

BASIC COLOUR TERMS IN FIVE FINNO-UGRIC LANGUAGES AND ESTONIAN SIGN LANGUAGE: A COMPARATIVE STUDY

Mari Uusküla¹, Liivi Hollman¹ and Urmas Sutrop^{1,2}

*¹Institute of the Estonian Language, Tallinn,
and the ²University of Tartu*

Abstract. In this paper we compare five Finno-Ugric languages – Estonian, Finnish, Hungarian, Udmurt and Komi-Zyrian – and the Estonian Sign Language (unclassified) in different aspects: established basic colour terms, the proportion of basic colour terms and different colour terms in the collected word-corpora, the cognitive salience index values in the list task and the number of dominant colour tiles in the colour naming task. The data was collected, using the field method of Davies and Corbett, from all languages under consideration, providing a distinctive foundation for linguistic comparison. We argue that Finno-Ugric languages seem to possess relatively large colour vocabularies, especially due to their rich variety of word-formation types, e.g. the composition of compound words. All of the languages under consideration have developed to Stage VI or VII, possessing 7 to 11 lexicalised basic colour terms. The cognitive salience index helps to distinguish primary and secondary basic colour terms, showing certain comprehensive patterns which are similar to Russian and English.

Keywords: colour vocabulary, basic colour term, comparative lexicology, field method, cognitive salience index, Estonian, Finnish, Hungarian, Komi-Zyrian, Udmurt, Estonian Sign Language

1. Introduction

Everyone seems to be interested in colour terms: linguists, poets, psychologists, philosophers, semioticians and even the average person. Linguists not only like them because colour terms form a well-defined and clear-cut semantic field and because there are good linguistic tool-kits available for research, but colour terms also form an extremely useful and simple model for examining the major contradictory theories in linguistics

and anthropology in the 20th century and the third millennium: relativism and universalism.

In 1969, after Brent Berlin and Paul Kay published their seminal work “Basic color terms: their universality and evolution”, it seemed that universalism had finally won and there was no place for relativism in linguistics and anthropology. The following years have shown that the universalists had only won a battle, but not the war, which formed a good basis for an ongoing discussion between relativists and universalists. Today most universalists accept some elements of weak relativism (e.g. see Regier et al. 2009).

The theory of Berlin and Kay is one of the most influential and discussed theories in linguistics and anthropology. A myriad of papers have been published in this tradition, covering numerous languages on every continent and in every language family. However, there have only been a few studies on colour terms in the Finno-Ugric languages until recently. Currently, most empirical and theory-driven study in the field of basic colour term research in the Finno-Ugric languages is carried out by the colour research group at the Institute of the Estonian Language in Tallinn, Estonia. Our research group has studied basic colour terms in the following Finno-Ugric languages: Estonian, Finnish, Hungarian, Udmurt and Komi-Zyrian. In addition, we have also done fieldwork to establish basic colour terms in the unclassified Estonian Sign Language and several Indo-European languages, e.g. Czech, Italian and Lithuanian, as well as in an Altaic language – Turkish (Indo-European and Altaic languages are not dealt with in this article).

Our colour research group works in the tradition of the Universalist school of basic colour term research established by Brent Berlin and Paul Kay in 1969. We use the same methodological framework, but a simplified field method (to collect data, developed by Ian Davies and Greville Corbett 1994, 1995). The field method is developed further and complemented with a cognitive salience index (Sutrop 2001).

Although there have been no profound fieldwork-based studies on Finno-Ugric colour terms until recently, some histori-

cal information on colour names in Finno-Ugric could be found in old dictionaries and all kinds of word lists. For example, the outstanding Czech scholar Jan Amos Komenský (Comenius) presented the Hungarian colour names *fejér* ‘white’, *fekete* ‘black’, *vörös* ‘red’, *sárga* ‘yellow’, *zöld* ‘green’, and *kék* ‘blue’ in the 1685 quadrilingual edition (Latin, German, Hungarian and Czech) of his pictorial dictionary *Orbis Sensualium Pictus* (The Visible World in Pictures). In 1786 Peter Simon Pallas listed, in the second volume of the *Vocabularia comparativa*, names for black, white, red and green in two hundred European and Asian languages (1786: 247-260). *Vocabularia* contains colour terms for the following Finno-Ugric languages and dialects: Hungarian, Čud (word lists for Vod, Ingrian or Ingrian-Finnish), Estonian, Karelian, Olonets (Livvi-Karelian), Kola-Sami, Komi-Zyrian, Komi-Permyak, Erzya-Mordvin, Moksha-Mordvin, Mari, Udmurt, different Mansi and Khanty dialects, and several different Samoyed languages and dialects. In a footnote, Pallas explained that all Samoyed people do not discriminate yellow and blue from green, and many other people name green and blue by using the same colour term (1786: 289). The Hungarian colour terms listed by Pallas were *fehir* ‘white’, *fekete* ‘black’, *veres* and *piros* (both ‘red’) and *zöld* ‘green’.

The only Finno-Ugric language mentioned by Berlin and Kay in their monograph in 1969 was Hungarian. Unfortunately, they were able to consult only one Hungarian subject living in the United States and some out-of-the-date Hungarian dictionaries as their sources. In consequence, the Hungarian results can be considered controversial: some words are misspelled and it is stated that there are two basic terms for red in Hungarian, *piros* and *vörös* (1969: 95). At present there is a rich literature on Hungarian colour terms focussed on two terms for red (see Uusküla 2008a, 2011a, and Uusküla and Sutrop 2010).

The first authors who dealt with Finno-Ugric languages, and who were inspired by the theory of Berlin and Kay (1969), were Michael Katzschmann and the Finnish linguist Mauno Koski. Katzschmann published a short note on the analysis of the Northern Mansi colour terms (1979). Koski published the

extensive monograph “The names for colours in Finnish and closely-related languages” (1983). Unfortunately, the data from the Finnic languages were not collected by using the empirical field method, but were instead drawn from dictionaries, corpora and dialect archives. The first empirical study on a Finno-Ugric language was done on Mansi colour terms by Katalin Sipőcz in 1994. She worked in the tradition of the theory of Berlin and Kay, using non-standard methods and colour stimuli to collect data.

In 1997 Robert E. McLaury and his colleagues published an empirical analysis of the Hungarian reds, also in the tradition of Berlin and Kay. According to the vantage theory founded by McLaury, they concluded that there was only one basic colour term for red in Hungarian: *piros* can be considered basic, whereas *vörös* is a non-basic colour term. As their task was to create their own vantage theory as a refinement of the original theory of Berlin and Kay, they tried to interpret the term *vörös* as a term for darker red than described by the basic term *piros*. Such an interpretation was suggested as a tendency by their empirical data. Unfortunately, their sample of native speakers was too small.

The rest of the literature on Finno-Ugric colour names includes some papers discussing the theory of Berlin and Kay to some extent (e.g. Futaky 1981), misinterpreting the theory (e.g. Turunen 2002) or following the theory partially without a clear manifestation of it (e.g. Rakin 1990).

The task of this article is to provide a lexical comparison of our empirical fieldwork-based data on Finno-Ugric languages published in earlier articles, monographs and dissertations (see for more detailed background information Uusküla and Sutrop, eds. 2011). First, we briefly describe the field methods we used to collect the data, also providing the background information on our subjects. Then we compare the colour term corpora collected in Estonian, Finnish, Hungarian, Udmurt, Komi-Zyrian and Estonian Sign Language. In the results and discussions section, we compare the number of basic colour terms in the studied languages, and discuss the cognitive salience index values

and rankings of the basic colour terms in each language and between languages. For a better understanding and lexical comparison with another language group (Indo-European), we use the data and results for Russian and English from the studies by Davies and Corbett (1994, 1995), also calculating the cognitive salience indices for these languages. Next, we compare the empirically collected Finno-Ugric colour-term corpora, presenting arguments regarding the richness of the colour vocabulary by calculating the percentage of basic terms in the corpora. Complexity of (compound) colour terms will be calculated by the complexity index. After that, we explain the dominance of the basic colour terms and the consensus levels of the subjects. Finally, we draw some general conclusions based on our findings.

2. Method

The basic colour terms survey consisted of two tasks and a colour vision test. The methodological and theoretical procedure of Ian Davies and Greville Corbett, i.e. their field method, was used to collect and analyse the data (Davies and Corbett 1994, 1995, and Corbett and Davies 1995).

The list task. The subjects were asked to name as many colours as they knew. All terms were written down in the order in which the subjects listed them; in the case of the Estonian Sign Language, the answers were video-recorded. The list task was complemented with a cognitive salience index (Sutrop 2001).

Colour vision test. After the list task, the colour vision of all participants was examined by using the City University colour vision test (Fletcher 1980, 1998).

The colour-naming task. This involved showing the subjects 65 colour squares, one square at a time, in a random sequence. The colours were shown indoors in sufficient daylight on a grey base. Each subject was asked to name the colours of the square tiles.

Stimuli. In the colour-naming task, we used the standard set of 65 tiles selected from the Color-Aid range of coloured pa-

pers by Ian Davies and his colleagues (1992). Each tile was a 5x5 cm rigid wooden or cardboard square covered with coloured paper; for more details and colour coordinates, see Sutrop (2002: 59-61).

Participants. Participants from Estonia, Finland and Hungary, and from the Udmurt and Komi-Zyrian Republics of the Russian Federation were native speakers of the titular languages. In Estonia, signers of Estonian Sign Language also participated. The sample sizes, the composition of the sample by sex and by age, and mean age (in years) are shown in Table 1. All samples were selected to be as heterogeneous as possible. The interviewers selected their subjects with great care, inviting as many people with different backgrounds as possible to take part in the interviews (see Section 3). The subjects were told that the interviews and questions would be about language and fairly easy to answer. The topic of colours was not mentioned before the list task to avoid any priming effects. Most of the subjects completed the three tasks at home, at work or in public libraries; university students and schoolchildren were questioned at

Table 1. The composition of a sample by language.

