DISSERTATIONES
LINGUISTICAE
UNIVERSITATIS
TARTUENSIS
32

KAIDI RÄTSEP

Colour terms in Turkish, Estonian and Russian: How many basic blue terms are there?

DISSERTATIONES LINGUISTICAE UNIVERSITATIS TARTUENSIS
32

DISSERTATIONES LINGUISTICAE UNIVERSITATIS TARTUENSIS
32

KAIDI RÄTSEP

Colour terms in Turkish, Estonian and Russian: How many basic blue terms are there?

UNIVERSITY OF TARTU
Press
1632

University of Tartu, Institute of Estonian and General Linguistics

Dissertation accepted for the commencement of the degree of Doctor of Philosophy on October 11st, 2018 by the Committee of the Institute of Estonian and General Linguistics, Faculty of Philosophy, University of Tartu

Supervisors: Professor Urmas Sutrop, University of Tartu
Associate Professor Mari Uusküla, Tallinn University

Opponent: PhD Anetta Kopecka, Université de Lyon

Commencement: December 14th, 2018 at 14.15, Jakobi 2–438, Tartu

This study has been supported by the Graduate School of Linguistics, Philosophy and Semiotics; funded by the European Social Fund and European Regional Development Fund (University of Tartu ASTRA Project PER ASPERA).

European Union
European Regional
Development Fund

Investing
in your future

ISSN 1406-5657
ISBN 978-9949-77-919-2 (print)
ISBN 978-9949-77-920-8 (pdf)

University of Tartu Press
www.tyk.ee

ACKNOWLEDGEMENTS

The experimental colour data found in the thesis can be used for a wide array of applications. They can be used to investigate language-specific terms, compare the naming of stimuli in Turkish, Estonian and Russian, investigate the grouping of colours and the strategies employed by language guides, and much more. I hope that the reader will find the data useful for generating new avenues of thought, and that the thesis will add to the knowledge and research of others.

The journey to my PhD thesis began with a misheard word. About a decade ago I was in a grammar lecture when I heard the lecturer say: "Today we will be talking about colours". Unfortunately, or fortunately, it was about verbs since the Estonian *värvid* 'colours' and *verbid* 'verbs' can in fact be mistaken, but I still remember the elation I felt when I thought it would be about colours. For my bachelor thesis *A Comparison of Estonian and Turkish Basic Color Terms and the Underlying Mechanism for Cultural Divisions of Color* (Rätsep 2006) I first learned about Brent Berlin and Paul Kay's idea of universal colour terms. My BA thesis opponent was Urmas Sutrop, who later became the advisor for both my master's thesis and my doctoral dissertation. Under his guidance, I gained valuable opportunities to do fieldwork for my research. Mari Uusküla, who is my thesis co-advisor, later joined me on my PhD journey. The importance of their work for colour research in general and for my thesis especially cannot be overstated. I am grateful for the time and thought that reviewers Anetta Kopecka, PhD and Anders Steinvall, PhD have invested in this thesis. Their constructive criticism is very much appreciated and wherever applicative their suggestions have been taken into consideration.

My master's thesis *Turkish basic and non-basic colour terms* (Rätsep 2010) resulted in my first ever published article *Preliminary research on Turkish basic colour terms with an emphasis on blue* (Rätsep 2011) at the colour conference *Progress in Colour Studies* (PICS) 2010 volume (Biggam 2011), which transitioned itself into this thesis. Brent Berlin and Paul Kay's basic colour terms paradigm (Berlin and Kay 1999) is the theoretical standpoint of the thesis; however, the focus is on the blue category and the number of possible basic level categories in it in Turkish, Estonian and Russian. At the PICS 2012 conference I analysed my Estonian results in the presentation "Sorting and naming blue: an Estonian case study" (Rätsep 2012). I am grateful for the opportunity to participate and the chance it gave me to present, and above all to meet researchers with a similar interest in colour from all over the world.

The work in question would not come to fruition without the impetus and help given by my supervisors Prof. Urmas Sutrop and Dr. Mari Uusküla, to whom I am greatly indebted. The research was also supported by the grants "Areal or universal: Basic colour terms in the Baltic Sea, Central-European and Mediterranean areas" (Estonian Science Foundation grant no. 8168, grant holder Mari Uusküla) and "Word and concept in the development of Estonian vocabulary and place names" (Ministry of Education and Research grant no.

50037s10, grantholder Urmas Sutrop). Additionally, I have worked on the colour names in Estonian place names, most notably *must* 'black' and *valge* 'white'. Publications on colours in Estonian place names include *The colour term 'black' in Estonian place names* (Rätsep 2012) and *Must 'black' and valge 'white' in Estonian place names: Their naming motives in folk etymology* (Rätsep 2016).

The work would have not been possible without the help of native speakers of Turkish, Russian and Estonian. My language guides spent their valuable time helping me in my research for no other reason than good will. For that I am immeasurably thankful.

To get reliable data, not only the interviewee but also the interviewers themselves need to be fluent or native speakers. My Russian skills are geared towards the literary language and I do not consider myself fluent in conversation, so I am grateful to Olga Titova, whom I instructed on conducting the interviews and together with whom we conducted the first interview with the Russian speakers. For the most part the Russian interviews were conducted by Olga Titova under Mari Uusküla's grant, but the data transcription and analysis were my own contribution. The Turkish part of the thesis benefited greatly from the advice from Ankara University professors Nese Özden, and Halil Ibrahim Usta.

For the technical know-how I am deeply grateful to David Bimler from Massey University, New Zealand. David provided invaluable help in analysing the data and offered insightful comments and literary suggestions for this thesis. We have a co-authored draft of an article *Estonian case study: Sorting and naming the BLUE* in progress. The legibility of the text was greatly improved by Robin Hazlehurst, who edited the main body of text, and Geda Paulsen, who edited the Estonian summary.

I have also had help from professors from the University of Tartu, whose lectures and coursework taught me the technological skills needed to analyse the data. Help with my programming skills came from University of Tartu lecturers Pärtel Lippus, Kristel Uiboaed, and Aki-Juhani Kyröläinen.

I will conclude by noting my gratitude to my former colleagues from the Institute of the Estonian Language, most particularly Vilja Oja, and to friends and family members for their support.

CONTENTS

List of figures

List of tables

INTRODUCTION

This thesis analyses the colour vocabulary of Turkish, Estonian and Russian. The division of the blue spectrum is scrutinised through the paradigm of basic colour terms theory.

A part of the thesis is dedicated to a review of the previous research. The previous studies summarised were selected because they look at the possibility of division within what is usually one colour category. The review is intended to familiarise the reader with Brent Berlin's and Paul Kay'sbasic colour terms theory by presenting research that pushes the boundaries of this theory. The revisions to the basic colour term paradigm theoretically allow within-category division through the use of 'light' and 'dark' modifiers or allow for a larger number of basic colour terms than the original upper limit of eleven, but it is not usual to find that the theoretical possibility is actually supported by experimental study. The results from the empirical tests are in some cases accompanied by examples from previous research, and so the literature review and the empirical research are not presented separately, but rather support each other where necessary.

This thesis continues from the earlier research, mainly the work by Emre Özgen and Ian R.L. Davies on Turkish basic colour terms (1998) and part of Urmas Sutrop's dissertation on Estonian basic colour terms (2002). Özgen and Davies prove the relevance of an extra category experimentally, but dismiss it on theoretical grounds. In Sutrop's experimental study a possible division of the Estonian 'blue' category is suggested and possible Russian influence is proposed, but no further fieldwork was done by Sutrop to verify or dismiss the possibility.

The aim of the thesis is to continue the previous research on Turkish, Estonian and Russian colour terms. The methodology and the selection of stimuli is focused on the blue area of the spectrum and it is hoped that this will help answer the question of whether there is more than one 'blue' category. Previous research on Russian has proven that *goluboj* 'light blue'[1] is a separate category from *sinij* 'blue', so if the division is possible in Russian, the question arises of whether it could be in Estonian too. The Turkish study was undertaken to further pinpoint the stimulus named *lacivert* 'dark blue' and the overall position of *lacivert* 'dark blue'within the previously established Turkish basic colour terms. The list task and naming tasks run earlier were completed again to re-validate or invalidate the position of *lacivert*'dark blue', this time with additional stimuli selected to target the blue area of colour space. A sorting task

---

[1] Russian uses the Cyrillic alphabet. The collected Russian data were transliterated to Latin (online, using ISO 9:1995 standard). It is meant for those who are more familiar with Latin alphabet, although the letters used in the transliteration do not correspond to their English pronunciation. For readers who are familiar with Russian, the transiteration will look umseemly.

was conducted to answer whether there is a possible division of the ‘blue’ category in Estonian.

Including a sorting task is very rare in Estonian colour term research (see Kalda 2014). Sutrop had previously done the list task and the naming task, but the selection of stimuli was different for the naming task. The selected stimuli for the Estonian and Russian field tests were previously used by David Bimler and Mari Uusküla (2014) to ascertain the division in the Italian ‘blue’ category. Uusküla’s tests support there being more than one ‘blue’ term in Italian, and so the selection of stimuli is suitable for testing whether there is more than one ‘blue’ category in Estonian and Russian.

The thesis contains the data from several field tests. The list and naming tasks were conducted in Turkish (N=56), Estonian (N=39), and Russian (N=30), and the sorting tasks in Estonian and Russian. Stimuli were used in the sorting and naming tasks and the hypothesis meant that the stimuli selection included an abundance of bluish stimuli.

It is hypothesised that if the Turkish *lacivert* ‘dark blue’, Estonian *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’, and Russian *goluboj* ‘light blue’ do not share a type-of relationship with the established basic colour terms *mavi*, *sinine* and *sinij* ‘blue’ and are separate basic colour terms from them, then all their quantifiable measures should be relatively similar to previously establishes basic colour terms. The null-hypothesis is that if the Turkish *lacivert* ‘dark blue’and Estonian *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’ are not basic colour terms, then their basic quantifiable measures as colour traits will not match those of previously established basic colour terms. Quantifiable variables are mostly operationalised by frequencies and a multi-method approach was used in a range of field tests. The most important quantifiable measures in the list task are frequency and mean position, those in the sorting task are the consensual formation and subsequent naming of stimuli groups, and those in the naming task are the consensus of stimulus naming by participants.

If there is more than one category of ‘blue’ in Turkish and Estonian, this would support the weak relativist view of colour terms, but if the null-hypothesis is supported by there being only one category of ‘blue’ in Turkish and Estonian, it would support the existence of universal colour categories.

By and large the main finding of the thesis supports the previous research, which suggests that Russian *goluboj* ‘light blue’ is a basic colour term that is separate from *sinij* ‘blue’. The Turkish colour term *lacivert* ‘dark blue’ lacked the high-level consensus to separate itself irrevocably from *mavi* ‘blue’, but that may in part be due to all the modifiers being included as they were used by the language guides. The Estonian *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’ lack the high level of consensus necessary to separate them from *sinine* ‘blue’ in the sorting task, while the high frequency of *helesinine* ‘light blue’ in the naming task may arise because of the large number of bluish stimuli, or possible because of priming from the preceding sorting task.

The new information in the thesis that I consider to be the most valuable is:

1) *a new variation of the list task scatter plot*: The thesis introduces a new way of displaying the list task results in a graph, see Figures 3, 5, and 12. The two main parameters of the list task are represented by the mean position of the term on the x-axis and the frequency in percentages on the y-axis. Each dot on the scatter plot represents a different list task term, and the terms with the highest frequency are labelled. This creates a clear and simple picture of the task.
2) *comparable Estonian and Russian sorting task data*: The data are most comprehensively represented in multidimensional scaling analysis, which presents a pictorial overview of the sorting task categories. Additionally, Estonian sorting data are directly comparable with the Russian sorting task results, which present a case of division in the 'blue' category.

The thesis comprises two main parts, the introduction to the theory and the relevant language examples from previous empirical studies, and the results from the data collected and the findings and conclusions. The introductory part includes an introduction to the universalist theory of basic colour terms and its main works, and it discusses relevant possible exceptions in different languages that supplement the experimental data. This selection from previous studies is not intended to be definitive or exhaustive, but rather is intended to present a selection of empirical data. The hypothesis and the research question that the experimental data were sourced to answer are presented. The first part concludes with the section on methods and participants.

The main part of the thesis reports the results of experimental data from three different languages, Turkish, Estonian and Russian. The results are presented by language, and for each language by the tasks. Each language has an introductory section on the relevant colour terms and the different issues inherent to that language. This is meant to orient the reader to the colour terms of that language and possible points of departure from the theory. Data analysis of each language ends with a section for the findings of the separate tasks.

The comments section reviews some of the weaknesses found in analysing the data, whether from a methodological, observational, or analytical standpoint. This is an attempt to synthesise the theoretical standpoint, the methods used, the data results and the research question into one cohesive unit. The comments section, which ends the third main part of the thesis, highlights some of the weak points of the conclusions drawn. The thesis also includes a list of tables and figures.

I. THEORY, RESEARCH QUESTION AND METHODOLOGY

Chapter 1

1.1 Theory of basic colour terms: background and the present position

Nineteenth-century researchers, among them the classicist and politician William Gladstone and the ophthalmologist Hugo Magnus, were aware that not all languages reflect identical lexical classifications of colour, but scholars viewed differences in colour lexicons in evolutionary terms. In the early twentieth century the linguistic and cultural relativity approach of Edward Sapir and B. L. Whorf gained support, and the relativist view became the established doctrine in the 1950s and 1960s, supported by cross-linguistic research (MITECS s.v. ‘Color categorization’).

The mainstream relativist approach, which stated that languages divided colours arbitrarily and that these linguistic divisions shape the way that speakers perceive colour (Lenneberg and Roberts 1956)}, was opposed by the proponents of the universalist approach, most significantly by Brent Berlin and Paul Kay. In 1969, a shift in the focus of anthropological linguistics from relativism to universalism occurred with the publication of Brent Berlin and Paul Kay’s *Basic Color Terms: Their Universality and Evolution* (Berlin and Kay 1999). The underlying impetus for the theory arose when Berlin and Kay perceived how common colour terms could be translated from language to language with ease. Although Berlin and Kay relied heavily on previous research, it was the publication of their monograph that brought widespread support – and criticism – to the universalist approach in colour term research.

Berlin and Kay proposed that “…a total universal inventory of exactly eleven basic colour categories, *white, black, red, green, yellow, blue, brown, purple, pink, orange,* and *grey,* exists from which eleven or fewer basic color terms of any given language are always drawn” (Berlin and Kay 1999: 2). Berlin and Kay’s original data encompassed 98 languages (Berlin and Kay 1999: 45) and they had experimental data for twenty of them (1999: 42), so additional non-experimental data were needed for 78 languages. For these languages they relied heavily on literary sources, mostly dictionaries, and previous research. Berlin and Kay used synchronic data to form a diachronic hypothesis. They argued that the sequence of basic colour terms represents “not only a distributional statement for contemporary languages but also the chronological order of lexical encoding of basic color categories in each language” (1969: 4) and they interpret the chronological order as a sequence of evolutionary stages.

> "In sum, our two major findings indicate that the referents for the basic color terms of all languages appear to be drawn from a set of eleven universal perceptual categories, and these categories become encoded in the history of a given language in a partially fixed order." (Berlin and Kay 1969: 5)

The temporal-evolutionary ordering presented in the original monograph (1969: 4) has since become largely outdated, while the idea of basic colour terms remains central to the paradigm.

Berlin and Kay's definition of basicness has four primary characteristics (1969: 6):

1) It is *monolexemic*, i.e. the meaning of the term cannot be predicted from the meaning of its parts;
2) Its signification is not included in that of any other colour term;
3) Its application is not restricted to a narrow class of objects;
4) It must be psychologically salient for participants, which means that the term has:
   a) a tendency to occur at the beginning of elicited lists of colour terms;
   b) stability of reference across participants and across occasions of use;
   c) occurrences in the idiolects of all participants.

Research into the third criteria has shown that in some European languages *beige* is contextually constrained (Martin Eessalu and Mari Uusküla 2013), so the Estonian *beež* 'beige' for example, does not fit the third criteria of a basic colour term, because it is mostly used for inanimate objects. It is generally assumed that basic colour terms correspond to basic level terms, but Steinvall (2002: 61) argues that some terms e.g. *blond(e)* could exist both at the basic level and as a colour term that does not automatically qualify as a basic colour term, even though *blond(e)* is frequent enough to be considered basic level.

There were a further four subsidiary criteria (Berlin and Kay 1969: 6–7):

5) The doubtful form should have the same distributional potential as the previously established basic colour terms.
6) Colour terms that also signify the name of an object that characteristically has that colour are suspect.
7) Recent foreign loan words could be suspect.
8) In cases where lexemic status is difficult to access, morphological complexity is given some weight as a secondary term.

Sarapik (2000: 13) suggests that the subsidiary criteria of an object having the same name and colour would eliminate *oranž* 'orange' if the position of *oranž* was suspect from the first primary characteristics.

Subsequent field studies have led to a reconceptualisation of the evolutionary sequence (Kay 1975; Kay and McDaniel 1978). The authors of *The World Color Survey* (Kay *et al.* 2010) say that the first major amendment to the temporal-evolutionary ordering, which followed mainly from the work of

Eleanor Rosch (nee Heider) (Heider 1971; Rosch Heider 1972; Heider 1972; Rosch 1973) and Berlin and Kay themselves, was the understanding that "two term systems in fact contain one term for black, green, and blue and other so-called cool colors and one term for white, red, yellow and other 'warm' colors" (Kay *et al.* 2010: 3).

Kay and McDaniel (1978) update the theory by introducing *fuzzy sets* into the colour terms paradigm, stating that colour categorisation is in general a matter of degree, and so colour categories are best regarded as fuzzy sets. They also suggested the possibility of composite categories, which are fuzzy unions of at least two fundamental response categories, *red, yellow, green, blue, black, white* (Kay and McDaniel 1978: 630). Furthermore, and most relevantly to this thesis, they indicated that the original upper limit of 11 basic colour terms may be arbitrary:

> "Identity with the six fundamental response functions is the basis of the primary basic color categories black, white, red, yellow, green, and blue. Fuzzy unions of fundamental response categories are the basis of the four composite basic-color categories light-warm, dark-cool, warm, and cool (grue). Fuzzy inter-sections of fundamental response categories are the basis of at least five derived basic color categories-brown, pink, purple, orange, and grey. Thus where B[erlin] & K[ay] described eleven universal basic color categories of a single logical type, there are in fact at least fifteen basic color categories of three types (McDaniel 1974, MS), distinguished by the relations which their semantic structures bear to the visual system's fundamental neural response categories for color." (Kay and McDaniel 1978: 637)

The possible elimination of the upper limit of eleven basic colour terms is relevant because in Russian *sinij* 'blue' and *goluboj* 'light blue' are both basic, meaning there are twelve basic colour terms. If either *helesinine* 'light blue' or *tumesinine* 'dark blue', or both, were to prove to be basic colour terms in Estonian, then the upper limit of eleven would be surpassed, as it would by the Turkish *lacivert* 'dark blue'.

Kay, Berlin and Merrifield (1991: 17) used the experimental data from *The World Color Survey* to reduce the number of composite categories from 63 to the eight composite categories (1978) found in *The World Color Survey*: red/white/yellow, green/blue/black, yellow/green/blue, red/yellow, green/blue, white/yellow, and yellow/green, blue/black. The possible ninth composite category yellow/green/blue/black was not attested.

A language can evidently have two or even three basic terms for one colour region, and opinion is divided on how to relate this to the original theory. While Kay and McDaniel (1978) had doubts about the accuracy of the upper limit of eleven basic colour terms, most researchers, even if they find it reasonable to raise the upper limit beyond eleven basic colour terms, usually find reassurance in the Berlin and Kay criteria for basic colour terms, which more often than not help to disqualify possible basic colour terms. The criteria, which were based on the data analysed, are not immutable, and it may be better to look at them as guidelines rather than irrefutable facts. Mylonas and MacDonald (2015) suggest, for example, extending the basic colour terms in English to 13 to include *lilac* and *turquoise*, while Zimmer (1982) suggests that the German *türkis* 'turquoise' could be a basic colour term.

To reiterate the point, the theory of basic colour terms argues that there is a limited number of universal colour words that designate general colour categories. While the original notions that basic colour terms appear in a strictly ordained sequence called the *temporal-evolutionary ordering of basic colour terms* (see Berlin and Kay 1969: 4) and that the upper limit of possible basic colour terms is eleven (see Kay and McDaniel 1978) have become somewhat antiquated, the adherence to the original criteria of the basic colour terms is sometimes still rigid. There are studies that advocate a more data-driven approach, using the definition of basic colour terms as a guideline, and some researchers even combine the universalist and relativist notions within one study. This could be part of a more general trend towards quantitative methods. According to Gries (2013: 4) the number of studies using quantitative methods has been increasing in all areas of linguistic research, and the field of linguistics is experimenting a paradigm shift towards more empirical methods.

The rest of this chapter gives a brief review of the most relevant literature and introduces the main points made by the most notable advocates and critics of the theory. The literature reviewed is chosen for its importance to the theory as it explains the most important updates and approaches to the theory of basic colour terms. It finishes with various comments on the original definition of basic colour terms and offers a preliminary view of how the exceptions could be situated within the theoretical paradigm.

The approval of the mainstream scientific community has swung like a pendulum between the relativist and universalist viewpoints. For an excellent recent introduction to the debate on language and thought through the example of colour, see the article *Language and thought: Which side are you on, anyway?* (Regier *et al.* 2010).

The basic colour terms theory, which argues that there are a certain number of colour universals, was introduced by anthropologist Brent Berlin and linguist Paul Kay in *Basic color terms: their universality and evolution* (1969). A considerable amount of research has been generated by the basic colour terms theory from the earliest critics to the newest studies. Although the research from the last decade indicates a trend towards cross-linguistic comparison or in some instances offers a more data-driven approach, the theory still generates research debating the relativist or universalist nature of the colour terms.

Research by Terry Regier and Paul Kay (2003; 2007; 2009), who use computational methods, retains a more balanced view between the universalist and relativist colour paradigms. On the matter of colour naming and perception they conclude that the Whorf hypothesis (1941), which proposes that the semantic categories of our native language filter our view of the world, is half right. Regier and Kay (2009) conclude that in the domain of colour naming and perception language affects colour perception primarily in the right visual field, probably through the activation of the language regions of the left hemisphere, and colour naming across languages is shaped by both universal forces and language-specific forces, or local determinants. They comment that neither of these findings was anticipated by either side of the universalist-relativist argument, and the idea that Whorf was half right, but only half, does not particularly suit either side. (Regier and Kay 2009)

In a series of studies, Regier and Kay applied computational methods to *The World Color Survey* and related data and demonstrated that universal constraints beyond grouping by similarity operate in colour naming across the world's languages. Kay and Regier (2003) questioned whether colour categories across languages tend to cluster in colour space at rates greater than chance would allow. They used computational methods to answer whether the empirically observed dispersion in *The World Color Survey* data was significantly less than would be expected by chance in their randomised computer simulation. The resulting randomised theoretical version of *The World Color Survey* dataset was compared to the actual empirical dataset. Their analysis revealed that the actual World Color Survey categories are clustered across languages to a greater degree than would happen by chance, and , moreover, that the categories cluster near those in the data of Berlin and Kay (1969) to a degree greater than chance would suggest. It should be noted that *The World Color Survey* (2010) is more or less dedicated to the idea of universal colour terms, so it would be surprising indeed if the data did not support the universalist paradigm. Kay and Regier's research is only one example that bridges the division between the universalist and relativist viewpoints, which is too often painted in stark black and white, and Regier, Kay, and Khetarpal (2007) draw on the proposal of Jameson and D'Andrade (1997). For multidimensional scaling of colour terms from *The World Color Survey* see Bimler (2011). There are many studies that either support the pro-universalist evidence or pro-relativist evidence, and in some rare cases the evidence may be interpreted either way. While some researchers tend to prefer one approach or the other, the newer data-driven research is slightly less theory-bound.

Jameson and D'Andrade (1997: 313) argue that the irregularities of the perceptual colour space provide an informational advantage for the first six of the Berlin and Kay evolutionary sequence in dividing categories so that the category foci are maximally different from each other, as their additive complementary hues are approximately opposite each other in colour space. Jameson and D'Andrade consider that the best illustration of the irregularity of the perceptual colour space is the non-perfect spherical shape of the diagrammatic representation of the Munsell colour solid with one quarter removed, "where the shape of the color solid deviates markedly from that of a sphere" (1997: 312). They say the most informative naming system is that with the maximum distance between established terms in colour space (1997: 312), and they consider it plausible that "people are using the *perceptual* structure of the stimulus space which is directly available to them" (1997: 313). In place of the neurological explanation in the Berlin and Kay evolution of the first six terms, Jameson and

D'Andrade (1997: 313) interpret those terms as successively maximising the information in the perceptual colour space (see also Boster 1986).

The research of Franklin and Davies (2004) has a pro-universalist slant for example. Their study suggests that categorical colour perception is present in pre-linguistic infants and thus indicates that some form of innate category boundaries exist. This implication is not well received by proponents of the relativist view, which gains support from one of the most frequently cited research papers on the case of Berinmo in Papua New Guinea, which supports the relativist position regarding colour terms. Roberson, Davies, and Davidoff (2000) found that Berinmo colour category boundaries differ from those of English and, furthermore, that native speakers exhibit categorical perception between native colour categories, i.e. Berinmo speakers mark the boundary between two Berinmo colour categories and English speakers note the boundary between English *yellow* and *green*. That native language suggests native colour boundaries was indicated earlier by Kay and Kempton (1984), who also found that these results did not apply when the spontaneous activation of colour names was inhibited, which suggests that the colour names direct the categorical perception of colour. This suggestion is confirmed by newer research, most notably Özgen and Davies (2002), Roberson, Davies and Davidoff (2000), and Winawer *et al.* (2007). If it is true both that language affects colour perception and infants have colour boundaries pre-language, then the rigid division between the relativist and universalist views of colour becomes less and less strict.

The theory of basic colour terms is opposed by proponents of natural semantic metalanguage (NSM), most notably Anna Wierzbicka and Cliff Goddard (Goddard and Wierzbicka 1994), who promote an approach based on the concepts of Andrzej Bogusławski. Wierzbicka maintains that there are a certain number of irreducible concepts of human conceptual universals that are essential to communication. To describe the human conceptualisation of colour needs the human conceptual universal of SEE to be employed. Wierzbicka believes that colour, and equally therefore basic colour terms, is not universal (Wierzbicka 1996: 287–288), and to Wierzbicka it is clear that "a definition of 'colour' would have to be based on the concept of SEEing" (1996: 299). She favours the approach of natural prototypes, where the most salient example of a colour perceptually must become linked conceptually with a noticeable reference point, such as black with dark, white with light, blue with sky or sea, green with vegetation, red with blood, yellow with the sun, or brown with the ground. In support of Swadesh's (1972: 205) speculation that fire and light are at the beginning of the human conceptualisation of colour, Wierzbicka proposes a hypothesis which correlates light, sun and fire in all their aspects, including brilliance and luminosity, with the macro-white category (1996: 322). Macro-colour categories (see Kay 1975: 260) are of dubious value to Wierzbicka, because "it is quite obvious: if a word is used to describe not only black, but also brown, grey, or dark-blue objects, then it cannot possibly mean only 'black'" (1996: 287). Wierzbicka (1996: 329–230) points to semantic change, where locally salient or particularly important references may change in time, as a colour word "may dissociate itself from its etymon altogether and attach itself exclusively, in the speakers' linguistic consciousness, to a different, more salient perceptual model", e.g.with the Russian *goluboj,* which has moved from pigeon to sky. Wierzbicka states that "the choice between linguistic arbitrariness and neurophysiological determinism in colour categorisation is a false one" (1996: 334).

The proponents of the relativist view have criticised the assumptions made by the colour universalists, so it is unsurprising that the study that can function as a reference guide to the universalist angle, *The World Color Survey* (2010), is also used as a reference point for their criticism. *The World Color Survey* is a large-scale collaborative project by William Merrifield and the Summer Institute of Linguistics, which was the work of decades. It has almost no references to any field-studies done by outside researchers who were not part of the project. It was authored by Paul Kay, Brent Berlin, Luisa Maffi, William R. Merrifield and Richard Cook and started in 1976 to check and expand the findings of Berlin and Kay (1969) in a full-scale field study. Described by C.L. Hardin as "the largest and most systematic color-term database ever assembled", it consists of denotations of colour terms in 110 minor and tribal languages. According to the information in Hardin's foreword it is therefore the most fundamental response to the critics of basic colour theory. (Kay *et al.* 2010)

The following quote is a concise summary of the base assumption made by the universalists. The authors sum up the long-standing assumption about the original model:

> "BK[2] operated on a tacit assumption, which was retained in the UE[3] models throughout the 70s, 80s and 90s, that every language *has* a basic color term system, in the sense that every language has a small set of words, or word senses, of pure color meaning whose significata partition the subjective color space." (Kay *et al.* 2010: 2–3)

The previous assumption has been explicitly criticised by Maffi (1990) and Levinson (2000) and implicitly by Lyons (1995), Lucy (1997), and Barbara Saunders and van Brakel (1997). According to Paul Kay (MITECS) the theory has mostly been challenged by anthropologists (Hickerson 1971; Durbin 1972; Collier 1973) on experimental grounds, but has mostly been embraced by psychologists (Brown 1976; Miller and Johnson-Laird 1976; Ratliff 1976).

Color and Cognition in Mesoamerica (1997) by Robert E. MacLaury, his magnum opus, introduces vantage point theory, which is a model of a method of categorisation that "a person employs to construct any category, to use it, to change it, or to recall it" (MacLaury 1997: 494). The English colour category

---

[2] BK = Berlin and Kay (1969/1999).
[3] UE = universals and evolution model(s) of basic colour terms.

"red" consists of the minimum three coordinates: the inherently fixed image of the purest red envisionable (R) plus inherently mobile emphases on similarity (S) and difference (D), where the similarity shows the extent of the category range and the difference curtails the extent of the category range and delimits the range at a margin (MacLaury 1997: 495–496). MacLaury's research included the Hungarian reds *piros* and *vörös*, and he considers the latter a term of cultural importance (1997: 499). For more details on Hungarian red terms see also Uusküla and Sutrop (Uusküla and Sutrop 2010; 2007), Uusküla (2008), Uusküla and Sutrop (eds) (2011), Benczes and Tóth-Czifra (2014).

In her dissertation Mari Uusküla (2008) is one of the many authors to have tried to redefine a basic colour term, but her definition is backed up by a large amount of fieldwork, and so while remaining faithful to Brent Berlin and Paul Kay's tradition, she redefines the basic colour term:

> "A basic colour term is a semantically consistent and psychologically salient term, which appears in the idiolects of all language speakers. It has a tendency to occur in the beginning of elicited colour term lists. In reference to a certain colour, native speakers use the term consistently. Its meaning is not included in the meaning of other basic colour terms. In some exceptional cases the term may be restricted to a narrow class of objects, but is granted the basic status if it meets the criteria of psychological salience in the language/culture under consideration." (Uusküla 2008: 29)

The definition uses the phrase "semantically consistent" to describe basic colour terms and the possibility that a basic colour term may not be monolexemic is reflected in Uusküla's research (Bogatkin-Uusküla and Sutrop 2007; Uusküla 2007; Sutrop 2000; 2011). Grossmann and D'Achille (2016: 22) have a particularly well-articulated, jargon-free list of the characteristic features of basic colour terms: from the morphological point of view basic colour terms are not complex; from the semantic point of view they are not transparent, they are not hyponyms of other terms and their application is not restricted to specific classes of entities; and from the psychological point of view they are salient to speakers (Grossmann and D'Achille 2016: 22). They continue with a comparison of nonbasic colour terms that can be analysed morphologically, are semantically transparent and are generally more recent than the basic colour terms (Grossmann and D'Achille 2016: 22).

The following chapters focus on experimental studies investigating the possible division of a colour category into not one but two basic colour terms.

The last approach to be discussed in this sub-chapter is the diachronic evidence aptly described by Carole Biggam, whose historical approach is applicable beyond the synchronic limitations of the theory of basic colour terms. *The Semantics of Colour: A Historical Approach* (2012) is a detailed, almost all-encompassing book on colour driven by a fascination for the classification, labelling and communication of colour.

The book reviews the criteria for basic colour terms, but the self-confessed historically-biased view of the author means that special interest is paid to the criteria “which can be applied in historical studies” (Biggam 2012: 22). Biggam’s chapter 3 (2012: 21–41) is dedicated to dissecting the criteria and offers nineteen “potentially useful criteria” (2012: 41). However many criteria there are, Biggam (2012: 43) concludes that the researcher should firstly never assume that the basic colour terms are obvious and secondly in some cases consider the colour terms as a spectrum of vocabulary ranging from high frequency to rarity, rather than imposing strict restrictions between basic and non-basic terms.

Biggam (2012: 33–41) considers that among the original criteria of Berlin and Kay, (i) hyponymy, (iii) contextual restriction, and (viii) morphological complexity are particularly valuable in historical studies, but she also suggests the additional criteria of frequency of occurrence in speech and text, type modification for primary basic colour terms, and embedded expressions. Biggam also reviews the criteria which are meant for modern, living languages. Among the original Berlin and Kay criteria, these are (i) a non-predictable or monolexemic character, for psychological salience, and (iv) elicited lists, consensus, consistency and idiolectal evidence, and also the four original secondary criteria of derivational morphology, homonymy, recent loan words and morphological complexity. Additional suggestions from other researchers inspire potentially useful criteria suggested by Biggam (2012: 33–41):
1) expression length;
2) frequency in texts;
3) frequency in speech;
4) response time;
5) type modification;
6) domains of expressive culture;
7) embedded expressions;
8) cultural-historical significance.

The last criterion, cultural-historical significance, is of particular interest for this thesis because it is represented in Russian (2012: 40–41) by *goluboj* ‘light blue’. If basic status were denied to *goluboj*, that would mean that the colour of certain blue entities could not be described at all by a basic term with cultural salience (Biggam 2012: 40), because Paramei’s research (2005; 2007) has shown *goluboj* ‘light blue’ to be ‘culturally basic’. Using a Russian corpus study on the constraints in taxonomic class combinability, Rakhilina and Paramei (2011) suggest a complementary linguistic criteria, which is the combinability of the colour term with the names of both artefacts and natural objects.

Kerttula (2007: 152–153) uses primacy, frequency, application and derivational productivity to assess the “relative basicness of colour terms” from dictionaries and corpora. If colour is the primary sense of a term, and if it is a hyperonym, then the primacy of the colour word is of the highest of four possible levels; frequency means the number of occurrences and application means the number of referents, which usually correlates with frequency; and derivational productivity refers to the number of derivates and compounds formed from it (Kerttula 2007: 152–153). Kerttula (2007: 156, 161) argues that qualifiers like *dark* or *pale* are an integral part of colour term evolution. Twenty colour terms with the highest values of relative basicness include words which are both colour names and qualifiers to other colour names: *white* (38), *red* (35), *black*, *yellow*, *green*, *blue*, *brown* (29), *purple*, *pink* (28), *grey*/*gray* (27), *dark* (25), *golden* (24), *pale*, *silver* (23), *orange*, *crimson* (20), *rose* (19), *ochre*, *tawny*, and *violet* (17). In the Finnish *Turun Sanomat Corpus* (1999) *vaaleansininen* 'light blue' (13) is "rather established", coming second after *vaaleanpunainen* 'light, pale red' (55), which Kerttula (2007: 160) interprets as further proof of the Finnish preference for cool colours. Oja (2002: 254) notes that the semantic fields of the words expressing the notions of 'light', 'bright' and 'dark' are not alike in different languages, as the equivalents for the Estonian word *helesinine* are the English *light blue* and *bright blue*, the Finnish v*aaleansininen* and *heleänsininen*, the German *hellblau*, and the Russian *светло-синий* [*svetlo-sinij*] and *голубой* [*goluboj*].

The aim of this theoretical introduction is to orient the reader in the paradigm of basic colour terms, so that it is then possible to understand the proposed exceptions to the theory and the approach taken in explaining them. For an in-depth look into the theory of basic colour terms, see the original monograph *Basic color terms: their universality and evolution* (Berlin and Kay 1999), or the empirical studies collected for decades as the handbook *The World Color Survey* (Kay *et al.* 2010).

1.2 Previous studies and the research question

This next section aims to provide examples of experimentally-proven relevant research into the basic colour terms theory paradigm of Berlin and Kay. It discusses empirical studies of languages that may have more than eleven basic colour terms, especially if they having more than one basic colour term for designating the blue colour space. The focus on there being more than one general term for blue arises from the proposed hypothesis, which dictated the selection of languages for the experimental part of the thesis.

Whether or not a term is considered basic is dependent on the data; when the data suggest that a term may be exceptional, the chance of it being accepted as such is almost always diminished by the criteria of the basic colour terms. Although the defining criteria have been debated heatedly, most researchers stay true to the original work.

There are instances of studies where the modifiers have been removed from the data that can be analysed, but this alters the results, since it is usually stated that removing the modifiers could be seen simply as a part of the methodology. The some-or-all approach of modifiers omitted or not allowed seems to be favoured in some works like the study *Evolution of semantic systems* (EoSS) by the Max Planck Institute for Psycholinguistics (Majid, Jordan, and Dunn 2013).

The EoSS research is designed to reveal the evolutionary path of terms in many languages, in which case the preferred choice is to remove the modifiers. I study the basic colour terms in three languages. Basic colour terms are usually simple terms with high frequency of use in the list task and high levels of agreement between participants in the stimuli tasks. Using basic colour terms is also indicative of their status, because basic colour terms can take more modifiers than other terms. For example, *beige* is a simple term, which is listed in all three languages in this thesis, but it is not combined with modifiers as often as the established basic colour terms are.

If the modifiers were to be removed from the data before the analysis, then all the terms would be more similar to basic colour terms. I see the approach of removing all modifiers as bordering on biased, because removing the modifiers renders all the terms more similar to basic colour terms and raises the overall consensus. If all the answers given are written down exactly as they were uttered by the participants, e.g. Estonian *veripunane* (Nom *veri*) and *verepunane* (Gen *vere*) 'blood red', which I recorded and counted as different terms in the analysis, then the overall consensus of the data is reduced. The basic colour terms reveal themselves as separate from the others by being simple terms, in most cases, that are used with higher frequency and are combined with modifiers more than other terms are.

Removing the modifiers from the data also makes comparison between studies slightly more unreliable. To uncover whether the original data were changed somehow for the data analysis may require careful reading, though some cases, Davies *et al.* (1997: 190–191, 194–195) for example, provide two tables of summarised labelling or display the frequencies at which a simple form was modified in brackets. However, in the Turkish study (Özgen and Davies 1998) there is no way of gauging of how the reduction of modified terms affected the study. In this pertinent case of a previous study of Turkish colour terms, it is written that:

> "The majority of terms offered in all cases were monolexemic. However, it was also common to combine the simple form with a general modifier such as *acik* 'light' and *koyu* 'dark'. The most frequent occurrence of a simple form plus modifier was *acik mavi* 'light blue'. Here we have collapsed all such constructions onto the simple form." (Özgen and Davies 1998: 925)

That no modifiers were omitted from the analysable data accounts for some of the differences from the previous research (compare to Özgen and Davies 1998). The approach of retaining all the modifiers in the data can be considered something of a departure from the standard, if such a thing exists. It is debatable whether all modifiers were removed, or only *acik* 'light' and *koyu* 'dark', but it has undoubtedly influenced the results by increasing the consensus, and possibly increasing the consensus in the naming task considerably. When comparing results, it is useful to keep in mind that the different authors may modify the original data differently before data analysis. In *The World Color Survey* there is a reference to modified terms having been "subsumed under the forms presented here" for the Yup'ik (Eskimo-Aleut) colour data (Kay *et al.* 2010: 575), which suggests that the original data had modifiers that were omitted. It is unclear if the same modus operandi is used elsewhere, or throughout, but it is a clear indication that not all modifiers are presented in *The World Color Survey* (2010).

1.2.1 Turkish *mavi* and *lacivert*

This section gives a brief introduction to the words *mavi* 'blue' and *lacivert* 'dark blue' (Rätsep 2011: 354). *Mavi* 'blue' is derived from the Arabic words *mā'* 'water' and *māhī* (*māwī*) 'watery'(Wehr 1979: 1094). The colour meaning is a Turkish addition, where the second meaning of *mâvi* blue' in the Ottoman language comes from the colour of water (Devellioglu and Güneycal 2006: 556, 573). For the designation of dark blue, almost a bluish black, there is a specific word in Turkish, *lacivert* 'dark blue', which is borrowed from Persian *lāzhuward* 'Lapis lazuli', *lāzhuwardī* 'of a blue colour, of the colour of lapis lazuli' (Steingass 1892: 1111). The French *azur* comes from the Persian *lazuward* 'lapis lazuli,' which in turn comes from the Arabic *lazard*, which is related to the Arabic *azraq* 'blue' (Borg 1999).

The most relevant previous study of Turkish basic colour terms was presented by Emre Özgen and Ian Davies (1998). The authors conducted three experiment-based studies of Turkish colour terms. In the first experiment, 80 children aged 8–14, 118 students aged 19–25, and 35 adults aged 20–38 completed a time restricted written list task with five minutes to 'Write down as many colour terms as you can'. In reporting the list task results the authors comment that they have collapsed all the simple terms used with a general modifieriers, i.e. *acik* 'light' and *koyu* 'dark' onto the simple form (1998; 925). This trend seems to continue through the whole article, with all modifiers eliminated. In the second experiment "a subset of the child and adult samples" from the previous experiment with 17 children and 33 adults altogether took part in the colour naming task conducted using the general method of Davies and Corbett (1995) for establishing basic colour terms. Özgen and Davies report that "measures of salience and consensus derived from the two tasks converge to suggest that Turkish has 12 basic color terms" (1998: 919). Besides the list and colour naming tasks for establishing Turkish basic colour terms, Özgen and Davies performed a third experiment where 125 university students aged 19–24, 109 of them female and 15 male, were tested during a class. They were asked to "write down as many kinds of *mavi* as they could think of" and after they had finished that list to "write down whether *lacivert* was a kind of *mavi*". The results showed that 57% of the participants included *lacivert* 'dark blue' in their lists of types of *mavi* 'blue', and moreover, 85.5% regarded *lacivert* 'dark blue' as a kind of *mavi* 'blue'. (Özgen and Davies 1998: 942). These results suggest that *lacivert*'dark blue' violates Brent Berlin and Paul Kay's type-of criteria for basicness, which states that basic colour term signification is not included in that of any other colour term (Kay and Berlin 1999: 6).

Their overall consensus was higher because they collapsed all modified constructions, e.g. *acik mavi* ‘light blue’ onto the simple form (Özgen and Davies 1998: 925), but even accounting for the removal of all modifiers the BV tile adult sample (N=33) consensus was incredibly high at 93.93%; for the child sample it was much lower though still over 50% at 56.25%, but the stimulus BVB gained dominance with 66.67% of the adult sample agreeing on *lacivert* ‘dark blue’ (1998: 937). The summary that “the safest conclusion is that Turkish has 11 basic colour terms” (1998: 951) was probably reached for two reasons:

1) 85.5% of university students (N=125) in a classroom answered ‘yes’ to “whether *lacivert* is a kind of *mavi*” (1998: 942), therefore violating the non-inclusion characteristic of a basic colour term (Berlin and Kay 1999: 5–6);
2) the child sample consensus was weaker, e.g. listing (N=80) consensus of 47.5% and naming (N=16) consensus of 56.25% (1998: 926, 933).

My research had lower consensus for *lacivert* ‘dark blue’, suggesting that the basicness of *lacivert* by stimulus naming consensus is weaker than in previous studies, but the inclusion of all modifiers is a significant factor that negatively affects consensus.

Şahin (1998) also carried out a large-scale study (N=322) on Turkish colour terms. The participants were instructed to match the Munsell stimuli to the colour names supplied with eight basic and 24 non-basic colour names, including *lacivert*‘dark blue’. Şahin (1998: 167) considers *lacivert* to be ‘navy blue’, a non-basic colour term that is known by 98.8% of the respondents. Şahin (1998: 176) remarks on how *lacivert* was picked up as quickly as a basic colour term and that it can be a potential basic colour term. In a newer, co-authored article, the authors relate how the colour samples constituting the range for *lacivert* show statistically significant (p-value: .001) similarities to the range for *gece mavisi* ‘night blue’ (Şahin Ekici, Yener and Camgöz 2006: 474).

1.2.2 Estonian *sinine*, *helesinine* and *tumesinine*

The Estonian Etymological Dictionary (Metsmägi, Sedrik and Soosaar 2012: 474) defines *sinine* as the colour of the “cornflower, linseed blossom, cloudless sky”, with the etymological derivation given doubtfully as being from the Proto Iranian *šīn* “blue; green” and Persian *χašīn* “bluish, bluey”. Blue has the same root as the Estonian *sinine* ‘blue’ in many of the related languages, e.g. the Livonian *si’ņņi*, Votic *sinin*, Finnish *sinine*, Ingrian *sinniin*, Aunus-Karelian *sinine*, Ludian *šińińe*, Veps *sińińe*, Erzya *seń*, and Moksha *śeńəm* (Metsmägi, Sedrik and Soosaar 2012: 474).

The Estonian basic colour term for ‘blue’ is *sinine*. Sutrop (Sutrop 2002: 73) remarks how the existence of two basic terms for blue in Russian, with *goluboj* ‘light blue’ complementing *sinij* ‘blue’, could have influenced the Estonian concept of blue so that it was divided into two or even three separate sub-concepts. Specifically, Sutrop considers the Russian *sinij* ‘blue’ and the Estonian term *sinine* ‘blue’ to be homonymous, and comments on how “Russian influence appears to be the destabilising factor on the Estonian concept of ‘blue’” (Sutrop 2002: 217). In fact, in some Karelian, Vepsian and Votic dialects the ‘blue’ terms have changed to mimic the Russian categorisation of ‘blue’ (Oja 2007: 207).

Triin Kalda (2014) conducted interviews with Estonian, Finnish and Finnish-Estonian participants for a comparative analysis of their colour lexicon that included a sorting task. Kalda interviewed Estonians (N=20), Finns (N=22), and Finnish-Estonians (F=20), who completed list, naming and sorting tasks. In the naming and sorting tasks 65 Color-Aid Corporation stimuli were used (for the selection of stimuli see Davies and Corbett 1995) . Neither *helesinine* ‘light blue’, nor *tumesinine* ‘dark blue’ were among the most frequent (F ≥ 2) group names in the Estonian and Finnish-Estonian sorting tasks (Kalda 2014: 53–54).

Kalda (2014: 58) lists *helelilla* ‘light purple’, *heleroheline* ‘light green’, *tumepruun* ‘dark brown’, *tumeroheline* ‘dark green’, *helesinine* ‘light blue’, and *helepruun* ‘light brown’ among basic colour term candidates, but disqualifies them because their meaning is evident from their components. In the combined analysis *helelilla* ‘light purple’, and *heleroheline* ‘light green’ pass the naming and sorting task thresholds, which I find quite remarkable (Kalda 2014: 58–59). The modified browns, *helepruun* ‘light brown’ and *tumepruun* ‘dark brown’, did not pass the sorting task threshold, and were generally grouped together (Kalda 2014: 59). The assumption that several pink and purple groups would be formed in the sorting task was not confirmed, and the results of the sorting task indicate that reds and pinks, blues and purples, and yellows and oranges were grouped together (Kalda 2014: 77).

Sutrop (1995; 2000; 2011) conducted an unconstrained listing task and colour naming task (65 Colour-aid stimuli, N=80). The listing percentage for *helesinine* ‘light blue’ was 35%, as it was listed 28 times by 80 participants, and the percentage for *tumesinine* ‘dark blue’ was 27.5% (F=22), while in the subsequent naming task *helesinine* did not gain dominance, and *tumesinine* had one dominant stimulus, BV (blue-violet), with 52.5% of consensus in naming. The conclusion reached by Sutrop is based on the definition of the basic term, which eliminates both *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’ because their meaning is deducible from their components and so we may delete the terms from the list of basic colour term candidates (Sutrop 2011: 80). However, he comments on how the basic term *sinine* ‘blue’ is psychologically highly salient, while *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’ also have some basic traits, both clearing one threshold for basicness. The term for light blue cleared a hurdle according to the frequency measure and the term for dark blue was dominant for the tile BV at the 50 % consensus level (Sutrop 1995: 164).

1.2.3 Russian *sinij* and *goluboj*

Berlin and Kay (1969) originally proposed two exceptions to the basic colour term upper limit of eleven. The two languages concerned were Russian with *goluboj* 'light blue' and Hungarian with *vörös* 'red'. While the Russian "second basic blue" *goluboj* has mostly been accepted as basic or culturally basic (Paramei 2005), Uusküla (2008; Uusküla and Sutrop (2007) argues against *vörös* being basic in Hungarian.

Supported by personal communication from D. Slobin, Berlin and Kay tentatively suggest the following 12 terms as basic in Russian: *belyy*[4] 'white', *chërnyy* 'black', *krasnyy* 'red', *zelënyy* 'green', *siniy* 'blue', *goluboy* 'light blue', *korichnevyy* 'brown', *purpurnyy* 'purple', *rozovyy* 'rose, pink', *kirpichnyy* 'orange' and *seryy* 'grey' (Berlin and Kay 1999: 98–99). Later research suggests *fioletovyj* 'purple' and *oranževyj* 'orange' as basic terms, considering the use of *kirpčnyj* to be erroneous, or suggesting the use of *oranzhevyj* (Frumkina and Mikhejev 1983: 55; Corbett and Morgan 1988: 27). The results of the five-minute list task (N=31) by Morgan and Corbett (1989) and Morgan (1993) support the basic status of both *sinij* 'dark blue' and *goluboj* 'light blue' (93.5% listing frequency, rank 4), and the authors state that the prominent position of *sinij* and *goluboj* 'light blue'reinforces the view that each of these terms is basic (Morgan and Corbett 1989: 140).

Inconclusive interviews by Berlin and Kay (1999: 36) suggested that for some Russian participants *sinij* 'blue' marks two categories, one including *goluboj* and the other contrasting with it *goluboj* 'light blue'. Thus, depending on context, *sinij* can mean 'blue' or 'dark blue'. Davies and Corbett (1994: 87) conclude that *goluboj* 'light blue' and *sinij* 'dark blue' denote "nonoverlapping regions of color space rather than *goluboj* 'light blue' being included in the domain of *sinij* 'dark blue', as Berlin and Kay originally thought". In the list task of Davies and Corbett (1994) *goluboj* was listed by 94.8% of the participants, was ranked 4.5 and had a mean position of 7.5 (rank 5), while in the naming task (N=54) the stimulus BGB T3 was dominant with consensus of 72.2% and a specificity index of 0.57. They conclude that providing their measures are valid indicators of basicness, then we must accept that both terms are basic (Davies and Corbett 1994: 87).

> This claim for basic status of the 12 terms given above is most clearly supported by the data from the list task. These terms are each offered by more people than any other term, and there is a reasonably clear "step" between the least frequent of these twelve – *seryj* 'grey' – which was offered by 69% of the Sample, and the most frequent of the remaining terms – *sirenevyj* 'mauve' – which was offered by 52% of the sample (Davies and Corbett 1994: 86).

---

[4] Berlin and Kay used an older convention for transliterating the Cyrillic, thus the terms all end "-yy" rather than "-yj" compared to ISO 1995 tranlsiteration. Different authors may use different tranliterations.

Provided that “colour-term inventories can be extended beyond the upper limit of 11” (Davies and Corbett 1994: 86–87), the data of Davies and Corbett support the claim of 12 basic colour terms in Russian: *belyj* ‘white’, *černyj* ‘black’, *krasnyj* ‘red’, *zelenyj* ‘green’, *želtyj* ‘yellow’, *sinij* ‘blue’, *goluboj* ‘light blue’, *koričnevyj* ‘brown’, *fioletovyj* ‘purple’, *rozovyj* ‘pink’, *oranževyj* ‘orange’ and *seryj* ‘grey’.

Research by Andrews (1994) also supports both *sinij* and *goluboj* as “bona-fide basic terms in standard Russian”, but adds a new facet, namely that this treatment is fixed by adulthood, because “among the younger émigré adults, however, there is definite evidence of semantic shift, the result of interference from English blue”. Studying Russian children aged three to six, Davies at al. (1998: 412) found that *goluboj* vied for primary status in the list task, but in the colour term production task it belonged with the derived terms, e.g. *sinij* and *goluboj* were “offered relatively frequently on the list task, but scored relatively poorly on the comprehension task” (Davies et al. 1998: 413). Laws, Davies and Andrews (1995) compared English and Russian blues in a colour discrimination task, a pair-similarity judgement task and a sorting task, but “little evidence was found for any of the predicted differences in behaviour”. They attributed this partly to “English participants also attending to differences between dark and light blues” (Laws, Davies 1995: 59). Their sample size was very small though, with between five and nine respondents for each task and language. However, Winawer *et al.* (2007) tested 26 native Russian speakers and 24 native English speakers, who completed a colour discrimination task with verbal and special interference blocks, and the results indicated that Russian-speakers were faster to distinguish between the two colours if they fell into different linguistic categories in Russian, one *sinij* and the other *goluboj*, than if the two colours were from the same category (Winawer *et al.* 2007: 7783). This category advantage was eliminated by a verbal dual task, but not by a spatial one, suggesting that the language had an effect. English-speakers tested on the identical stimulus did not show a category advantage under any condition, meaning colour discrimination performance “differs across language groups as a function of what perceptual distinctions are habitually made in a particular language” (Winawer *et al.* 2007: 7783).

1.2.4 Other Slavic and Eastern European languages

Hippisley (2001) repots that Ukrainian and Belarusian also have two basic terms for blue, but both lack a basic term for orange:

> “In examining whether the innovative second BLUE category is unique to Russian, or whether it is a broader characteristic of Russian’s family, Slavonic, we have produced some evidence, using a recognised psycholinguistic test, that both Ukrainian and Belarusian have two basic terms for BLUE. The first term is inherited from East Slavonic *sinij* denoting ‘(dark) blue’, and the second is a

> borrowing from Polish *błękitny*, which in both languages means 'light blue'". (Hippisley 2001: 175)

In the list task as a test of psychological salience Hippisley (2001: 165) suggests that Belarusian does not have a basic term for orange but has two basic terms for blue, *sini* 'blue' and *blakitny* 'light blue', which also appears to be basic (85.7% list task percentage[5], average place on the list 6.2). Hippisley (2001: 161) also considers that in Ukrainian *blakytnyj* 'light blue' is a second basic blue beside *synij* 'blue', as *blakytnyj* 'light blue' had a high list task (N=34) percentage of 88%, a high rank of 10.0, and a high average place on the list at 7.3:

> "The results of the list task for both Ukrainian and Belarusian strongly suggest that these languages have two basic terms for BLUE. Ukrainian has *synij* 'blue' and *blakytnyj* 'light blue', and Belarusian has *sini* 'blue' and *blakitny* 'light blue'. This in the context of both languages having all the Berlin and Kay basic color terms, except for ORANGE". (Hippisley 2001: 168)

Vasyl' Starko (2013: 150) proposes that the colour system of Ukrainian is special in that it has as many as three basic words, *synij*, *blakytnyj* and *holubyj*, for what is generally called *blue* in English. Starko (2013: 153) used *Ukraïn 'skyj asociatyvnyj slovny* 'A Ukrainian Associative Dictionary' by Martinek (2007) to present the results of free association experiments (N=200). The results indicate that the sky is the dominant association for all three colour terms *blakytnyj* (47%), *holubyj* (30%) and *synij* (23%), but the noun *blakyt'* 'light-blue colour' has an even stronger connection (66%) with *nebo* 'sky' as the reference object (Starko 2013: 153). In the sub corpus of imaginative prose in the online *Corpus of the Ukrainian Language* (CUL) the three colour names *blakytnyj* (F=530, 11%), *holubyi* (10%) and *synij* (F=1090, 12%) are most frequently associated with the colour of the sky, where *blakytnyj* often refers to the sky on a very sunny day and *holubyi* to clear sky, while *synij* can designate lighter and darker shades of blue with bright, unclouded, clear, sunlit or saturated sky (Starko 2013: 161) concludes that while the sky is by far the most dominant reference point for all three colour terms, their reference models include multiple auxiliary exemplars and the reference models of *blakytnyj*, *holubyj* and *synij* in Ukrainian are more nuanced than is suggested by previous research.

According to Wierzbicka (1996: 326), one common way of extending the set of basic colour concepts is to use mixed colours, e.g. Polish twelfth basic colour concept *granatowy*, which is decomposable roughly into 'blue' and 'black'. One component in mixed colour concepts always seems to be more salient than the other, e.g. *granatowy* is an offshoot of *niebieski* 'blue' (Wierzbicka 1990: 326).

[5] Not all participants (N=28) listed Belarusian as their first language, but of those that did 88%, listed *blakitny*.

Danuta Stanulewicz (2010; Stanulewicz, Komorowska and Pawłowski 2014), in agreement with MacLaury (2001) states that *niebieski* 'blue' is the basic term among the Polish set of terms for blue. In the list task (N=100, time constraint five minutes) 95% of participants listed *niebieski* 'blue', 86% *błękitny* '(sky) blue', and 78% *granatowy* 'navy blue' (Stanulewicz 2010: 90). Remarkably, over half of the participants (54%) defined *błękitny* '(sky) blue' as *odcień niebieskiego* '[a] shade of blue'(Stanulewicz 2010: 187). If *błękitny* '(sky) blue' is treated as a lighter shade of *niebieski* 'blue', then it does not comply with the non-inclusion criteria of a basic colour term. Stanulewicz (2010: 193) uses MacLaury's vantage theory to explain a model with the semantic ranges of two terms referring to a single colour category, where one range, in this case *niebieski*, is larger and more centrally focused than the other, *błękitny* '(sky) blue'. It remains unclear how the vantage theory might negate the non-inclusion criterion of a basic colour term. Stanulewicz (2010: 193) concludes that *błękitny* '(sky) blue' apparently has two senses, one that is nearly the same as that of *niebieski* 'blue' while the other represents a shade of it. The results of Stanulewicz's (2010: 186) Polish association task (N=200) corroborate the idea that the sky is the prototypical reference point for *niebieski* 'heavenly; pertaining to the sky' (23.5%) and *błękitny* '(sky) blue' (20.5%), and in this they are similar to Starko's Ukrainian results (2013). Stanulewicz *et. al.* (2014: 268) state that *błękitny* 'light blue' may be treated as a semi-basic colour term, which is frequently employed when blue is used with positive connotations, as in *błękitna krew* "blue blood", and that *błękitny* 'light blue' can be used interchangeably with *niebieski* 'blue' (Stanulewicz, Komorowska and Pawłowski 2014: 269).

Kopecka and Janic (2016) consider colour denomination to be closely associated with specific cultural practices. They studied different modes through which colours are conceptualised in the artistic and the decorative domains in Polish. They (Kopecka and Janic 2016: 4–6) found 21 lexical constructions used in the artistic domain including *ultramaryna niebieska* 'ultramarine blue' (N=9) and *błękit Rembrandta* ''light blue' [of] Rembrandt' (N=4), and 23 constructions in the decorative domain including *błękit na fortepianie* 'piano blue' (N=1). *Błękit* 'light blue' (N=11, a noun used in both domains) and *niebieska* 'blue' (N=2, an adjective used in artistic domain) are listed among the key terms in both domains (Kopecka and Janic 2016: 8). In artistic samples (Kopecka and Janic 2016: 10) the second most numerous theme after intensity was pigment and *błękit kobaltowy* 'cobalt 'light blue' was the most frequent example (N=16), while *błękit Rembrandta* (N=6) was connected to the representation of a personal theme, and *błękit nieba* 'sky blue' (N=2) to the theme of celestial bodies.

