, the measure derived from $SEP_{p,ij}$, uses the mean of the terms' separations, averaged across

only those participants in whose lists the terms i and j co-occur:

$$ADJ(i, j) = \exp \left[\left(\sum_p SEP_{p,ij} \right) / c_{ij} \right] \quad (3)$$

Adjacency matrices were analysed with multidimensional scaling (MDS) using PROXSCAL software within SPSS 28.0, with ordinal transforms (ranks) of adjacencies among the most salient terms processed by the non-metric MDS. An MDS analysis enabled to represent a semantic map of elicited CTs as points in a low-dimensional Euclidean space: the closer together the items are in the list, the closer the corresponding points in the spatial representation. Interpreting the dimensions enabled the researchers to reveal semantic attributes that respondents use (frequently unaware of these) while recalling colour terms.

We computed a 2D and a 3D solution, comparing S -Stress (badness-of-fit) indices between the two. Further, k -means hierarchical cluster analysis (HCA) was applied to CT-point coordinates of the 2D solution by applying Ward's (1963) algorithm implemented in MATLAB (Mathworks, Natick, MA).

Results

The number and variety of elicited colour terms. In total, participants produced 2675 items, 337 of which were unique colour names. For the full list of the elicited terms see Table S2. The elicited list included monolexemic BCTs and non-BCTs (e.g. *rosso* 'red', *indaco* 'indigo', *avorio* 'ivory'). Not infrequently compound terms were also offered (e.g. *verde acqua* 'sea-green'; *grigio topo* 'mouse grey'; *giallo limone* 'lemon yellow'), as well as terms with achromatic modifiers (e.g., *rosa pastello* 'pastel pink'; *olmo chiaro* 'light elm'; *viola scuro* 'dark purple') or suffixed terms (e.g. *biancastro* 'whitish', *giallo aranciato* 'yellow orangish', *marroccino* 'brownish'). In addition, we observed a plethora of terms with a metaphorical reference to objects (e.g. *verde pistacchio* 'pistachio green'; *vinaccia* 'wine-red, colour of grape marc/pomace') or geographic locations (*giallo Napoli* 'Naples yellow'; *fumo di Londra* 'London smog'), and metonymic names entrenched in Italian [e.g. *carta da zucchero* 'powder blue' (previously the colour of sugar paper); *terra di Siena* 'sienna'].

In individual lists the number of items varied between 15 and 58, with an average list containing 30.06 colour terms, indicating the richness of the colour-term inventory of Tuscan speakers. The full list of the elicited terms (Table S2) prompts another observation—that the participants offered a great variety of monolexemic terms and morphologically modified colour names, which suggest both the awareness of perceived shades and the need for nuanced lexicalisation of these shades.

The frequency of colour terms. Table 1 shows the list of Italian colour terms ($N_{FCTs} = 96$) that were offered by at least five participants. In Table 1, in English, we gloss *blu* as 'dark blue', *azzurro* as 'medium blue' and *celest* as 'light/sky blue', according to our findings of the denotative meanings of the three 'blue' terms for Tuscan speakers (Del Viva et al., 2022). Table 1 presents colour-term frequency, F (the number of participants who listed the term), the mean position of the term, mP, and the Cognitive Salience Index, S , for each term accompanied by the rank according to each of the three measures (R_F , R_{mP} , R_S , respectively). Note that in Table 1 the terms are listed according to their frequency rank, R_F , in descending order. (F , $F\%$ and R_F for the full list of the elicited terms are presented in Table S2.)

As is apparent from Table 1, the Italian counterparts of the 11 BCTs proposed by Berlin and Kay (1969/1991) occupy the top positions. It is worth noting that *blu* ($R_F = 4$) is the most frequent and salient among the three 'blue' terms, whereas *celest* and *azzurro* are relegated to lower ranks ($R_F = 13$ and $R_F = 14$ respectively). Prominent in frequency are also *fucsia* 'fuchsia'

Table 1 List of the most frequent Italian colour terms for Tuscan speakers.

Colour term	English gloss	F	F%	R _F	mP	R _{mp}	S-index	R _S
nero	black	87	97.8	1.0	9.89	5	0.09889	5
rosso	red	86	96.6	2.0	4.99	2	0.19371	2
giallo	yellow	85	95.5	4.0	4.86	1	0.19656	1
blu	dark blue	85	95.5	4.0	5.82	3	0.16400	3
viola	purple	85	95.5	4.0	11.07	8	0.08627	7
verde	green	83	93.3	7.0	5.90	4	0.15797	4
rosa	pink	83	93.3	7.0	11.19	9	0.08332	8
bianco	white	83	93.3	7.0	12.14	11	0.07679	9
arancione	orange	82	92.1	9.0	10.41	6	0.08847	6
marrone	brown	81	91.0	10.0	15.09	22	0.06033	12
grigio	grey	77	86.5	11.0	14.17	15	0.06106	11
fucsia	fuchsia	73	82.0	12.0	14.45	17	0.05675	14
celeste	light/sky blue	65	73.0	13.0	12.52	12	0.05832	13
azzurro	medium blue	63	70.8	14.0	10.78	7	0.06568	10
lilla	lilac	62	69.7	15.0	14.71	18	0.04736	15
beige	beige	54	60.7	16.0	20.48	45	0.02962	19
verde acqua	sea-green	53	59.6	17.0	15.68	24	0.03798	16
bordeaux	claret	49	55.1	18.0	18.98	38	0.02901	20
argento	silver	47	52.8	19.0	22.74	63	0.02322	23
oro	gold	45	50.6	20.0	21.73	57	0.02326	22
ocra	ochre	42	47.2	21.5	14.88	21	0.03171	17
magenta	magenta	42	47.2	21.5	15.21	23	0.03102	18
indaco	indigo	40	44.9	23.0	19.00	39	0.02365	21
porpora	cardinal red	36	40.4	24.0	18.17	33	0.02227	24
turchese	turquoise	27	30.3	25.0	16.74	26	0.01812	25
amaranto	amaranth	25	28.1	26.0	17.64	31	0.01592	27
verde scuro	dark green	23	25.8	27.0	14.78	20	0.01748	26
verde chiaro	light green	21	23.6	28.0	16.57	25	0.01424	28
violetto	violet	20	22.5	30.0	20.50	46	0.01096	30
giallo canarino	canary	20	22.5	30.0	21.10	49	0.01065	31
blu notte	night <i>blu</i>	20	22.5	30.0	21.20	52	0.01060	32
crema	(yellowish-)cream	17	19.1	32.0	24.35	71	0.00784	37
rosa antico	antique pink	16	18.0	33.0	22.44	62	0.00801	34
ciano	cyan	15	16.9	35.0	14.07	14	0.01198	29
giallo ocra	ochre yellow	15	16.9	35.0	18.60	36	0.00906	33
verde smeraldo	emerald green	15	16.9	35.0	21.47	54	0.00785	36
grigio chiaro	light grey	14	15.7	37.5	27.43	83	0.00574	43
giallo fluo	fluorescent yellow	14	15.7	37.5	28.64	89	0.00549	47
verde petrolio	petrol green	13	14.6	40.0	22.92	64	0.00637	40
marrone chiaro	light brown	13	14.6	40.0	25.38	74	0.00575	42
bianco sporco	off-white	13	14.6	40.0	28.15	87	0.00519	52
blu scuro	dark <i>blu</i>	12	13.5	44.0	16.92	27	0.00797	35
avorio	ivory	12	13.5	44.0	21.08	48	0.00640	39
verde militare	military green	12	13.5	44.0	23.75	67	0.00568	45
rosso mattone	brick red	12	13.5	44.0	24.67	72	0.00547	49
grigio scuro	dark grey	12	13.5	44.0	25.75	75	0.00524	50
rosa cipria	powder pink	11	12.4	47.0	21.82	58	0.00566	46
blu elettrico	electric <i>blu</i>	10	11.2	49.0	21.50	55	0.00523	51
panna	(whitish-)cream	10	11.2	49.0	23.00	65	0.00489	53
marrone scuro	dark brown	10	11.2	49.0	27.40	82	0.00410	62
rosso scuro	dark red	9	10.1	54.5	13.44	13	0.00752	38
tortora	taupe	9	10.1	54.5	17.78	32	0.00569	44
acquamarina	aquamarine	9	10.1	54.5	18.44	34	0.00548	48
bronzo	bronze	9	10.1	54.5	21.11	50	0.00479	55
pesca	peach	9	10.1	54.5	22.22	60	0.00455	58
verde bosco	forest green	9	10.1	54.5	23.89	68	0.00423	60
carta da zucchero	powder blue	9	10.1	54.5	24.33	70	0.00416	61
grigio topo	mouse grey	9	10.1	54.5	28.11	86	0.00360	69
blu chiaro	light <i>blu</i>	8	9.0	64.0	14.75	19	0.00609	41
terra di siena	sienna	8	9.0	64.0	18.88	37	0.00476	56
verde bottiglia	bottle green	8	9.0	64.0	19.25	40	0.00467	57
giallo limone	lemon	8	9.0	64.0	20.00	43	0.00449	59
ceruleo	cerulean	8	9.0	64.0	22.13	59	0.00406	63
giallo paglierino	straw yellow	8	9.0	64.0	25.88	77	0.00347	70
rosa pastello	pastel pink	8	9.0	64.0	27.88	85	0.00322	72
rosa fluo	fluorescent pink	8	9.0	64.0	28.38	88	0.00317	75

Table 1 (continued)

Colour term	English gloss	F	F%	R _F	mP	R _{mP}	S-index	R _S
verde pastello	pastel green	8	9.0	64.0	29.63	91	0.00303	77
verde fluo	fluorescent green	8	9.0	64.0	31.63	93	0.00284	82
giallo chiaro	light yellow	8	9.0	64.0	32.86	94	0.00274	85
antracite	anthracite	7	7.9	71.5	21.29	53	0.00370	66
rosso fuoco	fire red	7	7.9	71.5	21.71	56	0.00362	68
sabbia	sand	7	7.9	71.5	25.86	76	0.00304	76
giallo pastello	pastel yellow	7	7.9	71.5	29.57	90	0.00266	88
rosso chiaro	light red	6	6.7	78.5	17.17	28	0.00393	65
verde pisello	pea green	6	6.7	78.5	18.50	35	0.00364	67
salmone	salmon	6	6.7	78.5	21.17	51	0.00318	74
blu oltremare	ultramarine <i>blu</i>	6	6.7	78.5	22.33	61	0.00302	78
arancione chiaro	light orange	6	6.7	78.5	23.33	66	0.00289	80
smeraldo	emerald	6	6.7	78.5	24.00	69	0.00281	83
rame	copper	6	6.7	78.5	25.17	73	0.00268	87
rosso sangue	blood red	6	6.7	78.5	26.50	79	0.00254	89
rosa chiaro	light pink	6	6.7	78.5	27.00	80	0.00250	90
arancione fluo	fluorescent orange	6	6.7	78.5	34.33	96	0.00196	94
senape	mustard	5	5.6	90.0	11.60	10	0.00484	54
ottanio	teal	5	5.6	90.0	14.20	16	0.00396	64
verde menta	mint green	5	5.6	90.0	17.40	29	0.00323	71
rosa salmone	salmon pink	5	5.6	90.0	17.60	30	0.00319	73
blu cobalto	cobalt <i>blu</i>	5	5.6	90.0	19.40	41	0.00290	79
arancione scuro	dark orange	5	5.6	90.0	19.60	42	0.00287	81
tiffany	tiffany blue	5	5.6	90.0	20.40	44	0.00275	84
blu cielo	sky <i>blu</i>	5	5.6	90.0	20.80	47	0.00270	86
mattone	brick-coloured	5	5.6	90.0	26.00	78	0.00216	91
nocciola	hazelnut	5	5.6	90.0	27.20	81	0.00207	92
blu navy	navy <i>blu</i>	5	5.6	90.0	27.80	84	0.00202	93
giallo scuro	dark yellow	5	5.6	90.0	31.60	92	0.00178	95
rosa scuro	dark pink	5	5.6	90.0	33.20	95	0.00169	96

The most frequent Italian colour terms (N_{FACTS} = 96) with English glosses. Each term is accompanied by its absolute frequency (F), relative frequency (F%), mean position (mP) in the list, and Cognitive Salience Index, S, as well as the corresponding ranks based on the term's frequency (R_F), mean position (R_{mP}), and cognitive salience (R_S), respectively. The terms are listed in descending order according to the frequency. The terms constituting the first limb of the Zipf-function (see below) are highlighted by dark grey and those of the second limb by light grey.