Language	Language classification	Total number of subjects	Females/ Males	Age range	Mean age
Estonian	Finno-Ugric, Finnic	80	53/27	16-72	29
Finnish	Finno-Ugric, Finnic	68	42/26	10-75	39.5
Hungarian	Finno-Ugric, Ugric	125	66/59	9-82	36
Udmurt	Finno-Ugric, Permic	125	76/49	9-80	43.4
Komi-Zyrian	Finno-Ugric, Permic	51	37/11	11-81	49.4
Estonian Sign Language	Unclassified	50	26/24	15-74	43

the university or at school, and deaf subjects were mostly interviewed at clubs for the deaf. The results of the subjects who had colour vision deficiencies were excluded from the following analysis. All subjects participated on a voluntary basis.

Interviewers. Estonian speakers were tested by a native speaker (Sutrop). Finnish and Hungarian speakers were tested by a fluent Finnish and Hungarian speaker (Uusküla). Udmurt speakers were tested by a native speaker (Elena Ryabina). Komi-Zyrian speakers were tested by a fluent Komi-Zyrian speaker (Ryabina). Estonian Sign Language results were gathered by a hearing, fluent Estonian Sign Language signer (Hollman) and recorded by a deaf cameraman.

3. The subjects: background description

A questionnaire with some personal data was filled in by the interviewer usually immediately before the beginning of the list task, or after all the tasks were completed. Personal data included the sex and age, place of birth and place of residency of the subject; we also asked about the subject's education and professional work, and whether our subjects knew any dialects of their mother tongue (sometimes the interviewer also asked about the knowledge of foreign languages, e.g. the interviewer ascertained whether a Finnish subject was a bilingual Finnish-Swedish speaker or a monolingual Finnish speaker and whether s/he was able to speak any other foreign language, such as English). The questionnaire for the deaf ESL users also included information about the hearing status of their parents and siblings and access to sign language in early childhood. No personal names were added to the data on confidentiality purposes.

Estonian. The Estonian data was collected from 19 March to 5 April 1995 in Tallinn and Tartu, Estonia (Sutrop 1995: 799). Eighty subjects took part in the experiment, 53 women and 27 men, whose age ranged from 9 to 72 years (the mean age was 29 years) (Sutrop 2002: 57). All subjects had different dialect and professional backgrounds. Two interviewees were Estonian-

Russian bilinguals. There were students, biologists, linguists, writers, scientists, retired people etc. The subjects were born or raised in the following locations in Estonia: Kohtla-Järve, Alutaguse, Võru County, Läänemaa, Pärnu, Tartu and Tallinn (Sutrop 1995: 799).

Finnish. The Finnish data was collected in two stages: from 1-7 August 2005 in Helsinki, Turku and Lempäälä, and 7-14 September 2006 in Helsinki, Espoo and Tuusula. Sixty-eight participants took part in the list and colour-naming tasks, 42 female and 26 male subjects. In fact, there were 69 subjects, but the answers given by one subject could not be considered as the interviewer detected some colour vision anomalies in the colour vision test (Fletcher 1980) (Uusküla 2007: 371). In 2005 twenty-nine subjects took part in the interviews, and in 2006 there were 39 subjects in total. The subjects were either born or raised in different locations in Finland (see Uusküla 2007: 371), and had different dialect backgrounds; one subject was Finnish-Finnish Sign Language bilingual.

Hungarian. The Hungarian data was also collected in two stages: 22-31 October 2002 in Budapest and from 19 October to 5 November 2003 in Budapest, Balassagyarmat, Győr, Debrecen, Ipolyvece, Dejtár, Pécs and Budaörs. The total number of subjects was 125, among them 66 women and 59 men, whose age ranged from 9 to 82 years (the mean age was 36 years). The subjects had different occupations: there were students, engineers, medical doctors, school teachers, hair dressers, machinists, etc. Twenty-eight people claimed to be able to speak one of the Hungarian dialects, among them Székely, West Hungarian, North-east Hungarian, West Danube, Danube-Tisza and other dialects, but they used Standard Hungarian during the interview. All interviewees considered Hungarian to be their first language; one person was Hungarian-Finnish bilingual (Bogatkin 2005: 57-58).

Udmurt. The Udmurt data was collected by Elena Ryabina in 2007 and 2008 in the Republic of Udmurtia, the Agryzsky District of the Republic of Tatarstan, and the Tatyshlinsky District of the Republic of Bashkortostan (Ryabina 2011c: 195). The

total number of subjects was 125, among them 76 female and 49 male, whose ages ranged from 9-80 years (the mean age was 43 years). All subjects were first language Udmurt speakers, only a few elderly subjects were monolingual Udmurt speakers, and the others spoke Russian as their second language fluently. Ryabina divided the Udmurt subjects into three groups according to dialect differences: the Southern Udmurts, the Northern Udmurts and the speakers of a dialect from the area beyond the Kama River. A more detailed description of the Udmurt sample is provided in Ryabina (2011c: 37-39).

Komi-Zyrian. The Komi-Zyrian data was also collected by Elena Ryabina in 2008 and 2009 in Syktyvkar, Vizinga and the Kortkerossky and Koygorodsky Districts of the Republic of Komi. The total number of Komi-Zyrian subjects was 51, among them 37 women and 14 men, between 11 and 81 years of age (the mean age was 49 years) (Ryabina 2011b: 195). They were all native Komi-Zyrian speakers, and most of them had very good knowledge of Russian. The data on Komi-Zyrian subjects can be found in Ryabina (2011c: 39-40).

Estonian Sign Language. The data on ESL was collected during summer 2005. A total of 50 subjects from Tallinn, Tartu, Pärnu and Võru were interviewed, among them 24 men and 26 women, between the ages of 15 and 74, with an average age of 43 years. Most of the subjects (36 people) had studied in the Porkuni Deaf School, the main school for the deaf in Estonia for decades; twelve had graduated from the Tartu Hiie School (an oral school for deaf and hard of hearing students) and one had attended a local mainstream school. Thirty-eight subjects were profoundly deaf and twelve had some residual hearing. Although all the subjects used ESL during the survey, only eight were from deaf families and had access to ESL in early childhood. As is characteristic of deaf communities, most of the subjects (38) were from hearing families with no deaf parents or siblings, and had started to learn ESL between two to fourteen years of age (Hollman 2010: 58-60).

4. Data: the collected word-corpora

The number of collected colour terms in the list and colour-naming tasks is presented in Table 2. The sum of all collected terms, as well as the number of different colour terms, was counted for both tasks separately and in the whole colour term survey. The fourth column represents the mean length of the list in the first task.

Table 2. The number and distribution of the collected colour-word corpora by language. The number of subjects per language is indicated in brackets.

Language	List task			Colour naming task		Total	
	Sum	Dif-ferent terms	Mean length of the list	Sum	Dif-ferent terms	Sum	Dif-ferent terms
Estonian (80)	1,515	285	18.94	5,197	638	6,712	759
Finnish (68)	1,506	332	22.15	4,370	855	5,876	1,014
Hungarian (125)	2,851	413	22.81	8,066	990	10,917	1,148
Udmurt (125)	2,122	278	16.98	7,915	1,138	10,037	1,231
Komi-Zyrian (51)	612	109	12.00	3,121	481	3,733	514
Estonian Sign Language (50)	681	109	13.62	3,374	696	4,055	718

Estonian. The Estonian word-corpus consisted of 6,712 colour words in total, of which 759 were different. All morphologically modified terms and phonetic variants were counted separately. In the list task, the subjects were able to list 1,515 colour terms, among these 285 different ones. The average length of the list was 18.94 colour terms. In the list task, 80 subjects named 5,197 colour terms, of which 638 were different. There were only three occasions when the subjects could not find an appropriate name for a colour tile. Of the 285 different colour

terms in the list task, 121 were not mentioned in the colour-naming task (Sutrop 2002: 67). However, the subjects used 474 new colour terms to name colour stimuli which they had not used in the first task.

Finnish. The collected data for Finnish resulted in a word-corpus containing 5,876 colour terms and 1,014 different entries. In the list task, the subjects gave 1,506 colour names, including 332 different ones. The average number of named colours was 22.15. The lowest number of named colour terms in the list task was eight, offered by a 75-year-old farmer in Lempäälä, while one 68-year-old retired woman (also from Lempäälä) was able to list 53 different colour terms (Uusküla 2007: 373). In the colour-naming task, the subjects named the 65 coloured tiles in 4,370 ways, among them 855 terms were different (Uusküla 2007: 378). Usually male subjects experienced greater colour-naming difficulties than female subjects did. One 40-year-old house painter from Espoo was unable to name eight stimuli, while another 55-year-old house painter from Lempäälä did not know how to name six colour tiles. We can conjecture that they probably knew the technical code of a stimulus, but were unable to give appropriate names to the tiles. They usually described the colour tile they saw with an expression, such as “this is a mixture of blue and green blended with some black”, and then gave up, saying “but I really don’t know the name of this colour”.

Hungarian. The total number of responses in the Hungarian word-corpus was 10,917 colour words: the list task data included 2,851 colour terms, among these 413 different ones, while the total number of terms in the colour-naming task was 8,066 out of 8,125 possible answers (125×65), and among these there were 990 different terms (Bogatkin 2005). In the list task, the average list consisted of 22.81 entries. The lowest number of listed terms was nine, offered by a 42-year-old cleaning lady in Győr, who had attended school for eight years; the highest number of colour terms in the list task was 61, offered by a 38-year-old female gymnasium teacher. In the colour-naming task, there were 59 occasions where the subjects were unable to give appropriate names for colour tiles (the colour tile with the

Color-aid code ORO T3 was unnamed twelve times, the tiles RO T3 and ROR S3 were not named six times etc.). Generally, male subjects experienced greater difficulties with colour-naming than women did. One 45-year-old businessman and one 23-year-old male student did not know how to name seven colour tiles, while one 67-year-old retired man and one 26-year-old male historian did not name four colour tiles. In comparison, only one female subject, a 50-year-old librarian, had colour-naming difficulties, leaving seven colour tiles unidentified (Bogatkin 2005: 62-65). None of the other subjects had significant colour-naming difficulties.