1.2.5 Italian *blu*, *azzurro* and *celeste*

There are a number of notable pieces of experimental research that focus mostly on the blue category and the colour terms for blue, as this could provide exceptions to the original sequence of basic colour terms of Berlin and Kay. A part of Jodi Sandford's research (2011) on Italian blue terms includes a test that focuses on Berlin and Kay's hyponymy restrictions for basic colour terms, i.e. a basic colour term should not be a hyponym of any other colour term. She terms it a "kind of blue" survey, where subjects (N=30) were asked to combine the three terms, *blu*, *azzurro* and *celeste* in pairs, saying whether *azzurro* is a kind of *blu* for example. (Sandford 2011: 282)

The responses indicate that while 93% of Italians agreed that *azzurro* is a kind of *blu* and 100% agreed that *celeste* is a kind of *blu*, *blu* itself is not a type of *azzurro* nor *celeste* (Sandford 2011: 284). These results put *azzurro* and *celeste* subservient to *blu*, which could violate the hyperonym rule of Berlin and Kay's basic colour term definition. Corpora research by Grossmann and D'Achille (2016: 43) supports the basic colour term status of both *azzurro* and *blu* in contemporary Italian. In fact, Grossmann and D'Achille (2016) argue that *celeste*, which is "analyzable from a morphological point of view and semantically transparent" cannot by definition be considered a basic colour term. Maltese has a twelve-term colour system where *čelesti* 'light blue' from the Italian *celeste* is objectified alongside *blū* from the English *blue* or the Italian *blu* (Borg 2007: 263).

Mari Uusküla's (2014; see also Bimler and Uusküla 2014) research into the Tuscan dialect contained a list task, a naming task, a sorting task, and best-example and collocation tasks. The main conclusion reached is that Italians habitually denote the blue region with at least two salient terms and can choose to designate the lighter blues with either *celeste* or *azzurro* depending on the dialectal background of the speaker (Uusküla 2014: 76). She comments that "it is not uncommon for a native speaker to apply all three terms, including *blu*, to different denodata [denotata]" (Uusküla 2014: 76). In the list task all three Italian blues had a listing percentage over the halfway mark with *blu* at 90% (mean position = 6.15), *azzurro* at 76% (mp = 9.41), and *celeste* at 62% (mp = 9.02). Despite its high listing percentage, *azzurro* had no dominant stimuli in the naming task, while *blu*'s dominant tile was BVB with 54% consensus and *celeste*'s was BGB T3 with 57% consensus. In the 55 Color-Aid stimuli sorting task the percentages of the groups were decidedly lower with the highest being 35% for the *blu* group, while 16% of participants had both the *blu* and *celeste* groups and 11% grouped the blues into *blu* and *azzurro* or three ways into *blu, celeste* and *azzurro*. The subsequent best-example task had similarly low consensus, with *blu* having the highest agreement rate at 30%, while 11% of participants agreed on the best example of *azzurro*. In the collocation task the percentages were high for all three blue terms, but only *principe azzurro* 'Prince Charming' (100%) and *sangue blu* 'royal blood (line)' (89%) were idiomatic. The top three collocations for *blu* were *scuro* 'dark' (97%), *penna* 'pen' (92%), and *sangue* 'blood' (89%); for *azzurro* they were *principe* 'prince' (100%), *cielo* 'sky' (76%), and *occhi* 'eyes' (70%); and for *celeste* the most common collocations were *paradiso* 'paradise' (78%), *maglia* 'T-shirt' (73%), and *chiaro* 'light' (70%) (Uusküla 2014).

The dialectal background is of great importance for Italian. Paramei, D'Orsi and Menegaz (2014) conducted an unconstrained colour naming task and a best-example task with a small sample of participants (N=13, Alghero (Sardinia); N=15, Verona). Their results again provide evidence that for Italians at least two colour terms are necessary to name the blue area, *blu* and *azzurro*, with both terms behaving as basic colour term (Paramei, D'Orsi and Menegaz 2014: 33; Paggetti, Menegaz and Paramei 2015). According to Paramei, D'Orsi and Menegaz, *celeste* is a contender for a third basic colour term for the Alghero speakers (2014: 27).

1.2.6 Spanish *celeste* and Catalan *blau marí* and *blau cel*

Spanish presents a complex situation with dialectal differences and influences from and upon other languages. A study of Guatemalan Spanish by Harkness (1973), who tested both children (N=6) and adults (N=9) in a 40-stimuli task suggests *celeste* 'light blue' as a basic colour term:

> "Basic color terms in Spanish include the whole Stage VII list, with the addition (at least in the community sampled) of "celeste," or 'light blue'. (Harkness 1973: 177)

A search from *The World Color Survey* for two basic terms for the blue category revealed several examples of loans from Spanish, e.g. Aquacatec, a Mayan dialect that has separate terms for light blue *celest(e)* and dark blue *xew*, and no term for the blue category as a whole (Kay *et al.* 2010: 73–74); Kaqchikel Mayan, which has two Spanish loans *celeste* and *azul* (Kay *et al.* 2010: 135–136); and Chiquitano or Chiquito, where two Spanish borrowings have basic status for blue, *(n)asuru* "blue" and *celeste* 'light blue' (Kay *et al.* 2010: 181–182). There is evidence that the Peruvian Spanish *celeste* 'light blue' may also be basic (Bolton 1978).

In two-minute, one-word list tasks in three Spanish dialects – Castilian (N=47), Mexican (N=97), and Uruguayan (N=57) – eleven basic colour terms appeared in Spanish and Mexican, but twelve in Uruguayan Spanish (Lillo *et al.* 2018). The listing percentage for *azul* 'blue' was very high at 98% for Castilian, 94% for Mexican, and 93% for Uruguayan, while *celeste* "sky blue" was listed by 81% of Uruguayan university students (Lillo *et al.* 2018: 5), which was also the most important inter-dialectal difference in the extremes naming task and the boundary delimitation task (N=30) (Lillo *et al.* 2018: 15), where the Uruguayan *celeste* "sky blue" probably emerged because of the "influence of specific linguistic-cultural factors", e.g. use in flags or by Italian immigrants (Lillo *et al.* 2018: 17).

In contrast to previous studies, the results of Paggetti and Menegaz (2012) suggest that at least in some dialects of Spanish *celeste* is not a basic term. Paggetti and Menegaz (2012) consider *azul* to be the only basic blue colour term present in the Spanish language because reaction times in the Stroop experiment indicated that the time required to name the light and dark blue colours was not statistically different when these are used to display the *azul* term, while a statistically significant difference is observed when any other different colour-name combination, e.g. *azul* and red.

The example of Spanish Catalan is especially relevant in the context of another language, Italian. Paramei *et al.* (2014) propose that exposure to a different dialect can affect the colour vocabulary of a language. They use references to various sources of Italian dialects and suggest that an Italian dialect of Algherese Catalan might have been influenced by Spanish Catalan:

> *Celeste* may be considered a contender for a third 'blue' basic colour term for this sample exposed to Algherese Catalan dialect, a dialect that might have been influenced by the two Catalan terms for 'blue', *blau marí* 'navy blue' and *blau cel* 'sky blue' but the status of *celeste* ('relative basicness') seems to be markedly lower than that of *blu* and *azzurro*. (Paramei, D'Orsi and Menegaz 2014: 33)

Although Berlin and Kay (1969) consider Catalan a problematic example, it is not because of the blue area of the colour space (Davies, Corbett and Margalef 1995), but rather as it is the only example in their data where the status of black as a basic colour term is questionable (Berlin and Kay 1999: 42). The reason for "some doubt" with Catalan is Corson, whose informant insisted that *negre* 'black' was a kind of *gris* 'grey' (Berlin and Kay 1999: 42, 93). That sparked the research interest in other Romance languages, most notably Italian and Catalan, that may have at least two terms for blue (see Kristol 1980).

A study of Catalan colour terms by Davies *et al.* (1995) indicates that *blau marí* 'navy blue' and *blau cel* 'sky blue' have basic traits, though they are in violation of the original Berlin and Kay monosemantic and non-inclusion criteria:

> "The 'extra' blue terms of Catalan are the most basic of the non-basic terms, and they match the extra basic blue terms of Russian". (Davies, Corbett, and Margalef 1995: 47)

Children (N=40) had lower listing percentages for *blau marí* 'navy blue' (40%) and *blau cel* 'sky blue' (15%) than adults (N=40), whose percentages for listing both terms were quite high at 80% for both terms, while the positions of the terms in the lists were relatively equal, ranging from 10.8 to 12.6. In the naming task *blau marí* 'navy blue' and *blau cel* 'sky blue' were both used by over half of the adult respondents to name at least one tile (Davies, Corbett, and Margalef 1995: 36).

1.2.7 Greek *galázio*

Studies by Irwin (1974) and Maxwell-Stewart (1981) suggest that Ancient Greek had two basic term for blue. According to a large study by Androulaki *et al.* (2006), Modern Greek has twelve basic terms, including *galázio* 'light blue':

> "Taken overall, these data indicate that Modern Greek has twelve basic colour terms including two terms for blue — [γalázjo] and [blé] — and that these are glossed most appropriately as 'light blue' and 'dark blue'." (Androulaki *et al.* 2006: 39)

Androulaki *et al.* (2006) compared the domains of the Russian terms *sinij* 'dark blue' and *goluboj* 'light blue' from Moss *et al.* (1990) and the Turkish terms *lacivert* 'dark blue' and *mavi* 'light blue' from Özgen and Davies (1998) with those of Greek terms *blé* 'blue' and *galázio* 'light blue'. They found that the major difference between the two blue terms in the three languages is in lightness CIE L* axis, with a smaller difference in the chromatic plane (Androulaki *et al.* 2006: 36):

> "The lightness boundary is somewhere below about 37, and possibly as low as 30 for Turkish, the Russian boundary is probably at about 45. There are five stimuli named *mavi* 'light blue' in Turkish that are named *sinij* 'dark blue' in Russian. Both Russian and Turkish differ from Greek, where the boundary is at about 60."

Androulaki *et al.* (2006: 27) add a caveat in that the Color-Aid stimuli under-represent the *galázio* 'light blue' region, leaving gaps in the corresponding v* values between about –45 and –33 and above –20. The authors are convinced that these missing regions would include good examples of *galázio* 'light blue', and had they included these regions, *galázio* would probably have been used more frequently and, perhaps, without modifiers (Androulaki *et al.* 2006: 27).

Androulaki *et al.* (2006: 37) point to the small number of marginal terms, e.g. *laδí* 'olive green' and *béz* 'beige' that had high scores on many of the indexes of basicness, and comment on how they are "happy to regard basicness as a continuum, and certainly, for some speakers at least, these terms have high salience".

Bilingualism usually influences colour data, and this adds further influential factors. Athanasopoulos (2009) conducted a study on bilingual Greeks. Two groups of bilinguals with an advanced level (N=10) and an intermediate level (M=10) of English proficiency were tested in a naming task with 160 fully saturated Munsell colour chips. After the naming task, participants were shown the full Munsell array and asked which stimulus is the best example of *ble,* the darker shade in the blue region of colour space, and the best example of *galázio*, the lighter shade of blue (Athanasopoulos 2009: 87). Within-group naming agreement for *ble* was 77% for intermediate bilinguals and 82% for advanced bilinguals; for *galázio* the agreement was 67% for intermediate bilinguals and

65% for advanced bilinguals (Athanasopoulos 2009: 87). The category foci placement shows that the majority of intermediate bilinguals place *ble* one step away from the *blue* focus both in lightness and hue, while the majority of the advanced bilinguals shift the *ble* focus towards the *blue* (Athanasopoulos 2009: 88).

1.2.8 Arabic *samawee* and Persian *narwa*

Borg considers that the Urban Arabic vernaculars spoken in Cairo, Beirut, and Jerusalem, etc. approximate the eleven basic colour term paradigm (Borg 2007: 263), which means he considers *samawee* non-basic; Bulakh agrees for the Old-Ethiopic *samaywī* 'sky-blue', derived from *sama* 'sky' (Bulakh 2007: 250). Research by Al-Rasheed and colleagues (Al-Rasheed *et al.* 2011; Al-Rasheed 2014) suggests that the Arabic terms *zeatee* 'oil green' and *samawee* 'light blue' may merit further research:

> "Although, *zeatee* 'oil green', in the child results, and *samawee* 'light blue', in the adult data were dominant at 50% for one tile, most other possible basic colour terms achieved higher dominance scores, the specificity scores (~ 0.30) were low. These two terms may merit further investigation." (Al-Rasheed *et al.* 2011: 14)

The colour study included a list task (253 children, 200 adults) and a naming task (61 children, 60 adults). Neither *samawee* 'light blue' nor *khuhlie* 'dark blue' had high listing percentages, with the maximum being 40.6% for the child sample of *samawee*, but it was notable that in the naming task *samawee* had a 66.7% consensus by the adult sample for the stimulus BGB-T3 (SI 0.35). (Al-Rasheed *et al.* 2011)

Al-Rasheed (2014) collected data from 57 native Arabic speakers, all university students, who grouped the 320 colours by similarity and the results showed *samawee* 'light blue' had the next highest claim to being basic after the 11 probably basic colour terms.

A recent large-scale study on Persian colour terms revealed that there may be regional differences that affect the number of basic colour terms (Kandi *et al.* 2014). All the cities (N=200) – Tehran (N=50), Rasht (N=30), Yazd (N=30), Isfahan (N=30), Shiraz (N=30), Mashhad (N=30) – had black, red, green, purple, brown, pink and orange as their basic colour terms (Kandi *et al.* 2014: 6), but overall the Persian language had high levels of agreement for 11 colour terms, while there were lower levels of agreement for the terms cream, navy, and quince flower (Kandi *et al.* 2014: 9). It should be noted that the participants were "requested to avoid using 'light' or 'dark' terms and to avoid descriptive phrases like 'the color of blood clots'" (Kandi *et al.* 2014: 3). The term of particular interest, *narwa* 'navy' was "regionally basic" only in Isfahan instead of 'grey' and Mashhad, where 'blue' and 'yellow' were not basic (Kandi *et al.*

2014: 6). The authors comment on several deviations from the theory, such as the absence of terms for blue and yellow in Mashhad, grey in Isfahan, and white and grey in Shiraz, and the existence of a term for navy in Isfahan and Mashhad, and one for quince flower in Shiraz (Kandi *et al.* 2014: 9).

1.2.9 Further examples

Khomeriki, Kezeli and Lomashvili (2009) researched the categorical boundaries between the Georgian categories of the colours *tsiteli* ‘red’ and *vardisperi* ‘pink’, and *lurji* ‘blue’ and *tsisperi* ‘light blue’:

> “Among Georgian speakers the etymology of the pink and light blue terms does not influence determination of the boundaries. For observers *vardisperi* (pink) and *tsisperi* (light blue) are merely light red and light blue. They are not associated with the components of these compounds *vardi* (rose) and *tsa* (sky).” (Khomeriki, Kezeli and Lomashvili 2009: 139)

They conclude that the negative result of the experiment, which was done to find the categorical perception between *lurji* ‘blue’ and *tsisperi* ‘light blue’, may arise because it is difficult to choose two such very closely related colours from different categories that would have an identical perceptual boundary for a statistically significant number of subjects (Khomeriki, Kezeli and Lomashvili 2009: 139).

Several case studies that have found more than one basic level word for blue are published in *Värvinimede raamat* (“A book of colour terms”) edited by Mari Uusküla and Urmas Sutrop (2011). One of the cases is the Northern Udmurt language, where a Bulgarian loan *chagyr* ‘light blue’ is considered basic by Ryabina (2011: 263). Northern Udmurt has eight basic colour terms, ranked by Sutrop’s cognitive salience index[6], and these are *gord* ‘red’, *vož* ‘green’, *čuž* ‘yellow’, *lįz* ‘blue’, *śe̬d* ‘black’, *č́agįr* ‘light blue’, *purįś* ‘grey’ and *te̬dį̬* ‘white’ (Ryabina 2011):

> “The inclusion of *чагыр* [*chagyr*] ‘light blue’ amongst basic color terms is an interesting phenomen, which does not correspond to the temporal-evolutionary ordering of Berlin and Kay and definitely requires further research.” (Ryabina 2011: 267)

6 Salience in the following tables and figures refers to the cognitive salience index designed by Urmas Sutrop (2001; 2002), which combines two list task parameters naming frequency (F) and mean position (mP) and so reflect the tendency of basic colour terms to occur at the beginning of the elicited lists, and to occur in the usage of most participants. For a detailed introduction see 1.4.1 List task.

Indeed, *chagyr* 'light blue' had a high list task (N=31) percentage of 87%, a high mean position of 7.11 and the cognitive salience index of 0.122, and in the naming task *chagyr* was the dominant name for the stimulus BGB T3 (Ryabina 2011: 265). This Color-Aid stimulus, BGB T3, seems to be quite likely to become dominant for light blue if the language has a high enough naming percentage for it. In fact, Ryabina proposes that in Russian and Northern Udmurt, the terms for the lighter *goluboj* and *čagi̮r* had the same location in colour space, as they corresponded to colour sample BGB T3 (Ryabina 2011: 200; see also Rjabina 2011).

Lithuanian is another example of a language where light blue gained dominance in the naming task with Color-Aid stimulus BGB T3. Research by Pranaitytė (2011: 298) indicates that the Lithuanian *žydra* 'light blue' may be basic alongside *mėlyna* 'blue', but the language does not have a basic term designation for pink.

The most interesting examples in *The World Color Survey* (Kay *et al.* 2010) are those that provide a reasoning for why a particular term was or was not counted as a basic term. Here is the example of the Oto-Manguean language Amuzgo and how basicness was gauged in *The World Color Survey*:

> "For blue, two terms are used by all speakers: *tsa* [3] and *tsjo'* [2], the former being the most extensively and consistently used, while the latter is confined to naming dark blues with relatively low consensus. /---/ For *tsjo*' [2], the investigator indicates that it is related to the word for 'sky'. In spite of its general use, it is not altogether clear whether it can be considered a basic term, since it shows fairly low overall consensus in its term map, and appears only at 64% agreement in the aggregates, when the other blue term is fully established. On the other hand, its range is not included in that of *tsa* [3], and most individuals who use both terms seem to make a light blue/dark blue distinction." (Kay *et al.* 2010: 89–90)

There are other blue examples, e.g. Chumburu in the Niger-Congo language family, where the emerging term for blue *blu*, presumably borrowed from the English, is still competing with a black/blue term *kidʒidʒii* (Kay *et al.* 2010: 189); or the Nilo-Saharan language Murle, where a minority of speakers have developed *nyapus* '(light) blue' as a term for blue that is strongly biased towards light blues, while for those speakers *colai* is either a clear green term or a restricted *grue* (Kay *et al.* 2010: 419).

Lin *et al.* (2001), who researched Taiwanese Mandarin colour terms, use the term recessive for a basic term that has a cognitive overlay with the more dominant basic term:

> "The terms may take on slightly different meanings, because they name different points of view on the same perceptual reality, forming a cognitive overlay. In such a framework, only the dominant term would be basic, the recessive term nonbasic, and this would reduce the number of Mandarin basic colour terms from sixteen to eleven" (Lin *et al.* 2001: 47).

Their 200-sample naming task (N=60) revealed that Taiwanese Mandarin has 16 basic names, which include two reds (*hong* and *ju*), two oranges (*jyu* and *chen*), two greens (*liuh* and *ching*), two blues (*lan* and *diann*), and two browns (*zong* and *hur*), though the second term of each of the five pairs may be considered recessive, and hence not a basic name (Lin *et al.* 2001: 207).

An example of the opposite strategy of fitting the theory around the data comes from Wierzbicka (1996). Wierzbicka (1996: 313) cites the description of urban Thai given by Diller and Juntanamalaga (forthcoming [in 1996]), who propose two terms for blue, *fáa*, literally 'sky', and *nam-ŋən*, literally 'silver-tarnish', the darker blue of the Thai flag. Katemake *et al.* (2015) report five previous studies about Thai colour naming, three of them by linguists and written in Thai. A colour name identification process (N=20) by Engchuan (2003) also supported the occurrence of two basic blue terms in Modern Thai (Katemake et 2015). Katemake *et al.* (2015) report that they counted *fa* 'sky blue' among the basic colour terms because it was used to represent blue before the word *nam-ngen* 'blue' was known. The results of three experiments of colour naming show that *fa* 'sky blue' is more frequent than *nam-ngen* 'blue', which indicates that *fa* 'sky blue' might be more appropriate for consideration in the Berlin and Kay eleven basic colour terms than *nam-ngen* 'blue'. (Katemake *et al.* 2015)

For Wierzbicka, the position of the [English] focal blue, which is supposedly determined by universal human neurophysiology, in the no man's land between urban Thai *fáa* and *nam-ŋən*, "highlights the irreducible gap between neuro-physiology and meaning" (Wierzbicka 1996: 313). She speculates that sooner or later the Thais will develop a conceptual category corresponding to *blue*, but deems it most likely that they will follow a path similar to that of Russian (Wierzbicka 1996: 313).

Don Dedrick (1996) argues that the central problems of the theory arise from the difficulty of linking the linguistic, psychological, and physiological domains, which are conceptually and empirically disparate. Dedrick (1996) suggests that an account which stands between the biological and the cultural is the answer, rather than attributing colour naming either to biology or culture.

The blue category is not the only colour region where there could be a partition. For a possible partition of reds see Hungarian (see Maclaury, Almási and Kövecses 1997; Uusküla and Sutrop 2010; Benczes and Tóth-Czifra 2014). The division of *pink* is a later addition to the research (Vejdemo *et al.* 2015; Frenzel-Biamonti 2011).

This concludes the review of the literature. The next section focuses on describing the research question, followed by a description of the method.

Chapter 2

1.3. Research question

From Berlin and Kay's theory of basic colour terms and previous research into Turkish, Estonian and Russian basic colour terms arise the questions of whether the behaviour of one 'blue' is universal, and how the category of 'blue' might be divided. There being only one blue category reinforces the universalist view of colour terms, while the appearance of more than one 'blue' category, especially in the sorting task, supports a more relative approach, or a weak relativist approach.

The aim of my research is to examine three languages – Turkish, Estonian and Russian – and see how their most frequent terms for blue compare to each other on the scale of basic colour term traits, and in the colour categories of a sorting task. It is assumed that there is a certain order of strength in the basic traits of each of the blue terms. Previous research shows for example that it is reasonable to suggest that the Russian *goluboj* 'light blue' is in fact a basic term alongside *sinij* 'blue'. The Turkish *lacivert* 'dark blue' shows some characteristics of a basic colour term, but opinions differ as to whether to disqualify them to meet the non-inclusion criterion of Berlin and Kay's definition (see Özgen and Davies 1998), or to consider their dominance in the naming task as weaker than expected for a basic colour term (see also Rätsep 2011). The Turkish *lacivert* 'dark blue' shows enough basic traits to make it a strong contender for basic colour term status. The comparison of the Russian *goluboj* 'light blue' and the Turkish *lacivert* 'dark blue' with the Estonian terms *helesinine* (literally 'light blue') and *tumesinine* (literally 'dark blue') is important because the Estonian *helesinine* and *tumesinine* are not by definition basic colour terms. The parts of the terms form their meaning and they contain another, fully-established, basic colour term *sinine* 'blue' (for previous research on Estonian see Sutrop (1995; 2000; 2011).

The main research aims of the thesis are:

1) to examine the colour terms of Turkish, Estonian and Russian in general;
2) to focus on the basic colour term traits of the Russian *goluboj* 'light blue', the Turkish *lacivert* 'dark blue', and the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' in the context of other terms in the languages.

Some assumptions were made from the previous research into the three languages, and these formed the aims of the thesis and influenced the hypothesis. Sutrop proposes that Russian having two basic terms for blue, *sinij* and *goluboj*, may influence, or perhaps even destabilise, the Estonian concept of blue (Sutrop 2002: 73, 213). It is expected that the Russian *goluboj*, which is considered a basic colour term, exhibits more basic colour term traits than the Turkish *lacivert*, while the extent of the basic traits of the Estonian *helesinine* and *tumesinine* is to be determined from the analysis of the empirical data. I will draw parallels with the results for the Russian *goluboj* 'light blue' and compare them with the results from the Estonian *sinine* 'blue', *helesinine* 'light blue' and *tumesinine* 'dark blue'.

Null-hypothesis: It is hypothesised that the Turkish *lacivert* 'dark blue' shares a type-of relationship with *mavi* 'blue' and the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' are a type of *sinine* 'blue'. If the Turkish *lacivert* 'dark blue' and the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' are neither basic colour terms nor separate categories in the sorting task for Estonian terms, then quantifiable measures of their basic colour traits will not match those of previously established basic colour terms.

Hypothesis: If the Turkish *lacivert* is a separate basic colour term from *mavi* and the Estonian *helesinine* and *tumesinine* are separate from *sinine*, then their quantifiable measures should be relatively similar to those for established 'blue' basic colour terms. Therefore, the terms in question should have the same quantitative measures in all three languages. The Turkish *mavi* and *lacivert* and the Estonian *sinine*, *helesinine*, *and tumesinine* should have similar measures and be comparable to the previously established two Russian basic colour terms *sinij* and *goluboj.*

In short, there being one 'blue' category offers support to the universalist view, while the emergence of more than one 'blue' category supports the weak relativist approach to colour. The next question is how the hypothesis should be quantitatively proven or disproven using the experimental data. Quantifiable variables are mostly operationalised by frequencies. In a list task this is done by frequency, mean position, and Sutrop's cognitive salience index; in the sorting task by formation of separate stimuli groups; and in the naming task by consensus of naming by participants. The extent of basic traits of the terms is studied with the support of experimental data that give a comparison of the frequency of the Russian *goluboj*, the Turkish *lacivert* and the Estonian *helesinine* and *tumesinine* in the list task and their dominance in the naming task, and a comparison of the Estonian and Russian colour categories in the sorting task. *The World Color Survey* supplies a very useful guideline that directs much of the analysis of the study, in that the strength of the basic status of a term can only be assessed relatively and in comparison to the other terms in the same language (Kay *et al.* 2010: 21).

1.4 Methods and participants

The three studies were not conducted concurrently. The Turkish research is the earliest study. The Turkish study was conducted for my master's thesis "Turkish basic and non-basic colour terms" (2010, supervisors Otto Jastrow and Urmas Sutrop) at The Estonian Institute of Humanities at Tallinn University. I started my bachelor studies in Oriental philology, specialising in Turkology. The MA research was meant to replicate methodologically the study by Özgen and Davies with more stimuli in the dark blue area of colour space in the naming task. The methodology for the list and naming tasks comes from the practical field method for identifying basic colour terms published by Davies and Corbett (1995). Their selection of 65 stimuli was used with an additional 17 stimuli from the dark blue area of colour space.

The Estonian and Russian research was conducted later, during my doctoral studies. Estonian is my native language, and I have studied Russian, the largest minority language in Estonia. Olga Titova, a native Russian speaker, interviewed the Estonian-Russian participants. She was able to note semantic nuances and collect interesting vernacular terms, some of which I would probably have missed.

I was aware of the difference in Russian between the two blues, *goluboj* 'light blue' and *sinij* 'blue', in the theoretical sense. Since I had previously studied the possible exception of *lacivert* 'dark blue' in Turkish, I decided to expand my study to the Estonian and Russian 'blue' categories. Since completing the Turkish study I had learned new methods of using the non-verbal sorting task to ascertain colour categories. The stimuli selection was used in a previous research paper by David Bimler and Mari Uusküla (Bimler and Uusküla 2014; Bimler, Kirkland and Uusküla 2015; Uusküla 2014) for identifying the Italian 'blue' categories, so the selection of stimuli was already established.

The tasks and the selected stimuli are different in the three languages because in transitioning from my master studies to my doctoral studies I learned and implemented new methodological and analytical approaches to data. Ideally all three languages would have the same methodology, stimuli and number of participants. The participants were all volunteers and were not compensated in any way. They freely gave their time to participate in the studies. Since unpaid volunteers were used, the number of participants is directly correlated to the ability to gain new participants. In all three languages the minimum number of participants (20–25) was met.

No comparable Estonian and Russian free-sorting tasks using colour stimuli had previously been conducted in Estonia using the same method, instructions and stimuli and analysed identically, as far as I am aware.

The methods employed include a list task and a naming task for Turkish and a list task, a sorting task and a naming task for Estonian and Russian. The sorting task reflects the non-verbal colour grouping behaviour of the participants, while the list task illustrates the semantic memory, and the naming task illustrates how the colour names are used for specific stimuli. The tasks were designed to confirm the basic colour terms of Turkish, Estonian and Russian in general and to focus on the extent of basic colour term traits for the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' and the Turkish colour term *lacivert* 'dark blue'. This research compares the indices of basicness for the Russian *sinij* 'blue' and *goluboj* "light blue', the Turkish *mavi* 'blue' and *lacivert* 'dark blue', and the Estonian *sinine* 'blue', *helesinine* 'light blue' and *tumesinine* 'dark blue'. The basic colour term criteria exclude the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue'. The research shows whether any threshold of basicness is passed in the list, sorting and naming tasks. Previous studies (see Sutrop 2000;

Özgen and Davies 1998) have shown that some terms, i.e. Estonian *helesinine* 'light blue', *tumesinine* 'dark blue' and Turkish *lacivert* 'dark blue' that do not meet the basic colour terms criteria may still show some traits of basicness.

1.4.1 List task

The following section gives a short description of the list task, then introduces the most important quantitative measures for the analysable data, e.g. *frequency* (F) and *mean position* (mp), which are combined with the number of participants in Sutrop's *cognitive salience index* (referred to as Salience, see Sutrop 2001). These are the comparable parameters that form the basis of the list task analysis as they are featured in all three list task tables and figures.

The list task, also known as an elicitation task, can be used in all fields of the lexicon. The list task comes in several different varieties, see Cooke (1994). For an introduction to the list task and frequencies, see Archer (2009). The only criterion is the availability of an umbrella term or of a hyperonym, and in this case the availability of the hyperonym 'colour' makes it possible to instruct the participant to list all the colours they know.

The list task shows the importance of words in the semantic memory. The more general, basic and commonly used words are found at the top of the list and their importance is reflected in higher overall frequencies and lower mean positions than are found for more specific and less frequently used words. To get a picture of the domain of colour words, the list task asked participants to list all the colours they knew.

The research analyses three parameters of the list task: frequency, mean position and salience**.** Frequency (F) is the number of times the terms were used by all the participants. Here it is presented either in absolute numbers or in percentages (%) for comparison. The second parameter is mean position (mp), which shows the average position of the term and whether it was elicited in the beginning, middle or end of the lists by the language guides. Generally, the mean position and the frequency are inversely related, so the lower the mean position, the higher the frequency.

To help in visualising the list task data, a graph was created following Edward Tufte's (2001: 91–105) principles for the data-ink ratio, which advocates clear, simple graphs in which ink is used to represent new data. While the graph is simple, its shape illustrates the nature of the data (see Figure 3, Figure 5, and Figure 12).

Recognising that frequency and average rank are both reflections of the same underlying property of salience, some researchers like to combine the two into a single measure. Sutrop's (2001: 263) *cognitive salience index* combines the two list task parameters of frequency and mean position independently of how long the list in question is. The ideal most salient term has a cognitive salience index of 1 and a term that is not mentioned at all has an index of 0. Its calculation formula is: S = F / (N x mp), where S is the cognitive salience index; F is the frequency of use in the list task; N is the number of participants; and mp is the mean position. Sutrop's cognitive salience index has a tendency to overemphasise the one or two most salient items (Thompson and Juan 2006: 400):

> "/…/ Sutrop's salience index statistic tends to produce distributions in which one or two most salient items have much higher values than other items, and thus, appear as outliers in the domain."

Bimler and Uusküla (2014: 182) show that using the logarithmic function smooths the cognitive salience index by transforming an exponential decline into a roughly linear form.

Although Sutrop's index is not affected by the length of the list, like the earlier Smith index was, some authors prefer Smith's index, but unfortunately calculating it requires the full listing data, and it cannot be used with published data where the authors have only published the frequency of each term and its mean position.

An unrestrained list task was conducted with Estonian (N=39), Russian (N=30) and Turkish (N=56) participants. The interviews were conducted by a native or fluent speaker of the language and participants were asked to list all the colours they knew. The participants were instructed in Turkish to: "Bütün bildiginiz renkleri söyleyebilirmisiniz lütfen." In Estonian, *Palun loetlege kõik värvid, mida te teate.* In Russian, *Požaluysta skažite vse tsveta, kotoryye vy znayete.*

The list task was followed by a naming task for the Turkish speakers, as the method developed by Davies and Corbett (1995) requires. For the Estonian and Russian speakers, the sequence of tasks was altered to avoid bias in the sorting task, and so it ran list task, sorting task and naming task.

1.4.2 Naming task

Some form of naming task has been a part of the basic colour term paradigm since the very beginning. *The World Color Survey*, for example, used a classic combination of list and naming tasks. The stimuli set in the task for Estonians and Russians was different from the Turkish set of stimuli. In the naming task the stimuli were shown and named separately, not all at once as in the preceding free-sorting task for Estonian and Russian.

Unlike the list task, which requires there to be an umbrella term, the naming task does not need a hyperonym, and so it can be used to analyse domains that do not have a domain name. The naming task uses stimuli to generate answers and using colour stimuli in paper form makes the task reproducible and comparable.

For stimuli tasks the name most frequently given to a stimulus (*nmf*) is a simple indicator of consensus, which while crude, is useful when used comparatively. A term can be the most frequent term for a stimulus even if it is used by a small percentage of users, and a stimulus can have more than one nmf term if the terms have the same frequency (see Al-Rasheed 2014: 1719). The nmf results vary considerably depending on the stimuli, methodology and language. The range of stimuli per term is dependent on the number and selection of stimuli, but it helps to pinpoint the borders and position of the term within the stimuli presented. It is very difficult for example to find a colour stimulus that participants agree to name *beige* (see Eessalu and Uusküla 2013). The range of stimuli can also help to ascertain if the stimuli selection is indeed evenly spread on the colour spectrum, or if one area is over-represented. In the Estonian and Russian naming task the range of stimuli reveals the number of stimuli corresponding to 'light blue' to be particularly over-represented, but it also shows how large of an area of the colour spectrum 'light blue' can represent. Basic colour terms should have a larger range of stimuli than non-basic colour terms. The latter have a smaller range, often consisting only of a couple of stimuli, because they are more specific by nature.

The criteria for judging whether a term is basic or not (Kay *et al.* 2010: 21) can be applied to the naming task: 1) the highest consensus for a stimulus, or nmf – named most frequent; 2) the range of stimuli per term; and probably 3) the nmf plot and the use of a multidimensional scaling (MDS) plot.

I prefer to present the analysis visually in plot form (see Figures 3, 5 and 12), but a classic parameter often included in the tables, and also included here, is the *specificity index* (see Tables 2, 6 and 10). The specificity index is a ratio of the total frequency of use for each term and the total frequency for those stimuli where a term was dominant (Davies and Corbett 1995: 79), found from dominance frequency divided by the total frequency. The specificity index (SI) is also an indicator of consensual naming for each stimulus. As with the consensus, which is the percentage of consensual terms given to stimuli, the higher the specificity index, the higher the naming consensus. Generally, if there is a consensus of 50% or more between language guides on naming a stimulus, then it is most probably a basic colour term. If the percentage of those either listing the term or naming the stimulus with that term is less than 50, then the basic status of the colour word is questionable, though the *dominance threshold* of a consensus among participants of over 50% in naming a stimulus is dependent on the overall level of consensus in the data. The parameters for the naming and sorting tasks overlap. For the tasks with stimuli, which are the naming and sorting tasks, graphical representations of the analysis are included unless the colour coordinates were unavailable, as was the case with the Turkish additional stimuli.

The first task for all three languages was the list task. In the naming task the participants were shown stimuli in random order and asked what colour was on the stimulus shown. In Turkish, *Bu ne renk?* in Estonian, *Mis värvi see on?* and in Russian, *Eto kakoj svet?*.

Which task followed depended on the languages. For Turkish the naming task followed the list task without the sorting task, while Estonian and Russian had three tasks, list, sorting and naming.

1.4.3 Sorting task

Sorting tasks have rarely been used in the previous research into Turkish, Estonian[7] and Russian colour terms. It prolongs the interview time and requires more specialised data analysis than is classically needed for the naming task. It should, however, reflect the non-verbal grouping behaviour associated with the terms, or categories in this case, rather than just naming them.

The following section describes the sorting task, also termed a *pile sort* task by Weller and Romney (1988: 20–25). For an introduction to the sorting task see *Systematic data collection* by Weller and Romney (1988).

In a single sort task participants are mostly asked to sort visual stimuli, e.g. cards with the name of an item, pictures, or drawings, into piles so that similar items are in piles together (Weller and Romney 1988: 20–21). The participant decides what criteria are the most salient and determines similarity. Participants can make as few or as many groups as they wish out of the stimuli, or they may be asked to make a specific number of piles. Pile sort tasks are sometimes constrained to control for individual differences in creating larger, more generic categories, where those who use this strategy are called "lumpers", or smaller categories with finer distinctions, as preferred by "splitters" (Weller and Romney 1988: 22). The study by Rosenberg and Park Kim (1975) indicates that by its nature, the pile sort task cannot accommodate many dimensions or much discrimination at once. Splitters may try to use multiple dimensions at once, thus resulting in a large number of small piles. (Weller and Romney 1988: 24)

Weller and Romney (1988: 25) note several strengths of the sorting task method. It is easy to administer and only one of each stimulus is need in a field setting. Participants enjoy sorting things into piles and then talking about them. The authors suggest it is an appropriate method for studying relations between items, and the number of items can be large at over 100. Weller and Romney (1988: 25) refer to other authors to show how reliable the sorting method is. Burton (1975) demonstrates that "the unconstrained, unnormalised pile sort data perform quite well" and the results of medium-size samples of 30–40 are generally reliable and demonstrate a high degree of stability across replications (Romney *et al.* 1979).

The weaknesses of the sorting task (Weller and Romney 1988: 25) can be subverted by circumstances, e.g. stimuli are not words, so it can be used for non-literates; the number of participants is over twenty, so the results obtained are stable.

The sorting task looks for differences between subjects in their sorting. Romney and Weller (1988: 25–26) cite earlier research by Boorman and Arabie (1972; Arabie and Boorman 1973), and Boorman and Olivier (1973), who found that the difference between lumpers and splitters is so great that it

---

[7] Triin Kalda (2014) conducted interviews with Estonian, Finnish and Finnish Estonian participants for a comparative analysis of their colour lexicon, including the sorting task.

overwhelms all other differences among pile sorts. However, with a constrained sorting task, comparisons of individuals can be made (see Truex 1977). Rosenberg and Park Kim (1975) compared the dimensions subjects used in "single sort" and "multiple sort" tasks and the results indicate that when participants were allowed a single sort, they usually ignored the most obvious dimension or meaning (Weller and Romney 1988: 24).

The data are usually analysed using a similarity matrix constructed for each participant and then the matrices of all the participants are combined. Romney and Weller explain the construction of the matrix:

> For example, if we collected data on the similarity of seven items and a respondent put items A, B, and C together in a pile; D and E in a pile; and left F and G by themselves /---/ we would create a 7 by 7 table to tabulate similarity among the items. Since A, B, and C are categorised together, A and B are similar, B and C are similar, and A and C are similar. Since D and E are also put together in a pile, D and E are considered to be similar. Thus each pair would get "a point of similarity." This is indicated in the table with a one. For this individual, all other pairs are "dissimilar" and are recorded as zeros. Similarity matrices are tabulated for each individual and then combined across people. The similarity matrix can then be analyzed with a descriptive method such as hierarchical clustering or multidimensional scaling. (Weller and Romney 1988: 22)

Pile sort data are coded as 0's and 1's so that items that are grouped together in a pile are coded as 1's and items not grouped together are 0's (Weller and Romney 1988: 25).

The task has similarly been used to help ascertain basic colour terms, or rather categories, in previous research, see Jraissati *et al.* (2008), Roberson *et al.* (2005), and Frumkina (1984). Most notably, the same sorting task method and stimuli were previously used by Bimler and Uusküla (2014), and Uusküla (2014), who found there was no methodological flaw in the selection of mostly blue-purple stimuli with several yellowish distractor stimuli, since in Italian more than one blue category emerged in the sorting task, although the consensus was also very low.

The sorting procedure has been applied to the analogous question of the Russian blue terms (Laws, Davies and Andrews 1995) as one way of eliciting judgements of colour dissimilarity. More generally, Roberson, Davies, Corbett and Vandervyver (2005) analysed free-sorting and naming of colours from a cross-cultural perspective, testing speakers of 17 different languages. They used 65 Color-Aid stimuli spread evenly across the colour space, varying in lightness and saturation as well as hue. For the free-sorting task the participants were instructed to group the stimuli so that 'ones that looked similar were placed together in a way that members of a family go together' (Roberson *et al.* 2005: 94). The analysis of the results indicates that there were both large cross-cultural similarities in grouping behaviour and within-language variation between participants. Roberson *et al.* (2005: 87, 119) suggest that this is best explained by a combination of a ‘universal principle of grouping by similarity’ and ‘culture-specific category salience’, which is sometimes differential.

In the sorting task Estonian and Russian participants were presented with all the stimuli and asked to sort them by similarity, so that similar stimuli were grouped together. In Estonian they were asked, *Palun jagage need sarnasuse alusel gruppidesse*, and in Russian, *Požaluysta razdelite ikh na gruppy osnovannyye na skhodstve.*

The stimuli were presented all at once on a flat surface in random order in natural daylight with no direct sunlight or overcast sky. The participants were told that there were no restrictions, so they could sort the stimuli into as many groups as they liked and put as many or as few tiles into each group as they desired. After they had completed the sorting, the participant was asked to name the groups. The group names were recorded exactly as spoken by the participant. After the free-sorting task, the participants were asked to name the stimuli one by one for the naming task, for which the stimuli were presented separately on a neutral grey background. None of the tasks had a time limit.

To give a pictorial representation of the sorting task *multidimensional scaling analysis* (MDS) was used, which illustrates how the Estonian and Russian language guides grouped the stimuli. This allows the grouping behaviour of the two sets of participants to be compared. I feel it was necessary to include this brief introduction of the method since it is decisive in explaining the results of the sorting task, and therefore the use of colour categories.

The multidimensional scaling analysis and various other tasks that require statistical software was done in an R environment within which statistical techniques were used. R is an open source programming language (a GNU project) and software environment for statistical computing and graphics. One of its main strengths is the ease with which well-designed publication-quality plots can be produced. Working efficiently with R requires an appropriate editor, and RStudio was recommended by Borg, Groenen, and Mair (2013: 63) amongst others, and so it was chosen and used. RStudio is an integrated development environment (IDE) for R. R can be extended via packages, and the most heavily-used package for multidimensional scaling analysis was *smacof.* The *ggplot2* package, an enhanced data visualisation package for R, was employed for the graphics (R Core Team 2012; RStudio Team 2016; Leeuw and Mair 2009; Groenen and van de Velden 2016; Wickham 2009). A key feature of R is that outputs of analyses are stored as R objects like lists or matrices that can then be used for further processing (Borg, Groenen and Mair 2013: 21). The software is very useful, as instead of copying the multi-dimensional scaling plot coordinates to and from a ‘.txt’ file for example, which can lead to formatting errors, the user can simply use the stored data without introducing human error along the way.

Multidimensional scaling (MDS) is very much a visualisation technique (Borg and Groenen 2005: 543), and it originated as a psychological model on how people form judgments about similarity or preferential choice (Borg, Groenen and Mair 2013: vi). The purpose of multidimensional scaling analysis is to visualise proximity data optimally (Borg, Groenen, and Mair 2013: 79). Multidimensional scaling is based on a proximity matrix, which is usually derived from variables measured on objects as input entities (Groenen and van de Velden 2016). For multidimensional scaling visualisation the starting matrix is that of dissimilarities, rather than similarities, where large dissimilarities are represented by large distances and small dissimilarities by small distances (Groenen and van de Velden 2016: 2). When interpreting a multidimensional scaling solution it is assumed that the closer two datapoints are, the more similar the objects they represent are ('Metric and Nonmetric MDS' 2005: 203), so the closer two points are in the multidimensional scaling plane, the higher the correlation of the variables they represent is (Borg, Groenen and Mair 2013: 2). Multidimensional scaling always searches for coordinate values of n points in m dimensions, whose distances represent the given proximities as precisely as possible, or optimally (Borg, Groenen, and Mair 2013: 21). The dissimilarities are mapped on a low-dimensional spatial representation (Groenen and van de Velden 2016).

The basic implementation of multidimensional scaling within the smacof package is symmetric, with options for ratio, interval, ordinal, and spline transformations of the proximities (Groenen and van de Velden 2016: 1). Multidimensional scaling is either metric, such as airline distances between cities in miles as a symmetric input matrix, or nonmetric. Metric models represent various properties of the data related to the algebraic operations of addition, subtraction, multiplication and division, while in *nonmetric*, if $p12 = 5$ and $p34 = 2$ for example, an ordinal model reads this only as $p12 > p34$ (assuming here that the data are dissimilarities) and constructs the distances $d12$ and $d34$ so that $d12 > d34$. ('Metric and Nonmetric MDS' 2005: 203)

Sorting task data is nonmetric, which translates into the ordinal ratio option in multidimensional scaling. Borg, Groenen, and Mair (2013: 83) caution against assuming that dissimilarity data are Euclidean distances. They state that the only instance where it may be possible that the numerical responses observed are at least distance-like values occurs when participants are asked directly to rate pairwise dissimilarities (Borg, Groenen, and Mair 2013: 83). In the free-sorting task of this thesis the participants were not asked to rate pairwise dissimilarities. Borg, Groenen, and Mair (2013: 10) suggest that multidimensional scaling is useful for uncovering latent dimensions of judgement that can help in inferring the attributes that respondents assign to the subject in question. The authors then observe the problem with distance models, where objects are first conceived "as points in a psychological space that is spanned by the subjective attributes of the objects", as it is hardly ever known what attributes a participant assigns to the objects under consideration (Borg, Groenen, and Mair 2013: 10). Multidimensional scaling analysis represents proximities, as precisely as possible, as distances (Borg, Groenen, and Mair 2013: 60). Distances always exist in the pictorial representation of multidimensional scaling analysis (Borg, Groenen, and Mair 2013: 60), but it is inadvisable to interpret data directly as distances, and rather an optimal re-scaling should be allowed for (Borg, Groenen, and Mair 2013: 83).

When the data are correlatively represented then the first two steps of the classical multidimensional scaling, also known as Torgerson scaling or Torgerson–Gower scaling, can be skipped. In this case it is not necessary to square the dissimilarity data $\Delta^{(2)}$, nor to convert the squared dissimilarities to scalar products through double centring. This means that one point must be picked as an origin so that scalar products can be computed before principal component analysis proceeds, or alternatively the scalar products can be converted to distances. Principal component analysis begins with a data matrix of *n* cases, say, participants, and *k* variables such as stimuli. The objective of the method is to reduce the *k* variables to a much smaller set of *m* new variables, which are linear combinations of the original variables, but it is presumed that the new variables are sufficient to explain most of the variance in the data (Borg and Groenen 2005: 519).

The sorting task is a co-occurrence data matrix, which is typically aggregated over subjects. The result of the sorting task can be expressed for each subject by a 55 × 55 incidence matrix, or a 51 × 51 one if yellowish stimuli are excluded, with the entry 1 wherever its row and column entries are sorted into the same group, and 0 elsewhere. The pairs of objects that are in the same group have a dissimilarity of 0, which was originally a similarity of 1 in the co-occurrence matrix, and those in different groups have a dissimilarity of 1 (Borg and Groenen 2005: 114). Co-occurrence data incidence matrices are typically aggregated over individuals so that the aggregate proximity matrix contains in its cells the frequencies with which two objects were sorted into the same group (Borg and Groenen 2005: 127).

In this case a correlation matrix is also a similarity matrix, where the highest possible score equates to the number of subjects, which is 39 for Estonian and 30 for Russian. The maximum similarity score between two stimuli means that all the subjects grouped those two stimuli together, while a 0 in the similarity matrix indicates that the two stimuli were never sorted into the same group. Analysing the data via MDS means that instead of 30 or so different numerical indexes, or correlations, there is a simple *visual* representation of the empirical interrelations (Borg, Groenen, and Mair 2013: 3). The incidence matrix can be considered a proximity matrix of dichotomous (same–different) data (Borg and Groenen 2005: 126).

The *mds* function is the simplest MDS-smacof version of the package. It solves the stress target function for symmetric dissimilarities by taking the majorisation approach in *smacof*, and reports the Stress-1 value (squared) (Borg, Groenen, and Mair 2013: 104). The formal goodness of a multidimensional scaling solution can be measured by computing the solution's Stress value, which is zero when the solution is perfect. The minimum requirement is that the model's Stress value be less than the Stress expected for random data, but Borg, Groenen, and Mair advise that when the Stress value of a particular multi-dimensional scaling solution is evaluated, it should be assessed in the context of various parameters and contingencies, such as the number of data points, the dimensionality of the multidimensional scaling solution space, and the reliability of the data. (Borg, Groenen, and Mair 2013: 26)

To run the analysis of multidimensional scaling (MDS), the *smacof* package was used with the *RStudio* interface in the *R* programming environment. Smacof is an acronym for Scaling by MAjorising a COmplicated Function (De Leeuw and Heiser 1980). The optimisation method used by smacof is called Majorisation. The basic idea of this method is that a complicated goal function is approximated in each iteration by a less complicated function which is easier to optimise (Borg, Groenen, and Mair 2013: 85). For more details on how this method is used to solve multidimensional scaling problems, see De Leeuw and Mair (2009) or Borg and Groenen (2005). However, with well-structured data, different multidimensional scaling models yield solutions that do not differ much (Borg, Groenen, and Mair 2013: 77).

The default in most programs uses what is called the *primary approach to ties*, where ties, or equal data values, can be broken, meaning that equal proximities need not be mapped into equal distances (Borg, Groenen, and Mair 2013: 38). They (2013: 38) claim that the primary approach to ties is usually more meaningful in terms of the data, but it results in some pairs having the same proximity values irrespective of whether a participant truly conceives that the given stimuli are equally similar. The authors remark that no respondent can really make reliable distinctions, for example on a 55-stimuli scale, and therefore equal ratings should not be interpreted too closely (Borg, Groenen, and Mair 2013: 38).

According to Borg, Groenen, and Mair (2013: 74–75), a frequent problem in multidimensional scaling applications is the question of what to do with points that do not fit into an interpretation. If a plot cannot be partitioned in a theoretically pleasing way with simple, although sometimes very curvy, partitioning lines, then a solution may be accepted where some data points reside in a partition in which classification would not put them. Theoretically displeasing data points, which eschew the presumed partition, or theoretical outliers are often ignored (Borg, Groenen, and Mair 2013: 74). The authors purpose that the most common solution for disturbing points, is to deal with them poorly by eliminating them and explaining them away in substantive terms (Borg, Groenen, and Mair 2013: 74). They are fiercely opposed to this method of eliminating items that do not fit into a unidimensional structure (Borg, Groenen, and Mair 2013: 74, 105). In the multidimensional scaling analysis of the Estonian and Russian sorting and naming tasks, four distractor stimuli are not included because they were meant to be just that, distractors, and they are not part of the blue-purple colour continuum.

1.4.4 Stimuli

Using colour stimuli in paper form makes the task reproducible and comparable, but there are several drawbacks to the approach. The stimuli are context-free, so in some cases the shape or form of the stimulus may restrict the possible answers. Hollman (2010: 93) gives an example from Estonian sign language, where the sign for RED 1 would most commonly be used for an abstract colour stimulus, while the sign RED 2 would be most likely to be used for fluids, so the use of abstract stimuli can be constrictive. The shape of the stimulus can also be an obstacle if the colour name is not suited for a square-shaped stimulus, e.g. cattle terms can be quite practical and specific for example, making them harder to use for an abstract stimulus. The opposite seems to be the case for beige, which is rather abstract by itself, and it is hard to pinpoint a specific colour stimulus that is applicable to beige.

The Color-Aid stimuli was used for all the tasks, but the selections of stimuli differed. The Color-Aid corporation website has an introductory booklet ('The New Color-Aid Booklet' 2006), which describes the Color-Aid Full Set of 314 colours and can be summarised here.

The selection of stimuli in the Turkish study is from the set of 220 coloured papers made by the Color-Aid Corporation. Altogether there are 82 stimuli, with 65 standardised tiles for ascertaining basic colour vocabulary in the field (Davies *et al.* 1992) and an additional 17 for the specific purpose of studying the blue region. The basis of the stimuli is the Ostwald colour system, where the main features of colour are the hue, which is the colour tone; its tint, or the content of white; and shade, or the content of black. The Colour Aid 220 set has six basic colours designated by their first letters, which are yellow, orange, red, violet, blue and green and 18 transitional tones, e.g. BGB – blue-green-blue. Every chromatic colour hue has four different levels of white content (T1-T4) and three different levels of black content (S1-S3). There is also a scale for achromatic colours from white to grey to black. The Color-Aid system itself contains several additional colours designated with colour names, e.g. Sienna.

The Estonian and Russian stimuli were from the newer selection of 314 coloured papers in the Full Set of the Color-Aid Corporation. Most notably they differentiate between cold tones, (designated c) and warm tones (w). The 314 Full Set contains 295 chromatic colours and 19 achromatic colours. The hue symbols are similar, so R is red and RV is red-violet. The category symbols include full colour (HUE), extra hue (EX), shade (S), dark shade (DS), tint (T), light tint (LT), pastel (P), and light pastel (LP). So BG-P2-3 = blue-green pastel two, three; Bw-T4 = warm blue, tint four. For a visual representation of the 65 selected stimuli in the Turkish naming task see Figure 1, which does not include the 17 additional stimuli from blue section, because I have not found their recorded coordinates.

From the practical perspective of the researcher, the stimuli are convenient for fieldwork since they are light, small, easily packable, and reusable for a significant number of participants with a modicum of care. However, both selections of stimuli suffer the same drawbacks (see 2.5 Comments).

The Turkish participants (N=56) took part in a naming task where a "standard selection" of 65 coloured tiles (see Figure 1) suggested by Davies and Corbett (1995) from the Color-Aid Corporation 220 Standard Set was used. The 65 tiles were originally chosen by Davies and Corbett because they "formed a coarse, but evenly spread sample of colour space" (1995: 27). This constriction was used for the sake of expedience and so that large numbers of participants could be tested in everyday situations on the street, at home or at work. The tiles consisted of the Color-Aid coloured paper glued to a 5 × 5 × 0.2 cm piece of cardboard.

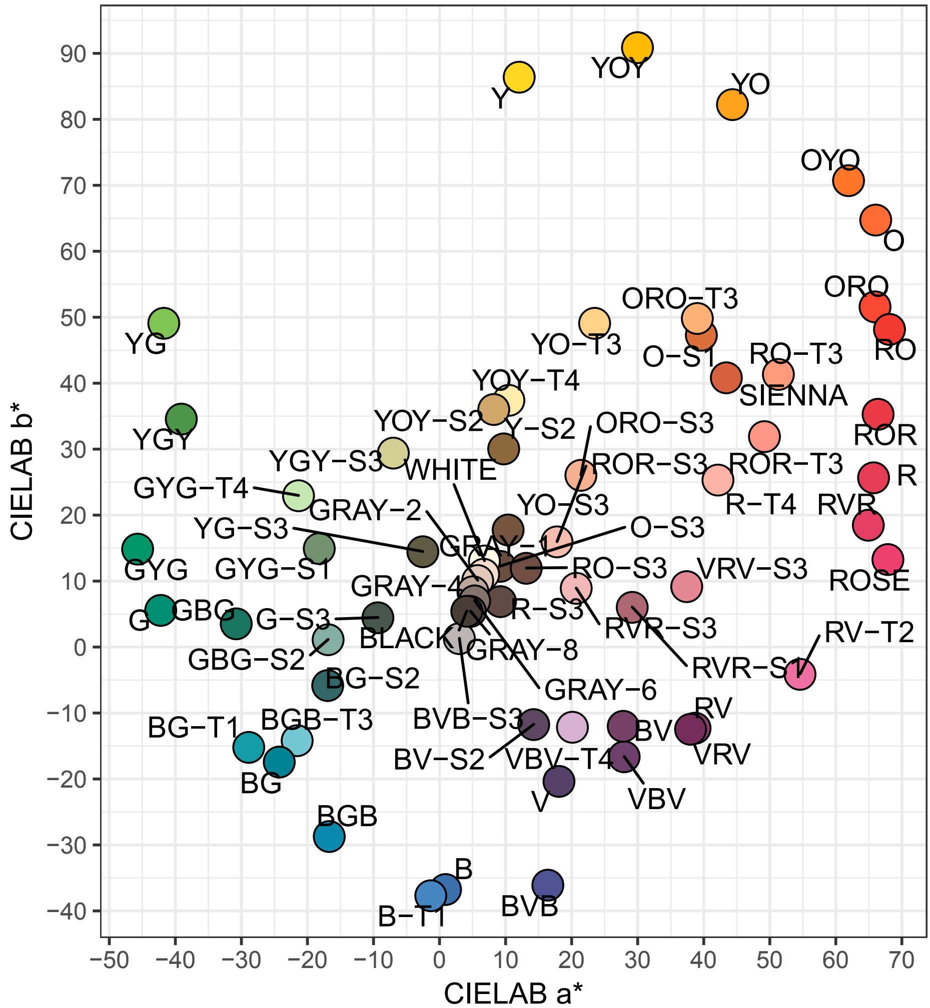

Figure 1. 65 standard stimuli in CIELAB a*– b* plane.

The L* axis – separating lighter from darker stimuli – should be imagined as perpendicular to the page.

Figure 1 depicts 65 standard stimuli with C.I.E. xyY coordinates from the stimuli selected by Davies and Corbett, but the figure does not feature all the stimuli used, with the 17 additional ones not depicted in Figure 1 as I was unable to source all the coordinates needed. The area of high concentration of stimuli in the middle of Figure 1 contains eleven stimuli, which are mostly monochromatic stimuli of white, grey and black. Moving upwards from the left of Figure 1 these are BVB S3, BLACK, GREY 8, 6, 4, 2, 1, WHITE, O S3, R S3, and RO S3.

To ascertain the position of *lacivert* 'dark blue' in the hierarchy of Turkish colour terms, 17 additional tiles were selected from the purple-blue region of colour space. The 17 extra tiles used in the naming task were: BV T1, BV T2, BV S1, BVB T1, BVB T2, BVB T3, BVB S1, B T2, B T3, B T4, B S1, B S2, B S3, BG T2, COBALT BLUE, NAVY BLUE, and CYAN BLUE. The additional tiles selected for the naming task covered the whole blue range of the Color-Aid tiles and most of the purple-blue region, with three supplementary tiles completing the selection. Participants were shown 82 tiles, 65 of them standard and 17 additional tiles, randomly one after another one on a neutral grey background in natural daylight.