(R_F = 12), *lilla* 'lilac' (R_F = 15) and *beige* (R_F = 16), as well as the compound *verde acqua* 'sea-green' (R_F = 17), salient for Italian speakers due to the prominence of the sea in their "visual diet".

Figure 1 shows the function of frequencies of the elicited colour terms (N_{FACTS} = 96), which follows a roughly exponential decline. Guided by Borgatti (1998), we searched for natural breaks in the distribution. In the function (Fig. 1), we discern two smaller breaks—between *fucsia* 'fuchsia' (R_F = 12) and *celeste* 'light blue' (R_F = 13), and between *lilla* 'lilac' (R_F = 15) and *beige* (R_F = 16). The former break conceivably manifests the hurdle between BCTs and non-BCTs, although estimation of the status of high-frequency CTs, with R_F > 11, would be better informed in combination with other indices considered below. A third, moderate break can be observed between *porpora* 'cardinal red' (R_F = 24) and *turchese* 'turquoise' (R_F = 25).

The Zipf-function. The Zipf-function presents frequency data on a logarithmic scale, to render the decline roughly linear. According to Ferrer i Cancho and Solé (2003), in large language corpora the Zipf-function is expected to have two segments, whereby two exponents divide words that form a kernel lexicon from a less popular lexicon used for specific communication. In the present context, two Zipf-function segments would be expected to reflect a division between BCTs and non-BCTs. However, inspired by findings in recent studies of colour lexicons in other languages (Brown et al., 2016; Jakovljević and Zdravković, 2018; Lindsey and Brown, 2014), we undertook the division of the Zipf-function into three segments. The rationale is that the division of the function beyond the first segment reveals the prominence of colour terms beyond the traditional BCT/non-

BCT dichotomy (Lindsey and Brown, 2021), differentiating the terms frequently used in daily communication and possibly emerging BCTs (second segment) from the terms that are tertiary, not commonly used (third segment).

Considering the above consideration pertinent to our data, we fitted these (Fig. 2) with the following formula:

$$\log \text{Popularity}_i = \min(a * \log \text{Rank}_i + b, c * \log \text{Rank}_i + d, e * \log \text{Rank}_i + f) \tag{4}$$

The best-fitting constants (Eq. (1)) were obtained by a least-squares criterion and described as a double-power law function (overall R² = 0.9). Figure 2 shows the three-segment Zipf-function, with each limb characterised by its slope.

The first limb has the slope close to zero (-0.0295); it includes 10 BCTs named by almost all participants (F ≥ 91.0%): *nero* 'black', *rosso* 'red', *giallo* 'yellow', *viola* 'purple', *bianco* 'white', *blu* 'dark blue', *rosa* 'pink', *verde* 'green', *arancione* 'orange', and *marrone* 'brown' (in Table 1, highlighted by dark grey).

The second descending limb has a shallow slope -0.9409; it corresponds to CTs with lesser frequencies (86.5% ≥ F ≥ 40.4%), ranked 11–24 (in Table 1, highlighted by light grey). A BCT *grigio* 'grey' (F = 86.5%) and two 'blue' terms, *celeste* 'light blue' (F = 73.0%) and *azzurro* 'medium blue' (F = 70.8%), basic for Tuscan speakers (Del Viva et al., 2022), are better fitted within the group of high-frequency non-BCTs. Among the latter, as Fig. 2 shows, the high popularity terms are *fucsia* 'fuchsia' (F = 82.0%), *lilla* 'lilac' (F = 69.7%), *beige* (F = 60.7%), *verde acqua* 'sea-green' (F = 59.6%), *bordeaux* 'claret' (F = 55.1%), *ocra* 'ochre'

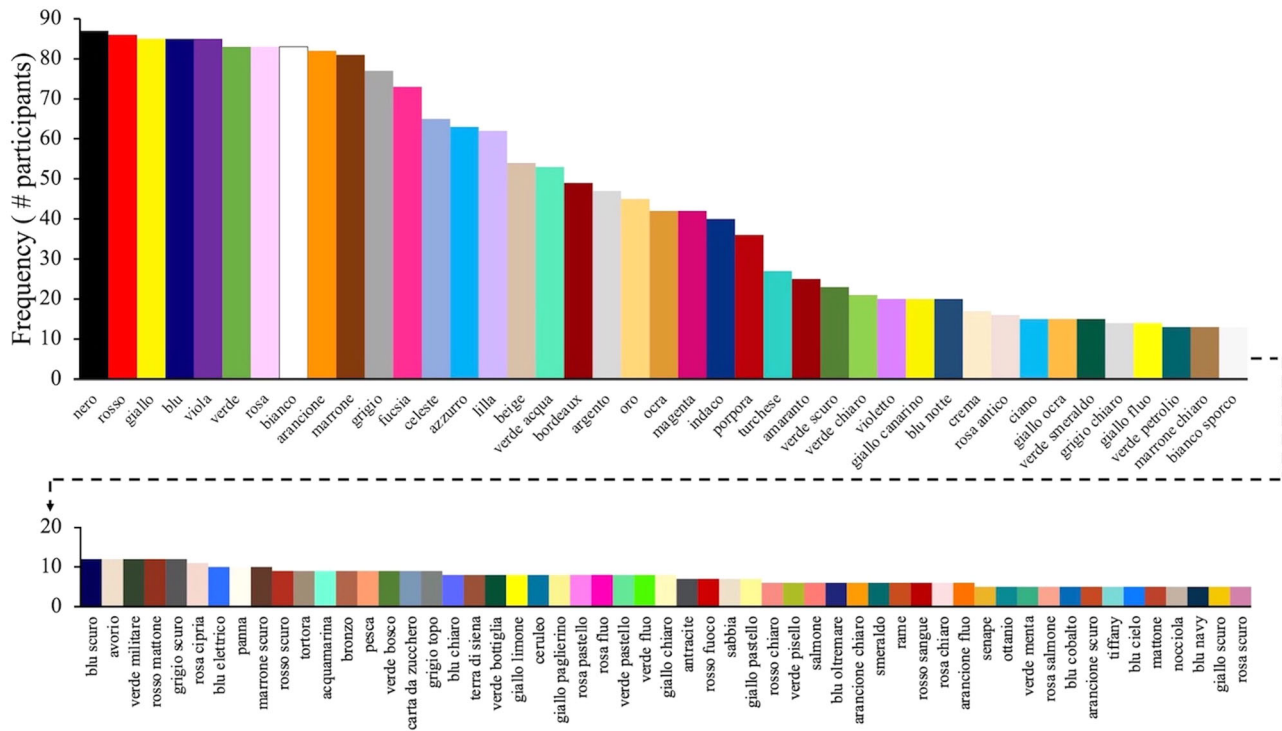


Fig. 1 Frequency of colour terms. Frequency of CT (rendered in the corresponding colour) ranked 1-96 and was produced by at least five participants.

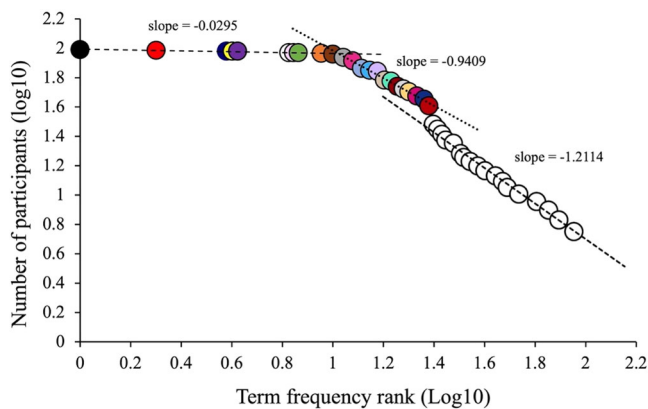


Fig. 2 Zipf-function. Colour term popularity diagram. Limb slopes were fitted by a trilinear equation using a least-squares criterion. Data points of the first two limbs are corresponding-colour coded by CT meaning. Limb 1: 10 BCTs (ranks 1-10); Limb 2: three less popular BCTs (*grigio*, *celeste*, *azzurro*) and highly popular non-BCTs (ranks 11-24); Limb 3 (ranks 25-96; empty circles): low-popularity non-BCTs.

($F = 47.2\%$) and *magenta* ($F = 47.2\%$). This group also includes two terms of metallic brilliance, *oro* ‘gold’ and *argento* ‘silver’. All other non-BCTs, with low popularity ($30.3\% \geq F \geq 5.6\%$), ranked 25-96, form the third limb with a steeper slope of -1.2144 , with rare (tertiary) colour terms offered by few participants.

The mean position of colour terms. We explored the mean position (mP) of the terms, i.e. the tendency to be listed towards the beginning of the elicited list. This measure usually manifests a gap between the six primary and five (or more) secondary BCTs, thus, distinguishing these two BCT types (e.g. Hippisley et al.,

2008). As can be seen in Table 1, five Italian primary BCTs have the highest mP: *giallo* ‘yellow’ 4.86; *rosso* ‘red’ 4.99; *blu* ‘dark blue’ 5.82; *verde* ‘green’ 5.90, and *nero* ‘black’ 9.89, with respective R_{mP} varying from 1 through 5. However, the $mP = 12.14$ of BCT *bianco* ‘white’ is lower; according to the corresponding rank ($R_{mP} = 10$), it falls within R_{mP} -range of secondary BCTs (for Tuscan speakers): *arancione* ‘orange’ ($R_{mP} = 6$), *azzurro* ($R_{mP} = 7$), *viola* ‘purple’ ($R_{mP} = 8$), *rosa* ‘pink’ ($R_{mP} = 9$), *celeste* ($R_{mP} = 11$) and *grigio* ‘grey’ ($R_{mP} = 13$). Another “anomaly” is the relatively low rank of the secondary BCT *marrone* ‘brown’ ($R_{mP} = 18$), which is lower than those of four frequent non-BCTs, *ciano* ‘cyan’ ($R_{mP} = 12$), *fucsia* ‘fuchsia’ ($R_{mP} = 14$), *lilla* ‘lilac’ ($R_{mP} = 15$), *ocra* ‘ochre’ ($R_{mP} = 17$) and the compound *verde scuro* ‘dark green’ ($R_{mP} = 16$).

The Cognitive Salience Index. The Cognitive Salience Index, S , which combines both F and mP measures (Sutrop, 2001), provides an additional criterion for discriminating basic and non-basic CTs. Table 2 presents the list of Italian colour terms ($N_{HS} = 32$) with the highest Cognitive Salience Index, $S \geq 0.01$. It is apparent that the list of the most salient terms (highlighted by darker grey) includes 13 BCTs, i.e. the 11 BCTs identified by Berlin and Kay (1969/1991), and two additional ‘blue’ terms, *azzurro* ($R_S = 10$) and *celeste* ($R_S = 13$), basic for Tuscan speakers (Del Viva et al., 2022). As illustrated by Fig. 3, the cognitive salience function is championed by the four primary chromatic BCTs. These are followed by a noticeable drop-off in salience between *verde* ‘green’ ($S = 0.158$) and *nero* ‘black’ ($S = 0.099$). The Cognitive Salience Index of *bianco* ‘white’ ($S = 0.077$, $R_S = 9$) indicates that for young Tuscan Italians this term is “less basic” than the other primary BCTs.

Notably, the S -index of *fucsia* ‘fuchsia’ ($S = 0.057$, $R_S = 14$) is only slightly lower than that of *celeste* ($S = 0.058$, $R_S = 13$). Thereafter, there is a perceptible S -index gap at non-basic *lilla* ‘lilac’ ($S = 0.047$, $R_S = 15$) and another gap at *verde acqua* ‘sea-green’ ($S = 0.038$, $R_S = 16$), followed by a gradual S -index decrease of

Table 2 List of the Italian colour terms with the highest Cognitive Salience Index (S).