Udmurt. The Udmurt subjects used a total of 10,037 colour terms, among these 1,231 different ones. All morphological variants, such as *gord* ‘red’ and *gordales* ‘reddish’, were counted as different terms, whereas the Udmurt colour term survey did not distinguish phonetic variants. In the list task, the total number of responses across subjects was 2,121, of which 278 were different colour terms, while the mean number of listed colour terms was 16.9. Commonly, Southern and Kama River dialect speakers were able to list more colour terms (the average number of terms was 17.9) than the speakers of the Northern dialect (the average number of terms was 14). There was one subject who could only name nine colour terms: a 49-year-old cleaning lady; all other subjects named at least 10 colour terms in the list task. In the colour-naming task, 52 subjects did not know appropriate names for 65 colour tiles on 172 occasions.

Komi-Zyrian. The Komi-Zyrian subjects named a total of 3,733 colour words, of which 718 were different. The total number of colour terms in the list task was 612, among them 109 different ones. The subjects listed 12 colour terms on average, while the shortest list, that of an 11-year-old pupil, contained only five colour terms. At least three further subjects were able to list only six colour terms each. The longest list contained 34 colour terms, named by a 45-year-old female student. In the colour-naming task, the subjects named 3,121 of 3,315 possible colour terms (65 x 51). Thirty-nine subjects faced some colour-naming difficulties, leaving 194 colour tiles unnamed. They

were mainly elderly women; one 60-year-old pensioner was not able to name 17 colour tiles (Ryabina 2011c: 75).

Estonian Sign Language. In the ESL colour terms survey, the subjects named a total of 681 colour terms in the list task and 3,374 colour terms in the colour-naming task. The list task data included 109 different colour signs and the data from the colour-naming task included 696 different terms. The average number was 13.62 terms; subjects from deaf families and from hearing families with early access to ESL listed more terms than the average. The shortest list was given by a 73-year-old man from a hearing family who probably did not learn sign language before going to school. In the colour-naming task, he mostly used Estonian words instead of ESL signs, showing that Estonian had probably remained his first language, although he used ESL to communicate in the deaf community. The longest lists were given by a female tailor from a hearing family but with early access to sign language, and a male artist, also from a hearing family. Of the 109 different signs named in the list task, 51 signs were mentioned only once, and 58 were named at least twice. In the colour-naming task, more compounds were used (the average being 1.79 signs) and only 225 signs out of 696 were mentioned at least twice. In ten cases out of 3,250 (50 subjects x 65 colour tiles), the subjects were not able to name the tiles (Hollman 2010: 60-64, 75). These difficulties were mainly experienced by female subjects from hearing, non-signing families.

5. Results and discussion

The data was collected using the same field method for all languages under examination and can thus be easily compared. We will compare the results within the Finno-Ugric language group and partly with English (Davies and Corbett 1995) and Russian (Davies and Corbett 1994).

5.1. The basic colour terms in the Finno-Ugric languages and Estonian Sign Language

The number and inventory of basic colour terms in the Finno-Ugric languages have been established in our previous studies (Sutrop 1995, 2000, 2002, Hollman 2010, Hollman and Sutrop 2011, Uusküla 2007, 2008a, Sutrop and Bogatkin-Uusküla 2005a, 2005b, and Uusküla and Sutrop 2007, 2010). The results of Udmurt and Komi-Zyrian have been taken from the studies by Elena Ryabina (2011a, 2011b, 2011c). Due to significant dialect distinctions in the Udmurt language, Southern (SU) and Northern (NU) dialects have developed different basic colour terms (see Table 3). In order to establish the basic colour terms, we have adopted the analysis introduced by Ian Davies and his colleagues (Davies et al. 1992), which they also applied in their numerous other studies (Davies and Corbett 1994, 1995, Davies et al. 1995, etc.) In this article, we will only comment on the cognitive salience index in the list task – an important addition made by Sutrop (2001) – which underlines the psychological salience of a colour term, and the dominance index in the colour-naming task, which tends to show the consensus of the language speakers. These two indices are constituted to reflect the most important, fourth criterion of the Berlin and Kay basic colour term criteria concerning the psychological salience of a colour term. The cognitive salience index in the list task shows whether a colour term has a tendency to emerge at the beginning of the elicited lists in combination with the appearance in the idiolects of all language speakers, while the dominance index in the colour-naming task emphasises the stability of reference across speakers and across occasions of use (Berlin and Kay 1969: 6-7).

Table 3 shows the inventory of basic colour terms in five Finno-Ugric languages and Estonian Sign Language, according to the original Berlin and Kay basic colour term evolutionary order. The number of basic colour terms is separately shown in the last row. The basic colour terms have been established relying on the results of the two tasks, the list task and the colour-nam-

Table 3. The list and number of basic colour terms (BCTs) in Estonian, Finnish, Hungarian, Komi-Zyrian, Udmurt, Komi-Zyrian and Estonian Sign Language (ESL). SU – Southern Udmurt, NU – Northern Udmurt.

Berlin and Kay original order	Estonian	Finnish	Hungarian	Udmurt	Komi-Zyrian	ESL
black	must	musta	fekete	şed	şed	BLACK
white	valge	valkoinen	fehér	teđj	ježđj	WHITE
red	punane	punainen	piros	gord	gerd	RED
yellow	kollane	keltainen	sárga	čuž	kólkviž	YELLOW
green	roheline	vihreä	zöld	vož	turunviž	GREEN
blue	sinine	sininen	kék	lijz/ čágyr (NU)	lez	BLUE
brown	pruun	ruskea	barna	kureń (SU)	-	BROWN
grey	hall	harmaa	szürke	purjś	rud	GREY
pink	roosa	vaalean-punainen	rózsa-szín	łemłet' (SU)	-	PINK/PURPLE
purple	lilla	-	lila	-	-	-
orange	oranž	oranssi	narancs-sárga	nap-čuž (SU)	-	-
Number of BCTs	11	10	11	10 (SU)/ 8 (NU)	7	9

ing task, and using the combined analysis according to Davies and Corbett (1994, see also Sutrop 2002: 87, Bogatkin 2005: 71, Uusküla 2007: 388, and Ryabina 2011c: 68-69, 82).

The eleven basic colour terms in Estonian are *must* 'black', *valge* 'white', *punane* 'red', *kollane* 'yellow', *roheline* 'green', *sinine* 'blue', *pruun* 'brown', *hall* 'grey', *roosa* 'pink', *lilla* 'purple' and *oranž* 'orange' (Sutrop 1995, 2000, 2002). Finnish possesses ten basic colour terms: *musta* 'black', *valkoinen* 'white', *punainen* 'red', *keltainen* 'yellow', *vihreä* 'green', *sininen* 'blue', *ruskea* 'brown', *harmaa* 'grey', *vaaleanpunainen* 'pink', and *oranssi* 'orange' (Uusküla 2007, 2008). Finnish has not evolved a basic colour term for purple, as several words can express this colour: *violetti*, *liila*, *lila*, *sinipunainen*, etc. (Uusküla 2007: 392, Kerttula 2007). However, it might be worth carrying out some additional fieldwork in Finland to examine recent changes in the development of the Finnish colour vocabulary, as the slot for purple may have been filled by now. According to the collected data, we predict that this colour term might be *violetti*.

The inventory of basic colour terms in Hungarian consists of eleven and not twelve basic colour terms, as suggested by Berlin and Kay (1969: 95, for discussion see MacLaury et al. 1997, Forbes and Kiss 1999, Kiss and Forbes 2001, Bogatkin-Uusküla and Sutrop 2005a, 2005b, Uusküla and Sutrop 2010, Bogatkin 2005, and Uusküla 2011a). These colour terms are: *fekete* 'black', *fehér* 'white', *piros* 'red', *sárga* 'yellow', *zöld* 'green', *kék* 'blue', *barna* 'brown', *szürke* 'grey', *rózsaszín* 'pink', literally 'rose colour(ed)', *lila* 'purple' and *narancssárga* 'orange', literally 'orange yellow' (Bogatkin 2005, Bogatkin-Uusküla and Sutrop 2005a, Uusküla and Sutrop 2010, Uusküla 2011a).

The basic colour terms in Udmurt have been established by Elena Ryabina, who divided her subjects into three dialect subgroups (see Section 3): Southern Udmurts, Northern Udmurts and speakers of a dialect from the area beyond the Kama River. Ryabina claims that this distribution is justifiable as the colour vocabulary, and therefore the set of basic colour terms, depends on the spoken dialect. According to Ryabina, the number of basic colour terms in Southern Udmurt is ten – *šed'* 'black', *teđi*

‘white’, *gord* ‘red’, *čuž* ‘yellow’, *vož* ‘green’, *liz* ‘blue’, *puris* ‘grey’, *lemlet* ‘pink’, and *nap-čuž* ‘orange’ – while in Northern Udmurt there are only eight lexicalised basic colour terms – *šed* ‘black’, *teđj* ‘white’, *gord* ‘red’, *čuž* ‘yellow’, *vož* ‘green’, *liz* ‘blue’, *puris* ‘grey’, and *čagir* ‘light blue’ (Ryabina 2011a, 2011c). Although *lel* ‘pink’ was cognitively salient in the list task, it failed to achieve the 50% dominance in the colour-naming task, and has therefore been excluded from the basic colour term list. Interestingly, *čagir* ‘light blue’ gained dominance in the colour naming task, as over 50% of the Northern Udmurt subjects described the colour of tile BGB T3 as *čagir* ‘light blue’. For detailed analysis, see Ryabina (2011a, 2011c: 86-87).