In the Estonian and Russian sorting and naming tasks 55 colour stimuli (see Figure 2) were used, again from the Color-Aid Corporation and concentrating on the blue-green-purple neighbourhood of colour space (for the use of the same stiimuli see Bimler and Uusküla (2014); Uusküla (2014). Bimler and Uusküla (2016) have used the author's Estonian data in their article along with data on five other languages, Russian, Italian, English, Lithuanian and Udmurt.

The stimuli were matt-finished coloured papers from the Color-Aid Full Set. The complete range contains 314 colours, which would have been impractical for fieldwork conditions. Given this constraint, and following Laws *et al.* (1995), 55 papers were selected (for stimuli selectiona and coordinates see Bimler and Uusküla 2014: 335), with 51 centred on the entire blue range of the colour space extending into green and purple regions, plus four stimuli from the yellow region to act as distractors and as a starting point during sorting.

These selected papers were mounted on 5 × 5 cm square tiles. Most relevantly for the analysis, since the Russian *goluboj* 'light blue' is considered to be a light and cool tone, the selected stimuli contained both warm-tinged (Color-Aid code Bw) and cool-tinged (Bc) blue tones, and these are distinguished in the Color-Aid set by the letters 'w', and 'c'.

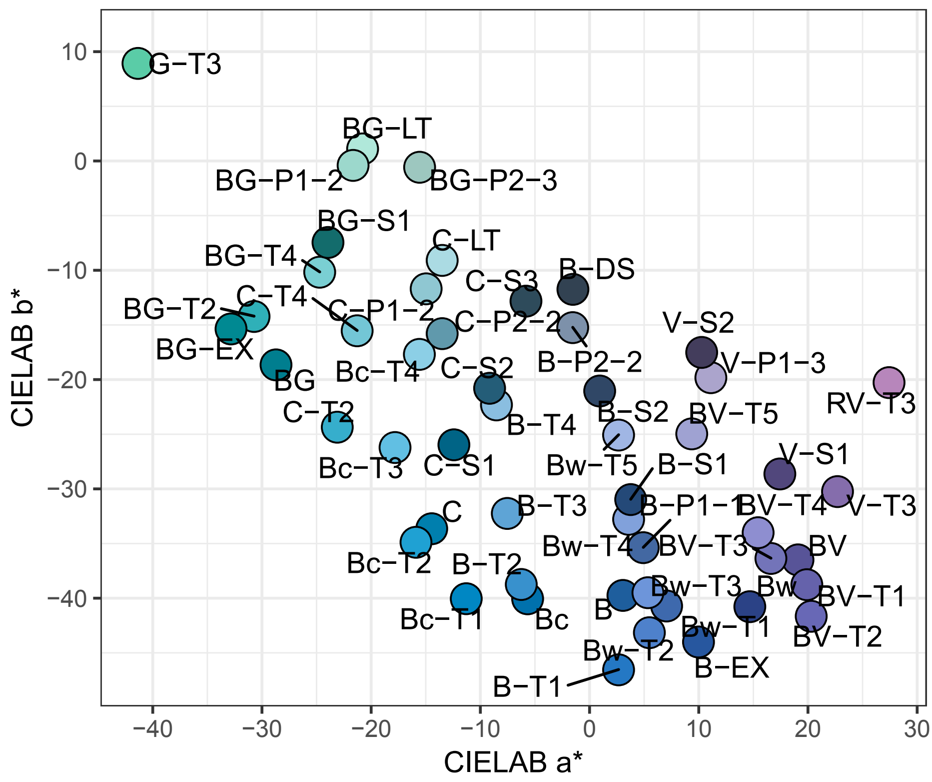

Figure 2. 51 non-yellow stimuli. Coordinates from Bimler and Uusküla (2014).

The L* axis – separating lighter from darker stimuli – should be imagined as perpendicular to the page.

1.4.5 Participants

All the participants were recruited randomly on a voluntary basis. They received no reward for their participation. Altogether, more than 130 participants were interviewed. I tested all the Turkish and Estonian participants, and the Russian participants were tested by a Russian native speaker. The length of time needed to complete the task varied greatly between participants.

The Turkish data were collected earlier than the Estonian and Russian data. There was a sorting task in the Estonian and Russian data, while in the Turkish data the participants completed the list task and the naming task, which are sufficient for assessing the number of probable basic colour terms. In the naming task, the selection of stimuli was even expanded to include additional blue tiles. The sorting task is sometimes used in addition to the list and naming tasks, but usually only the list task and naming task are completed.

Turkish participants (N=56). The participants were all native speakers of Turkish and the interviews with them were done by a fluent or native speaker.

The terms given were written down by a native or fluent Turkish speaker in the way they were said by the native participants. The list and naming tasks using the recognised field method of Davies and Corbett were conducted in Ankara and Antalya on March 17–23 and July 12–26, 2007. An intrinsic part of the field method is the subject's colour vision, which was tested with the *City University Colour Vision Test* (Fletcher 1998), which let the interviewer determine whether or not the subject had normal colour vision. For a newer web-based colour vision test, see the City University London Dynamic Colour Vision Test ('City University Dynamic Colour Vision Test' 2015; for a detailed explanation see Barbur, Harlow, and Plant 1994). For the results of the Turkish list task with data from all 60 participants see Rätsep (2011). In fact, four participants did not pass the test and their answers were not included in the colour naming part of the data. The results of 56 subjects were used.

There were 30 females with a mean age of 28.7 and 26 males with a mean age of 35.6 who completed the oral list task and continued on to the stimuli naming task. The youngest subject was a 14-year-old schoolgirl and the oldest a 79-year-old former schoolteacher. Young adults aged 20–36 formed the largest age group at 68%, of whom half, or 34% of the total, were university students in the 20–23 age bracket, and the other half were aged 25–36. Participants aged over 40 were 27% of the total, while 5% were younger than 20. The age groups represented least were the elderly and teenagers. It should be noted that a large percentage of the participants were full-time university students in the middle of their studies.

Estonian participants (N=39). Forty-one participants participated in the list task but the data from two participants were omitted from the analysis. The first of these participants did not pass the City University Colour Vision Test, and the second completed only the list task. The participants were interviewed between August and November 2010.

The data are not a balanced representation of population, because 30 subjects were female and only 9 were male. The two omitted participants were male. The average age of the male participants was 37 years, while the average age of the female participants was 46. The youngest male respondent was 21 years old and the oldest 57, while the youngest female participant was 17 and the oldest 86.

Most of the participants were born in and lived in the capital, Tallinn. Although the majority of participants had completed higher education, 13 participants had only finished high school and one subject had only basic education. The most common foreign language they spoke was English, followed Russian, Finnish and German, and 13 participants listed Russian as their first foreign language, 10 participants gave it as their second, nine participants as their third and three participants as their fourth foreign language.

Russian participants (N=30). It must be noted that all the Russian participants were Estonian Russians, emigres whose language may differ from that of Russians living in Russia. The participants were interviewed between August 2011 and February 2012. Thirty participants participated in the tasks, 14 of them male and 16 female. The male participants were aged 18 to 62 and their mean age was 44.0 years, while the females were aged 15 to 75 with a mean of 47.4 years. The majority of the participants, 19 of them, had completed higher education, seven had followed secondary vocational education, three had secondary education, and one subject was still in school in the 9th year. There were eight monolingual Russian speakers and the other participants claimed knowledge of at least one other language. The languages most commonly spoken by the participants in their self-assessment were Russian, Estonian and English, claimed by seven participants, while three participants listed Russian and English, and three participants listed Russian and Estonian. Thirteen participants did not list Estonian. Most of the Estonian-born Russian participants, 13 of them, were born in the Estonian capital Tallinn, and the two others were born in Võru and Narva. Many participants were born in the Russian Federation, two of them in Leningrad and others in Moscow, Perm, Bryansk Oblast, Kirov Oblast, Sverdlovsk Oblast, Lomonossov, Smolensk, Chelyabinsk and Vladivostok. Some participants were born neither in Estonia nor Russia, but in Kustanai in Kazakhstan, Minsk in Belarus, Riga in Latvia, and Mohyliv-Podilskyi in Ukraine.

1.4.6 Data preparation

The data were entered as they were said by the participants, which lowered the consensus considerably, so for example the Turkish *camgöbegi* 'pale bluish green' is formed from *cam* 'glass, bottle' and *göbek*, *göbegi* 'bottom; belly', and was listed in five different forms as *cam göbegi*, *camgöbegi mavi*, *camgöbegi mavisi* 'blue' and *camgöbegi yesil* 'green' (F=1), and *camgöbegi* (F=2). This was true for all languages and tasks. One rather obvious consequence of this is that it produces a considerable variability in terms, but it also points to the need not only to state clearly which original data were omitted, but also, if possible, the percentages of change that doing so produced.

The list task data were examined for double entries, and the naming task was examined for non-applicable answers such as 'I do not know'. Data were omitted from the list task if a participant gave them twice, like Turkish participant no 34, who listed *mor* 'purple' twice, in fourth place and fourteenth place. The second listing of purple was deleted, since the presumption is that every participant lists a term once. In the naming task there are a possible of $56 \times 82 = 4592$ answers, but on seven occasions the answer was a non-applicable *I do not know*. Stimuli Y –S2, RO-T3, and BG-T1 were each not named by one participant, while one participant did not name four stimuli, YO-T3, RVR-S1, RV, and GBG-S2, so the Turkish naming task had seven *I do not know* answers.

II. RESULTS

The results are a complement to previous colour research. The Turkish data dovetail nicely on the results of Özgen and Davies (1998), adding to the discussion of the position of *lacivert* 'dark blue' in the Turkish colour terms. Although the stimuli selection contains additional tiles, the standard selection of 65 stimuli is comparable with many other datasets from other languages. It not only illustrates the naming of the same stimuli in different languages, but also highlights the nuances of each particular language, within the same dataset and across datasets.

The Russian dataset has the same data collection method and is therefore comparable to the Estonian data, but as intended it presents a different case due to the basic status of *goluboj* 'light blue' in Russian. The three datasets aim to present a case where the different use of the blue terms is displayed, with the Russian *goluboj* firmly established as a basic colour term, while the Turkish *lacivert* can be considered a borderline basic colour term. In Estonian, however, the modified blue terms *helesinine* 'light blue' and *tumesinine* 'dark blue' do not fit the established criteria of a basic colour term, but in some instances they do meet the thresholds expected of a basic colour term.

The results are presented by language for Turkish, Estonian and Russian.

Table 1 presents the data in numerical form. The list task was the first task and no stimuli were used in it, but the participants were asked to list all the colours they knew. In the stimuli tasks the numbers are more dependent on the number of language guides, the colour range of the stimuli and the number of them. In the sorting task, which was completed by the Estonian and Russian language guides, the stimuli were predominantly bluish-purple since they were used to target the blue region of colour space. In the naming task for Estonian and Russian the same stimuli were used as in the preceding sorting task. In Turkish the list task was followed by the naming task. The selection of stimuli for the naming task contained more different colours than were used in the Estonian and Russian naming tasks. The overall numbers found from the naming tasks vary accordingly, but the number of different terms is fairly similar in all languages.

Plural forms were also used to name the stimuli in the Estonian and Russian sorting tasks. In Estonian there were 36 plural forms, in Russian 15. The concatenated forms are marked in brackets in Table 1.

The overall number of list task answers is stable across languages and the number of participants. A frequent term is listed by many participants, but counts as one for the number of different terms. The number of different terms is lowest for Turkish at 163, and about equal for Estonian at 336 and Russian at 294. The smaller number of unique terms for Turkish is reflected in the higher average for each listed term of 6.0, while Estonian, with a mean per term of 3.4, and Russian with a mean of 3.1 have more terms listed by one participant only.

Table 1. Numeric overview of the collected data

	Tur list	Est list	Rus list	Est sort	Rus sort	Tur name	Est name	Rus name
Terms (overall)	978	1145	917	2145	1650	4565	2143	1649
Terms (different)	163	336	294	208 (172)	185 (170)	508	575	618
Mean per term	6.0	3.4	3.1	10.3	9.0	9.0	3.7	2.7
Words	1165	1157	1042	3451	3884	6787	2411	2354
Mean word length per term	1.6	1	1.3	1.6	2.4	2.1	1.4	2.0
Characters	6491	10072	8770	28837	30623	36515	25094	22459
Mean no.of characters per term	9.5	11.0	11.8	14.4	19.9	12.0	15.4	18.4

Turkish is a Turkic language, which is characterised by vowel harmony and extensive agglutination, while Estonian has lost its vowel harmony. Estonian is from the Finnic branch of the Uralic language family and is mostly agglutinative. Russian is from the eastern branch of the Slavic family of languages in the Indo-European language family. Russian was transliterated from the Cyrillic alphabet. Russian has gender, e.g. and differentiates between the feminine *temnaja birûza* and masculine *temno-birûzovyj* ‘dark turquoise’.

The number of words in Table 1 was calculated by counting the whitespaces in a row, which depends on orthography. For example *koyu mavi* ‘dark blue’ is written separately in Turkish, while a hyphen is used in the Russian *temno-sinij*, or *tëmno-sinij* ‘dark blue’ (with an ‘ë’). In the Estonian *tumesinine* ‘dark blue’ (*tume* ‘dark’ and *sinine* ‘blue’) form one word. When whitespaces are counted in a data row then only the Turkish *koyu mavi* ‘dark blue’ counts as two words. In list tasks the scale of the number of words is more or less equal. Counting whitespaces to count words fails for counting the number of words in the Estonian naming task as it gives 2411. The failure arises because compound words are written together as one, like *akvamariinsinine* ‘aquamarine blue’, *hallikassinine* ‘greyish blue’, *eresinine* ‘bright blue’, and *rukkilillesinine* ‘cornflower blue’. The mean number of characters in a term accurately reflects the length of the terms given, and so reflects the complexity of the terms if the length of a term is correlated to complexity. The Turkish naming task has over a six thousand words, and the Russian naming task over 2,300 and this is also reflected in the mean number of words per term, where only the Russian sorting task with an average of 2.4, the Russian naming task with 2.0, and the Turkish naming task with 2.1 average two or more words per term.

‘Characters’ in Table 1 are human-readable characters counted with the *nchar* function in R base. The Estonian *väga külm valkjas helesinine* ‘very cold whitish light blue' counts four words and twenty eight characters (including spaces), but *türkiishallikas* 'turquoise greyish' counts as one word and fifteen characters. The Estonian and Russian list tasks have more characters and longer terms than the Turkish list task. The mean number of characters per term – not per word – indicates the long, descriptive term names given in the Russian sorting task, at an average of 19.9 characters per term and naming task at an average of 18.4. The mean number of characters per term is lowest in the list tasks.

The following paragraphs briefly describe the types of terms listed, which included simple non-basic terms, modifiers, suffixes, compound terms or even expressions and the most frequent examples of them. Simple, non-basic terms are usually internationally used words, though even some expressions were used in more than one language, e.g. 'wet asphalt' in Estonian and Russian. The compound terms were the hardest to classify since there was a lot of variation. Sometimes the compound terms included a modifier or a suffix, or a compound term could be formed from two or more parts, e.g. *elevandiluuvärvi* 'ivory, literally elephant bone colour'. The types of terms used in the language depended on the grammar and orthography of the language, but some descriptive, compound terms, e.g. 'sky blue', 'blood red', 'grass green', 'snow white' and 'coal black' are almost universal.

There were 39 different simple terms listed 227 times in the Turkish list task (see Table 2). Eleven basic colour terms were listed 531 times. The most frequent simple, non-basic colour terms that were listed by more than one participant in total 207 times were *lacivert* 'dark blue' (38), *lila* 'lilac' (26), *bordo* 'burgundy' (24), *eflatun* 'mauve' (23), *bej* 'beige' (17), *turkuaz* 'turquoise' (16), *krem* 'cream' (11), *ela* 'hazel' (9), *haki* 'khaki' (8), *kizil* 'red' (6), *fusya* 'fuchsia', *füme* 'smoke' (5), *kiremit* 'brick', *menekse* 'violet', *turanj* 'orange' (3), *bronz* 'bronze', *dore* 'golden', *gömüs* 'silver', *kestane* 'chestnut', *leylak* 'lilac', and *petrol* 'petroleum' (2). The Turkish *kursun* 'lead' takes an Arabic *ī* in *kursuni* 'lead-coloured, grey' (2).

Acik 'light' was listed 30 times and *koyu* 'dark' 20 times for nine different terms and these were the most frequent modifiers. *Acik* 'light' was used for *acik mavi* 'light blue' (6), *acik pembe* 'light pink' (5), *acik kirmizi* 'light red', *acik sari* 'light yellow', *acik yesil* 'light green' (4), *acik kahve* 'light coffee' (3, non-BCT), *acik mor* 'light purple' (2), *acik haki* 'light khaki' (non-BCT) *acik kizil* 'light red' (1, non-BCT). *Koyu* 'dark' was used in *koyu yesil* 'dark green' (5), *koyu kirmizi* 'dark red' (4), *koyu mavi* 'dark blue' (3), *koyu mor* 'dark purple', *koyu pembe* 'dark pink' (2), *koyu kahverengi* 'dark brown', *koyu sari* 'dark yellow', *koyu siyah* 'dark black' (1), and *koyu kizil* 'dark red' (1, non-BCT). Other modifiers included *kirik* 'broken' (2) in *kirik beyaz* 'broken white', and *ucuk* 'pale' in *ucuk pembe* 'pale pink' (1).

Of 94 compound terms, 31 were listed more than once and 63 by only one participant. See also section 2.1 Introductory comments on Turkish colour terms. Some of them are written solid and some separate. The most frequent compound terms are *yavruagzi* 'peach' (9), *fistik yesili* 'pistachio green' (8), *gök mavisi* 'sky blue', *kavunici* 'light pinkish yellow' (6), *cimen yesili* 'grass green', *kiremit rengi* ‘brick coloured’, *sampanya rengi* ‘champagne coloured’, *visnecürügü* ‘purple brown’ (5), *deniz mavisi* ‘sea blue’, *parlament mavisi* ‘parliament blue’ (4), *cagla yesili* ‘almond green’, *fildisi* ‘ivory’, *gülkurusu* ‘dried rose’, and *petrol mavisi* ‘petroleum blue’ (3).

For an in-depth analysis of the morphological complexity of Estonian colour terms see Vilja Oja’s dissertation “Linguistic studies of Estonian colour terminology” (2001) and Urmas Sutrop’s “The vocabulary of sense perception in Estonian” (2002: 137–182).

The list task produced 33 simple, non-basic Estonian colour terms (see Table 6), most of which were listed by only one participant. There were 11 simple, non-basic terms listed by more than one participant. These were *beež* ‘beige’ (27)*, türkiis* ‘turquoise’ (10), *purpur* ‘purple’ *indigo* ‘indigo’ (6), *ooker* ‘ochre’ (5), *violett* ‘violet’, *kuld* ‘gold’, *hõbe* ‘silver’ (3), *terrakota* ‘terracota’, *pronks* ‘bronze’, and *bordoo* ‘bordeaux’ (2).

Most of Estonian list task terms are not simple as either they are compound words formed from mostly two or sometimes three words, or they used suffixes. Suffixes were also used in basic colour terms. Suffixed, non-basic one word terms with the suffix *-ne* include: *hõbedane* ‘silvery’*, kuldne* ‘golden’, *violetne* ‘violet’ (12), *purpurne* ‘purplish’ (2), *metalne* ‘metalic’, and *vaskne* ‘coppery’ (1); with *-jas*: *valkjas* ‘whitish’ (2), and *kahkjas* ‘sallow’ (1). The suffix *-jas* is more often used in a compound word, e.g. *valkjashall* ‘whitish grey’, *mustjas-sinine* ‘blackish blue’ (3). Of 31 terms used with the suffix *-kas,* only *kreemikas* ‘creamy’ was a one word term. The other terms with the *-kas* suffix were compounds with two words, and the most frequent *-kas* compound terms were *sinakasroheline* ‘bluish green’ (5), *sinakashall* ‘bluish grey’ (4), *hallikasmust* ‘greyish black’, *hallikasvalge* ‘greyish white’, *kollakasroheline* ‘yellowish green’, *punakaspruun* ‘reddish brown’, *rohekassinine* ‘greenish blue’ (3), *hallikas-roheline* ‘greyish green’, *hallikassinine* ‘greyish blue’, *lillakassinine* ‘purplish blue’, *pruunikasmust* ‘brownish black’, *rohekaskollane* ‘greenish yellow’, *rohekasmust* ‘greenish black’, and *sinakasmust* ‘bluish black’ (2).

The modifiers *hele* ‘light’ and *tume* ‘dark’ were again the most frequently used and they were used to modify twelve different terms. Eleven basic colour terms were used with these modifiers, and one simple, non-basic term *beež* ‘beige’. *Hele* produced *helesinine* ‘light blue’ (24), *heleroheline* ‘light green’ (14), *helepunane* ‘light red’ (12), *helekollane* ‘light yellow’ (10), *helehall* ‘light grey’ (7), *helelilla* ‘light purple’, *helepruun* ‘light brown’, *heleroosa* ‘light pink’ (5), *helebeež* ‘light beige’, *helemust* ‘light black’, *heleoranž* ‘light orange’, and *helevalge* ‘light white’ (1). *Tume* gave *tumesinine* ‘dark blue’ (16), *tumepunane* ‘dark red’, *tumeroheline* ‘dark green’ (11), *tumehall* ‘dark grey’, *tumekollane* ‘dark yellow’, *tumeroosa* ‘dark pink’ (8), *tumelilla* ‘dark purple’, *tumepruun* ‘dark brown’ (5), *tumebeež* ‘dark beige’ (2), *tumemust* ‘dark black, *tumeoranž* ‘dark orange’, and *tumevalge* ‘dark white’ (1). Modifiers that were listed more than once included *kahvatu* ‘pale’, *määrdunud* ‘dirty’ (4), *erk* ‘bright’ (3), *kärts* ‘bright, shocking’, and *sügav* ‘deep’ (2).

In the Estonian list task the two-part compound without modifiers was the most common type. There were 177 different two-word compounds without modifiers (365). The most frequent compound terms without modifiers were *taevasinine* 'sky blue' (13), *potisinine* 'pot blue' (11), *meresinine* 'sea blue', *samblaroheline* 'moss green' (10), *süsimust* 'coal black', *tibukollane* 'chick yellow' (9), *türkiissinine* 'turquoise blue' (8), *mereroheline* 'sea green' (7), *beebiroosa* 'baby pink', *kirsipunane* 'cherry red', *lumivalge* 'snow white', *purpurpunane* 'purple red', *sidrunikollane* 'lemon yellow', *veripunane* 'blood red', and *vesihall* 'water grey' (6). Of the twenty listed three-word compounds only *telliskivipunane* 'brick red' (4) was listed by more than one participant. In the list task 14% of the Estonian participants used a *hele* 'light' modifier and 13% a *tume* 'dark'one, while only 6% of the Russian participants did so. 'Light' and 'dark' were the most frequent modifiers.

Twelve basic colour terms were counted 311 times in the Russian list task (see Table 10). Most of the one-word, non-basic terms were suffixed (394, 110 different terms, 54 listed once). The most frequent simple non-basic colour words without a suffix (59) were *haki* 'khaki', *ochra* 'ochre' (7), *bordo* 'bordeaux', *indigo* 'indigo (4), *kadmij* 'cadmium', *kobalt'* 'cobalt', *marengo* 'marengo', *terrakota* 'terracota' (3), *kinovar'* 'cinnabar', *lazur'* 'azure', *mahagon* 'mahogany', and *siena* 'sienna' (2). To form an adjective, a suffix is added depending on the gender of the word (usually *svet* 'colour'), e.g. *bordo, bordovyj* 'bordeaux, bordeaux [coloured]'. The most common one word non-basic terms with an adjectival suffix (341) were *birûzovyj* 'turquoise' (17), *beževyj* 'beige' (16), *purpurnyj* 'purple' (15), *sirenevyj* 'lilac' (13), *bordovyj* 'bordeaux', *lilovyj* 'mauve' (12), *salatovyj* 'lettuce' (11), *alyj* 'scarlet', *limonnyj* 'lemony', *zolotyj* 'golden' (9), *malinovyj* 'rasberry', *višnëvyj* 'cherry' (8), *buryj* 'brown red', *serebrânyj* 'silver' (7), *izumrudnyj* 'emerald' (6).

The modifiers *svetlyj* 'light' (1), *chistyj* 'clean' (1), *jasnyj* 'clear' (1), *bledno* 'pale' (2), *grâzno* 'dirty' (1), *jarko* 'bright' (3), *matovo* 'matt' (1), and *pastel'no* 'pastel' (1) were listed, but the most frequent modifiers were *svetlo* 'light'(27) and *tëmno* 'dark' (26). *Svetlo* 'light' was used with all the basic colour terms except black and white. *Svetlo-zelënyj* 'light green' (5), *svetlo-koričnevyj* 'light brown', *svetlo-žëltyj* 'light yellow' (4), *svetlo-seryj* 'light grey', *svetlo-sinij* 'light blue' (3), *svetlo-krasnyj* 'light red' (2), *svetlo-fioletovyj* 'light purple', *svetlo-goluboj* 'light 'light blue'', *svetlo-oranževyj* 'light orange', *svetlo-rozovyj* 'light pink' (1) all appeared, as did the non-basic terms *svetlo-sinij metallik* 'light blue metallic' and *svetlo-zelënyj metallik* 'light green metallic' (1). *Tëmno* 'dark' was used with green (6), blue (5), red (4), brown (3), purple, yellow (2), orange, and grey (1, BCTs), and the same non-basic colour terms *sinij metallik* 'blue metallic' and *zelënyj metallik* 'green metallic' (1).

There were 91 compounds (118) in the list task. Here I had difficulty separating compounds from expressions. The most frequent compound terms included *cvet morskoj volny* 'sea wave colour' (10), *ul'tramarin* 'ultramarine' (7), *nebesno-goluboj* 'sky 'light blue'' (3), *akvamarin* 'aquamarine', *belosnežnyj* 'snow-white', *morskoj volny* 'sea wave', *ognenno-krasnyj* 'fiery red', *okis'* *hroma* ‘chromium oxide’, *slonovoj kosti* ‘ivory’, *snežno-belyj* ‘snow white’, *staraja roza* ‘old rose’, *žemchužnyj* ‘pearl’, and *vinno-krasnyj* ‘wine-red’ (2). Of the compound terms, 78 were named once, and 16 terms were classed as expressions. Three were listed more than once *mokryj asfal’t* ‘wet asphalt’ (3), *krasnoe derevo* ‘red tree’, and *kofe s molokom* ‘coffee with milk’ (2).

2.1 Introductory comments on Turkish colour terms

Turkish[8] has undergone and is still undergoing a lot of turnover, with some terms remaining the same while archaic terms are replaced with newer, more fashionable loanwords. Sometimes the archaic versions appear in the lists[9], and these have counterparts in other Turkic languages. While the colour terms for green and yellow have retained their form, others have not, so the term *kara* ‘black’ has the current basic colour term *siyah*, *ak* ‘white’ has the current basic colour term *beyaz*, *kizil* ‘red’ has the current basic colour term *kirmizi*, *gök* ‘sky; celestial’ is used most often in compound form and the current basic colour term is *mavi*, and *boz* ‘ash grey’ has the current basic colour term *gri*. The old terms for red, black and white are frequently used in many fixed forms, e.g. *Kizilayi* ‘Red Crescent’, *Karadeniz* ‘Black Sea’, or *Akbas* ‘Akbash region’, but they appear to have been replaced in the mental colour lexicon and are rarely used to name colour stimuli.

There are also newer loanwords that are not so frequently represented in Turkic languages. The Turkic dictionary mostly provides different stems for brown, purple, pink, orange and grey (Öztopcu 1999), which is consistent with the proposed evolution of basic colour terms (Berlin and Kay 1999). The Turkish *kahverengi* ‘brown’ (literally coffee + coloured) derives from the Arabic *qahwa* ‘coffee’ (Wehr and Cowan 1979: 930) and Persian *rang* ‘colour’ (Steingass 1892: 588). *Pembe* ‘pink’ can also be considered a Persian loan from *penbe* (*paṃba*, *puṃba*) (Steingass 1892: 256), which designated cotton in Otto-man Turkish (Devellioglu and Güneycal 2006: 857). The orange fruit is either *turuncu* ‘bitter orange’ or *portakal* ‘sweet orange’, although the colour of orange, or the basic term for orange is the former, *turuncu*. The most notable of the French loans is the basic colour term *gri* ‘grey’ from the same meaning as the French *gris*, but others include *füme* ‘smoke coloured’ from *fumée* ‘smoke’, and *somon* ‘salmon’ from *saumon* ‘salmon (fish); salmon coloured’ (Kann and

---

8 Although Turkish uses the Latin alphabet, it contains some letters, i.e. *ç*, *ğ*, *ı*, *ö*, *ş*, *ü* not found in English. These letters – except *ü* and *ö* which are used in Estonian and well - were transliterated.

9 The participants rarely used these older archaic terms in the list or naming tasks and each term appeared only once or twice. *Kızıl* ‘red’ is an exception, as it was elicited seven times in the list task. *Gök* cannot be considered an exception since it was not used separately, but rather in a compound word *gök mavisi* (*gökmavisi*) ‘sky blue’; the grammatically incorrect vernacular form *gök mavi* was also used by participants.

Leesi 1999: 299). Other words, e.g. *turkuaz* ‘turquoise’, *bej* ‘beige’, *sampanya* ‘champagne’, *krem* ‘cream’, and *bordo* ‘bordeaux’ etc. are internationally used loanwords.

To better understand the Turkish list task and even the naming task, a short overview of the language-specific, or perhaps culture-specific colour terms (Rätsep 2011: 355–356), e.g. *yavruagzi* ‘peach’, *kavunici* ‘light pinkish yellow’, *camgöbegi* ‘pale bluish green’, *visnecürügü* ‘purple brown’ and others is included.

Yavruagzi is used mostly by female language guides because of the nature of the word. *Yavru* can be used as a nickname for new-borns, and in the vernacular also for a beautiful young adolescent or a child, and for baby animals and birds, e.g. kittens. The second half of the compound term *agiz (-agzi*) means ‘mouth’, so literally it is ‘baby’s mouth’, although in dictionaries it is usually described as a “pinkish orange colour” (Bezmez *et al.* 2001: 811). The translation given to this term varies, as Özgen and Davies (1998: 926) use ‘peach’, while Şahin Ekici, Yener, and Camgöz (2006: 471) translate it as ‘salmon’. In the naming task, stimulus ORO T3 was named *yavruagzi* by Turkish participants, while in the Estonian task (see Sutrop 2000) it was named either *roosa* ‘pink’ by 16%, or *oranž* ‘orange’ by 10% (Rätsep 2011: 376). In the Turkish naming task female participants used *yavruagzi* three times as frequently as males, but *kavunici* was used by female and male participants in equal measure. *Yavruagzi* is a Turkish colour term, but in other languages the compound ‘child’s mouth’ is used in a more figurative meaning. Estonian *lapsesuu* ‘child’s mouth’ is mostly used figuratively for a childishly direct way of speaking (Langemets 2009), perhaps comparably to the American English, where the idiom ‘out of the mouths of babes’ is used to for wisdom spoken by young and inexperienced people.

Kavunici is also a colour word that could be termed language-specific. It literally designates the inner part of the melon, specifically the fleshy innards. The colour of this melon flesh is usually described as pale yellowish orange (Kornfilt 2010: 523) or also as melon pink and yellowish pink (Bezmez *et al.* 2001: 442). Özgen and Davies (1998: 926) use the gloss ‘orange’.

Another compound term is *camgöbegi*, which designates a bluish-greenish colour. It is translated as aquamarine (Bezmez *et al.* 2001: 129). It consists of *cam* ‘glass, bottle’ and *göbek* (*göbegi*) ‘bottom; belly’, which relates to the shade apparent when looking through a coloured glass bottle. According to Prof. Örücü (University of Glasgow) the English gloss ‘bottle green’ might be misleading because the glass in question is of the antique variety, which was usually of a light turquoise blue colour.

Visnecürügü designates the colour of the sour cherry (*Cerasus vulgaris*) that is overly ripe and has gone sour or has bruising. Dictionary translations offer purplish brown, oxide brown or purple as possible glosses (Bezmez *et al.* 2001: 792). The Turkish word *visne* ‘sour cherry, morello cherry’ is considered a Bulgarian loan (*Güncel Türkce Sözlük*; Nişanyan 2007 s.v. *visne*) probably from the word *vjshn/a* ‘cherry (tree and fruit)’ (Jotov and Ponomareva 1959: 49). The Greek term *visiní* ‘cherry’ was relatively rare in the Classical Greek corpora (18 instances) and quite rare in the Greek naming task, but three language guides did name NCS stimuli as this, and it was used once for a Munsell stimulus (Androulaki *et al.* 2006: 11, 20, 30). Coincidentally there is a parallel with the dialectal Southern Estonian word *visna-*, *visnapuu* (tree) 'cherry tree', which is an older Russian loan from the word *вишня* [*vishnja]* 'cherry tree, cherry' with the same meaning, which has come to Russian from Proto Slavic (Blokland 2009: 452).

Kiremit 'clay roofing tile' leans more towards light brown or brown than to a brick colour, although it has been glossed as 'brick colour'. *Parlament mavisi* 'parliament blue, cobalt blue' (participants used either *parlament* or *parlement*) was used only by young female adults. This may be because the term is used primarily in the fashion industry. *Gülkurusu* (literally rose + dried) designates a pinkish colour with a purplish undertone, which develops when a pink rose dries. (Rätsep 2011: 355–356)

2.1.1 Turkish list task

This section starts with a short numerical overview and short list of the terms that made up the bulk of the data. The list task analysis contains two parts, with a graphical overview of the list percentage and mean position (see Figure 3) and a table overview (Table 2) of the same data with the terms listed and glossed, and with Sutrop's cognitive salience index also included. Table 2 provides more detailed information on the results. The end of the section compares the salience with previous research and the subchapter ends with a conclusion as to whether the list task supports the status of *lacivert* 'dark blue' as a basic colour term or not.

Usually a basic colour term is a simple word, but if it is not then by Berlin and Kay's definition of a basic colour term it should not be predictable from its parts, as with Turkish *kahverengi* 'brown', which is literally 'coffee-coloured'.

Sharing a type-of relationship violates the second primary characteristic of Berlin and Kay's definition of a basic colour term. A term may be in violation of the second primary criterion even if it does not literally include another term. According to Özgen and Davies (1998: 942) *lacivert* 'dark blue' is a type of *mavi* 'blue'.

The participants (N=56) listed 163 colour names 978 times. There are a small number of terms that were listed by a large percentage of the participants. The frequencies of ten colour words reached 490, which is half of the overall frequency: *yesil* 'green', *siyah* 'black', *sari* 'yellow', *beyaz* 'white', *mavi* 'blue', *kirmizi* 'red', *kahverengi* 'brown', *turuncu* 'orange', *mor* 'purple', and *pembe* 'pink'. These were the terms listed most often. Together with those ten terms, the frequencies of a further sixteen terms formed approximately three quarters of the list data. The bottom half of the list is allocated to 85 colour words that were listed only once. Some participants produced several terms that they were the only one to list.

Frequency (F) is the first list task measurement. It is also presented in percentage form (%), which supports the comparison of the Turkish, Estonian and Russian list tasks. This helps to orient the reader better, because it is the percentage value of the participants who used the term. The next measure is mean position (mp). Mean position suggests whether the term was listed at the beginning of the lists with a low mean position, or at the middle of the list with a higher mean position. The mean position is connected to the length of the lists. These two parameters of the list task are shown in increasing detail throughout the subchapter.

In Figure 3 each circle represents a listed term, and the high-frequency terms are labelled. Mean position is the variable on the x-axis and listing percentage is on the y-axis. The closer to the bottom right of the figure the term is, the higher its frequency and the lower its mean position, thus increasing the chances of it being a basic colour term. Figure 3 can be considered a pictorial representation of the Turkish list task.

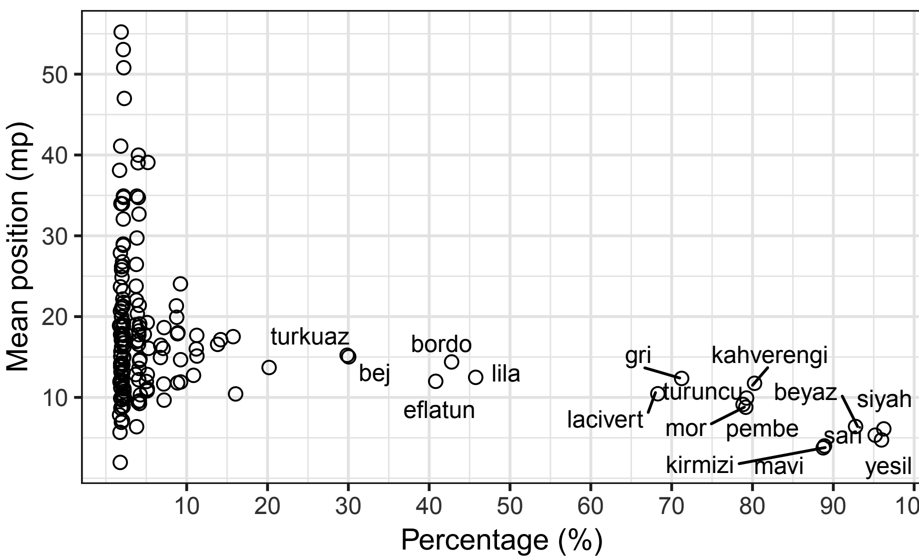

Figure 3. Turkish list task terms by percentage (≥ 30% labelled) and mean position.

Figure 3 is an attempt to follow some of the graphical design principles of Tufte (2001: 105), by showing the data, maximising the data-ink ratio, erasing non-data-ink and redundant data-ink, revising, and editing. This is but a simple iteration of Tufte's principles.

These graphs, used in Figure 3 for the Turkish list task, Figure 5 for the Estonian one, and Figure 12 for the Russian one, portray the list task better than if only one variable is displayed. They are meant to help illustrate the list task data in advance of the descriptive table.

From right-to-left in Figure 3 are the terms with the highest listing percentage and lowest mean position, and 17 terms with a listing percentage of

30% or higher are labelled. Starting from the right-most *yesil* 'green' and *siyah* 'black' (96%) with the highest frequency and moving to the left are *sari* 'yellow' (95%) and *beyaz* 'white' (93%). With an identical listing percentage (89%) and mean position, *mavi* 'blue' and *kirmizi* 'red' are situated on top of one another and are only visible because the data are staggered to give visibility for all the data points. After these six terms there is a slight drop in frequency and a rise in the mean positions in Figure 3, and next come four terms in close proximity. Of these four terms, *kahverengi* 'brown' (80%), *turuncu* 'orange', *mor* 'purple', and *pembe* 'pink' (79%), the latter three have similar mean positions, with *mor* having the lowest mean position and thus being placed lower on the y-axis. After a slight gap come *gri* 'grey' (71%), and *lacivert* 'dark blue' (71%). *Lacivert* 'dark blue'is a possible basic colour term, while the first eleven are all previously confirmed Turkish basic colour terms. There is a noticeable divide between the first twelve terms and the following five terms. The listing percentages of these five terms, *lila* 'lilac' (46%), *bordo* 'burgundy' (43%), *eflatun* 'mauve' (41%), *turkuaz* 'turquoise', and *bej* 'beige' (30%) are below the 50% mark or "listed by half of participants" in Figure 3, illustrating the sharp divide between the first *twelve* terms, and all the following terms.

Additional details from the Turkish list task are presented in Table 2, which adds Sutrop's cognitive salience index, or salience, to the list task parameters. Glosses, the cognitive salience index and ranks are shown in Table 2, which includes all terms with a frequency of five or over. All the listed terms were included when computing frequency (F, %), mean position (mp) and salience (S), but in the ranking order the terms that were listed only once were omitted. For example, *cingene pembesi* 'shocking pink, gypsy pink' and *süt beyaz* 'milk white' were both listed by only one participant, but since the terms were placed second on their list, then the salience score was relatively high at S = 0.0083, which is comparable to that of *fistik yesili* 'pistachio green' (F = 8, S = 0.0084).

The addition of salience in Table 2 only switches the positions of *gri* 'grey' (S = 0.0575) and *lacivert* 'dark blue' (S = 0.0641) from where they were in Figure 3, with the basic colour term *gri* exhibiting a slightly lower cognitive salience index than *lacivert*'dark blue'. Regrettably, Özgen and Davies (1998: 943) have not provided the list task mean position in their article, only their mean position ranking, so it is not possible to calculate a cognitive salience index from their results and compare the two sets of Turkish fieldwork data.

Taken together with a mean position of 10.6, the salience indicates *lacivert* 'dark blue' as the twelfth Turkish basic colour term, if the upper limit of eleven basic colour terms is discarded. This conclusion comes from the analysis of only the list task data.

The previously researched Turkish basic colour terms are confirmed with *lacivert* 'dark blue'included, because the results of frequency (68% ranked 12th), mean position (10.5 ranked 12th) and cognitive salience index (0.0653, ranked 11th) support the status of *lacivert* 'dark blue'as a basic colour term in the list task.

Table 2. Turkish list task terms by frequency (≥ 5), mean position and salience.
F – frequency, mp – mean position, Salience – cognitive salience index.

Turkish list task (N = 56)	Gloss	F	Rank	%	mp	Rank	Salience	Rank
yesil	green	54	1.5	96	4.9	3	0.1972	3
siyah	black	54	1.5	96	6.4	5.5	0.1518	5
sari	yellow	53	3	95	5.6	4	0.1695	4
beyaz	white	52	4	93	6.4	5.5	0.1450	6
mavi	blue	50	5.5	89	3.8	1.5	0.2350	1
kirmizi	red	50	5.5	89	3.8	1.5	0.2325	2
kahverengi	brown	45	7	80	11.5	21	0.0698	10
mor	purple	44	9	79	8.8	8	0.0896	7
pembe	pink	44	9	79	9.2	10	0.0858	8
turuncu	orange	44	9	79	10.2	14	0.0772	9
gri	grey	40	11	71	12.4	29	0.0575	12
lacivert	dark blue	38	12	68	10.6	16.5	0.0641	11
lila	lilac	26	13	46	12.3	28	0.0376	13
bordo	burgundy	24	14	43	14.3	35	0.0300	15
eflatun	mauve	23	15	41	11.8	23.5	0.0347	14
turkuaz	turquoise	17	16.5	30	15.2	40.5	0.0200	16
bej	beige	17	16.5	30	15.2	40.5	0.0199	17
krem	cream	11	18	20	13.8	34	0.0142	19
ela	hazel	9	19.5	16	10.6	16.5	0.0152	18
yavruagzi	peach	9	19.5	16	17.4	52	0.0092	20
haki	khaki	8	21.5	14	16.6	48	0.0086	21.5
fistik yesili	pistachio green	8	21.5	14	17	50	0.0084	23
acik mavi	light blue	6	24.5	11	12.5	30	0.0086	21.5
kavunici	light pinkish yellow	6	24.5	11	15.3	42	0.0070	27
gok mavisi	sky blue	6	24.5	11	16.3	45.5	0.0066	28
kizil	red	6	24.5	11	18	54.5	0.0060	30.5
acik pembe	light pink	5	30.5	9	12	26	0.0074	24.5
koyu yesil	dark green	5	30.5	9	12	26	0.0074	24.5
fusya	fuchsia	5	30.5	9	14.8	38	0.0060	30.5
fume	smoke	5	30.5	9	18.2	57.5	0.0049	34.5
kiremit rengi	brick coloured	5	30.5	9	18.2	57.5	0.0049	34.5
cimen yesili	grass green	5	30.5	9	19.8	64	0.0045	40.5
sampanya rengi	champagne coloured	5	30.5	9	21.4	66	0.0042	43
visnecürügü	purple brown	5	30.5	9	24.2	70	0.0037	49

2.1.2 Turkish naming task

Altogether the participants (N=56) listed 508 terms[10], of which a relatively large number, 284, were named only once. The name was written down exactly as it was said, which increased the variance of the terms considerably and therefore lowers the consensus.

The blue area of colour space had additional stimuli, which are not depicted in Figure 4 as I could not acquire the coordinates for them, and this heavily distorts the image of the distribution of stimuli. Consequently, with its larger quantities of blue stimuli the naming task also had a lot of modifiers.

Table 3 shows the total frequency, dominant frequency, number of tiles for which the term in question was the most frequently named term (*nmf* – name most frequent), and the *specificity index* (SI) for the most common terms in the naming task ranked by the total frequency. Table 2 includes the responses to all the stimuli, including the additional blue stimuli. For an overview with data on the 65 standard stimuli see Rätsep (2011). The additional stimuli are not depicted in Figure 4.

The specificity index (SI) is a ratio of the total frequency of use for each term and the total frequency for those tiles for which a term was dominant (Davies and Corbett 1995: 79), or dominance frequency divided by the total frequency. The specificity index (SI) is the most accurate indicator for basicness here, as it shows not only how many times a term was used, but how consensual the use was for each stimulus. The most common terms are ranked by their total frequency. The term *hardal sarisi* 'mustard-yellow' did not have the required frequency and was left out of Table 3, but it was nevertheless named most frequent for Color-Aid tile YOY-S2; *bej* 'beige' was also a low frequency term, though it was the most frequent name for ORO-S3.

Table 3 heavily reflects the selection of stimuli, which included a large area of the blue colour region. Since the participants were shown the stimuli one by one without any knowledge of the standard set of 65 stimuli and the additional non-standard bluish tiles, the data included here reflect the true proportions of the blue stimuli. Most notable are the high frequencies of occurrence for *acik mavi* 'light blue' (F = 202, SI = 0.139), *lacivert* 'dark blue' (F = 141, SI = 0.206), and *koyu mavi* 'dark blue' (139), which were expected, but also rather remarkable is the high frequency of the modified greens, *acik yesil* 'light green' (F = 125, SI = 0.224) and *koyu yesil* 'dark green' (F= 102).

Table 4 is quite large because it gives the most detailed overview of the Turkish naming task. The 65 standard stimuli results are comparable with those of Özgen and Davies (1998), see also Rätsep (2011a; 2011b).

---

[10] Technically 509 terms were listed, including seven blank "I do not know" answers.

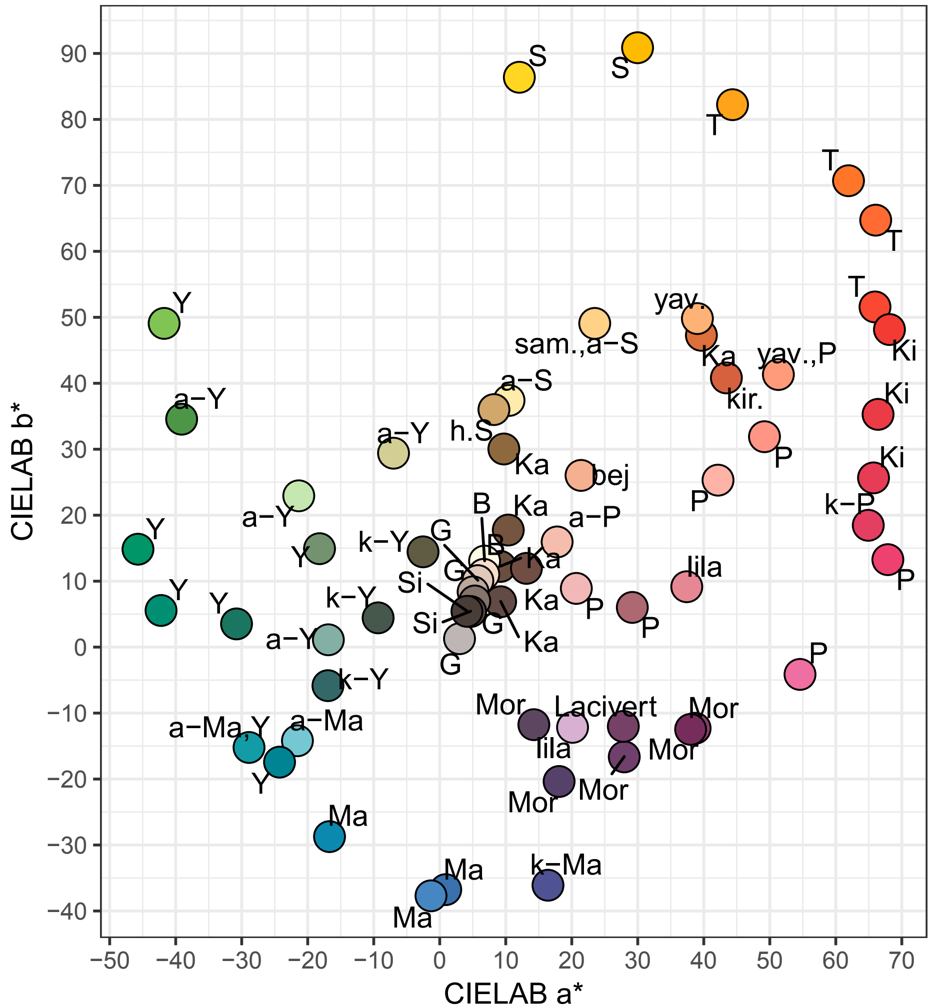

Figure 4. Turkish naming task most frequent stimuli names for 65 standard stimuli.

Most frequent stimuli names labelled: a – *acik* 'light', B – *beyaz* 'white', *bej* – 'beige', G – *gri* 'grey', h.S – *hardal sarisi* 'mustard yellow', Ka – *kahverengi* 'brown', Ki – *kirmizi* 'red', *Lacivert* – 'dark blue', lila – 'lilac', Ma – *mavi* 'blue', *Mor* – 'purple', P – *pembe* 'pink', S – *sari* 'yellow', Si – *siyah* 'black', T – *turuncu* 'orange', Y – *yesil* 'green', sam. – *sampanya* 'champagne', yav. – *yavruagzi* 'peach', kir. – *kiremit′* 'brick'.

Table 3. Turkish naming task terms ranked by specificity index.
F – frequency, DomF – dominant frequency, nmf – no. of tiles for which the term was named most frequently, SI – specificity index.

Term	**Gloss**	**F**	**DomF**	**nmf**	**SI**
siyah	black	99	94	2	0.949
sari	yellow	105	82	2	0.781
beyaz	white	56	40	2	0.714
gri	grey	145	99	4	0.683
mor	purple	233	144	8	0.618
kirmizi	red	130	80	3	0.615
turuncu	orange	114	70	4	0.614
mavi	blue	397	171	8	0.431
kahverengi	brown	149	62	6	0.416
yesil	green	251	67	7	0.267
pembe	pink	135	31	7	0.230
acik yesil	light green	125	28	4	0.224
lacivert	'dark blue'	141	29	4	0.206
acik mavi	light blue	202	28	8	0.139
koyu mavi	dark blue	139		1	
koyu yesil	dark green	102		3	
koyu pembe	dark pink	63		1	
lila	lilac	63		3	
eflatun	mauve	61			
acik pembe	light pink	59		1	
yavruagzi	peach	48		2	
acik mor	light purple	45		1	
turkuaz	turquoise	42			
acik gri	light grey	38			
kavunici	light pinkish yellow	36			
koyu sari	dark yellow	32			
koyu mor	dark purple	31			
sampanya	champagne	30			
acik sari	light yellow	29		2	
koyu kahverengi	dark brown	28			
kiremit rengi	brick coloured	28		1	

While Figure 4 depicted only the 65 standard stimuli and their most frequent names, Table 3 includes the two most frequent names for each of the stimuli used[11]. The most frequent term and the second most frequent term in Table 4 can be the same basic colour term with or without a modifier, or a specific term.

The specific terms usually come second after the term with the highest frequency for each stimulus. Two exceptions were *hardal sarisi* 'mustard yellow', which was the most frequently named term for stimulus YOY-S2 (F=6), and *kiremit rengi* 'brick colour' for Sienna Brown (F=8). The second most commonly named non-basic terms were *eflatun* 'mauve' (stimuli RVR-S1, VRV, V, VBV, VBV-T4; for VRV-S3, BV-T2* also given as *acik eflatun* 'light mauve'), *lila* 'lilac' (stimuli VRV-S3, VBV-T4; for RVR-S3 *acik lila* 'light lilac'), *bej* 'beige' (ORO-S3), *haki* 'khaki' (Y –S2), *kiremit* 'brick' (Sienna Brown) or *kiremit rengi* "brownish orange" (first for Sienna Brown; O-S1), *sampanya* 'champagne' (YO-T3), *turkuaz* 'turquoise' (BG, BG-T1, BG-T2*), *kavunici* 'light pinkish yellow' (YO-T3, OYO, O, ORO-T3, RO-T3), and *yavruagzi* 'peach' (ORO-T3, ORO-S3, RO-T3, for ROR-S3 also *toz pembe* 'powder pink'). A few basic terms modified by something other than 'light' or 'dark' were also named second most frequently: *askeri yesil* 'army green' (YG-S3), *fistik yesili*' pistachio green' (YGY), *kirli beyaz* 'dirty white' (White), and *kirli sari* 'dirty yellow' (YOY-S2). (Rätsep 2011)

The most common compound terms with *mavi* 'blue' were *gök mavisi* 'sky-blue', *deniz mavisi* 'sea-blue', *buz mavisi* 'ice-blue', *gece mavisi* 'night-blue', *petrol mavi(si)* 'petroleum blue' and *parlament mavisi* 'cobalt-blue' (Rätsep 2011: 378). The most common compound with *pembe* 'pink' was *toz pembe* 'powder pink', and also frequent were *seker pembesi* 'sugar-pink' and *ucuk pembe* 'faded pink'. A rather unusual colour word is *cingene pembe(si)*, grammatically *cingenepembesi* and literally 'gypsy pink', which designates a 'shocking pink, (coloured) a very bright shade of pink' (Bezmez *et al.* 2001: 167). For *sari* 'yellow' the most frequent compound terms were *hardal sarisi* 'mustard-yellow', *civciv sarisi* 'chick-yellow', *kirli sari* 'dirt yellow', *limon sarisi* 'lemon-yellow', *sampanya sarisi* 'champagne-yellow' and *ucuk sari* 'faded yellow'. For *gri* 'grey' the most frequent compound term was *metalik gri* 'metallic grey'; for *mor* 'purple' it was *patlican moru* 'aubergine-purple'; for *kirmizi* 'red' it was *parlak kirmizi* 'bright pink'. Even *beyaz* 'white' was modified, most often with *kirli* 'dirty', but also *kirik beyaz* 'broken white', *kir beyaz* 'dirty white' and *koyu beyaz* 'dark white'. (Rätsep 2011: 378). For a more detailed approach to gender differences in Turkish colour data see Rätsep (2013).

[11] Additional stimuli used in the Turkish list task are marked with '*', for more details see 1.4.4 Stimuli.

Table 4. Turkish naming task most frequent terms given to a stimulus
F – frequency, % – percentage (≥ 50% in bold), * – additional stimuli.