Colour term	English gloss	F	F(%)	R _F	mP	R _{mP}	S	R _S
giallo	yellow	85	95.5	4.0	4.86	1	0.19656	1
rosso	red	86	96.6	2.0	4.99	2	0.19371	2
blu	dark blue	85	95.5	4.0	5.82	3	0.16400	3
verde	green	83	93.3	7.0	5.90	4	0.15797	4
nero	black	87	97.8	1.0	9.89	5	0.09889	5
arancione	orange	82	92.1	9.0	10.41	6	0.08847	6
viola	purple	85	95.5	4.0	11.07	8	0.08627	7
rosa	pink	83	93.3	7.0	11.19	9	0.08332	8
bianco	white	83	93.3	7.0	12.14	10	0.07679	9
azzurro	medium blue	63	70.8	14.0	10.78	7	0.06568	10
grigio	grey	77	86.5	11.0	14.17	13	0.06106	11
marrone	brown	81	91.0	10.0	15.09	18	0.06033	12
celeste	light/sky blue	65	73.0	13.0	12.52	11	0.05832	13
fucsia	fuchsia	73	82.0	12.0	14.45	14	0.05675	14
lilla	lilac	62	69.7	15.0	14.71	15	0.04736	15
verde acqua	sea-green	53	59.6	17.0	15.68	20	0.03798	16
ocra	ochre	42	47.2	21.5	14.88	17	0.03171	17
magenta	magenta	42	47.2	21.5	15.21	19	0.03102	18
beige	beige	54	60.7	16.0	20.48	27	0.02962	19
bordeaux	claret	49	55.1	18.0	18.98	25	0.02901	20
indaco	indigo	40	44.9	23.0	19.00	26	0.02365	21
oro	gold	45	50.6	20.0	21.73	31	0.02326	22
argento	silver	47	52.8	19.0	22.74	32	0.02322	23
porpora	cardinal red	36	40.4	24.0	18.17	24	0.02227	24
turchese	turquoise	27	30.3	25.0	16.74	22	0.01812	25
verde scuro	dark green	23	25.8	27.0	14.78	16	0.01748	26
amaranto	amaranth	25	28.1	26.0	17.64	23	0.01592	27
verde chiaro	light green	21	23.6	28.0	16.57	21	0.01424	28
ciano	cyan	15	16.9	32.0	14.07	12	0.01198	29
violetto	violet	20	22.5	30.0	20.50	28	0.01096	30
giallo canarino	canary yellow	20	22.5	30.0	21.10	29	0.01065	31
blu notte	night <i>blu</i>	20	22.5	30.0	21.20	30	0.01060	32

Terms (N_{HS} = 32) with S > 0.01 ranked (RS) in descending order. Each term is also accompanied by its absolute frequency (F), relative frequency (F%), mean position (mP) in the list, and the corresponding ranks based on frequency (R_F) and mean position (R_{mP}).

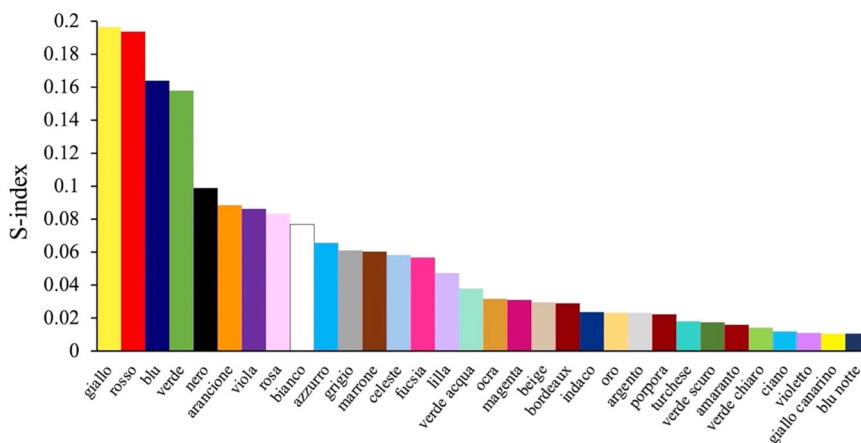


Fig. 3 Cognitive salience of colour terms. Colour terms (rendered in the corresponding colour) with the highest Cognitive Salience Index, S > 0.01 (N_{HS} = 32) ranked in order of decreasing S.

other non-BCTs. We also remark that the lower end of the cognitive salience function (Fig. 3) includes elaborated forms of three of the most salient BCTs, namely, two *verde* ‘green’ terms with modifiers related to lightness, *scuro* ‘dark’ and *chiaro* ‘light’, and two “noun-clad adjectives”, *giallo canarino* ‘canary yellow’ and *blu notte* ‘night *blu*’, that specify the basic colour by the shade of a nominal referent, i.e. the two CT-modification forms most frequently used in Italian (Grossmann and D’Achille, 2019).

In addition to Sutrop’s (2001) measure of cognitive salience S, we followed Bimler and Uusküla’s (2021) logarithm of both F and mP estimates: when values are plotted in decreasing order against successively less salient items, S tends to decline in a roughly exponential manner. The log-transformation enhances the linearity of the distribution, helping to identify separate sub-distributions within it. Figure 4 shows a scatterplot of the most salient Italian colour terms (N_{HS} = 32) representing the

“priority” (“earliness”) of the CT in the elicited list as a function of frequency (“prevalence”) of each term. The horizontal axis, Frequency, is $-\log(p_i)$, where $p_i = F/89$ (the denominator is the total number of lists here). The vertical axis, Priority, is $-\log(mP_i)$ and represents the term’s mean position, in descending order. Note that mP_i was calculated for colour terms ($N_{FCTS} = 96$) listed by at least five participants, as indicated in Table 1.

In Fig. 4, the term’s frequency, or “prevalence”, p_i decreases from left to right, while its “priority” decreases from the top downward. At the top left, one finds Italian BCTs that first come to mind when one is asked to “think of colour names”. From there, other terms follow a steep linear segment: while frequency drops off only slightly, mP_i decreases at a greater pace. After that initial decline, the sequence of terms meets an “elbow,” a transition to a second linear segment, less steep because the progressive drop-off in frequency is greater. Toward the right, the distribution of terms flares out. Notably, in so expressed cognitive

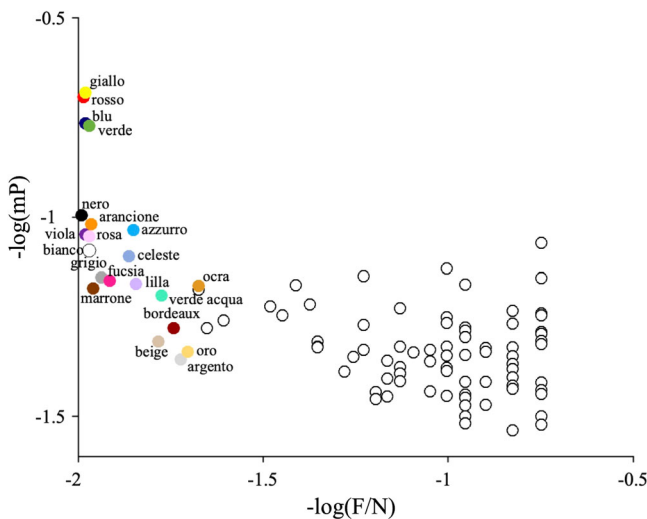


Fig. 4 Relationship between mean position and frequency of salient colour terms. Scatterplot of the most salient Italian colour terms ($N_{HS} = 32$) of Tuscan speakers: x-axis, $-\log(p_i)$, represents the term’s frequency, in descending order; y-axis, $-\log(mP_i)$, represents the term’s mean position, in descending order. Points for the 13 BCTs and 8 most frequent non-BCTs are coloured in correspondence with the term; unfilled circles indicate the remaining frequent non-BCTs listed in Table 2.

salience function, the observed “elbow” is around *fucsia* which takes a borderline position between the 13 BCTs, abutting to the least basic *grigio* and *marrone*, and the most frequent non-BCTs (*lilla*, *verde acqua*, *ocra*, *bordeaux* and *beige*).

The semantic map of the elicited colour terms. For analysis of the conceptual closeness of the elicited terms, we included only CTs with a Cognitive Salience Index $S > 0.01$ ($N_{HS} = 2$; Table 2, Fig. 5), whose association measure $ADJ(i,j)$ could be estimated with confidence. Using MDS, we computed a semantic map, which represents inter-term adjacencies, as a 2D solution (Fig. 5a). Its Stress value, 0.11312 (or 11%), is beyond the 10% cut-off of a “satisfactory” solution (Kruskal, 1964). We retained, however, the 2D solution for ease of interpretation and display, and, as well, to enable comparison with a 2D semantic map of Italian CTs obtained previously (Uusküla and Bimler, 2016).

Spatial representation of the 2D semantic map enables the researcher to recognise clusters of associated colour names that tend to prime each other and to appear in the lists in close succession. The horizontal axis D1 can be identified as a gradient of cognitive salience. It ranges from primary chromatic BCTs at the left to a tier of non-BCTs at the right, including the marginal cases of *argento* ‘silver’ and *oro* ‘gold’. Between these extremes lie secondary BCTs, frequent non-BCTs and modified terms.

The vertical axis D2 appears to reflect the gradient in chromatic content—lightness or desaturation as characteristic qualities of denoted colours. This distinction (dark–light, vivid–unsaturated) is, however, observed not throughout the whole semantic map but within individual clusters, implying that in sub-lists the consecutive appearance of closely linked terms may follow the opposite (semantic) direction of the achromatic gradient (cf. Bimler and Uusküla, 2018).

A tree diagram (dendrogram), the outcome of the HCA, is plotted in Fig. 5b. The clusters at the highest and intermediate agglomerative levels are superimposed upon the spatial model (Fig. 5a) by enclosing the clustered items within loops. Note that the dendrogram presents a $k = 8$ solution; in our choice of k we were leaning upon the number of clusters in the HCA solution for Italian in Uusküla and Bimler (2016). These clusters highlight the ‘chunking’ of terms, which tend to emerge as self-contained sub-lists within the listing sequence. Clustering, however, cannot show the parallelism of the internal structure and the relationships between chunks.

To improve “goodness-of-fit” of the semantic map, we computed a 3D solution (Fig. 6). Compared to the 2D map, its Stress = 0.06854 (7%) is a noticeable improvement, falling into

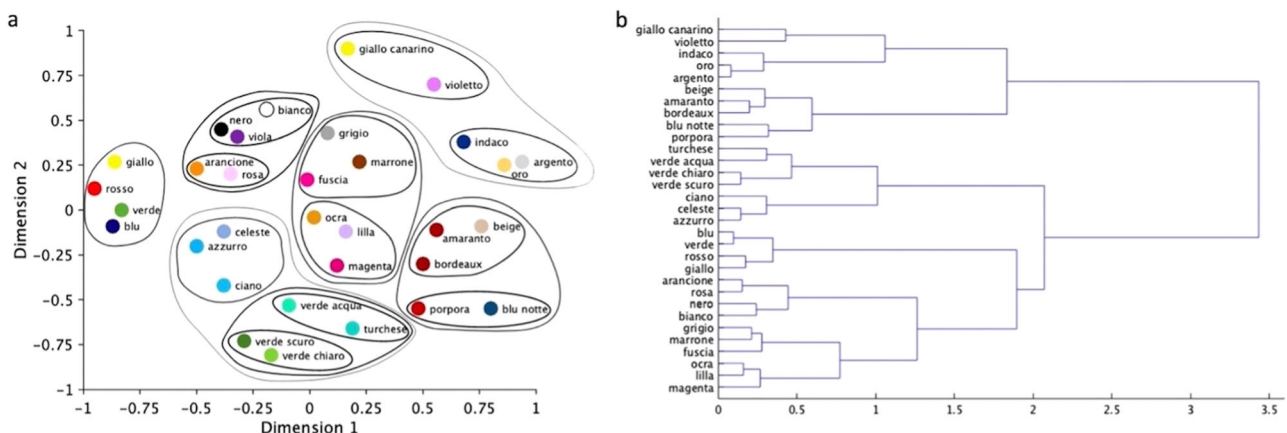


Fig. 5 Semantic map (2D solution) and dendrogram of salient colour terms. Representation of inter-term adjacencies for the Italian colour terms with the highest salience indices, $S > 0.01$ ($N_{HS} = 32$). **a** Semantic map (2D solution) with embedded loops reflecting the dendrogram (**b**).

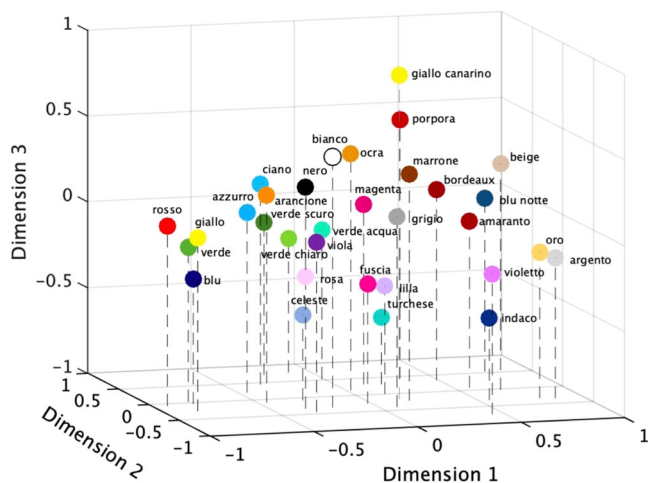


Fig. 6 Semantic map (3D solution) of salient colour terms. Semantic map (3D solution) of inter-term adjacencies for the Italian colour terms with the highest salience indices, $S > 0.01$ ($N_{H5} = 32$).

