

Research Article

Martin Alldrick*

Colour terms and bilingualism: An experimental study of Russian and Tatar

<https://doi.org/10.1515/opli-2025-0050>

received June 10, 2024; accepted April 3, 2025

Abstract: Research has shown that bilingual speakers of languages that differ in categorisation concepts experience mediation between the two languages. This phenomenon extends to colour categorisation. Previous research has focused on studying bilingual speakers from a migratory background. This study takes an alternative approach and looks at speakers who live in a bilingual community, namely Tatar–Russian bilingual speakers in Tatarstan, and employs an experimental methodology using the salience criterion. The data show that Tatar has 6 basic colour terms (in contrast to 12 basic colour terms in Russian). The results further reveal that whilst Tatar–Russian speakers are able to maintain a distinction between the colour categorisation systems of each language, they nevertheless experience interference in their categorisation of colour terms in Russian when compared to monolingual Russian speakers.

Keywords: bilingualism, categorical concept, colour, Russian, Tatar

1 Introduction

Previous research has shown that bilingual speakers of languages that differ in categorical concepts experience mediation between the two languages; that is to say speakers do not maintain a distinct categorical system for every language they speak. Much work has been conducted in the area of categories with studies showing cognitive shifts in the representation of colour in bilinguals (Ervin 1961, Caskey-Sirmons and Hickerson 1977, Andrews 1994, Athanasopoulos 2008, 2011, Athanasopoulos et al. 2010a, b). Many recent studies have largely focused on respondents who have acquired the L2 later in life and have moved from a speech community with a high degree of colour differentiation to a speech community with a lower degree of colour differentiation. As such, three variables remain understudied: the effect of bilingualism on colour categorisation in speakers who acquire both L1 and L2 as children; the effect of bilingualism on colour categorisation in speech communities where the two languages naturally co-exist (Ryabina 2011, Rätsep 2018); and the L2 influence of a language with a higher degree of colour differentiation on an L1 language of a lower level of colour differentiation.

This contribution to the discussion examines these three variables. It reports on an experimental study of speakers of Russian and Tatar and the effect of bilingualism on colour categorisation. It provides a description of the basic colour term inventory of Tatar and examines whether bilingual speakers are able to maintain separate basic colour term systems discretely in different languages. It looks at the basic colour term system at a global level and also with specific reference to the ‘blue’ part of the spectrum.

* **Corresponding author: Martin Alldrick**, Surrey Morphology Group, University of Surrey, Guildford, GU2 7XH, United Kingdom, e-mail: m.alldrick@surrey.ac.uk

1.1 Basic colour terms

1.1.1 Basic colour terms as a linguistic universal

The theory of basic colour terms, articulated in its original form by Berlin and Kay (1969/1991), takes a universalist position on the categorisation of colour terms in the lexicon and is based on an experimental survey of 98 different languages. Their findings show that the denotata of colour foci differ as much between speakers of the same language as between speakers of different languages. These findings also show that there is a universal and discrete way of dividing up the colour spectrum. This undermines the relativist hypothesis that a speaker's language objectively affects colour perception.

Central to their theory is the concept of the basic colour term. According to Berlin and Kay, such terms are monolexemic, not subsumed under the meaning of any other basic colour term, applicable to a wide range of objects, and psychologically salient. As such, basic colour terms can be understood to be the set of high-frequency colour terms from which the signification of all other colour terms is derived. Furthermore, Berlin and Kay (1969/1991) propose that the emergence of basic colour terms in a language is governed by universal constraints. Their study put forward seven stages of evolution from Stage I (languages with just two basic colour terms) to Stage VII (languages with eight to eleven basic colour terms). As illustrated in Figure 1, it is a necessary condition for a language to have a colour term for 'red' if it is to have a colour term for 'green'. Similarly, Berlin and Kay (1969/1991) argue in favour of a universal constraint of a maximum of 11 basic colour terms.

Further studies into the development of basic colour terms have suggested revisions to the original proposal articulated by Berlin and Kay (1969/1991). Kay and McDaniel (1978) argue in favour of a bifurcated approach that better represents the physiological realities of colour vision. They divide up the evolutionary sequence into two sections, which have been referred to as 'primary basic' (Stage I to Stage V) and 'secondary basic' (Stage VI and beyond) by Davies and Corbett (1994). The emergence of primary basic colour terms is attributed not to the successive encoding of colour foci, but to the successive differentiation between composite colour categories, whilst the emergence of secondary colour terms is caused by the encoding of foci located at the intersection of basic colour terms (Kay and McDaniel 1978). Such an analysis allows for the emergence of a greater number of basic colour terms than originally proposed by Berlin and Kay (1969/1991).

Debate has emerged over the extent that colour terminology is determined by universal constraints. Researchers such as Burgess et al. (1983) and Roberson et al. (2000) have observed that certain languages (in the instance of Tarahumara and Berinmo, respectively) do not fit the theoretical straitjacket initially proposed by Berlin and Kay (1969/1991). They argue that the colour categorisation system of a given language is not determined by universal constraints, but rather by the cultural experiences of the speakers of a given language (Roberson et al. 2000). Such a hypothesis has been termed relativist.

Further investigations (Kay and Regier 2003, Roberson 2004, Regier et al. 2007, Kay 2015, Lindsey and Brown 2021) have since arrived at a third position, commonly referred to as the weak relativist hypothesis or *conceptualism* (Douven and Paramei 2023), according to which there are certain minimal constraints that constrain the development of language terms in each language. These minimal constraints can be accounted for by the absence of arbitrary differences between the colour lexicons of different languages; nevertheless, there is sufficient variation between languages in order to argue that colour categorisation is not innate (Lindsey and Brown 2021). Such an approach acknowledges the role of cultural experiences in the development of colour terminology, subject to minimal guiding principles. Kay and Regier (2003) propose the notion of privileged points upon which languages tend to centre their basic colour term inventory. Kay (2015) adds that another one of these minimal constraints is a prohibition on the creation of idiosyncratic colour terms composed of parts of the colour spectrum separated by another part that already constitutes a basic colour term.

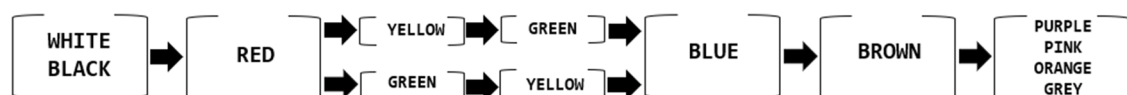


Figure 1: Universal evolution of colour terms according to Berlin and Kay (1969/1991).

1.1.2 ‘Blue’ as a basic colour term

One such intersection not accounted for in Berlin and Kay’s (1969/1991) analysis is the emergence of distinct terms for ‘dark blue’ and ‘light blue’. Though rare, such a distinction at the basic colour term level has been attested in a variety of unrelated languages (Paramei and Bimler 2021) including Russian, Greek, and Turkish. Russian has *sinij* (‘dark blue’) and *goluboj* (‘light blue’), Greek has *ble* (‘light blue’) and *ghalazio* (‘light blue’), and Turkish has *lacivert* (‘dark blue’) and *mavi* (‘light blue’) (Morgan and Corbett 1989 for Russian; Coventry et al. 2006 for Greek; Özgen and Davies 1998; Rätsep 2018 for Turkish). Research shows that this further categorical development is not an irregularity but rather an inherent part of the evolution of colour terms. Kertulla (2004) shows that the blue and purple parts of the colour spectrum are foremost candidates for further categorical and lexical differentiation. Similarly, Uusküla and Bimler (2016) found, in their cross-linguistic typological study of conceptual mapping of colour terms, that categorical and lexical splitting of the blue part of the spectrum is a regularity rather than an anomaly in the development of colour terms.

Russian has often been considered the prototypical example of a language with a split basic colour term system for ‘blue’. Although Berlin and Kay (1969/1991) tentatively defined *goluboj* (‘light blue’) as a subset of *sinij* (‘dark blue’), they allow for the distinct possibility of *goluboj* being a twelfth basic colour term in Russian. Further research (Morgan and Corbett 1989, Davies and Corbett 1994, Paramei 2004, 2007) has established the basic colour term status of both *goluboj* (‘light blue’) and *sinij* (‘dark blue’). The debate has emerged over the exact nature of the basic colour term system for ‘blue’ given that *sinij* (‘dark blue’) fulfils the non-derivational criterion for basicness established, whilst *goluboj* (‘light blue’) does not. However, given that Berlin and Kay (1969/1991) themselves argue that their criteria are merely diagnostic (sufficient) and not necessary, greater weight can be placed on the salience criteria as determined in experimental studies. Salience studies (Morgan and Corbett 1989, Davies and Corbett 1994) have shown that *goluboj* (‘light blue’) appears with the five most salient colour terms making it highly unlikely that native speakers do not use a split basic colour term system for ‘blue’. Such a conclusion is supported by further etymological and cultural evidence. Baxilina (1975) mentions that both terms had emerged very early on within the development of the Russian language, by at least the eleventh century. Similarly, Frumkina (1984) reports on Russian speakers being surprised by the fact that other languages did not make the obligatory *sinij/goluboj* distinction; so ingrained is this colour distinction in the minds of Russian speakers. This removes any ambiguity about the fact that Russia employs a split basic colour term system for ‘blue’, with neither *sinij* (‘dark blue’) nor *goluboj* (‘light blue’) being expressible in terms of the other.