Tile	Term	F	%
Y	*sari* ‘yellow’	49	**88**
Y –S2	*kahverengi* ‘brown’	11	20
	haki ‘khaki’	7	13
YOY	*sari* ‘yellow’	33	**59**
	koyu sari ‘dark yellow’	12	21
YOY-T4	*acik sari* ‘light yellow’	13	23
	krem ‘cream’	9	16
YOY-S2	*hardal sarisi* ‘mustard-yellow’	6	11
	kirli sari ‘dirty yellow’	5	9
YO	*turuncu* ‘orange’	10	18
	sari ‘yellow’	9	16
	koyu sari ‘dark yellow’		
YO-T3	*acik sari* ‘light yellow’	9	16
	sampanya ‘champagne’		
	kavunici ‘light pinkish yellow’	5	9
YO-S3	*kahverengi* ‘brown’	25	45
	acik kahverengi ‘light brown’	6	11
OYO	*turuncu* ‘orange’	36	**64**
	kavunici ‘light pinkish yellow’	6	11
O	*turuncu* ‘orange’	34	**61**
	kavunici ‘light pinkish yellow’	3	5
O-S1	*kahverengi* ‘brown’	14	25
	kiremit rengi ‘brick coloured’	10	18
O-S3	*kahverengi* ‘brown’	29	**52**
	koyu kahverengi ‘dark brown’	7	13
ORO	*turuncu* ‘orange’	17	30
	kirmizi ‘red’	11	20
ORO-T3	*yavruagzi* ‘peach’	12	21
	kavunici ‘light pinkish yellow’	7	13
ORO-S3	*bej* ‘beige’	7	13
	yavruagzi ‘peach’	6	11
RO	*kirmizi* ‘red’	45	**80**
	acik kirmizi ‘light red’	4	7

Tile	**Term**	**F**	**%**
RO-T3	*pembe* 'pink'	8	14
	yavruagzi 'peach'		
	kavunici 'light pinkish yellow'	5	9
	acik pembe 'light pink'		
RO-S3	*kahverengi* 'brown'	26	46
	koyu kahverengi 'dark brown'	10	18
ROR	*kirmizi* 'red'	35	**63**
	acik kirmizi 'light red'	5	9
ROR-T3	*pembe* 'pink'	17	30
	yavruagzi 'peach'	10	18
ROR-S3	*acik pembe* 'light pink'	9	16
	yavruagzi 'peach'	4	7
	toz pembe 'powder pink'		
R	*kirmizi* 'red'	18	32
	koyu pembe 'dark pink'	8	14
R-T4	*pembe* 'pink'	19	34
	acik pembe 'light pink'	16	29
R-S3	*kahverengi* 'brown'	33	**59**
	koyu kahverengi 'dark brown'	6	11
RVR	*koyu pembe* 'dark pink'	15	27
	pembe 'pink'	11	20
RVR-S1	*pembe* 'pink'	7	13
	acik mor 'light purple'	5	9
	eflatun 'mauve'		
RVR-S3	*pembe* 'pink'	9	16
	acik lila 'light lilac'	7	13
	acik pembe light pink'		
RV	*mor* 'purple'	13	23
	koyu pembe 'dark pink'	9	16
RV-T2	*pembe* 'pink'	31	**55**
	koyu pembe 'dark pink'	5	9
VRV	*mor 'purple'*	34	**61**
	eflatun 'mauve'	6	11
VRV-S3	*lila* 'lilac'	23	41
	acik eflatun 'light mauve'	5	9
	eflatun 'mauve'		

Tile	**Term**	**F**	**%**
V	*mor* ‘purple’	43	**77**
	eflatun ‘mauve’	4	7
VBV	*mor* ‘purple’	33	**59**
	eflatun ‘mauve’	7	13
VBV-T4	*lila* ‘lilac’	17	30
	eflatun ‘mauve’	10	18
BV	*lacivert* ‘dark blue’	22	39
	mor ‘purple’	13	23
BV-T1*	*mor* ‘purple’	13	23
	mavi ‘blue’	10	18
BV-T2*	*lila* ‘lilac’	10	18
	mor ‘purple’		
	acik mor ‘light purple’		
	acik eflatun ‘light mauve’	4	7
	eflatun ‘mauve’		
BV-S1*	*mor* ‘purple’	34	**61**
	koyu mor ‘dark purple’	6	11
BV-S2	*mor* ‘purple’	18	32
	lacivert ‘dark blue’	11	20
BVB	*koyu mavi* ‘dark blue’	18	32
	mavi ‘blue’	12	21
BVB-T1*	*mavi* ‘blue’	26	46
	acik mavi ‘light blue’	14	25
BVB-T2*	*acik mavi* ‘light blue’	20	36
	mavi ‘blue’	19	34
BVB-T3*	*acik mavi* ‘light blue’	21	38
	mavi ‘blue’	13	23
BVB-S1*	*mavi* ‘blue’	15	27
	koyu mavi ‘dark blue’	8	14
BVB-S3	*gri* ‘grey’	30	**54**
	acik gri ‘light grey’	7	13
B	*mavi* ‘blue’	30	**54**
	koyu mavi ‘dark blue’	12	21
B-T1	*mavi* ‘blue’	43	**77**
	koyu mavi ‘dark blue’	5	9
B-T2*	*acik mavi* ‘light blue’	27	48
	mavi ‘blue’	16	29

Tile	**Term**	**F**	**%**
B-T3*	*acik mavi* 'light blue'	21	38
	mavi 'blue'	19	34
B-T4*	*acik mavi* 'light blue'	28	**50**
	mavi 'blue'	8	14
B-S1*	*mavi* 'blue'	19	34
	koyu mavi 'dark blue'	17	30
B-S2*	*lacivert* 'dark blue'	17	30
	koyu mavi 'dark blue'	15	27
B-S3*	*lacivert* 'dark blue'	29	**52**
	koyu mavi 'dark blue'	10	18
BGB	*mavi* 'blue'	36	**64**
	koyu mavi 'dark blue'	5	9
BGB-T3	*acik mavi* 'light blue'	18	32
	mavi 'blue'	13	23
BG	*yesil* 'green'	16	29
	turkuaz 'turquoise'	5	9
	mavi 'blue'		
BG-T1	*yesil* 'green'	9	16
	acik mavi 'light blue'		
	turkuaz 'turquoise'	8	14
BG-T2*	*acik mavi* 'light blue'	13	23
	turkuaz 'turquoise'	10	18
BG-S2	*koyu yesil* 'dark green'	20	36
	yesil 'green'	13	23
GBG	*yesil* 'green'	27	48
	koyu yesil 'dark green'	11	20
GBG-S2	*acik yesil* 'light green'	17	30
	yesil 'green'	11	20
G	*yesil* 'green'	33	**59**
	koyu yesil 'dark green'	13	23
G-S3	*koyu yesil* 'dark green'	20	36
	yesil 'green'	16	29
GYG	*yesil* 'green'	34	**61**
	koyu yesil 'dark green'	7	13
GYG-T4	*acik yesil* 'light green'	28	**50**
	yesil 'green'	8	14

Tile	**Term**	**F**	**%**
GYG-S1	*yesil* ‘green’	27	48
	acik yesil ‘light green’	11	20
YG	*yesil* ‘green’	19	34
	acik yesil ‘light green’	10	18
YG-S3	*koyu yesil* ‘dark green’	19	34
	askeri yesil ‘army-green’	7	13
YGY	*acik yesil* ‘light green’	17	30
	fistik yesili’ pistachio-green’	9	16
YGY-S3	*acik yesil* ‘light green’	22	39
	yesil ‘green’	6	11
GRAY-1	*beyaz* ‘white’	13	23
	gri ‘grey’	12	21
GRAY-2	*gri* ‘grey’	23	41
	acik gri ‘light grey’	13	23
GRAY-4	*gri* ‘grey’	38	**68**
	acik gri ‘light grey’	4	7
GRAY-6	*gri* ‘grey’	31	**55**
	koyu gri ‘dark grey’	13	23
GRAY-8	*siyah* ‘black’	39	**70**
	koyu kahverengi ‘dark brown’ *acik siyah* ‘light black’ *koyu gri* ‘dark grey’	3	5
BLACK	*siyah* ‘black’	55	**98**
WHITE	*beyaz* ‘white’	40	**71**
	kirli beyaz ‘dirty white’	5	9
ROSE RED	*pembe* ‘pink’	12	21
	kirmizi ‘red’	11	20
SIENNA BROWN	*kiremit rengi* ‘brick colour’	13	23
	kahverengi ‘brown’ *kiremit* ‘brick’	8	14
Cobalt Blue*	*mavi* ‘blue’	28	**50**
	koyu mavi ‘dark blue’	11	20
Navy Blue*	*lacivert* ‘dark blue’	24	43
	koyu mavi ‘dark blue’	10	18
Cyan Blue*	*mavi* ‘blue’	34	**61**
	koyu mavi ‘dark blue’	8	14

Some basic terms had more than one tile with consensus of 50% or more. Amongst the high frequency names for stimuli were *siyah* 'black' (98% for tile Black, 70% for Gray 8), *beyaz* 'white' (71% for tile White) and *gri* 'grey' (68% for Gray 4, 55% for Gray 6, 54% for BVB-S3). Other basic colour terms that had a high consensus percentage for at least one tile were *sari* 'yellow' (88% for Y, 59% for YOY), *kirmizi* 'red' (80% for RO, 63% for ROR), and *mor* 'purple' (77% for V, 61% for VRV, BV-S1*, 59 for VBV). *Mavi* 'blue' had several tiles dominant, with B-T1 the highest with 77% consensus, followed by BGB (64%), Cyan Blue* (61%), and B (54%). Three stimuli were borderline dominant at 50% (F=28), and these were Cobalt Blue* for *mavi* 'blue', B-T4 for *acik mavi* 'light blue', and GYG-T4 for *acik yesil* 'light green'. The basic terms that had a lower consensus value were *turuncu* 'orange' (64% for OYO, 61% for O), *yesil* 'green' (61% for GYG, 59% for G), *kahverengi* 'brown' (59% for R-S3, 52% for O-S3) and *pembe* 'pink' (55% for RV-T2). The additional tile B-S3 was dominant for *lacivert* 'dark blue'. It was the only additional stimulus that gained consensus of over 50% for *lacivert*'dark blue', but the non-contested basic term *pembe* 'pink' also had only one dominant tile with a similarly low consensus of 55% for RV-T2.

2.1.3 Conclusion of the Turkish results

The aim here is to establish the Turkish basic colour terms and ascertain the position of *lacivert* 'dark blue' (see Rätsep 2011a, 2011b). I conducted two field tests to find out if the position of *lacivert* 'dark blue' as the twelfth Turkish basic colour term would be supported or refuted by the fieldwork results. I used 82 tiles, 65 standard ones and 17 additional purple-blue, in the colour naming task to establish more precisely the foci of *lacivert* 'dark blue'.

The list task frequency was high for *lacivert*'dark blue', but even with the additional tile B-S3, which attained dominance, *lacivert* 'dark blue'would still be placed in the position of a probable basic colour term.

Table 5 indicates the very low listing percentages of 11% for *acik mavi* 'light blue' and 5% for *koyu mavi* 'dark blue'. This is especially relevant because they are the literal equivalents of the Estonian *helesinine* and *tumesinine* and the Russian *svetlo sinij* and *temno sinij*. The basic term *mavi* 'blue' is undisputedly supported by the list task results, while *lacivert* 'dark blue' has weaker support for its basic colour term status.

Table 5. Overview of Turkish *mavi* and *lacivert*
F – frequency (%), mp – mean position, S – cognitive salience index, N – no of participants using the terms (%), StR – number of stimuli named with a term.

Term	**Gloss**	**List task (N=56)** **F (%) – mp – S**	**Naming task (N=56)** **N (%) – F (F/N) – StR**
acik lacivert	light 'dark blue'		10 (18%) – 17 (1.7) – 10
acik mavi	light blue	6 (11%) – 12.5 – 0.0086	29 (**52%**) – 202 (7.0) – 22
koyu lacivert	dark 'dark blue'		11 (20%) – 17 (1.5) – 8
koyu mavi	dark blue	3 (5%) – 13 – 0.0041	26 (46%) – 139 (5.3) – 19
lacivert	'dark blue'	38 (**68%**) – 10.6 – 0.0641	28 (**50%**) – 141 (5.0) – 16
mavi	blue	50 (**89%**) – 3.8 – 0.2350	30 (**54%**) – 397 (13.2) –29

Naming task indicators are heavily influenced by the selection of stimuli, which had a disproportionate number of bluish stimuli. This circumstance upped the numbers for blue terms, especially for *acik mavi* 'light blue' (52%) and *koyu mavi* 'dark blue' (46%), and *mavi* 'blue' (54%) and *lacivert* 'dark blue' (50%). On the other hand, it also forced some participants to be more specific by using compound terms which were modified by 'light' or 'dark'. Juxtaposing with the list task results advocates for a weak basic status for *lacivert* 'dark blue' and a strong, confirmed basic status for *mavi* 'blue'. In any case the consensus in the colour naming task for *lacivert* 'dark blue' was unexpectedly low with only one additional tile gaining 52% dominance. This can be seen as *lacivert* 'dark blue' having a weak claim for being a basic colour term or being comparably weak alongside the established basic colour term *pembe* 'pink'.

The terms that gained a specificity index as shown in Table 3, are *siyah* 'black', *sari* 'yellow', *beyaz* 'white', *gri* 'grey', *mor* 'purple', *kirmizi* 'red', *turuncu* 'orange', *mavi* 'blue', *kahverengi* 'brown', *yesil* 'green', *pembe* 'pink', and *lacivert* 'dark blue'. The latter was situated between two modified terms *acik yesil* 'light green' and *acik mavi* 'light blue'.

There were terms that had the highest frequency for one or more tiles, nmf – named most frequent, but not high enough consensus (> 50%), and these were *koyu mavi* 'dark blue' (F= 139, nmf = 1 stimulus), *koyu yesil* 'dark green' (F=102, nmf = 3), *koyu pembe* 'dark pink' (F = 63, nmf = 1), *lila* 'lilac' (F = 63, nmf = 3), *acik pembe* 'light pink' (F = 59, nmf = 1), *yavruagzi* 'peach' (literally 'baby-mouth') (F = 48, nmf = 2), *acik mor* 'light purple' (F = 45, nmf = 1), *acik sari* 'light yellow' (F = 29, nmf = 2), and *kiremit rengi* 'brick coloured' (F = 28, nmf = 1). These terms, however, had many instances of the modifiers *acik* 'light', and *koyu* 'dark'.

The terms with modifiers, e.g. *acik yesil* 'light green', *koyu yesil* 'dark green', and *acik mavi* 'light blue' are similar to the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' in that these are modified terms which do include another basic term in their name, and so are not applicable to the definition of a basic colour term. However, the values in the Turkish list task suggest some basic traits.

2.2 Introductory comments to Estonian colour terms

The following comments are included to provide general context for the Estonian experimental data. Compound terms were also common in the Turkish results, and the importance of compound terms for the Estonian results is underlined in the following paragraphs. From the first list task table (see Table 6) to the conclusion of the sorting results, the combinatory nature of Estonian colour terms is evident. The use of compound terms is also common in other Finno-Ugric languages like Hungarian (Uusküla and Sutrop 2007), Finnish (Uusküla 2007), Udmurt and Komi (Ryabina 2011).

Another previous author, Vilja Oja (1998: 32) emphasises not only the importance of contact with other Finnic languages and the similarities of some traits of their colour words systems, but also the role of German:

> "Like Estonian, both Finnish and German have colour names with modifying suffix (e.g., Germ. *grünlich*, Est. *rohekas*, Fin. v*ihreähkö* 'greenish'). A great number of compound colour names in German and Finnish have the same structure as in Estonian (e.g., Germ. *schneeweiss*, Est. *lumivalge*, Fin. *lumivalkoinen* 'snowy white'; Germ. *dunkelbraun*, Est. *tumepruun*, Fin. *tummanruskea* 'dark brown'; Germ. *goldfarbig*, Est. *kullakarva*, Fin. *kullanvärinen* 'golden' etc.)"

Most of Oja's (1998) database of Estonian colour adjectives, which contains about 1,400 different colour terms are "either compound words or phrasal units". All compound words can be analysed in two parts, with the first attributive part and the final base component, either of which can occur "separately as a root word, suffixed derivative, compound word, or word combination" (Oja 1998).

For some terms there is a preference or prototypicality for a certain colour, e.g. *potisinine* 'pot's blue', *mürkroheline* 'poison green', *tibukollane* 'chick yellow', *süsimust* 'coal black', *kirsipunane* 'cherry red', or *lumivalge* 'snow white'. However, what may seem to be a fixed term for one participant, e.g. *meresinine* 'sea blue' (10) and *mereroheline* 'sea green' (7), may not be so for another. *Türkiis* 'turquoise' (10) was only combined with blue in *türkiissinine* 'turquoise blue' (8), possibly because *türkiisroheline* 'turquise green' is not so frequently used.

> "We cannot always pinpoint just one single hyperonym for each particular color name. Some terms function as hyponyms of several equal general terms. For example, Est. *türkiis* 'turquoise' is considered by some informants 'blue,' by other 'green'. /---/ (Oja 2007: 19)

A loanword from the German *Türkis* (Wiedemann 1973: 1238), French *turquoise* and Turkish *turkuaz* (Vääri, Kleis and Silvet 2012: 1163), it is defined as a 'sky-blue or greenish opaque mineral, a semi-precious stone' (Langemets 2009). The Estonian *türkiis* 'turquoise' does not seem to be as prevalent as the Russian equivalent *birûzovyj*.

The search from *The Corpus of Estonian Literary Languages* (2014) attained 57 tokens for *türkiis*, mostly in the contexts of fashion, semi-precious stones and jewellery. The search also yielded 93 tokens for *türkiissinine* 'turquoise blue', mostly in the contexts of water, fashion and clothing. The context of water can be either artificial, such as an aquarium or swimming pool, or natural as in a river, the Atlantic or Indian oceans, the Käsmu, Black or Mediterranean seas, or a lagoon. Only two tokens were found for *türkiisroheline* 'turquoise green', the first for a *türkiisroheline* 'turquoise green' Fender Stratocaster guitar and the second for an accessory with a *türkiisroheline* 'turquoise green' belt. Steinvall (Steinvall 2002: 154), who calculated the semasiological salience of a nominal, which is the ratio of the frequency of a term within a specific field to the overall frequency of that term, found that *turquoise* shows a clear preference for the CLOTHING domain, while all the terms signifying pale purple like *lavender*, *lilac* and *mauve* are most frequently used in the domain of PLANTS. English nonbasic colour terms have a tendency towards the domain of CLOTHING, e.g. textiles, clothes, interior decoration, cosmetics, vehicles; nonbasic colour terms are extensively used only in the domain of PLANTS in the somewhat artificial context of horticulture (Steinvall 2002: 141,151).

While some other terms are either prototypical or can be guessed, the specific term *potisinine* from *pott* 'pot' + *sinine* 'blue', 'indigo-blue' (Saagpakk 1992: 651), is more oblique. It designates either the colour (hue) which is attained by dissolving indigo in fermented urine or a dull-shaded darkish blue (Langemets 2009), but the shade can vary depending on the concentration of the dye and thus it "may be light or dark, greyish or bright blue" (Oja 2007: 205). But the pot is only the vessel for the dye:

> "A number of modern townspeople believe it is an ugly colour of a greyish or purplish hue, as they associate the *pott*-component either with the bottom of a kettle or with a blue enamel vessel. Actually, indigo blue was used to get the purest blue possible. Its lightness or darkness depended – as usual in the dyeing of textiles – on how strong the solution was." (Oja 2000: 14)

That illustrates how some relatively frequently used terms have become opaque to modern users.

For a non-basic colour term, *beež* 'beige' appears frequently in the list task. Its infrequency in the naming task may be due to stimuli selection or more frequent usage in listing than naming, but previous study has shown that *beež* 'beige' is most often used for specific objects instead of abstract stimuli. Cross-linguistic analysis of data from Czech, Italian, Spanish, Estonian, Finnish and Hungarian colour listing and naming tasks indicates that "beige is a context-specific term used rather in reference to specific objects (e.g. clothes, shoes, leather or furniture) than to colour tiles per se" (Eessalu and Uusküla 2013: 175). Most literary examples in the *Eesti keele seletav sõnaraamat* (Langemets 2009) describe inanimate objects like *beež mantel, pluus, auto* 'beige coat, blouse, car', *beežid kingad* 'beige [plural] shoes', *suvila värviti beežiks* 'the summer house was painted beige' and *talle sobib beež* 'beige suits him/her'. In the list task, *beež* had only three compound terms: *helebeež* 'light beige' (F=2), *tumebeež* 'dark beige' (2) and *tuhkbeež* 'ash beige' (1). The low mean position of *beež* 'beige' in the Estonian list task (see Figure 5, Table 6) decreases the chances of it being a basic term candidate.

From the theoretical viewpoint following Brent Berlin and Paul Kay's basic term theory, it was hypothesised that the frequency, mean position and salience of the Estonian *helesinine* 'light blue' would be lower than those of the Russian *goluboj* 'light blue' in Estonia, because in Russian *goluboj* is considered basic (Paramei 2005; Winawer *et al.* 2007) and in Estonian *helesinine* is not considered a basic term (Sutrop 2002).

A basic colour term is not predictable from its parts (Berlin and Kay 1999: 6–7), so Finnish *vaaleanpunainen* is not 'light + red', but 'pink' (Uusküla 2007; Uusküla 2008: 389) and Turkish *kahverengi* is not 'coffee + coloured', but 'brown' (Özgen and Davies 1998). Its significance is not included in that of any other colour term, e.g. Russian *goluboj* 'light blue' and the Turkish *lacivert* 'dark blue' (Özgen and Davies 1998). These characteristics of a basic colour term are not applicable to the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue'. *Helesinine* and *tumesinine* are psychologically salient, which also means that they occur at the beginning of list tasks and are in the vocabulary of all or most language speakers. In the list task, the basic colour term should be listed roughly by more than half of the participants and ideally before non-basic terms.

The Russian *goluboj* fits the criteria for a basic colour term, therefore the comparison of the Estonian non-basic modified term *helesinine* and the Russian basic unmodified *goluboj* is of special interest because of the long-standing Russian language influence over its neighbouring country's language, Estonian (Hint 1990; 1991; Must 2000; Blokland 2009; Zabrodskaja 2009). For speculation about how the Russian blue categories could have destabilised and possibly split the Estonian concept of *sinine* 'blue' into sub-concepts, see Sutrop (2002: 73, 213). The opposite should also be considered, as Estonian could have influenced Russian in Estonia so that *goluboj* would be less salient for Estonian Russians in their list task than for Russians from Russia. Further comparison was made with the data reported earlier from the list tasks of Estonian and Russian.

2.2.1 The Estonian list task

This section is similar to the Turkish list task structure, in that it starts with a short numerical overview and a short list of the terms that made up the bulk of the data. The list task analysis contains two parts, with a graphical overview of frequency and mean position in Figure 5 and a table overview of the same data in Table 6, with the terms listed included and glossed and with Sutrop's cognitive salience index. This provides more detailed information on the results.

The end of the section compares the salience with previous research and concludes with whether the list task supports *helesinine* 'light blue' and *tume-sinine* 'dark blue' having basic status or not.

The participants (N=39) listed 336 colour names 1145 times and 126 terms were listed more than once for a total of 933 times. Of these, 27 terms had a total frequency of 571, which is about half the overall frequency and 88 terms with a frequency of 857 supplied about three quarters of the overall frequency. At 376, roughly a third of the overall frequency was for eleven terms *punane* 'red', *sinine* 'blue', *kollane* 'yellow', *roheline* 'green', *hall* 'grey', *oranž* 'orange', *valge* 'white', *must* 'black', *pruun* 'brown', *roosa* 'pink', and *lilla* 'purple'. A quarter of overall frequency is for the lowest frequency terms, the 38 terms which were listed by two participants, and the 210 terms listed by only one participant.

Figure 5 illustrates the Estonian list task in the way Figure 3 does the Turkish list task, and each circle in Figure 5 represents a listed term. The mean position and listing percentage are the variables. The terms listed most often (≥30%) are labelled. The closer to the bottom right of the figure a term is, the higher its frequency and the lower its mean position, increasing the chances of it being a basic colour term. The infrequently listed words taper off before the halfway point of 50%. Figure 5 offers a pictorial representation of the Estonian list task data.

Over 30% of participants listed the terms labelled in Figure 5: *punane* 'red', *sinine* 'blue', *kollane* 'yellow' (95%), *roheline* 'green' (92%), *valge* 'white', *hall* 'grey', *oranž*[12] 'orange' (87%), *must* 'black' (85%), *pruun* 'brown' (82%), *lilla* 'purple', *roosa* 'pink' (79%), *beež* 'beige' (67%), *helesinine* 'light blue' (59%), *tumesinine* 'dark blue' (38%), *heleroheline* 'light green' and *taevasinine* 'sky blue' (33%).

[12] In Figure 5 'ž' was replaced with '*z*'. *Oranž* 'orange' and *beež* 'beige' were transliterated for graphic purposes.

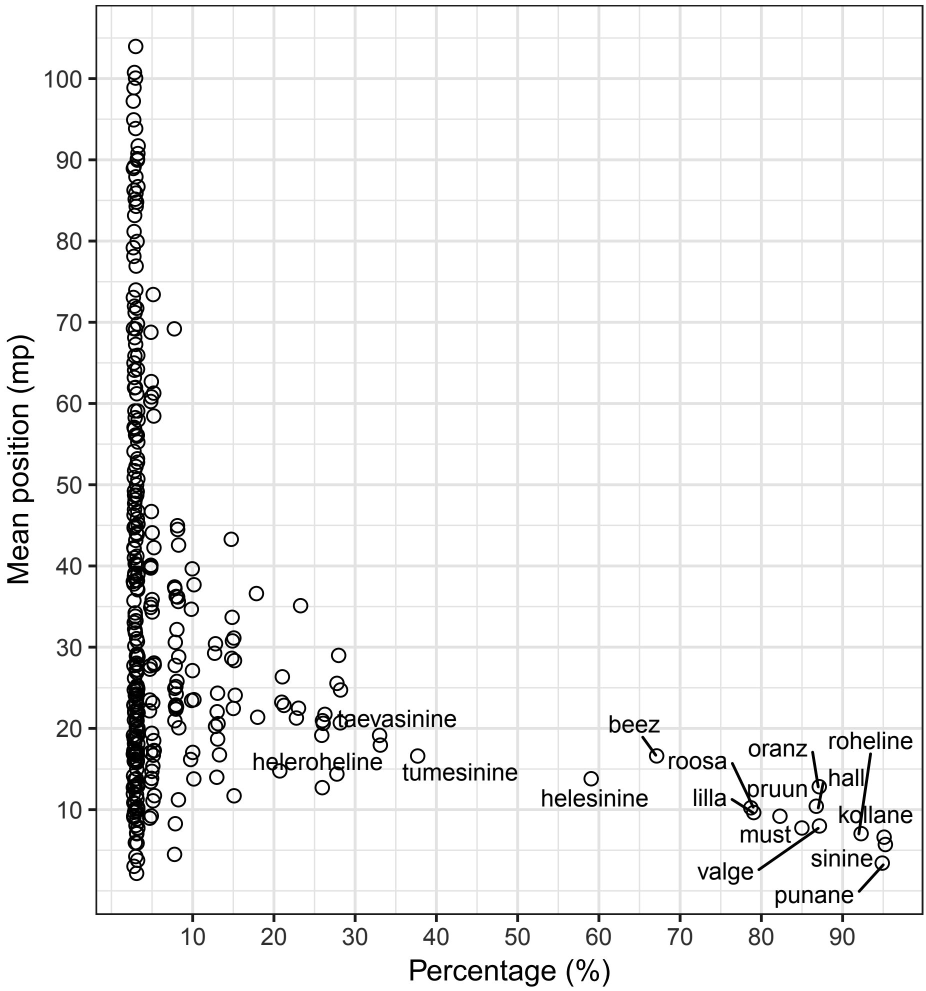

Figure 5. Estonian list task terms by percentage (≥ 30% labelled) and mean position.

In the bottom right of Figure 5 are the terms with the highest frequency and lowest mean position, leading with *punane* 'red' (mp = 3.2), *sinine* 'blue' (mp = 5.8) and *kollane* 'yellow' (mp = 6.4, 95%). The lower mean position of red 'punane' corresponds to a lower position on the y-axis of Figure 5. These three are followed by *roheline* 'green' (92%, mp = 7), which is trailed by *valge* 'white' (mp = 7.7), *hall* 'grey' (mp = 10.2), and *oranž* 'orange' (mp = 12.8, 87%). These three, which are almost on top of each other, are followed by the four basic colour terms *must* 'black' (85%), *pruun* 'brown' (82%), *lilla* 'purple' and *roosa* 'pink' (79%). The latter two data points are very close to each other differing only slightly in their mean positions. After these eleven basic colour terms there is a large gap in Figure 5 before the two in-between terms *beež* 'beige' (67%) and *helesinine* 'light blue' (59%). *Beež* and *helesinine* are literally situated between the lower and higher frequency colour terms, the higher frequency terms being the basic colour terms.

The half-way point on the x-axis at 50% in Figure 5 creates a sharp division between the high frequency terms at the bottom right and the rest of the terms, which form a triangle at the left side of the graphic. The left side of the plot, where the frequency is 1, shows a heavily dotted line going upwards, which perfectly illustrates the abundance of terms listed only once. It also indicates that mean position alone is not a valid indicator of basicness for low frequency terms. *Tumesinine* 'dark blue' (38%), *heleroheline* 'light green' and *taevasinine* 'sky blue' (33%) are also labelled.

Figure 5, the pictorial overview of Estonian list task, is followed by a table-format overview in Table 6. As did the Turkish list task data table, Table 6 provides additional information for the Estonian list task data by including glosses, the cognitive salience index and ranks.

Looking at Table 6 shows the large quantity of colour words with a modifying component in Estonian, especially *hele* 'light' and *tume* 'dark'. Some compound terms have a reference object as the attributive part, usually a familiar one like the Estonian *smaragdroheline*, and Russian *izumrudnyj*, *izumrudno-zelënyj* 'emerald [green]'. Components can sometimes be interchangeable, e.g. *meresinine* 'sea blue' (F = 10) and *mereroheline* 'sea green' (7). *Türkiis* 'turquoise' (F = 10) was only combined with blue, e.g. *türkiissinine* 'turquoise blue' (8).

The frequency (F) and mean position (mp) were computed with all 1145 terms listed, which covered 336 different terms. Of these, 210 terms listed only once were omitted when the rankings were computed for Table 6 because their mean positions and salience skew the data. However, the terms listed only once were included in Figure 5, where they form a dotted column at the left of the figure.

While *beež* 'beige' and *helesinine* 'light blue' were both listed by over half of the participants, the percentage listing of *tumesinine* 'dark blue' was 21% lower than that of light blue. Sutrop's (2002: 149) listing percentage difference for the modified blues, *helesinine* and *tumesinine* was smaller at 7.5%. Overall, there were slight differences in list task frequencies (see Sutrop 2000: 149–150). Here *violetne* 'violet' (F=11, 28%) was preferred over *violet* (F=3.8%). *Taevassinine* 'sky blue' and *veinpunane* 'wine red' did not appear here in the nominative, but only in the genitive as *taevasinine* and *veinipunane*. No colour terms with *neoon* 'neon' were listed in Sutrop's list task table, while here *neoonroheline* 'neon green', *neoonkollane* 'neon yellow', *neoonroosa* 'neon pink' and *neoonoranž* 'neon orange' were given.

Table 6. Estonian list task terms by frequency (≥ 5), mean position and salience.
F – frequency, % – percentage, mp – mean position, Salience – cognitive salience index.

Estonian list task (N=39)	**Gloss**	**F**	**Rank**	**%**	**mp**	**Rank**	**Salience**	**Rank**
punane	red	37	2	95	3.2	1	0.2925	1
sinine	blue	37	2	95	5.8	3	0.1625	2
kollane	yellow	37	2	95	6.4	4	0.1487	3
roheline	green	36	4	92	7	5	0.1319	4
valge	white	34	6	87	7.7	6	0.1136	5
hall	grey	34	6	87	10.2	13	0.0852	8
oranž	orange	34	6	87	12.8	20	0.0683	11
must	black	33	8	85	7.9	7	0.1066	6
pruun	brown	32	9	82	9.3	11	0.0881	7
lilla	purple	31	10.5	79	9.6	12	0.0824	9
roosa	pink	31	10.5	79	10.3	14	0.0770	10
beež	beige	26	12	67	16.7	32	0.0399	13
helesinine	light blue	23	13	59	13.8	22	0.0428	12
tumesinine	dark blue	15	14	38	16.9	34	0.0227	14
heleroheline	light green	13	15.5	33	18.2	39	0.0183	17
taevasinine	sky blue	13	15.5	33	18.9	42	0.0176	18
helepunane	light red	11	19	28	14.4	26	0.0196	16
violetne	violet	11	19	28	20.4	47.5	0.0139	20
hõbedane	silvery	11	19	28	24.8	73	0.0114	27
kuldne	golden	11	19	28	25.5	75	0.0111	28
potisinine	pot blue	11	19	28	28.8	86.5	0.0098	31
tumepunane	dark red	10	24	26	12.6	19	0.0204	15
tumeroheline	dark green	10	24	26	19.2	43	0.0134	22
türkiis	turquoise	10	24	26	20.4	47.5	0.0126	24
samblaroheline	moss green	10	24	26	21.1	51	0.0122	25
meresinine	sea blue	10	24	26	22	55.5	0.0117	26
helekollane	light yellow	9	28	23	21.3	52.5	0.0108	29
tibukollane	chick yellow	9	28	23	22.6	59	0.0102	30
süsimust	coal black	9	28	23	35	98	0.0066	43
türkiissinine	turquoise blue	8	31.5	21	15	27.5	0.0137	21
tumeroosa	dark pink	8	31.5	21	22.9	61	0.0090	33.5
tumekollane	dark yellow	8	31.5	21	23.4	66	0.0088	35
vanaroosa	old pink	8	31.5	21	26.2	77	0.0078	37
tumehall	dark grey	7	34.5	18	21.3	52.5	0.0084	36

Estonian list task (N=39)	Gloss	F	Rank	%	mp	Rank	Salience	Rank
mereroheline	sea green	7	34.5	18	36.7	104	0.0049	56.5
purpur	purple	6	40	15	11.8	18	0.0130	23
kirsipunane	cherry red	6	40	15	22.5	58	0.0068	41.5
indigo	indigo	6	40	15	24	69.5	0.0064	44
helehall	light grey	6	40	15	28.2	84	0.0055	52
vesihall	water grey	6	40	15	28.8	86.5	0.0053	53
sidrunikollane	lemon yellow	6	40	15	30.7	90.5	0.0050	55
veripunane	blood red	6	40	15	31.2	92	0.0049	56.5
beebiroosa	baby pink	6	40	15	33.5	94	0.0046	59
lumivalge	snow white	6	40	15	43	114	0.0036	70
purpurpunane	purple red	5	49	13	14.2	25	0.0090	33.5
ooker	ochre	5	49	13	16.8	33	0.0076	38
heleroosa	light pink	5	49	13	18.8	41	0.0068	41.5
tumepruun	dark brown	5	49	13	20.2	45	0.0063	45.5
helepruun	light brown	5	49	13	20.8	49	0.0062	47
veinipunane	wine red	5	49	13	21.8	54	0.0059	49
helelilla	light purple	5	49	13	24.6	71	0.0052	54
sinakasroheline	bluish green	5	49	13	29	88	0.0044	62
neoonroheline	neon green	5	49	13	30.6	89	0.0042	64

The list task results also confirmed that the two blues, *helesinine* (F=23; 59%) and *tumesinine* (F=15; 38%), were the most frequent of the light and dark modified pairs. The next most frequent light-dark modified pairs were light and dark green, red, yellow, grey, pink, purple, brown and beige.

According to Sutrop's cognitive salience index the twenty most salient Estonian terms (see Table 6) are *punane* 'red' (S = 0.2925), *sinine* 'blue' (0.1625), *kollane* 'yellow' (S = 0.1487), *roheline* 'green' (S = 0.1319), *valge* 'white' (S = 0.1136), *must* 'black' (S = 0.1066), *pruun* 'brown' (S = 0.0881), *hall* 'grey' (S = 0.0852), *lilla* 'purple' (S = 0.0824), *roosa* 'pink' (S = 0.077), *oranž* 'orange' (S = 0.0683), *helesinine* 'light blue' (S = 0.0428), *beež* 'beige' (S = 0.0399), *tumesinine* 'dark blue' (S = 0.0227), *tumepunane* 'dark red' (S = 0.0204), *helepunane* 'light red' (S = 0.0196), *heleroheline* 'light green' (S = 0.0183), *taevasinine* 'sky blue' (S = 0.0176), *violett* 'violet' (S = 0.0165), and *violetne* 'violet' (S = 0.0139). *Violett* 'violet', ranked nineteenth by cognitive salience index, is not included in Table 6, because it was only listed by three participants (F = 3.8%), but its mean position of 4.7 elevated its salience score to the top 20.

As indicated by previous studies, the results confirm 11 basic colour terms in Estonian: *punane* 'red', *sinine* 'blue', *kollane* 'yellow' (95%), *roheline* 'green' (92%), *valge* 'white', *hall* 'grey', *oranž* 'orange' (87%), *must* 'black' (85%), *pruun* 'brown' (82%), *lilla* 'purple', and *roosa* 'pink' (79%). The basic terms were followed in frequency by *beež* 'beige' (listed by 67% of participants), *helesinine* 'light blue' (59%), *tumesinine* 'dark blue' (38%), *heleroheline* 'light green' and *taevasinine* 'sky blue' (33%).

Despite being a modified term and thus a non-basic term by the original definition, the frequency, mean position and salience in the list task place *helesinine* 'light blue' together with *beež* 'beige' between basic and non-basic colour terms, indicating a relative basicness in the list task.

2.2.2 The Estonian sorting task

To ascertain whether the Estonian BLUE category shows any tendency to split into sub-groups, a free-sorting, or grouping, task and a colour-naming task were conducted. Following from Sutrop's earlier research (2001, 2002), it was hypothesised that the Russian term *goluboj* 'light blue' has influenced the Estonian BLUE category as well, causing *helesinine* 'light blue' or *tumesinine* 'dark blue' to acquire 'basicness' (Kerttula 2007), supplementing or even supplanting the unmodified *sinine* 'blue'. The modified terms *helesinine* 'light blue' and *tumesinine* 'dark blue' are not by definition basic colour terms (Berlin and Kay 1999: 6).

Here the Estonian free-sorting and colour naming results are analysed to indicate the nature and strength of any category salience variations and are overlaid on perceptual similarities to modify the structure of the Estonian blue category or categories. This research uses 55 colour stimuli, again from the Color-Aid Corporation, concentrating on the blue-green-purple neighbourhood of colour space (Bimler, Kirkland and Uusküla 2015; Bimler and Uusküla 2014; Uusküla 2014). In addition to a qualitative examination, the data are analysed with multidimensional scaling (MDS) and hierarchical clustering analysis (HCA).

A matrix of stimuli co-occurrence was created in which each cell entry – identified by its row and column indicating a pair of stimuli – designated the number of participants who placed that pair of stimuli in the same group. A high score for a pair of frequently co-occurring stimuli corresponds to high similarity.

Multidimensional scaling analysis and hierarchical cluster analysis were used to plot the results. Multidimensional scaling constructs a pictorial representation of the elements to summarise the values in a (dis)similarity matrix, so that distances between points reflect the dissimilarities between stimuli. In the non-metric form of MDS, distances are not a linear function of dissimilarities, but instead model the rank order of dissimilarities (Woods, Fletcher and Hughes 2003: 262).

In hierarchical clustering the stimuli are located at the end-points of branches as the leaves in a tree model or dendrogram, with the dissimilarities represented by the distances along branches between one leaf and another. The dendrogram can be imagined as a compromise across individual sorting structures, and Ward's method of agglomerative clustering was applied using the R statistics programme and RStudio (R Core Team 2012; RStudio Team 2016). Various packages were also used (Wickham 2012; 2007; 2011; Ross *et al.* 2013; Fox and Weisberg 2011; Adler, Murdoch and others 2014). Two outstanding resources that helped render the analysis into graphics were *R Graphics Cookbook* and *R in Action: Data Analysis and Graphics with R* (Chang 2013; Kabacoff 2015), which also clearly and effectively explained the principles of the analysis used.

The pictorial representation of the sorting task starts with stimuli mapped on the LAB colour space with each tile labelled with its most frequent term in Figure 6. Although variance, which is high, and consensus, which is low, are not pictured, Figure 6 allows for a simplified, but still accurate overview of the sorting task. In Figure 6 each stimulus is represented with a dot, and the size of the dot is correlated with frequency, so the higher the frequency as a percentage, the larger the stimulus dot.

Figure 6 depicts the non-yellow stimuli in the CIELAB colour space labelled with the most frequent group names. The most frequent stimuli names are either *roheline* 'green', *sinine* 'blue' or *lilla* 'purple', with three exceptions of *helesinine* 'light blue' (stimuli BG P2 3), *tumesinine* 'dark blue' (B S2), and *must* 'black' (B DS); two stimuli had no consensus (BG T4, C S3).

Altogether 311 groups were made, with participants each creating on average eight groups, or seven groups when the yellow distractor stimuli are excluded. More than half the participants (N=22) divided the stimuli into between five and eight groups, forming the peak of a skewed distribution. The long tail of this distribution consisted of 11 splitters with a low tolerance for dissimilarity within a group, and they created smaller but correspondingly more numerous groups, producing 10–29 of them. At the other extreme, six participants were lumpers, with relatively few groups, as they created two, three or four.

The most popular group names are listed in Table 7 together with how many participants used each name, how many times it was applied to stimuli overall (frequency), and the average group size. The names assigned to the groups reflect the selection of the stimuli for the blue-purple region of colour space.

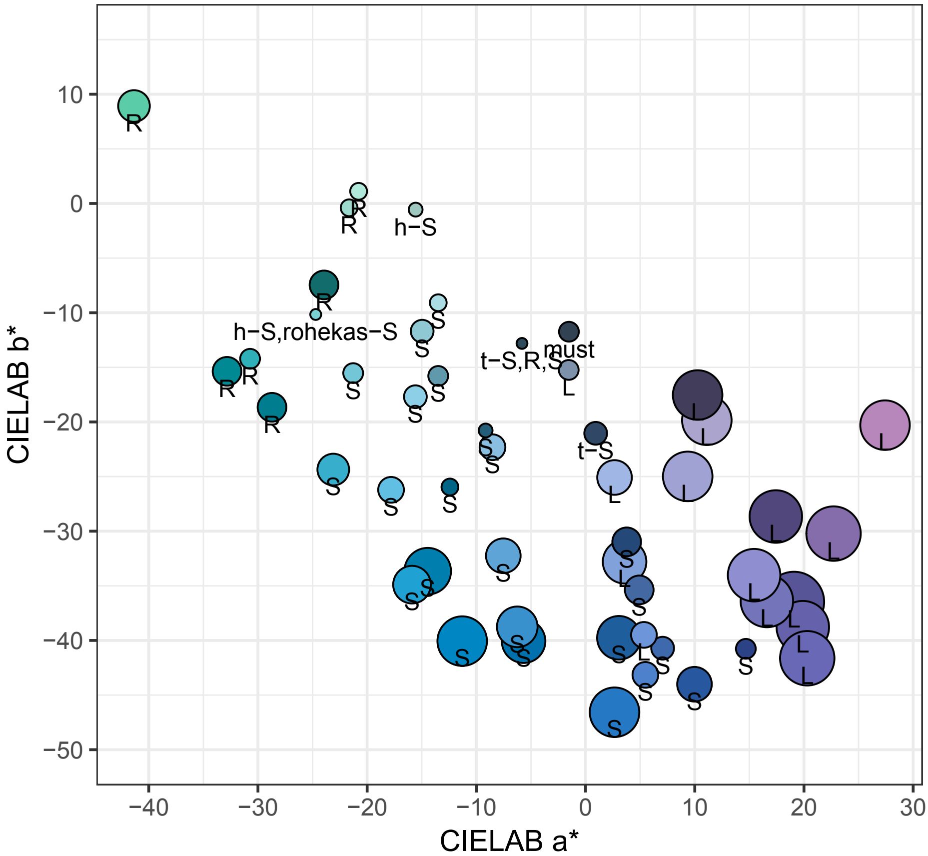

Figure 6. Estonian **s**orting task by most frequent stimulus name.

Most frequent stimuli names labelled: *rohekas* – ‘greenish’, R – *roheline* ‘green’, S – *sinine* ‘blue’, h-S – *helesinine* ‘light blue’, t-S – *tumesinine* ‘dark blue’, L – *lilla* ‘purple’, *must* – ‘black’.

Table 7. Estonian sorting task terms by participants' usage (≥ 6, 11%)

Term	Gloss	N	%	F	DomF	SI	No. of tiles (range)	No. of dominant tiles	Average group size
lilla	purple	21	54	255	20	0.760	26	1	12.1
*kollane**	yellow	20	51	75	57	0.078	4	3	3.8
sinine	blue	17	44	267			44		15.7
roheline	green	11	28	77			20		7.0
tumesinine	dark blue	7	18	35			15		5.0
helesinine	light blue	6	15	57			19		9.5
lillakas-sinine	purplish blue	6	15	34			20		5.7
helelilla	light purple	6	15	29			12		4.8
türkiis	turquoise	6	15	27			12		4.5
must	black	6	15	6			1		1.0

* distractor stimuli

More than half of the participants, agreed on naming one group *lilla* 'purple' (54%) and one *kollane* 'yellow' (51%). At the level of individual stimuli, these two terms became dominant for three tiles placed in a group labelled *kollane*, and one tile (BV) placed in a group labelled *lilla* by at least half the participants. Subsequent terms, beginning with *sinine* 'blue' (44%), occurred to only a minority of the participants as suitable names for their groups. Blue was followed by *roheline* 'green' (28%), *tumesinine* 'dark blue' (18%) and *helesinine* 'light blue' (15%). Six participants (15%) also made groups for *lillakassinine* 'purplish blue', *helelilla* 'light purple', *türkiis* 'turquoise' and *must* 'black', with a long tail of less frequent groups. This lack of consensus perhaps reflects the high variance in the number and composition of the groups.

Groups named *sinine* 'blue' were the largest on average with a mean of 16 stimuli per group. *Sinine* was thus the most frequent name, with 267 instances of a stimulus sorted into this group. *Sinine* also showed the widest range, with the highest number of stimuli (44) receiving that name at least once. Other frequent names were *lilla* 'purple' (255), *roheline* 'green' (77) and *kollane* 'yellow' (75). Six participants labelled one of their groups as *helesinine* 'light blue', and these groups had an average size of 9.5 and ranged across 19 stimuli in total. Seven participants formed a group they labelled *tumesinine* 'dark blue', and these averaged five stimuli in size and covered 15 stimuli in their total range. *Helesinine* was more frequent (57) than *tumesinine* (35), but both frequencies were low in comparison with *sinine* group name (267). There is little to support *helesinine* and *tumesinine* as categories as B T2 was sorted into *sinine* 'blue' 13 times, into *helesinine* 'light blue' twice, and once each into the groups *tumesinine* 'dark blue', *lilla* 'purple' and *heleroheline* 'light green'. This was an exception to a general rule that when some participants labelled a tile *sinine*, others used *helesinine* or *tumesinine*.

Next come the plots of the results for multidimensional scaling and hierarchical cluster analysis. Both aim to construct a pictorial representation of the task. First are the multidimensional scaling analysis results.

In multidimensional scaling the pictorial representation of the elements reflects the values in a (dis)similarity matrix, so that the distances between points reflect the dissimilarities between stimuli. However, the results of the multidimensional scaling analysis offer more support for the null-hypothesis, as the analysis does not support the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' as separate categories from *sinine* 'blue'. Figure 7 depicts the Estonian sorting task quite accurately, as supported by the table overviews.

It is assumed that the multidimensional scaling plane exhibits the essential structure of the (dis)similarity data, and interpreting this psychological map then raises the question of what psychologically meaningful dimensions span this space. The computer program assigns "Dimension 1" and "Dimension 2", which are the principal axes of the point configuration, but any other system of two coordinate axes also spans the plane. Consequently, Borg, Groenen, and Mair (2013: 11) postulate that the coordinate system should be found that is most plausible in psychological terms. One implied assumption of this thesis is that the colour stimuli used can be, and often are, divided by the light and dark axis. In trying to interpret the underlying psychological factors that govern the sorting behaviour by scrutinising the MDS model, I find it helpful to rotate an imaginary light-and-dark axis – horizontally or diagonally – on the model to see if there is a discernible division between light and dark. Borg, Groenen, and Mair (2013: 11) caution that these interpretations are meant as hypotheses about the attributes that the respondents use when they generate their similarity judgments, and whether this explanation is valid cannot be checked, because multidimensional scaling only suggests that this is a model that is compatible with the observations.

Figure 7 depicts all the non-distractor stimuli used in the sorting task, showing the disproportionate number of bluish stimuli, which may have forced some of the participants to group the non-bluish stimuli more tightly and be more specific for the blue stimuli. Conversely, it may simply reflect how the participants group the stimuli named with basic terms, in this case *kollane* 'yellow', *roheline* 'green' and *lilla* 'purple', more frequently than they do stimuli with non-basic group names.

Unfortunately, the analysis does not reveal a concentrated cluster of blue, but rather it indicates a continuum of blue stimuli. The pictorial representation of the sorting task (see Figure 7) does not show support for the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' bring separate groups.

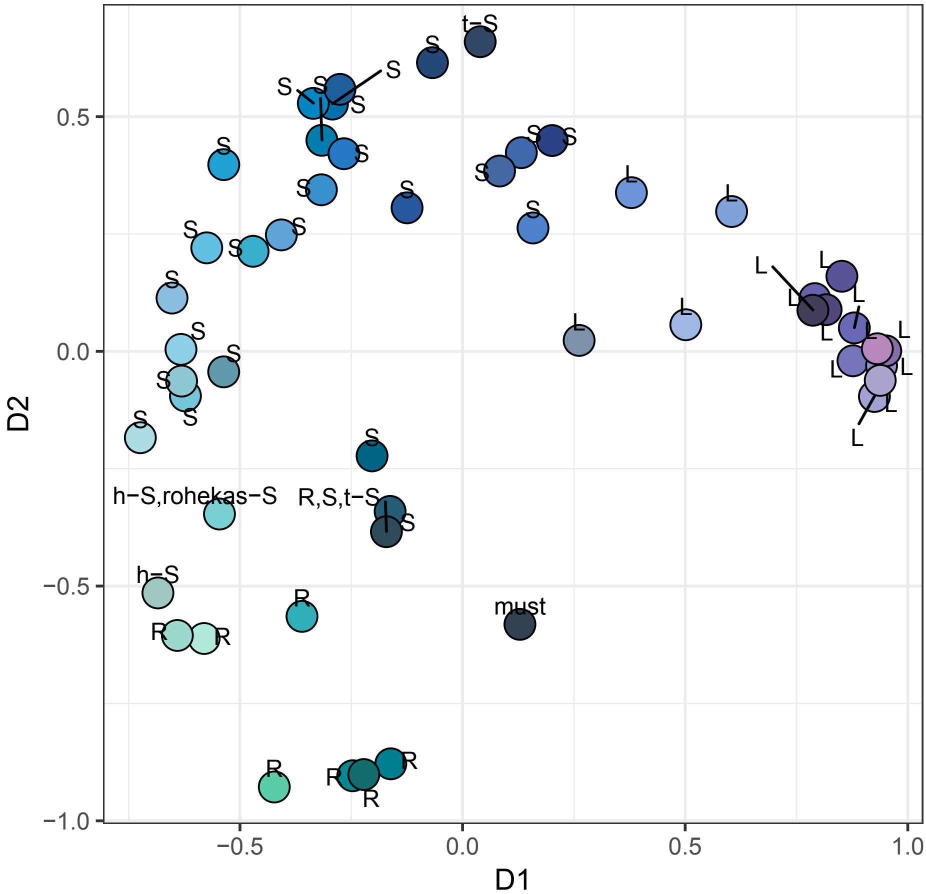

Figure 7. Estonian sorting task multidimensional scaling analysis.

Most frequent stimuli names labelled: S – *sinine* 'blue', L – *lilla* 'purple', R – *roheline* 'green', K – *kollane* 'yellow', H – *hele* 'light', T – *tume* 'dark', *rohekas* 'greenish', *must* 'black'.

The succeeding hierarchical cluster analysis reflects the results of the multidimensional analysis. Borg, Groenen, and Mair (2013: 74) caution that selecting different criteria in cluster analysis can lead to vastly different clusters. They state that this means cluster analysis is not a method for validating a multidimensional scaling analysis solution, but rather it forms groupings of points that tend to surface in a similar way to multidimensional scaling analysis solutions (Borg, Groenen, and Mair 2013: 74).

However, cluster analysis in Figure 8 supports the idea of the blue as a continuum, rather than a set of separate categories. The *tumesinine* branches out from the blue category among the darker purplish stimuli, while *helesinine* 'light blue' is among the lighter, greenish stimuli.

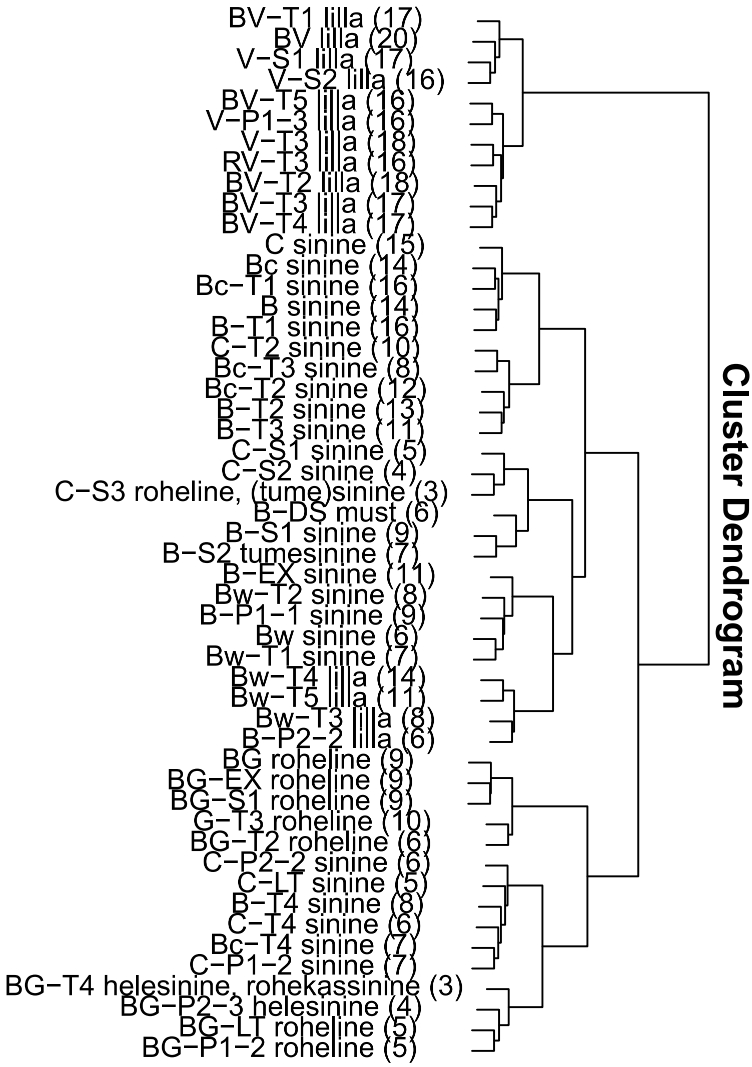

Figure 8. Estonian sorting task hierarchical cluster analysis.

In the end, *helesinine* 'light blue' and *tumesinine* 'dark blue' did not gain high enough frequencies in the sorting task, which is also evident from the comparison in Tables 14 and 15, which shows the overall low frequencies, and also the lack of consensus for *helesinine* 'light blue' and *tumesinine* 'dark blue' in the sorting task.

The summary output of the principal components analysis indicates that 41.8% of the variance is explained by the first component, 67.9% by the first two components, 81.5% by the first three, 87.2% by the first four, and 91% by the first five components. Inside these frames are the clusters and the cluster 's branches are the individual stimuli with their Color-Aid stimulus name, most frequent name (nmf) and frequency (F). Of the five clusters, three have the same most frequent names, *lilla* ‘purple’, *sinine* ‘blue’, and *roheline* ‘green’, while the two blue clusters have more than one most frequent name.

In Figure 8 the *lilla* ‘purple’ cluster with its two sub-clusters is relatively large at 11 stimuli, but all the stimuli within it have high frequencies of 16–20, meaning there is a consensus of 41–51% on the group name, which is the term bequeathed to the group after the sorting. Below the *lilla* ‘purple’ cluster is quite a large cluster with two large sub-clusters. One subcluster, which branches into two, even has a consensual name, *sinine* ‘blue’. The other part of the large cluster has the darker stimuli with blue, purple or black tones. It includes stimulus B-S2 *tumesinine* ‘dark blue’ (F=7), which is the only stimulus with *tumesinine* as its most frequently given group name. The cluster analysis representation in Figure 8 indicates another cluster for blue in the bottom of the figure. This low-frequency cluster of ten stimuli represents the lighter tones, including the one stimulus, BG-P2-3, which had nmf *helesinine* ‘light blue’ (F=4) and two *roheline* ‘green’ stimuli. Notably, all the warm stimuli with the Color-Aid notation of ‘w’ in the stimulus name, are in the other bigger blue cluster, which seems to be divided into warm and cool (denoted ‘c’) tones. Sutrop (2002: 73) notes that the Russian *goluboj* ‘light blue’ marks “only the light and cool tones of blue”. In fact, even the multidimensional scaling representation reflects the cool and warm continuum represented in the Color-Aid Corporation stimuli. *Sinine* ‘blue’ does not have a consensually named cluster, but rather two consensually named sub-clusters with two separate blue clusters.

A brief summary of the Estonian sorting task results can state that the analysis of the sorting data shows the hypothesis of Estonian having several separate categories for blue, with *sinine* ‘blue’, *helesinine* ‘light blue’ and *tume-sinine* ‘dark blue’, is not supported by the quantitative measures obtained from the task, notably the frequency of participants who sorted the stimuli and later named the groups *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’. This conclusion coincides that of with Bimler and Uusküla (2016), who used the author’s Estonian data with permission in their article along with data for five other languages, Russian, Italian, English, Lithuanian and Udmurt. Using the same Estonian data, they created an index that quantified the separation of light and dark blue and the strength of the category boundary in six languages. They found that the index was the lowest for English and Estonian, while Lithuanian, Russian and Italian values were high. On the provision that the clustering of the blue stimuli is taken as an additional indicator of basicness, Bimler and Uusküla (2016) conclude, like earlier research, that *goluboj* “‘light blue’ is a separate basic colour category in Russian and Italian, and [the data] further indicate that light blue terms are basic in Udmurt and Lithuanian, but not Estonian”.

2.2.3 The Estonian naming task

While the structure of how the Estonian naming task results are presented copies that of the presentation of the sorting task results, it must be stressed that the tasks are different because in the sorting task all the stimuli was presented together, and were grouped non-linguistically, so the naming of the groups came after all the stimuli were sorted. In the subsequent naming task, the stimuli were the same, but they were presented on a one-by-one basis, and they were named one after another. In short, the sorting task is reflective of grouping behaviour, while the naming task mirrors the names given to stimuli. Therefore, multidimensional scaling analysis is somewhat ill-suited for the naming task, but the most frequent stimuli names plotted in Figure 9 and the table with the most frequent names in Table 8 help in understanding the naming task results.

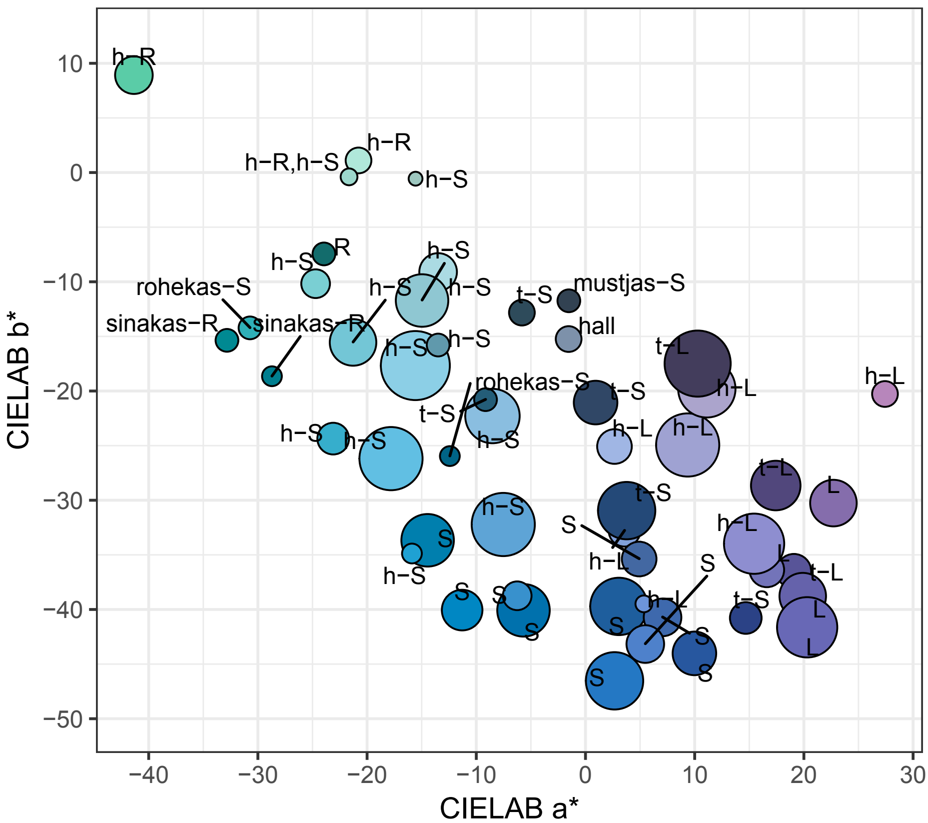

Figure 9. Estonian naming task by most frequent stimulus name.

Most frequent stimuli names labelled: h – *hele* 'light, *hall* 'grey', L – *lilla* 'purple', *must* 'black', *mustjas* 'blackish', R – *roheline* 'green', *rohekas* 'greenish', S – *sinine* 'blue', *sinakas* 'bluish', t – *tume* 'dark'.

The pictorial representation of the sorting task starting with stimuli mapped on the LAB colour space with each non-yellow stimulus labelled with its most frequent term is shown in Figure 9. Figure 9 allows for a simplified but still accurate overview of the naming task, reflecting the one-by-one presentation and naming of the stimuli with a higher number of different modifiers in the plot.

Each dot in Figure 9 represents a stimulus and the larger the dot, the higher the frequency. *Hele* 'light' (h) and *tume* 'dark' (t) are more prominent in the named most frequent stimulus names in Figure 9, especially in *helesinine* 'light blue' and *tumesinine* 'dark blue', but also in *helelilla* 'light purple' and *tumelilla* 'dark purple'. Given that purple stimuli, for example, had a smaller range and lower frequencies of use than the blue stimuli did, it appears that a more concise set of stimuli is easier for participants to name and thus the result is more consensual.

The following paragraphs give a brief numerical overview, with Table 8 conveying the parameters of the most frequent terms in the naming task. In this task the stimuli were shown and named separately, not all at once as in the preceding sorting task. There were 2145 possible combinations from 55 stimuli multiplied by 39 participants. The number of colour names given by each participant ranged from 3 to 47, with an average of 31.4, or 28.3 if the yellow items are excluded, and 19% of the names given were used only once. Comparison of the most frequent terms for each stimulus indicates a preference for modified terms in the naming task, which is to be expected, as the stimuli were shown and named one by one, although the consecutive order of the tasks might also have influenced the naming task.