Kruskal's (1964) "satisfactory" solution category. In the 3D map, D1 reflects, too, the cognitive salience gradient; the distinction in chromatic content is also apparent but represented by D3.

D2 seems to be a convolution of the term's salience and the "economy" of its linguistic form: it distinguishes BCTs and salient non-BCTs (*fucsia*, *oro*, *argento*) (negative coordinates) from short (two-phoneme) non-BCTs (e.g. *beige*, *lilla*, *ocra*), whose D2-values are around zero; in comparison, on the positive D2 semi-axis, three-phoneme non-BCTs (e.g. *ciano*, *turchese*, *magenta*) have lower values, while compound terms have higher values (*verde acqua*, *verde scuro*, *blu notte*).

We comment on some conceptual relationships among the elicited terms revealed by the 3D map. For Tuscan speakers of Italian, who possess three basic 'blue' terms (Del Viva et al., 2022), the most inclusive is the primary term *blu*, while *azzurro* 'medium blue' and *celeste* 'light blue' are both mapped close to the secondary basic terms *viola* 'purple' and *rosa* 'pink'; also, *celeste* has stronger connotations of lightness and desaturation than *azzurro*.

Further, note that D1 is best interpreted as the 'priority' aspect of salience, rather than as 'frequency': although offered relatively infrequently, a term can still lie to the left of D1, if those respondents who listed it, did so among high-salience terms, early in their lists. This can be illustrated by the non-BCT *ciano* 'cyan' that was associated with and gravitated towards the two basic 'blue' terms, *azzurro* and *celeste*.

Another tendency is exemplified by the cluster of derived terms grouped around the common source BCT *verde* 'green' which all appear to be listed early; coalescing with these is the non-BCT *turchese* 'turquoise', which is located within a sector of qualified greens as its closest neighbours. We observe that, in Italian, 'blue' and 'green' are particularly generative, with their derived forms characterised by either the chromatic content, or hue 'nuance', or admixture of other colours. In the semantic map, we also observe relatively small distances between pairs of CTs, whose best exemplars, in a colour-naming experiment, were found to be closely located in the colour space (Albertazzi and Da Pos, 2017), such as *rosa-fucsia*, *celeste-turchese*, *beige-oro*.

It is noteworthy that the modified or noun-clad terms are typically listed in chunks pointing to a systematic attempt to exhaust all variants of (say) 'green' before moving on to (say) 'blue'. One more aspect to note is the clustering of terms that denote colours with subtle differences and which are, in effect,

alternative forms for a specific shade, like *amaranto*, *bordeaux* and *porpora* in this data. If participants listed only one or the other form, but in similar contexts, this resulted in adjacent points in the MDS solution.

The distinction in the gradient of chromatic content (D3) prevails within the cluster of the cardinal-hue BCTs, as well as within clusters of secondary BCTs and of frequently occurring non-BCTs. We note that *viola* 'purple' appears next to *nero* 'black' as if conceptualised by its darkness; *ocra* 'ochre' and *magenta* were listed with the BCTs *marrone* 'brown' and *grigio* 'grey' apparently associated with these by achromatic content. *Beige*, salient in the Italian listing data, as in other languages (Eessalu and Uusküla, 2013), has a high position on D3, comparable to that of *bianco*, *nero*, *ocra* and *marrone* suggesting that participants focussed on the achromatic connotations of the concept. Note also that in some clusters the direction of the gradient is inverted, i.e. the terms denoting darker shades have more positive D3-coordinates (e.g. *azzurro-celeste* or *verde scuro-verde chiaro*).

Discussion

The reported colour term inventory was obtained for young Tuscan speakers exposed to dialects spoken in and near Florence. We are aware that the results are confined by a regional (Tuscan) boundary and an age boundary. Within these boundaries, our results can be contrasted and compared with equivalent data from the colour-term inventories of other Italian regions and/or other generations. In this context, we mention the findings of Wieling et al. (2014, p. 674) on dialect levelling, specifically, that "Tuscan dialects overlap most closely with standard Italian in the area around Florence", and that younger, urban and higher-educated speakers—a cohort recruited for the present study—use lexical forms more likely to match the standard language.

The richness of the Italian colour-term inventory. We start by addressing the factors that are considered to drive the development of the colour-term inventory in a language. As surmised by Berlin and Kay (1969/1991), the fine-grainedness of a colour naming system is associated with societal complexity. Furthermore, more recent evidence has been obtained which shows that the emergence of fine-grained meanings corresponding to colours and shades of lightness is driven by the degree of interest in colour in the local culture and, hence, by the need (or otherwise) for efficient communication about it (Kemp et al., 2018).

In the present study, among Tuscan speakers, we observed a great variety and richness of colour terms. Elicited within 5 min, the total number of unique colour names, 337, is high. Note that it is comparable to the 310 unique terms elicited in Florence (Tuscany) with no time constraint (Uusküla and Bimler, 2016).

Also informative are numbers relating to the elicited lists: in the present study, the number of items varied between 15–59 in individual lists, with an average list containing 30.16 colour terms. It is worth noting that under no time constraint, Italian participants produced on average 19.86 terms (Uusküla and Bimler, 2016, p. 61). The lower productivity measure in the latter study is likely to reflect the fact that participants' oral responses were recorded by the experimenter, as opposed to, in the present study, the participants themselves typing the colour names.

Since the free-list length is a crude measure of "cultural competence" in the explored domain (Borgatti, 1998), the high productivity measures in the present study indicate the chromonymic richness of the (Tuscan) Italian language, comprising a great variety of colour names, which implies both an awareness of perceived shades and their nuanced lexicalisation. One can think of at least two factors behind the recorded

chromonymic richness. One is the variety of colour terms in Classical Latin, from which were inherited, in modified forms, Modern Italian *nero, rosso, verde, giallo, ceruleo, celeste*, etc. (Kristol, 1980). Moreover, due to cross-cultural contacts in the Middle Ages and later history, Italian adopted colour terms from Old High German (e.g. *bianco, grigio*); from French, in particular in the 17th–18th centuries (*azzurro, blu, arancione, marrone, bordeaux* etc.) (André, 1949) and, more recently, from English (e.g. *beige, cian, salmone*).

Furthermore, nuanced colour naming, to a certain degree can be explained by the linguistic means of lexical refinement available in Italian, in particular, (i) a great variety of suffixes expressing attenuative, approximate and/or evaluative meanings of the denoted hue, such as *-astro, -ette, -iato, -iccio, -iere, -ino, -igno, -ognolo* (whose meanings would be translated by English *-ish* or French *-âtre*); intensifying (*-one*), and elative (*-issimo*) suffixes; (ii) derivatives, such as denominal and deverbal adjectives; (iii) a variety of achromatic modifiers and object-qualifiers; (iv) complex compounds, with two or more colour terms; and (v) multiword expressions with a higher degree of expressiveness (Grossmann and D’Achille, 2019). In contemporary Italian, compounding is the most productive means of codifying different values of hue, lightness and saturation (Grossmann and D’Achille, 2019).

The (i)–(iv) morphological devices for enlarging the inventory of colour terms in Italian can be illustrated by examples from Tables 1, S2 and Fig. 1. Examples of suffixed terms are *biancastro* ‘whitish’, *giallo aranciato* ‘yellow orangish’, *rosino* ‘pinkish’, *bluette* ‘medium-bluish’. Furthermore, we observe great productivity and variety of compounds, in particular, with denominal derivatives and object-qualifiers. Specifically, the compound *verde aqua* ‘sea-green’ ($R_S = 16$) is among the most salient non-BCTs; broadly used (basic) terms with object-qualifiers are *giallo canarino* ‘canary yellow’ ($R_S = 31$), *verde petrolio* ‘petrol green’ ($R_S = 40$) and *rosso mattone* ‘brick red’ ($R_S = 49$). Among the examples of expressive compounds are *grigio fumo di Londra* ‘London-smog-grey’, *grigio canna di fucile* ‘gun-barrel grey’ and *rosso Ferrari* ‘Ferrari red’ (although such cases were offered in low numbers).

The terms that were named by two or more participants, 171, we consider conventional and culturally salient, following Smith et al. (1995). Among the culturally salient terms there are both lexically simplex and complex examples; among the simplex non-BCTs are, e.g., *oro* ‘gold’, *bordeaux* ‘claret’, *amaranto* ‘amaranth’, *violetto* ‘violet’ (an intensifying form of *viola* ‘purple’), *pesca* ‘peach’; examples of lexically complex non-BCTs are *rosa antico* ‘antique pink’, *verde smeraldo* ‘emerald green’, *terra di Siena* ‘sienna’. Note though that both categories are semantically complex (transparent), as defined by Smith et al. (1995), since they have real-world-object referents.

The conventional terms are complemented by novel terms mentioned by a few respondents or only once, “unique words ... that individuals invoke to meet particular needs and circumstances” (Casson, 1994, p. 8). The full list (Table S2) of such shows that some novel terms are “culturally-idiosyncratic” [e.g. *viola vinaccia* ‘wine-coloured purple’; *verde bandiera* ‘(Italian) flag green’; *rosso pompeiano* ‘Pompeian red’]. Furthermore, among the recurring novel terms are compounds containing the qualifiers ‘fluorescent’ or ‘phosphorescent’. Some other terms reveal the influence of Anglophone culture being Italian calques of English CTs frequently used in fashion to convey nuances in tonality, lightness or saturation, such as *blu royal* ‘royal blue’; *rosa shocking* ‘shocking pink’, *verde militare* ‘military green’ (Arcangeli, 2020). Also, it is hardly surprising that among our young Italian respondents, the influence of fashion lingo on CTs is revealed by the names of cosmetic products or cosmetic brand

names [e.g. *rosso cremino* ‘Cremino (a brand-name) red’; *rouge noir* (branded nail polish colour)].

We cannot but highlight a great variety of ‘blue’ terms and those straddling the boundary of the BLUE and GREEN categories, which point to the sea and sky as highly salient natural referents (Grossmann, 1988; Uusküla et al., 2016). Among the conventional terms are, e.g., *indaco* ‘indigo’, *turchese* ‘turquoise’, *blu elettrico* ‘electric blue’, *carta da zucchero* ‘powder blue’, *ceruleo* ‘cerulean’. Among novel terms are, e.g., *blu bimbo* ‘baby blue’, *rosso magenta* ‘magenta red’; *celeste metallico* ‘metallic celeste’.

In the elicited lists we also observe extensive use of various achromatic modifiers (cf. Grossmann and D’Achille, 2019): along with the most frequent *chiaro* ‘light’ and *scuro* ‘dark’, the participants occasionally offered *pastello* ‘pastel’, *opaco* ‘opaque’, *vivo* ‘vivid’, *intense* ‘intense’, *sbiadito* ‘bleached’ and *tenuo* ‘pale’. Among the relatively novel inclusions were modifiers, such as *metallico* ‘metallic’, *accesso* ‘bright (lit.) excessive’, *fluo* ‘fluorescent’, and *fosforescente* ‘phosphorescent’, that denote vibrant (high-reflectance) colours.

It is worth noting that “the number of lightness terms within a language represents an optimal solution to the problem of describing the variation in reflectances encountered within the visual environment” in order to achieve effective communication (Baddeley and Attewell, 2009, p. 1105). These authors found that the majority of languages possess just three basic lightness terms —‘light’, ‘dark’ and ‘grey’. In comparison, fewer languages, possess a lightness inventory of five basic terms, including ‘light grey’ and ‘dark grey’, to reflect the distribution of reflectances in the natural world. Contemporary (Tuscan) Italian, with cognitively salient *grigio chiaro* ‘light grey’ ($R_S = 43$) and *grigio scuro* ‘dark grey’ ($R_S = 50$), is likely to be one of those five-term lightness languages.

Finally, the list of the elicited terms (Table S2), when compared with the terms produced in a colour-naming experiment in Verona (Paggetti et al., 2016), prompts another observation of diatopic differences between the Tuscany and Veneto regions; namely, that the Tuscan and Verona samples differ in their inventories of elaborated colour terms, whereby some lexically complex terms are specific to one participant sample but were not offered at all by the other, e.g. *viola fiorentina* (Florence) vs. *celeste colore a spirito* (Verona).