The colour system of Turkish has been posited as an inversion of the split-‘blue’ system found in Russian. It has been reported that Turkish has two basic colour terms for ‘blue’ – *mavi* (‘light blue’) and *lacivert* (‘dark blue’) – with *mavi* (‘light blue’) being the more salient of the two terms. Unlike Russian where both basic colour terms for ‘blue’ are native, in Turkish both proposed basic colour terms for ‘blue’ are loanwords with *mavi* being borrowed from Arabic and *lacivert* being borrowed from Persian (Rätsep 2018). Özgen and Davies’ (1998) investigation of Turkish colour terms found there was a minimum of 11 basic colour terms with the possibility of an additional basic colour term for ‘dark blue’. Salience rankings placed *mavi* (‘light blue’) in second place and *lacivert* (‘dark blue’) in twelfth place. The ambiguous position of *lacivert* leaves open the debate whether Turkish behaves like Russian in terms of having a split basic colour term system for ‘blue’ or whether it has only a single basic colour term. In the event that Turkish does have a split basic colour term system, ‘dark blue’ rather than ‘light blue’ is considered the dominant colour term.

1.2 Colour and bilingualism

The notion that language mediates the interpretation and categorisation of objective reality has been termed the linguistic relativity hypothesis, sometimes also referred to as the Sapir–Whorf hypothesis (Athanasopoulos and Aveledo 2012). Such a theory has found an experimental underpinning in a variety of experimental and cognitive studies, including Lucy (1993), Levinson (1996), and Slobin (1996). Until recently, investigation of the Sapir–Whorf hypothesis has been focused on monolingual contexts. However, there has been growing interest

in exploring linguistic relativity from the perspective of bilingualism and multilingualism, given that it is in fact the bilingual (multilingual) speaker that represents the default human experience (Cook 2002). Given that the majority of people globally speak more than one language, the assumption of monolingualism should be seen as an exception prevalent in certain Western contexts rather than a universal tendency.

It has been argued that bilingual speakers have a distinct form of cognitive processing compared to monolingual speakers which is subject to a different set of constraints (Grosjean 1989, Hunt and Agnoli 1991). The nature of the bilingualism affecting cognition is mediated by several variables including specific language proficiency, general language proficiency, age of language acquisition, amount of language use, interactional setting, and length of stay in Lx speaking country/community (Athanasopoulos 2011). Since the turn of the twenty-first century, the scientific discourse has seen an increase in the number of studies investigating cognitive-linguistic processing in bilingual speakers and the implication this has on the linguistic relativity hypothesis.

A starting point to test the interaction between bilingualism and cognitive processing is phonological output. Amengual's (2016) study of Spanish–Catalan bilingual speakers identified elements of interference between the phonological systems of the two languages. This interference was especially prominent in those words which were cognate across both languages. Such a finding serves to support the viewpoint that cognitive processing in each language is not discrete but rather closely interconnected.

Similarly, Athanasopoulos' (2007) study investigated the influence of bilingualism on the disambiguation of categorical concepts. When examining whether Japanese-English bilinguals categorise mass and count nouns based on shape or material, he found that bilingual respondents occupy a middle ground between English monolinguals, who categorise by shape, and Japanese monolinguals, who categorise by material.

The studies by Amengual (2016) and Athanasopoulos (2007) serve to show the wider applicability of bilingualism influencing linguistic processing to the debate.

The role of bilingualism and language contact resulting in shifts in colour categorisation has long been attested (Brown and Lenneberg 1954, Ervin 1961, Caskey-Sirmons and Hickerson 1977).

The consequences of language contact on the development of colour categorisation systems can be categorised into two broad groups. First are changes to the entire colour categorisation system of a given language as a result of language contact. Second are changes that take place to the colour categorisation systems of bilingual speakers in contrast to monolingual speakers of both one and other languages. The following sections look at both of these categories.

1.2.1 Basic colour term categorisation and bilingualism

The notion that bilingual speakers of two languages with two distinct colour categorisation systems will adopt some sort of hybrid form between the two as a result of language mediation is well attested (Brown and Lenneberg 1954, Ervin 1961). Brown and Lenneberg's (1954) study on English-Zuni (a North American language isolate (Campbell 2000)) bilingualism focused on the 'yellow'–'orange' part of the spectrum, which is represented distinctly in English but as a single term in Zuni. The results show that the ability of Zuni speakers to distinguish between the two terms is correlated with the speaker's knowledge of English. Similarly, Ervin's (1961) study of Navajo shows that bilingual Navajo-English speakers undergo a semantic shift in their categorisation of the basic colour terms foci in comparison to their monolingual Navajo counterparts.

Language contact can also trigger wholesale changes within the colour classification system of a given language. Berlin and Kay's (1969/1991) description of changes in the colour classification system of Tzeltal serves as a prime example of this. Tzeltal was originally a Stage IV language with a unified term for both 'green' and 'blue'. Under the influence of Spanish, Tzeltal has borrowed the Spanish term for 'blue' *azul* into its lexicon, such that the two-colour terms are now distinct and Tzeltal is a Stage V language.

Such augmentation of the colour classification system has more recently also been witnessed occurring in real time. An investigation of Himba (Mylonas et al. 2022) examined changes to the colour classification system over a 15-year period from 2005 to 2020. Initial investigations of Himba in 2005 observed just five basic colour terms. By the time of further investigations in 2020, the number of basic colour terms had increased to seven

with the addition of terms for ‘blue’ and ‘brown’. Mylonas et al. (2022) attribute this reconfiguration of the basic colour term system of Himba to increased contact with outside communities.

1.2.2 ‘Blue’ in contact

Andrews (1994) provides one of the earliest studies on the effect of language immersion on basic colour term categorisation with his investigation of the split-‘blue’ basic colour term system amongst the Russian émigré community in the United States. Adult émigrés still made a distinction between *sinij* (‘dark blue’) and *goluboj* (‘light blue’) as equally salient basic colour terms distinct from one another. Young émigrés, however, had borrowed the English single-‘blue’ categorisation system, with *sinij* (‘dark blue’) becoming the overarching term for all shades of blue and *goluboj* becoming a specialised term for shades of ‘light blue’.

Andrews (1994) does not specify the age at which the members of the young émigré sample moved to the United States, merely stating that they did so in ‘childhood or adolescence’. As such, it can be assumed that most, if not all, of the sampled respondents, moved to an English-speaking environment after the age of 6 years, which is considered to be cut-off point for the development of relevant cognitive mechanisms and neural networks; bilingual speakers older than this will start to exhibit age of acquisition effects (Wattendorf and Festman 2008). These findings show that language immersion can result in a reconsolidation of a split-‘blue’ system to a single-‘blue’ system. The effect of age on the findings shows that the initial age of language immersion and the consequential age of acquisition effects do have an effect on which speakers will experience a shift in their usage of Russian colour terminology.

Studies of monolingual Greek speakers and bilingual Greek-English speakers living in the United Kingdom (Athanasopoulos 2008) show that bilingual speakers differ from monolingual speakers in the manner in which they categorise the colour space. As with Andrews (1994), there is a difference between the English and Greek categorisation of the colour space. Bilingual speakers experienced a shift in colour prototypes and alteration in the manner in which they judged perceptual similarity between different colours. The extent to which this was the case is correlated with the length of time the speaker spent acculturated in the second language country, in this case, the United Kingdom. As such, the results show that concepts of colour categorisation are not fixed, but rather flexible, and are subject to change as a result of both linguistic and cultural influence.

Such effects are further corroborated by studies of late-stage Italian-English bilinguals residing in the United Kingdom (Paramei et al. 2016, Douven and Paramei 2023). Similar to Greek, Italian maintains a basic colour term distinction between *blu* (‘dark blue’) and *azzurro* (‘light blue’), with a third, subsidiary, term *celeste* for very light shades of blue. Both studies (Paramei et al. 2016, Douven and Paramei 2023) identified bi-directional shifts in the colour categorisation of bilinguals indicating that both languages influence the cognitive processing of colour categorisation. One difference of note between Italian and Greek was found in the area of the perceived lightness of the English term *blue*. Italian-English bilinguals shifted their understanding of English *blue* to match their already existing notions of Italian *blu*. This contrasts with Greek where Greek-English bilinguals shifted their prototype of the Greek *ble* in the direction of the English *blue*. Paramei et al. (2016) explain this finding on the basis of homophony between English *blue* and Italian *blu*, which does occur in the English-Greek context.

The studies on both Greek and Italian point towards an integration of the bilingual colour lexicon. Nevertheless, both studies examine late-stage bilinguals functioning in a monolingual society. This limits the extent to which these studies can be extended to fully bilingual speech communities.

Ryabina’s (2011, 2014) studies of the basic colour term systems of Finno-Ugric languages (Komi and Udmurt) have shed much light on the role of language contact on the development of basic colour term systems. In both Komi and Udmurt, dialects of each language which have had longer prolonged contact with Russian have borrowed more substantively from its basic colour term structure. The findings show that the Russian split-basic colour term system for ‘blue’ has been borrowed in Northern Udmurt (Ryabina 2011) demonstrating that a split-‘blue’ basic colour term system can be acquired through language contact. The findings further show that Northern Udmurt has the lowest number of basic colour terms of the languages studied demonstrating that the split-basic colour term system for ‘blue’ can be acquired independently of the

development hierarchy (Figure 1). As such the development of a split-basic colour term system for ‘blue’ is not necessarily predicated on a system consisting of 12 basic colour terms (cf. Russian, Turkish, Greek, etc.). Ryabina (2014) further finds that languages can acquire other basic colour terms from split-‘blue’ basic colour term systems without acquiring the split-‘blue’ distinction itself. According to Ryabina (2014), Komi has historically been associated with Stage III of the Berlin and Kay’s (1969/1991) basic colour term system with a single term for both ‘yellow’ and ‘green’ (*viž*). Komi-Permyak, which has experienced prolonged contact with Russian, has adopted the Russian term *zelenyj* as a term for ‘green’, whilst Komi-Zyrian, which has experienced less contact with Russian, is currently in the process of developing its own native word for ‘green’ (*turunviž*).