The most frequent terms in the colour naming task (see Table 8) were *helesinine* 'light blue' (F= 211), *sinine* 'blue' (F= 180), *helelilla* 'light purple' (F= 117), *lilla* 'purple' (F= 107), *tumesinine* 'dark blue' (F= 80), *kollane* 'yellow' (F= 52) and *tumelilla* 'dark purple' (F= 51). *Helesinine* 'light blue' has become marginally more salient than *sinine* 'blue', reversing the case in the listing and sorting tasks. Three stimuli grouped under *helesinine* 'light blue' did gain dominance, but so did stimuli in the *helelilla* 'light purple' and *tumelilla* 'dark purple' groups (see Table 8).

Dominance is taken to mean more than half the number of participants, and the precise number can be quite a significant factor in the analysis of the data. The calculation of dominance uses the requirement that a term is applied by only the majority of participants, and so dominance is usually taken to be over 50% of the participants (> 50%). The precise number of the dominance threshold is quite an influential factor with the Estonian (N=39) data, especially in the naming task, as 39 divided by two is 19.5, so that the dominance frequency should be twenty and over.

The results with the dominance threshold at 20 are that the names given to the dominant stimuli were *tumelilla* 'dark purple' (SI = 0.431), *helelilla* 'light purple' (SI = 0.350), *helesinine* 'light blue' (SI = 0.308) and *lilla* 'purple' (SI = 0.187). Only the latter is a basic colour term, but it has the lowest specificity index. Surprisingly, given that a stimulus for the *lilla* 'purple' group gained dominance, the *sinine* 'blue' group had no dominant tiles.

If the cut-off point were to be lowered by half a point to 19, then the naming task would gain four more dominant stimuli and, more importantly, three more dominant stimuli *names*, these being *sinine* 'blue' and *tumesinine* 'dark blue', and also *kollane* 'yellow' from the distractor stimuli. *Tumesinine* is the only nonbasic term, while *sinine* and *kollane* are previously established Estonian basic colour terms. Of the initial four terms *tumelilla* 'dark purple', *helelilla* 'light purple', *helesinine* 'light blue' and *lilla* 'purple', only *helelilla* gained a dominant stimulus by the addition of V-P1-3 meaning that it had two dominant stimuli with one borderline (F=19) dominant stimulus. The three new names with the borderline dominant frequency at nineteen are *kollane* 'yellow' (stimulus Y), *sinine* 'blue' (two stimuli, B and B-T1) and *tumesinine* 'dark blue' (stimulus B-S1). It can be argued that including the borderline dominant tiles and names like this could be influenced by the hypothesis and that the rules are being bent to include data which would support the hypothesis. Equally though, lowering the threshold by one point from twenty to nineteen should not have much of an influence on the number of dominant *names*. Here, clearly, is a case where lowering the dominance threshold by one point, does affect which names are included as dominant.

The names that gained dominance when the threshold was lowered by one are marked with '*' in Table 8, which gives all the dominant stimuli names from the Estonian naming task. The names ranked by specificity index in Table 8 are *helelilla* 'light purple' (SI=0.513), *tumelilla* 'dark purple' (SI=0.431), *kollane* 'yellow' (SI=0.365), *helesinine* 'light blue' (SI=0.308), *tumesinine* 'dark blue' (SI=0.238), *sinine* 'blue' (SI=0.211), and *lilla* 'purple' (SI=0.187). The low specificity scores for the basic colour terms *lilla* 'purple' and *sinine* 'blue' could be attributed to the higher specificity scores of *helelilla* and *tumelilla* and *helesinine* and *tumesinine*. *Lilla* 'purple' had the highest percentage of participants using the term at 90%, followed by *helelilla* 'light purple' (87%), *helesinine* 'light blue', *sinine* 'blue'(85%), *kollane* 'yellow' (74%), *tumelilla* 'dark purple' (69%) and *tumesinine* 'dark blue' (67%). *Helesinine* 'light blue' had the highest overall frequency (F=211), the largest range of stimuli at 27 and the largest average group size with 6.4, while *helelilla* 'light purple', which also had three dominant stimuli, had an overall frequency of 117, a stimuli range of 15 and an average group size of 3.4. *Helelilla* 'light purple' has the highest specificity index due to the relatively high frequency for all the stimuli given that name. While the average group size column in Table 8 is more appropriate for the sorting task, it helps to illustrate the point of *helesinine* 'light blue' (range 27, average group size 6.4) and *blue* 'sinine' (range 24, average group size 5.5) having the largest range of stimuli given that name and the largest average group sizes, which lowered their specificity.

Table 8. Estonian naming task dominant stimuli names (at 49%) by specificity index N – number of participants, % – percentage of participants, F – frequency, DomF – dominant frequency, SI – specificity index.

Term	Gloss	N	%	F	Dom F	SI	No. of tiles	No. of dominant tiles	Average group size
helelilla	light purple	34	87	117	60	0.513	15	3	3.4
tumelilla	dark purple	27	69	51	22	0.431	4	1	1.9
*kollane**	yellow	29	74	52	19	0.365	4	1	1.8
helesinine	light blue	33	85	211	65	0.308	27	3	6.4
*tumesinine**	dark blue	26	67	80	19	0.238	13	1	3.1
*sinine**	blue	33	85	180	38	0.211	24	2	5.5
lilla	purple	35	90	107	20	0.187	18	1	3.1

* – dominance gained from lowering the threshold from 20 to 19.

To describe the high variance in the naming task more accurately, the terms that were used by one quarter or more of the participants in the naming task but with a low consensus should be noted. In Figure 9 there are some names most frequent for stimuli that are absent from Table 8. This is because *nmf*, named most frequent, for a stimulus, does not necessarily mean a high frequency count for the stimulus. Many names had a frequency of over a quarter of the participants (N≥ 10, as 39 divided by four is 9.75) using that term with a low consensus. Twelve terms had a frequency of over 25%, but low named most frequent scores for any one stimulus. These were *rohekassinine* (54% participant usage; nmf BG-T2 with 7 participants naming it so), *heleroheline* 'light green' (51%; nmf G-T3 F=12), *lillakassinine* 'purplish-blue' (51%; Bw-T1 (7)), *taevasinine* 'sky-blue' (46%; Bc-T2, Bc-T3 (4)), *sinakaslilla* 'bluish-purple' (44%; Bw-T2, Bw-T2, Bw-T3, B-P1-1 (3)), *meresinine* 'sea-blue' (38%; C-S2 (4)), *hallikassinine* 'greyish-blue' (36%; C-P2-2 (6)), *sinakasroheline* 'bluish-green' (36%; BG-EX (6)), *potisinine* "indigo-blue" (31%; B-EX, Bc, B-S1 (2)), *roheline* 'green' (28%, BG-S1 (7)), *türkiis* 'turquoise' (28%; BG (5)), and *hall* 'grey' (26%; B-P2-2 (8)). *Mereroheline* 'sea-green' and *sinakashall* 'bluish-grey' (N=9, 23%) for example fell just below the quarter threshold. The dominance threshold counts the number of participants. Some of the terms and their frequencies are also shown in Figure 11, which shows the naming data results for the hierarchical cluster analysis.

The following frequencies do not differentiate by participant, but rather give the number of terms that had low frequencies. There are 412 names given to a stimulus once, 59 names given twice, 33 given three times, 11 given four times, seven given five times, eight given six times, six given seven times, two given eight times, which were *akvamariin* ‘aquamarine’ and *mustjassinine* ‘blackish-blue’, and two more given nine times, which were *ilussinine* ‘pretty-blue’ and *rohekashall* ‘greenish-grey’.

Multidimensional scaling analysis is by default not best suited for a naming task, because in the naming task the stimuli were named separately, while the figure resulting from multidimensional scaling analysis shows the pictorial representation of the grouping behaviour, as in the sorting task. The format for naming data is equivalent to sorting data, as a co-occurrence matrix can be created in which each cell entry is the number of participants who labelled those two stimuli with the same name.

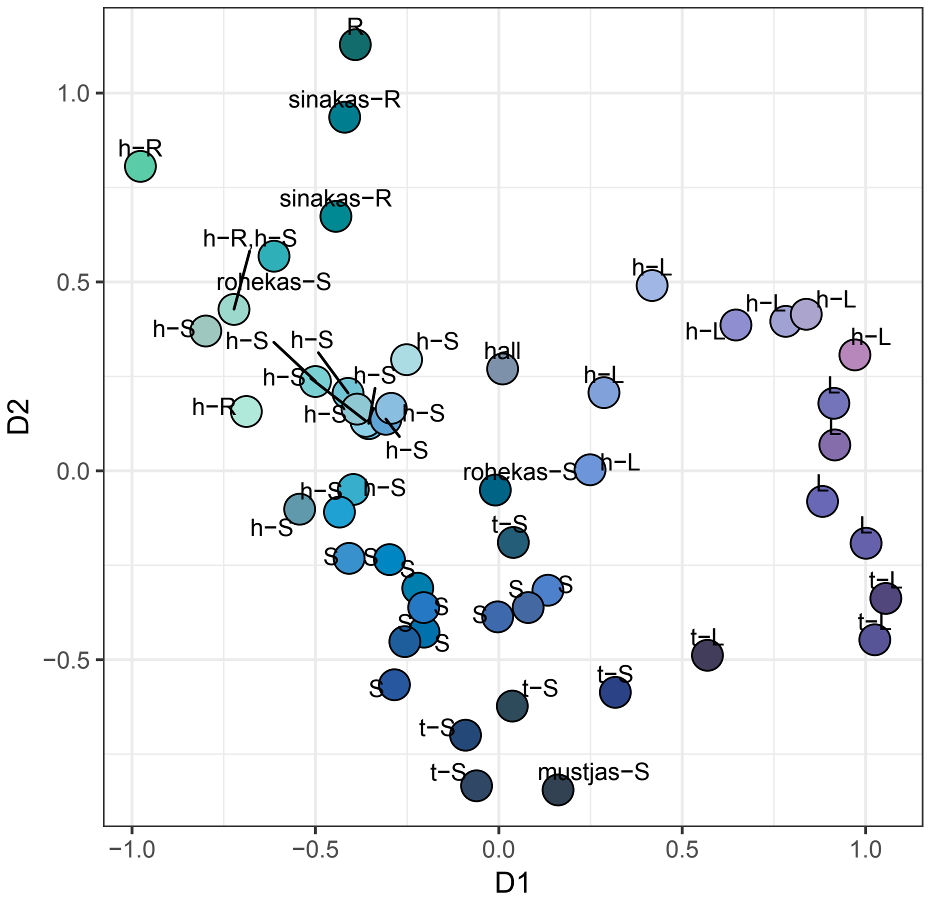

Figure 10. Estonian naming task multidimensional analysis.

Most frequent stimuli names labelled: *hall* – ‘grey’, h – *hele* ‘light’, L – *lilla* ‘purple’, *mustjas* ‘blackish’, R – *roheline* ‘green’, *rohekas* ‘greenish’, S – *sinine* ‘blue’, *sinakas* ‘bluish’, t – *tume* ‘dark’

Because of the greater specificity of naming, more terms were used with fewer stimuli per term, so the co-occurrence values are smaller than in the grouping task. Nevertheless, MDS transforms the matrix into a comparable 2D representation of the stimuli (see Figure 10).

In the Estonian naming task, the summary output of the principal components analysis reaches above 90% far more slowly, as only the thirtieth component explains 90% of the variance. The scree plot indicates the importance of the first two components, which account for 18% and 15% of the variance, followed by the third with 8%, and the fourth, and the fifth with 5% each. The first five components account for only about 52% of the variance.

Several smaller clusters that have the same most frequent name can be observed in hierarchical cluster analysis (Figure 11), but only two larger clusters in which stimuli have the same most frequent name, these being *sinine* 'blue' and *helesinine* 'light blue'.

The *lilla* 'purple' cluster has three small sub-clusters with stimuli V-S2, BV, V-S1 for *tumelilla* 'dark purple'; Bw-T5, BV-T4, BV-T5, V-P1-3 for *helelilla* 'light purple'; and BV-T1, BV-T2, V-T3, BV-T3 for *lilla*, with an additional *helelilla* stimulus RV-T3 being added to the *lilla* 'purple' sub-cluster. The stimuli other than stimulus RV-T3 (F=8) are all higher frequency (F≥11). The next cluster in Figure 12 is the *sinine* 'blue' cluster with nine stimuli with the same most frequent name and high frequency in relation to the naming task, again F≥11. The seven stimuli for *helesinine* 'light blue' have an even higher lowest frequency (F≥12) than those for the *sinine* 'blue' cluster. Next to the *helesinine* 'light blue' cluster is a sub-cluster of darker blues, including named most frequent *tumesinine* 'dark blue' for Bw, B-S1, B-S2, C-S2, C-S3; *rohekas-sinine* 'greenish blue' for C-S1; and *mustajassinine* 'blackish blue' for B-DS.

The last cluster is the largest with 16 stimuli containing different lighter shades in several different sub-clusters. The largest cluster includes named most frequent for *helesinine* 'light blue' for C-P2-2 (F=17), C-T2 (10), Bc-T2 (6), BG-P2-3 (4), and BG-T4 (9); B-T2 (9) for *sinine* 'blue'; B-P2-2 for *hall* 'grey' (8); Bw-T3 (5) and Bw-T4 (10) for *helelilla* 'light purple'; BG-S1 for *roheline* 'green' (7); BG (6) and BG-EX (7) for *sinakasroheline* 'bluish green'; BG-LT (8) and G-T3 (12) for *heleroheline* 'light green'; BG-T2 (7) for *rohekassinine* 'bluish green'; and stimulus BG-P1-2 which had a frequency of five for both *heleroheline* 'light green' and *helesinine* 'light blue'.

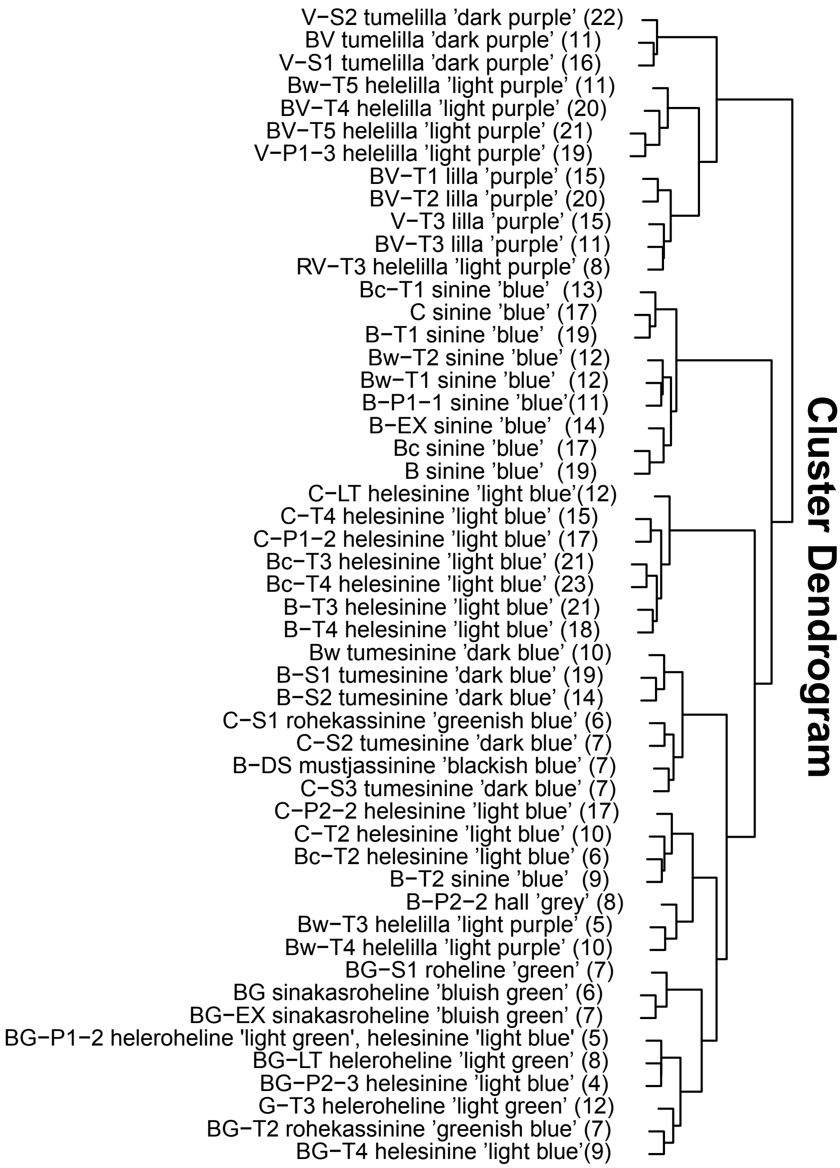

Figure 11. Estonian naming task hierarchical cluster analysis.

It is quite hard to pinpoint the factors which elevated the *hele* ‘light’ and *tume* ‘dark’ modified terms to dominance in the naming task. Whether it was the preceding sorting task, the selection of stimuli, the unconstrained nature of the tasks, or some other unnamed factor, the terms that gained dominance in the Estonian naming task were *helelilla* ‘light purple’ (SI=0.513), *tumelilla* ‘dark purple’ (SI=0.431), *kollane* ‘yellow’ (SI=0.365), *helesinine* ‘light blue’ (SI=0.308), *tumesinine* ‘dark blue’ (SI=0.238), *sinine* ‘blue’ (SI=0.211), and *lilla* ‘purple’ (SI=0.187). The dominant stimuli name threshold was lowered to

49%, from 20 to 19, and this meant *kollane* 'yellow', *tumesinine* 'dark blue' and *sinine* 'blue' gained literally borderline dominance. *Kollane* 'yellow' and *sinine* 'blue' are formerly established basic colour terms.

It is debatable whether the dominance of both light and dark modified purples, *helelilla* 'light purple' and *tumelilla* 'dark purple', and blues, *helesinine* 'light blue' and *tumesinine* 'dark blue', alongside the basic colour terms *lilla* 'purple' and *sinine* 'blue' diminishes the dominance gained by *helesinine* 'light blue' and *tumesinine* 'dark blue'.

2.2.4 Conclusion of Estonian results

While the list task results confirm the eleven previously established basic colour terms, the Estonian list task plot in Figure 5 illustrates the in-between nature of *beež* 'beige' and *helesinine* 'light blue'. These two terms are located between the established basic colour terms and the rest of the data over on the left-hand side of the plot. There are many indicators of basicness in the list task or in literary, even corpus, data (see Biggam 2012), but the plot gives a concise overview of the list task at a glance. The pictorial overview of the data by the mean position and frequency is supported by the cognitive salience index, which uses these same parameters together with the number of participants.

Although it is a modified term and thus a non-basic term by the original definition, the frequency, mean position and salience of *helesinine* 'light blue' place it, together with *beež* 'beige', between the basic and non-basic colour terms, indicating a relative basicness in the list task.

The usage percentages in the three tasks were above half for the list and naming tasks, but below half in the sorting task, even for *sinine* 'blue' at 44% (see Table 9). In the list task, all three blues had a larger percentage of participants using the term, with *helesinine* at 59%, *sinine* 'blue' at 95%, and *tumesinine* 'dark blue' at 38%, but the percentages from the sorting task participants were far lower at 15%, 44% and 18% respectively. The naming task percentages, however, were similar to those of the list task with all three registering more than half at 85% for *helesinine* and *sinine* 'blue', and 67% for *tumesinine* 'dark blue'.

A glance at Table 9 shows that the sorting task consensus numbers are drastically lower than those for the list and naming tasks. One factor causing the low consensus in the sorting task may be that the participants did not know that the sorting of stimuli would be followed by a request to name the groups. Therefore, they had no special reason to group together stimuli that could easily be encompassed under a single name. This meant there was no reason for the groups to be aligned with the boundaries of colour categories. Naming the groups after the grouping of stimuli was completed in the sorting task was a challenging task that forced both lumpers with two or three heterogeneous groups and splitters with 21 groups to coin ad-hoc non-basic descriptions for them, which unsurprisingly generated a wide diversity of names.

Table 9. Overview of Estonian *helesinine*, *sinine* and *tumesinine*
F – frequency (%), mp – mean position, S – cognitive salience index, N – no of participants using the terms (%), StR – number of stimuli named with a term.

Term	**List task (N=39)** **F (%) – mp – S**	**Sorting task (N=39)** **N (%) – F (F/N) – StR**	**Naming task (N=39)** **N (%) – F (F/N) – StR**
helesinine	23 (**59%**) – 13.8 – 0.0428	6 (15%) – 57 (9.5) – 19	33 (**85**%) – 211 (6.4) – 27
sinine	37 (**95**%) – 5.8 – 0.1625	17 (44%) – 267 (15.7) – 44	33 (**85**%) – 180 (5.5) – 24
tumesinine	15 (38%) – 16.9 – 0.0227	7 (18%) – 35 (5.0) – 15	26 (**67**%) – 80 (3.1) – 13

In particular, a high frequency of modified compounds was observed, which can be considered a distinguishing feature of Estonian colour vocabulary (Sutrop 2002: 72–73) as any hue, shadow, tint, intensity or darkness-lightness can be denoted with one compound colour word in Estonian (Õim 1983: 27). For participants who were splitters in the free-sorting task, it was inevitable that they would use the names of the general form ‘light blue’ and ‘dark blue’ to label the groups they formed, whether or not these terms were particularly salient. Even so, few participants used *helesinine* or *tumesinine*, with only six participants for *helesinine* and seven for *tumesinine*. This does not support them as categories. The hierarchical clustering and multidimensional scaling analysis solutions do not provide evidence for subcategories either. They indicate that participants focused on lightness as a criterion for sorting, but delimited their groups at different lightness thresholds, so that the empirical colour map from the MDS possessed a lightness dimension but was not polarised into concentrated groups of light or dark blue.

Various factors combined to make naming task accentuate the use of the modified terms *tumesinine* and especially *helesinine*, and the number of dominant stimuli, making it harder to interpret the results. The tendency towards specification may have been increased by priming from the preceding free-sorting task. The appearance of two unequivocally non-basic terms, *helelilla* ‘light purple’ and *tumelilla* ‘dark purple’, gaining dominance means that this is not convincing evidence for the formation of a ‘light blue’ category in Estonian. When the same tasks and stimuli were applied to Italian (Uusküla 2014), evidence emerged for several BLUE categories in *blu, azzurro* and *celeste.*, but the consensus was also low in their study. Bimler and Uusküla (2016) have used the author’s Estonian data with my permission in their article along with data for five other languages Russian, Italian, English, Lithuanian and Udmurt. Using the same Estonian data, they created an index that quantified the separation of light and dark blue and the strength of the category boundary in six languages. They found that the index was the lowest for English and Estonian, while values were high for Lithuanian, Russian and Italian.

In the sorting task the number of participants who named a group *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’ was so low that it indicates that only one BLUE basic colour term exists in Estonian: *sinine* ‘blue’.

2.3 Introductory comments on Russian colour terms

The first introductory paragraphs on Russian colour terms start with some notes on transliteration and translation with examples, before highlighting some of the most particular colour names from the list task. The comments end by noting some of the characteristics of the data.

Russian uses the Cyrillic alphabet. The terms used in the data collected for this study are all transliterated from the Cyrillic alphabet to Latin, but the preferred transliteration varies between authors. Thus the transliteration in previous works (Davies and Corbett 1994; Moss *et al.* 1990; Laws, Davies and Andrews 1995; Morgan 1993; Andrews 1994; Corbett and Morgan 1988) may differ from the transliteration provided here. Here ISO 9:1995 transliteration[13] was used. As the Cyrillic alphabet may be unfamiliar to the eye of a reader who is used to the Latin alphabet, so in reverse the Latin transliteration looks unseemly to a Russian reader. It should also be noted that the letters used in the transliteration here do not correspond to their pronunciation in English, and so the reader can only hope to pronounce the Russian words thus transliterated if they already have a good knowledge of Cyrillic and of Russian.

Once the terms have been transliterated, the next step is to gloss them. As previously seen in introductory sections to the Turkish and Estonian terms, the translation preferences of authors may also vary. One easily understandable issue is the use of Russian adjectives, which can encompass not only a specific colour term, but also add an emphasis that makes the term into a colour term. In English translation sometimes 'colour' is added, e.g. *bordovyj* 'bordeaux [red]', *limonnyj* 'lemon[-coloured]', *kirpičnyj* 'brick[-coloured]', *buryj* 'brown red', and *kofejnyj* 'coffee[-coloured]'. Some terms are translated as 'explanations', e.g. *buryj* 'brown red', *ryžij* 'ginger', *palevyj* 'pale yellow', or *marengo* 'greyish black'. There is also a relatively large register of purples, such as *purpurnyj* 'purple', *sirenevyj* 'lilac', *lilovyj* 'mauve', and even *višnëvyj* 'cherry', which resembles the Turkish *visne* 'sour cherry, morello cherry' and the Estonian dialect *visnapuu*, and *malinovyj* 'rasberry'. *Alyj* 'scarlet' is closer to the bright red part of the spectrum.

The next terms are grouped together here, which may make them appear more relevant, but there is a small number of colour terms that could be related to coffee or to other drinks, *kremovyj* 'creamy', *kofejnyj* 'coffee[-coloured]' and *moločnyj* 'milky', and some terms related to semi-precious stones and metals, *serebrënyj* 'silver-plated', *serebristyj* 'silvery', *bronzovyj* 'bronzy', *mednyj* 'coppery', and *izumrudnyj* 'emerald' and *ântarnyj* 'amber'. Notable not only for its frequency, but also for its construction, is *cvet morskoj volny* 'sea wave colour', which is remarkably persistent through all three tasks. Some of these examples also reflect the more vernacular nature of the terms, which makes these

---

[13] For transliteration an online resource was used ('Translit RU/EN - Russian Transliteration and Spell Checker').

data somewhat different from previous research, which rarely displayed the vernacular terms. The study specifically targets Estonian Russian-speakers (N=30), a linguistic minority, and the results may differ from those in the literature.

The most notable difference in borderline basic terms is *birûzovyj* 'turquoise', which is remarkably frequent, like in previous research, especially in the list task. The frequent use of the non-basic *birûzovyj* 'turquoise' has been noted earlier too (Davies and Corbett 1994: 81). Using the Russian National Corpus, Rakhilina and Paramei (2011: 125) illustrate the expansions of the taxonomic constraints of *birûzovyj* 'turquoise':

> "According to the RNC, the term *birjuzovyj* was already used in Russian in the mid-eighteenth century but originally meant only "made of turquoise" (in relation to "stone", "ring", "necklace", etc). The mid-nineteenth century registered the usage of *birjuzovyj* predominantly in the sense of "turquoise-coloured", as related to the colour of a "collar" or "skirt" and, rarely, of "sky" or "sea", By the beginning of the twentieth century not only had the frequency of *birjuzovyj* usage increased significantly (by 50%) but also its combinability expanded – from nouns for artefacts (e.g. "fabric", "carpet") to those for "eyes" and natural surfaces (e.g. "water", "sky")." (Rakhilina and Paramei 2011: 128)

Using a term for eyes, water and sky indicates a trend toward basicness (see Kerttula 2007, for the concept of relative basicness) in Russian, but does not apply for the Estonian *türkiis* 'turquoise'. *Birûza* was not listed in the list task.

Another expected result from the previous research is that there will be a lack of modifiers, especially in comparison to the Estonian data. The most important difference expected in Russian is the prevalence of *goluboj* 'light blue', which is predicted to have high visibility across all three tasks, manifesting as a separate blue category from *sinij*[14] 'blue' in the sorting task.

> "An additional ANOVA analysis confirmed a highly significant category boundary effect within the blue area in the Russian group. The findings from the cross-cultural studies indicate that the color grouping by Russian observers reveals a small-scale language modulation because of the availability of the additional blue color category and term." (Paramei 2005: 22–23)

For a visual view of the Russian blues in the sorting task see Paramei (2005: 16) She used the work by Frumkina (1984: 59) to form a structure (Paramei's Figure 1) of the Russian blue colour terms derived from the free-sorting task. It illustrates the structure of Russian blue terms found from previous research into Russian sorting task. It is in Russian, but fortunately the glosses are added. It indicates the separate clusters for *sinij* and *goluboj*. (Paramei 2005: 16)

---

[14] The non-basic Arabic term *ṣīnī* 'Chinese' eventually became a term for 'blue' and yielded a series of cognate colour terms in several Near Eastern languages, the similarity of which to the Russian *sinij* 'dark blue' Borg (2007: 278) considers intriguing.

To summarise, *goluboj* 'light blue' is expected to present itself in the Russian data as a separate category, while the use of modifiers is expected to be less frequent than in the Estonian data, and perhaps even than in the Turkish data, and arguably some differences may appear from previous Russian research because this data targets Estonian Russian-speakers.

2.3.1 Russian list task

The structure of this section is like that of the Estonian and Turkish list task introductions. The introduction starts with a short numerical overview and a short list of the terms that made up the bulk of the data. The list task analysis contains two parts, a graphical overview of the frequency and mean position (see Figure 12) and a table overview (Table 10) of the same data with the terms listed and glossed, and with Sutrop's cognitive salience index included. These provide more detailed information on the results. The end of the section compares the salience with previous research and concludes with a discussion of goluboj 'light blue'.

The participants (N=30) listed 294 colour names 917[15] times, and 108 terms were listed more than once for a total of 731 times. In 186 cases a term was listed by only one participant, making 20% of the total number of terms listed or 63% of the different terms. There were 75 terms that were elicited three or more times, 54 terms four or more times, 40 five or more times, and 20 terms ten or more times.

This meant that 25 terms were listed by at least a quarter of the participants (≥ 8), fifteen terms by half of them (≥ 15), and ten terms by three-quarters (≥ 23). Technically the basic term *koričnevyj* 'brown' (F=22, 73%) is on the borderline for the three quarters criterion, while the basic colour term *rozovyj* 'pink' (F=19) fell below the criterion of three quarters. The number of terms listed by a quarter of the participants is relatively large at 25, but this is perhaps due to the smaller sample size (N=30). Although 12 terms form 34% of the data with *goluboj* 'light blue' providing 3%, this percentage would be even higher if only simplified, non-modified data were analysed.

[15] The original count was 918, but a control table showed that participant no 6 listed *oranževyj* 'orange' twice in fifth, and eleventh place. The second listing was removed to avoid duplication.

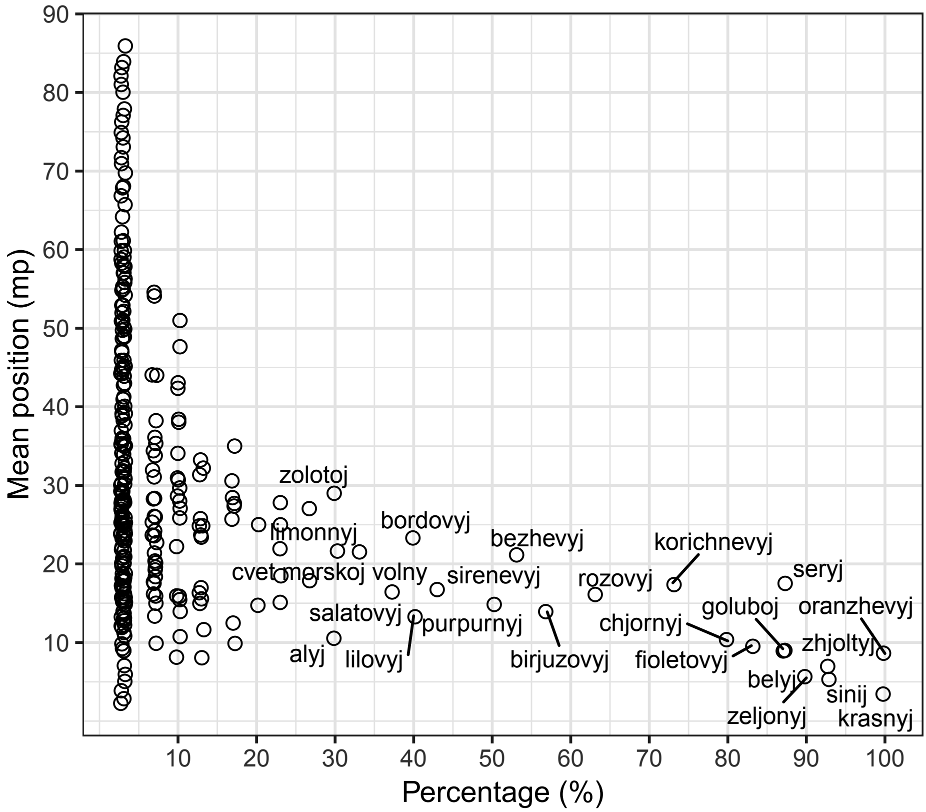

Figure 12. Russian list task terms by percentage (≥ 30% labelled) and mean position.

However, in order to present the most frequent Russian list task terms here are the 25 terms that were listed by a quarter of the participants: *krasnyj* ‘red’, *oranževyj* ‘orange’ (F=30, 100%), *sinij* ‘blue’, *žëltyj* ‘yellow’ (93%), *zelënyj* ‘green’ (90%), *belyj* ‘white’, *goluboj* ‘light blue’, *seryj* ‘grey’, (87%), *fioletovyj* ‘purple’ (83%), *čërnyj* ‘black’ (80%), *koričnevyj* ‘brown’ (73%), *rozovyj* ‘pink’ (63%), *birûzovyj* ‘turquoise’ (57%), *beževyj* ‘beige’ (53%), *purpurnyj* ‘purple’ (50%), *sirenevyj* ‘lilac’ (43%), *bordovyj* ‘bordeaux [red]’, *lilovyj* ‘mauve’ (40%), *salatovyj* ‘lettuce’ (37%), *cvet morskoj volny* ‘sea wave colour’ (33%), *alyj* ‘scarlet’, *limonnyj* ‘lemon[-coloured], *zolotoj* ‘golden’ (30%)[16], *višnëvyj* ‘cherry’, and *malinovyj* ‘raspberry’ (27%). Out of 917 times that words were listed, these listed terms accounted for 50% of the Russian list data by number.

Each data point in Figure 12 represents a colour word from Russian. Like in the Turkish list task figure, Figure 3, and the Estonian list task, Figure 5, the data points are staggered with the R function *ggrepel* (Slowikowski 2017) to avoid overlapping and to show the data points for all the different terms. This shows how the frequent colour words tend to have lower mean positions and that many colour words were listed infrequently.

16 These terms are labelled in Figure 12.

In Figure 12 *krasnyj* 'red' and surprisingly *oranževyj* 'orange' (F=30, 100%) both had the highest possible listing percentage, but *krasnyj* 'red' with an mp of 3.7 comes in a lower position in the right-most corner because it has a lower mean position than *oranževyj* 'orange' (mp = 8.6). These first two are followed by *sinij* 'blue' (mp = 5.2), *žëltyj* 'yellow' (93%, mp = 7.1) and *zelënyj* 'green' (90%, mp = 5.5). The data points for *belyj* 'white' (mp = 8.7) and *goluboj* 'light blue' (mp = 9.2, 87%) are extremely close as there is only a slight difference in their mean positions, with *seryj* 'grey' (mp = 17.3) above them as it has a higher mean position. These are trailed by *fioletovyj* 'purple' *čërnyj* 'black', and *koričnevyj* 'brown', which finish the high frequency grouping of colour terms. The high frequency group includes *goluboj* 'light blue' (87%, mp = 9.2), but not *rozovyj* 'pink' (63%, mp = 15.9). Particularly significant is the situation of *goluboj* 'light blue', which is firmly situated in the centre of the basic terms and so by frequency and mean position is nowhere near the borderline.

It should be noted that the separation between the high frequency grouping in the Russian list task and the other terms is located at a higher listing percentage value of around 70% along the x-axis in Figure 12 than it is in the corresponding Turkish and Estonian data, where the separation is nearer to 50%. In comparison to the list task plots for Turkish (see Figure 3) and Estonian (see Figure 5), which have a larger divide between the high and low frequency colour terms, the Russian list task plot (see Figure 12) has more in-between terms. The lower frequency grouping is spearheaded by the basic colour term *rozovyj* 'pink' (63%), while other terms (≥30%) that are more stretched are *birûzovyj* 'turquoise' (57%), *beževyj* 'beige' (53%), *purpurnyj* 'purple' (50%), *sirenevyj* 'lilac' (43%), *lilovyj* 'mauve', *bordovyj* 'bordeaux [red]' (40%), *salatovyj* 'lettuce' (37%), *cvet morskoj volny* 'sea wave colour' (33%), *alyj* 'scarlet', and *limonnyj* 'lemon[-coloured]' (30%). The number of terms labelled is higher than in the Turkish or Estonian data.

The twenty most salient Russian colour terms ranked by Sutrop's cognitive salience index (see Table 10) are *krasnyj* 'red' (S = 0.2703), *sinij* 'blue' (S = 0.1790), *zelënyj* 'green' (S = 0.1642), *žëltyj* 'yellow' (S = 0.1313), *oranževyj* 'orange' (S = 0.1163), *belyj* 'white' (S = 0.1001), *goluboj* 'light blue' (S = 0.0939), *fioletovyj* 'purple' (S = 0.0864), *čërnyj* 'black' (S = 0.0784), *seryj* 'grey' (S = 0.0500), *koričnevyj* 'brown' (S = 0.0417), *birûzovyj* 'turquoise' (S = 0.0406), *rozovyj* 'pink' (S = 0.0397), *purpurnyj* 'purple' (S = 0.0335), *lilovyj* 'mauve' (S = 0.0298), *alyj* 'scarlet' (S = 0.0278), *sirenevyj* 'lilac' (S = 0.0257), *beževyj* 'beige' (S = 0.0254), *salatovyj* 'lettuce' (S = 0.0227), and *bordovyj* 'bordeaux [red]'[17] (S = 0.0173).

---

[17] *Bordo* 'bordeaux' (F = 4) and *bordovyj* 'bordeaux [red]' (F=12) are counted as separate terms. The former is a cognitive salience index outlier, because it has low frequency *bordo* 'bordeaux' (S = 0.0162) (F=4) but appears at positions 3, 4, 12 and 14, thus showing that Sutrop's cognitive salience index can be less reliable for low-frequency terms. The Estonian *kuldne* 'golden' is in a similar position.

Table 10. Russian list task terms by frequency (≥4), mean position and salience
F – frequency, % – listing percentage, *mp* – mean position, *Salience* – cognitive salience index.

Russian list task (N=30)	**Glosses**	**F**	**Rank**	**%**	**mp**	**Rank**	**Salience**	**Rank**
krasnyj	red	30	1.5	100	3.7	1	0.2703	1
oranževyj	orange	30	1.5	100	8.6	7	0.1163	5
sinij	blue	28	3.5	93	5.2	2	0.1790	2
žëltyj	yellow	28	3.5	93	7.1	4	0.1313	4
zelënyj	green	27	5	90	5.5	3	0.1642	3
belyj	white	26	7	87	8.7	8	0.1001	6
goluboj	'light blue'	26	7	87	9.2	9	0.0939	7
seryj	grey	26	7	87	17.3	38	0.0500	10
fioletovyj	purple	25	9	83	9.6	10	0.0864	8
čërnyj	black	24	10	80	10.2	12.5	0.0784	9
koričnevyj	brown	22	11	73	17.6	40	0.0417	11
rozovyj	pink	19	12	63	15.9	30	0.0397	13
birûzovyj	turquoise	17	13	57	13.9	20	0.0406	12
beževyj	beige	16	14	53	21	49	0.0254	18
purpurnyj	purple	15	15	50	14.9	23	0.0335	14
sirenevyj	lilac	13	16	43	16.8	36	0.0257	17
lilovyj	mauve	12	17.5	40	13.4	18	0.0298	15
bordovyj	bordeaux	12	17.5	40	23.2	56	0.0173	20
salatovyj	lettuce	11	19	37	16.2	34.5	0.0227	19
cvet morskoj volny	sea wave colour	10	20	33	21.4	50	0.0156	23
alyj	scarlet	9	22	30	10.8	15	0.0278	16
limonnyj	lemon coloured	9	22	30	21.7	52	0.0138	26
zolotoj	golden	9	22	30	29.1	82	0.0103	33
višnëvyj	cherry	8	24.5	27	17.8	41	0.0150	25
malinovyj	raspberry	8	24.5	27	26.9	72	0.0099	34
ohra	ochre	7	28	23	15.4	27	0.0151	24
ul'tramarin	ultramarine	7	28	23	18.3	43	0.0128	29
serebrânyj	silvery	7	28	23	22.1	53	0.0105	32
haki	khaki	7	28	23	25	64	0.0093	36
buryj	'brown red'	7	28	23	28	76.5	0.0083	39
tëmno-zelënyj	dark green	6	31.5	20	14.7	22	0.0136	27
izumrudnyj	emerald	6	31.5	20	25	64	0.0080	41

Russian list task (N=30)	Glosses	F	Rank	%	mp	Rank	Salience	Rank
svetlo-zelёnyj	light green	5	36.5	17	10.2	12.5	0.0163	21
tёmno-sinij	dark blue	5	36.5	17	12.4	17	0.0134	28
ryžij	ginger	5	36.5	17	25.8	68.5	0.0065	45.5
serebristyj	silver-plated	5	36.5	17	27.4	74	0.0061	49
kirpičnyj	brick coloured	5	36.5	17	27.6	75	0.0060	50
zolotistyj	golden	5	36.5	17	28.6	80	0.0058	51
kremovyj	creamy	5	36.5	17	30.8	85	0.0054	54.5
kofejnyj	coffee	5	36.5	17	35.2	95	0.0047	59
bordo	bordeaux	4	47.5	13	8.2	6	0.0162	22
salatnyj	lettuce	4	47.5	13	11.5	16	0.0116	31
bolotnyj	marsh	4	47.5	13	15.2	25	0.0087	37
lazurnyj	azure	4	47.5	13	15.5	28	0.0086	38
svetlo-žёltyj	light yellow	4	47.5	13	16.2	34.5	0.0082	40
indigo	indigo	4	47.5	13	17	37	0.0078	42
travânoj	grass	4	47.5	13	23.5	58	0.0057	52
slivovyj	plum	4	47.5	13	23.8	60	0.0056	53
svetlo-koričnevyj	light brown	4	47.5	13	24.8	62	0.0054	54.5
ântarnyj	amber	4	47.5	13	25	64	0.0053	56
tёmno-krasnyj	dark red	4	47.5	13	25.8	68.5	0.0052	57
bronzovyj	bronzy	4	47.5	13	31.5	88	0.0042	63
molochnyj	milky	4	47.5	13	32.5	90	0.0041	65
mednyj	coppery	4	47.5	13	33.2	91	0.0040	66

In this study the listing percentage of *goluboj* was 87%, but in previous studies it was reportedly higher at 93.5% (Morgan and Corbett 1989) or 94.8% (Davies and Corbett 1994; 51.3% in the child sample). The frequency rank was higher in previous research, with rank = 3 in Morgan and Corbett (1989), rank = 4 in Morgan (1993), and rank = 4.5 in Davies and Corbett (1994). In Corbett and Davies (1988, Table 7, *goluboj* rank = 6) and the Davies *et al.* (1998) child sample the rank of *goluboj* 'light blue'(rank = 7) was close to that in this sample. Mean positions were also reported lower, with mp = 6.0 in the Davies *et al.* (1998) child sample and mp = 7.5 in Davies and Corbett (1994).

Table 10 does contain *svetlo* 'light' and *tёmno* or *temno* 'dark' modifiers in *tёmno-zelёnyj* 'dark green' (20%), *svetlo-zelёnyj* 'light green', *tёmno-sinij* 'dark blue' (17%), *svetlo-žёltyj* 'light yellow', *svetlo-koričnevyj* 'light brown', and *tёmno-krasnyj* 'dark red' (13%), though the frequency is relatively low.

As indicated by previous studies of Russian (Corbett and Morgan 1988; Davies and Corbett 1994), the Russian list task frequency from Figure 12 and Table 10 confirms twelve Russian basic colour terms: *krasnyj* 'red', *oranževyj* 'orange', *sinij* 'blue', *žëltyj* 'yellow', *zelënyj* 'green', *belyj* 'white', *goluboj* 'light blue', *seryj* 'grey', *fioletovyj* 'purple', *čërnyj* 'black', *koričnevyj* 'brown' and *rozovyj* 'pink'. Three non-basic colour terms had a listing percentage of over or equal to half, and these were *birûzovyj* 'turquoise' (57%), *beževyj* 'beige'(53%), and *purpurnyj* 'purple' (50%). *Beige* has no suitable stimulus in the sorting and naming tasks, but *birûzovyj* 'turquoise' does factor in in the following tasks.

2.3.2 The Russian sorting task

The sorting task has several more aspects which may have influenced the results. This makes it important when analysing the results to pinpoint which aspect affected the results the most. In some cases, an educated guess is needed as to which aspects of the sorting task matter most, whether stimuli selection, non-restricting instructions to the participants, non-verbal grouping before naming the groups, or something else.

The instructions for the sorting task make no mention of colour, and the participants are simply instructed to group the stimuli by similarity. The implicit expectation is that the participant will sort the stimuli by colour, but it is not explicitly stated that the sorting criterion for the stimuli is colour. After the list task, which primes the participants colour vocabulary, it is quite unlikely that the participant would not comply with the implicit instruction to sort by colour, but the decision on the criteria for sorting is left to the participant, whether to use the darkness or lightness of the shade for example, the feeling evoked by the stimuli, the name of the colour category that takes in the most stimuli, or another angle.

To orient the reader in the sorting task, a simple plot of the stimuli labelled with their most frequent group names was constructed. Since the Estonian and Russian data were collected using the same methods and stimuli, it is suggested that Figure 6 and Figure 13 be viewed side by side, as they depict the stimuli in the CIELAB colour space labelled with their most frequent names. All the figures for the Estonian and Russian plots are comparable in this way.

Figure 13 depicts the non-yellow stimuli in CIELAB colour space labelled with the most frequent group names. The most frequent stimuli names are *sinij* 'blue', *žëltyj* 'yellow' (not depicted), *fioletovyj* 'purple' and *goluboj* 'light blue'. Although this is a simple figure of the most frequent group names, comparison to the Estonian (see Figure 6), where there was only one stimulus where the group name *helesinine* 'light blue' was dominant, shows that the corresponding group name in Russian *goluboj* 'light blue' is in a category of its own.

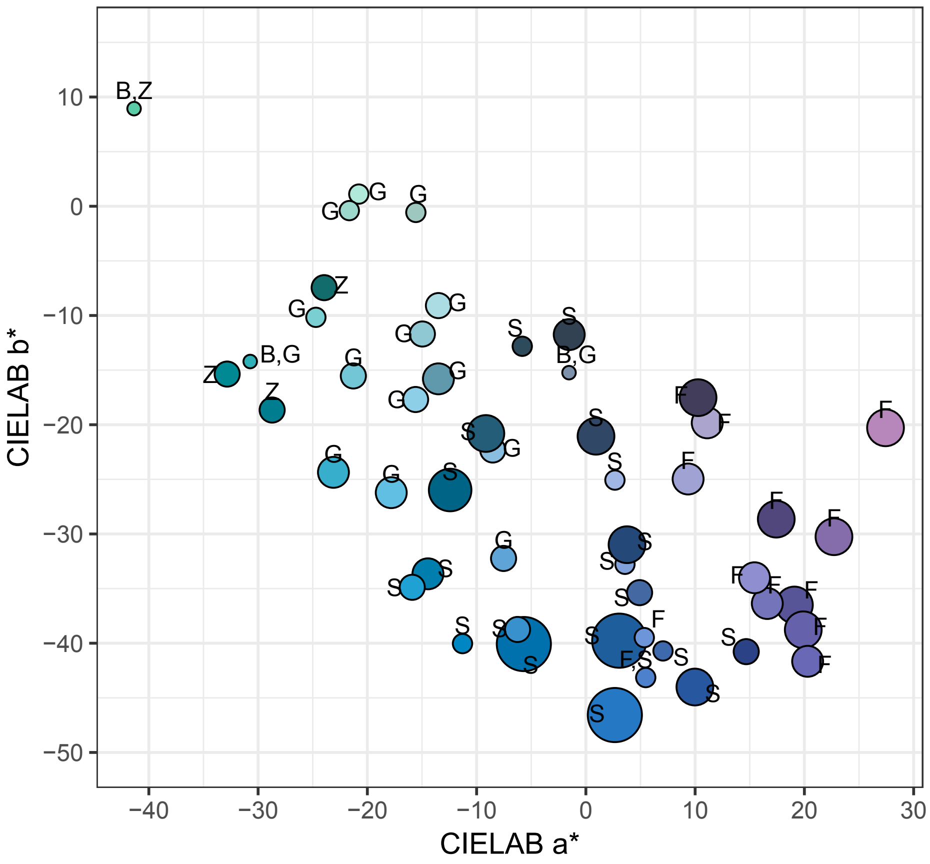

Figure 13. Russian sorting task by most frequent stimulus name.

Most frequent stimuli names labelled: F – *fioletovyj* ‘purple’, G – *goluboj* ‘light blue’, S – *sinij* ‘blue’, Z – *zelënyj* ‘green’.

In Figure 13 the size of the dots, which represent the stimuli, is correlated with the frequency of naming, so the larger the dot, the higher the frequency. In Figure 13 ‘G’ marks the stimuli that were most frequently named *goluboj* ‘light blue’. Here is where a comparison with the corresponding Estonian plot, Figure 6, is invaluable. In the Estonian data (see Figure 6), *helesinine* ‘light blue’ is a not a frequent stimulus name or named most frequent, whereas in the Russian data (see Figure 13), the corresponding *goluboj* ‘light blue’ is much more visible.

Table 11 presents the data from the Russian sorting task in table form. The most prominent aspect in Table 11 is the most frequent names given to groups by participants. *Sinij* ‘blue’ was given as a name to a group of stimuli by the highest number of participants (10) and with the highest frequency (F=124), and it covered the largest range of stimuli (39) and had the largest average group size (12.4). In second place, nine participants named a group *žëltyj* ‘yellow’, but the range of stimuli and the average group size were very low (4), which demonstrate the unintended distinction of the distractor stimuli in the sorting task. Relatively large numbers of participants also named groups *fioletovyj* 'purple' (N=8) and *goluboj* 'light blue' (N=6). While their range and average group size are similar, *fioletovyj* 'violet' has a higher frequency (F=80) than *goluboj* 'light blue' (F=68). However, in comparison to the frequencies of the following group names *birûzovyj* 'turquoise' (F=45), *zelënyj* 'green' (F=24), and *temno-sinij* 'dark blue' (F=18), which were each given by four participants, the frequency of *goluboj* 'light blue' is high.

Table 11. Russian sorting task by no of participants (N > 1)

Term (in Singular)	**Gloss**	**No. of participants**	**Frequency**	**No. of tiles (range)**	**Average group size**
sinij	blue	10	124	39	12.4
žëltyj	yellow	9	36	4	4
fioletovyj	purple	8	80	22	10
goluboj	'light blue'	6	68	20	11.3
birûzovyj	turquoise	4	45	32	11.3
zelënyj	green	4	24	13	6
temno-sinij	dark blue	4	18	16	4.5
temnyj	dark	2	24	22	12
sirenevyj tuman	lilac smoke	2	20	16	10
lilovyj	lilac	2	18	17	9
grâzno-sinij	dirty blue	2	8	6	4
svetlo-sinij	light blue	2	12	10	6

Due to the low consensus, all the stimuli names given by more than one participant are represented in Table 11. This low consensus in the sorting task might be better called 'high variability' or 'broad inter-participant differences' in naming the sorted stimuli groups, and it is also reflected in the subsequent multidimensional scaling analysis plot. Ideally the multidimensional scaling plot should reveal separate groups. The following multidimensional scaling plot of the Russian sorting task shows the grouping of the stimuli in Figure 14.

It could be speculated that the four yellowish distractor stimuli exhibit an extremely strong isolating effect. It is possible that the ratio of experimental stimuli, with 51 non-yellow tiles, to distractors, with 4 yellow stimuli, is too small. They are so far separated from the rest of the stimuli that all the non-yellow stimuli form an elongated horseshoe-shaped even spread of stimuli, rather than distinct groups. Most importantly for the comparison, the Russian *goluboj* 'light blue' has a more defined partition in the MDS spread in Figure 14 with a slight separation visible from *sinij* 'blue' and from *zelënyj* 'green'.

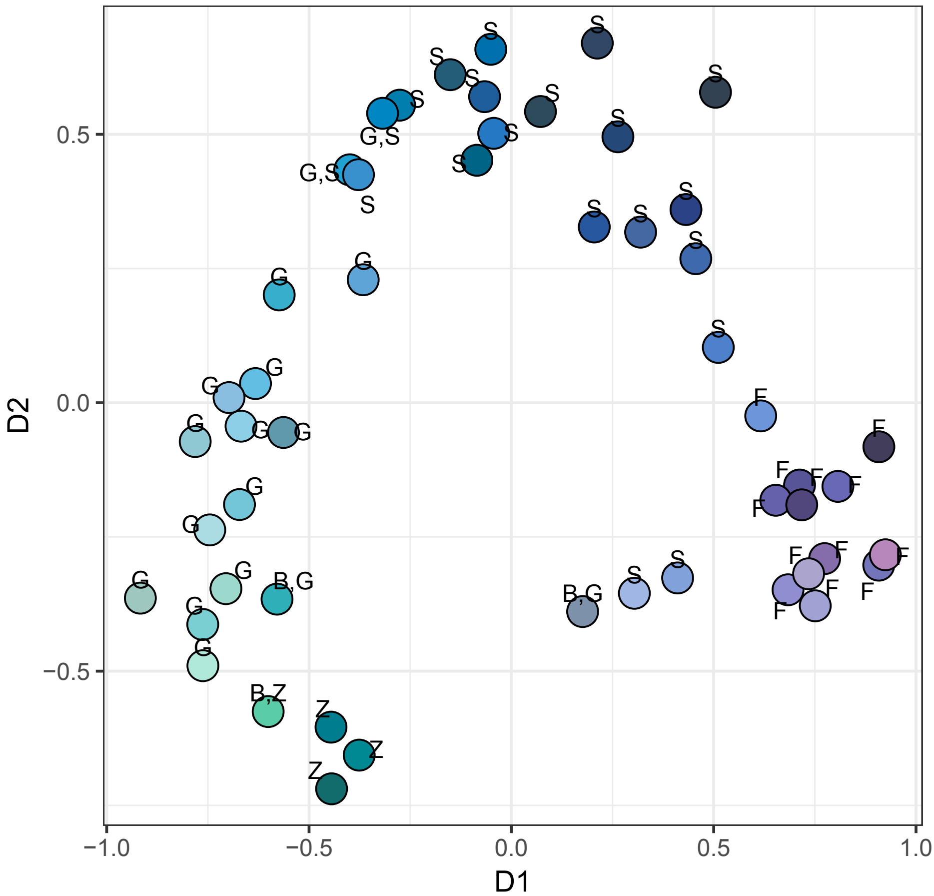

Figure 14. Russian sorting task multidimensional analysis.
Most frequent stimuli names labelled: F – *fioletovyj* ‘purple’, G – *goluboj* ‘light blue’, S – *sinij* ‘blue’, Z – *zelënyj* ‘green’.

The output of the principal components analysis suggests that the first component counts for 44.3% of the variance, the first two cumulatively for 70.1%, the first three for 79.5%, the first four for 84.3, the first five for 87.8%, and the first six for 90.4% of the variance. The scree plot has a very steep decline after the first two components, and starts to level out at the third component.

In the hierarchical cluster analysis of the Russian sorting task (Figure 15) there are three big clusters, one with mostly *fioletovyj* ‘purple’ most frequent names, one with almost all *sinij* ‘blue’, and one containing many stimuli that were most frequently sorted into the *goluboj* ‘light blue’ group. The hierarchical cluster analysis plot (see Figure 15) reveals that the four yellowish distractor stimuli may not be the only ones to exhibit isolating tendencies.

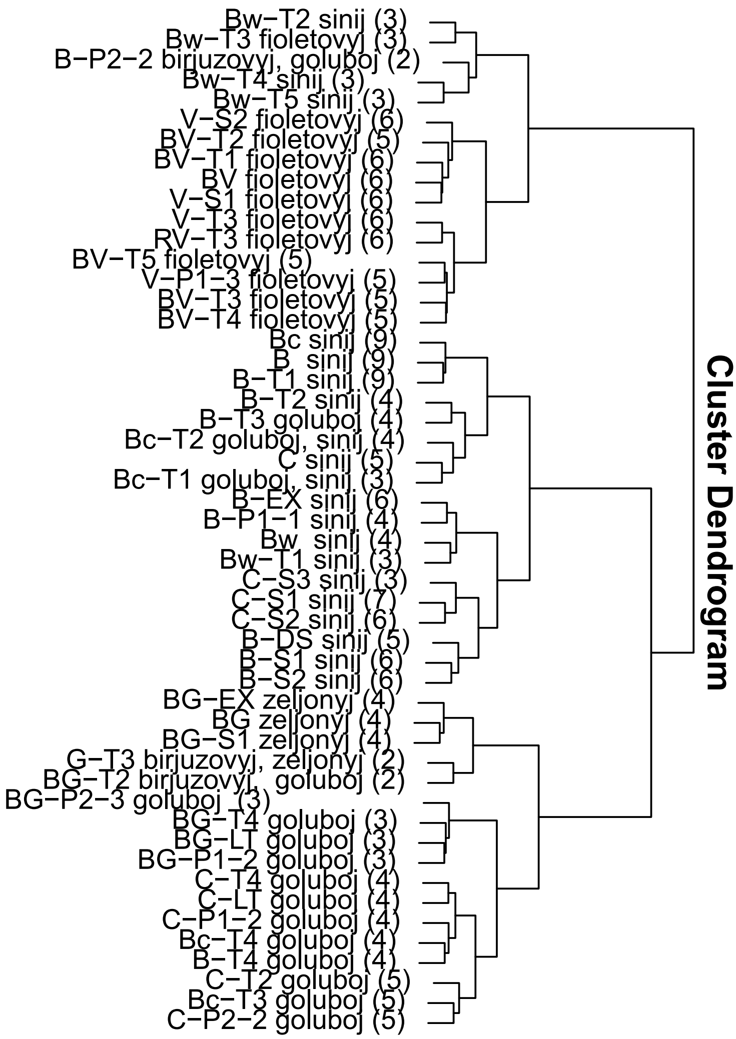

Figure 15. Russian sorting task hierarchical cluster analysis.

That *goluboj* 'light blue' forms a sub-cluster in Figure 15, in which all the stimuli are sorted and named under the term named most frequent (nmf), is significant in comparison to the Estonian sorting task, where only one stimulus had a *helesinine* most frequent name, rendering the possibility of a *helesinine* cluster or category rather unlikely.

2.3.3 The Russian naming task

The significant lack of agreement in naming the stimuli in the Russian naming task presented a problem with the dominance threshold, which will be discussed after Figure 16. Figure 16 presents the most frequent names given to the stimuli and is not dependent on the level of agreement between the participants on naming the stimuli. The figure for the Russian naming task *named most frequent* (*nmf*), Figure 16, is comparable to the corresponding Estonian naming task *nmf* figure. The description of the Russian naming task results continues with multidimensional scaling analysis (Figure 17), hierarchical cluster analysis (Figure 18), and an overview of Russian 'blue' terms (Table 13).