Italian basic colour terms (for Tuscan speakers). Cognitive salience indices, S (Table 2, Figs. 3, 4), provide evidence that for Tuscan speakers the BCT inventory is augmented to 13: along with the primary basic *blu* ‘dark blue’ ($R_S = 3$), it includes two secondary basic ‘blue’ terms *azzurro* ($R_S = 10$) and *celeste* ($R_S = 13$), in accord with the previous finding in a psycholinguistic study on the naming of the blue area (Del Viva et al., 2022). This finding confirms Kay and McDaniel’s (1978, p. 641) conjecture that the development of BCT lexicons may extend beyond 11 terms, although rather than being “a theoretical inevitability”, the augmentation is dependent on specific language history, as demonstrated by Baronchelli et al.’s (2015) computational model, or, in other words, additional BCTs can be “culturally basic” (cf. Paramei, 2005).

The enrichment of the Italian basic colour lexicon denoting the blue area is apparently driven by salient environmental features such as the presence of the sea and the visibility of a blue sky (Giacalone Ramat, 1967; Jossierand et al., 2021; Kristol, 1979; Philip, 2003; Uusküla, 2014). For Tuscan speakers, the emergence of the “triple blues” is likely to result from another factor underscored by Jossierand et al. (2021)—by cultural complexity in this region’s society that stimulated the linguistic refinement of

the blue area. Specifically, in the Middle Ages, a complex woad dyeing technology was developed in the woollen cloth industry, as attested by the Florentine treatise *Trattato d'Arte della Lana* (1419; cit. in Cardon, 1992). It meticulously regulated the qualities of a dyed cloth: the dyed samples, arranged from the darkest to the lightest shades of blue, significantly varied in their price and had the corresponding names, among these *azzurini*, *cilestrini per Roma* ('for Roman taste'), *cilestrini al modo nostro* ('in our [Florence] fashion').

We also remark on the ranking of other Italian BCTs. Among the five primary BCTs, S-rankings for *rosso* 'red' ($R_S = 2$) and *nero* 'black' ($R_S = 5$) are similar to those in Hungarian (Uusküla and Sutrop, 2007) and Serbian (Jakovljević and Zdravković, 2018; Krimer-Gaborović and Jakovljević, 2022). However, *bianco* 'white' ($R_S = 9$) is lower in salience than that estimated for other languages, typically ranking 1–6 (e.g. Corbett and Davies, 1995; Hippisley et al., 2008). This deviation reflects a relatively late recollection of *bianco* in the elicited lists, $mP = 12.35$, compared to other languages [cf. $mP = 7.74$ for Estonian (Sutrop, 2001); $mP = 6.35$ for Hungarian (Uusküla and Sutrop, 2007); $mP = 6.21$ for Serbian (Jakovljević and Zdravković, 2018); or $mP = 8.44$ for Castilian Spanish (Xu et al., 2023)].

We observe that in the listing data collected in Florence earlier (Uusküla and Bimler, 2016) *bianco* ranked 5; in comparison, in the Verona psycholinguistic study *bianco* ranked 10 (Paggetti et al., 2016). It is noteworthy that three Italian samples in the compared studies varied with regard to the participant age range: 11–80 years (mean 39) in Uusküla and Bimler (2016); 20–37 years (mean 24) in Paggetti et al. (2016), and 20–31 years (mean 23) in the present one. One could speculate that *bianco* is less salient for young Italian speakers, whose environment of abundant colourful artefacts may cause them to consider that 'white' is not a colour term *stricto sensu*.

Another "anomaly" in the present data is the high frequency of *viola* ($R_F = 4$; $R_S = 7$), comparable to the frequency ranks of the primary BCTs. This indicates the cultural significance of purple for Tuscan speakers: since 1929 it has been the colour of the banner and paraphernalia of the Florence football club (*viola Fiorentina*; Prizio and Signoria, 2017).

Furthermore, the present outcomes for *fucsia* (Table 2, Figs. 1–4) deserve closer exploration of the term's linguistic features. In particular, we observe (Table 2) that the cognitive salience of *fucsia* ($S = 0.05675$; $R_S = 14$) is only slightly lower than that of *celeste* ($S = 0.05832$; $R_S = 13$). In frequency, *fucsia* ($R_F = 12$) overtakes both *celeste* ($R_F = 13$) and *azzurro* ($R_F = 14$), as illustrated by Fig. 1. In the Zipf-function, it is part of the second limb comprising lower frequency secondary BCTs and frequent non-BCTs (Fig. 2). With regard to "earliness" of listing, the mean position of *fucsia* ($R_{mP} = 13$) is higher than that of *marrone* 'brown' ($R_{mP} = 18$).

The borderline position of *fucsia*, between the BCTs and frequent non-BCTs, is also apparent when assessed using a modified salience index that combines logarithms of F (prevalence) and mP (priority) measures (Fig. 4): in the function "elbow" that separates BCTs and non-BCTs, *fucsia* gravitates to the least-salient BCTs *grigio* and *marrone* but is also close to the next most frequent non-BCTs, *lilla*, *verde acqua*, *beige*, *ocra* and *bordeaux*, with slightly lower salience. In the 2D semantic map (Fig. 5a), *fucsia* clusters with *grigio* and *marrone*, which suggests that its status is approaching basicness.

Notably, in the data elicited in Florence earlier (Uusküla and Bimler, 2016), the frequency of *fucsia* ($R_F = 14$) was similar to that in the present study. In the log-transformed cognitive salience (Bimler and Uusküla, 2021), *fucsia* takes a midway position between basic *grigio* and non-BCTs *lilla*, *beige* and *ocra*. In the colour-naming task with speakers of Verona (Veneto

region), according to its frequency and naming consistency *fucsia* ranked 11 ahead of the BCTs *grigio* and *nero* (Paggetti et al., 2016). It is also worth noting that *fucsia* is known to very young Italians: in a colour-naming task almost half of children as young as 3–6 years (in Cremona, Lombardy) offered this term; moreover, according to its frequency *fucsia* ranked 13 after the 12 Italian BCTs (Maccalli and Rizzi, 2009). In the colour space, the denotative meaning of *fucsia* fills in the "denotative gap" between *rosa* and *viola*, whereby the best exemplar of *fucsia* is close to that of the BCT *rosa*, and, also, to those of the non-BCTs *porpora* and *magenta* (Albertazzi and Da Pos, 2017).

The relatively high *fucsia* cognitive salience is probably the consequence of frequent use in the field of textiles and fashion (Arcangeli, 2020; Marellò and Onesti, 2016). We gather that under the influence of fashion 'parlance', the term's counterpart in English, *fuchsia*, is ranked (relatively) high for US speakers, too, with $R_F = 12$ in a listing task (Taft and Sivik, 1997) and $R_F = 25$ in colour-term usage, when participants assigned labels to Munsell coloured swatches (Lindsey and Brown, 2014). Assessed by a composite 'index of basicness', for British English speakers *fuchsia* ranked 8 among the most frequent monolexic non-BCTs (Mylonas and MacDonald, 2016).

The cognitive salience of *fucsia* in Italian, as observed in the present study, as well as psycholinguistic measures reported earlier, i.e. the consistency of naming samples and the word's distinct denotative "niche" (insertion) between the PINK and PURPLE BCCs, are likely to indicate that it is an emerging BCT. It is possible though that this "fancy" term is more frequent in the younger generation's parlance (cf. Biggam, 2012), and that its status in the colour inventory is characteristic of younger Italian speakers who have more advanced BCT systems than older speakers (cf. Kay, 1975). Indirectly this is supported by the observation that in a colour-naming task, for Russian speakers aged 20–49, the frequency of *fuksiä*-usage ranked 15–17, whereas for those aged 50 and over the term frequency ranked 24–29 (Griber et al., 2021).

A semantic map of (Tuscan) Italian colour terms. Figures 5a and 6 present the semantic maps, obtained using MDS, that represent adjacency, or co-occurrences of colour terms in the elicited lists. Although we interpret list adjacency in terms of 'similarity', this is a conceptual, not a perceptual connection, and reflects the pattern of associations and inter-relationships among the terms. Dimensions of the MDS maps reflect conceptual themes at a higher level of abstraction; in addition, clusters of mutually associated terms can be recognised, and listed in one another's company.

Such clusters may be related by CT semantic similarity and contiguity, as well as by the number of features the two items share (Uusküla and Bimler, 2016). Also, CTs are often related as opposites, i.e. in a contrastive way, with antonymous pairs suggested by the intuitive arrangement of terms for the opponent primary colours (Conklin, 1973). Cultural associations and collocations are also likely to contribute to the mutual priming of colour terms (Ronga et al., 2014).

Closely associated words tend to prime each other and to appear in the lists in close succession. Here the model is that listing a term involves activating the corresponding node in the participant's semantic network, causing the activated node to prime adjacent nodes in the network, thereby making it more likely that they will also be listed.

The 3D semantic map appears to reflect three competing criteria by which terms were sequenced. In particular, Italian speakers tended to follow a salience gradient of colour terms (D1): after listing the cardinal-hue primaries, respondents tended

to name secondary BCTs, frequent non-BCTs and ‘popular’ derived terms. Furthermore, the linguistic “economy” of the term, reflected by D2, appears to matter, too: following the BCTs, listed earlier are salient short hyponyms; listing segues to longer entrenched non-BCTs and derived terms, with lower popularity. This observation calls to mind the Brevity Law formulated by Zipf (1965; cit. in Durbin (1972, p. 270), namely, that lengthier words tend to have lower frequency. Finally, D3 indicates that speakers also made separate clusters of fully chromatic concepts—colour terms *stricto sensu*—and of unsaturated or desaturated concepts defined primarily by lightness rather than by hue. Furthermore, within the clusters differing in the level of cognitive salience, speakers pursue the chromatic content theme.

Frequent Italian non-BCTs as indicators of the lexical refinement of the colour space. Apart from the BCTs, the present data show a rich lexicon of hyponyms—indicators of the need for the communication of perceived colour nuances (cf. Gibson et al., 2017; Zaslavsky et al., 2019, 2022). We conjecture that this communicative need is stipulated by the opulent environmental “colour diet” of Italian speakers (cf. Komarova and Jameson, 2008), as well as by their exposure to an artistic culture with profuse employment of different pigments (Ronga and Bazzanella, 2015) and copious variation of colour in artefacts (cf. Jossierand et al., 2021), including textile manufacture, fashion and advertising.

In some detail, we consider three colour space areas, mostly at the intersections of traditional BCCs, that were shown to be particularly prone to further lexical differentiation of non-BCTs (Kerttula, 2002). In addition, the exploration of Italian high-frequency non-BCTs was motivated by recent findings of emerging BCTs in several languages, e.g. *teal*, *peach*, *lavender* and *maroon* in American English (Lindsey and Brown, 2014); *lilac* and *turquoise* in British English (Mylonas and MacDonald, 2016), and *sirenevyyj* ‘lilac’ and *birúzovyyj* ‘turquoise’ in Russian (Griber et al., 2021; Paramei et al., 2018b).

The BLUE-GREEN area. As addressed above, lexical refinement is productive in the blue area and at the BLUE-GREEN category boundary, with the frequently offered hyponyms *indaco* ‘indigo’ ($R_S = 21$), *turchese* ‘turquoise’ ($R_S = 25$) and *ciano* ‘cyan’ ($R_S = 29$). It is noteworthy that *indaco* ($R_S = 17$) and *turchese* ($R_S = 19$) were among the 30 most frequent non-BCTs elicited in Florence in 2014 from speakers aged 11–80 (mean age 39); however, *ciano* was not part of that CT group (Uusküla and Bimler, 2016).

This reminds one of Biggam’s (2012) observation that non-BCTs frequently occurring in the younger generation’s parlance are “fancy” object-derived loanwords with colour references transparent for this generation’s speakers, as was the case in the present study, but not necessarily transparent for older speakers. While estimating focal colours of Italian CTs in a blue-green quadrant of the colour circle, in their choice of non-BCTs, *turchese* and *ciano*, Albertazzi and Da Pos (2017) probably assumed a ‘novel’ colour vocabulary on the part of their participants (university students).

Apart from inter-generational differences, in our previous study (Paramei et al., 2018b) we recorded a diatopic variation in the choice of ‘blue’-hyponyms: in a blue-area naming experiment, *indaco* ($R_F = 13$) and *turchese* ($R_F = 19$) were among the 40 most frequent terms for speakers of Alghero (the north-west coast of Sardinia) but not for speakers of Verona (the Veneto region) (Paramei et al., 2018a). The frequent appearance of *indaco* and *turchese* in the present data hints at the “coastal” variation of Tuscan speakers’ vocabulary. The finding is in accord with Regier

et al., (2016) conclusion that the colour lexicon reflects an interaction of local cultural communicative needs with environmental factors.