According to the fieldwork conducted by Ryabina, Komi-Zyrian possesses seven basic colour terms: *šed* ‘black’, *ježid* ‘white’, *gerd* ‘red’, *kolkviz* ‘yellow; egg yolk’, *turunviž* ‘green’, *lez* ‘blue’, and *rud* ‘grey’ (Ryabina 2011c: 72-83). Yellow and green are not well distinguished in the Komi-Zyrian language (Ryabina 2011b: 209), probably due to its historical development. Etymologically, Komi-Zyrian yellow and green both originate from the same word stem and the meaning of *viž* can be either yellow or green (Wichmann 1942: 338).

Estonian, Finnish, Hungarian and Udmurt have all reached the highest stage, Stage VII, according to the basic colour term evolutionary order (Berlin and Kay 1969: 4), whereas Komi-Zyrian, having derived seven lexicalised basic colour terms, is a language at Stage VI, with an additional lexicalised basic term for grey. According to Berlin and Kay, grey can evolve “as a wild-card” at any stage starting from Stage IV (Berlin and Kay 1969: 45). Kay and McDaniel revised the evolutionary scheme by suggesting that grey might occur as a separate category from Stage IIIa onwards, or even earlier (1978: 639-640).

Estonian Sign Language possesses nine basic colour terms, according to the data – BLACK, WHITE, RED, GREEN, YELLOW, BLUE, BROWN, GREY and PINK/PURPLE – being a Stage VII language (Hollman 2010: 109). The sign PINK/PURPLE covered the area of both *pink* and *purple*, being more frequent in cases of brighter variants and in the red area of the

colour space, while another sign PURPLE, not having a basic term status in ESL according to the survey because of its low frequency, appeared in the transition area of red and violet and violet and blue (Hollman 2010: 101). There was a great variety of signs used to denote *orange* in the list task: seven different signs were named more than twice, and in the colour-naming task eight different signs were used, most of them occurring only once for one tile (Hollman 2010: 80). All signs for basic colour terms in ESL are shown in Appendix 1.

5.2. The cognitive salience index as a simple measure for ranking basic colour terms

Corbett and Davies suggested that some basic colour terms can have higher status in the Berlin and Kay evolutionary hierarchy and can therefore be considered “more basic” than the other basic colour terms (1995: 301). They applied different measures and techniques to their data to prove the hypothesis, emphasising their different functions. Although we applied most of the measures to our data in our previous studies, e.g. frequency of occurrence in the list and in the colour-naming tasks, consistency of use across occasions and across subjects, etc., we have found a relatively simpler way to establish basic colour terms: we rely on the value of the cognitive salience index. The cognitive salience index was designed by Urmas Sutrop by unifying two parameters in the list task: naming frequency and mean position of a term (Sutrop 2001). If a colour term has been named frequently and listed towards the beginning of the elicited list, it has a better chance of competing for basic term status. The cognitive salience index can be used to distinguish between basic colour terms and non-basic ones. Moreover, it is also a useful measure for ranking basic colour terms according to their basicness, and it tends to discriminate “primary” and “secondary” basic colour terms.

Figure 1 represents the value of the cognitive salience index for eleven (basic) colour terms: red, blue, yellow, green,

white, black, brown, grey, purple, pink and orange, in Estonian, Finnish, Hungarian, Udmurt, Komi-Zyrian and Estonian Sign Language. Not all languages possess the full set of basic colour terms (see Table 3). Figure 1 explicates the threshold for the terms to obtain the basic colour status.

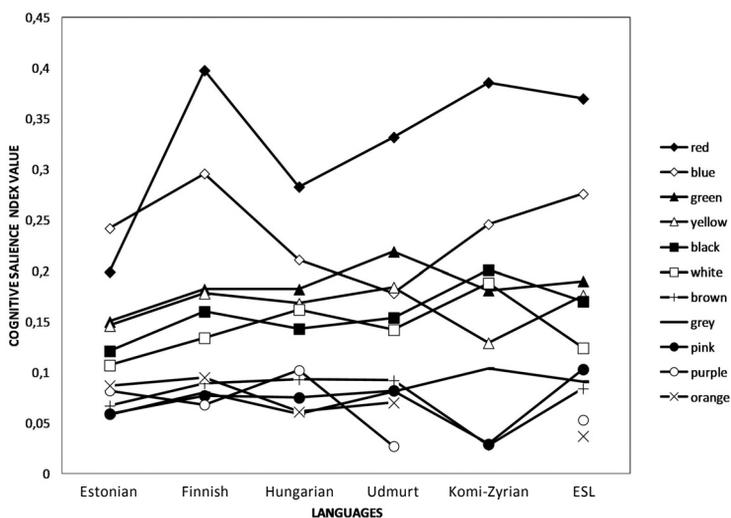


Figure 1. The salience of 11 (basic) colour terms in five Finno-Ugric languages and ESL.

The figure clearly illustrates the usefulness of the cognitive salience index in distinguishing the primary and secondary basic colour terms: values of the primary terms always score higher. We have not included any further colour terms in the figure on purpose while the values of colour terms only drop. Figure 1 displays the importance of the colour term red: the most salient colour term in four Finno-Ugric languages and ESL. However, Estonian differs from the other languages by having blue as the most salient term. While red and blue are the most salient basic colour terms according to the cognitive salience index, other basic terms also follow a certain pattern: yellow and green, and black and white tend to appear together, with a slight

difference between their values, and brown, pink, purple, orange and grey have an inclination to occur together with relatively lower index values than all previously named colour terms. The appearance of brown, pink, purple, orange and grey after red, blue, green, yellow, black and white accords well with the original basic colour term evolution hierarchy established by Berlin and Kay (1969). The greater difference with the evolutionary ordering hierarchy is that the cognitive salience index in most languages has the two achromatic colours, black and white, falling after red, blue, yellow or green, with some exceptions: Udmurt, Komi-Zyrian and ESL. Komi-Zyrian possesses two competitive colour terms, *viž* ‘yellow’ and *kol’kviž* ‘yellow, egg yolk’, to fill the gap for yellow. They were listed almost with equal frequency, but *viž* tended to be named towards the beginning of the list and therefore has a higher mean position and higher cognitive salience index values. (For further discussion see Section 5.3.) Udmurt fails to follow the pattern because it possesses a second term for blue, the highly salient Northern Udmurt colour term *čagir* ‘light blue’ (Ryabina 2011a), and therefore the cognitive salience index value for basic blue *līz* drops significantly. Estonian Sign Language tends to follow the pattern perfectly, with one slight exception: the values of black and yellow coincide and fall in the same place.

This means that the evolutionary order, i.e. the order of the emerging of the basic colour terms is different from their synchronic salience. In general, bright saturated colours come to subjects’ mind before black and white in the elicited lists. The cognitive salience index of the first six basic colour terms in five Finno-Ugric languages and Estonian Sign Language is shown in Figure 2.

For the calculation of the cognitive salience index for Russian, the data from Davies and Corbett (1994), and for the calculating index for English, the data from Davies and Corbett (1995) was applied. The analysis of Russian reveals exactly the same pattern, where primary basic colour terms have higher scores than the secondary basic colour terms and the two achromatic colours, with only one specification: black and white tend

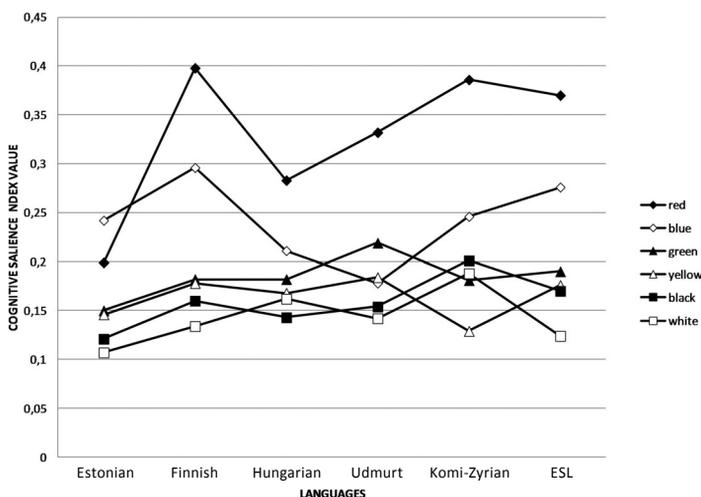


Figure 2. The cognitive saliency index of the “primary” basic colour terms in five Finno-Ugric languages and ESL.

to have lower values on the mean position level than the secondary basic colour terms for purple and orange (Davies and Corbett 1994: 73-75), which means that they are named before black and white. This, in turn, leads to lower cognitive saliency index values for black and white. Consequently, black and white have lower saliency index values than purple and orange, and Russian tends to show a slight exception to the pattern, fulfilling other requirements.

The pattern is better illustrated in Figure 3, which represents the cognitive saliency curve of English and Russian. Both languages roughly follow the same pattern as Finno-Ugric languages. According to the study by Davies and Corbett, Russian has two basic colour terms to express blue: *sinij* and *goluboj* (1994) affirmed in numerous further studies (see Paramei 2005). We have used only one basic colour term for blue in Russian, *sinij*, for the cognitive saliency index calculation. Therefore, it is no surprise that Russian blue has a lower cognitive saliency index value. This might explain why *sinij* scores somewhat lower than terms for blue in other languages (see Figure 3). The case is

comparable to Komi-Zyrian yellow and Hungarian red: all languages with a category split have lower cognitive salience index values for both colour terms expressing the same hue.