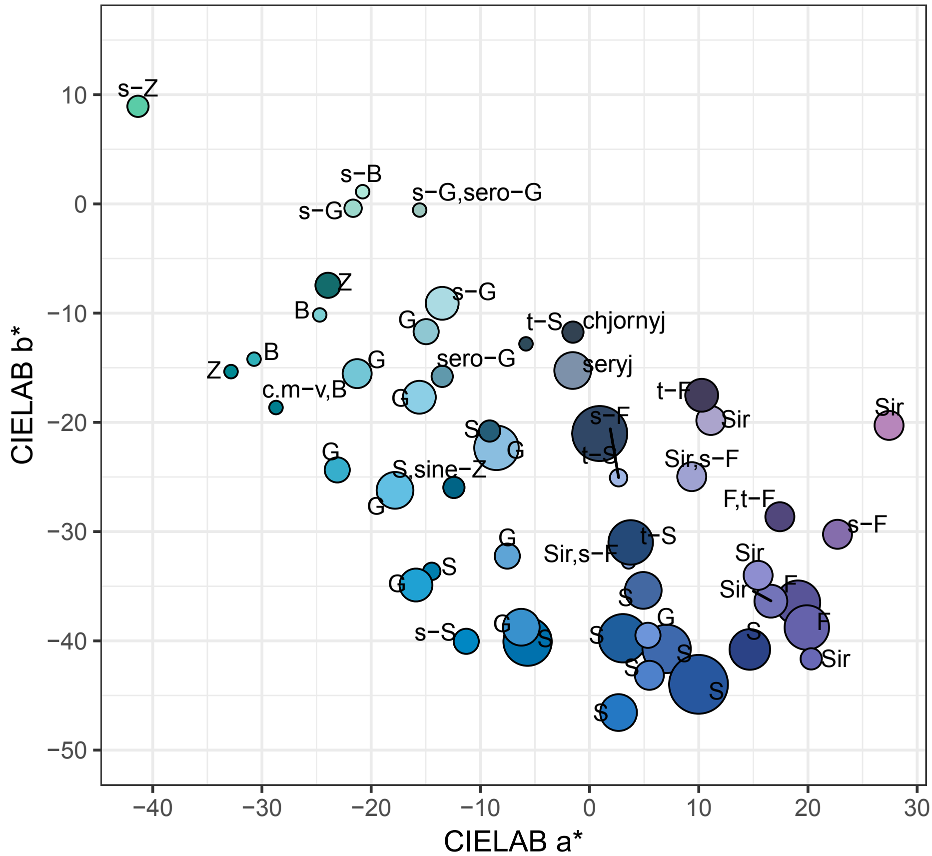

Figure 16. Russian naming task by most frequent stimulus name.

B – *birûzovyj* 'turquoise', c. m-v – *cvet morskoj volny* 'sea wave colour', *čërnyj* 'black', F – *fioletovyj* 'purple', G – *goluboj* 'light blue', Sir – *sirenevyj* 'lilac', *sero* 'greyish', *seryj* 'grey', *sine* 'blueish', Z – *zelënyj* 'green', s – *svetlo* 'light', t – *tëmno* 'dark'.

In Figure 16 the size of the dot representing a stimulus is in correlation with the frequency.

It is notable that *goluboj* 'light blue' was used with the modifiers *svetlo* 'light' and *tëmno* 'dark', which is characteristic of basic colour terms. It would be highly unusual to combine the Estonian *helesinine* with another *hele* 'light' modifier as *hele helesinine* 'light light-blue' for example but this is entirely possible for the Russian *goluboj* 'light blue' as *svetlo* and *tëmno goluboj* 'light' and 'dark' 'light blue' are used in the naming task, and the modifier *sero* 'grey(ish)' is also used. This is an indicator of the strength of the basic colour term traits of the Russian *goluboj*. *Sirenevyj* 'lilac', abbreviated to Sir in Figure 16, is also a named most frequent term in the naming task, but it did not have a named most frequent representation in the sorting task.

In the Russian naming task participants did not reach a high degree of agreement on naming the stimuli. The agreement in the Russian naming task was so low that only one stimulus, B-EX, gained dominance at 50%. Half of the participants (15 out of 30) named B-EX *sinij* 'blue', but that was the only agreement at the 50% dominance level. The analysis of Estonian data indicated that lowering the threshold by just one point can broaden the dominant stimuli selection and possibly even add a dominant name. If the dominance threshold is lowered by one point from 15 to 14, giving 47%, then the stimulus B-S2 *temno-sinij* 'dark blue' (DomF = 14) is added to B-EX *sinij* 'blue' (DomF = 15).

Table 12. Russian naming task by frequency.
F – frequency, No of tiles (range), DomF – dominant frequency (≥ 47%, F=14), SI – specificity index

Naming task	**Gloss**	**F**	**DomF**	**No of dominant tiles**	**No of tiles (range)**	**SI**
sinij	blue	131	15	1	26	0.115
goluboj	'light blue'	104			24	
fioletovyj	purple	59			16	
temno-sinij	dark blue	57	14	1	16	0.246
sirenevyj	lilac	57			14	
svetlo-goluboj	light 'light blue'	41			14	
svetlo- fioletovyj	light purple	36			12	
svetlo-sinij	light blue	32			18	
temno-goluboj	dark 'light blue'	29			17	
žëltyj	yellow	25			4	
birûzovyj	turquoise	21			10	
sero-goluboj	grey 'light blue'	19			6	
seryj	grey	18			8	
temno-fioletovyj	dark purple	18			5	
želtyj	yellow	16			4	
zelënyj	green	16			8	
lilovyj	mauve	16			9	
sine-zelenyj	blue-green	15			6	

Only *sinij* 'blue' gained a dominance (F=15) with one stimulus (B-EX), but even that was borderline. The yellow distractor stimuli were salient. The addition of B-S2 *temno-sinij* 'dark blue' (DomF = 14) raises the number of dominant stimuli to two out of 55, which is very low.

The low naming consensus is not an obstacle in the multidimensional scaling analysis for the Russian naming task, see Figure 17, where the separation between light and dark is noticeable along an imaginary diagonal line (/). Although the separation between light and dark purple is more evident, the blue terms are displayed gradually from the darkest shade of blue to the lightest blue.

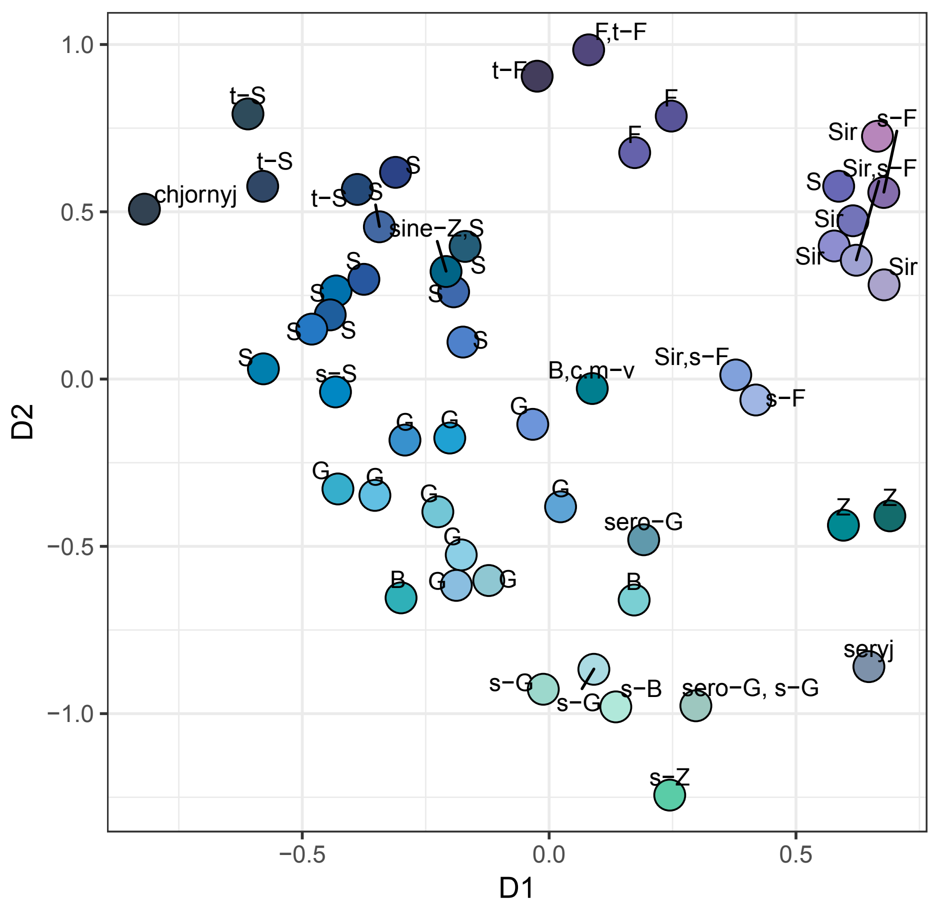

Figure 17. Russian naming task multidimensional analysis.

Most frequent names labelled: B – *birûzovyj* 'turquoise', c. m-v – *cvet morskoj volny* 'sea wave colour', *čërnyj* 'black', F – *fioletovyj* 'purple', G – *goluboj* 'light blue', *sinij* 'blue', Sir – *sirenevyj* 'lilac', *sero* 'greyish', *seryj* 'grey', *sine* 'bluish', Z – *zelënyj* 'green', s – *svetlo* 'light', t – *temno* 'dark'.

The multidimensional scaling analysis for the Russian naming task depicted in Figure 17 does not offer any clear cut results. The division between *fioletovyj* 'purple' and *sirenevyj* 'lilac' is evident, which it was not when the stimuli were mapped by CIE coordinates in Figure 16. Figure 17 can be compared to the Estonian naming data multidimensional scaling analysis in Figure 10, which is a little less spread out.

In the hierarchical cluster analysis for the Russian naming task, see Figure 18, there are three groups. Two of them, which may correspond to the light and dark categories, are quite large and have several subgroups each, while the third group, which contains most frequently named *sirenevyj* 'lilac' stimuli, is quite small and compact.

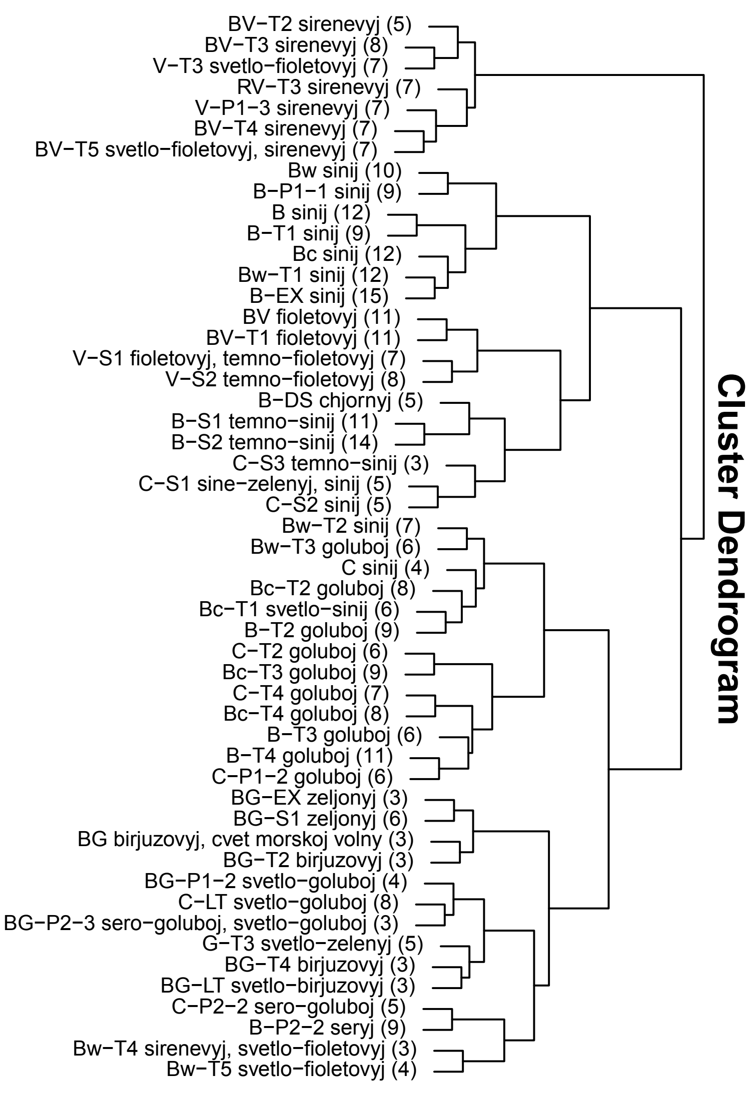

Figure 18. Russian naming task hierarchical cluster analysis.

Like in the output of the principal component analysis for the Estonian naming task, the variance reaches 50% for a much larger number of components than in the sorting task. In the principal components analysis for the Russian naming task, the first component accounts for 14.2% of the variance, the first two components cumulatively for 27.1%, the first three for 34.3%, the first four for 39%, the first five for 43.5%, the first six for 47.3%, and the first seven for 50.8%. It is only the first 32 components taken cumulatively that account for 90% of the variance. The scree plot also has a steep decline after two components, but not as steep as the Estonian naming task scree plot. The Russian scree plot levels out slightly more slowly at the fourth component and almost becomes flat at the eighth component.

The consensus in the Russian naming task was so low that only two stimuli had a consensus of approximately half of the participants, stimuli B-S2 *temno-sinij* 'dark blue' (DomF = 14), and B-EX *sinij* 'blue' (DomF = 15). The modifiers *svetlo* 'light' and *tëmno* 'dark' were more frequently used in the naming task, and the term *sirenevyj* 'lilac' (F=57) became distinguished from *fioletovyj* 'purple'.

2.3.4 Conclusion of the Russian results

In the Russian list task plot (see Figure 12), *goluboj* is found among the other eleven basic colour terms, while other terms stretch the data points more across the plot. In fact the basic terms *koričnevyj* 'brown' and *rozovyj* 'pink' are slightly separate from the rest of the basic terms, bridging the gap between basic and non-basic terms, the most notable of which are *birûzovyj* 'turquoise', *beževyj* 'beige' and *purpurnyj* 'purple'. The basic status of *goluboj* is supported in all three Russian tasks.

Table 13. Overview of Russian *sinij* and *goluboj*
F – frequency (%), mp – mean position, S – cognitive salience index, N – no of participants using the terms (%), StR – number of stimuli named with a term.

Term	**List task (N=30)** **F (%) – mp – S**	**Sorting task (N=30)** **N (%) – F (F/N) – StR**	**Naming task (N=30)** **N (%) – F (F/N) – StR**
sinij	28 (**93%**) – 5.2 – 0.1790	10 (33%) – 124 (12.4) – 39	24 (**80%**) – 131 (5.5) – 26
goluboj	26 (**87%**) – 9.2 – 0.0939	6 (20%) – 68 (11.3) – 20	24 (**80%**) – 104 (4.3) – 24
svetlo sinij	3 (10%) – 8.0 – 0.0125	2 (7%) – 12 (6.0) – 10	11 (37%) – 32 (2.9) – 18
temno sinij	5 (17%) – 12.4 – 0.0134	4 (13%) – 18 (4.5) – 16	18 (**60%**) – 57 (3.2) – 16
svetlo goluboj	1 (3%) – 18.0 – 0.0019		15 (**50%**) – 41 (2.7) – 14
temno goluboj			11 (37%) – 29 (2.6) – 17

In the sorting task *sinij* ‘blue’ had the highest number of participants (10), followed by *žëltyj* ‘yellow’ (N=9), *fioletovyj* ‘violet’ (N=8) and *goluboj* ‘light blue’ (N=6). In comparison to *birûzovyj* ‘turquoise’, *zelënyj* ‘green’ (basic term), and *temno-sinij* ‘dark blue’ (N=4), which were given by four participants, the frequency of *goluboj* ‘light blue’ is high. In the naming task *goluboj* ‘light blue’ came second in overall frequency behind only *sinij* ‘blue’. The named most frequent (*nmf*) count, which means that the term is the highest frequency term for a stimulus, was also high for *goluboj* ‘light blue’. The consensus was low for all the terms in the Russian sorting and naming tasks, but the *nmf* count was very high for *goluboj* in comparison with the Estonian *helesinine* ‘light blue’. *Goluboj* ‘light blue’ was the highest frequency term for 17 stimuli in the sorting task and 10 in the naming task. The visible presence of *goluboj* is also seen in the pictorial representations of the multidimensional scaling analysis (see Figure 17). The figures for the Russian multidimensional scaling analysis show a continuous arc for the *goluboj* stimulus, whereas the Estonian multi-dimensional scaling analysis figures are more scattered.

A remark should be made about the turquoise term, because the previously noted (Davies and Corbett 1994: 81, Rakhilina and Paramei 2011: 125) high frequency of the non-basic term *birûzovyj* ‘turquoise’ was strongly supported by the results of the list task. The relatively low frequency and lower salience of the Estonian counterpart *türkiis* ‘turquoise’ only strengthens the position of the Russian *birûzovyj* ‘turquoise’.

2.4 Estonian and Russian most frequent stimulus names

The most frequent stimuli names are in the purple region of the Estonian and Russian sorting tasks, which is *lilla* in Estonian and *fioletovyj* in Russian. The expected difference is in stimuli sorted into the Russian group named *goluboj* ‘light blue’, which most often has the Estonian equivalent named most frequent *sinine* ‘blue’, for example stimuli B T3, B T4, C T2, C T4, Bc T3, Bc T4, C LT, C P1 2, C P2 (see Table 14). In the sorting task the Russian *goluboj* also had equivalents with names other than *sinine* ‘blue’, as it was also *roheline* ‘green’ (stimulus BG P1 2, perhaps BG T2 as the Russian nmf is *birûzovyj* ‘turquise’). The participants found stimulus B P2 2 rather difficult to group, which is evident from the low consensus, but also gave it quite different names, using the Estonian *lilla* ‘purple’, and the Russian *goluboj* ‘light blue’ and *birûzovyj* ‘turquise’. Stimulus BG P2 3 on the other hand was named with the Estonian *helesinine* ‘light blue’ and the Russian *goluboj* ‘light blue’. There were four stimuli that were termed sinij ‘blue’ by Russian participants, but which gained different Estonian names, as stimuli Bw T4 and Bw T5 were grouped into the *lilla* ‘purple’ category, stimulus B DS was *must* ‘black’, and stimulus B S2 was *tumesinine* ‘dark blue’. All in all, the most frequent terms in the sorting task were most different where the Russian *goluboj* ‘light blue’ was equivalent to the Estonian *sinine* ‘blue’.

Table 14. Estonian and Russian sorting task most frequent stimuli names

Stimuli	Estonian sorting task	%	Russian sorting task	%
Yw	*kollane* 'yellow'	**49**	*žëltyj* 'yellow'	30
Y	*kollane* 'yellow'	**49**	*žëltyj* 'yellow'	30
Yc	*kollane* 'yellow'	46	*žëltyj* 'yellow'	30
G T3	*roheline* 'green'	26	*zelënyj* 'green'	7
			birûzovyj 'turquoise'	7
BG	*roheline* 'green'	23	*zelënyj* 'green'	13
BG T2	*roheline* 'green'	15	*goluboj* 'light blue'	7
			birûzovyj 'turquoise'	7
BG T4	*helesinine* 'light blue' *rohekas* 'greenish' *sinakasroheline* 'bluish green' *sinine* 'blue'	8 8 8 8	*goluboj* 'light blue'	10
C	*sinine* 'blue'	38	*sinij* 'blue'	17
C T2	*sinine* 'blue'	26	*goluboj* 'light blue'	17
C T4	*sinine* 'blue'	15	*goluboj* 'light blue'	13
Bc	*sinine* 'blue'	36	*sinij* 'blue'	30
Bc T1	*sinine* 'blue'	41	*sinij* 'blue'	10
			goluboj 'light blue'	10
Bc T2	*sinine* 'blue'	31	*sinij* 'blue'	13
			goluboj 'light blue'	13
Bc T3	*sinine* 'blue'	21	*goluboj* 'light blue'	17
Bc T4	*sinine* 'blue'	18	*goluboj* 'light blue'	13
B	*sinine* 'blue'	36	*sinij* 'blue'	30
B T1	*sinine* 'blue'	41	*sinij* 'blue'	30
B T2	*sinine* 'blue'	33	*sinij* 'blue'	13
B T3	*sinine* 'blue'	28	*goluboj* 'light blue'	13
B T4	*sinine* 'blue'	21	*goluboj* 'light blue'	13
Bw	*sinine* 'blue'	15	*sinij* 'blue'	13
Bw T1	*sinine* 'blue'	18	*sinij* 'blue'	10
Bw T2	*sinine* 'blue'	21	*sinij* 'blue'	10
Bw T3	*lilla* 'purple'	21	*fioletovyj* 'purple'	10
Bw T4	*lilla* 'purple'	36	*sinij* 'blue'	10
Bw T5	*lilla* 'purple'	28	*sinij* 'blue'	10
BV	*lilla* 'purple'	**51**	*fioletovyj* 'purple'	20
BV T1	*lilla* 'purple'	44	*fioletovyj* 'purple'	20
BV T2	*lilla* 'purple'	46	*fioletovyj* 'purple'	17
BV T3	*lilla* 'purple'	44	*fioletovyj* 'purple'	17

Stimuli	Estonian sorting task	%	Russian sorting task	%
BV T4	*lilla* 'purple'	44	*fioletovyj* 'purple'	17
BV T5	*lilla* 'purple'	41	*fioletovyj* 'purple'	17
V T3	*lilla* 'purple'	46	*fioletovyj* 'purple'	20
RV T3	*lilla* 'purple'	41	*fioletovyj* 'purple'	20
B EX	*sinine* 'blue'	28	*sinij* 'blue'	20
Yc EX	*kollane* 'yellow'	**49**	*žëltyj* 'yellow'	30
BG EX	*roheline* 'green	23	*zelënyj* 'green'	13
BG LT	*roheline* 'green	13	*goluboj* 'light blue'	10
C LT	*sinine* 'blue'	13	*goluboj* 'light blue'	13
B DS	*must* 'black'	15	*sinij* 'blue'	17
BG S1	*roheline* 'green'	23	*zelënyj* 'green'	13
BG P1 2	*roheline* 'green'	13	*goluboj* 'light blue'	10
BG P2 3	*helesinine* 'light blue'	10	*goluboj* 'light blue'	10
C S1	*sinine* 'blue'	13	*sinij* 'blue'	23
C P1 2	*sinine* 'blue'	18	*goluboj* 'light blue'	13
C S2	*sinine* 'blue'	10	*sinij* 'blue'	20
C P2 2	*sinine* 'blue'	15	*goluboj* 'light blue'	17
C S3	*sinine* 'blue'	8	*sinij* 'blue'	10
	tumesinine 'dark blue'	8		
	roheline 'green'	8		
B S1	*sinine* 'blue'	23	*sinij* 'blue'	20
B P1 1	*sinine* 'blue'	23	*sinij* 'blue'	13
B S2	*tumesinine* 'dark blue'	18	*sinij* 'blue'	20
B P2 2	*lilla* 'purple'	15	*goluboj* 'light blue'	7
			birûzovyj 'turquoise'	7
V S1	*lilla* 'purple'	44	*fioletovyj* 'purple'	20
V P1 3	*lilla* 'purple'	41	*fioletovyj* 'purple'	17
V S2	*lilla* 'purple'	41	*fioletovyj* 'purple'	20

The range of differences in the most frequent stimuli names in the naming task is notable. The most probable reason for the variance is that the naming task stimuli were presented one by one and after the sorting task, which almost enforced specific names. It is therefore quite dangerous to judge the named equivalents (Table 15) because the Estonian *sinine* 'blue' can be found to be equivalent to the Russian *sinij* 'sinine' (stimuli C, Bc, B, B T1, Bw T1, Bw T2, B EX, B P1 1), *svetlo-sinij* 'light blue' (stimulus Bc T1) and *goluboj* 'light blue' (stimulus B T2). Of course, even in the naming task the same most frequent stimulus names match, so that the Russian *sinij* 'blue' equivalent is mostly *sinine* ‘blue’ (C, Bc, B, B T1, Bw T1, Bw T2, B EX, B P1 1). For stimuli Bw and C S2 the Russian *sinij* ‘blue’ was named darker in Estonian as *tumesinine* ‘dark blue’.

Table 15. Estonian and Russian naming task most frequent stimuli names

Stimuli	Estonian naming task	%	Russian naming task	%
Yw	*kollane* ‘yellow’	13	*oranževyj* ‘orange’	20
	tumekollane ‘dark yellow’	13		
Y	*kollane* ‘yellow’	**49**	*žëltyj* ‘yellow’	30
Yc	*kollane* ‘yellow’	33	*limonnyj* ‘lemony’	23
G T3	*heleroheline* ‘light green’	31	*svetlo-zelenyj* ‘light green’	27
BG	*sinakasroheline* ‘bluish green’	15	*cvet morskoj volny* ‘sea wave colour’	10
			birûzovyj ‘turquise’	10
BG T2	*rohekassinine* ‘greenish blue’	18	*birûzovyj ‘turquise’*	10
BG T4	*helesinine* ‘light blue’	23	*birûzovyj ‘turquise’*	10
C	*sinine* ‘blue’	44	*sinij* ‘blue’	13
C T2	*helesinine* ‘light blue’	26	*goluboj* ‘light blue’	20
C T4	*helesinine* ‘light blue’	38	*goluboj* ‘light blue’	23
Bc	*sinine* ‘blue’	44	*sinij* ‘blue’	40
Bc T1	*sinine* ‘blue’	33	*svetlo-sinij* ‘light blue’	20
Bc T2	*helesinine* ‘light blue’	15	*goluboj* ‘light blue’	27
Bc T3	*helesinine* ‘light blue’	**54**	*goluboj* ‘light blue’	30
Bc T4	*helesinine* ‘light blue’	**59**	*goluboj* ‘light blue’	27
B	*sinine* ‘blue’	**49**	*sinij* ‘blue’	40
B T1	*sinine* ‘blue’	**49**	*sinij* ‘blue’	30
B T2	*sinine* ‘blue’	23	*goluboj* ‘light blue’	30
B T3	*helesinine* ‘light blue’	**54**	*goluboj* ‘light blue’	20
B T4	*helesinine* ‘light blue’	46	*goluboj* ‘light blue’	37
Bw	*tumesinine* ‘dark blue’	26	*sinij* ‘blue’	33
Bw T1	*sinine* ‘blue’	31	*sinij* ‘blue’	40
Bw T2	*sinine* ‘blue’	31	*sinij* ‘blue’	23
Bw T3	*helelilla* ‘light purple’	13	*goluboj* ‘light blue’	20
Bw T4	*helelilla* ‘light purple’	26	*svetlo-fioletovyj* ‘light purple’	10
			sirenevyj ‘lilac’	10
Bw T5	*helelilla* ‘light purple’	28	*sirenevyj* ‘lilac’	13
BV	*tumelilla* ‘dark purple’	28	*fioletovyj* ‘purple’	37
BV T1	*lilla* ‘purple’	38	*fioletovyj* ‘purple’	37

Stimuli	Estonian naming task	%	Russian naming task	%
BV T2	*lilla* 'purple'	**51**	*sirenevyj* 'lilac'	17
BV T3	*lilla* 'purple'	28	*sirenevyj* 'lilac'	27
BV T4	*helelilla* 'light purple'	**51**	*sirenevyj* 'lilac'	23
BV T5	*helelilla* 'light purple'	**54**	*svetlo-fioletovyj* 'light purple'	23
			sirenevyj 'lilac'	23
V T3	*lilla* 'purple'	38	*svetlo-fioletovyj* 'light purple'	23
RV T3	*helelilla* 'light purple'	21	*sirenevyj* 'lilac'	23
B EX	*sinine* 'blue'	36	*sinij* 'blue'	**50**
Yc EX	*kollane* 'yellow'	38	*žëltyj* 'yellow'	33
BG EX	*sinakasroheline* 'bluish green'	18	*zelënyj* 'green'	10
BG LT	*heleroheline* 'light green'	21	*svetlo-birûzovyj* 'light turquoise'	10
C LT	*helesinine* 'light blue'	31	*svetlo-goluboj* 'light 'light blue''	27
B DS	*mustjassinine* 'blackish blue'	18	*čërnyj* 'black'	17
BG S1	*roheline* 'green'	18	*zelënyj* 'green'	20
BG P1 2	*helesinine* 'light blue'	13	*svetlo-goluboj* 'light 'light blue''	13
	heleroheline 'light green'	13		
BG P2 3	*helesinine* 'light blue'	10	*svetlo-goluboj* 'light 'light blue''	10
			sero-goluboj 'grey 'light blue''	10
C S1	*rohekassinine* 'greenish blue'	15	*sine-zelenyj* 'bluish green'	17
			sinij 'blue'	17
C P1 2	*helesinine* 'light blue'	44	*goluboj* 'light blue'	20
C S2	*tumesinine* 'dark blue'	18	*sinij* 'blue'	17
C P2 2	*helesinine* 'light blue'	18	*sero-goluboj* 'grey 'light blue''	17
C S3	*tumesinine* 'light blue'	21	*temno-sinij* 'dark blue'	10
B S1	*tumesinine* 'light blue'	**49**	*temno-sinij* 'dark blue'	37
B P1 1	*sinine* 'blue'	28	*sinij* 'blue'	30
B S2	*tumesinine* 'dark blue'	36	*temno-sinij* 'dark blue'	**47**
B P2 2	*hall* 'grey'	21	*seryj* 'grey'	30
V S1	*tumelilla* 'dark purple'	41	*temno-fioletovyj* 'dark purple'	23
			fioletovyj 'purple'	23
V P1 3	*helelilla* 'light purple'	**49**	*sirenevyj* 'lilac'	23
V S2	*tumelilla* 'dark purple'	**56**	*temno-fioletovyj* 'dark purple'	27

In the purple region Estonians used *hele* ‘light’ or *tume* ‘dark’ as modifiers for *lilla* ‘purple’, while Russian participants used those modifiers for *fioletovyj* ‘purple’, but not for *sirenevyi* ‘lilac’. *Fioletovyj* ‘purple’ and *sirenevyj* ‘lilac’ are sometimes used for the same stimulus, which lowers the consensus, but *sirenevyj* seems for be used for a lighter purple, like the Estonian *helelilla* ‘light purple’.

2.5 Summary of results

As in previous studies, the Russian goluboj ‘light blue’ is found to be a separate category from sinij ‘blue’, which illustrates some support for the weak relativist view on colour. It means that there being two ‘blue’ categories is not only theoretically possible, but previously experimentally proven. If there is only one ‘blue’ category in Estonian, for example, then the universalist view of Berlin and Kay’s theory of basic colour terms is offered more support by the Estonian data. The focus on the possibility of there being more than one general term for blue comes from the proposed hypothesis, which dictated the selection of languages for the experimental part of the thesis.

The three datasets aim to present a case where the different use of the blue categories is displayed, with the Russian *goluboj* firmly established as a basic level colour term, while the Turkish *lacivert* is a borderline case of a basic colour term. In Estonian, however, the non-basic blue terms *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’ do not fit the established criteria for a basic colour term, but in some instances they meet the thresholds expected of a basic colour term.

In Figure 19 each dot represents a significant blue term and the size of the dot is dependent on the listing task percentage. The basic colour terms for ‘blue’, which are the Estonian *sinine*, Russian *sinij*, and Turkish *mavi*, have the largest listing percentage of the blue terms in all the languages. The dotted line in Figure 19 marks the halfway point at 50%. The “second blue” basic colour terms were all listed by more than half of the participants. The basic colour terms for ‘blue’ are followed by the Russian basic colour term *goluboj* ‘light blue’, the Turkish probable basic colour term *lacivert* ‘dark blue’, and the Estonian modified term *helesinine* ‘light blue’. The Estonian *tumesinine* ‘dark blue’ was not listed by more than half of the participants.

The Estonian term *helesinine* ‘light blue’ has strong indicators in the list task (see Figure 3 and Figure 19) but does not show the grouping behaviour required for a basic colour term in the sorting task. Indeed the Estonian *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’ did not gain high enough frequencies in the sorting task for them to be considered basic colour terms, at only six and seven participants respectively. Analysis of the sorting data indicates the hypothesis of Estonian having several separate categories for blue, namely *sinine* ‘blue’, *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’, is not supported. That is, the analysis of the results offers more support for the null-hypothesis. The null-hypothesis of only one 'blue' category supports the universal viewpoint on colours.

The Russian *goluboj* 'light blue' exhibits the characteristics of a fully-developed basic colour term, while the Turkish *lacivert* 'dark blue' shows weaker indicators of basic traits, most especially quite low agreement in the naming task.

Russian with *sinij* 'blue' and *goluboj* 'light blue' and Turkish with *mavi* 'blue' and *lacivert* 'dark blue' divide their 'blues' into two basic colour terms, while Estonian divides one basic colour term – *sinine* 'blue' – into sub-concepts using *hele* 'light' and *tume* 'dark' as modifiers.

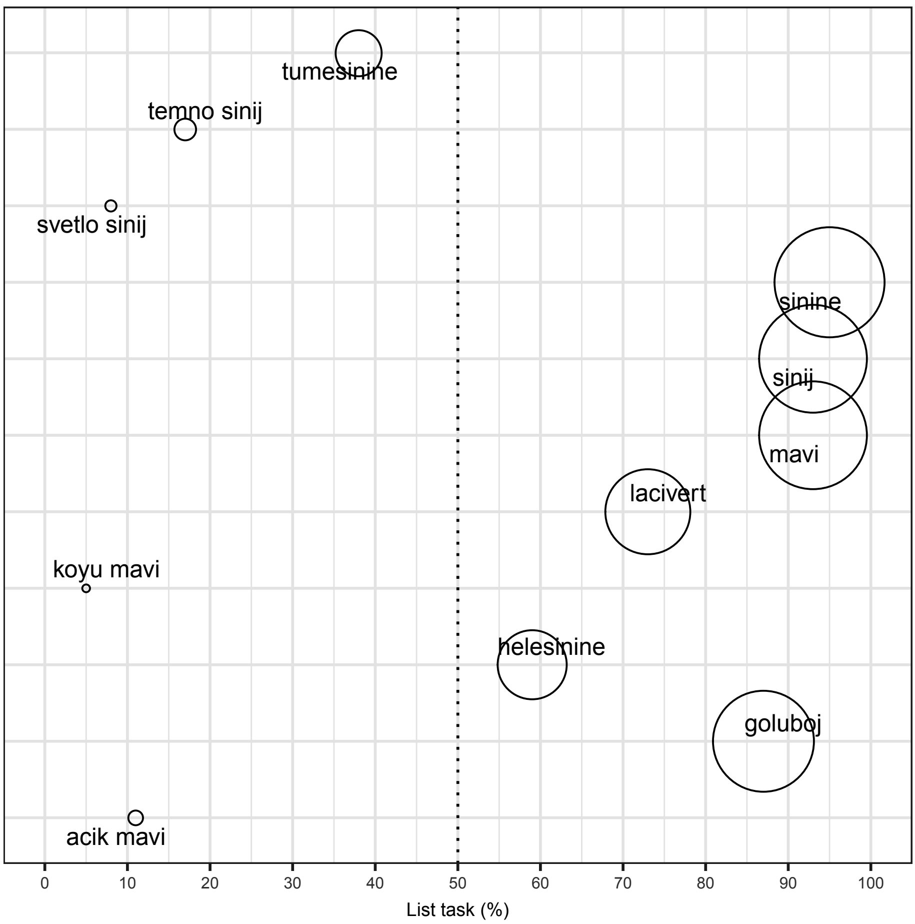

Figure 19. Listing percentage of frequent blue terms in Turkish, Estonian and Russian.

To help in visualising the list task data, graphs were created with the frequency of the listed colour terms on the x-axis and the mean position on the y-axis, which illustrate the nature of the data. The Turkish list task plot, Figure 3, indicates a noticeable divide between basic and non-basic term frequency and mean position count. *Lacivert* 'dark blue' is included in the basic colour term division. Listed by fewer than half of the participants are the non-basic colour terms *lila* 'lilac', *bordo* 'burgundy' and *eflatun* 'mauve', which are quite distant from the first twelve terms. These three terms probably have some traits in common with the basic colour terms. The Estonian list task results confirm the eleven previously established basic colour terms. The Estonian list task plot, Figure 5, illustrates the in-between nature of *beež* 'beige' and *helesinine* 'light blue'. Despite being a modified term and thus a non-basic term by the original definition, the frequency, mean position and salience of *helesinine* 'light blue' in the list task place the word together with *beež* 'beige' between the basic and non-basic colour terms, indicating a relative basicness in the list task. The list task supports *helesinine* being put in a separate category.

The Russian list task plot, Figure 12, has more in-between terms than Turkish or Estonian, and these are the basic colour term *rozovyj* 'pink' with a list task percentage of 63%, followed by *birûzovyj* 'turquoise', *beževyj* 'beige', *purpurnyj* 'purple', *sirenevyj* 'lilac', *bordovyj* 'bordeaux [red]', *lilovyj* 'mauve', and *salatovyj* 'lettuce'. *Goluboj* 'light blue' is firmly situated in centre of the basic terms in Figure 12 and thus is by frequency and mean position nowhere near the borderline in the list task.

In the Turkish list task, 18 terms out of 180 different terms accounted for 69% of the frequencies of the Turkish list task data. These 18 terms, ranked by the percentage of participants who listed them, are *yesil* 'green' (97%), *siyah* 'black', *sari* 'yellow' (93%), *beyaz* 'white' (90%), *kirmizi* 'red' (88%), *mavi* 'blue' (87%), *mor* 'purple', *kahverengi* 'brown', (80%), *turuncu* 'orange', *pembe* 'pink' (78%), *gri* 'grey' (72%), *lacivert* 'dark blue' (68%), *lila* 'lilac', *bordo* 'burgundy' (43%), *eflatun* 'mauve' (40%), *bej* 'beige' (30%), *turkuaz* 'turquoise' (28%), and *krem* 'cream' (20%). The previously researched and established eleven Turkish basic colour terms are confirmed by the list task results. According to these list task results *lacivert* 'dark blue' should be included, because its frequency is above half at 68%, its mean position of 10.5 is ranked twelfth and its cognitive salience index of 0.0653 is ranked eleventh above the basic colour term *gri* 'grey'. These parameters are comparably weak for a basic colour term. *Mavi* 'blue' has a listing percentage of 87% and a very low mean position of 3.8. However, the colour terms *lila* 'lilac' and *bordo* 'burgundy', which follow *lacivert* 'dark blue', have a far lower listing percentage at 43%.

In the Estonian list task, 23 terms out of 339 listed words accounted for 48% of the Estonian list data frequency. These 23 terms, ranked by the percentage of participants who listed them, are *punane* 'red', *sinine* 'blue', *kollane* 'yellow' (95%), *roheline* 'green' (92%), *valge* 'white', *hall* 'grey', *oranž* 'orange'(87%), *must* 'black' (85%), *pruun* 'brown' (82%), *lilla* 'purple', *roosa* 'pink' (79%), *beež* 'beige' (67%), *helesinine* 'light blue' (59%), *tumesinine* 'dark blue' (38%), *heleroheline* 'light green', *taevasinine* 'sky blue' (33%), *helepunane* 'light red', *violetne* 'violet', *hõbedane* 'silvery', *kuldne* 'golden', and *potisinine* 'pot blue' (28%). The list task results confirm 11 basic colour terms in Estonian: *punane* 'red', *sinine* 'blue', *kollane* 'yellow', *roheline* 'green', *oranž* 'orange', *valge* 'white', *must* 'black', *hall* 'grey', *lilla* 'purple', *pruun* 'brown' and *roosa* 'pink'.

In the Russian list task, 25 terms out of 294 different words listed accounted for 50% of the Russian list data. These 25 terms are *krasnyj* 'red' (F=30, 100%), *oranževyj* 'orange' (97%), *sinij* 'blue', *žëltyj* 'yellow' (93%), *zelënyj* 'green' (90%), *belyj* 'white', *seryj* 'grey', *goluboj* 'light blue' (87%), *fioletovyj* 'purple' (83%), *čërnyj* 'black' (80%), *koričnevyj* 'brown' (73%), *rozovyj* 'pink' (63%), *birûzovyj* 'turquoise' (57%), *beževyj* 'beige' (53%), *purpurnyj* 'purple' (50%), *sirenevyj* 'lilac' (43%), *bordovyj* 'bordeaux [red]', *lilovyj* 'mauve' (40%), *sala-tovyj* 'lettuce' (37%), *cvet morskoj volny* 'sea wave colour' (33%), *zolotoj* 'golden', *limonnyj* 'lemon[-coloured]', *alyj* 'scarlet'(30%), *višnëvyj* 'cherry', and *malinovyj* 'raspberry' (27%). Bordering the criterium of one quarter are *ohra* 'ochre', *buryj* 'brown red', *serebrânyj* 'silvery', *ul'tramarin* 'ultramarine', and *haki* 'khaki' (F=7, 23%). The Russian list task confirms 12 basic colour terms by frequency: *krasnyj* 'red', *oranževyj* 'orange', *sinij* 'blue', *žëltyj* 'yellow', *zelënyj* 'green', *goluboj* 'light blue', *belyj* 'white', *fioletovyj* 'purple', *čërnyj* 'black', *seryj* 'grey', *koričnevyj* 'brown' and *rozovyj* 'pink'.

The characteristics of the basic colour terms exhibited in the field tests are most evident in use by participants in percentages. In the list task for example, the percentage of participants who used the term in question is illustrative, but in stimuli tasks, the percentage of participants agreeing on a name for a stimulus is the preferred choice, rather than a simple usage percentage. If consensus is low in the stimuli task, then the usage percentage for each term can also be useful.

If participants agree on naming stimuli with a consensus of over 50%, then the term is most probably a basic colour term. If the percentage of either listing the term or naming the stimuli with that term is less than 50%, then the basic status of the colour word is questionable. If possible, the results of the list task should be compared to previous studies of basic colour terms to see if a term is considered a basic colour term, or with the results of the next task where possible.

It is notable that *goluboj* 'light blue' was used with the modifiers *svetlo* 'light' and *tëmno* 'dark', which is characteristic of basic colour terms. It highly unusual to combine the Estonian *helesinine* with another *hele* 'light' modifier in *hele helesinine* 'light light-blue' for example, but for the Russian *goluboj* 'light blue' it is entirely possible, and *svetlo* and *tëmno goluboj* 'light' and 'dark 'light blue' are used in the naming task.

Table 16. Most frequent blue terms by participant use (%)

	'Blue'			'Blue2'		'Light blue'			'Dark blue'		
	mavi	*sinine*	*sinij*	*lacivert*	*goluboj*	*acik mavi*	*hele-sinine*	*svetlo sinij*	*koyu mavi*	*tume-sinine*	*temno-sinij*
List	**93**	**95**	**93**	**73**	**87**	11	**59**	8	5	38	17
Sorting	–	44	33	–	20	–	15	7	–	18	13
Naming	**100**	**85**	**80**	**75**	**80**	**88**	**85**	37	**71**	**67**	**60**

The listing percentages for *acik mavi* 'light blue' (11%) and *koyu mavi* 'dark blue' (5%) are low (see Table 16) in comparison to the Estonian *helesinine* (59%) and *tumesinine* (38%), but they are comparable to the Russian *svetlo sinij* (8%) and *temno sinij* (17%). In the naming task the Turkish *acik lacivert* 'light 'dark blue'' (20%) and *koyu lacivert* 'dark 'dark blue'' (21%) were also used, as were the Russian *svetlo* and *tëmno goluboj* 'light' (50%) and 'dark 'light blue'' (37%).

The Estonian and Russian sorting task found that presenting the coloured stimuli with the most frequent stimulus name labelled offers a visual overview of the stimuli task, where comparison between languages is also possible. For example, the Russian *goluboj* 'light blue' is much more numerous in the sorting task (for Russian see Figure 13, 14) than the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' (see Figure 6, 7). In Estonian one *helesinine* 'light blue' stimulus was named most frequent in the sorting task, while in Russian *goluboj* 'light blue' is literally in a category of its own.

The figures for the Estonian and Russian naming tasks, Figure 9 and Figure 17, are more detailed because the participants had just completed grouping the stimuli and then naming the groups for the sorting task, and the naming task required the participant to name each stimulus one at a time. The Turkish naming task, which had different stimuli, can be seen in Figure 4.

In the Estonian naming task *helesinine* was more salient than *sinine*, reversing the situation in the listing and sorting tasks, but the appearance of two stimuli for unequivocally non-basic terms *helelilla* 'light purple' and *tumelilla* 'dark purple' gaining dominance suggests that the *helesinine* 'light blue' stimulus gaining dominance, with over 50% agreement by the participants, is not convincing evidence for the formation of a 'light blue' category in Estonian.

Lowering the dominance threshold by one point, can influence which *names* are possible dominant. If the cut-off point were to be lowered by half a point to nineteen, then *tumesinine* 'dark blue' would gain dominance in the Estonian naming task for stimulus B-S1. The agreement in the Russian naming task was so low that only one stimulus, B-EX, gained dominance at 50%. If the dominance threshold is lowered by one point, from 15 to 14, making 47%, then stimulus B-S2 *temno-sinij* 'dark blue' (DomF = 14) is added to B-EX *sinij* 'blue' (DomF = 14). The addition of B-S2 *temno-sinij* 'dark blue' (DomF = 14) increases the number of dominant stimuli to two out of 55, which is very low.

2.6 Comments

By null-hypothesis if the Turkish *lacivert* 'dark blue' and Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' are not basic colour terms, then the quantifiable measures of their basic colour traits will not match those of previously established basic colour terms.

The Estonian list task results indicate that *helesinine* 'light blue' has some traits that are attributed to basic colour terms, high frequency first and foremost, but both *helesinine* 'light blue' and *beež* 'beige' are by frequency between the eleven Estonian basic colour terms and the rest of the terms (see Figure 5). The frequency for the Turkish *lacivert* 'dark blue' is higher than that of *gri* 'grey', a previously established basic colour term, and it is placed within the basic colour terms (see Figure 3). The Russian *goluboj* 'light blue' has the strongest basic traits, as it is situated with the other basic colour terms (see Figure 12).

In the sorting task the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' did not gain high enough frequencies at six and seven participants respectively and featured weakly in the sorting task analysis, see Figures 6, 7 and 8. The Russian *goluboj* also had low frequencies in the sorting task, but in comparison to the Estonian, where there was only one stimuli where the group name *helesinine* 'light blue' was dominant, the corresponding group name for the Russian *goluboj* 'light blue' is literally in a category of its own (see Figures 13, 14, 15).

Based on the analysis of the sorting data, the hypothesis of Estonian having several separate categories for blue, namely *sinine* 'blue', *helesinine* 'light blue' and *tumesinine* 'dark blue', is not supported. That is to say, the analysis of the results offers more support for the null-hypothesis.

Overall there are some possible limiting factors that could affect the results and comparison with previous research, notably the exclusion of modifiers. If any or all modifiers are removed from the analysable data, then the extent to which the data were manipulated should be explained and examples given of how it was done, preferably before the results are analysed. Ideally this would cover how far the manipulation of the data altered the results, even if the exclusion of modifiers was part of the required methodology. Authors may modify the original data before data analysis, and in *The World Color Survey* there is a dataset, where the data were "subsumed under the forms presented here" (Kay *et al.* 2010: 575), which suggests that the data were modified to some degree, but unfortunately no clear examples, or numbers, were given. In earlier Turkish list task results the authors state that they collapsed all the simple terms used with a general modifier like *acik* 'light' or *koyu* 'dark' onto the simple form. This trend seems to continue through the whole article, eliminating all modifiers (Özgen and Davies 1998). I believe that if the original data are modified in any form, then this should be documented with examples, and numbers. If I were to remove the *acik* 'light' (F = 737) and *koyu* 'dark' (F = 565) modifiers from the Turkish naming task, then it would increase the number of dominant stimuli by more than a quarter. More importantly, it would add a dominant stimulus ('Navy Blue') for *lacivert* 'dark blue' (F = 29). That would strengthen the case for *lacivert* 'dark blue' having basic colour term status. Overall dominance would increase if the data were to be collapsed into the "simple form". As previously stated, the extent to which modifiers are used varies from language to language, and 14% of the Estonian participants used *hele* 'light' and 13% *tume* 'dark' modifiers, while only 6% of the Russian participants did so.

Another factor that may distort the results is the sequence of the tasks, in which the question of priming arises. The list task is meant for the mental colour vocabulary, the free-sorting of stimuli for ascertaining colour categories, and the one-by-one naming of stimuli for ascertaining consensual colour terms. In the field it is most practical to test as many participants as possible in as many tasks as possible. If possible, the tasks should be separate, especially the sorting and naming tasks. In ideal conditions only one task should be performed by each participant, or the selection should be divided, so that half of participants do the list and sorting tasks, and the other half do the list and naming tasks. For further research, a sorting task should be conducted where half the participants know in advance that they have to name the groups they sort, and half of the participants do not know of the naming requirement. A small sample was also tested with the naming task preceding the sorting task, but a larger sample would be beneficial to show the extent of the priming. Group-naming after completing the grouping of stimuli in the sorting task was a challenging task, forcing both lumpers with two or three heterogeneous groups and splitters with 21 groups to coin ad-hoc non-basic descriptions for them, which unsurprisingly generated a wide diversity of names.

It is possible that the sequence of tasks and the selection of stimuli are the most important factors that influence both the results and the interpretation of results. Unquestionably, if the results of the naming task were taken in isolation, then *helesinine* 'light blue' would be considered a basic colour term, but the sequence of tasks and the selection of stimuli need to be taken into consideration. The 65 colour tiles were originally chosen by Davies and Corbett because they "formed a coarse, but evenly spread sample of colour space" (1995: 27). Androulaki *et al.* (2006: 27) add the caveat that the Color-Aid stimuli under-represent the *γalázjo* 'light blue' region in Greek, leaving gaps. The authors are convinced that these missing regions included good examples of *γalázjo* 'light blue', and had they included these regions, *γalázjo* would probably have been used more frequently and, perhaps, without modifiers (2006: 27). The Color-Aid stimulus BGB T3 seems to be favoured to become dominant for light blue if the language has a high enough naming percentage for it. Ryabina proposes that in Russian and Northern Udmurt, the terms for a lighter blue, *goluboj* and *chagyr* respectively, had the same location in colour space, and they corresponded to colour sample BGB T3 (Ryabina 2011a: 200; see also Rjabina 2011, and Bimler and Uusküla 2016). Lithuanian is another example of a language where light blue gained dominance in the naming task with Color-

Aid stimulus BGB T3. Research by Pranaitytė (2011: 298) indicates that the Lithuanian *žydra* 'light blue' may be basic alongside *mėlyna* 'blue'.

The results of the sorting task and the naming task depend on the even distribution of the selected stimuli. To confirm or deny the hypothesis, there was a disproportionate number of bluish stimuli. It is possible that the ratio of experimental stimuli, with 51 non-yellow tiles, to distractors, with four yellow stimuli, is too small. If the selection of stimuli is seen as a methodological flaw, it does not exclude the possibility of there being more than one blue colour category. The same tasks and stimuli were applied to Italian (Uusküla 2014), where evidence for there being several blue categories emerged for *blu, azzurro* and *celeste.* It could be speculated that the four yellowish distractor stimuli used force a strong isolating effect. However, in the Estonian naming task the purple stimuli had a smaller range and lower frequencies of use than the blue stimuli, which made for a more concise set of stimuli that was apparently easier for the participants to name, and thus more consensual. The selection of stimuli may have forced some of the participants to group the non-bluish stimuli more tightly and the larger number of bluish stimuli presented may have led them to make different groups for blue in the sorting task and increase specificity in the naming task.

Two unexpected but fairly significant points were gained from the analysis of the empirical tasks. The first is the question of dominance thresholds, which, while undoubtedly influenced by the possible priming effect of the sequence of tasks and the disproportionate number of bluish stimuli, is still worth considering in future research.

The exclusion of modifiers is one issue which raises the dominance or consensus level. Another, which increases the consensus, is lowering the consensus threshold. While the threshold is not usually lowered below 50%, it can be useful to explore whether lowering the threshold by a margin of, say, one participant would result in an increase in possible basic colour terms. In the Estonian (N=39) data, especially in the naming task, the dominance threshold was revealed to be an influential factor. The dominant frequency should be twenty and over, as 39 divided by two is 19.5, but lowering the threshold by one point from twenty to nineteen should not have had much of an influence on the number of dominant *names* in the sorting and naming tasks. However, if the cut-off point is lowered by half a point to nineteen, then in the Estonian naming task *tumesinine* 'dark blue' would gain dominance (stimulus B-S1). In the Russian naming task, if the dominance threshold is lowered by one participant to 47%, then the colour term *temno-sinij* 'dark blue' gains dominance with stimulus B-S2. Dark blue gains dominance in both the Estonian and Russian naming tasks if the dominance threshold is reduced by one participant. That one point, or one participant, is very much dependent on the number of participants. Here the phenomenon appears in both Estonian and Russian for 'dark blue', but it may influence a smaller sample more. The selection of stimuli may also be influential.

In future research, isolating the sorting task would be a much better choice. Doing only the sorting task would eliminate the possibility of any priming effect. The sorting task instruction was to group similar stimuli together. The implicit assumption was that the participants would sort the stimuli into groups by colour terms. After the list task, which primes the participants' colour vocabulary, it is quite unlikely that the participant will not comply with the implicit instruction to sort by colour, but the decision on the criteria of sorting is selected by the participant, so whether to use overall or comparative darkness or lightness of the stimuli, the feeling evoked, the colour category which is the most inclusive of most stimuli, or something else. The strategy chosen by the participants is influenced by the information given. The participants sorted the stimuli into groups and only after the groupings were finished, were they instructed to *name* the groups. The strategy used for sorting the stimuli was made more from a visual point of view than from a linguistic point of view, meaning participants had more than one option for grouping stimuli together under one, single, inclusive colour term. In fact, one participant remarked that the groups would have been different if it had been known beforehand that the groups would be named.

A rather offhand remark, which nevertheless raises the important point of how different the groups would have been if the participants had been instructed to group and name the groups before sorting the stimuli into piles. This could be dealt with in several ways, so for example, another sorting exercise could be conducted, where the participants know beforehand that they have to name the groups. This would offer valuable comparison material and insight into sorting behaviour.

While the results on the whole support the null-hypothesis, several important datapoints emerged from this research. First and foremost was plotting the list task results by mean position and frequency as a percentage, with the highest frequency terms labelled (see Figures 3, 5 and 12). This offers a simple, but accurate, overview of most of the list task data at one glance. That should not be undervalued, since it helps to illustrate not only the salience of the terms, but also the structure of the list task data.

Secondly, the fact that the Estonian and Russian data is directly comparable because the methodology was identical. It makes is possible to view Russian *sinij* 'blue' and *goluboj* 'light blue' with Estonian *sinine* and *helesinine* (and *tumesinine* 'dark blue').

SUMMARY

Empirical tests were used for all three languages, with a list task to elicit the mental colour lexicon and a naming task to ascertain the naming of the stimuli, and a sorting task to categorise the stimuli in Estonian and Estonian Russian, a linguistic minority in Estonia. Of interest from a theoretical standpoint, following Berlin and Kay's theory of the universalist colour vocabulary for basic colour terms, is the use of blue terms in the three languages.

The theory of basic colour terms argues that there are a limited number of universal colour words that designate general colour categories. Brent Berlin and Paul Kay (1969) proposed the hypothesis that there is a certain number of universally used colour terms. They used their data to generate the criteria for basic colour terms and these criteria are sometimes rigidly adhered to even though they were originally created more as guidelines (see Biggam 2012: 22).

Grossmann and D'Achille (2016) summarise the criteria of basic colour terms elegantly, saying that basic colour terms are neither morphologically complex nor semantically transparent, they are not hyponyms of other terms, their application is not restricted to specific classes of entities, and they are psychologically salient to speakers .

The World Color Survey (2010) supplies a very useful guideline that directs much of the analysis in this study, in that the strength of the basic status of a term can only be assessed relatively and in comparison to the other terms in the same language (Kay *et al.* 2010: 21). Whether a term is basic or not can be judged in field tests by the fieldworker's expert opinion and considering (Kay *et al.* 2010: 21) the percentage of speakers that use the term; the range of stimuli named by the term on the modal array; the level of agreement at which the term first appears on the naming aggregates, where higher levels mean the term is more established; and the relative clarity of the definition of the term map. These criteria for judging whether a term is basic or not (Kay *et al.* 2010: 21) can be applied to the naming task by finding

1) the highest consensus for a tile, or *nmf* – named most frequent;
2) the range of stimuli the term covers; and possibly
3) the nmf mapped on the plot and the multidimensional scaling plot.

The characteristics exhibited by the basic colour terms in the field tests are most evident in the percentages for usage by participants, such as the percentage of participants who used the term in question in the list task. In the stimuli tasks, the percentage of participants agreeing on a name for a stimulus is the preferred choice, rather than a simple usage percentage. If the overall consensus is low in the stimuli task, then the usage percentage for each term can also be useful.

If participants agree on the naming stimuli with a consensus of over 50%, then the term is most probably a basic colour term. If the percentage of those listing the term or naming a stimulus with that term is less than half, then the basic status of the colour word is questionable. If possible, the results of the list task should be compared to those of previous studies of basic colour terms to see if a term is considered a basic colour term, or where possible to the results of following task.

Biggam (2012: 43) remarks that the researcher should firstly never assume that the basic colour terms are obvious and secondly in some cases should consider the colour terms as a spectrum of vocabulary ranging from high frequency to rarity, rather than imposing strict restrictions between basic and non-basic terms.

It was hypothesised that the Turkish *lacivert* 'dark blue' does not share a type-of relationship with *mavi* 'blue', the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' are not just a kind of *sinine* 'blue' and neither is the Russian *goluboj* 'light blue' a type of *sinij* 'blue'. If *lacivert* is a separate category from *mavi*, and *helesinine* and *tumesinine* differ from *sinine*, then their quantifiable measures should be relatively similar to those of established 'blue' basic colour terms. In this case the 'blue' terms would have the same quantitative measures in all three languages. Conversely, if the Turkish *lacivert* 'dark blue', and the Estonian *helesinine* 'light blue' and *tumesinine* 'dark blue' are not separate categories, then quantifiable measures of their basic colour traits will not match those of previously established basic colour terms.

A multi-method approach was used in the form of different field tests in which data were collected in a list task to scrutinise the semantic memory, a sorting task assessed the categories of blue in the groupings of stimuli, and a naming task analysed the one-by-one naming of colour stimuli. Using the sorting task allowed the use of colours to be examined by categories rather than by colour terms. Quantifiable variables are mostly operationalised by frequencies, and the variables in the list task are frequency and mean position, or salience, in the sorting task they are the formation of separate stimuli groups, and in the naming task the variable is the agreement on naming by the participants.