1.3 Tatarstan as a bilingual language community

According to the 2010 census, 14.27% of the population of the Russian Federation is bilingual, with Tatar being the second most spoken language within the country with around 4.5 million speakers (3.15% of the population) (Rosstat 2010). Tatar is a Turkic language belonging to the Kipchak branch (Poppe 1965) and, together with Russian, is the official language of Tatarstan (Konstitucija 1991). As such, Tatar is particularly suitable for investigation, as it retains a critical mass of speakers and is unrelated to Russian. This reduces the effect of confounding variables such as language prestige and genetic links.

Tatarstan is a multi-ethnic autonomous republic within the Russian Federation. It has a population of almost 3.8 million people and Tatar is the first language of approximately 53% of the population (Rosstat 2010). There is a distinct divide between the Russian and Tatar communities in terms of their language preference, with 94% of Tatars in Tatarstan giving Tatar as their first language, whilst 99% of Russians gave Russian as their first language (Rosstat 2010). The nature of bilingualism in Tatarstan is fundamentally asymmetrical, with Tatar speakers also being fluent in Russian, and most native Russian speakers not speaking Tatar. This has been attributed to the observation that participation in Russian society “requires the ability to communicate in Russian” (Faller 2011, 14). Although Tatar language learning was made compulsory in Russian language medium schools in 1998, the effect it had on the Tatar language competency of Russian speakers was minimal (Wigglesworth-Baker 2015). Unlike many other minority languages within the Russian Federation, however, Tatar remains a working language within certain functional domains and is not reduced to the role of a symbolic official language (cf. Zamyatin 2014). As such, there are still substantive groups of people within the Tatar speech community for whom Tatar functionally remains a first language.

The current linguistic environment has a language community that has been brought up functionally bilingually from an early age and can be contrasted with a monolingual language community living within the same geographic area. This sample contrasts with other studies which look at bilingual speech communities where the second language was acquired in adulthood. Furthermore, previous studies have focused on speakers moving from a two-blue language system to a one-blue language system. The sociolinguistic situation in Tatarstan allows for the study of native speakers of a one-blue language system interacting within a two-blue language environment.

2 Methodology

The experiment consisted of two tasks, a language proficiency task and a colour elicitation task. Both tasks were made available using Google Forms software and could be completed remotely. Participants were recruited from two universities in Kazan.

The University of Texas at Austin Bilingual Language Profile (Birdsong et al. 2012) was adapted for local conditions and used to assess the language usage of respondents. This survey assesses the age of language acquisition and use of language across different domains and allows for language dominance to be standardised in a quantifiable manner.

In consultation with a lecturer at Kazan Federal University, the original Russian text was altered to make it sound more authentic and certain questions were consolidated for ease of completion. After a pilot study, certain questions remained ambiguous, and these were rephrased. The Russian was translated into Tatar by a postgraduate student at Kazan Federal University. The translation was back-translated into Russian by another bilingual speaker; no discrepancies of note were observed. The questionnaire was presented in a bilingual format to avoid an implicit suggestion one way or another.

The basic colour term inventory of individual speakers was tested using the salience criterion set out by Berlin and Kay (1969/1991). The methodology for assessing was derived from previous studies on salience, such as Battig and Montague (1969), Morgan and Corbett (1989), and Davies and Corbett (1995). The salience value of individual lexical items was calculated using a cognitive salience index (Sutrop 2001) that allows for the comparison of word lists of different lengths. The formula for the cognitive salience index is as follows:

$$S = F/(NmP),$$

where S is the salience, F is the frequency, N is the number of subjects, and mP is the mean position.

In contrast to other approaches (Morgan and Corbett 1989), this index does not take into account the list length provided by the respondent.

The colour listing task consisted of two questions: one asking the respondents to name all the colours and shades of colours in Tatar and one asking the respondents to name all the colours and shades of colours in Russian. This follows Ervin (1961) who interviewed English-Navajo bilinguals first in Navajo and only then in English. It was considered important that respondents were asked to list all shades of colours as well as colours themselves, as this meant they would not restrict themselves to naming only the most common colour terms in each language. No time limit was imposed on the respondents because the intention was to identify the basic colour terms of Tatar using Sutrop's cognitive salience index (Sutrop 2001) rather than by the time/frequency metric used by Morgan and Corbett (1989).

A total of 88 respondents filled out both the Bilingual Language Profile and the colour elicitation task. The average respondent used Tatar 45% of the time, and Russian 41% of the time (the remaining 16% being allocated to other languages, predominantly English and/or other (non-Tatar) minority languages of the Russian Federation).

Of the respondents to the linguistic suitability survey and colour survey, 55 were considered to have met the necessary linguistic suitability criteria; this meant, namely, that the participant reported using Russian or Tatar over 50% of the time in their day-to-day life. These participants formed the population sample for the study. In this respect, the sample size was similar to that in other similar studies such as Ervin (1961), Morgan and Corbett (1989), and Abdramanova (2017).

Out of the sample (Table 1), 38 participants were included in the Tatar group (self-reported speaking Tatar $x > 50\%$ of the time), and 17 were included in the Russian group (self-reported speaking Russian $x > 50\%$ of the time). The demographic breakdown for both groups was similar with the average age for both sample groups being 18.6 years. Out of the Tatar sample group, 26 reported their place of origin as Kazan, 11 as elsewhere in

Table 1: Demographic and bilingual language survey data for Tatar and Russian population samples

	Tatar sample	Russian sample
Sample size	38	17
No. of men	3	2
No. of women	35	15
Average age	18.61	18.63
Standard deviation (age)	1.23	0.96
Self-reported usage of Tatar (% of time)	65%	17%
Self-reported competence in Tatar (1–6)	4.63	2.57
Self-reported usage of Russian (% of time)	31%	62%
Self-reported competence in Russian (1–6)	4.42	4.6

Tatarstan, and 1 as Mari El. For the Russian sample group, 14 participants reported their place of origin as Kazan and 3 as elsewhere in Tatarstan. A gender bias was observed in both sample groups with the Tatar sample consisting of 35 women and 3 men, and the Russian sample consisting of 15 women and 2 men. This gender bias was to be expected given the higher likelihood of women filling in experimental surveys than men (Curtin et al. 2000, Singer et al. 2000).

Both sample groups have a similar experience with their dominant language, with the Tatar sample group using Tatar 65% of the time, and the Russian sample group using Russian 62% of the time. The two sample groups rated their Russian language competence similarly (Tatar: 4.4 versus Russian: 4.6; on a scale of 1–6). The Tatar sample rated their Tatar language competence (on the same scale) as good as, if not better than, their Russian language competence. The Russian sample, on the other hand, rated their Tatar language competence as significantly lower than their Russian language competence. As such, it is possible to label the Tatar sample as functionally bilingual in both Tatar and Russian, and the Russian sample as functionally monolingual in just Russian.

3 Results

Only the responses from the Tatar sample were examined for the Tatar language part of the colour survey. In total, 39 distinct Tatar colour terms were obtained. The median number of terms elicited was 10. All of the six most common colour terms were listed by at least 32 out of the 38 participants (84.2%). A significant difference in salience was observed between the six most frequently listed colour terms, and the remainder of the colour terms listed. This suggests that Tatar is a stage V language.

For the Russian language part of the colour survey, responses from both the Tatar and Russian samples were included. The Tatar group provided a total of 34 colour terms in Russian (median, 10; range, 3–27), while the Russian group provided a total of 35 colour terms (median, 14; range, 9–18). The first 12 colour terms listed (the basic colour terms of Russian) were the same for both the Tatar sample and the Russian sample.

3.1 Research question I: The basic colour term inventory of Tatar

The responses from the Tatar-Russian bilingual sample were examined for this research question. In all, a total of 39 colour terms in the Tatar language were obtained. The median number of terms elicited was 10 (range, 5–18). An overview of their salience scores is provided in Figure 2. The first six terms were listed by at least three-quarters of the respondents, and the first eight by at least half.

The results show a substantial difference between the salience value of the sixth term – *zängär* ‘blue’ (0.16) – and the seventh term – *sory* ‘grey’ (0.07); giving cause for the boundary of salience and non-salience (and by extension, basic colour term status) to be placed between these two terms. Working on the basis of Berlin and Kay’s (1969/1991) notion that salient colours are more likely to be basic and non-salient terms are more likely to not be basic, it can be concluded that Tatar is a Stage V language with six basic colour terms, namely (in order of salience): *kyzyl* ‘red’; *sary* ‘yellow’; *yaşel* ‘green’; *ak* ‘white’; *kara* ‘black’; *zängär* ‘blue’. This is in line with Abdramanova’s (2017) findings on the Tatar language’s linguistically close relative Kazakh which has the same number of basic colour terms.

The basic colour term inventory falls in line with wider expectations of how colour term systems should develop. The seventh most salient colour term identified was ‘grey’ (*sory*). Various linguists (Kay and McDaniel 1978, MacLaury et al. 1992) have identified ‘grey’ as a ‘wild card’ term that can emerge at any basic colour term evolution stage. The first two Tatar terms for ‘brown’, generally considered a pre-requisite for a Stage VI language, are positioned in the eleventh and twelfth positions, respectively (*körän* and *koñgyrt*).