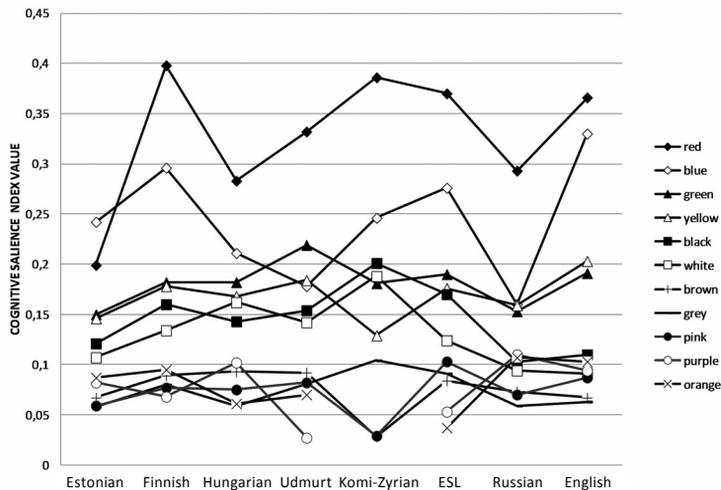


Figure 3. The cognitive salience index of 11 (basic) colour terms in five Finno-Ugric languages, ESL, Russian and English.

5.3. The composition of word-corpora and the richness of colour language in the Finno-Ugric languages

Although the assignment of basic colour terms in a language is interesting and important, basic colour terms form only a small part of colour vocabulary (Sutrop 2011: 47). Our data was collected in order to establish the basic colour terms in a language and therefore the distribution of basic and non-basic colour terms should be in favour of basic colour terms. However, the basic colour terms form a smaller part of the whole collected corpus than hypothesised: in the five Finno-Ugric languages studied and in Estonian Sign Language, the proportion

of basic colour terms is lower than 45%. Figure 4 presents the composition of colour-word corpora in Estonian, Finnish, Hungarian, Udmurt, Komi-Zyrian and ESL, with emphasis on the share of basic and non-basic colour terms. The dark grey parts of the columns indicate the percentage of basic colour terms. Here we have counted only “pure” basic colour terms: if a colour term has been modified or compounded with the basic colour term stem, it is left out of the analysis. The set and number of basic colour terms is represented in Table 3 and discussed in Section 5.1. The exact sizes of colour-word corpora are the following: the Estonian corpus contains a total of 6,712 colour terms, of which 2,922 are basic colour terms (44%) and 3,790 are non-basic colour terms (56%); the Finnish corpus has 5,876 colour terms in all, 2,016 basic colour terms (34%) and 3,860 non-basic colour terms (66%); the Hungarian corpus consists of a total of 10,917 colour words, of which 3,820 are basic (35%) and 7,097 (65%) non-basic; the Udmurt corpus has 10,037 colour

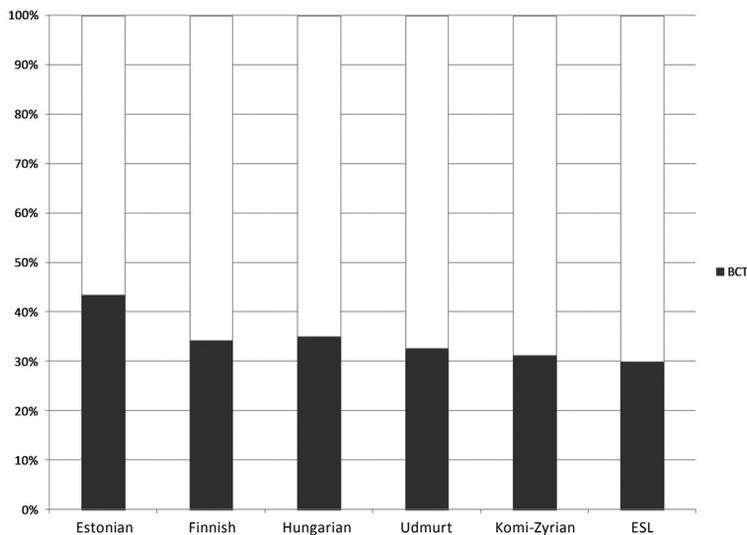


Figure 4. The proportion of basic colour terms and non-basic colour terms in five Finno-Ugric languages and ESL in the list task and colour-naming task combined.

words in all, 3,274 (33%) basic terms and 6,763 (67%) non-basic terms; Komi-Zyrian corpus has 3,733 colour terms in all, 1,165 basic terms (31%) and 2,568 (69%) non-basic terms; and the corpus of Estonian Sign Language consists of a total of 4,055 colour signs, 1,216 basic signs (30%) and 2,839 non-basic signs (70%). The proportion of basic colour terms in the collected data stays between 30% and 44%.

Figure 5 shows the proportion of basic colour terms in the list task, and Figure 6 shows the share of basic colour terms in the colour-naming task. These two figures show that fewer basic colour terms occur in the colour-naming task, probably because there are different ways to name the presented stimuli: with a basic colour term, with a compound, with a modified term, with a special non-basic colour term, with a descriptive phrase, etc. The proportion of basic colour terms in the colour-naming task is relatively lower in Udmurt, Komi-Zyrian and Estonian Sign Language than in the total word-corpora, which may be caused by the smaller number of basic colour terms in these languages (for comparison, see Figure 4 and Figure 6). This statement is logical, while the more basic colour terms language possesses,

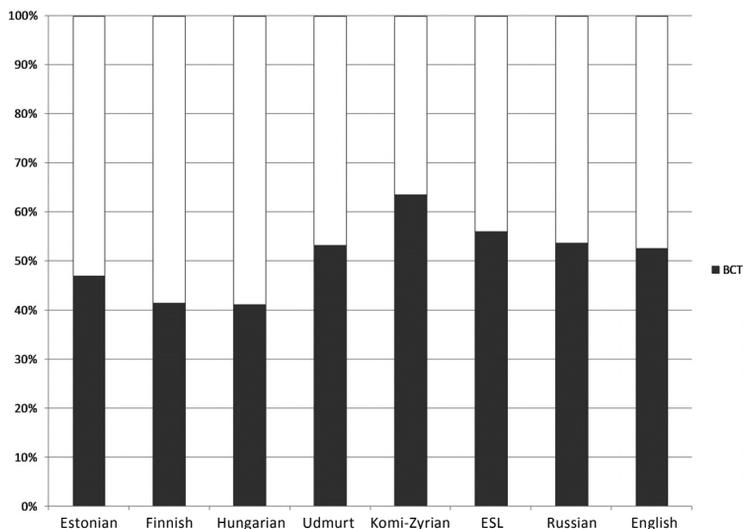


Figure 5. The proportion of basic colour terms in the list task.

the more occurrences of basic colour terms could be ascertained in the corpus.

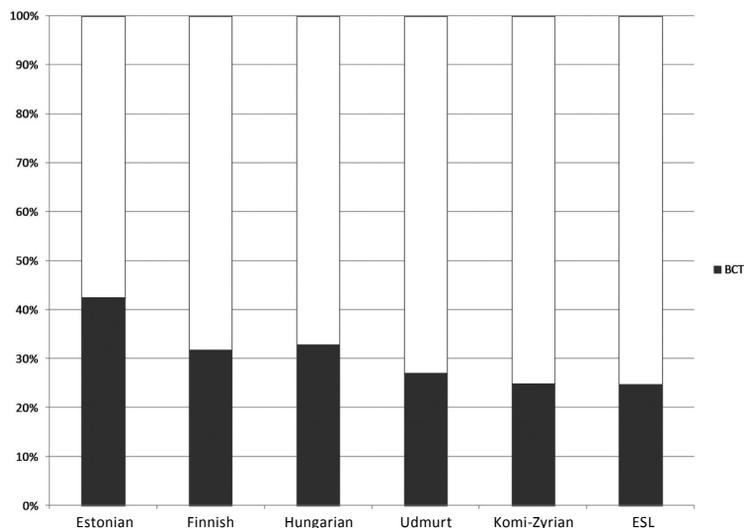


Figure 6. The proportion of basic colour terms in the colour-naming task.

Barratt and Kontra compared the colour vocabulary of American English and Hungarian and found that “the use of *sötét/dark* and *világos/light* among our subjects revealed that the Hungarian terms are used frequently while American English relies either on the basic color terms alone (for darker shades) or on other non-basic color words (for lighter shades)” (1996: 113). Finno-Ugric languages apparently possess only a few specific non-basic colour terms and these are used less frequently than in the Indo-European languages (for comparison, see Czech data in Uusküla 2008b). In contrast, Finno-Ugric languages use compound colour terms, as also noticed by Barratt and Kontra. In fact, there are numerous possibilities of expressing colour nuances with detailed descriptions, referring to the intensity, lightness, darkness or saturation of a colour. These issues are discussed in-depth in future papers. The rich variety of word-

formation forms make it possible to express all these variations grammatically.

The number of compound words in our word-corpora is relatively high, especially if we compare the results with the data of Russian or English collected using exactly the same field method (Davies and Corbett 1994, 1995). We have also calculated the complexity index, which shows how many parts the colour terms in the corpora consist of on average. Sutrop describes the complexity index as follows:

“C.I. is a complexity index. $C.I. = P/N$, where P is the number of simple words in terms (compounds) and N is the number of terms, e.g. if we have three terms: green, red, yellow, then $P = N$ and $C.I. = 3/3 = 1$, but if we have yellow-green, red, green, then $P = 4$, and $N = 3$; $C.I. = 4/3 = 1.33$. The complexity index ignores morphological complexity, e.g. yellowy or bluish” (Sutrop 2002: 63).

Table 4 shows the complexity index values in the five Finno-Ugric languages and Estonian Sign Language. On average, the complexity index values fall between 1.2 and 1.51 in the list task and are significantly higher in the colour-naming task. In this table, the complexity index has been counted for all given answers, not different words. The study of Estonian colour terms by Sutrop shows that the complexity index has a tendency to rise if counted for different words (2002: 133). The sizes of the word corpora are presented in Table 2.