Results

The Turkish participants (N=56) listed 163 colour names 978 times. Ten of the colour terms that were listed accounted for 490, or about half, of these listings, and these ten were *yesil* 'green', *siyah* 'black', *sari* 'yellow', *beyaz* 'white', *mavi* 'blue', *kirmizi* 'red', *kahverengi* 'brown', *turuncu* 'orange', *mor* 'purple', and *pembe* 'pink'. These were the terms listed most often. The least common of the basic colour terms was not in these ten most frequent terms and was *gri* 'grey', which has similar listing percentage (71%) and mean position (mp = 12.4) to *lacivert* 'dark blue' (68%, mp = 10.6), which is not an established basic colour term. It is, however, a frequent colour term. When the Turkish list terms are plotted by frequency and mean position (Figure 5), the division of terms is evident, and there are three distinct groupings. The first contains twelve high frequency and low mean position terms that correlate with the eleven Turkish basic colour terms and *lacivert* 'dark blue', a possible basic colour term; the second grouping contains the five in-between terms *lila* 'lilac', *bordo* 'burgundy', *eflatun* 'mauve', *turkuaz* 'turquoise', *bej* 'beige'; and the third is all the remaining low frequency and high mean position terms.

Turkish participants completed a list task and a naming task. In the naming task participants were asked to name coloured stimuli one by one. The stimuli were randomly shown to the participants one after another on a neutral grey background in natural daylight. The stimuli consisted of the Color-Aid coloured paper glued to a 5 × 5 × 0.2 cm piece of cardboard. A standardised set of 65 coloured tiles suggested by Davies and Corbett (1995) was used from the Color-Aid Corporation 220 Standard Set. This constrained set of stimuli was used for expedience and to allow large numbers of participants to be tested in the field. In addition to the standardised set of 65 stimuli, 17 more stimuli from the purple-blue region of colour space were selected to ascertain the status of *lacivert* 'dark blue'.

The participants gave 508 different names, and there was a large drop in frequencies after fifteen terms. These high frequency terms were *mavi* 'blue', *yesil* 'green', *mor* 'purple', *acik mavi* 'light blue', *kahverengi* 'brown', *gri* 'grey', *lacivert* 'dark blue', *koyu mavi* 'dark blue', *pembe* 'pink', *kirmizi* 'red', *acik yesil* 'light green', *turuncu* 'orange', *sari* 'yellow', *koyu yesil* 'dark green', and *siyah* 'black'. The basic colour term *beyaz* 'white' was not as frequent and was ranked twentieth by frequency. The frequency parameter is supported by the specificity index. The specificity index shows the degree of agreement between participants in naming a stimulus with a term. For example, while the overall frequency of *beyaz* 'white' was the lowest among the basic colour terms, its specificity index placed third it after *siyah* 'black' and *sari* 'yellow'. All the basic colour terms gained specificity. Alongside the eleven basic colour terms, three non-basic colour terms also had a specificity index, *acik yesil* 'light green', *lacivert* 'dark blue', and *acik mavi* 'light blue'. A term can be high frequency and low consensus, like *koyu mavi* 'dark blue' or *koyu yesil* 'dark green'. Named most frequent refers to the term with the highest frequency for each stimulus. For example, *koyu mavi* had one stimulus where it was the term given most frequently, while *koyu yesil* was the most frequent term for three stimuli. The most frequent for each stimulus can also mean low frequency when participants do not reach consensus because multiple different names are given, like *koyu pembe* 'dark pink', *lila* 'lilac', *acik pembe* 'light pink', *yavruagzi* 'peach', *acik mor* 'light purple', and *acik sari* 'light yellow'.

The naming task frequency and the specificity index for the possible basic colour term *lacivert* 'dark blue' (F = 141, SI = 0.206) were comparable with those of the basic colour term *pembe* 'pink' (F = 135, SI = 230). Several interpretations of this are possible as *lacivert* 'dark blue'can be considered a weak basic colour term or a basic colour term candidate. With or without a theory-bias, the list and naming tasks put *lacivert* 'dark blue'as a high frequency Turkish colour term.

The Estonian participants (N=39) listed 336 colour names 1145 times and 126 terms were listed by more than one participant. About half the overall frequency was provided by 27 terms and 376 of the 1145 frequencies were for eleven terms *punane* ‘red’, *sinine* ‘blue’, *kollane* ‘yellow’, *roheline* ‘green’, *hall* ‘grey’, *oranž* ‘orange’, *valge* ‘white’, *must* ‘black’, *pruun* ‘brown’, *roosa* ‘pink’, and *lilla* ‘purple’. The basic colour terms form a tight group when plotted by frequency and mean position. The basic colour term grouping is followed by two non-basic terms, *helesinine* ‘light blue’ (59%, mp = 13.8) and *beež* ‘beige’ (67%, mp = 16.7), which are between the eleven established basic colour terms and the third group of remaining terms. The third group starts with *tumesinine* ‘dark blue’ (38%), *heleroheline* ‘light green’ and *taevasinine* ‘sky blue’ (33%).

The Estonian and Russian participants completed three tasks, which were a list task, a sorting task and a naming task. The sorting task is reflective of grouping behaviour, while the naming task mirrors the names given to stimuli. The Estonian and Russian sorting and naming tasks used 55 stimuli from Color-Aid Corporation. The stimuli were specifically selected to test the blue-green-purple neighbourhood of colour space (Bimler and Uusküla 2014; Uusküla 2014). Of the 55 stimuli, 51 were non-yellowish, while four stimuli from the yellow region were distractors and functioned as a starting point during sorting. Matt-finished coloured papers from the Color-Aid Full Set of 314 colours were mounted on 5 x 5 cm square tiles. The selected stimuli contain both warm-tinged (Color-Aid code Bw) and cool-tinged (Bc) blue tones, as well as blue-greens and blue-purples, which is relevant for the Russian *goluboj* ‘light blue’, which previous research has shown to be cool-tinged. The selection was not meant to establish other basic colour terms, but to ascertain whether *helesinine* ‘light blue’ and *tumesinine* ‘dark blue’ formed separate categories from *sinine* ‘blue’, an established basic colour term. The data were analysed with multidimensional scaling and hierarchical cluster analysis.

The stimuli were all presented together in the sorting task and one-by-one in the subsequent naming task. In the free-sorting task the participants were instructed to group the selected stimuli so that ‘ones that looked similar were placed together in a way that members of a family go together’ (Roberson *et al.* 2005, 94). The data are first analysed with a qualitative examination, then with multidimensional scaling and hierarchical cluster analysis. Multidimensional scaling constructs a pictorial representation of the elements to summarise the values in a (dis)similarity matrix, so that the distances between points reflect the dissimilarities between stimuli modelled non-linearly after the rank order of dissimilarities (Woods, Fletcher and Hughes 2003: 262). The dendrogram from hierarchical cluster analysis can be imagined as a compromise across individual sorting structures.

In the Estonian sorting task 311 groups were named, with participants each creating on average eight groups, or seven groups if the yellow distractor stimuli are excluded. More than half the participants (N=22) divided the stimuli into between five and eight groups. There were also 11 splitters with a low tolerance for dissimilarity within a group, who created smaller but correspondingly more numerous groups, producing 10–29 of them, and six lumpers, who created relatively few groups, making two, three or four.

The lack of consensus perhaps reflects the high variance in the number and composition of the groups. More than half of the participants, agreed on naming one group *lilla* 'purple' (54%) and one *kollane* 'yellow' (51%). At the level of individual stimuli, these two terms became dominant for three tiles placed by at least half the participants in groups labelled *kollane* and one tile (BV) placed in groups labelled *lilla*. Subsequent terms, beginning with *sinine* 'blue' (44%), occurred to only a minority of participants as suitable names for their groups. Blue was followed by *roheline* 'green' (28%), *tumesinine* 'dark blue' (18%) and *helesinine* 'light blue' (15%) and six participants (15%) also made groups for *lillakassinine* 'purplish blue', *helelilla* 'light purple', *türkiis* 'turquoise' and *must* 'black'. *Sinine* was the most frequent name, with 267 instances of a stimulus being sorted into this group. *Sinine* also showed the widest range, with 44 stimuli receiving that name at least once. Other frequent names were *lilla* 'purple' (255), *roheline* 'green' (77) and *kollane* 'yellow' (75). Six participants labelled one of their groups as *helesinine* 'light blue', and these groups had an average size of 9.5 and ranged across 19 stimuli in total, while seven participants formed a group they labelled *tumesinine* 'dark blue', and these averaged five stimuli in size and took in 15 stimuli in their total range. The low number of participants who sorted stimuli into a group they named either *helesinine* 'light blue' or *tumesinine* 'dark blue' suggests the chances are low of them being separate categories from *sinine* 'blue'.

In the multidimensional scaling analysis plot of the Estonian stimuli with their most frequent group names, there is a distinct grouping for *lilla* 'purple', while the blue region is a continuum from lighter blue tones, including green and *helesinine* 'light blue', to darker tones and *tumesinine* 'dark blue', but no distinct grouping within the bluish region can be detected.

In the sorting task the stimuli were presented together in a pile and were grouped non-linguistically. After the participant had finished grouping the stimuli by similarity, they were asked to name the groups. This was followed by a naming task, where the stimuli were presented one-by-one in a random order and were named on a one-by-one basis.

In the Estonian naming task 2145 names were given. On average 31 different names were given by each participant. A preference for modified terms was noticed either because one stimulus was shown at a time and then named, or perhaps because the consecutive order of the tasks might have had an influence. Modifiers present in most frequent names include *hele* 'light and *tume* 'dark', and also *mustjas* 'blackish', *rohekas* 'greenish', and *sinakas* 'bluish'. The consensus per stimulus is higher than in the previous sorting task, with consensus for *helesinine* 'light blue', *sinine* 'blue', *helelilla* 'light purple', *tumelilla* 'dark purple', and *lilla* 'purple'.

The most frequent terms in the Estonian colour naming task were *helesinine* 'light blue' (F= 211), *sinine* 'blue' (F= 180), *helelilla* 'light purple' (F= 117), *lilla* 'purple' (F= 107), *tumesinine* 'dark blue' (F= 80), *kollane* 'yellow' (F= 52) and *tumelilla* ‘dark purple’ (F= 51). *Helesinine* ‘light blue’ had become marginally more salient than *sinine* ‘blue’, reversing the situation from the listing and sorting task.

If half or more of the participants agreed on naming a stimulus with a term, then the stimulus gained dominance. The dominance threshold is half the number of participants, which should be twenty (39 / 2 = 19.5), in which case the terms *tumelilla* ‘dark purple’, *helelilla* ‘light purple’, *helesinine* ‘light blue’ and *lilla* ‘purple’ were dominant. Only *lilla* ‘purple’ is a basic colour term. If the dominance threshold is lowered to nineteen then *kollane* ‘yellow’ (stimulus Y), *sinine* ‘blue’ (two stimuli, B and B-T1) and *tumesinine* ‘dark blue’ (stimulus B-S1) also become dominant. *Kollane* ‘yellow’ and *sinine* ‘blue’ are basic colour terms and with stimuli representative of the full-spectrum have been dominant at much higher levels. *Tumesinine* ‘dark blue’ is not a basic colour term and it is debatable which factors are responsible for the disparity which is caused by lowering the threshold by such a minimal amount.

Lilla ‘purple’ (SI=0.187) and *sinine* ‘blue’ (SI=0.211) had lower specificity scores than *helelilla* ‘light purple’ (SI=0.513) and *tumelilla* ‘dark purple’ (SI=0.431), and *helesinine* ‘light blue’ (SI=0.308) and *tumesinine* ‘dark blue’ (SI=0.238). *Lilla* ‘purple’ had the highest percentage of participants using the term at 90%, followed by *helelilla* ‘light purple’ (87%), *helesinine* ‘light blue’, *sinine* ‘blue’ (85%), *kollane* ‘yellow’ (74%), *tumelilla* ‘dark purple’ (69%) and *tumesinine* ‘dark blue’ (67%). *Helesinine* ‘light blue’ had the highest overall frequency (F=211), the largest range of stimuli (27) and largest average group size (6.4). *Helesinine* ‘light blue’ (range 27, average group size 6.4) and *sinine* ‘blue’ (range 24, average group size 5.5) had the largest range of stimuli and largest average group sizes because of the selection of stimuli, but this also lowered their specificity. In the multidimensional scaling analysis plot *helesinine* ‘light blue’ is the most densely grouped, while the blues and the purples are more dispersed.

Many factors made the use of modifiers prevalent in the naming task. I recommend further tests where half the participants complete the sorting task only, and the other half complete the naming task only. The selection of stimuli should include an equal number of lighter and darker shades of all the basic colour terms, and also an equivalent number of shades for beige and turquoise with some indeterminable shades as distractors.

Previous research has given the status of basic colour term to the Russian *goluboj* ‘light blue’, and so it is presumed that *goluboj* will present the strongest measurable quantities of being a separate category from *sinij* ‘blue’.

The Russian participants (N=30) listed 294 colour names 917 times. A quarter of the participants listed 25 terms, which count for 50% of the overall frequencies. Half of the participants listed fifteen terms: *krasnyj* ‘red’, *oranževyj* ‘orange’ (F=30, 100%), *sinij* ‘blue’, *žëltyj* ‘yellow’ (93%), *zelënyj* ‘green’ (90%), *belyj* ‘white’, *goluboj* ‘light blue’, *seryj* ‘grey’, (87%), *fioletovyj* ‘purple’ (83%), *čërnyj* ‘black’ (80%), *koričnevyj* ‘brown’ (73%), *rozovyj* ‘pink’ (63%), *birûzovyj* ‘turquoise’ (57%), *beževyj* ‘beige’ (53%), and *purpurnyj* ‘purple’

(50%). The latter three are high frequency terms, but not basic colour terms. Plotted by frequency and mean position, eleven colour terms are slightly separate from the rest of the data in the Russian list task, although the gap is not large between the first eleven terms and the lowest placed basic colour term *rozovyj* ‘pink’, which spearheads the rest of the terms.

The number of Russian participants (N=30) is smaller than the number of Estonian participants (N=39), and though consensus is low in both sorting tasks, it is lower in the Russian sorting task. Low consensus, or to be more precise ‘high variability’ or ‘high inter-participant differences’ in naming the sorted stimuli, is also reflected in the subsequent multidimensional scaling analysis plot because the stimulus was named one by one in the naming task..

The most frequent stimuli names were *sinij* ‘blue’, *želtyi* ‘yellow’, *fioletovyj* ‘violet’ and *goluboj* ‘light blue’. *Sinij* ‘blue’ has the highest number of participants (N=10), the highest frequency (F=124), the largest range of stimuli (39) and the largest average group size (12.4). In second place, nine participants named a group *želtyi* ‘yellow’, but it contained only the four distractor stimuli. *Fioletovyj* ‘purple’ (N=8, F=80) and *goluboj* ‘light blue’ (N=6, F=68) are also featured group names. The frequency of *goluboj* ‘light blue’ is high in comparison to the frequencies of the group names given by four participants, *birûzovyj* ‘turquoise’ (F=45), *zelënyj* ‘green’ (F=24), and *temno-sinij* ‘dark blue’ (F=18)‘light blue’.

Goluboj ‘light blue’ is the most frequent group name for 13 stimuli: BG-T4 (F=3), C-T2 (F=5), C-T4 (F=4), Bc-T3 (F=5), Bc-T4, B-T3, B-T4 (F=4), BG-LT (F=3), C-LT (F=4), BG-P1-2, BG-P2-3 (F=3), C-P1-2 (F=4), C-P2-2 (F=5). It also shares most frequent for B-P2-2 and BG-T2 with *birûzovyj* ‘turquise’ (F=2) and for Bc-T1 and Bc-T2 with *sinij* ‘blue’ (F=2).

In the multidimensional scaling analysis the Russian *goluboj* ‘light blue’ has a more defined partition from *sinij* ‘blue’. *Sinij* ‘blue’ forms two groups, the first with some *goluboj* ‘light blue’ stimuli included and the second without, but both *sinij* ‘blue’ groups are fairly spread out, forming a sort of bridge between the *goluboj* ‘light blue’ group at one end and the *fioletovyj* ‘purple’ group at the other. In the *fioletovyj* ‘purple’ and *žëltyj* ‘yellow’ clusters in the hierarchical cluster analysis, all stimuli in both groups have the highest consensus, whereas the targeted bluish stimuli show slightly less strong clustering. *Goluboj* ‘light blue’ forms a sub-cluster in which all the stimuli are sorted and named with the term named most frequent.

In the Russian naming task the participants did not reach a high degree of agreement on naming the stimuli. The significant lack of agreement in naming the stimuli in the Russian naming task presented a problem with the 50% dominance threshold, which was attained by only one stimulus B-EX *sinij* ‘blue’. The analysis of Estonian data indicated that lowering the threshold by just one point can broaden the dominant stimuli selection and possibly even add a dominant name. If the dominance threshold is lowered by one point, from 15 to 14, or 47%, then the stimulus B-S2 *temno-sinij* ‘dark blue’ (DomF = 14) is added to B-EX *sinij* ‘blue’ (DomF = 15).

The most frequent names are *sinij* 'blue' (F = 131, 26 stimuli), *goluboj* 'light blue' (F=104, 24 stimuli), *fioletovyj* 'purple' (F=59, 16 stimuli), *temno-sinij* 'dark blue' (F=57, 16 stimuli), *sirenevyj* 'lilac' (F=57, 14 stimuli), and *svetlo-goluboj* light 'light blue' (F=41, 14 stimuli). *Sirenevyj* 'lilac' is also a 'named most frequent' term, which did not have a 'named most frequent' representation in the sorting task. *Goluboj* 'light blue' was used with the modifiers *svetlo* 'light' (F=41, 14 stimuli) and *tëmno* 'dark' (F=29, 17 stimuli), which is an indicator of the strength of the basic colour term traits of the Russian *goluboj*.

In the figure for the multidimensional scaling analysis of the Russian naming task, the separation between light and dark is noticeable. The *fioletovyj* 'purple' group marks the darker shades of purple, and lighter shades are a mix of *svetlo-fioletovyj* 'light purple', and *sirenevyj* 'lilac'. The darker shades of *sinij* 'blue' are closely clustered, while *goluboj* 'light blue' is more spread out. In the hierarchical cluster analysis for the Russian naming task, two groups are quite large and may correspond to the light and dark categories, and they have several subgroups each. The darker cluster has three sub-clusters, *sinij* 'blue', a dark purple to purple grouping, and a black and dark blue to blue grouping. The lighter cluster is the largest with 27 stimuli and many different most frequent names. The frequency is generally lower than in the darker cluster. The third group, containing the mostly frequently named *sirenevyj* 'lilac' stimuli, is small and compact.

KOKKUVÕTE

Türgi, eesti ja vene keele värvisõnavara: Millised on põhivärvinimed?

Väitekirjas analüüsitakse türgi, eesti ja vene keele värvisõnavara, täpsemalt sinise värviruumi jagunemist neis keeltes. Kas keeles võib olla vaid üks sinine? Eelnevad uurimused kinnitavad, et sinine võib jaguneda kaheks: vene keeles on omaette kategooriad kahele sinise kategooriale, *sinij* 'sinine' ja *goluboj* 'helesinine'. Järelikult on ka teisi võimalusi peale üheainsa sinise kategooria, mis omakorda viitab universaalsete klasside olemasolule keeles ning toetab keelelise relatiivsuse teooriat, nn nõrka relatiivset vaadet värvisõnavarale.

Kas ka eesti sinine jaguneb alakategooriateks? Milline on eesti ja vene keele sinise kategooria võrdlus? Kas türgi keeles on samuti kaks sinise kategooriat, nagu eelnevad uurimused (vt Özgen ja Davies 1998) kinnitavad? Need on tähtsaimad küsimused, millele väitekirjas vastust otsitakse. Uurimisküsimuste analüüsiks viidi läbi välitööd: loetelu-, sorteerimis- ja nimeandmiskatse. Sorteerimis- ja nimeandmiskatses kasutati stiimuleid, mis olid spetsiaalselt valitud sinise värviruumi uurimise tarbeks.

Ülesehituselt koosneb väitekiri kahest suurest osast: teoreetilisest ja praktilisest. Teooriaosa sisaldab põhivärvinime teooria tutvustust, eelnevate uurimuste välitööde näiteid, uurimusküsimuse, meetodi, stiimulite ja keelejuhtide kirjeldust. Töö praktiline osa sisaldab türgi, eesti ja vene keele katsete tulemusi ning tulemuste analüüsi koos eesti ja vene keele katsete tulemuste võrdlusega. Väitekirja lõpuosas on lisaks esitatud kommentaarid ning ingliskeelne ja eestikeelne kokkuvõte.

Teooria tutvustus ja näited eelnevatest välitöödest

Selles väitekirjas kasutatud meetodit on kasutatud Brent Berlini ja Paul Kay nn põhivärvinimede teooria raames. Põhivärvinimede idee tulenes arvamusest, et teatud värvinimed on universaalsed. Seda teooriat võib nimetada ka universaalsete värvinimede teooriaks, sest põhivärvinimed tähistavad erinevates keeltes sarnaseid värve. Põhivärvinimede teooria loojad Brent Berlin ja Paul Kay avaldasid värvi ja keele suhete käsitluses klassikaks muutunud monograafia „Basic Color Terms: Their Universality and Evolution" (1969) justkui vastukaaluks tollal valitsenud keelelise relatiivsuse suunale.

Põhivärvinimede teooria järgi on igas keeles piiratud arv üldises kasutuses olevaid värvinimesid. Tavaliselt on keeles 2–11 põhivärvinime. Teooria loojate Berlini ja Kay originaaldefinitsiooni kohaselt (1999: 5–7) on põhivärvinimel neli esmast tunnust:

1) see on monolekseemne, st tähendus ei tulene nimetuse mõne sõnaosa tähendusest;

2) see ei tohi tähistada värvi, mida tähistab juba mingi teine värvinimi;
3) see ei tohi olla kasutatav vaid piiratud hulga objektide tähistamiseks;
4) see peab olema keelejuhtidele psühholoogiliselt silmatorkav ehk kõigile keelekasutajatele üldtuntud ja kõnealuse värvi tähistamisel esikohal.

Kui eesti *sinine* vastab kõigile neljale esmasele tunnusele, siis näiteks *helesinine* teeb seda ainult osaliselt. Esiteks on *helesinise* osiste *hele* ja *sinine* tähendus läbipaistev ning teiseks on *sinine* ülemmõiste *helesinise* suhtes, st *helesinine* on teatud tüüpi *sinine. Helesinine* on üldtuntud, kuid arvatavasti tähistab ta keelejuhtide jaoks vaid teatud osa *sinisest.* Kuigi türgi nn teine sinine – *lacivert* 'tumesinine' – ei sisalda *mavi* 'sinist', peavad keelejuhid seda teatud tüüpi siniseks ning seega ei vasta ta definitsiooni tüüpi-tunnusele (vt Özgen ja Davies 1998). Kolmas tunnus välistab piiratud kasutusega sõnad nagu *blond* või *beež*. Näiteks *beeži* kasutatakse enamasti elutute objektide tähistamiseks (vt Eessalu ja Uusküla 2013), seega on *beež* piiratud kasutusega, kuigi loetelukatse kasutussageduse ja keskmise positsiooni järgi on ta põhivärvinime kandidaat. Neljas tunnus käsitleb psühholoogilist silmatorkavust. Psühholoogiline silmatorkavus värvinimede puhul tähendab seda, et põhivärvinimed esinevad suurema osa keelejuhtide sõnavaras ning neid loetletakse värvide nimekirja alguses.

Kui värvinime põhivärvinime staatus on pärast nelja esmast tunnust ikka veel kahtluse all, siis kasutatakse Berlini ja Kay (1999: 5–7) täiendkriteeriume, mille järgi:
5) kahtlusalusel vormil peaks olema samasugune distributsioonipotentsiaal nagu leitud põhivärvinimedel;
6) kahtlased on värvinimed, mis on samas ka neile iseloomuliku värviga objektide nimed nt *vask*;
7) kahtlased võivad olla hiljutised laenud;
8) oluline on värvinime morfoloogiline komplekssus.

Kokkuvõtlikult on põhivärvinimi on kasutajatele psühholoogiliselt silmapaistev, kuid see ei ole morfoloogiliselt kompleksne (analüüsitav) ega semantiliselt läbipaistev, ka ei ole põhivärvinimi teiste nimede alammõiste ega piiratud kasutusalaga (Grossmann ja D'Achille 2016: 22).

Berlini ja Kay põhivärvinimede teooriat on aastakümnete jooksul rohkesti täiendatud ja kritiseeritud, eriti karmi kriitikatule alla on sattunud just ülalmainitud põhivärvinime tunnused. Algselt on põhivärvinimede maksimaalseks arvuks antud 11. Kuigi Kay ja McDanieli (1978) uurimus viis põhivärvinimede ülemlimiidi üle 11, on paljud uurijad jäänud kindlaks algsetele põhivärvinimede tunnustele, millele mittevastavuse korral ei loeta kahtlusalust värvinime põhivärvinimeks isegi sel juhul, kui see vastab välitööde katsete põhjal põhivärvinimele. Mylonas ja MacDonald (2015) soovitavad inglise põhivärvinimedele lisada *lilac* 'sirel[ililla]' ja *turquoise* 'türkiis', Paramei (2005) nimetab vene *goluboj*'d 'helesinine' kultuuriliselt põhivärvinimeks, türgi *lacivert* 'tumesinine' kuulub Özgeni ja Daviesi (1998) uurimuse põhjal *mavi* 'sinine' alla jne. Palju potentsiaalseid põhivärvinimesid on jäetud kõrvale esmastele tunnustele mittevastavuse tõttu. Carol Biggam (2012: 43) arvab, et mõnel juhul võiks värvinimesid vaadelda pigem skaalal kõrge esinemissagedusega kuni harvaesinevad, mitte aga kehtestada rangeid piiranguid põhivärvinimede ja mitte-põhivärvinimede vahel. Arvutuslingvistika meetodeid kasutades on Terry Regier ja Paul Kay (2003; 2007; 2009) jõudnud järeldusele, et värvinimetamist ja -tunnetust mõjutavad nii universaalsed kui keelespetsiifilised tegurid, mis ei ole nende sõnul meeltmööda ei värviuniversalistidele ega värvirelativistidele.

Värviuniversalistidest Berlini ja Kay põhivärvinimede teooria põhine kogumik „The World Color Survey" (2010) on suuremahuline ja pikaaegne koostööprojekt, kuhu on kokku kogutud 110 keele värvinimed. „The World Color Survey" (2010) ja Berlini ja Kay originaalteos „Basic Color Terms: Their Universality and Evolution" (1969) on teooria kriitikutele andnud palju materjali. Üks suurimaid kriitikaallikaid on Berlini ja Kay eeldus, igas keeles on olemas väike arv sõnu, mida kasutatakse vaid värvi tähenduses ning mis jagavad subjektiivse värviruumi osadeks (Kay *et al.* 2010: 2–3). See eeldab, et igas keeles on sõna värvi tähistamiseks, st saab küsida „Mis värvi see on?". Seda eeldust on kritiseeritud nii otsesõnu (vt Maffi (1990) ja Levinson (2000)) kui kaudselt (vt Lyons (1995), Lucy (1997), Barbara Saunders ja van Brakel (1997)). Eksperimentaalsetel alustel on kriitikat teinud enamasti antropoloogid (Hickerson 1971; Durbin 1972; Collier 1973), kuid selle teooria kõige tulisemad kriitikud on psühholoogiataustaga (Brown 1976; Miller ja Johnson-Laird 1976; Ratliff 1976).

Väitekirja teooriaosa on eelnevatest välitöödest valitud näited, mis viitavad sinise kategooria võimalikule jagunemisele mitme ülemmõistega tähistatavaks värviruumiks. Üks tähtsaimaid varasemaid uurimusi on Emre Özgeni ja Ian R. L. Daviesi (1998: 928) türgi keele värvinimede uurimus, mille järgi on türgi keeles 11 põhivärvinime. Sinise kategooria puhul on kasutusel kaks värvinime *mavi* 'sinine' ja *lacivert* 'tumesinine'. Keelejuhtide andmetel on *lacivert* 'tumesinine' teatud tüüpi *mavi* 'sinine', rikkudes sellega põhivärvinime tüüpi-tunnust, kuigi *lacivert*'i kasutus on ülekaalukas, konsensuslik ja spetsiifiline (Özgen ja Davies 1998: 919) ehk sarnane põhivärvinime kasutusele. Emre Özgen ja Ian R. L. Davies (1998) viisid läbi ajaliselt piiratud kirjaliku loetelukatse (80 last, 118 üliõpilast, 35 täiskasvanut), mille järel osa loetelukatse sooritanud keelejuhtidest (17 last, 33 täiskasvanut) osalesid ka 65 stiimuliga nimeandmiskatsel. Nende katsete esiletuleku ja konsensuse põhjal järeldasid Özgen ja Davies (1998, 919), et türgi keeles on 12 põhivärvinime, st lisaks *mavi* 'blue' kategooriale ka *lacivert* 'tumesinine', kuid kolmanda katse tulemused võimaldasid neil *lacivert*'i tumesinise põhivärvinimena välistada, kuna 57% vastanuist määratles *lacivert*'i siniste tüüpide hulka ning 86% vastanuist pidas *lacivert*'i teatud tüüpi *mavi*'ks 'sinine'. (Özgen ja Davies 1998: 942). Şahin (1998) palus keelejuhtidel (N=322) sobitada stiimulid etteantud 32 värvinimega (sh 8 põhivärvinime). Tulemuste põhjal järeldab Şahin (1998: 167), et kuigi *lacivert* (tema tõlge sellest sõnast inglise keelde on 'navy blue') on teada 99% keelejuhtidele, ei ole see põhivärvinimi. Ta nendib, et *lacivert* 'tumesinine' omandati põhivärvinimega samal kiirusel ja et see võib olla potentsiaalne põhivärvinimi (Şahin 1998: 176).

Urmas Sutropi (1995, 2000, 2002) eesti keele loetelu- ja nimeandmiskatse (N = 80, 65 Colour-Aid stiimulit) viitab üheteistkümnele põhivärvinimele: *valge*, *must*, *punane*, *kollane*, *roheline*, *sinine*, *pruun*, *hall*, *roosa*, *lilla*, ja *oranž*. *Helesinist* loetles 35% keelejuhtidest ja *tumesinist* 28%, nimeandmiskatse konsensus oli *helesinise* puhul piiripealne ning vaid ühe *tumesinise* stiimuli puhul veidi üle poole (53%). Seega on nii *helesinisel* kui *tumesinisel* põhivärvinimele omaseid jooni (Sutrop 1995: 164). Sutrop (2002: 73) viitab võimalusele, et vene sinise kategooria jagunemine *sinij* 'sinine' ja *goluboj* 'helesinine' vahel võib olla mõjutanud eesti sinise kontseptsiooni sedasi, et see jaguneb samuti kaheks või isegi kolmeks eraldi alamõisteks. Ta nendib, et eesti *sinise* kontseptsiooni puhul on vene mõju destabiliseeriv faktor (Sutrop 2002: 217).

Berlin ja Kay (1969) viitasid algselt kahele võimalikule põhivärvinime ülempiiri (≤ 11) erandile ehk vene keele *goluboj*'le 'helesinine' ja ungari *vörös*'ile 'punane', st vene keeles jagunevat kaheks sinine ja ungari keeles punane. Vene keel on üks tuntumatest näidetest keelest, milles on kaksteist põhivärvinime. See on Berlin ja Kay põhivärvinime teooria järgi erandlik, sest seal on üks põhivärvinimi rohkem kui algse teooria kohaselt oleks võimalik. Vene keele põhivärvinimedeks pakkusid Berlin ja Kay kahteteist värvinime: *белый* [18] 'valge', *черный* 'must', *красный* 'punane', *зеленый* 'roheline', *синий* 'sinine', *голубой* 'helesinine', *коричневый* 'pruun', *пурпурный* 'lilla', *розовый* 'roosa', *кирпичный* 'oranž' ja *серый* 'hall' (Berlin ja Kay 1999: 98–99). Hilisemad uuringud viitavad põhivärvinimedena *фиолетовый* 'lillale' ja *оранжевый* 'oranžile' (Frumkina ja Mikhejev 1983: 55; Corbett ja Morgan 1988: 27). Vene keele kahte sinist *sinij* 'sinine' ja *goluboj* 'helesinine' on Berlini ja Kay põhivärvinimeteooria seisukohalt uurinud mitmed autorid (vt Andrews 1994; Corbett ja Morgan 1988a; Davies ja Corbett 1994; Davies *et al.* 1998, Laws; Davies ja Andrews 1995; Morgan 1993; Morgan ja Corbett 1989; Moss *et al.* 1990; Winawer *et al.* 2007; Paramei 2005; Paramei 2005). Galina Paramei (2005) nimetab *goluboj*'d 'helesinine' „kultuuriliselt põhivärvinimeks" (inglise keeles *culturally basic colour terms*).

Slaavi keeltes on veel näiteid nn kahest sinisest, sh valgevene *sini* 'sinine' ja *blakitny* 'helesinine' ning ukraina *synij* 'sinine' ja *blakytnyj* 'helesinine' (Hippisley 2001: 168), millele Starko (2013: 150) lisab veel kolmanda sinise kategooria *holubyj*; Stanulewiczi (2010: 190) poola keele uurimus lisab põhivärvinimena *niebieski* 'sinine'. Elena Ryabina (2011b: 267) uurimusest selgub, et põhjaudmurdi *chagyr* 'helesinine' on põhivärvinimi, kuigi tegemist on arvatavasti bulgaaria laensõnaga. Ryabina väitel okupeerivad vene *goluboj* 'helesinine' ja põhjaudmurdi *chagyr* 'helesinine' värviruumis sama punkti vastates Color-Aidi stiimulile BGB T3 (Ryabina 2011a: 200; vt ka Rjabina 2011). Seesama stiimul oli dominantne ka leedu keele nimeandmiskatses, kus *žydra* 'helesinine' võib olla *mėlyna* 'sinine' kõrval põhivärvinimi (Pranaitytė 2011: 298).

---

[18] Berlin ja Kay kasutasid vanamoelist inglisepärast ladina transliteratsiooni, siin esitan erandkorras vene keele värvinimed vene tähtedega. Venekeelsed sõnad on mujal väitekirjas läbivalt translitereeritud ladina tähestikku vene-inglise transkriptsiooni järgi.

Itaalia keele mitme sinise põhivärvinimedena kasutamise puhul rõhutatakse murdealade tähtsust (Uusküla 2014; Bimler ja Uusküla 2014; Paramei, D'Orsi ja Menegaz 2014: 33; vt ka Paggetti, Menegaz ja Paramei 2015). Kuigi Mari Uusküla uurimuses (2014) pidasid itaallased *azzurro*'t ja *celeste*'t teatud tüüpi *blu*'ks 'sinine', viitab tema korpuseuuring tänapäeva itaalia keeles siiski kahele sinise põhivärvinimele – *azzurro* ja *blu*. Üle poole itaalia keelejuhtidest loetles kõiki kolme sinist *blu* (90%), *azzurro* (76%) ja *celeste* (62%), kuid nimeandmiskatse konsensus oli üle poole vaid kahel värvinimel, *blu* 'sinisel' (stiimul BVB, konsensus 54%) ja *celeste* 'helesinisel' (BGB T3 57%). Katalaani värvinimesid *blau marí* 'meresinine' ja *blau cel* 'taevasinine' peetakse põhivärvinime kandidaatideks (Davies, Corbett ja Margalef 1995: 47), mis võivad olla mõjutanud isegi ühte itaalia katalaani dialektidest, kus *celeste* tõusis hispaania katalaani 'meresinise' ja 'taevasinise' mõjul kolmanda sinise põhivärvinime kandidaadiks (Paramei, D'Orsi ja Menegaz 2014: 33).

Eespool oli itaalia keele puhul märgitud, et murded võivad mõjutada põhivärvinimede arvu, kuid näiteks pärsia keele puhul rõhutavad Kandi *et al.* (2014) regionaalseid erinevusi, mis mõjutavad nende uurimuse kohaselt põhivärvinimede arvu. Keelejuhtide põhivärvinimede hulgas oli nii puuduvaid põhivärvinimesed nagu sinine ja kollane Mashhadis, hall Esfahanis, valge ja hall Shirazis kui ka uusi lisandusi: *narwa* 'tumesinine' Esfahanis ja Mashhadis ning *gul bihī* 'küdoonia õis' Shiraziz (Kandi *et al.* 2014: 9).

Tänapäeva kreeka keeles on kaksteist põhivärvinime, sh kaks põhivärvinime sinise jaoks: *blé* 'tumesinine' ja *galázio* 'helesinine' (Androulaki *et al.* 2006: 39). Kakskeelseid kreeklasi uurides leidis Athanasopoulos (2009: 87), et kakskeelsuse taseme tõusuga kasvab inglise *blue* ja kreeka *ble* tunnetuslik lähendus nende kahe värvinime vahel.

Uurimusküsimus

Väitekirja uurimisteema on türgi, eesti ja vene värvisõnavara ning uurimusküsimus sinist tähistavate põhivärvinimede arv nendes keeltes.

Põhivärvinimi on värvi tähistav ülemmõiste, mis on enamiku keelejuhtide sõnavaras ning kasutuses. Kuid kas näiteks sinise puhul on olemas ainult üks põhivärvinimi – *sinine*? Vene keel on siinkohal näide nõrgast keelelisest relatiivsusest, kuna selles keeles on olemas kaks sinise kategooriat. Keeleline universalism aga toetab vaid ühte kategooriat. Mida näitavad aga katsetulemused?

Tuginedes eelnevatele uurimustele eeldatakse, et vene keeles on põhivärvinimed nii *sinij* 'sinine' kui *goluboj* 'helesinine', st eeldatakse, et vene *sinij* ja *goluboj* kvantitatiivsed näitajad on võrreldavad.

Uurimust motiveerib hüpotees, et eesti ja türgi keeles on rohkem kui üks sinise põhivärvinimi. Kui türgi *lacivert* 'tumesinine' on põhivärvinimi koos *mavi*'ga 'sinine' ning eesti *helesinine* ja *tumesinine* põhivärvinimed koos *sinisega*, siis nende värvinimede katsetulemused peaksid olema sarnased eelnevalt tõestatud põhivärvinimedele. Nullhüpoteesi kohaselt ei ole türgi *lacivert* 'tumesinine' ega eesti *helesinine* ja *tumesinine* eraldi kategooriad, mis tähendaks, et nende värvinimede mõõdetavad tunnused ei ole võrreldavad eelnevalt tõestatud põhivärvinimede tunnustega.

Peamine värvinimede mõõdetavate tunnuste operaator tulemuste analüüsis on sagedus. Loetelukatses mõõdetakse värvinimede sagedus, keskmine positsioon, sorteerimiskatses hinnatakse stiimuligruppide moodustamist ning nimeandmiskatses konsensust.

Katsed

Kõigis kolmes keeles viidi läbi loetelukatse, kus keelejuhtidel paluti nimetada kõik värvid, mida nad teavad. Stiimulikatsed olid türgi keele puhul erinevad. Nimeandmiskatses paluti keelejuhtidel nimetada, mis värvi on stiimul. Türgi keele nimeandmiskatses kasutati 65-t Color-Aidi värvipaberil stiimulit, mida on põhivärvinimede kindlakstegemiseks korduvalt kasutatud, ning 17 lisastiimulit (kokku 82 stiimulit). Lisastiimulid olid valitud sinisest värviskaalast. Eesti ja vene keeles viidi läbi nii sorteerimis- kui ka nimeandmiskatse. Sorteerimiskatses paluti keelejuhtidel jagada stiimulid sarnasuse alusel gruppidesse. 55 Color-Aidi stiimulit ei esindanud kogu spektrit, kõige rohkem oli stiimuleid sinisest ja lillast alast, kuid ka kollane, roheline ja must olid esindatud. Eesti ja vene keele katsed on viidud läbi täpselt sama meetodi ja stiimulitega.

Katsetel olid erinevad eesmärgid. Loetelukatse – milles keelejuhtidel paluti loetleda kõik värvid, mida nad teavad – näitab keelejuhtide mentaalset leksikoni. Sorteerimiskatse – milles keelejuhtidel paluti sorteerida stiimulid sarnasuse alusel gruppidesse – viitab värvikategooriatele. Nimeandmiskatse – milles keelejuhtidel paluti stiimuleid ühe kaupa nimetada – näitab kuidas värvinimesid kasutatakse.

Eesti ja türgi keele eelnevad uurimused (vt Sutrop 2000; Özgen ja Davies 1998) on näidanud, et teatud värvinimed, mis ei vasta mõnele põhivärvinime tunnusele, võivad siiski täita teisi põhivärvinime tunnuseid. Potentsiaalsetest nn teise sinise kategooria värvinimedest ei vasta eesti *helesinine* ja *tumesinine* kõigile põhivärvinime definitsioonile esitatavatele tunnustele, mis ei tähenda seda, et need katsetel ei või ületada põhivärvinimedele esitatud künniseid.

Loetelukatses paluti keelejuhtidel loetleda kõik värvid, mida nad teadsid. Loetelukatset saab kasutada mistahes kategooria nimede loetelu saamiseks. Ainuke piirang on ülemmõiste, nagu näiteks värv, olemasolu. Loetelukatse ei olnud ajaliselt piiratud. Kõik vastused märgiti üles nii, nagu keelejuht neid loetles, nt *veripunane* ja *verepunane* loeti erinevateks sõnadeks. Intervjueerija oli kas emakeele kõneleja või valdas antud keelt soravalt. Loetelukatses on eespool üldisema iseloomuga ja sagedasti kasutatavad sõnad. Kvantitatiivselt väljendub see loetelukatses sõna sageduses ja keskmises positsioonis. Sagedus ja keskmine positsioon on loetelukatse kaks kõige tähtsamat mõõdetavat tunnust, mida kombineerib Sutropi kognitiivse esiletuleku indeks (vt Sutrop 2001). Nende mõõdetavate tunnuste võrdlus on loetelukatse analüüsi aluseks. Mida suurem on sagedus ja madalam on keskmine positsioon, seda suurem on sõna tähtsus loetelukatses. Loetelukatse joonis, mille x-teljel on sagedus ja y-teljel keskmine positsioon, illustreerib andmestikku (vt jooniseid 3, 5, 12).

Türgi keele puhul järgnes loetelukatsele nimeandmiskatse, eesti ja vene keele puhul oli loetelukatse ja nimeandmiskatse vahel veel sorteerimiskatse. Nimeandmiskatses kasutatakse vastuste saamiseks stiimuleid. Värvipaberiga stiimulite kasutus võimaldab katset korrata ja katsetulemusi võrrelda. Stiimulikatsetes on stiimulile üheks mõõdetavaks tunnuseks antud kõige sagedasem nimi. Kõige sagedasema nime sagedus võib suures ulatuses kõikuda. Kui keelejuhtide vastused küsimusele „Mis värvi see [stiimul] on?" varieerusid suures ulatuses, siis võib stiimulile antud kõige sagedasem nimi ja seega ka konsensus olla väga madala sagedusega. Stiimulile nime andmisel ei nõustunud keelejuhid mitte kunagi sajaprotsendiliselt, st 100%-list konsensusust ei esinenud. Üldjuhul peetakse nimeandmiskatses 50% künnist ületavat konsensust piisavaks. Teisisõnu, dominantseks nimetatakse stiimulit juhul, kui üle poole keelejuhtidest olid stiimuli nimetamisel üksmeelsed. Keelejuhtide konsensuslik (≥ 50%) stiimuli nimetamine on tähtis nimeandmiskatse ja sorteerimiskatse mõõdetav tunnus. Kui üle poole keelejuhtidest nõustub stiimulile antud nime osas, siis on suure tõenäosusega tegemist põhivärvinimega või põhivärvinime kandidaadiga. Spetsiifilisusindeks on stiimuli konsensusliku nimetamise näitaja, kus stiimulile antud nime dominantne sagedus (≥ 50% konsensus) jagatakse kogusagedusega.

Sorteerimiskatse sooritasid eesti ja vene keelejuhid, kellel paluti sorteerida 55 Color-Aidi stiimulit sarnasuse alusel gruppidesse. Kui keelejuht oli stiimulite grupeerimise lõpetanud, palus intervjueerija tal gruppidele nimi anda. Keelejuhi otsustada on, millise kriteeriumi (v.a sarnasus) alusel ta grupid moodustab ning kui mitmesse gruppi ta stiimulid jagab. Weller ja Romney (1988: 25) kiidavad, et sorteerimiskatset on lihtne läbi viia, keelejuhtidele üldiselt meeldib stiimuleid gruppidesse sorteerida ja stiimuleist rääkida, ning soovitavad seda objektidevaheliste suhete uurimiseks sobiva meetodina. Sorteerimiskatse andmeanalüüsiks koostatakse tavaliselt iga keelejuhi (mitte)sarnasusemaatriks, misjärel kõigi keelejuhtide maatriksid liidetakse. 55 x 55 maatriksis said sarnasuspunkti need stiimulid, mis grupeeriti ühte gruppi ja nulli need, mida keelejuht ei grupeerinud ühte gruppi. Sorteerimiskatset on põhivärvinimede uurimiseks kasutanud vene keele puhul Frumkina (1984), kuid metoodilisest seisukohast on tähtsaim Bimleri ja Uusküla (2014) (vt ka Uusküla 2014) uurimus, kus kasutati neidsamu 55 stiimulit, mida siingi sooritatud katsetes. Seega on itaalia sorteerimiskatses kasutatud stiimulid samad, mis siin eesti ja vene sorteerimiskatses st need sorteerimiskatsed on võrreldavad. Kuna itaalia keeles ilmnes rohkem kui üks sinise kategooria, siis ei saa peamiselt sinise-lilla värvigammast (koos nelja kollaka täitestiimuliga) valitud stiimulivalikut metodoloogiliselt valeks lugeda, kuigi ka nende keelejuhtide konsensus oli sorteerimiskatses väga madal.

Sorteerimiskatse analüüsimiseks on eelkõige kasutatud mitmemõõtmelist skaleerimist (inglise keeles *multidimentional scaling* ehk MDS). Mitmemõõtmeline skaleerimine sai alguse psühholoogilisest mudelist, mis kujutas, kuidas inimesed moodustavad sarnasushinnanguid või eelisvalikuid (Borg, Groenen ja

Mair 2013: vi). Mitmemõõtmelise skaleerimise analüüsi produkt on visuaalne (Borg ja Groenen 2005: 543), mille eesmärgiks on optimaalselt kujutada lähedusmõõte (Borg, Groenen ja Mair 2013: 79), kus suured erisused (siin stiimulite vahel) väljenduvad suurte vahemaadena joonisel ning sarnasused väikeste vahemaadena (Groenen ja van de Velden 2016: 2). Sorteerimiskatse andmetes koostatakse stiimulite koosesinemisemaatriks, st keelejuhi andmetest moodustakse 55 × 55 maatriks, kus 1 esinemusmaatriksis näitab, et selle real ja veerul olevad stiimulid sorteeriti samasse gruppi ning 0, et need kaks stiimulit sorteeriti keelejuhi poolt kahte erinevasse gruppi (Borg ja Groenen 2005: 114). Keelejuhtide maatriksid liidetakse kokku, kuid selle numbrimaatriksi asemel esitab mitmemõõtmeline skaleerimine empiiriliste suhete andmeid lihtsal visuaalsel kujul (Borg, Groenen ja Mair 2013: 3). Mitmemõõtmelise skaleerimise analüüsi läbiviimiseks kasutati statistikatarkvara R.

Stiimulid

Stiimulid olid türgi nimeandmiskatses erinevad eesti ja vene katsetes kasutatuist. Türgi nimeandmiskatses kasutati 82 stiimulit Color-Aid Corporationi 220-st standardkomplektist. 65 valitud stiimulitest (vt joonis 1) peetakse põhivärvinimede välitingimustes kindlakstegemisel standardvalikusse kuuluvaiks kuuluvaiks (Davies et al. 1992; Davies ja Corbett 1995: 27), kuid lisaks kasutati 17 stiimulit sinise värviruumist BV T1, BV T2, BV S1, BVB T1, BVB T2, BVB T3, BVB S1, B T2, B T3, B T4, B S1, B S2, B S3, BG T2, COBALT BLUE, NAVY BLUE, CYAN BLUE. Kõik valitud värvipaberid kleebiti 5 × 5 × 0,2 cm suurusele vineeritükile. Eesti ja vene sorteerimis-ja nimeandmiskatses kasutatud stiimulid olid piiratuma värvivalikuga. Kasutati Color-Aid Corporationi uuemat värvipaberivalikut, mille 314-osalisest täiskomplektist valiti stiimuliteks 55 värvipaberit (vt joonis 2). Nende värviliste paberite puhul eristatakse külmi (tähistatud tähega ‘c’) ja sooje tooni (‘w’), mis on eriti olulised vene *goluboj* ‘helesinise’ puhul, kuna seda peetakse külmaks värvitooniks. David Bimler ja Mari Uusküla on kasutanud sama sinise, rohelise ja lilla värviskaalale kontsentreeritud stiimulivalikut uurimaks itaalia keele sinise kategooria jagunemist värviruumis (Bimler ja Uusküla 2014; Uusküla 2014). Bimler ja Uusküla (2016) on oma töös kasutanud minu eesti värvinimede katsete andmestikku.

Keelejuhid

Keelejuhid värvati vabatahtlikkuse alusel. Neile ei hüvitatud katses osalemist mingil moel. Kokku testiti üle 130 osaleja. Autor intervjueeris türgi ja eesti keelejuhte; vene keelejuhte küsitles vene keeles Olga Titova. Keelejuhtidelt küsitud eluloolisi andmeid nagu sugu, vanus, elukoht jt kasutati ebaisikuliselt keelejuhtide kogumi iseloomustamiseks.

Türgi andmestikust kasutati 56 keelejuhi (30 naist, 26 meest) andmeid (märts-juuli 2007), kuna 4 keelejuhti ei läbinud *City University Colour Vision Test*i, mida kasutati värvinägemise kontrolliks (Fletcher 1998; Fletcher 1984). Kõigi 60 keelejuhi loetelukatse andmete analüüsist vt Rätsep (2011: 136–139). Naissoost keelejuhtide keskmine iga oli 29 aastat ja meeskeelejuhtidel 36 aastat. Keelejuhtidest 34% moodustasid üliõpilased (vanus 20–23 aastat), 34% olid noored täiskasvanud (25–36 aastat), 27% keelejuhtidest olid üle 40-aastased ja 5% keelejuhtidest olid alla 20-aastased.

Eesti katsetes (august-november 2010) osales 41 keelejuhti, kuid analüüsis kasutati 39 (30 naiskeelejuhti, 9 meeskeelejuhti) värvipimedustesti läbinud keelejuhi andmestikku. Meeskeelejuhid (keskmine vanus 37 aastat) olid naiskeelejuhtidest (keskmine iga 46 aastat) nooremad. Enamik keelejuhtidest oli sündinud ja elas Tallinnas. Kõige sagedamini loetletud keeleoskuse hulgas olid inglise, vene, soome ja saksa keel.

Vene keele andmestiku puhul on tegemist Eestis räägitava vene keelega. Vene keelejuhte oli 30 (16 naiskeelejuhti, 14 meeskeelejuhti). Vene keelejuhtide puhul oli tegemist emigrantidega, kelle keelekasutus võib erineda Venemaal elavate keelejuhtide keelekasutusest. Neid küsitles Olga Titova 2011. aasta augustist kuni 2012. aasta veebruarini. Keskmine vanus oli meeskeelejuhtidel 44 aastat ja naiskeelejuhtidel 47 aastat. Enamik keelejuhte (19) olid omandanud kõrgema hariduse. 8 keelejuhti olid ükskeelsed vene keele rääkijad, kuid kõige sagedamini, 7 keelejuhi puhul, nimetasid nad osatavate keeltena vene, eesti ja inglise keelt. Eesti keele oskust ei loetlenud 13 keelejuhti. Eestis sündinuist olid enamik sündinud Tallinnas, kuid paljud keelejuhid olid sündinud Vene Föderatsioonis. Mõned vene keelejuhid ei olnud pärit ei Eestist ega Venemaalt.

Andmete madala konsensuse üheks põhjuseks võib pidada andmesisestust, mille kohaselt kõik liited jäid andmestikku nii, nagu keelejuhid olid neid nimetatud. Loetelukatsete tulemustest eemaldati topeltkirjed – nt türgi keelejuht nr 34 loetles kaks korda *mor*'i 'lilla' (neljandal ja neljateistkümnendal kohal). *City University Color Vision Test*i (Fletcher 1998) mitteläbinud keelejuhtide andmeid ei analüüsitud.

Tulemused

Türgi loetelukatse

Türgi keelejuhid (N=56) loetlesid värvinime 978 korda. Kümne kõige sagedamini loetletud värvinime sagedused moodustasid peaaegu poole loetelukatse kogusagedusest (490 978-st). Need värvinimed olid *yesil* 'roheline', *siyah* 'must' (loetlesid 96% keelejuhtidest), *sari* 'kollane' (95%), *beyaz* 'valge' (93%), *mavi* 'sinine', *kirmizi* 'punane' (89%), *kahverengi* 'pruun' (80%), *turuncu* 'oranž', *mor* 'lilla', and *pembe* 'roosa' (79%). Umbes kolm neljandikku loetelukatse kogusagedusest moodustasid 26 kõige enam loetletud värvinime.

Joonis 3 kujutab värvinime keskmist positsiooni x-teljel ja loeteluprotsenti y-teljel ning kõige kõrgema sagedusega värvinimed (≥ 30%) on tähistatud sildiga, sh *gri* ‘hall’, *lacivert* ‘tumesinine’ (71%), *lila* ‘lilla’ (46%), *bordo* ‘bordoo’ (43%), *eflatun* ‘lilla’ (41%), *turkuaz* ‘türkiis’, *bej* ‘beež’ (30%). Ma pean joonist 3 loetelukatse visuaalseks väljenduseks, kuna see kujutab loetelukatse kahte tähtsaimat parameetrit, sagedust ja keskmist positsiooni. Jooniselt eristuvad hästi üle poole keelejuhtide loetletud kaksteist kõrge sagedusega värvinime ülejäänud, madalama loeteluprotsendiga värvinimedest. Tabelis 2 on lisatud Sutropi kognitiivse esiletuleku indeks.

Loetelukatse tulemuste järgi on *lacivert* ‘tumesinine’ põhivärvinimi koos *mavi*’ga ‘sinine’, mis on eelnevalt tõestatud põhivärvinimi.

Türgi nimeandmiskatse

Nimeandmiskatses nimetasid türgi keelejuhid 508 värvinime vastuseks 82-le neile ükshaaval näidatud stiimulile. 284 värvinime nimetas vaid üks keelejuht.

Lacivert ‘tumesinine’ saavutas 52%-lise konsensuse, mida võib pidada madalaks või isegi piiripealseks. Kõige kõrgema konsensusprotsendiga (≥ 50% keelejuhtidest, vt tabel 4) värvinimed olid *siyah* ‘must’ (stiimul BLACK 98%, GRAY-8 70%), *sari* ‘kollane’ (Y 88%, YOY 59%), *kirmizi* ‘punane’ (RO 80%, ROR 63%), *mor* ‘lilla’ (V 77%, VRV, BV-S1* 61%, VBV 59%), *beyaz* ‘valge’ (WHITE 71%), *mavi* ‘sinine’ (B-T1 77%, BGB 64%, Cyan Blue* 61%, B 54%, Cobalt Blue* 50%), *gri* ‘hall’ (GRAY-4 68%, GRAY-6 55%, BVB-S3 54%), *turuncu* ‘oranž’ (OYO 64%, O 61%), *yesil* ‘roheline’ (GYG 61%, G 59%), *kahverengi* ‘pruun’ (R-S3 59%, O-S3 52%), *pembe* ‘roosa’ (RV-T2 55%), *lacivert* ‘tumesinine’ (B-S3* 52%), *acik mavi* ‘helesinine’ (B-T4* 50%), *acik yesil* ‘heleroheline’ (GYG-T4 50%).

Stiimulile antud kaks kõige sagedasemat nime on esitatud tabelis 4, kus on näha ka eelkõige värvinimede *acik* ‘hele’ ja *koyu* ‘tume’ konsensust vähendav mõju, kuna andmetest pole eemaldatud laiendeid ega liitsõnu üheosaliseks tehtud. See võib olla otsustava tähtsusega, sest türgi katsete tulemused viitavad *lacivert*’i ‘tumesinine’ põhivärvinime nõrgale staatusele. *Lacivert*’i ‘tumesinine’ 52%-line konsensusprotsent nimeandmiskatses on oodatust madalam, aga siiski võrreldav põhivärvinime *pembe* ‘roosa’ konsensusega (55%).

Eesti loetelukatse

Eesti keelejuhid (N=39) loetlesid 336 värvinime 1145 korda. Umbes kolmandiku kogusagedusest moodustasid 11 (põhi)värvinime sagedused *punane*, *sinine*, *kollane* (95%), *roheline* (92%), *valge*, *hall*, *oranž* (87%), *must* (85%), *pruun* (82%), *lilla*, *roosa* (79%). 88 kõige enam loetletud värvinime moodustasid umbes kolm neljandikku loetelukatse kogusagedusest. Eesti loetelukatse tulemused on esitatud joonisel 5 (vt ka tabel 6), kus põhivärvinimedele järgnevad sildistatud (≥ 30%) kõrge sagedusega värvinimed *beež* (67%), *helesinine* (59%), *tumesinine* (38%), *heleroheline*, *taevasinine* (33%). Seega on *beeži* ja *helesinise* loeteluprotsent üsna kõrge, kuid siiski väiksem madalaima loeteluprotsendiga põhivärvinimede *lilla* ja *roosa* (79%) omast.

Eesti sorteerimiskatse

Sorteerimiskatses esitati keelejuhtidele korraga kõik stiimulid, mis paluti sarnasuse alusel gruppidesse jagada ning pärast gruppidele nimi anda. Sorteerimiskatsele järgnes nimeandmiskatse, kus needsamad stiimulid esitati keelejuhtidele ükshaaval ning paluti neile üksteise järgi nimi anda.

Sorteerimiskatses osalesid nii eesti kui vene keelejuhid. Katsetes kasutati samu stiimuleid, seega on sorteerimiskatse tulemused võrreldavad. Sorteerimiskatse stiimulitele antud kõige sagedamatest nimedest annavad ülevaate joonis 6 eesti sorteerimiskatsest ja joonis 13 vene sorteerimiskatsest. Joonistel 6 ja 14 on stiimulid kantud LAB-koordinaatide abil värviruumi ning iga stiimul (v.a kollakad täitestiimulid) on sellele antud kõige sagedasema nime sildiga tähistatud. Näiteks eesti sorteerimiskatses oli vaid üks stiimul, BG P2 3, millele keelejuhid andsid kõige sagedasemaks nimeks *helesinine* (10% keelejuhtidest). Keelejuhtide arv, kes stiimulile antud grupinime andsid, kajastub stiimulit kujutava ringi suurusena, st mida suurem on stiimulit kujutav ring, seda suurem on keelejuhtide hulk, kes nimetasid stiimulit kõige sagedamini selle nimega. Näiteks vene sorteerimiskatses, kus konsensus oli madal, on kõige suurema, 30%-lise konsensusega, kõige sagedasema antud nimega stiimulid Bc, B, B T1, mis nimetati pärast sorteerimist nimega *sinij* 'sinine'. Kui keelejuhid kasutasid stiimuli nimetamisel mitmeid eri nimesid, on kõige sagedamini antud värvinimi madala konsensusega. Siiski annavad kõige sagedasemate nimede siltidega varustatud stiimulid ülevaate nimedest, mida keelejuhid katses kasutasid, eriti eesti *helesinise* ja vene *goluboj* puhul, sest *goluboj* 'helesinine' on kõige sagedasema antud nimena palju nähtavam kui eesti *helesinine* (vrd joonised 6 ja 13, vt tabel 16).