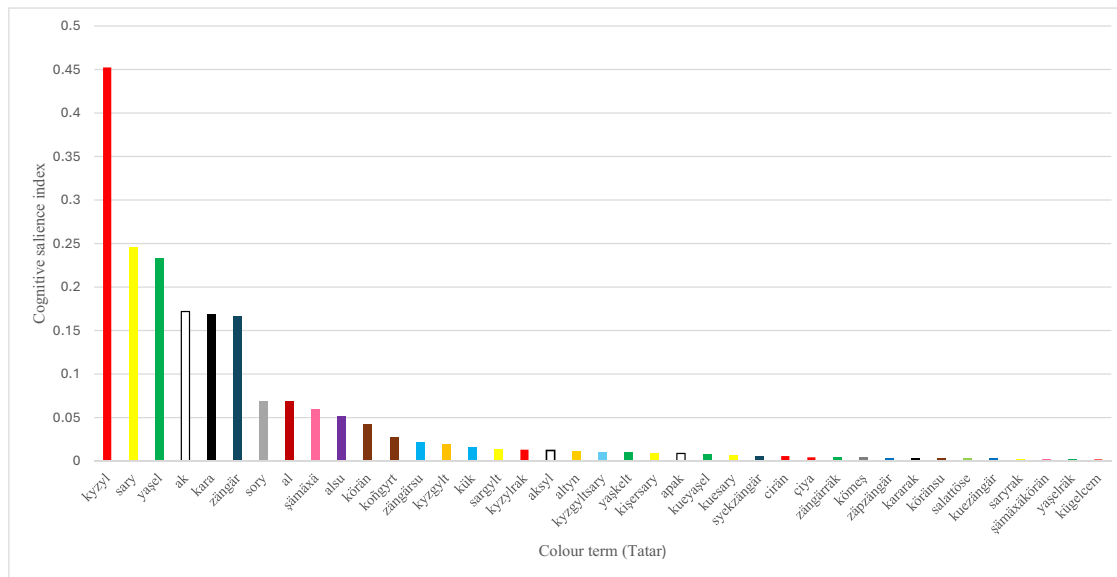


Figure 2: Cognitive salience index scores for Tatar colour terms (colour coding is purely illustrative) – for a full list of Tatar colour terms with glosses and cognitive salience index scores please refer to the Appendix.

3.2 Research question II: Do Tatar-Russian bilingual speakers and Russian monolingual speakers have the same basic colour term inventory in Russian?

The responses from both the Tatar and the Russian groups were considered for this question. The Tatar group provided 34 colour terms in Russian (median, 10; range, 3–27), while the Russian group provided 35 colour terms (median, 14; range, 9–18).

When tested for the basic colour term inventory of Russian, both demographic groups provided the same 12 colour terms (though in different orders) as the most salient. There were, namely: *krasnyj* ‘red’; *belyj* ‘white’; *černyj* ‘black’; *sinij* ‘dark blue’; *želtij* ‘yellow’; *zelenyj* ‘green’; *fioletovij* ‘purple’; *goluboj* ‘light blue’; *rozovij* ‘pink’; *oranževij* ‘orange’; *koričnevij* ‘brown’; *serij* ‘grey’.

An observation of the data (Figure 3) shows a marked difference between the first and second set of six colour terms elicited from the Tatar-Russian bilingual sample. None of the colour terms in the second set (*fioletovij* ‘purple’); *goluboj* ‘light blue’); *rozovij* ‘pink’); *oranževij* ‘orange’); *koričnevij* ‘brown’); *serij*

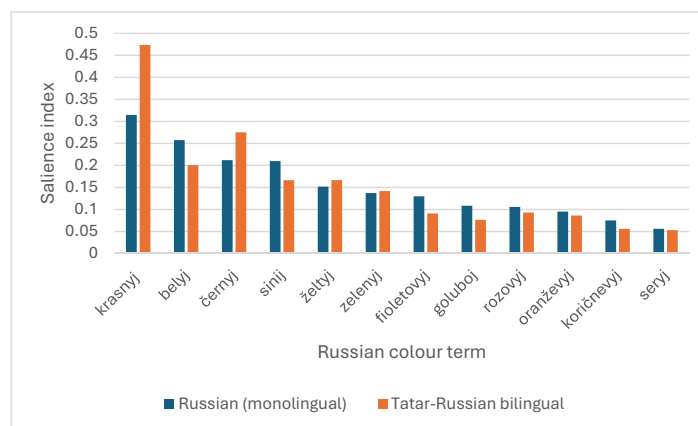


Figure 3: Salience of Russian basic colour terms according to Russian (monolingual) and Tatar-Russian (bilingual) speakers.

Table 2: Frequency (%) at which ‘dark blue’ and ‘light blue’ were reported in Russian or Tatar by Tatar-Russian bilingual respondents ($n = 38$)

Reporting language			
Russian		Tatar	
‘Dark blue’ <i>sinij</i>	‘Light blue’ <i>goluboj</i>	‘Dark blue’ <i>zängär</i>	‘Light blue’ <i>kük</i>
76.3	55.3	84.2	15.8

(‘grey’)) have a cognitive salience index score of greater than 0.01 and Tatar-Russian bilingual sample cognitive salience index score for all of these lexical items is lower than the Russian monolingual sample score.

Under a purely binary approach, two outcomes can be hypothesised: that there is interference from Tatar and that Tatar-Russian bilingual speakers will express fewer basic colour terms in Russian than monolingual Russian speakers, or that there is no interference from Tatar and that Tatar-Russian bilingual speakers and Russian monolingual speakers will have the same basic colour term inventory for Russian. No significant difference was observed in the basic colour term inventory for Russian between the sample groups ($P = 0.26$).

This shows that speakers can maintain different basic colour term systems for different languages. Even though Tatar lacks 6 of Russian’s 12 basic colour terms (‘pink’; ‘purple’; ‘orange’; ‘light blue’; ‘brown’; ‘grey’), Tatar-Russian bilingual respondents were still able to accord them some form of basic colour term status in Russian but not in Tatar.

3.3 Research question III: Do Tatar-Russian bilingual speakers express ‘blue’ in the same way across both languages?

The responses from the Tatar-Russian bilingual sample were considered for this question. Table 2 shows the frequency at which each colour term was expressed in either Russian or Tatar. ‘Dark blue’ was always the more salient colour.

The data in Table 3 were analysed using the Student t -test. They show that the Tatar term for ‘(dark) blue’ – *zängär* – is significantly ($P = 9.52 \times 10^{-12}$) more salient than the Tatar term for ‘light blue’ – *kük*. This serves to reconfirm (as outlined in Section 3.1) that Tatar only has one basic colour term for ‘blue’. The data in Table 3 likewise show that there was no significant difference ($P > 0.01$) in the Tatar-Russian bilinguals’ salience scores for the two Russian terms for ‘blue’.

These results demonstrate that bilingual speakers are able to maintain some form of a distinction between one ‘blue’ and split-‘blue’ type languages. Furthermore, the maintenance of these two ‘blue’ categorisation systems has had no effect on the introduction of the split-‘blue’ system, in contrast to other languages in

Table 3: Expression of ‘dark blue’ and ‘light blue’ in Russian or Tatar by Tatar-Russian bilingual respondents: Comparison of frequencies using the Student t -test

Criterion	Comparison	Statistical analysis	
		<i>P</i> value	Conclusion
Within language			
Russian	‘dark blue’ (<i>sinij</i>) vs ‘light blue’ (<i>goluboj</i>)	0.243	Not significant
Tatar	‘dark blue’ (<i>zängär</i>) vs ‘light blue’ (<i>kük</i>)	9.52×10^{-12}	Significant
Between languages			
‘Dark blue’	Russian (<i>sinij</i>) vs Tatar (<i>zängär</i>)	0.109	Not significant
‘Light Blue’	Russian (<i>goluboj</i>) vs Tatar (<i>kük</i>)	2.27×10^{-4}	Significant

Table 4: Frequency (%) at which ‘dark blue’ and ‘light blue’ were reported in Russian by monolinguals and Tatar bilinguals

Demographic group			
Monolingual Russian (<i>n</i> = 17)		Bilingual Tatar/Russian (<i>n</i> = 38)	
‘Dark blue’ <i>sinij</i>	‘Light blue’ <i>goluboj</i>	‘Dark blue’ <i>sinij</i>	‘Light blue’ <i>goluboj</i>
100	94.1	76.3	55.3

contact with Russian such as Komi Permyak or Northern Udmurt (Ryabina 2011, 2014), showing that such developments are not obligatory.

3.4 Research question IV: Do Tatar-Russian bilingual speakers and Russian monolingual speakers express ‘blue’ in the same way in Russian?

The responses from both the Tatar and the Russian groups were considered for this question. Table 4 shows the frequency at which each colour term for ‘blue’ in Russian was expressed by the Tatar-Russian bilingual sample or the Russian monolingual sample.

As shown in Table 5, for both sample groups, the difference in frequency between the two Russian terms for ‘blue’ was not statistically significant ($P > 0.01$), again further confirming that Russian has two basic colour terms for ‘blue’. A noticeable difference was, however, observed in the frequency of elicitations between the two sample groups.

This significant difference between the responses of the Tatar-Russian bilingual sample and the Russian monolingual sample shows that Tatar-Russian bilingual speakers experience interference from the ‘blue’ colour structure of Tatar when they speak Russian.

4 Discussion

Three key findings emerge from the data which merit further discussion: the structure of the basic colour term inventory in Tatar and the similarities and differences between it and the colour term inventories of other related languages; the imperviousness of the Tatar colour lexicon to Russian linguistic influences, and the link with bilingualism; and linguistic interference in the categorisation of colour terms.