Table 4. The complexity index values in the list task, colour-naming task and the two tasks combined in five Finno-Ugric languages and ESL.

Task	Estonian	Finnish	Hungarian	Udmurt	Komi-Zyrian	ESL
List	1.40	1.42	1.51	1.29	1.25	1.2
Colour-naming	1.59	1.64	1.77	1.66	1.55	1.8
Combined	1.55	1.58	1.70	1.59	1.50	1.7

Table 4 indicates that colour terms in the Finno-Ugric languages tend to be compound words rather than simple words or, more precisely, compound words dominate over simple words in our surveys. By a compound, we mean all words that are made up of more than one word stem; some of these should instead be called word phrases as the spelling (one word or more) depends on the local orthographic tradition of a language. To illustrate this point, we will offer some numerical totals and examples. All examples have a form word1-word2-word3 etc. to facilitate reading. These words might be spelled differently in accordance with the local orthographic rules. The Finnish corpus consists of a total of 5,876 colour words, containing 3,172 compound words or compound phrases and 2,704 simple words. Compound words or phrases are divided in the following way: 2,925 compounds or phrases of two stems, for example *tumman-keltainen* ‘dark yellow’, 238 compounds or phrases consisting of three stems, for example *tumman-vaalean-punainen* ‘dark pink’ and nine consisting of four stems, such as *kesä-yö-taivaan-sininen* ‘summer night’s sky blue’. In the Hungarian corpus, we also came across some compounds or phrases consisting of five simple words or stems, such as *narancs-sárgás-rózsaszínes-piros* ‘orangish pinkish red’. It seems that the speakers of the Finno-Ugric languages are quite creative when asked to name colours; their gift was especially evident in the colour-naming task, where there variety of given responses multiplied significantly. In fact, it is interesting that non-prototypical colours tend to obtain higher variation in subject’s answers than prototypical colours. Figure 7 and Figure 8 present the number of different colour terms (including compounds and phrases) in the whole word-corpora and in the colour-naming task in the five Finno-Ugric languages and ESL.

Figure 7 shows that the share of different colour terms does not exceed 20% in the word-corpora (list and colour-naming task combined). In the colour-naming task, however, the proportion of different terms increases 1-3% (see Figure 8).

We also compared the usage of different colour terms in the list task with Russian and English. The data was taken from studies by Davies and Corbett (1994, 1995). Unfortunately,

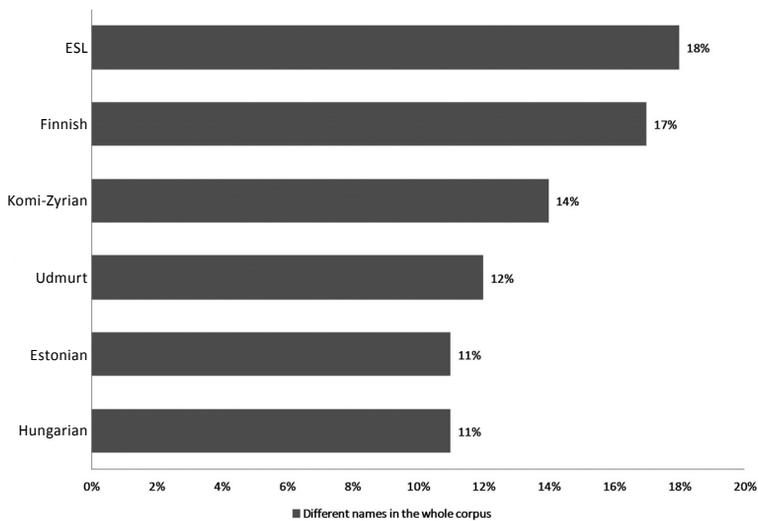


Figure 7. The percentage of different colour names in the colour-word corpora in five Finno-Ugric languages and ESL.

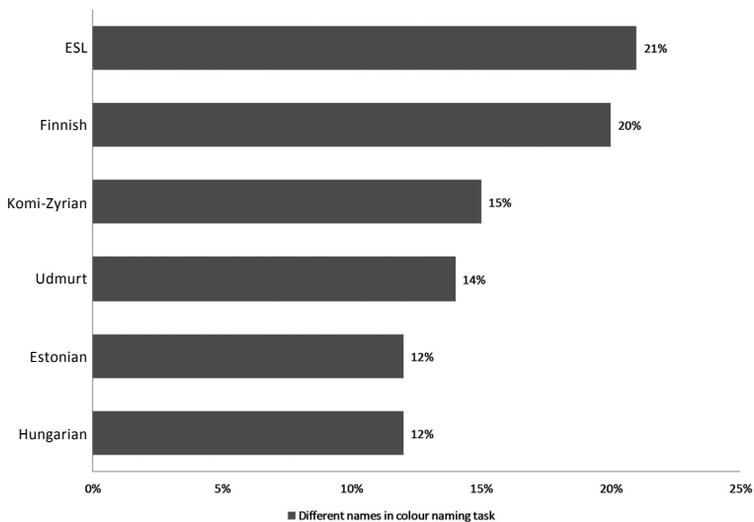


Figure 8. The percentage of different colour terms in the colour-naming task in five Finno-Ugric languages.

Davies and Corbett did not count the number of different colour terms in the colour-naming task, which makes the calculation of different names in the second task and overall impossible. In the list task, the proportion of different colour terms in Komi-Zyrian and Estonian seems to be significantly higher than in the colour-naming task or in the two tasks combined. Figure 9 also demonstrates that Finnish and Estonian speakers (the Finnic group) listed more different colour terms than the speakers of other languages.

If we take a look at the data from the two Indo-European languages, we immediately notice the contrast between Russian and other languages: Russian subjects named fewer distinctive colour terms than the speakers of other languages. However, the English corpus is comparable to that of the Finno-Ugric languages (see Figure 9).

To conclude, the Finno-Ugric languages have a large colour vocabulary, partly due to the rich word-formation tools

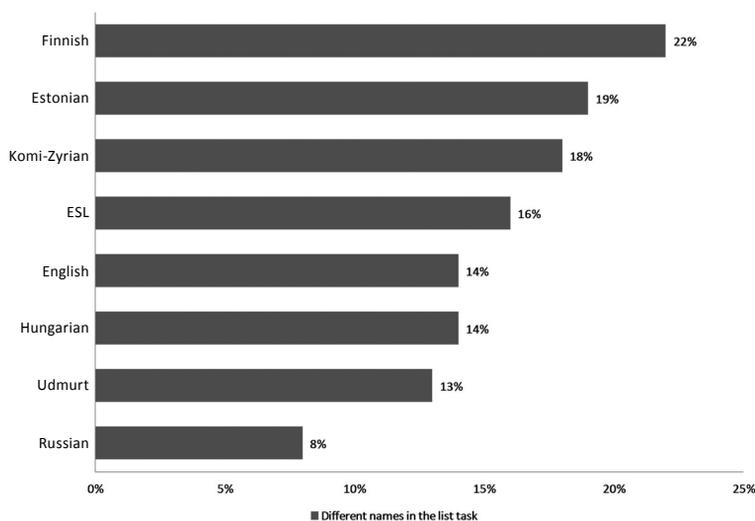


Figure 9. The percentage of different colour terms in the list task in five Finno-Ugric languages and ESL, in comparison with English and Russian.

at their disposal. However, it is interesting to observe how the speakers of Finno-Ugric languages vary their responses. Subsequently, we will discuss how extreme variation in given responses results in low consensus percentages.

5.4. Dominant colour tiles and the consensus percentage among the subjects

There are different measures available to distinguish the candidates for the basic colour term status in the colour-naming task. Firstly, we can count how often colour terms were named in the task, which shows the frequency of occurrence of a colour term. Secondly, we are interested in how many times the presented colour tiles were named with the same colour name or, in other words, we can count the dominance of a colour term, using certain threshold levels. If a colour term seems to dominate over one or more colour tiles, it has a better chance to apply for basic term status. The dominance of a colour term, in turn, demonstrates the agreement or consensus level of the language speakers. If a colour term shows at least 50% dominance, it is established among the basic colour terms (Davies and Corbett 1994).

Subsequently, we will discuss the issue, showing how Finno-Ugric language speakers have a large variety of colour terms at their disposal. For example, an Estonian subject may look at two or three stimuli of a blue hue and name them all as *sinine* 'blue'. If a fourth stimulus is presented to her, she tries to seek out another colour adjective and specify her answer according to the proportion of saturation, lightness or darkness of the stimulus, using other words besides the basic colour term *sinine* 'blue'. If the viewed colour tile is somewhat lighter than the stimuli she has called *sinine* 'blue', she might want to name the presented tile *helesinine* 'light blue'. In other words, she will try to specify her answer according to what she sees exactly, thus we observed a strong priming effect. In the colour naming task, there are at least 16 blue tiles and over half of them will most probably be named using some other simple colour adjec-

tive, compound word or colour phrase than the basic term for blue. We noticed this tendency in all of the researched Finno-Ugric languages. In the discussion below, we compare the data of Finno-Ugric languages and Estonian Sign Language with two Indo-European languages, Russian and English.

Table 5 shows the most dominant tiles for 11 basic colour term categories and the percentage of the dominance in the five Finno-Ugric languages and Estonian Sign Language. The colour terms of Estonian Sign Language are glossed into English; the exact signs can be found in Appendix 1. Table 5 shows only the most dominant basic colour term for each colour category. If a language does not possess a dominant basic colour term of at least 50% consensus level (more than 50% of the subjects have

Table 5. Dominant colour tiles and the percentage of dominance in Estonian, Finnish, Hungarian, Udmurt, Komi-Zyrian and Estonian Sign Language (ESL).