The observed Italian lexical refinement of the BLUE and BLUE-GREEN areas concurs with similar elicitation-task findings in other languages, in which ‘turquoise’ terms occur particularly frequently, e.g. *turkos* in Swedish and *seledynowy* in Polish, *turquoise* in American English (Taft and Sivik, 1997), *türkiz* in Hungarian (Uusküla and Sutrop, 2007), *turquesa* in Castilian Spanish (Xu et al., 2023), and *birjuzovyyj* in Russian (Uusküla and Bimler, 2016). Also, colour-naming studies indicate frequent use of various hyponyms conveying certain blue shades, e.g. *teal* in American English (Lindsey and Brown, 2014), *turquoise* in British English (Mylonas and MacDonald, 2016; Sturges and Whitfield, 1997), *kon* ‘indigo’ in Japanese (Kuriki et al., 2017), or *turquesa* in Spanish (Xu et al., 2023).

The RED-PINK-PURPLE area. The elicitation task reveals that refinement by hyponyms of the areas denoted by the BCTs *rosso* ‘red’, *viola* ‘purple’ and *rosa* ‘pink’ is lexically productive. In particular, apart from *fucsia*, the relatively high-frequency ranks of *lilla* ‘lilac’ ($R_S = 15$) and *violetto* ‘violet’ ($R_S = 30$) are probably due to their perceptual proximity and, thus, semantic association with *viola*, the symbolic colour of Florence. Other frequently listed hyponyms are *magenta* ($R_S = 18$), *bordeaux* ‘claret’ ($R_S = 20$), *porpora* ‘cardinal red’ ($R_S = 24$) and *amaranto* ‘amaranth’ ($R_S = 27$). *Lilla*, *bordeaux*, *amaranto*, *porpora* and *violetto* were found, too, among the 30 most salient terms elicited from Florence speakers (Uusküla and Bimler, 2016). In the colour-naming experiment carried out in Verona (Paggetti et al., 2016), *lilla*, *bordeaux*, *vinaccia*, *magenta*, *amaranto* and *violetto* ranked between 15 and 22 on consistency, corroborating the suggestion that this area of the colour space is highly refined for Italian speakers, including beyond Tuscany.

The Italian lexical refinement of this area concurs, too, with recurring findings in elicitation task results for ‘lilac’-counterparts in other languages, e.g. *lila* in Swedish (Taft and Sivik, 1997); *lila* in Serbian (Jakovljević and Zdravković, 2018); *lila* in Spanish and *danzi* in Mandarin Chinese (Xu et al., 2023); *lilla* in Finnish, Latvian, and *sirenevyyj* in Russian (Uusküla and Bimler, 2016). The frequently listed Italian *bordeaux* ‘claret/burgundy’, the French calque (Claidière et al., 2008), also finds counterparts in other languages, e.g. *bordo* in Hungarian (Uusküla and Sutrop, 2007), Turkish (Uusküla and Bimler, 2016) and Serbian (Jakovljević and Zdravković, 2018); *bordovyyj* in Russian (Corbett and Morgan, 1988), Ukrainian and Belarusian (Hippisley et al., 2008); *burdeos* in Spanish (Xu et al., 2023).

The lexical prominence of *magenta* in Italian demonstrated here is similar to that in elicited lists in other languages, such as American English (Lindsey and Brown, 2014; Taft and Sivik, 1997), and Spanish (Uusküla and Bimler, 2016; Xu et al., 2023). Italian *porpora* ‘cardinal red’, in comparison, has a “high grade of semantic specialisation”: frequently it is used in an ecclesiastic context or in relation to church positions (Marello and Onesti, 2016, p. 100). In this context it is similar to its counterparts in Russian, *purpurnyyj* (Davies and Corbett, 1994; Paramei et al., 2018b), Estonian *purpur*, Latvian *purpurs*, Lithuanian *purpurine* (Uusküla and Bimler, 2016), and Serbian *purpurna* (Jakovljević and Zdravković, 2018).

In colour-naming studies in other languages ‘magenta’ and counterparts of ‘lilac’, ‘violet’, ‘claret/burgundy’ or ‘cardinal red’ are also frequently offered (e.g. American English: Lindsey and Brown, 2014; British English: Mylonas and MacDonald, 2016; Sturges and Whitfield, 1997; Russian: Griber et al., 2021; Paramei et al., 2018b; Spanish: Xu et al., 2023; Mandarin Chinese: Xu et al., 2023).

The *YELLOW* area periphery. In the present data hyponyms that denote colours at the “hard-to-name” fringes of the *YELLOW* category are relatively high-salience—*ocra* ‘ochre’ and *beige*, where this BCC borders with the *RED* and *PINK* categories. In line with the Levinson (2000) emergence hypothesis, in Modern Italian (and many other languages) a need arises to label colours in this less-salient region of the colour space since it lacks a BCT.

Ocra denotes shades of dull red tending to reddish-brown, at the borders with the *ORANGE*, *RED* and *BROWN* categories. The salience of *ocra* ($R_S = 17$) can be explained by the prominence of ochre in the “colour diet” of Italy: ochre-derived pigments have widespread utilisation, in multiple contexts—for painting house walls, ceramic articles, and to tan leather (Masset, 1980). We observe that ‘ochre’-counterparts, produced in elicitation and colour-naming tasks, are also salient in some other languages, e.g. *ochre* in British English (Mylonas and MacDonald, 2016); *okker* in Hungarian, *okers* in Latvian (Uusküla and Bimler, 2016), and *oker* in Serbian (Jakovljević and Zdravković, 2018).

Unsurprisingly, *beige* ($R_S = 19$) is also a relatively salient non-BCT. It is prominent among elicited Italian colour names found in Uusküla and Bimler’s (2016) study. Its denotation estimated in colour-naming tasks is rather vague—of pastel colours in the hard-to-name *YELLOW-BROWN-WHITE-RED-ORANGE* region (Albertazzi and Da Pos, 2017; Paggetti et al., 2016). The cognitive salience of Italian *beige* is similar to that of its cognates in other languages, e.g. in Estonian (Sutrop, 2001); American English, Polish and Swedish (Taft and Sivik, 1997); Hungarian (Uusküla and Sutrop, 2007); French (Claidière et al., 2008); Czech, Finnish, Latvian, Russian, Turkish (Uusküla and Bimler, 2016); Serbian (Jakovljević and Zdravković (2018); and Spanish (Lillo et al., 2018; Xu et al., 2023). The denotation of Italian *beige* is similar to those of its cognates in other languages, e.g. American English (Lindsey and Brown, 2014; Taft and Sivik, 1997); British English (Mylonas and MacDonald, 2016; Sturges and Whitfield, 1997); and many other European languages. In their cross-language psycholinguistic study of ‘beige’, Eessalu and Uusküla (2013) concluded that ‘beige’ is associated with a skin-tone colour and denotes a light yellowish-brownish colour but also includes pinkish and orangish nuances.

We remark that *beige* rose to prominence in the 1980s, in English and with cognate terms in many European languages under cultural-economic pressure connected with prestigious brands in fashion (clothing, shoes, leather products) and advertising. Arcangeli (2020) assumes that the term was instigated by the iconic beige fashion products of Burberry, while Paramei and Bimler (2021) suggest that it emerged at a time when IBM desktop computers known as *beige boxes* became a standard office fixture.

Finally, among salient Italian non-BCTs in this area and, close to *beige* and *ocra* in the semantic map, are designations of the metallic sheens *oro* ‘gold’ ($R_S = 22$) and, by association with it, *argento* ‘silver’ ($R_S = 23$). The salience of these terms in Italian is unsurprising if one bears in mind the cultural and economic significance of both referents, and the beauty of artefacts—art and artisan objects—to which Tuscan speakers are exposed. The popularity of these terms’ cognates in many other languages, referred to above, is, too, explained by similar factors in social, economic and cultural development.

The limitations of the study and the possible future direction of Italian colour naming research. A few limitations of the present study can be considered. First, data collection was undertaken in one specific region, Tuscany. In view of the notable dialectal variability in Italy (De Renzo, 2009), it would be worth replicating it in other regions of the country—northern, central

and southern. The plausibility of this suggestion is supported by recent findings concerning the regional variability of Italian ‘blue’ terms (basic status and denotative meaning) in Alghero (Sardinia), Verona (the Veneto region) and Tuscany (Del Viva et al., 2022; Paramei et al., 2018a).

Second, although the listing task was demonstrated to provide reliable measures for distinguishing BCTs and non-BCTs (Corbett and Davies, 1997), the frequency and cognitive salience indices do not always indicate a sharp cut-off between these (cf. Figs. 1–3). In order to resolve ambiguities in the status of individual Italian colour terms, complementary measures can be used. In particular, Corbett and Davies (1997) found that certain linguistic measures are illuminating for this purpose, such as frequency in texts. Beyond analyses of individual Italian corpora undertaken earlier (e.g. D’Achille and Grossmann, 2013; Philip, 2003), modern methods of computational linguistics could provide more elaborate estimates and, also, encompass various Italian corpora. Furthermore, as demonstrated by Corbett and Davies (1997), behavioural measures obtained in a colour-naming task, such as consistency, consensus and response times are instructive. For Italians, consistency in naming colours was, indeed, shown in recent colour-naming studies to be informative (Del Viva et al., 2022; Paggetti et al., 2016).

Third, our sample was gender-unbalanced, including 68 women and 21 men. Numerous studies (e.g. Krimer-Gaborović and Jakovljević, 2022; Lindsey and Brown, 2014; Paramei et al., 2018b, to name just a few) have demonstrated gender/sex differences in colour vocabulary, whereby women were shown to offer a greater variety of non-basic, in particular “fancy” colour terms than men. Future studies will need to examine the outcome of a balanced sample to ascertain possible gender differences in the inventory of Italian colour terms.

Finally, apart from diatopic variations of the colour lexicon in different regions of Italy, it is likely that one can find inter-generational differences between young, middle-aged and mature Italians (cf. recent findings for Russian by Griber et al., 2021; for Galician by Teixeira Moláns, in preparation).

Conclusions

In the present study, we employed the elicitation task to explore the Italian colour term inventory, with data collected in Tuscany. We are aware of notable dialectal variability in Italy (De Renzo, 2009), so the results presented here for young speakers of the Tuscan dialect are now available for comparison with the colour-term inventories of speakers from other Italian regions.

For ascertaining the Italian colour lexicon, we undertook a comprehensive analysis of individual terms’ frequency, the “earliness” of their listing (their mean position in the list), and their cognitive salience. We then reconstructed a semantic map of highly salient CTs (terms that are conceptually associated and appear in the lists in close succession). Our aims were to identify (Tuscan) Italian BCTs and, beyond these, high-frequency non-BCTs, some of which may be emerging as (culturally) basic. Furthermore, we were interested in determining the linguistic patterns of compound and modified expressions, which serve to improve communication efficiency, as well as to identify areas of colour space that undergo lexical refinement in Italian.

Based on the measures of elicitation productivity, we found that 10 universal Berlin and Kay BCCs, bar ‘blue’, have counterparts in (Tuscan) Italian. The present results have provided further evidence that three ‘blue’ terms are basic for Tuscan speakers—*blu*, *azzurro* and *celest*, confirming the “triple blues” identified in Del Viva et al.’s (2022) psycholinguistic study in Florence. Along with the augmented inventory of 13 BCTs, *fucsia*, a high-frequency non-BCT, is conceivably emerging as a basic

term, at least for young speakers of this region. This “core” colour inventory is extended by the great variety of other offered names, comprising abundant hyponyms, derived (suffixed and modified) forms, and compounded expressions that all indicate (Tuscan) Italian speakers’ “cultural competence” in the colour domain (cf. Hays et al., 1972; Josserand et al., 2021) and their need to communicate nuanced information about colour shades (cf. Gibson et al., 2017; Zaslavsky et al., 2019, 2022).

An exploration of the colour lexicon in diverse Italian dialects would be worth undertaking: beyond a purely academic interest in regional colour vocabularies, the evaluation of the denotative meanings of identical terms is essential, since within-language diversity would seem likely to impede the accurate communication of colour among speakers (cf. Brown and Lindsey, 2023).

The present findings of the processes of lexical refinement in the (Tuscan) Italian colour lexicon indicate that such processes are similar to those also observed in other modern languages, whose colour terminology is evolving, whereby a few terms, conventionally considered to be non-BCTs, become more salient and frequently used. In particular, *lilac* and *turquoise*, counterparts of Italian *lilla* and *verde acqua*, are argued to augment the British English BCT inventory (Mylonas and MacDonald, 2016); a similar tendency is observed in American English with regard to *lavender* and *teal* (Lindsey and Brown, 2014); in Russian, with the counterparts *sirenevyyj* and *birizovyyj* (Griber et al., 2021), and in Spanish, with *lila* and *turquesa* (Xu et al., 2023).