The results show a large degree of similarity with the basic colour term system of Kazakh, the only other Kipchak Turkic language that has been subjected to such investigation. Similarities are noted for the first five lexical items (according to Berlin and Kay’s (1969/1991) development hierarchy) and also both languages’ status

Table 5: Expression of ‘dark blue’ and ‘light blue’ in Russian by Russian (monolingual) and Tatar-Russian bilingual respondents: Comparison of frequencies using the Student *t*-test

Criterion	Comparison	Statistical analysis	
		P value	Conclusion
Within demographic group			
Monolingual	‘Dark blue’ (<i>sinij</i>) vs ‘light blue’ (<i>goluboj</i>)	0.332	Not significant
Bilingual		0.243	Not significant
Between demographic groups			
Dark blue (<i>‘sinij’</i>)	Russian (monolingual) vs Tatar-Russian (bilingual)	1.97×10^{-4}	Significant
Light blue (<i>‘goluboj’</i>)		3.12×10^{-4}	Significant

as Stage V languages; the extent of similarities between the two languages diminishes as colour terms grow less salient. The lexical similarity of the two languages is unsurprising given that according to Kononov's (1978) reconstruction of proto-Turkic colour terms, proto-Turkic was a Stage IV language. The key area of divergence between Tatar and Kazakh basic colour terms lies within the lexical terms used to describe the 'blue' part of the colour spectrum. Both languages have a single basic overarching term for 'blue' with a specific non-basic term for 'light blue'. Kazakh has reanalysed the term for 'sky' – *kök* – as 'blue' and developed a secondary subsidiary term for 'light blue' – *kögildir* – which is lexically derived from the word for 'blue'. Tatar has taken the opposite approach in using a distinct word (though also of Turkic origin – Axmatjanov 2001) for 'blue' – *zängär* – and has retained the word for 'sky' (*kük*) to mean 'light blue'. Tatar has 2 terms for 'brown' within the 12 colour terms with the highest cognitive salience index rating – *körän* and *koñgyrt* (in eleventh and twelfth position, respectively). These can be compared with two similar terms for 'brown' in Kazakh – *qoñyr* and *küren*. A similar form of lexical inversion as found in the colour 'blue' across Tatar and Kazakh can also be found in the colour 'brown'. In Tatar, *körän* is the default term for 'brown' with *koñgyrt* being lexically restricted to collocations involving hair or eye colour. In contrast, *qoñyr* is the default term for 'brown' in Kazakh, with *küren* being the lexically restricted item. Such a finding is unusual in light of Rakhilina and Paramei's (2011) analysis which predicts that such collocations emerge as a result of historic change. Given that both Tatar and Kazakh share a common historical ancestor, it would be natural to presume that the default colour term for 'brown' would be the same in both languages. As far as other important differences between Tatar and Kazakh are concerned, attention should also be drawn to the terms for 'pink' and 'purple' which are *šämäxä* and *alsu*, respectively, in Tatar, but *kyzğylt* and *külgin* in Kazakh. Furthermore, there is no equivalency between the Kazakh term *kyzğylt* ('pink') and the Tatar term *kyzğylt* ('orange'). The Tatar term *al* ('scarlet') does not correspond to the Kazakh term *ala* ('multicoloured') being instead derived from the adjectival Russian term for 'scarlet' *alij*, rather the term *ala* means 'multicoloured' in both languages.

Unlike the languages of other minoritised peoples within the Russian Federation (Ryabina 2011, 2014 for Komi and Udmurt), the basic colour term inventory of Tatar has remained unaffected by contact with Russian, both in terms of the number of basic colour terms and the actual lexemes used to describe various colour terms.

Based on the data collected in this study, Tatar has been identified as a Stage V language with six basic colour terms. Russian on the other hand is a Stage VII language with 12 basic colour terms.

As far as the effect of Russian language contact on Tatar colour terms is concerned, the findings show that Russian has had no influence on the development of colour terms in Tatar at either a structural or lexical level. There is a lack of Russian loan words within the Tatar colour inventory and Tatar has not acquired either a wider range of basic colour terms or the split-'blue' basic colour term system from Russian. This is in contrast to the other minoritised languages in the Russian Federation such as Komi or Udmurt, which are developing distinct colour terms outside of the structure proposed by Berlin and Kay (1969/1991) as a result of contact with Russian (Ryabina 2011, 2014).

The differences in the development of the colour term inventory of Komi and Udmurt on the one hand and Tatar on the other can, in large part, be attributed to differences in the sociolinguistic status of the languages rather than any genetic differences between them. Due to various factors, including length of language contact and the existence of a Tatar nation-state, the Tatar language has found itself to be more impervious to Russian language influence than other minoritised languages such as Komi and Udmurt.

As far as the length of cultural exposure is concerned, Russian-speaking peoples were making inroads into the Komi and Udmurt-speaking communities as early as the twelfth and thirteenth centuries, respectively (Avril 2006, Berdinskix 2002). In contrast, the Tatars were organised into a distinct state independent of Russian influence until the sixteenth century, whereupon it was subjugated by Russian forces in 1552 (Wigglesworth-Baker 2015).

It has been postulated (Faller 2011) that the Tatar people constitute a *sui generis* case study of a minoritised nationality within the Russian Federation, as they constitute a nation-state within the wider federation, as opposed to other minoritised peoples who lack such cultural–political organisation. Arutyunova and Zamyatin (2020) report that social activism against the 2018 education law increasing the status of the Russian language *vis-à-vis* recognised minority languages was strongest in Tatarstan and that the response by the federal

authorities in relation to the 2018 law was likewise largely focused on Tatarstan. Unlike other minoritised languages, including Komi and Udmurt (Ryabina 2011), the Tatar language already had both a written standard and literary tradition by the end of the nineteenth century, as well as a historical legacy of an ancestor state. Furthermore, differing religious beliefs between Tatar people, who are primarily Muslim, and Russian people who are primarily Christian allowed for further differentiation. This differentiation based on religion does not exist for Komi, Udmurt, and Russian people who are all coreligionists. A further aspect to consider is the size of the ethnic groups. Tatars constitute the second largest ethnic group in the Russian Federation numbering around 5 million people. This contrasts with the Komi and Udmurt peoples who number around 350,000 and 500,000 people respectively (Rosstat 2010). Finally, unlike other titular republics within the Russian Federation (including the Komi Republic and Udmurtia), from the 1990s until the late 2010s, Tatarstan was able to secure for itself far-reaching autonomy which included mandatory Tatar language lessons for students in Russian language medium schools. Such a combination of factors allows for Russian cultural influences to be kept at arm's length within the Tatar cultural milieu in a manner that is unavailable in Komi and Udmurt cultural settings.

The other relevant point for discussion is the link between bilingualism and colour categorisation. The linguistic situation of Tatarstan is of particular relevance. The Tatar language-speaking community is functionally bilingual (even if asymmetrically) in both Tatar and Russian and is exposed to both languages from an early age. This contrasts with other studies (e.g. Andrews (1994) for Russian–English contact; Athanasopoulos (2008) for Greek–English contact; Paramei et al. (2016) for Italian–English contact), which focus on bilingualism arising from a migratory setting, and often occurring at a late stage of development. Similarly, given the lack of language contact effects of Russian on Tatar, nor any genetic link between the two languages, there are no homophony effects unlike those observed in Italian–English language contact (Paramei et al. 2016, Douven and Paramei 2023).

The data show that although Tatar-Russian bilingual speakers are able to maintain a distinction between the colour term systems of both Tatar and Russian, some degree of interference nevertheless takes place.

It can be argued that the expression of 'blue' by Tatar-Russian bilingual speakers occupies a middle ground between the expression of 'blue' in Tatar and the expression of 'blue' in Russian by monolingual Russian speakers. This is evidence of Tatar interference in the expression of blue by Tatar-Russian bilinguals in Russian. Similar to the findings of Andrews (1994), the less dominant language, in this instance the speaker's L2, is the one being subject to influence, in this case by not having as clear a distinction in terms of *sinij* versus *goluboj*, as it would be with monolingual speakers. When these data are compared with those from similar studies concerning the role of bilingualism on colour categorisation, it shows that its effects are not restricted in terms of directionality (L1 has a split blue system and L2 has a single-'blue' system or vice versa), nor in terms of the nature of the bilingualism (naturally occurring bilingualism or bilingualism that results from migration).

When viewed within the wider context, several conclusions can be drawn on the link between semantic shifts in colour categorisation and language acquisition. However, it is important to note that these shifts did not alter the focus of the basic colour terms for blue and therefore undermine the conceptual understanding of there being two salient terms. Andrews (1994) found that the *sinij/goluboj* distinction solidifies in adulthood and further exposure to a single-'blue' system (in this instance English) does not result in loss in the perception of *sinij* and *goluboj* as two distinct salient categories. The results indicate that the split-'blue' system is sufficiently typologically unusual that it is difficult for non-native speakers to acquire it. Further research will be necessary to determine how this interacts with the process of child language acquisition. In the case of Tatar-Russian bilinguals, it is either acquired incompletely (due to limited exposure) at an early age or at a later date with Russian very much being a second language. This is in line with previous work conducted on the acquisition of a 'light blue'/'dark blue' distinction in second-language speakers of Russian. Vartanov and Nguyen's (1995) study of Vietnamese speakers' intuition of Russian terms for 'light blue' and 'dark blue' found that intuitions were strongly correlated with Russian language competence with speakers with high level Russian competence having almost native-like intuitions about the *sinij/goluboj* distinction. Nevertheless, even in the case of those with the highest level of Russian language competence, these intuitions were in some places imperfect, specifically, when distinguishing it from 'sea-blue' or other pale modifiers. This observation would go a fair way in explaining the interference in terms of the salience of the two-colour terms in comparison to monolingual Russian speakers. Such a hypothesis would go a long way in explaining how

the notion of there being two salient terms for blue in both Greek and Russian does not dissipate for speakers who enter a bilingual speech community in adulthood. Similarly, it would explain why the split-‘blue’ colour system of Russian has been repurposed into the single-‘blue’ colour system of English by young heritage speakers of Russian within the émigré community (Andrews 1994).

Looking at the wider phenomenon of colour categorisation and bilingualism, the data from this study can be integrated with previous investigations into the role of bilingualism and colour categorisation to identify four main frames in which bilingualism results (or does not result) in shifts in colour categorisation within the split-‘blue’ phenomenon.

At the one extreme are situations where a substantively marginalised single-‘blue’ language is in contact with split-‘blue’ language, speakers will adopt the split-‘blue’ system of the L2 into the L1. Evidence of this can be found in highly minoritised languages in the Russian Federation (e.g. Komi-Permyak and Northern Udmurt; Ryabina 2011, 2014), which have developed a split-‘blue’ system as a result of contact with Russian, in advance of other developments that would have been postulated by Berlin and Kay (1969/1991).