Colour category	Color-aid code of a best example	Language	Colour term	Percentage
Yellow	Y	Estonian	<i>kollane</i>	65%
		Finnish	<i>keltainen</i>	83%
		Udmurt	<i>čuž</i>	61%
		ESL	YELLOW	74%
	YOY	Komi-Zyrian	<i>kol̄kviž</i>	63%
Orange	OYO	Estonian	<i>oranž</i>	70%
		Finnish	<i>oranssi</i>	72%
		Hungarian	<i>narancs-sárga</i>	71%
		S-Udmurt	<i>nap-čuž</i>	53%
Brown	YO S3	Hungarian	<i>barna</i>	53%
		Finnish	<i>ruskea</i>	53%
	O S3	Estonian	<i>pruun</i>	74%
Red	RO	Finnish	<i>punainen</i>	68%
		Hungarian	<i>piros</i>	63%
		Udmurt	<i>gord</i>	56%
		ESL	RED	54%

	ROR	Estonian	<i>punane</i>	57%
		Komi-Zyrian	<i>gerd</i>	76%
Pink	R T4	Finnish	<i>vaalean-punainen</i>	57%
	ROR T3	Estonian	<i>roosa</i>	57%
Purple	RV	Estonian	<i>lilla</i>	54%
	VRV	Hungarian	<i>lila</i>	58%
Blue	B	Estonian	<i>sinine</i>	64%
		Komi-Zyrian	<i>lęz</i>	67%
		Udmurt	<i>lįz</i>	51%
	BGB	Estonian	<i>sinine</i>	64%
		Hungarian	<i>kék</i>	52%
		ESL	BLUE	58%
	BVB	Finnish	<i>sininen</i>	56%
Green	G	Estonian	<i>roheline</i>	75%
		Finnish	<i>vihreä</i>	62%
		Hungarian	<i>zöld</i>	60%
		Udmurt	<i>vož</i>	51%
		ESL	GREEN	54%
	GYG	Komi-Zyrian	<i>turunviž</i>	67%
Gray	GRAY 2	Estonian	<i>hall</i>	71%
	GRAY 4	Finnish	<i>harmaa</i>	58%
		Hungarian	<i>szürke</i>	72%
		Udmurt	<i>purįš</i>	70%
		Komi-Zyrian	<i>rud</i>	66%
		ESL	GRAY	70%
Black	BLACK	Estonian	<i>must</i>	89%
		Finnish	<i>musta</i>	85%
		Hungarian	<i>fekete</i>	84%
		Udmurt	<i>śęd</i>	74%
		Komi-Zyrian	<i>śęd</i>	84%
		ESL	BLACK	82%
White	WHITE	Estonian	<i>valge</i>	84%
		Finnish	<i>valkoinen</i>	59%
		Hungarian	<i>fehér</i>	62%
		Udmurt	<i>tęđi</i>	81%
		Komi-Zyrian	<i>ježid</i>	92%
		ESL	WHITE	66%

named a certain tile with one colour term) for a colour category, it has not been included in the table. For example, we excluded Hungarian yellow and Finnish purple from Table 5, because these terms were applied to a certain colour tile with less than 50% consensus.

We counted the number of dominant colour tiles (consensus level 50% or greater) for each Finno-Ugric language and Estonian Sign Language, and for comparison, for English (Davies and Corbett 1995) and Russian (Davies and Corbett 1994) (English and Russian data is not presented in Table 5). The overall number of dominant colour tiles is bigger than those presented in Table 5, because Table 5 only shows the number of dominant basic colour terms. In Estonian, 12 colour terms dominate for 23 colour tiles at the 50% consensus level, including 11 basic colour terms and *tumesinine* ‘dark blue’; in Finnish 11 colour terms dominate for 19 colour tiles: 10 basic colour terms (*violetti* ‘purple’ excluded) and *vaaleansininen* ‘light blue’ (Uusküla 2007: 380-382); in Hungarian, 10 colour terms dominate for 14 tiles: nine basic colour terms, excluding *sárga* ‘yellow’ and *rózsaszín* ‘pink’, and one compound, *sötétbarna* ‘dark brown’.

There are several reasons why Hungarian *sárga* ‘yellow’ fails to gain 50% dominance. *Citromsárga* ‘lemon yellow’ (given to tile Y 61 times, 48.4%) seemed to be more popular in the tile-naming task. In fact, *sárga* ‘yellow’ was most frequently given to colour tiles Y with a consensus percentage of 37.6, and YOY with a consensus percentage of 36.8. The basic colour term for yellow, *sárga*, in Hungarian, has a relatively weak status as a basic term, having a strong contender, *citromsárga* ‘lemon yellow’, used for the part of the colour spectre between greenish yellow and dark yellow (but not orange). Therefore, the usage of *sárga* is lower than expected, especially in the colour-naming task, where only two yellow colour tiles were presented to the subjects: Y – highly saturated bright yellow (mostly *citromsárga* for Hungarian speakers) and YOY – yellow with a slight overtone of orange (too orange to be described as *sárga* for Hungarian speakers) (see Section 5.3 and Table 5 for further explanation). Although *sárga* scored high in the list task and followed the

same pattern as the basic terms for yellow in other Finno-Ugric languages, with a comparatively high salience index score, *citromsárga* appeared in 13th place in frequency list, after 11 basic colour terms and *bordó* ‘bordeaux, claret’ by the same index (Bogatkin 2005: 59). The main difference between the two colour terms *sárga* and *citromsárga* seems to be the consistency of occurrence in the colour-naming task, which is proportionally high for *citromsárga* and low for *sárga*.

In the Udmurt language, only eight colour terms (yellow, orange, red, blue, green, white, grey and black) dominate, for only 11 colour tiles, at the 50% consensus level (Ryabina 2011a: 209-214). In Komi-Zyrian, we find 10 colour terms dominating for 17 tiles, including colour terms for dark blue and dark green (Ryabina 2011c: 75-77). If we compare the results in the Finno-Ugric languages and Estonian Sign Language with Russian or English, we find significant differences. In Russian, 13 colour terms dominate for 37 colour tiles, including 12 basic colour terms, and *sirenevyj*, which represents a variety of the purple hue. In English, exactly 11 basic colour terms dominate for 49 colour tiles out of 65.

Table 5 shows that dominance percentages are relatively low, even for the colour tiles BLACK and WHITE, especially when we compare the results with Russian or English (Davies and Corbett 1994, 1995). Russian subjects, for example, showed a 100% consensus when naming the colour tile WHITE (Davies and Corbett 1994: 77). Consensus percentages were also relatively low for blue, green and red tiles amongst all Finno-Ugric languages, which demonstrates both the richness of colour vocabulary, i.e. there are many possible ways to name one stimulus, and the disagreement of the subjects, i.e. what should be the (correct) name of a certain stimulus. The consensus percentages are relatively high for tiles representing the categories of white, black, grey and orange.

Dominant tiles can also be considered to form a colour foci in a language. As seen in Table 5, these do not correspond to each other between languages (see also Uusküla 2006, 2008a, 2011b, and Hollman 2010).

6. Conclusion

In this paper, the data on the basic colour terms in five Finno-Ugric languages (Estonian, Finnish, Hungarian, Komi-Zyrian and Udmurt) and in the Estonian Sign Language was discussed and compared with each other and with the data from two Indo-European languages (English and Russian). The collected data shows that the inventory of basic colour terms in the five Finno-Ugric languages and Estonian Sign Language has reached Stage VI (Komi-Zyrian) or Stage VII (Estonian, Finnish, Hungarian, Udmurt and Estonian Sign Language). Estonian and Hungarian have fully developed basic colour vocabularies, with the maximum number of 11 basic colour terms. Besides the basic terms, the analysed Finno-Ugric languages and ESL possess a wide variety of different colour terms, including compounds and morphologically marked colour words describing different shades, degrees of saturation and lightness or darkness of colours. The complexity index was used in order to ascertain the length of the words in corpora. The variety of the terms used is evident in the relatively high percentage of different terms and the low proportion of basic terms in the whole corpus, as well as in the low consensus compared to the colour term studies in Russian or English. The basic colour terms in the data range from 30% to 44% of the colour terms used in the empirical tasks. The comparative study shows that Finno-Ugric languages have rich colour vocabularies. However, we also noticed that male subjects experienced greater colour-naming difficulties than female subjects did. The cognitive salience index is useful in distinguishing the primary and secondary basic colour terms. The values of the primary terms always score higher. If there is a category split in a language (e.g. Russian, with two basic terms for blue), both colour terms for that category have lower cognitive salience index values. The most salient term in the Finno-Ugric languages, Estonian Sign Language, Russian and English is red, except in Estonian, where it is blue. An important finding is that the evolutionary order, i.e. the diachronic order of emergence of the basic colour categories and corresponding terms, is different from their synchronic salience.

Acknowledgements

The study was supported by the Estonian Ministry of Education and Research project SF0050037s10 and the Estonian Science Foundation grant no 8168. This support is highly appreciated. We are indebted to Geda Paulsen for her valuable comments. We would also like to thank Elena Ryabina for providing her data on Udmurt and Komi-Zyrian for analysis, and Lenno Uusküla for carrying out the statistical analysis.

Addresses:

Mari Uusküla
Institute of the Estonian Language
Roosikrantsi 6
Tallinn 10119, Estonia
E-mail: Mari.Uuskyla@eki.ee

Liivi Hollman
E-mail:liivi@viipekeeletoolid.ee

Urmas Sutrop
E-mail: Urmas.Sutrop@eki.ee

References

- Barratt, Leslie and Miklós Kontra (1996) "Matching Hungarian and English color terms". *International Journal of Lexicology* 9, 2, 102-117.
- Berlin, Brent and Paul Kay (1969) *Basic color terms: their universality and evolution*. Berkeley, CA: University of California Press.
- Bogatkin, Mari (2005) *Värvinimed ungari keeles: põhinimed, nende struktuur ja kujunemine*. Unpublished Master's Thesis. Tartu: Tartu Ülikooli uurali keelte õppetool.
- Bogatkin-Uusküla, Mari and Urmas Sutrop (2005a) "Tänapäeva ungari keele põhivärvinimed". *Keel ja Kirjandus* 7, 558-571.
- Bogatkin-Uusküla, Mari and Urmas Sutrop (2005b) "Kas ungari keeles on kaks punase värvi põhinime *piros* ja *vörös*?". *Emakeele Seltsi aastaraamat* 50, 93-110.
- Corbett, Greville G. and Ian R. L. Davies (1994) "Linguistic and behavioural measures for ranking basic colour terms". *Studies in Language* 19, 2, 301-357.