Among other non-BCTs that increase in frequency and salience across languages, including Italian, are two in the YELLOW-BROWN-WHITE area of the colour space—*beige* ‘beige’ (Claidière et al., 2008; Eessalu and Uusküla, 2013; Lindsey and Brown, 2014; Mylonas and MacDonald, 2016; Sturges and Whitfield, 1997; Taft and Sivik, 1997) and *ocra* ‘ochre’ (Jakovljević and Zdravković, 2018; Mylonas and MacDonald, 2016; Uusküla and Bimler, 2016). These are complemented by terms for the metallic sheens *oro* ‘gold’ and *argento* ‘silver’ that exhibit a high-frequency figure in all those languages whose colour vocabularies have so far been investigated.

High-frequency Italian terms: *fucsia*, *bordeaux*, *amaranto*, *magenta* and *porpora*, which lexically refine the purple area (Albertazzi and Da Pos, 2017; Paggetti et al., 2016), remind one of similar processes in American English (Lindsey and Brown, 2014), British English (Mylonas and MacDonald, 2016; Sturges and Whitfield, 1997), Russian (Griber et al., 2021; Paramei et al., 2018b), Castilian Spanish (Xu et al., 2023) and many other languages (Uusküla and Bimler, 2016).

The recorded cross-language evolution, with enrichment of the colour inventory, is apparently driven by the need for effective communication of perceived colour in the modern world with its increasing variety of coloured artefacts (cf. Gibson et al., 2017; Zaslavsky et al., 2019, 2022). Across the languages named above, including Italian, we observe that lexical differentiation of the colour space develops through each of the three processes (reviewed by Paramei, 2020; Paramei and Bimler, 2021): *category insertion* at the BLUE-GREEN category boundary (‘turquoise/teal’); *lexical partition* (fission) in the purple-red area (‘fuchsia’, ‘claret’, ‘magenta’ etc.); and *emergence* of a colour category in the “no man’s land” (‘beige’, ‘ochre’, ‘gold’, ‘silver’). As argued by Mylonas et al. (2022), critical for augmenting colour inventories is cultural transfer, i.e. colour names learned by individuals through interactions with other cultures, contexts and technological developments.

We conclude that the (Tuscan) Italian colour lexicon explored here can be considered a paradigmatic case of the close interaction of universal trends in the colour inventory, reflecting cognitive and perceptual biases, with culture-specific biases stipulated by the cultural history of Italian (cf. Baroncelli et al., 2015). Conceivably, the cross-language patterns in the colour inventory will further converge

as a result of the globalisation of travel, media and trade. As a consequence, due to language contacts, exposure to colour category distinctions and their lexicalisation will be increased in order to achieve cross-language communication efficiency (Xu et al., 2013).

Data availability

The datasets generated and analysed during the current study are available in the ZENODO repository: <https://doi.org/10.5281/zenodo.8177310>.

Received: 24 July 2023; Accepted: 14 November 2023;

Published online: 01 December 2023

References

- Albertazzi L, Da Pos O (2017) Color names, stimulus color, and their subjective links. *Color Res Appl* 42(1):89–101. <https://doi.org/10.1002/col.22034>
- André J (1949) Étude sur les termes de couleur dans la langue latine. Klincksieck (in French)
- Arcangeli M (2020) La fascinazione esotica nei colori della moda [The exotic fascination in the colours of fashion]. *Ling Cult Mediazione/Lang Cult Mediat* 7(2), 113–136 (in Italian). <https://www.ledonline.it/index.php/LCM-Journal/article/view/2301/1412>
- Baddeley R, Attewell D (2009) The relationship between language and the environment: information theory shows why we have only three lightness terms. *Psychol Sci* 20(9):1100–1107. <https://doi.org/10.1111/j.1467-9280.2009.02412.x>
- Baroncelli A, Loreto V, Puglisi A (2015) Individual biases, cultural evolution, and the statistical nature of language universals: the case of colour naming systems. *PLoS ONE* 10(5):e0125019. <https://doi.org/10.1371/journal.pone.0125019>
- Bazzanella C (2017) Riflessioni sui colori in italiano. Categorizzazione e varietà di forme [Colours in Italian: categorisation and variety of forms]. *Studi Gramm Ital* 36:181–200. Accademia della Crusca (in Italian)
- Berlin B, Kay P (1969/1991) Basic color terms: their universality and evolution. University of California Press
- Bernard HR (2006) Research methods in anthropology: qualitative and quantitative approaches, 4th edn. AltaMira Press
- Biggam CP (2012) The semantics of colour: a historical approach. Cambridge University Press
- Bimler D, Uusküla M (2014a) Clothed in triple blues”: sorting out the Italian blues. *J Opt Soc Am A* 31(4):A332–A340. <https://doi.org/10.1364/JOSAA.31.00A332>
- Bimler D, Uusküla M (2014b) Analysing listing data with multidimensional scaling to reconstruct the cognitive organization of the colour lexicon in 11 languages. In: Rossi M (ed) *Colour and colorimetry: multidisciplinary contributions*, vol XB. Maggioli Editore, pp. 179–190
- Bimler D, Uusküla M (2018) Individual variations in color-concept space replicate across languages. *J Opt Soc Am A* 35(4):B184–B191. <https://doi.org/10.1364/JOSAA.35.00B184>
- Bimler D, Uusküla M (2021) Two indices are better than one: building on Robbins, Nolan, and Chen (2017). *Field Methods* 33(1):42–51. <https://doi.org/10.1177/1525822X20945063>
- Bolton R (1978) Black, white, and red all over: the riddle of color term salience. *Ethnology* 17(3):287–311. <https://doi.org/10.2307/3773198>
- Borgatti SP (1998) Elicitation techniques for cultural domain analysis. In: Schensul JJ, LeCompte MD (eds) *The ethnographer’s toolkit*, vol 3. Altimira Press, pp. 1–26
- Borgatti SK (1999) Elicitation techniques for cultural domain analysis. In: Schensul JJ, LeCompte MD, Nastasi BK, Borgatti SK (eds) *Enhanced ethnographic methods: audiovisual techniques, focused group interviews, and elicitation techniques*. Altamira Press, pp. 115–151
- Brown AM, Isse A, Linsey DT (2016) The color lexicon of the Somali language. *J Vis* 16(5):14. <https://doi.org/10.1167/16.5.14>
- Brown AM, Lindsey DT (2023) The color communication game. *Sci Rep* 13:16006. <https://doi.org/10.1038/s41598-023-42834-3>
- Cardon D (1992) New information on the medieval woad vat. *Dyes in history and archaeology*, No. 10. Textile Research Associates, pp. 22–31
- Casson RW (1994) Russett, rose, and raspberry: the development of English secondary color terms. *J Linguist Anthropol* 4(1):5–22. <https://www.jstor.org/stable/43102893>
- Claidière N, Jaissati Y, Chevallier C (2008) A colour sorting task reveals the limits of the universalist/relativist dichotomy: colour categories can be both language specific and perceptual. *J Cogn Cult* 8(3):211–233. <https://doi.org/10.1163/156853708X358155>