In a situation where the usage and status of the languages is broadly equivalent, or (as indicated earlier) the colour categorisation system of one language has acquired permanence in the mind of the speaker as part of the process of language acquisition, one of two scenarios will emerge depending on the colour classification system of the speaker’s L1.

In situations where native speakers of a split-‘blue’ language (refer to Andrews 1994 for Russian; Athanasopoulos 2008 for Greek) acquire a single-‘blue’ language in later life, speakers continue to maintain the split-‘blue’ distinction in their L1. In Russian, the focus of the two basic colour terms for blue remains fixed (Andrews 1994), whilst in Greek the focus of the two-colour terms undergoes a semantic shift but the overall boundaries of the ‘blue’ part of the colour spectrum remain fixed. The disparity between the two languages can be attributed to the location of the respective colour terms within the colour spectrum. In any case, however, the actual notion of there being two basic colour terms for blue does not change.

Although similar in terms of their categorisation, there is one notable difference between the role of language contact, on the one hand in the Tatar-Russian scenario and the Greek/Russian-English on the other. When the L1 is a single-‘blue’ system and the L2 is a split-‘blue’ system, there is still a statistically significant degree of interference from the L1 into the L2. On the other hand, when the L1 is a split-‘blue’ and the L2 is a single-‘blue’ system, then no such interference exists on the basis that there is no additional lexical term that can compete for assignment as the headword for the colour category of blue. Such a phenomenon is evidenced in the instance of Tatar-Russian bilingualism hitherto discussed, where their L2 intuitions about the salience of the Russian split-‘blue’ system are not as clearly defined as with monolingual Russian speakers. In terms of understanding the bilingual nature of the Tatar speech community, this is notable as it confirms that Tatar continues to occupy the role of a working language within certain functional domains (cf. Edygarova 2014 for Udmurt where this is not the case).

In situations where the L1 is single-‘blue’ and the L2 is a minoritised language with a split-‘blue’ colour system, then a different form of interference between the languages takes place. In this instance, the dominant term for ‘blue’ becomes an overarching term for both categories and the second basic term for blue becomes demoted to a subsidiary term (Andrews 1994).

The data from this study can be contrasted with the findings proposed by Andrews (1994) in his study on English-Russian bilingualism who found that heritage speakers of Russian in the diaspora applied the English single-‘blue’ categorisation template on the Russian split-‘blue’ template. In the instance of Tatar-Russian bilingualism, the two templates remained distinct with Tatar-Russian bilingual speakers not subsuming *goluboj* (‘light blue’) under the wider category heading of *sinij* (‘dark blue’).

This difference in categorisation can largely be attributed to the differing sociolinguistic dynamics experienced. In the émigré community studied by Andrews (1994), English is the only official language and thus exerts a dominant position in the mind of the bilingual speaker. In contrast, Tatar is a co-official language within Tatarstan, and whilst Russian is used as the language of inter-community communication, Tatar is still used sufficiently within the public and private sphere, and children acquire both languages during their childhood. This greater exposure to and usage of both languages serves to keep the two categorisation systems distinct in each language.

Aside from the blue part of the colour spectrum, the wider interaction between the Tatar and the Russian systems of colour categorisation, especially within the wider bilingual setting, also merits discussion. The results of the surveys reveal that there is little overlap between the secondary colour term structures of Tatar and Russian (positions 7 to 12). This finding is reflected in differences between monolingual Russian speakers and bilingual Tatar-Russian speakers in their categorisation of Russian basic colour terms, with monolingual Russian speakers on average finding secondary basic colour terms more salient than Tatar-Russian bilingual speakers (monolingual Russian 0.095 vs Tatar-Russian bilingual 0.076). Although the findings show in a significant manner that Tatar-Russian bilingual speakers perceive Russian to have 12 basic colour terms, just like monolingual Russian speakers, there is a difference in how they perceive the salience of the secondary basic colour terms. This suggests that the interference that Tatar-Russian bilingual speakers experience with categorising distinct basic colour terms for ‘dark blue’ and ‘light blue’ extends across the entire basic colour term system of Russian.

These data seem to indicate that it is not entirely appropriate to speak of a strict dichotomy between basic and non-basic terms; rather a gradated approach is more suitable. Whereas both sample groups accord basic colour term status to all 12 standardly acknowledged basic colour terms in Russian, the Tatar-Russian bilingual sample accords a lower degree of basicness to the latter 6 lexical items than the Russian monolingual sample. This hints at interference in the processing of colour terms by Tatar-Russian bilinguals. Whilst they are able to maintain a difference between the 6-term colour system in Tatar and the 12-term system in Russian, interaction between the two languages has led to the latter 6 basic colour terms in Russian being accorded a lower level of salience than they enjoy among monolingual Russian speakers. An analogy can be drawn between this scenario and that of Italian-English bilingual speakers (as discussed by Paramei et al. (2016) and Douven and Paramei (2023)) where although they were still able to draw a structural distinction between the colour term system of Italian and English, the focal points of the system had undergone a shift as a result of cross-linguistic influence.

As has already been discussed as regards the categorisation of ‘blue’ in Russian by Tatar-Russian bilingual speakers, this scenario whereby such speakers are able to maintain distinct colour term categorisations for each language and nevertheless experience interference can in the large part be explained by sociolinguistic constraints. As has been illustrated in other instances (see, for example, Paramei 2007), it is possible for speakers to acquire the colour classification for another language. Given the extent of exposure Tatar-Russian bilingual speakers have to Russian, it is not unreasonable to assume that members of this speech community acquired the colour classification system for each language at an early age, and largely independently from the other. Nevertheless, due to Tatar being the more dominant language for Tatar-Russian bilingual speakers, in marginal cases, such as the classification of secondary basic colour terms, a tendency to revert to the classification system (in full or in part) is made manifest, leading to the interference observed as seen in both the terminology for the colour ‘blue’ and the wider colour classification system.

5 Conclusion

This study has shown that similar to other Kipchak languages, such as Kazkah, Tatar is a Stage V language per Berlin and Kay’s basic colour term hierarchy with six basic colour terms (*kyzyl* ‘red’; *sary* ‘yellow’; *yaşel* ‘green’; *ak* ‘white’; *kara* ‘black’; and *zängär* ‘blue’). Language contact with Russian has had no effect on the development of Tatar colour terms. As such, it has only one basic colour term for ‘blue’.

No significant difference was observed in the categorisation of basic colour terms in Russian by monolingual Russian speakers or by Tatar-Russian bilingual speakers. A degree of interference was however observed for Tatar-Russian bilingual speakers, especially for secondary basic colour terms. Similarly, both sample groups found both *sinij* ‘dark blue’ and *goluboj* ‘light blue’ to be salient in Russian; again, a degree of interference was observed in Tatar-Russian bilingual speakers in terms of their perception of the salient nature of the *sinij/goluboj* distinction. These findings suggest that there are elements of gradation, both to the representation of colour categorisation systems in general by bilingual speakers, and also specifically to

the representation of a split-‘blue’ system. Tatar-Russian bilingual speakers in large part recognise discrete colour categorisation systems for each language, subject to such interference effects already described. The findings of this article pose further questions about the language acquisition process behind both colour categorisation systems in general and the split-‘blue’ system more specifically.

Acknowledgments: I would like to thank Neil Bermel for his considerable help in setting up the experiment and useful comments on the draft and Greville Corbett for help in preparing the manuscript. I would also like to thank my contacts in Kazan, Guzelya Akhmetgaraeva, Raushaniya Nurmukhametova, and Ruslan Shakirov for their local knowledge and help in running the experiment. I am very grateful to two anonymous reviewers of *Open Linguistics* for their highly constructive comments and suggestions. Any errors are, of course, my own.

Funding information: The author states no funding involved.

Author contribution: The author confirms the sole responsibility for the conception of the study, presented results, and manuscript preparation.

Conflict of interest: The author states no conflict of interest.