- Davies, Ian R. L. and Greville G. Corbett (1994). "The basic color terms of Russian". *Linguistics* 32, 65-89.
- Davies, Ian R. L. and Greville G. Corbett (1995) "A practical field method for identifying basic colour terms". *Languages of the World* 9, 25-36.
- Davies, Ian R. L., C. Macdermid, Greville G. Corbett, H. McGurk, D. Jerrett, and P. Snowden (1992) "Color terms in Setswana: a linguistic and perceptual approach". *Linguistics* 30, 1065-1103.
- Davies, Ian R. L., Greville G. Corbett, Al Mtenje, and Paul Sowden (1995). "The basic colour terms of Chichewa". *Lingua* 95, 259-278.
- Forbes, Isabel and Gábor Kiss (1999) "Colour categorization and naming in French and Hungarian". In Christopher Beedham, ed. *Selected Proceedings of the XXXIst Annual Meeting of the Societas Linguistica Europaea, St Andrews 1998*, 181-188. Oxford: Pergamon.
- Futaky, István (1981) "Zur Herausbildung des Farbfeldes im Finnougrischen". In *Lakó-émlekkönyv*, 48-58. Budapest.
- Hollman, Liivi (2008) "Miks must on must ja valge valge: eesti viipekeele värvinimedest". *Keel ja Kirjandus* 51, 11, 847-862.
- Hollman, Liivi (2010) *Basic colour terms in Estonian Sign Language*. (Dissertationes Linguisticae Universitatis Tartuensis, 14.) Tartu: Tartu University Press.
- Hollman, Liivi and Urmas Sutrop (2011) "Basic color terms in Estonian Sign Language". *Sign Language Studies* 11, 2, 130-157.
- Kay, Paul and Chad K. McDaniel (1978) "The linguistic significance of the meanings of basic color terms". *Language* 54, 610-646.
- Kiss, Gábor and Isabel Forbes (2001) "Piros, vörös – red, rot, rouge". In Tamás Gecső, ed. *Kontrasztív szemantikai kutatások*, 190-199. (Segédkönyvek a Nyelvészet Tanulmányozásához, 11.) Budapest: TINTA Könyvkiadó.
- Kerttula, Seija (2007) "Relative basicness of color terms: modelling and measurement". In Robert E. MacLaury, Galina V. Paramei, and Don Dedrick, eds. *Anthropology of color: interdisciplinary multilevel modeling*, 151-169. Amsterdam and Philadelphia: John Benjamins.
- Komenský (Comenius), Jan Amos (1685) *Orbis sensualium pictus quadrilinguis*. Leutschoviae (Levoča): Brewer. After a Facsimile ed. Prague, 1979.
- Koski, Mauno (1983) *Värien nimitykset suomessa ja lähisukukielissä*. (Suomalaisen Kirjallisuuden Seuran toimituksia, 391.) Helsinki: Suomalaisen Kirjallisuuden Seura.
- MacLaury, Robert E., Judit Almási, and Zoltán Kövecses (1997) „Hungarian *piros* and *vörös*: color from points of view”. *Semiotica* 114, 1-2, 67-81.
- Pallas, Peter Simon (1786) *Linguarum totius orbis vocabularia comparativa. Pars secunda*. Petropoli: Schnoor. Facsimile ed. Hamburg: Buske, 1977.
- Paramei, Galina (2005) "Singing the Russian blues: an argument for culturally basic color terms". *Cross-Cultural Research* 39, 1, 10-34.

- Rakin, A. N. (1990) "Leksika cvetooboznačenja v permskix jazykax". *Fenno-Ugristica* 16 = *Acta et Commentationes Universitatis Tartuensis* 876, 112-121.
- Regier, Terry, Paul Kay, and Naveen Khetarpal (2009) "Color naming and the shape of color space". *Language* 85, 4, 884-892.
- Ryabina, Elena (2011a) "Osnovnye cvetooboznačenja v sovremennom udmurtskom jazyke". *Linguistica Uralica* 47, 3, 199-223.
- Ryabina, Elena (2011b) "Differences in the distribution of colour terms in colour space in the Russian, Udmurt and Komi-Zyrian languages". *Journal of Estonian and Finno-Ugric linguistics* 2, 2, 191-213.
- Ryabina, Elena (2011c) *Osnovnye cvetooboznačenja v permskix jazykax*. (Dissertationes Philologiae Uralicae Universitatis Tartuensis, 11.) Tartu: Tartu University Press.
- Sipőcz, Katalin (1994) *A vogul nyelv színnevei*. (Studia-uralo-altaica, Suppl. 3.) Szeged.
- Sutrop, Urmas (1995) "Eesti keele põhivärvinimed". *Keel ja Kirjandus* 12, 797-808.
- Sutrop, Urmas (2000) "The basic colour terms of Estonian". *Trames* 4, 1, 143-168.
- Sutrop, Urmas (2001) "List task and a cognitive salience index". *Field Methods* 13, 263-276.
- Sutrop, Urmas (2002) *The vocabulary of sense perception in Estonian: structure and history*. (Opuscula Fenno-Ugrica Gottingensia, 8.) Frankfurt am Main: Peter Lang.
- Sutrop, Urmas (2011). "Towards a semiotic theory of basic colour terms and the semiotics of Juri Lotman". In Carole Biggam, Carole Hough, Christian Kay, and David Simmons, eds. *Progress in colour studies*, 39-48. Amsterdam: Benjamins.
- Turunen, Rigina (2002) "Die Farbbezeichnungen im Mokscha-Mordwinischen". *Finnisch-Ugrische Forschungen* 57, 1-3, 167-194.
- Uusküla, Mari (2006) "Distribution of colour terms in Ostwald's colour space in Estonian, Finnish, Hungarian, Russian, and English". *Trames* 10, 2, 152-168.
- Uusküla, Mari (2007) "The basic colour terms of Finnish". *SKY Journal of Linguistics* 20, 367-397.
- Uusküla, Mari (2008a) Basic colour terms in Finno-Ugric and Slavonic languages: myths and facts. (Dissertationes Linguisticae Universitatis Tartuensis, 9.) Tartu: Tartu University Press.
- Uusküla, Mari (2008b) "The basic colour terms of Czech". *Trames* 12, 1, 3-28.
- Uusküla, Mari (2011a) "Terms for red in Central Europe: an areal phenomenon in Hungarian and Czech". In Carole Biggam, Carole Hough, Christian Kay and David Simmons, eds. *Progress in colour studies*, 147-156. Amsterdam: Benjamins.
- Uusküla, Mari (2011b) "Värviprototüübid eri keeltes". In Mari Uusküla and Urmas Sutrop, eds. *Värvinimede raamat*, 387-395. (Töid antropoloogilise ja etnolingvistika vallast, 5.) Tallinn: Eesti Keele Sihtasutus.

Uusküla, Mari and Urmas Sutrop, eds. (2011) *Värvinimed raamat*. (Töid antropoloogilise ja etnolingvistika vallast, 5.) Tallinn: Eesti Keele Sihtasutus.

Wichmann, Yrjö (1942) *Syrjänischer Wortschatz nebst Hauptzügen der Formenlehre*. (Lexica Societatis Fenno-Ugricae, 7.) Helsinki: Suomalais-ugrilainen Seura.

APPENDIX 1

Colour terms in Estonian Sign Language.



BLACK



WHITE



RED



YELLOW



GREEN



BLUE



BROWN



GRAY



PINK/PURPLE

Kokkuvõte. Mari Uusküla, Liivi Hollman ja Urmas Sutrop: Põhivärvinimed viies soome-ugri keeles ja eesti viipekeeles. Artiklis võrreldakse viie soome-ugri keele: eesti, soome, ungari, udmurdi ja sürjakomi ning eesti viipekeele värvisõnavara mitmest erinevast aspektist: põhivärvinimede arv ja koostis, põhivärvinimede ja erinevate värvinimede osakaal kogutud andmekorpustes, kognitiivse esiletuleku indeksi väärtus loetelukatses ja dominantsete värvitahvlite arv nimeandmiskatses. Andmete kogumiseks on kasutatud Daviese ja Corbett' välimeetodit, mis võimaldab andmeid omavahel võrrelda. Soome-ugri keeltele ja eesti viipekeelele on iseloomulik erakordselt rikas värvisõnavara. Tänu sõnamoodustusele on võimalik väljendada igale värvitoonile iseloomulikke nüansse, nt heledust, tumedust, eredust, küllastust jne. Kõik vaadelavad keeled asetsevad Berliini ja Kay

põhivärvinimede hierarhias VI-VII staadiumil ning nendes on 7-11 põhivärvinime. Primaarseid ja sekundaarseid põhivärvinimesid aitab eristada kognitiivse esiletuleku indeks, mis moodustab teatud seaduspärasusi või mustreid. Soome-ugri keeli vene ja inglise keelega võrreldes leidsime, et need mustrid tunduvad olevat universaalsed.

Märksõnad: värvisõnavara, põhivärvinimi, võrdlev leksikoloogia, välimeetod, kognitiivne esiletuleku indeks, eesti keel, soome keel, sürja-komi keel, udmurdi keel, ungari keel, eesti viipekeel