- Conklin HC (1973) Color categorization. *Am Anthropol* 75(4):931–942. <https://www.jstor.org/stable/673098>
- Corbett GG, Davies IRL (1995) Linguistic and behavioural measure for ranking basic colour terms. *Stud Language* 19(2):301–357. <https://doi.org/10.1075/sl.19.2.02cor>
- Corbett GG, Davies IRL (1997) Establishing basic color terms: measures and techniques. In: Hardin CL, Maffi L (eds) *Color categories in thought and language*. Cambridge University Press, pp. 197–223
- Corbett GG, Morgan G (1988) Colour terms in Russian: reflections of typological constraints in a single language. *J Linguist* 24(1):31–64. <https://doi.org/10.1017/S002226700011555>
- D’Achille P, Grossmann M (2013) I composti <colorati> in Italian tra passato e presente [The “coloured” compounds in Italian between past and present]. In: E. Casanova Herrero E, Calvo Rigual C (eds) *Actas del XXVI Congreso Internacional de Lingüística i Filología Románicas*. De Gruyter (in Italian), pp. 523–537
- Davies IRL, Corbett GG (1995) A practical field method for identifying probable basic colour terms. *Lang World* 9:25–36
- Davies I, Corbett GG (1994) The basic color terms of Russian. *Linguistics* 32:65–89. <https://doi.org/10.1515/ling.1994.32.1.65>
- De Mauro T (1983) *Storia linguistica dell’Italia unita* [Linguistic history of the unified Italy]. Laterza & Figli (in Italian)
- De Renzo F (2009) Per un’analisi della situazione sociolinguistica dell’Italia contemporanea. *Italiano, dialetti e altre lingue* [About an analysis of sociolinguistic situation in contemporary Italy. Italian, dialects and other languages]. *Italia* 85, 44–62 (in Italian)
- Del Viva MM, Mariani I, De Caro C, Paramei GV (2022) Florence “blues” are clothed in triple basic terms. *i-Perception* 13(5). <https://doi.org/10.1177/20416695221124964>
- Durbin, M. (1972). Basic terms – off color? *Semiotica*, 6(3), 257–278
- Eessalu M, Uusküla M (2013) The special case of *beige*: a cross-linguistic study. In Rossi M (ed) *Colour and colorimetry: multidisciplinary contributions*, vol. IX. B. Maggioli Editore, pp. 168–176
- Ferrer i Cancho R, Solé RV (2003) Least effort and the origins of scaling in human language. *Proc Natl Acad Sci USA* 100(3):788–791. <https://doi.org/10.1073/pnas.0335980100>
- Friendly ML (1977) In search of the M-Gram: the structure of organization in free recall. *Cogn Psychol* 9(2):188–249. [https://doi.org/10.1016/0010-0285\(77\)90008-1](https://doi.org/10.1016/0010-0285(77)90008-1)
- Giacalone Ramat A (1967) Colori germanici nel mondo romanzo [Germanic colour names in the Romance world]. *Atti e Memorie dell’Accademia Toscana di Scienze e Lettere “La Colombaria”*, Vol 32, pp. 107–211 (in Italian)
- Gibson E, Futrell R, Jara-Ettinger J, Mahowald K, Bergen L, Ratnasingam S, Gibson M, Piantadosi ST, Conway BR (2017) Color naming across languages reflects color use. *Proc Natl Acad Sci USA* 114(40):10785–10790. <https://doi.org/10.1073/pnas.1619666114>
- Griber YA, Mylonas D, Paramei GV (2021) Intergenerational differences in Russian color naming in the globalized era: linguistic analysis. *Humanit Soc Sci Commun* 8:262. <https://doi.org/10.1057/s41599-021-00943-2>
- Grossmann M (1988) *Colori e lessico: studi sulla struttura semantica degli aggettivi di colore in catalano, castigliano, italiano, romeno, latino ed ungherese* [Colours and lexicon: studies on semantic structure of colour adjectives in Catalan, Castilian, Italian, Romanian, Latin and Hungarian]. In: *Tübinger Beiträge zur Linguistik*, vol 310. Gunter Narr Verlag (in Italian)
- Grossmann M, D’Achille P (2019) Compound color terms in Italian. In: Raffaelli I, Katunar D, Kerovec B (eds) *Lexicalization patterns in color naming: a cross-linguistic perspective*. John Benjamins, pp. 61–79
- Hays DG, Margolis E, Naroll R, Perkins DR (1972) Color term salience. *Am Anthropol* 74(5):1107–1121. <https://www.jstor.org/stable/672967>
- Hippisley A, Davies I, Corbett GG (2008) The basic colour terms of Lower Sorbian and Upper Sorbian and their typological relevance. *Stud Language* 32(1):56–92. <https://doi.org/10.1075/sl.32.1.04hip>
- Jakovljević I, Zdravković S (2018) The colour lexicon of the Serbian language—a study of dark blue and dark red colour categories. Part 1: Colour-term elicitation task. *Psihologija* 51(2):197–213. <https://doi.org/10.2298/PSN160521002J>
- Josserand M, Meeussen E, Majid A, Dediu D (2021) Environment and culture shape both the colour lexicon and the genetics of colour perception. *Sci Rep* 11:19095. <https://doi.org/10.1038/s41598-021-98550-3>
- Kay P (1975) Synchronic variability and diachronic change in basic color terms. *Language* 50(3):257–270. <https://doi.org/10.1017/S0047404500006667>
- Kay P, McDaniel CK (1978) The linguistic significance of the meanings of basic color terms. *Language* 54(3):610–646. <https://doi.org/10.1353/lan.1978.0035>
- Kemp C, Xu Y, Regier T (2018) Semantic typology and efficient communication. *Annu Rev Linguist* 4:109–128. <https://doi.org/10.1146/annurev-linguistics-011817-045406>
- Kerttula S (2002) English colour terms: Etymology, chronology, and relative basicness. In: *Mémoires de la Société Néophilologique de Helsinki*, vol LX. Société Néophilologique
- Komarova NL, Jameson KA (2008) Population heterogeneity and color stimulus heterogeneity in agent-based color categorization. *J Theor Biol* 253(4):680–700. <https://doi.org/10.1016/j.jtbi.2008.03.030>
- Krimer-Gaborović S, Jakovljević I (2022) [A study of gender-related differences in colour naming in Standard Serbian language]. *Zb Matice Srp Filol Linguist* 65(1):47–75. (in Serbian)1081485/ms_zmsfil.2022.65.1.3
- Kristol AM (1979) Il colore azzurro nei dialetti italiani [The colour azzurro in Italian dialects]. *Vox Romanica* 38:85–99. (in Italian)
- Kristol AM (1980) Color systems in Southern Italy: a case of regression. *Language* 56(1):139–145. <https://doi.org/10.2307/412646>
- Kruskal JB (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29(1):1–27. <https://doi.org/10.1007/BF02289565>
- Kuriki I, Lange R, Muto Y, Brown AM, Fukuda K, Tokunaga R, Lindsey DT, Uchikawa K, Shioiri S (2017) The modern Japanese color lexicon. *J Vision* 17(3):1. <https://doi.org/10.1167/17.3.1>
- Levinson SC (2000) Yéli Dnye and the theory of basic color terms. *J Linguist Anthropol* 10(1):3–55. <https://doi.org/10.1525/jlin.2000.10.1.3>
- Lillo J, Prado-León L, González Perilli F, Melnikova A, Álvaro L, Collado J, Moreira H (2018) Spanish basic colour categories are 11 or 12 depending on the dialect. In: MacDonald LW, Biggam CP, Paramei GV (eds) *Progress in colour studies: cognition, language and beyond*. John Benjamins, pp. 59–82
- Lindsey DT, Brown AM (2014) The color lexicon of American English. *J Vis* 14(2):17. <https://doi.org/10.1167/14.2.17>
- Lindsey DT, Brown AM (2021) Lexical color categories. *Annu Rev Vis Sci* 7:605–631. <https://doi.org/10.1146/annurev-vision-093019-112420>
- Maccalli E, Rizzi A (2009) Color naming in età prescolare: un esperimento. In: Prestileo F, Rizzi A (eds) *Colore i colorimetria: contributi multidisciplinari*, Palermo, 7–9 October 2009, pp. 53–63 (in Italian)
- Marello C, Onesti C (2016) Colourful microstructures: how Italian dictionaries colour terms. In: Silvestre JP, Cardeira E, Villalva A (eds) *Colour and colour naming: crosslinguistic approaches*. Centro de Linguística da Universidade de Lisboa—Universidade de Aveiro, pp. 91–104
- Masset C (1980) Red ochre and human evolution: a case for discussion. *Current Anthropol* 21(5):638–639. <https://doi.org/10.1086/202541>
- Mitzenmacher M (2003) A brief history of generative models for power law and lognormal distributions. *Internet Math* 1(2):226–251. <https://doi.org/10.1080/15427951.2004.10129088>
- Morgan G, Corbett GG (1989) Russian colour term salience. *Russ Linguist* 13(2):125–141. <https://doi.org/10.1007/BF02551669>
- Mylonas D, Caparos S, Davidoff J (2022) Augmenting a colour lexicon. *Humanit Soc Sci Commun* 9:29. <https://doi.org/10.1057/s41599-022-01045-3>
- Mylonas D, MacDonald L (2016) Augmenting basic colour terms in English. *Color Res Appl* 41(1):32–42. <https://doi.org/10.1002/col.21944>
- Paggetti, G., & Menegaz, G. (2012). Is light blue (azzurro) color name universal in the Italian language? In K. N. Kutulakos (Ed.), *Trends and Topics in Computer Vision*, Vol. 6554 (pp. 90–103). Springer. https://doi.org/10.1007/978-3-642-35740-4_8
- Paggetti G, Menegaz G, Paramei GV (2016) Color naming in Italian language. *Color Res Appl* 41(4):402–415. <https://doi.org/10.1002/col.21953>
- Paramei GV (2005) Singing the Russian blues: an argument for culturally basic color terms. *Cross-Cult Res* 39(1):10–38. <https://doi.org/10.1177/1069397104267888>
- Paramei GV (2020) Color categorization: patterns and mechanisms of evolution. In: Shamey R (ed) *Encyclopedia of color science and technology*, 2nd edn. Springer
- Paramei GV, Birmler D (2021) Language and psychology. In: Steinvall A, Street S (eds) *A cultural history of colour*, vol 6. The modern age: from 1920 to present. Bloomsbury, pp. 117–134
- Paramei GV, D’Orsi M, Menegaz G (2014) ‘Italian blues’: a challenge to the universal inventory of basic colour terms. *J Int Colour Assoc* 13:27–35. <http://www.aic-colour-journal.org/>
- Paramei GV, D’Orsi M, Menegaz G (2018a) Diatopic variation in the referential meaning of the “Italian blues”. In: MacDonald LW, Biggam CP, Paramei GV (eds) *Progress in colour studies: cognition, language and beyond*, Ch. 5. John Benjamins, pp. 83–106
- Paramei GV, Griber YA, Mylonas D (2018b) An online color naming experiment in Russian using Munsell color samples. *Color Res Appl* 43(3):358–374. <https://doi.org/10.1002/col.22190>
- Philip GS (2003) *Collocation and connotation: a corpus-based investigation of colour words in English and Italian*. Unpublished PhD thesis, University of Birmingham
- Prizio S, Signoria L (2017) *Il romanzo della grande Fiorentina. Dal 1926 a oggi la storia del mito viola* [The novel of the great Fiorentina. From 1926 to today the history of the purple myth]. Newton Compton Publisher (in Italian)
- Quinlan MB (2019) The freelisting method. In: Liamputtong P (ed) *Handbook of research methods in health social sciences*. Springer, pp. 1431–1446

- Regier T, Carstensen A, Kemp C (2016) Languages support efficient communication about the environment: words for snow revisited. *PLoS ONE* 11(4):e0151138. <https://doi.org/10.1371/journal.pone.0151138>
- Regier T, Kay P, Khetarpal N (2007) Color naming reflects optimal partitions of color space. *Proc Natl Acad Sci USA* 104(4):1436–1441. <https://doi.org/10.1073/pnas.0610341104>
- Robbins MC, Nolan JM, Chen D (2017) An improved measure of cognitive salience in free listing tasks: a Marshallese example. *Field Methods* 29(4):395–403. <https://doi.org/10.1177/1525822X1727672>
- Ronga I, Bazzanella C (2015) The emergence of color categories: variance and invariance. *CEUR Workshop Proc* 1419:222–227. <http://ceur-ws.org/Vol-1419/paper0033.pdf>
- Ronga I, Bazzanella C, Strudsholm E, Salvati L (2014) Black as night or as a chimney sweep? Color words and typical exemplars. *Intercult Pragmat* 11(4):485–520. <https://doi.org/10.1515/ip-2014-0022>
- Sandford JL (2015) A cognitive linguistic usage perspective: what is Italian *Blu, azzurro, celeste*? Do English agree on BLUE semantics? *Cult Sci Colore—Color Cult Sci* 4:22–30
- Saunders BAC, van Brakel J (1997) Are there nontrivial constraints on colour categorization? *Behav Brain Sci* 20(2):167–228. <https://doi.org/10.1017/S0140525X97531426>
- Smith JJ, Furbee L, Maynard K, Quick S, Ross L (1995) Salience counts: a domain analysis of English color terms. *J Linguist Anthropol* 5(2):203–216
- Società Linguistica Italiana (1995) *Dialetti e lingue nazionali* [National dialects and languages]. Bulzoni (in Italian)
- Stausberg M (2022) Free-listing. In: Engler S, Stausberg M (eds) *The Routledge handbook of research methods in the study of religion*, 2nd edn. Routledge, pp. 554–564
- Sturges J, Whitfield TWA (1997) Salient features of Munsell colour space as a function of monolexic naming and response latencies. *Vis Res* 37(3):307–313. [https://doi.org/10.1016/S0042-6989\(96\)00170-8](https://doi.org/10.1016/S0042-6989(96)00170-8)
- Sutrop U (2001) List task and a cognitive salience index. *Field Methods* 13(3):263–276. <https://doi.org/10.1177/1525822X0101300303>
- Taft C, Sivik L (1997) Salient color terms in four languages. *Scand J Psychol* 38(1):29–34. <https://doi.org/10.1111/1467-9450.00005>
- Teixeira Moláns P. Galician colour semantics: an investigation of basic colour terms. PhD Thesis, The University of Glasgow, UK (in preparation)
- Uusküla M (2007) The basic colour terms of Finnish. *SKY. J Linguist* 20:367–397. <http://www.linguistics.fi/julkaisut/sky2007.shtml>
- Uusküla M (2014) Linguistic categorization of BLUE in Standard Italian. In: Anderson W, Biggam CP, Hough C, Kay CJ (eds) *Colour studies. A broad spectrum*. John Benjamins, pp. 67–78
- Uusküla M, Bimler D (2016) From listing data to semantic maps: cross-linguistic commonalities in cognitive representation of colour. *Folklore* 64:57–90. <http://folklore.ee/Folklore/vol64/colour.pdf>
- Uusküla M, Sutrop U (2007) Preliminary study of basic colour terms in modern Hungarian. *Linguist Ural* 43(2):102–122
- Vincent JM (1983) Categorizzazione e strategie di denominazione dei colori: aspetti metodologici e problemi relativi all'inglese e all'italiano [Categorisation and colour naming strategy: Methodological aspects and problems related to English and Italian]. In: Grupp di Lecce (ed) *Linguistica e antropologia*. Atto del XIV Congresso Internazionale della SLI, Roma. Bulzoni, pp. 161–171 (in Italian)
- Ward Jr. JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58(301):236–244. <https://doi.org/10.1080/01621459.1963.10500845>
- Weller SC (2014) Structured interviewing and questionnaire construction. In Bernard HR, Gravlee CC (eds) *Handbook of methods in cultural anthropology*, 2nd edn. Rowman & Littlefield Publishers, pp. 343–390
- Weller SC, Romney AK (1988) *Systematic data collection*. Sage Publications
- Wieling M, Nerbonne J, Montemagni S, Baayen RH (2014) Lexical differences between Tuscan dialects and Standard Italian: accounting for geographic and sociodemographic variation using additive mixed modelling. *Language* 90(3):669–692. <https://www.jstor.org/stable/24672042>
- Xu J, Dowman M, Griffith TL (2013) Cultural transmission results in convergence towards colour term universals. *Proc R Soc B* 280(1758):20123073. <https://doi.org/10.1098/rspb.2012.3073>
- Xu M, Zhu J, Benítez-Burraco A (2023) A comparison of basic color terms in Mandarin and Spanish. *Color Res Appl* 48(6):709–720. <https://doi.org/10.1002/col.22863>
- Zaslavsky N, Garvin K, Kemp C, Tishby N, Regier T (2022). The evolution of color naming reflects pressure for efficiency: evidence from the recent past. *J Language Evol*. <https://doi.org/10.1093/jole/lzac001>
- Zaslavsky N, Kemp C, Tishby N, Regier T (2019) Color naming reflects both perceptual structure and communicative need. *Top Cogn Sci* 11(1):207–219. <https://doi.org/10.1111/tops.12395>

Acknowledgements

The project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme, Grant Agreement No. 832813 GenPercept "Spatio-temporal mechanisms of generative perception". The University of Florence supported GVP as a Visiting Professor. David Bimler's advice on the multidimensional scaling and proofreading of an original manuscript is gratefully acknowledged. We are indebted to Carole Biggam for valuable comments on and proofreading of the final version of the paper. The authors thank Francesca Casti and Alessia Spadoni for their help in data collection, and all participants for their time and goodwill.

Author contributions

MMDV: conceptualisation, funding acquisition, data collection, writing—original draft, writing—review and editing. SC: data collection, formal analysis, writing—original draft, writing—review and editing. GVP: conceptualization, project administration, formal analysis, writing—original draft, writing—review and editing.

Competing interests

The authors declare no competing interests.

Ethical approval

The study has been approved by the local ethics committee ("Commissione per l'Etica della Ricerca", University of Florence, July 7, 2020, No. 111) and was conducted according to the guidelines of the Declaration of Helsinki.

Informed consent

All participants have signed the informed consent.


Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-023-02393-4>.

Correspondence and requests for materials should be addressed to Maria Michela Del Viva.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023