References

- Abdramanova, Saule. 2017. “Basic Color Terms in the Kazakh Language.” *Sage Open* 7 (2): 2158244017714829. doi: 10.1177/2158244017714829.
- Amengual, Mark. 2016. “Cross-Linguistic Influence in the Bilingual Mental Lexicon: Evidence of Cognate Effects in the Phonetic Production and Processing of a Vowel Contrast.” *Frontiers in Psychology* 7: 617–34. doi: 10.3389/fpsyg.2016.00617.
- Andrews, David R. 1994. “The Russian Color Categories Sinij and Goluboj: An Experimental Analysis of Their Interpretation in the Standard and Emigré Languages.” *Journal of Slavic Linguistics* 2 (1): 9–28. <https://www.jstor.org/stable/24599022>.
- Arutyunova, Ekaterina and Konstantin Zamyatin. 2020. “An Ethnolinguistic Conflict on the Compulsory Learning of the State Languages in the Republics of Russia: Policies and Discourses.” *The International Journal of Human Rights* 25 (5): 832–52. doi: 10.1080/13642987.2020.1804368.
- Athanasopoulos, Panos. 2007. “Interaction Between Grammatical Categories and Cognition in Bilinguals: The Role of Proficiency, Cultural Immersion, and Language of Instruction.” *Language and Cognitive Processes* 22 (5): 689–99. doi: 10.1080/01690960601049347.
- Athanasopoulos, Panos. 2008. “Cognitive Representation of Colour in Bilinguals: The Case of Greek Blues.” *Bilingualism Language and Cognition* 12 (1): 83–95. doi: 10.1017/s136672890800388x.
- Athanasopoulos, Panos. 2011. “2. Cognitive Restructuring in Bilingualism.” In *Multilingual Matters eBooks*, 29–65. doi: 10.21832/9781847693389-004.
- Athanasopoulos, Panos, Ljubica Damjanovic, Andrea Krajciová, and Miho Sasaki. 2010a. “Representation of Colour Concepts in Bilingual Cognition: The Case of Japanese Blues.” *Bilingualism Language and Cognition* 14 (1): 9–17. doi: 10.1017/s1366728909990046.
- Athanasopoulos, Panos, Benjamin Dering, Alison Wiggett, Jan-Rouke Kuipers, and Guillaume Thierry. 2010b. “Perceptual Shift in Bilingualism: Brain Potentials Reveal Plasticity in Pre-attentive Colour Perception.” *Cognition* 116 (3): 437–43. doi: 10.1016/j.cognition.2010.05.016.
- Athanasopoulos, Panos and Fraibet Aveledo. 2012. “Linguistic Relativity and Bilingualism.” In *Cambridge University Press eBooks*, 236–55. doi: 10.1017/cbo9781139035279.010.
- Avril, Yves. 2006. *Parlons Komi*. Paris: Editions L'Harmattan.
- Axmatjanov, Rifkat. 2001. *Tatar telemen kyskača tarixi-etimologik suzlege [Concise historical and etymological dictionary of the Tatar language]*. Kazan: Tatarstan Kitap Neşriyatı.
- Battig, William F. and William E. Montague. 1969. “Category Norms of Verbal Items in 56 Categories a Replication and Extension of the Connecticut Category Norms.” *Journal of Experimental Psychology* 80 (3 Pt.2): 1–46. doi: 10.1037/h0027577.
- Baxilina, Natalja. 1975. *Istorija cvetooboznacenij v russkom jazyke [History of colour terms in Russian]*. Moscow: Nauka.
- Berlin, Brent and Paul Kay. 1969/1991. *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.
- Berdinskix, Viktor. (2002). *Istorija goroda Vjatki [History of the city of Vjatka]*. Očerki – Kirov: Vjatskoe knižnoe izdatelstvo.
- Birdsong, David, Libby Gertken, and Mark Amengual. 2012. *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism*. <https://sites.la.utexas.edu/bilingual/>.
- Brown, Roger W. and Eric H. Lenneberg. 1954. “A Study in Language and Cognition.” *Journal of Abnormal & Social Psychology* 49 (3): 454–62. doi: 10.1037/h0057814.

- Burgess, Don, Willett Kempton, and Robert E. Maclaury. 1983. "Tarahumara Color Modifiers: Category Structure Presaging Evolutionary Change." *American Ethnologist* 10 (1): 133–49. doi: 10.1525/ae.1983.10.1.02a00080.
- Campbell, Lyle. 2000. *American Indian Languages: The Historical Linguistics of Native America*. Oxford: Oxford University Press.
- Caskey-Sirmons, Leigh and Nancy P. Hickerson. 1977. "Semantic Shift and Bilingualism: Variation in the Color Terms of Five Languages." *Anthropological Linguistics* 19 (8): 358–67. <https://www.jstor.org/stable/30027603>.
- Cook, Vivian. 2002. *Portraits of the L2 User*. Bristol: Multilingual Matters.
- Coventry, Kenny, Christos Mitsakis, Ian Davies, Julio Lillo Jover, Anna Androulaki, and Natalia Gómez-Pestaña. 2006. "Basic Colour Terms in Modern Greek: Twelve Terms Including Two Blues." *Journal of Greek Linguistics* 7 (1): 3–47. doi: 10.1075/jgl.7.03and.
- Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. "The Effects of Response Rate Changes on the Index of Consumer Sentiment." *Public Opinion Quarterly* 64 (4): 413–28. doi: 10.1086/318638.
- Davies, Ian and Greville Corbett. 1994. "The Basic Color Terms of Russian." *Linguistics* 32 (1): 65–90. doi: 10.1515/ling.1994.32.1.65.
- Davies, Ian and Greville Corbett. 1995. "A Practical Field Method for Identifying Probable Basic Colour Terms." *Languages of the World* 9: 25–36. <https://epubs.surrey.ac.uk/2090/>.
- Douven, Igor and Galina V. Paramei. 2023. "Optimality, Prototypes, and Bilingualism." In *Trends and Challenges in Cognitive Modeling*, 147–64. doi: 10.1007/978-3-031-41862-4_11.
- Edygarova, Svetlana. 2014. "The Varieties of the Modern Udmurt Language." *Finnisch-Ugrische Forschungen* 2014 (62): 376–98. doi: 10.33339/fuf.86085.
- Ervin, Susan M. 1961. "Semantic Shift in Bilingualism." *The American Journal of Psychology* 74 (2): 233. doi: 10.2307/1419408.
- Faller, Helen M. 2011. *Nation, Language, Islam: Tatarstan's Sovereignty Movement*. Budapest: Central European University Press.
- Frumkina, Revekka. 1984. *Cvet, smysl, sxodstvo. Aspekty psixolingvisticheskogo analiza [Colour, meaning, and similarity: Aspects of psycholinguistic analysis]*. Moscow: Nauka.
- Grosjean, François. 1989. "Neurolinguists, Beware! The Bilingual Is Not Two Monolinguals in One Person." *Brain and Language* 36 (1): 3–15. doi: 10.1016/0093-934x(89)90048-5.
- Hunt, Earl and Franca Agnoli. 1991. "The Whorfian Hypothesis: A Cognitive Psychology Perspective." *Psychological Review* 98 (3): 377–89. doi: 10.1037/0033-295x.98.3.377.
- Kay, Paul and Chad K. McDaniel. 1978. "The Linguistic Significance of the Meanings of Basic Color Terms." *Language* 54 (3): 610–46. doi: 10.2307/412789.
- Kay, Paul and Terry Regier. 2003. "Resolving the Question of Color Naming Universals." *Proceedings of the National Academy of Sciences* 100 (15): 9085–89. doi: 10.1073/pnas.1532837100.
- Kay, Paul. 2015. "Universality of Color Categorization." In *Cambridge University Press eBooks*, 245–58. doi: 10.1017/cbo9781107337930.012.
- Kertulla, Seija. 2004. "English Colour Terms: Etymology, Chronology, and Relative Basicness." PhD diss., University of Helsinki.
- Kononov, Andrej. 1978. Semantika cvetoooboznačeniiv tjurkskix jazikax [The semantics of colour terms in Turkic languages]. *Tjurkologičeskij sbornik* 159–79.
- Konstitucija Respubliki Tatarstan [Constitution of the Republic of Tatarstan]. 1991. Kazan: Minijust Tatarstana.
- Levinson, Stephen. 1996. "Relativity in Spatial Conception and Description." In *Rethinking Linguistic Relativity*, 177–202. Cambridge: Cambridge University Press.
- Lindsey, Delwin T. and Angela M. Brown. 2021. "Lexical Color Categories." *Annual Review of Vision Science* 7 (1): 605–31. doi: 10.1146/annurev-vision-093019-112420.
- Lucy, John Arthur. 1993. *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.
- MacLaury, Robert E., Gordon W. Hewes, Paul R. Kinnear, Jan B. Derogowski, William R. Merrifield, Barbara A. C. Saunders, James Stanlaw, Christina Toren, J. Van Brakel, and Roger W. Wescott. 1992. "From Brightness to Hue: An Explanatory Model of Color-Category Evolution [and Comments and Reply]." *Current Anthropology* 33 (2): 137–86. doi: 10.1086/204049.
- Morgan, Gerry and Greville Corbett. 1989. "Russian Colour Term Salience." *Russian Linguistics* 13 (2): 125–41. doi: 10.1007/bf02551669.
- Mylonas, Dimitris, Serge Caparos, and Jules Davidoff. 2022. "Augmenting a Colour Lexicon." *Humanities and Social Sciences Communications* 9 (1): 1–12. doi: 10.1057/s41599-022-01045-3.
- Özgen, Emre and Ian R. L. Davies. 1998. "Turkish Color Terms: Tests of Berlin and Kay's Theory of Color Universals and Linguistic Relativity." *Linguistics* 36 (5). doi: 10.1515/ling.1998.36.5.919.
- Paramei, Galina V. 2004. "Singing the Russian Blues: An Argument for Culturally Basic Color Terms." *Cross-Cultural Research* 39 (1): 10–38. doi: 10.1177/1069397104267888.
- Paramei, Galina V. 2007. "Russian 'Blues': Controversies of Basicness." In *John Benjamins Publishing Company eBooks*, 75–106. doi: 10.1075/z.137.07par.
- Paramei, Galina V., Mauro D'Orsi, and Gloria Menegaz. 2016. "Cross-Linguistic Similarity Affects L2 Cognate Representation: Blu vs Blue in Italian-English Bilinguals." *Journal of the International Colour Association* 16: 69–81. <https://hira.hope.ac.uk/id/eprint/1624/>.
- Paramei, Galina and David Bimler. 2021. "Language and Psychology." In *A Cultural History of Color, Vol. 6, The Modern Age: From 1920 to present*, 117–34. London: Bloomsbury.
- Poppe, Nicholas. 1965. *Introduction to Altaic Linguistics*. Weisbaden: Otto Harrassowitz.
- Rakhilina, Ekaterina and Galina Paramei. 2011. "Colour Terms: Evolution via Expansion of Taxonomic Constraints." In *New Directions in Colour Studies*, 121–32. Amsterdam/Philadelphia: John Benjamins.

- Rätsep, Kaidi. 2018. "Colour terms in Turkish, Estonian and Russian: How many basic blue terms are there?" PhD diss., University of Tartu. <https://core.ac.uk/download/pdf/187733469.pdf>.
- Regier, Terry, Paul Kay, and Naveen Khetarpal. 2007. "Color Naming Reflects Optimal Partitions of Color Space." *Proceedings of the National Academy of Sciences* 104 (4): 1436–41. doi: 10.1073/pnas.0610341104.
- Roberson, Debi, Ian Davies, and Jules Davidoff. 2000. "Color Categories Are Not Universal: Replications and New Evidence From a Stone-age Culture." *Journal of Experimental Psychology General* 129 (3): 369–98. doi: 10.1037/0096-3445.129.3.369.
- Roberson, Debi. 2004. "Color Categories Are Culturally Diverse in Cognition as Well as in Language." *Cross-Cultural Research* 39 (1): 56–71. doi: 10.1177/1069397104267890.
- Rosstat. 2010. *Vserossijskaja perepis' naselenija – Vol 4. Nacional'nyj sostav i vladenie jazykami, graždanstvo [Russian census – Vol. 4 Distribution of nationalities and languages]*.
- Ryabina, Elena. 2011. "Differences in the Distribution of Colour Terms in Colour Space in the Russian, Udmurt and Komi Languages." *Eesti Ja Soome-Ugri Keeleteaduse Ajakiri Journal of Estonian and Finno-Ugric Linguistics* 2 (2): 191–213. doi: 10.12697/jeful.2011.2.2.10.
- Ryabina, Elena. 2014. "A Study of Loan Colour Terms in the Komi Languages." *Eesti Ja Soome-Ugri Keeleteaduse Ajakiri Journal of Estonian and Finno-Ugric Linguistics* 5 (2): 91–116. doi: 10.12697/jeful.2014.5.2.05.
- Singer, Eleanor, John Van Hoewyk, and Mary P. Maher. 2000. "Experiments With Incentives in Telephone Surveys." *Public Opinion Quarterly* 64 (2): 171–88. doi: 10.1086/317761.
- Slobin, Dan I. 1996. "From 'Thought and Language' to 'Thinking for Speaking.'" In *Rethinking Linguistic Relativity*, 70–96. Cambridge: Cambridge University Press. <http://psycnet.apa.org/record/1996-98701-002>.
- Sutrop, Urmas. 2001. "List Task and a Cognitive Salience Index." *Field Methods* 13 (3): 263–76. doi: 10.1177/1525822x0101300303.
- Uusküla, Mari and David Bimler. 2016. "From Listing Data to Semantic Maps: Cross-Linguistic Commonalities in Cognitive Representation of Colour." *Folklore Electronic Journal of Folklore* 64: 57–90. doi: 10.7592/fejf2016.64.colour.
- Vartanov, Alexander and B. Thu Nguyen. 1995. "Formirovanie semantičeskogo cvetovogo prostranstva pri usvoenii inostrannyx jazykov [Formation of a semantic color space while learning foreign languages]." *Psikhologija* 14: 29–46.
- Wattendorf, Elise and Julia Festman. 2008. "Images of the Multilingual Brain: The Effect of Age of Second Language Acquisition." *Annual Review of Applied Linguistics* 28: 3–24. doi: 10.1017/s0267190508080033.
- Wigglesworth-Baker, Teresa. 2015. "Language Policy and Russian-Titular Bilingualism in Post-Soviet Tatarstan." PhD diss., University of Sheffield. <https://etheses.whiterose.ac.uk/10057/1/Language%20Policy%20and%20Russian%20Titular%20Bilingualism%20in%20Post-Soviet%20Tatarstan%20ethesis.pdf>.
- Zamyatin, Konstantin. 2014. "An Official Status for Minority Languages?: A Study of State Languages in Russia's Finno-Ugric Republics." PhD diss., University of Helsinki. <https://helda.helsinki.fi/server/api/core/bitstreams/dd4eb810-14b8-4d0c-abb0-3269340ca4d4/content>.

Appendix

A1 Participant instructions for the colour elicitation task

Tatar colour term section:

Татар телендә үзегез белгән БАРЛЫК төсләрне һәм төсләрнең төсмерләрен языгызчы. // Напишите, пожалуйста, ВСЕ названия цветов и оттенков цветов на татарском языке которые Вы знаете.

Russian colour term section:

Напишите, пожалуйста, ВСЕ названия цветов и оттенков цветов на русском языке которые Вы знаете. // Рус телендә үзегез белгән БАРЛЫК төсләрне һәм төсләрнең төсмерләрен языгызчы.

Table A1: Tatar colour terms elicited from Tatar-Russian bilingual respondents

Colour term	Gloss	Cognitive salience index	Total number of elicitations	Average list position
kızıl	red	0.4524	38	2.21
sarı	yellow	0.2461	35	3.74
yaşıl	green	0.2324	37	4.19
ak	white	0.1719	34	5.21
kara	black	0.1690	34	5.29
zängär	dark blue	0.1663	32	5.06
sory	grey	0.0688	22	8.41
al	scarlet	0.0688	20	7.65
şamäxä	purple	0.0591	16	7.13
alsu	pink	0.0519	12	6.08
körän	brown	0.0426	14	8.64
koñgyrt	brown	0.0267	8	7.88
zängärsu	blue	0.0216	8	9.75
kızgylt	reddish	0.0197	6	8.00
kük	light blue	0.0161	6	9.83
sargylt	yellowish	0.0134	5	9.80
kızylrak	red	0.0132	1	2.00
aksyl	off white	0.0122	5	10.80
altyn	gold	0.0113	3	7.00
kızgyltsarı	light blue	0.0105	4	10.00
yaşkelt	greenish	0.0103	4	10.25
kişersarı	orange	0.0091	3	8.67
apak	white	0.0088	1	3.00
kueyaşıl	dark green	0.0074	3	10.67
kuesarı	dark yellow	0.0062	2	8.50
syekzängär	dark blue	0.0058	2	9.00
cirän	auburn	0.0058	2	9.00
çıya	cherry red	0.0044	1	6.00
zängärräk	dark blue	0.0044	1	6.00
kömeş	silver	0.0038	1	7.00
zäpzängär	dark blue	0.0033	1	8.00
kararak	black	0.0029	1	9.00
köränsu	brown	0.0029	1	9.00
salattöse	light green	0.0029	1	9.00
kuezängär	dark dark blue	0.0026	1	10.00
sarırak	yellow	0.0024	1	11.00
şamäxäkörän	pink	0.0024	1	11.00
yaşelräk	green	0.0020	1	13.00
kügelcem	red	0.0020	1	13.00

Table A2: Russian colour terms elicited from Tatar-Russian bilingual respondents

Colour term	Gloss	Cognitive salience index	Total number of elicitations	Average list position
krasnyj	red	0.4737	36	2.00
černyj	black	0.2754	38	3.63
belyj	white	0.2006	36	4.72
želtyj	yellow	0.1668	30	4.73
sinij	dark blue	0.1664	29	4.59
zelenyj	green	0.1418	30	5.57
rozovyj	pink	0.0930	24	6.79
fioletovyj	purple	0.0909	25	7.24
oranževyj	orange	0.0861	24	7.33
goluboj	light blue	0.0764	21	7.24
koričnevij	brown	0.0559	17	8.00
seryj	grey	0.0532	17	8.41
beževyj	beige	0.0300	11	9.64
birjuzovyj	turquoise	0.0215	6	7.33
alyj	scarlet	0.0169	5	7.80
salatovyj	light green	0.0148	6	10.67
bordovyj	burgundy	0.0137	5	9.60
pudrovij	beige	0.0076	3	10.33
bolotnyj	mud brown	0.0055	2	9.50
lilovyj	violet	0.0050	2	10.50
fuksi	fuchsia	0.0038	1	7.00
purpurnyj	purple	0.0030	2	17.50
tsvet morskoi peny	turquoise	0.0029	1	9.00
tsvet xaki	khaki	0.0026	1	10.00
perlamutrovij	pearl	0.0024	1	11.00
malinovyj	crimson	0.0022	1	12.00
izumrudnyj	emerald	0.0020	1	13.00
višnevij	cherry red	0.0018	1	15.00
serebristij	silver	0.0015	1	17.00
kremovij	cream	0.0013	1	20.00
temno-krasnyj	dark red	0.0011	1	23.00
temno-zelenyj	dark green	0.0011	1	24.00
jarko-krasnyj	bright red	0.0010	1	26.00
jarko-sinij	bright dark blue	0.0010	1	27.00

Table A3: Russian colour terms elicited from monolingual Russian respondents

Colour term	Gloss	Cognitive salience index	Total number of elicitations	Average list position
krasnyj	red	0.3148	17	3.18
belyj	white	0.2576	17	3.88
černyj	black	0.2121	16	4.44
sinij	dark blue	0.2099	17	4.76
želtij	yellow	0.1518	17	6.59
zelenyj	green	0.1373	14	6.00
fioletovij	purple	0.1298	15	6.80
goluboj	light blue	0.1083	16	8.69
rozovij	pink	0.1058	13	7.23
oranževij	orange	0.0952	15	9.27
koričnevij	brown	0.0750	12	9.42
seryj	grey	0.0560	10	10.50
salatovij	light green	0.0390	7	10.57
sirenevij	lilac	0.0313	5	9.40
beževij	beige	0.0241	4	9.75
bordovij	burgundy	0.0224	4	10.50
birjuzovij	turquoise	0.0219	5	13.40
purpurnyj	purple	0.0156	3	11.33
izumrudnyj	emerald	0.0151	3	11.67
oxra	ochre	0.0138	2	8.50
mjatnyj	mint green	0.0131	2	9.00
lilovij	violet	0.0087	2	13.50
pastelnye ottenki	pastel colours	0.0074	1	8.00
indigo	indigo	0.0059	1	10.00
seroburomalinovij	purple	0.0059	1	10.00
bagrovij	crimson	0.0049	1	12.00
nefritovij	jade	0.0049	1	12.00
sapfir	sapphire	0.0049	1	12.00
fuksija	fuchsia	0.0049	1	12.00
xaki	khaki	0.0049	1	12.00
zolotoj	gold	0.0045	1	13.00
telesnyj	skin colour	0.0045	1	13.00
fuksi	fuchsia	0.0045	1	13.00
oksidxroma	chrome green	0.0042	1	14.00
ryžij	auburn	0.0035	1	17.00