Eesti sorteerimiskatses moodustati kokku 311 gruppi. Keskmine gruppide arv keelejuhi kohta oli 8. Üle poolte keelejuhtidest (N=22) moodustasid 5–8 gruppi. Olenevalt moodustatud gruppide arvust võib ülejäänud keelejuhid jagada kas „jagajateks" (inglise keeles *splitters*) või „koondajateks" (inglise keeles *lumpers*). Katses osales 11 keelejuhti, kes kasutasid „jagamise" strateegiat moodustades palju väiksearvulisi gruppe (10–29 gruppi), kuna nende grupisisese erinevuse tolerantsus oli väike. „Koondamise" strateegiat kasutasid 6 keelejuhti, kes moodustasid 2–4 gruppi. Madal konsensus viitab tõenäoliselt suurele erinevusele gruppide arvus ja koosseisus.

54% keelejuhtidest kasutasid grupinime *lilla* ja 51% grupinime *kollane*. Stiimuli tasandil olid *kollase* puhul dominantsed kolm stiimulit neljast ja *lilla* puhul üks stiimul, BV. Ülejäänud värvinimesid (vt tabel 7) kasutasid grupinimena alla poolte keelejuhtidest, need olid näiteks *sinine* (44%), *roheline* (28%), *tumesinine* (18%), *helesinine*, *lillakassinine*, *helelilla*, *türkiis*, *must* (15%). *Sinine* oli nii kõige sagedasem (F=267) kui ka kõige suurema stiimulite arvuga (44) grupinimi. Sagedasemate grupinimede hulgas olid veel *lilla* (255), *roheline* (77), *kollane* (75).

Helesinise gruppidesse sorteeriti 6 keelejuhi poolt kokku 19 stiimulit ja keskmine grupi suurus oli 10. *Tumesinise* gruppi sorteeriti 7 keelejuhi poolt 15 stiimulit, keskmine grupi suurus oli 5 stiimulit. Mitmemõõtmelise skaleerimise joonisel (joonis 7), kus stiimulid on märgitud neile antud kõige sagedasema värvinimega, on nähtav, kuidas *lilla*-nimelised stiimulid on rohkem eraldunud *sinistest* stiimulitest kui heledamad ja tumedamad sinised üksteisest; sinise stiimulid ei ole nii tihedas grupis kui *lilla* stiimulite peamine grupeering, vaid moodustab keti heledamatest sinistest (k.a *rohelise* stiimulid) tumedamateni.

Helesinist ja *tumesinist* grupinimena kasutanud keelejuhtide väikesearvulisus ei anna toetust nende eraldiseisvusele *sinisest* kategooriast.

Eesti nimeandmiskatse

Eesti ja vene nimeandmiskatse joonised (vt joonised 9 ja 16) on palju detailsemate nimedega kui sorteerimiskatse joonistel, sest keelejuhid olid just lõpetanud nende samade stiimulite mitteverbaalse grupeerimise ja verbaalse grupinimetamise.

Eesti nimeandmiskatses nimetasid keelejuhid stiimuleid 2145 korda. Keskmiselt kasutas keelejuht 31 erinevat värvinime. Kõige sagedasemad nimed eesti nimeandmiskatses olid *helesinine* (F=211), *sinine* (F=180), *helelilla* (F=117), *lilla* (F=107), *tumesinine* (F=80), *kollane* (F=52) ja *tumelilla* (F=51). *Helesinine* edestas nimeandmiskatses sageduselt *sinist*, sorteerimiskatses aga ei olnud *helesinine* sugugi nii sage.

Kui pool või rohkem keelejuhte nimetas stiimulit sama nimega, siis see stiimul on dominantne. Dominantsuse piirmääraks on pool keelejuhtide arvust (39 / 2 = 19,5), mille järgi oleksid dominantsed värvinimed *tumelilla*, *helelilla*, *helesinine* ja *lilla*. Nendest on lihtsõna ja põhivärvinimi vaid *lilla*. Kui aga dominantsusmäära alandada 19,5-lt 19-le, siis oleksid dominantsed ka *kollane* (stiimul Y), *sinine* (B, B-T1) ja *tumesinine* (B-S1). *Kollane* ja *sinine* on põhivärvinimed, mille dominantsus 65 standardvalikus oleva stiimuliga testides oleks olnud palju kõrgem. On üllatav, et dominantsusmäära ühe arvumäära võrra alandamine teeb dominantseks *tumesinise*, mis ei ole põhivärvinimi. Võimalik, et nimeandmiskatses on tegemist täpsustamisega, mille tarbeks moodustatakse palju liitsõnu, kasutades esiosadena eriti *hele-* ja *tume-*.

Helesinisel oli kõige suurem kogusagedus (F=211), stiimulite arv (27) ja suurim keskmise grupi suurus (6,4). *Helelillal* (SI=0,513), *tumelillal* (SI=0,431), *helesinisel* (SI=0,308) ja *tumesinisel* (SI=0,238) oli spetsiifilisusindeks kõrgem *lilla* (SI=0,187) ja *sinise* (SI=0,211) spetsiifilisusindeksitest. See on ebaharilik, sest tavaliselt on nimeandmiskatses kõige suuremad spetsiifilisusindeksid põhivärvinimedel, mis üldiselt välistab liitsõnade kasutamise. *Lillat* kasutas siiski 90% keelejuhtidest, sellele järgnesid *helelilla* (87%), *helesinine, sinine* (85%), *kollane* (74%), *tumelilla* (69%) ja *tumesinine* (67%).

Hierarhiline klasteranalüüs (joonis 11, vrd joonis 8) on palju „sopilisem”, kui sorteerimiskatses kuna mitmed stiimulid moodustavad väikesearvulisi alaklastreid. Ka 90%-list variatiivsust selgitab ligi 30 erineva komponendi koosmõju, millest viis esimest komponenti seletavad vaid 52% variatiivsusest. Vaid kahel suurimal klastril, *sinisel* ja *helesinisel*, on sama, kõige sagedamini antud nimi, kuigi ka *lillal* on üsna tugev klaster, mis koosneb *helelilla*, *tumelilla* ja *lilla* alaklastritest. Nendel kolmel suuremal klastril ongi kõige kõrgemad sagedused sagedamini antud nimede puhul. Ülejäänud kaks suuremat klastrit on madalama sagedusega, üks tumedamate ja teine heledamate toonide jaoks. Viimane, heledamate toonide klaster on klastritest suurim, sisaldades 16 stiimulit, millele antud sagedasemad nimed on erinevad.

Nimeandmiskatses nimetati stiimuleid ühekaupa ning mitmemõõtmeline skaleerimine aitab eelkõige visualiseerida gruppe. Eesti nimeandmiskatse mitmemõõtmelise skaleerimise joonisel (vt joonis 10) on näha kuidas mõned kõige sagedamini *helesiniseks* nimetatud stiimulid on tihedalt grupis, kuid ülejäänud stiimulite vahelised kaugused on veidi suuremad .

Vastupidi eelnenud sorteerimiskatsele on eesti nimeandmiskatses *helesinine sinisest* esiduvam. Kuna *helelilla* ja *tumelilla* olid samuti dominantsed, siis vähendab see *helesinise* dominantsuse kaalu. Ilmselt on tumesinise piiripealne dominantsus tingitud piiratud stiimulivalikust ja katsete järjekorrast tulenevast praimingust (inglise keeles *priming*). Vene nimeandmiskatses lisanduks dominantsuse piirmäära alandamisel ühe punkti võrra (15-lt 14-le) *sinij*'le 'sinine' (stiimul B-EX 50%) samuti *temno-sinij* 'tumesinine' (stiimul B-S2 47%).

Nimeandmiskatses on laiendite domineerimisele mitu võimalikku põhjust. Nende väljaselgitamiseks võib kaaluda katset, kus praimingu vältimiseks sooritavad pool keelejuhtidest ainult sorteerimiskatse ja pool vaid nimeandmiskatse. Stiimulivalikusse tuleks kaasata *türkiisile* ja *beežile* vastavad stiimulid, kuna loetelukatsete põhjal on nende näol tegemist põhivärvinimede kandidaatidega.

Eesti katsete tulemused ei toeta *helesinise* ja *tumesinise* staatust põhivärvinimedena, põhivärvinimena jääb kindlaks *sinine*. Loetelukatses on *helesinine* psühholoogiliselt esilduv, kuid sorteerimiskatses *helesinist* ja *tumesinist* grupinimena kasutanud keelejuhtide väikesearvulisus ei toeta neid *sinisest* eraldiseisvate kategooriatena. Nimeandmiskatses on laiendite *hele* ja *tume* esiletõus täheldatav ka teiste värvinimede, eriti *lilla*, puhul. See võib viidata eelneva sorteerimiskatse praimingule, kus kasutati samu stiimuleid.

Vene loetelukatse

Eelnevatele uurimustele toetudes eeldatakse, et vene *goluboj* 'helesinise' põhinime mõõdetavad suurused on kõige võrreldavamad teiste vene keelte põhivärvinimedega ning *goluboj* on selgelt eraldi *sinij* 'sinisest' kategooriast.

Eestivene keelejuhid (N=30) loetlesid 294 värvinime 917 korda. Veerand keelejuhtidest loetles 25 värvinime, mille sagedus moodustab poole värvinimede kogusagedusest. Pool keelejuhtidest loetles järgmised 15 värvinime: *krasnyj* 'punane', *oranževyj* 'oranž' (F=30, loeteluprotsent 100%), *sinij* 'sinine', *žëltyj* 'kollane' (93%), *zelënyj* 'roheline' (90%), *belyj* 'valge', *goluboj* 'helesinine', *seryj* 'hall', (87%), *fioletovyj* 'lilla' (83%), *čërnyj* 'must' (80%), *koričnevyj* 'pruun' (73%), *rozovyj* 'roosa' (63%), *birûzovyj* 'türkiis' (57%), *beževyj* 'beež' (53%), *purpurnyj* 'purpurne' (50%). Neist esimest kahteteist peetakse vene keele põhivärvinimedeks, kuid kolm viimast – *birûzovyj* 'türkiis', *beževyj* ‘beež’ ja *purpurnyj* ‘purpurne’ – ei ole põhivärvinimed, vaid lihtsalt kõrge esinemissagedusega värvinimed loetelukatses. Loetelukatse tulemused on kujutatud joonisel 12, kus värvinimede loeteluprotsent on kujutatud x-teljel ja nende keskmine positsioon y-teljel ning värvinimed loeteluprotsendiga 30% või üle selle on sildiga tähistatud. Jooniselt 12 on näha, et esimese üheteistkümne põhivärvinime ning ülejäänud loetelukatses loetletud värvinimede sageduste vahel on väike vahe. Üheteistkümnele värvinimele, sh *goluboj*’le ‘helesinine’ järgneb põhivärvinimi *rozovyj* ‘roosa’, millele järgnevad ülejäänud värvinimed.

Vene sorteerimiskatse

Vene keelejuhtide arv (N=30) on väiksem eesti keelejuhtide arvust (N=39) ning vene sorteerimiskatse konsensus on väike, st keelejuhtide vahelised erinevused on suured.

Kõige sagedasemad stiimulinimed on *sinij* ‘sinine’, *želtyi* ‘kollane’, *fioletovyj* ‘violetne’ ja *goluboj* ‘helesinine’. *Sinij*-nimelise grupi moodustas ja nimetas kõige suurem hulk keelejuhte (N=10), sel oli kõige kõrgem sagedus (N=124), suurim stiimulite hulk (39 stiimulit 55-st) ning suurim keskmise grupi suurus (12,4). 9 keelejuhti nimetas sorteeritud grupi *žëltyj* ‘kollaseks’ kasutades selleks kõiki nelja täitestiimulit. Sagedased grupinimed olid *fioletovyj* ‘lilla’ (N=8, F=80) ja *goluboj* ‘helesinine’ (N=6, F=68). Järgnevaid grupinimesid kasutas neli keelejuhti, need nimed olid *birûzovyj* ‘türkiis’ (F=45), *zelënyj* ‘roheline’ (F=24), *temno-sinij* ‘tumesinine’ (F=18). *Goluboj* ‘helesinine’ oli kõige sagedasem grupinimi 13 stiimuli puhul: BG-T4 (F=3), C-T2 (F=5), C-T4 (F=4), Bc-T3 (F=5), Bc-T4, B-T3, B-T4 (F=4), BG-LT (F=3), C-LT (F=4), BG-P1-2, BG-P2-3 (F=3), C-P1-2 (F=4), C-P2-2 (F=5). Eesti sorteerimiskatses olid nii *helesinine* (BG P2 3) kui ka tumesinine (B S2) ühele[19] stiimulile antud kõige sagedasemad nimed.

Vene mitmemõõtmelise skaleerimise joonisel (joonis 14) on suur osa *goluboj* ‘helesinise’ stiimulitest selgesti eristunud *sinij* ‘sinise’ grupi stiimulitest, kolmanda suurema grupeeringuna on näha *fioletovyj*’d ‘lilla’. Vähene keelejuhtide üksmeel stiimuligruppidele nime andmisel ei takista *goluboj* grupi eristumist kategooriana vene sorteerimiskatse mitmemõõtmelise skaleerimise joonisel. Uurimistöö hüpoteesi vaatenurgast on vene sorteerimiskatse hierarhilise klasteranalüüsi joonisel (joonis 15) kõige tähtsam see, et seal moodustub 12-stiimuline klaster, kus on neile stiimulitele antud sagedasemaks nimeks *goluboj* ‘helesinine’.

[19] Nii *helesinise* kui *tumesinise* puhul esines stiimul, kus kõige sagedasem nimi jagunes mitme nime vahel. Stiimul BG T4 grupeeriti gruppidesse nimega *helesinine, rohekas*, *sinakasroheline*, *sinine* ning C S3 kõige sagedasemaks grupinimedeks olid *sinine*, *tumesinine*, *roheline*. *Goluboj* ‘helesinine’ jagab sagedasema värvinime tiitlit järgmiste stiimulite puhul: B-P2-2 ja BG-T2 *birûzovyj* ’türkiis’ ning stiimulite Bc-T1 ja Bc-T2 puhul *sinij* ’sinine’.

Vene nimeandmiskatse

Vene nimeandmiskatses ei olnud keelejuhid üksmeelsed, mis kõige lihtsamal kujul väljendus selles, et vaid stiimul B-EX *sinij* 'sinine' ületas 50% künnise.

Eesti andmed näitasid, et dominantsuskünnise alandamine punkti võrra võib lisada dominantseid stiimuleid, kuid vene nimeandmiskatse puhul dominantsuskünnise langetamine 47%-ni lisas dominantse stiimulina vaid B-S2 *temno-sinij* 'tumesinine'. Dominantsust arvestatakse 50%-st alates, et oleks võimalik öelda, et enamik keelejuhte kasutas teatud värvinimesid. Vene nimeandmiskatse puhul saab seda öelda vaid kahe stiimuli ja värvinime puhul, B-EX *sinij* 'sinine' (50%) ja B-S2 *temno-sinij* 'tumesinine' (47%), viimase puhul sedagi mööndustega.

Kõige sagedasemad nimed on vene nimeandmiskatses *sinij* 'sinine' (F=131, 26 stiimulit), *goluboj* 'helesinine' (F=104, 24 stiimulit), *fioletovyj* 'lilla' (F=59, 16 stiimulit), *temno-sinij* 'tumesinine' (F=57, 16 stiimulit), *sirenevyj* 'sirel[ililla]' (F=57, 14 stiimulit) ja *svetlo-goluboj* 'helesinine' (F=41, 14 stiimulit). *Goluboj*'d kasutati koos laienditega *svetlo* 'hele' (F=41, 14 stiimulit) ja *tëmno* 'tume' (F=29, 17 stiimulit), mis viitab *goluboj* põhivärvinime omadusele. Nimeandmiskatses tuli esile stiimulitele kõige sagedasema nimena *sirenevyj* 'sireli[ililla]', mis sorteerimiskatses kõige sagedasema stiimulinimena ei esinenud.

Vene mitmemõõtmelise skaleerimise joonisel (joonis 17) on stiimulite eraldus hele-tumeda ettekujutataval joonel nähtav. *Fioletovyj* 'lilla' grupp tähistab lilla tumedamaid toone ning heledamas osas on *svetlo-fioletovyj* 'helelilla' ja *sirenevyj* 'sireli[lilla]'. Värvinime *sinij* 'sinine' tumedamad toonid on tihedalt grupeerunud, kuid *goluboj* 'helesinise' sagedasema nimega stiimulid on veidi enam hajutatud.

Hierarhilise klasteranalüüsi joonisel (joonis 18) on *sirenevyj* 'sireli[lilla]' ja *svetlo-fioletovyj* 'helelilla' klaster selgelt teistest stiimulitest eristunud. Sellesse klastrisse mittekuuluvad stiimulid on jagunenud kahte suurde klastrisse, mis jagab need tumedate toonide ja heledate toonide klastriteks. Tumedate toonide suurest klastrist on eraldatav *sinij* alaklaster, kus 7 stiimulile antud kõige sagedasem nimi ongi *sinij* 'sinine'. Peale siniste alaklastri on tumedate toonide klastris veel kaks alaklastrit. Üks neist alaklastreist koosneb kas *fioletovyj* 'lilla' või *temno-fioletovyj* 'tumelilla' sagedasemate nimedega stiimulitest ning teine alaklaster on segu mustast, tumesinistest ja sinistest. Kõige suurem klaster on aga heledate toonide päralt, kus üksmeelse sagedasema nimega on *goluboj* 'helesinise' klaster.

Vene katsete tulemused toetavad *sinij* 'sinise' ja *goluboj* 'helesinise' põhivärvinime staatust, nagu eelnevad uurimusedki.

Kommentaarid

Põhivärvinime omadused

Põhivärvinime definitsiooni nõrkus ilmneb värvinimede puhul, mis on kognitiivselt ühtsed, mida tunnetatakse tervikuna ning mitte nende osadena, nt türgi *kahverengi* 'pruun' (sõna-sõnalt kohvivärvi). Eesti keele värvinimed *helesinine* ja *tumesinine* ei sobi põhivärvinimede omaduste loendisse, sest nende tähendus on tuletatav nende liitsõnade osistest, millest üks on põhivärvinimi *sinine*. Tekib küsimus, kas põhivärvinime teooria raamistikus moodustab iga põhivärvinimi omaette kategooria? Kuna põhivärvinime ühe omaduse järgi ei tohi põhivärvinimi olla ennustatav oma osadest ning juhul, kui värvinimi ei ole oma osade summa, võib see viia ringargumentatsioonini. Isiklikus kirjavahetuses tõi David Bimler näiteks soome *vaaleanpunainen* 'roosa' (sõna-sõnalt 'helepunane') – kuna ta ei ole oma osadest ennustatav, siis ta on põhivärvinimi; kui ta on põhivärvinimi, siis ta pole ennustatav oma osadest.

Laiendite eemaldamine

Kui originaalandmestikust eemaldada analüüsiks laiendid, siis peaks olema selgitus, millises ulatuses see andmestikku mõjutas, võimalusel koos näidetega. Soovituslikult peaks originaalandmestiku ja analüüsitava andmestiku erinevused olema välja toodud enne analüüsi. Kui vähegi võimalik, peaks lisama, milline on nende kahe andmestiku erinevus, seda isegi juhul, kui näiteks laiendite eemaldamine originaalandmestikus kuulub meetodi juurde või on otsustatud lugeda meetodi juurde kuuluvaks.

Originaalandmestikku võib olla enne andmeanalüüsi muutetud. Näiteks „The World Color Survey" puhul on ühe andmestiku puhul mainitud, et andmed on kokku võetud esitatud vormi (Kay *et al.* 2010: 575), mis viitab võimalusele, et teisigi andmestikke on andmeanalüüsi staadiumiks muudetud, kuid kahjuks puuduvad arvandmed muutuste ulatuse kohta.

Özgen ja Davies (1998) märgivad türgi loetelukatse tulemustes, et nad panid üheks lihtsaks sõnaks kokku kõik lihtsõnalised värvinimed, millega koos kasutati üldist laiendit *acik* 'hele' ja *koyu* 'tume'. See meetod tundub läbivat tervet artiklit, milles esitatakse ka torgi nimeandmiskatse tulemused, mille konsensusprotsent oleks kindlasti olnud madalam, kui originaalandmestikku poleks muudetud.

Autor on arvamusel, et enne originaalandmestiku muutmist tuleb see fikseerida näidete ja arvandmetena. Näiteks *acik* 'hele' (F=737) ja *koyu* 'tume' (F=565) eemaldamisel käesoleva töö türgi nimeandmiskatsest oleks dominantsete stiimulite arv tõusnud enam kui veerandi võrra. Veelgi enam, sel juhul oleks lisandunud B-S3-le (52%) veel üks dominantne stiimul (stiimul Navy Blue, 52%) värvinime *lacivert*'i 'tumesinine' puhul. See tugevdanuks *lacivert*'i põhivärvinime staatust. Üldiselt oleks originaalandmestiku lihtsasse vormi viimine tõstnud keelejuhtide loeteluprotsenti ning sorteerimiskatse ja nimeandmiskatse konsensusust.

Sorteerimiskatse strateegia

Sorteerimiskatses paluti keelejuhtidel grupeerida stiimulid nii, et sarnased stiimulid oleksid ühes grupis, st jaotatud sarnasuse alusel. Vaikides eeldati, et keelejuhid grupeerivad stiimulid neile vastavate värvinimede järgi. Loetelukatse valmistab ette sorteerimiskatses stiimulite sorteerimist värvinimede järgi, kuid keelejuhi nimevalikut ei piirata. Pealegi valitakse sorteerimisstrateegia enne nimetamist, esimest korda kogu stiimulihulka koos nähes. Keelejuhi valida on niisiis, millest gruppide moodustamisel lähtutakse – kas üldisest tumedusest ja heledusest, värvikategooriast, mis hõlmab kõige enam stiimuleid või hoopis tundest, mida stiimulid loovad jne. Sorteerimisstrateegia valikul oli mõjuvaks teguriks loetelukatse praiming, kuid visuaalse aspekti tähtsust ei tohiks alahinnata.

Katsete järjekord

Esimesena sooritatavas loetelukatses stiimuleid ei kasutata – see on mõeldud värvisõnavara kindlakstegemiseks. Sorteerimis- ja nimeandmiskatseid tehes sooritavad keelejuhid need tihti üksteise järel, mis tingib sorteerimiskatse järel tehtava nimeandmiskatse suhtes praimingu. Logistilisest vaatenurgast on see paratamatu, sest välitingimustes on loogiline, et katsesse kaasatakse nii palju keelejuhte, kui võimalik, ning keelejuhid sooritavad nii palju erinevaid katseid, kui ajaliselt võimalik. Ideaalne oleks aga variant, kus keelejuht sooritab vaid ühe stiimulikatse. Kuigi sorteerimiskatses ja nimeandmiskatses kasutatakse tihti samu stiimuleid, on need katsed siiski erinevad, kuna sorteerimiskatses uuritakse sarnaste olemite, nt värvistiimulite, paberile trükitud sõnade jms kategoriseerimist mitteverbaalsel kujul, mis mõnel juhul avaldub verbaalselt alles pärast keelejuhi sorteerimistegevuse lõppu. Nimeandmiskatses aga uuritakse ühe stiimuli kaupa, kui konsensuslikult keelejuhid nimetavad objekte. Seega võiks võimalusel keskenduda kas sorteerimis-või nimeandmiskatsele, kuid mitte mõlemale, et praimingu võimalus üldse välistada.

Kui pärast sorteerimiskatset tehtava nimeandmiskatse puhul on praimingu oht, siis selle praimingu efekti ulatuse teadasaamiseks võiks teha katse, kus pooled keelejuhtidest sooritavad enne sorteerimiskatse ja pärast nimeandmiskatse ning pooled keelejuhtidest sooritavad nimeandmiskatse enne sorteerimiskatset.

Ka ei tohiks tähelepanuta jätta fakti, et värvistiimulitega sorteerimiskatses ei ole keelejuhid informeeritud sellest, et nad peavad sorteeritud stiimuligruppe pärast nimetama. Ühe keelejuhi sõnul oleks antud vastused olnud üsna teistsugused, kui teda oleks enne informeeritud sellest, et sorteeritud gruppidele tuleb pärast ka nimetus anda. Sellest lähtuvalt oleks edasine uurimus vajalik. Näiteks võiks jagada grupi keelejuhte pooleks ning pooltele öelda enne sorteerimiskatse algust, et sorteeritud grupid tuleb nimetada ning teistele mitte. Leian, et oleks huvitav teada, kas ja kui palju nimetamisest teadmine enne sorteerimist mõjutab gruppide moodustamist ja nimetamist.

Stiimulivalik

Daviesi ja Corbetti (1995: 27) algselt valitud 65 stiimulit katavad värviruumi hõredalt, kuid ühtlaselt (vt joonis 1). Androulaki *et al.* (2006: 27), kes uurisid kreeka *γalázjo* 'helesinise' kategooriat, aga leidsid, et Color-Aid stiimulid ei kata värviruumi ühtlaselt, jättes just nimelt helesinise alaesindatuks. Autorid on veendunud, et nendel stiimulitega katmata aladel olid head *γalázjo* 'helesinise' näited ning spekuleerivad, et *γalázjo* kasutus oleks olnud sagedasem

(Androulaki *et al.* 2006: 27). Helesinise sobivaks näiteks tundub aga olevat Color-Aidi stiimul BGB T3. Ryabina (2011a: 200; ka Rjabina 2011) uurimuses vastasid nii vene *goluboj* kui põhjaudmurdi *chagyr* 'helesinine' just sellele stiimulile. Leedu uurimus viitab samuti stiimulile BGB T3, mille puhul oli *žydra* 'helesinine' dominantne koos *mėlyna* 'sinisega' (Pranaitytė 2011: 298).

55-stiimuliline valik eesti ja vene sorteerimiskatses on kallutatud sinise-lilla-rohelise värviruumi suunas, millele on lisatud neli kollakat täitestiimulit. Võimalus on olemas, et täitestiimulite suhe stiimulitega on liiga väike, 51 : 4. Spekuleerida võib selle üle, kas täitestiimulid tekitasid tugeva isoleeriva efekti, mille ilmnemine ei olnud taotluslik. Sorteerimis- ja nimeandmiskatse stiimulid olidki mõeldud just sinise värviala uurimiseks.

Tulemused peegeldasid stiimulivalikut, mis tähendab, et jääb võimalus, et stiimulivalik sundis keelejuhte grupeerima mittesiniseid stiimuleid, nt *lilla* ja *kollane*, konsensuslikumalt kui siniseid stiimuleid, mille nimetamine muutus nimeandmiskatses eriti spetsiifiliseks, kuna siniseid stiimuleid oli proportsionaalselt teistest rohkem. Isegi stiimulivalikus metodoloogilist viga arvestades ei välista see mitmete siniste kategooriate moodustumist, mida tõestas itaalia keele puhul sama stiimulivalikuga Uusküla (2014), kus avaldus mitu sinise kategooriat, *blu*, *azzurro* ja *celeste*, kuigi ka nende konsensus oli sorteerimiskatses madal.

Dominantsuskünnis

Laiendite kõrvalejätmine on üks faktoritest, mis tõstab üldist konsensustaset. Kui andmestik viitab selle vajalikkusele võib teise võimalusena kaaluda dominantsuskünnise minimaalset alandamist nt ühe keelejuhi arvu võrra. Vene katses tähendas see näiteks dominantsuskünnise alandamist 47%-ni, mis oli ka tingitud üleüldisest madalast konsensusest. Eesti andmestiku puhul on 50% lävendiks 39 / 2=19,5 ehk 20. Kui vähendada lävend 20-lt 19-ni, on selle mõju tuntav eriti nimeandmiskatses, kus dominantse nimena lisandub *tumesinine* (stiimuli B-S1); ka vene nimeandmiskatses lisandub lävendi alandamisel *temno-sinij* 'tumesinine' (stiimul B-S2). Seega dominantsuskünnise alandamine punkti võrra lisab dominantsete nimede hulka nii eesti kui vene nimeandmiskatses tumesinine. Tegemist on ilmselgelt hüpoteesist kallutatud valikuga, mis sellegipoolest näitab, et dominantsuspiiri nihutamine võib tuua kaasa suuremaid või väiksemaid erinevusi, mida tasuks arvesse võtta. Näiteks nimeandmiskatses võiks üle kontrollida, kas mõned stiimulid ja stiimulinimed on dominantsuse piiri peal, ning sel juhul need ära märkida, sest madalama konsensusetaseme puhul võib tegemist olla põhivärvinime kandidaatidega.

Võtmesõnad: põhivärvinimed, välitöö, *sinine*, *helesinine*, *tumesinine*. eesti keel, türgi keel, vene keel

BIBLIOGRAPHY

Adler, Daniel, Duncan Murdoch, and others. 2014. Rgl: 3D visualization device system (OpenGL). Available online at http://CRAN.R-project.org/package=rgl, accessed on April 19, 2018.

Andrews, David R. 1994. “The Russian color categories sinij and goluboj: an experimental analysis of their interpretation in the standard and émigré languages.” Journal of Slavic Linguistics 2 (1): 9–28.

Androulaki, Anna, Natalia Gômez-Pestaña, Christos Mitsakis, Julio Lillo Jover, Kenny Coventry, and Ian Davies. 2006. “Basic colour terms in Modern Greek: twelve terms including two blues.” Journal of Greek Linguistics 7 (1): 3–47. Available online at http://booksandjournals.brillonline.com/content/journals/10.1075/jgl.7.03and, accessed on April 19, 2018.

Arabie, Phipps, and Scott A. Boorman. 1973. “Multidimensional scaling of measures of distance between partitions.” Journal of Mathematical Psychology 10 (2): 148–203. Available online at https://doi.org/10.1016/0022-2496(73)90012-6, accessed on April 19, 2018.

Archer, Dawn (ed.). 2009. What’s in a word-list?: investigating word frequency and keyword extraction (Digital Research in the Arts and Humanities). Farnham, England; Burlington, VT: Ashgate Pub. Ltd.

Athanasopoulos, Panos. 2009. “Cognitive representation of colour in bilinguals: the case of Greek blues.” Bilingualism: Language and Cognition 12 (01): 83–95. Available online at https://doi.org/10.1017/S136672890800388X, accessed on April 19, 2018.

Barbur, J. L., A. J. Harlow, and G. T. Plant. 1994. “Insights into the different exploits of colour in the visual cortex.” Proceedings of the Royal Society of London B: Biological Sciences 258 (1353): 327–334. Available online at https://pdfs.semanticscholar.org/f905/d61c4491ad47323aedf02ee9b30ecced9a26.pdf, accessed on April 19, 2018.

Benczes, Réka, and Erzsébet Tóth-Czifra. 2014. “The Hungarian colour terms piros and vörös: a corpus and cognitive linguistic account.” Acta Linguistica Hungarica 61 (2): 123–152. Available online at https://doi.org/10.1556/ALing.61.2014.2.1, accessed on April 19, 2018.

Berlin, Brent, and Paul Kay. 1969 [1999]. Basic color terms: their universality and evolution. The David Hume Series. Stanford, Calif: Center for the Study of Language and Information.

Bezmez, Serap, Richard Blakney, C. H Brown & Nilüfer Aydin. 2001. Redhouse büyük elsözlügü: İngilizce -Türkce, Türkce -İngilizce = The larger Redhouse portable dictionary : English-Turkish, Turkish-English. İstanbul: Redhouse Yayinevi.

Biggam, C. P. (ed.). 2011. New directions in colour studies. Amsterdam ; Philadelphia: John Benjamins Pub. Co.

Biggam, C. P., Christian Kay, and Nicola Pitchford (eds.). 2006. Progress in colour studies. Amsterdam; Philadelphia: J. Benjamins Pub.

Biggam, Carole Patricia. 2012. The Semantics of Colour: A Historical Approach. Cambridge University Press.

Bimler, David and Mari Uusküla. 2016. “A similarity-based cross-language comparison of basicness and demarcation of “blue” terms”. Color Research & Application. Available online at http://doi.wiley.com/10.1002/col.22076, accessed on August 22, 2016).

Bimler, David and Mari Uusküla. 2018. "A Points-of-View analysis of colour listing data from fourteen European languages". In Progress in colour studies: cognition, language and beyond, edited by L. W. MacDonald, C. P. Biggam and Galina V. Paramei, 223–239. Amsterdam; Philadelphia: John Benjamins Publishing Company.
Bimler, David L., John Kirkland, and Mari Uusküla. 2015. "Applying points-of-view analysis to individual variations in colour sorting data." Journal of Cognition & Culture 15 (1/2): 87–108.
Bimler, David, and Mari Uusküla. 2014. "'Clothed in triple blues': sorting out the Italian blues." Journal of the Optical Society of America 31 (4): 332–340. Available online at https://doi.org/10.1364/JOSAA.31.00A332, accessed on April 19, 2018.
Bimler, David. 2011. "Universal trends and specific deviations: multidimensional scaling of colour terms from the world color survey." In New directions in colour studies, edited by Carole P. Biggam, Carole Hough, Christian Kay, and David R. Simmons, 13–26. Amsterdam: John Benjamins Publishing Company. Available online at https://benjamins.com/catalog/z.167.05bim, accessed on April 19, 2018.
Blokland, Rogier. 2009. The Russian loanwords in literary Estonian. Veröffentlichungen Der Societas Uralo-Altaica, Bd. 78. Wiesbaden: Harrassowitz.
Bogatkin-Uusküla, Mari and Urmas Sutrop. 2007. Are there two basic terms piros and vörös for red in Hungarian? Emakeele Seltsi aastaraamat nr 50, 93–110.
Bolton, Ralph. 1978. "Black, white, and red all over: the riddle of color term salience." Ethnology 17 (3): 287–311. Available online at https://doi.org/10.2307/3773198, accessed on April 19, 2018.
Boorman, S.A. & P. Arabie. 1972. Structural measures and the method of sorting. In Roger N. Shepard (ed.), Multidimensional scaling: theory and applications in the behavioral sciences, vol. 1. New York: Seminar Press.
Boorman, Scott A., and Donald C. Olivier. 1973. "Metrics on spaces of finite trees." Journal of Mathematical Psychology 10 (1): 26–59. Available online at https://doi.org/10.1016/0022-2496(73)90003-5, accessed on April 19, 2018.
Borg, Alexander. 1999. The language of color in the Mediterranean: an anthology on linguistic and ethnographic aspects of color terms. Acta Universitatis Stockholmiensis 16. Stockholm: Almqvist & Wiksell International.
Borg, Alexander. 2007. "Towards a history and typology of color categorization in colloquial Arabic." In Anthropology of color, edited by Robert E. MacLaury, Galina V. Paramei, and Don Dedrick, 263–293. Amsterdam: John Benjamins Publishing Company. Available online at https://doi.org/10.1075/z.137.17bor, accessed on April 19, 2018.
Borg, Ingwer, and Patrick J. F. Groenen. 2005. Modern multidimensional scaling: theory and applications. 2nd ed. Springer Series in Statistics. New York: Springer.
Borg, Ingwer, Patrick J. F Groenen and Patrick Mair. 2013. Applied multidimensional scaling. Berlin; London: Springer.
Boster, J.S. 1986. "Can individuals recapitulate the evolutionary development of color lexicons?" Ethnology 25 (1): 61–74.
Brown, R. W. 1976. "Reference." Cognition 4: 125–153.
Bulakh, Maria. 2007. "Basic color terms from Proto-Semitic to Old Ethiopic." In Anthropology of color, edited by Robert E. MacLaury, Galina V. Paramei, and Don Dedrick, 247–261. Amsterdam: John Benjamins Publishing Company. Available online at https://doi.org/10.1075/z.137.16bul, accessed on April 19, 2018.
Burton, Michael L. 1975. "Dissimilarity measures for unconstrained sorting data." Multivariate Behavioral Research 10 (4): 409–423.

Chang, Winston. 2013. R Graphics Cookbook. O'Reilly Media. Available online at http://www.cookbook-r.com, accessed on April 19, 2018.
Collier, George A. 1973. "Review of basic color terms." Language 49: 245–48.
Cooke, Nancy J. 1994. "Varieties of knowledge elicitation techniques." International Journal of Human-Computer Studies 41: 801–849.
Corbett, Greville, and Gerry Morgan. 1988. "Colour terms in Russian: reflections of typological constraints in a single language." Journal of Linguistics 24 (1): 31–64. Available online at http://epubs.surrey.ac.uk/1330/, accessed on April 19, 2018.
Davies, I. R. L., C. Macdermid, G. G. Corbett, H. Mcgurk, D. Jerrett, T. Jerrett, and P. Sowden. 1992. "Color terms in Setswana: a linguistic and perceptual Approach." Linguistics 30 (6): 1065–1103. Available online at http://epubs.surrey.ac.uk/44140/, accessed on April 19, 2018.
Davies, Ian R. L., Greville G. Corbett, Harry McGurk, and Catriona MacDermid. 1998. "A developmental study of the acquisition of Russian colour terms." Journal of Child Language 25 (02): 395–417.
Davies, Ian R. L., Penny Roling, Greville G. Corbett, Fritz Xoagub and Jomo Xoagub. 1997. Color terms and color term acquisition in Damara. Journal of Linguistic Anthropology 7 (2): 181–207. Available online at http://doi.wiley.com/10.1525/jlin.1997.7.2.181, accessed on March 9, 2017).
Davies, Ian, and Greville Corbett. 1994. "The basic color terms of Russian." Linguistics 32: 65–89.
Davies, IanR L., and Greville G. Corbett. 1995. "A practical field method for identifying probable basic colour terms." Languages of the World 9 (1): 25–36. Available online at http://epubs.surrey.ac.uk/2090/, accessed on April 19, 2018.
Davies, IanR L., Greville G. Corbett, and José B. Margalef. 1995. "Colour terms in Catalan: an investigation of eighty informants, concentrating on the purple and blue regions." Transactions of the Philological Society 93: 17–49. Available online at http://epubs.surrey.ac.uk/1344/, accessed on April 19, 2018.
De Leeuw, Jan and Willem Heiser. 1980. "Multidimensional scaling with restrictions on the configuration". Multivariate Analysis 5 (1): 501–522. Available online at https://www.researchgate.net/publication/265360786_Multidimensional_scaling_with_restrictions_on_the_configuration, accessed on October 24, 2018).
Dedrick, Don. 1996. "Color language universality and evolution: on the explanation for basic color terms." Philosophical psychology 9 (4): 497–524. Available online at http://www.tandfonline.com/doi/abs/10.1080/09515089608573197, accessed on April 19, 2018.
Devellioğlu, Ferit, and Aydin Sami Güneyçal, eds. 2006. Osmanlica-Türkçe ansiklopedik lûgat: eski ve yeni harflerle. 23. baski. Sözlük Dizisi 1. Ankara: Aydin Kitabevi.
Durbin, M. 1972. "Basic cerms – off-color?" Semiotica 6: 257–278.
Engchuan, S. 2003. Color terms and the concept of color of the Thais in Sukhotai period and at the present time. Bangkok: Chulalongkorn University.
Fletcher, R. 1984. "The City University colour vision test." Ophthalmic and Physiological Optics 4 (3): 279–279. Available online at https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-1313.1984.tb00367.x, accessed on April 19, 2018.
Fletcher, Robert. 1998. The City University Colour Vision Test. Third edition. London: Keeler.

Fox, John, and Sanford Weisberg. 2011. An R companion to applied regression. 2nd ed. Thousand Oaks CA: Sage. Available online at http://socserv.socsci.mcmaster.ca/jfox/Books/Companion, accessed on April 19, 2018.

Franklin, Anna, and Ian R.L. Davies. 2004. "New evidence for infant colour categories." British Journal of Developmental Psychology 22 (3): 349–377. Available online at https://doi.org/10.1348/0261510041552738, accessed on April 19, 2018.

Frenzel-Biamonti, Claudia. 2011. "Rosa schätze – pink zum kaufen: stylistic confusion, subjective perception and semantic uncertainty of a loaned colour term." In New directions in colour studies, edited by Carole P. Biggam, Carole Hough, Christian Kay, and David R. Simmons, 91–104. Amsterdam: John Benjamins Publishing Company. Available online at https://benjamins.com/catalog/z.167.13fre, accessed on April 19, 2018.

Frumkina, R.M. 1984. Cvet, smysl, sxodstvo: aspekty psixolingvistic eskogo analiza [Coulor, meaning, and similarity: aspects of psycholinguistic analysis]. Moscow: Nauka.

Frumkina, R.M. and A.V. Mikhejev. 1983. Vozmožnosti sopostavitel'nogo izučenija leksiki v èksperimente (na materiale russkix prilagatel'nyxcvetooboznačenij) [The possibilities for a comparative study of vocabulary in an experiment (on Russian adjectives—color terms)]. Sapostavitelno Jezykoznanie 8, 51–63.

Goddard, Cliff, and Anna Wierzbicka, eds. 1994. Semantic and lexical universals: theory and empirical findings. Studies in Language Companion Series (SLCS), v. 25. Amsterdam; Philadelphia: J. Benjamins.

Gries, Stefan Thomas. 2013. Statistics for linguistics with R: a practical introduction. 2nd revised edition. Berlin: De Gruyter Mouton.

Groenen, Patrick J. F., and Michel van de Velden. 2016. "Multidimensional scaling by majorization: a review." Journal of Statistical Software 73 (8). Available online at https://doi.org/10.18637/jss.v073.i08, accessed on April 19, 2018.

Grossmann, Maria, and Paolo D'Achille. 2016. "Italian colour terms in the BLUE area: synchrony and diachrony." In Colour and colour naming: crosslinguistic approaches, edited by João Paulo Silvestre, Esperança Maria da Cruz Marreiros Cardeira, and Alina Villalva. Lisboa: Centro de Linguística da Universade da Lisboa : Universade de Aveiro.

Güncel Türkçe Sözlük. n.d. Türk Dil Kurumu Yayınları. Available online at http://www.tdk.gov.tr/index.php?option=com_gts&view=gts, accessed on April 19, 2018.

Harkness, Sara. 1973. "Universal aspects of learning color codes: a study in two cultures." Ethos 1: 175–200.

Heider, Eleanor R. 1971. "'Focal' color areas and the development of color names." Developmental Psychology 4 (3): 447–455.

Heider, Eleanor R. 1972. "Universals in color naming and memory." Journal of Experimental Psychology 93 (1): 10–20.

Heider, Eleanor Rosch. 1972. "Probabilities, sampling, and ethnographic method: the case of Dani colour names." Man, New Series, 7 (3): 448–466. Available online at https://doi.org/10.2307/2800917, accessed on April 19, 2018.

Hickerson, Nancy P. 1971. "Review of Berlin & Kay (1969)." International Journal of American Linguistics 37: 257–270.

Hint, Mati. 1990. "Vene keele mõjud eesti keelele." Akadeemia 2: 1383–1404.

Hint, Mati. 1991. “The changing language situation: Russian influences in contemporary Estonian.” Journal of Multilingual and Multicultural Development 12 (1–2): 111–116.

Hippisley, Andrew. 2001. “Basic BLUE in East Slavonic.” Linguistics Faculty Publications, 39 (1): 151–179. Available online at https://doi.org/10.1515/ling.2001.003, accessed on April 19, 2018.

Hollman, Liivi. 2010. “Basic color terms in Estonian sign language.” Doctoral thesis. Available online at http://dspace.utlib.ee/dspace/handle/10062/14769, accessed on April 19, 2018.

Irwin, Eleanor. 1974. Colour terms in Greek poetry. Toronto: Hakkert.

Jameson, Kimberly, and Roy G. D’Andrade. 1997. “It’s not really red, green, yellow, blue: an inquiry into perceptual color space.” In Color categories in thought and language. Cambridge University Press. Available online at https://www.cambridge.org/core/books/color-categories-in-thought-and-language/its-not-really-red-green-yellow-blue-an-inquiry-into-perceptual-color-space/D633A5960C9130A0F1C0DBD5A2C68238, accessed on April 19, 2018.

Jotov, Marin I., and Nina N. Ponomareva. 1959. Bolgarsko-russkij slovar’: okolo 29.000 slov. Moscow: Gosud. izd. inostr. i nacional’nych slovarej.

Jraissati, Yasmina, Coralie Chevallier, and Nicolas Claidière. 2008. “A colour sorting task reveals the limits of the universalist/relativist dichotomy: colour categories can be both language specific and perceptual.” Journal of Cognition and Culture 8 (3): 211–233. Available online at https://doi.org/10.1163/156853708X358155, accessed on April 19, 2018.

Kabacoff, Robert. 2015. R in action: data analysis and graphics with R. Second Edition. Manning Publications. Available online at https://www.manning.com/books/r-in-action-second-edition, accessed on April 19, 2018.

Kalda, Triin. 2014. “Eestlaste, soomlaste ja soomes elavate eestlaste värvileksikoni võrdlev analüüs.” Magistritöö, Tallinna Ülikool.

Kandi, S. Gorji, M. Amani Tehran, N. Hassani, and A. Jarrahi. 2014. “Color naming for the Persian language.” Color Research & Application, 40 (4): 352: 360. Available online at https://doi.org/10.1002/col.21887, accessed on April 19, 2018.

Kann, Kallista, and Lauri Leesi, eds. 1999. Prantsuse-eesti sõnaraamat =: Dictionnaire français-estonien. 4. trükk. Tallinn: Valgus.

Katemake, Pichayada, Dimitris Mylonas, Lindsay Macdonald, and Amra Prasithrathsint. 2015. “Comparison among three methods for Thai colour naming.” AIC Midterm Meeting. Tokyo, Japan, May 2015. Available online at https://www.researchgate.net/publication/282124349_Comparison_Among_Three_Methods_for_Thai_Colour_Naming, accessed on April 19, 2018.

Kay, Paul, and Chad K. McDaniel. 1978. “The linguistic significance of the meanings of basic color terms.” Language 54 (3): 610–646. Available online at http://muse.jhu.edu/journals/language/v054/54.3.kay.html, accessed on April 19, 2018.

Kay, Paul, and Terry Regier. 2003. “Resolving the question of color naming universals.” Proceedings of the National Academy of Sciences 100 (15): 9085–9089. Available online at https://doi.org/10.1073/pnas.1532837100, accessed on April 19, 2018.

Kay, Paul, and Terry Regier. 2007. “Color naming universals: the case of Berinmo.” Cognition 102 (2): 289–298. Available online at https://doi.org/10.1016/j.cognition.2005.12.008, accessed on April 19, 2018.

Kay, Paul, and Willett Kempton. 1984. "What is the Sapir-Whorf hypothesis?" American Anthropologist, 65–79. Available online at http://www.jstor.org/stable/679389, accessed on April 19, 2018.

Kay, Paul, Brent Berlin, and William Merrifield. 1991. "Biocultural implications of systems of color naming." Journal of Linguistic Anthropology 1 (1): 12–25. Available online at https://doi.org/10.1525/jlin.1991.1.1.12, accessed on April 19, 2018.

Kay, Paul, Brent Berlin, Luisa Maffi, William Merrifield, and Richard Cook. 2010. The world color survey. Stanford, California: Center for the Study of Language and Information Publications.

Kay, Paul. 1975. "Synchronic variability and diachronic change in basic color terms." Language in society 4 (3): 257–270. Available online at http://www.jstor.org/stable/4166830, accessed on April 19, 2018.

Kerttula, Seija. 2002. English colour terms: etymology, chronology, and relative basicness. Société néophilologique.

Kerttula, Seija. 2007. Relative basicness of color terms: Modeling and measurement. In Robert E. MacLaury, Galina V. Paramei & Don Dedrick (eds.), Anthropology of Color, 151–169. Amsterdam: John Benjamins Publishing Company. Available online at https://benjamins.com/catalog/z.137.11ker, accessed on May 27, 2018.

Khomeriki, Manana, Archil Kezeli, and Natela Lomashvili. 2009. "The revealing of color categorical boundaries among Georgian speakers." Bull. Georg. Natl. Acad. Sci 3 (3). Available online at http://science.org.ge/old/moambe/3-3/Khomeriki.pdf, accessed on April 19, 2018.

Kopecka, Anetta and Katarzyna Janic. 2016. Dénomination des couleurs artistiques et décoratives en polonais. In Wiesław Banyś, Katarzyna Kwapisz-Osadnik and Fabrice Marsac (eds.), Neophilologica, Vol. 28, 122–146. Katowice: Uniwersytet Śląski. Available online from https://libra.ibuk.pl/book/178049, accessed on June 18, 2018.

Kornfilt, Jaklin. 2010. Turkish (Descriptive Grammars). Transferred to digital print. 2010. London: Routledge.

Kristol, Andres M. 1980. "Color systems in Southern Italy: a case of regression." Language 56 (1): 137–145. Available online at https://doi.org/10.2307/412646, accessed on April 19, 2018.

Langemets, Margit, ed. 2009. Eesti keele seletav sõnaraamat. 2., täiendatud ja parandatud trükk. Tallinn: Eesti Keele Sihtasutus. Available online at http://www.eki.ee/dict/ekss/, accessed on April 19, 2018.

Laws, Glynis, Ian Davies, and Catherine Andrews. 1995. "Linguistic structure and non-linguistic cognition: English and Russian blues compared." Language and Cognitive Processes 10 (1): 59–94. Available online at https://doi.org/10.1080/01690969508407088, accessed on April 19, 2018.

Leeuw, Jan de, and Patrick Mair. 2009. "Multidimensional scaling using majorization: SMACOF in R." Journal of Statistical Software 31 (3): 1–30. Available online at http://www.jstatsoft.org/v31/i03/, accessed on April 19, 2018.

Lenneberg, Eric H., and John M. Roberts. 1956. The language of experience: a study in methodology. Indiana University Publications in Anthropology and Linguistics 13. Baltimore: Waverly Press.

Levinson, Stephen C. 2000. "Yélî Dnye and the theory of basic color terms." Journal of Linguistic Anthropology 10 (1): 3–55. Available online at https://anthrosource.onlinelibrary.wiley.com Available online at /doi/abs/10.1525/jlin.2000.10.1.3, accessed on April 19, 2018.

Lillo, Julio, Fernando González-Perilli, Lilia Prado-León, Anna Melnikova, Leticia Álvaro, José A. Collado & Humberto Moreira. 2018. Basic Color Terms (BCTs) and Categories (BCCs) in Three Dialects of the Spanish Language: Interaction Between Cultural and Universal Factors. Frontiers in Psychology 9. Available online at https://www.frontiersin.org/article/10.3389/fpsyg.2018.00761/full, accessed on May 26, 2018).

Lin, H., M. R. Luo, L. W. MacDonald and A. W. S. Tarrant. 2001. A cross-cultural colour-naming study. Part I: Using an unconstrained method. *Color Research & Application* 26 (1): 40–60. Available online at http://doi.wiley.com/10.1002/1520-6378%28200102%2926%3A1%3C40%3A%3AAID-COL5%3E3.0.CO%3B2-X, accessed on August 11, 2015).

Lucy, John A. 1997. "The linguistics of 'color.'" In Color categories in thought and language 320–346, edited by C. L. Hardin and Luisa Maffi. Cambridge: Cambridge University Press.

Lyons. 1995. "Color in language." In Colour: art & science, edited by Trevor Lamb and Janine Bourriau, 194–224. Cambridge: Cambridge University Press.

MacLaury, Robert E. 1997. Color and cognition in Mesoamerica: constructing categories as vantages. University of Texas-Pan American.

MacLaury, Robert E. 2001. "Color terms." In Language typology and language universals, edited by Haspelmath, Martin, König, Ekkehard, Oesterreicher, Wulf, Rabble, Wolfgang, 1227–1251. Berlin and New York: Walter de Gruyter.

MacLaury, Robert E., Galina V. Paramei, and Don Dedrick, eds. 2007. Anthropology of color: interdisciplinary multilevel modeling. Amsterdam ; Philadelphia: J. Benjamins Pub. Co.

Maclaury, Robert E., Judit Almási, and Zoltán Kövecses. 1997. "Hungarian piros and vörös: color from points of view." Semiotica 114 (1–2): 67–82. Available online at https://doi.org/10.1515/semi.1997.114.1-2.67, accessed on April 19, 2018.

Maffi, Luisa. 1990. "Somali color term evolution: grammatical and semantic evidence." Anthropological Linguistics 32: 316–334.

Majid, Asifa, Fiona Jordan, and Michael Dunn. 2013. "Evolution of semantic systems: a research initiative of the Max Planck Institute for Psycholinguistics." Max Planck Institute for Psycholinguistics. August 26, 2013. Available online at http://www.mpi.nl/departments/other-research/research-consortia/eoss, accessed on April 19, 2018.

Eessalu, Martin and Mari Uusküla. 2013. "The special case of beige: a cross-linguistic study." In Color and colorimetry multidisciplinary contributions, edited by Rossi, Maurizio, IX B:168–176. Rimini: Maggioli Editore.

Maxwell-Stuart, P. G. 1981. Studies in Greek colour terminology. Mnemosyne, Bibliotheca Classica Batava. Supplementum 65. Leiden: Brill.

Metsmägi, Iris, Meeli Sedrik, and Sven-Erik Soosaar, eds. 2012. Eesti etümoloogiasõnaraamat. Tallinn: Eesti Keele Sihtasutus. Available online at http://www.eki.ee/dict/ety/, accessed on April 19, 2018.

Miller, G. A., and P. Johnson-Laird. 1976. Language and perception. Cambridge, MA: Harvard University Press.

MITECS = Kay, Paul. Color Categorization. MITECS: The MIT Encyclopedia of the Cognitive Sciences. Available online at http://ai.ato.ms/MITECS/Entry/kay.html, accessed on April 19, 2018.

Morgan, Gerry, and Greville Corbett. 1989. "Russian colour term salience." Russian Linguistics 13 (2): 125–141. Available online at https://doi.org/10.1007/BF02551669, accessed on April 19, 2018.

Morgan, Gerry. 1993. "Basic colour terms: comparative results for French and Russian." Journal of French Language Studies 3 (01): 1–17. Available online at https://doi.org/10.1017/S0959269500000326, accessed on April 19, 2018.

Moss, A., I. Davies, G. Corbett, and G. Laws. 1990. "Mapping Russian basic colour terms using behavioural measures." Lingua 82 (4): 313–332. Available online at https://doi.org/10.1016/0024-3841(90)90068-V, accessed on April 19, 2018.

Must, Mari. 2000. Vene laensõnad eesti murretes. Tallinn: Eesti Keele Sihtasutus.

Mylonas, Dimitris, and Lindsay MacDonald. 2015. "Augmenting basic colour terms in English." Color Research & Application 41 (1): 32–42. Available online at https://doi.org/10.1002/col.21944, accessed on April 19, 2018.

Nişanyan, Sevan. 2007. Sözlerin soyağacı: çağdaş Türkçenin etomolojik sözlüğü. Third edition. İstanbul: Adam Yayınları. Available online at http://www.nisanyansozluk.com/, accessed on April 19, 2018.

Õim, Asta. 1983. Värvinimetuste moodustamisest eesti ja vene keeles. Keel ja Kirjandus 1. 26–32.

Oja, Vilja. 1998. "Database and computer analysis of the estonian colour adjectives." In Proceedings of the XVIth international congress of linguists : [CD-ROM], edited by Bernard Caron. Oxford: Pergamon-Elsevier Science Ltd.

Oja, Vilja. 2000. "Concept, word and place." In Koht ja paik =: Place and location, edited by Kaia Lehari and Virve Sarapik, 138–147. Eesti Kunstiakadeemia toimetised 8. Tallinn: Eesti Kunstiakadeemia. Available online at http://www.eki.ee/km/place/koht_1.htm, accessed on April 19, 2018.

Oja, Vilja. 2001. Linguistic studies of Estonian colour terminology. PhD thesis. Tartu: Tartu University Press.

Oja, Vilja. 2002. "Some Semantic problems in the translation of colour terms." In Proceedings of the tenth international symposium on lexicography, 253–260. Lexicographica. Series Maior 109. Tübingen: Niemeyer.

Oja, Vilja. 2007. "Color naming in Estonian and cognate languages." In Anthropology of color. interdistiplinary multilevel modeling, 189–209. Amsterdam, Philadelphia: John Benjamins B V Publ.

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): 919–956. Available online at https://doi.org/10.1515/ling.1998.36.5.919, accessed on April 19, 2018.

Özgen, Emre, and Ian R. L. Davies. 2002. "Acquisition of categorical color perception: a perceptual learning approach to the linguistic relativity hypothesis." Journal of Experimental Psychology: General 131 (4): 477–493. Available online at https://pdfs.semanticscholar.org/3683/92e28515db64bbd1371a5b1e047e2345e453.pdf, accessed on April 19, 2018.

Öztopcu, Kurtulus. 1999. Dictionary of the Turkic Languages: English, Azerbaijani, Kazakh, Kyrgyz, Tatar, Turkish, Turkmen, Uighur, Uzbek. Routledge.

Paggetti, Giulia, and Gloria Menegaz. 2012. “Is light blue (azzurro) color name universal in the Italian language?” In Trends and topics in computer vision, edited by Kiriakos N. Kutulakos, 90–103. Lecture Notes in Computer Science 6554. Springer Berlin Heidelberg. Available online at https://link.springer.com/chapter/10.1007/978-3-642-35740-4_8, accessed on April 19, 2018.

Paggetti, Giulia, Gloria Menegaz, and Galina V. Paramei. 2015. “Color naming in Italian language.” Color Research & Application. Available online at http://onlinelibrary.wiley.com/doi/10.1002/col.21953/full, accessed on April 19, 2018.

Paramei, G. V. 2005. “Singing the Russian blues: an argument for culturally basic color terms.” Cross-Cultural Research 39 (1): 10–38. Available online at https://doi.org/10.1177/1069397104267888, accessed on April 19, 2018.

Paramei, Galina V. 2007. Russian “blues”: Controversies of basicness. In Robert E. MacLaury, Galina V. Paramei & Don Dedrick (eds.), Anthropology of Color, 75–106. Amsterdam: John Benjamins Publishing Company. doi:10.1075/z.137.07par. https://benjamins.com/catalog/z.137.07par (24 October, 2018).

Paramei, Galina V., Mauro D’Orsi, and Gloria Menegaz. 2014. “‘Italian blues’: a challenge to the universal inventory of basic colour terms.” JAIC-Journal of the International Colour Association 13. Available online at http://jaic.jsitservices.co.uk/index.php/JAIC/article/view/155, accessed on April 19, 2018.

Pranaitytė, Simona. 2011. In Värvinimede raamat, edited by Mari Uusküla and Urmas Sutrop, 271–299. Töid antropoloogilise ja etnolingvistika vallast 5. Tallinn: Eesti Keele Sihtasutus.

R Core Team. 2012. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available online at http://www.R-project.org/, accessed on April 19, 2018.

Rakhilina, Ekaterina V., and Galina V. Paramei. 2011. “Colour terms: evolution via expansion of taxonomic constraints.” In New directions in colour studies, edited by Carole P. Biggam, Carole Hough, Christian Kay, and David R. Simmons, 121–132. Amsterdam: John Benjamins Publishing Company. Available online at https://benjamins.com/catalog/z.167.15rak, accessed on April 19, 2018.

Al-Rasheed, Abdulrahman S. 2014. “Further evidence for Arabic basic colour categories.” Psychology 05 (15): 1714–1729. Available online at https://doi.org/10.4236/psych.2014.515179, accessed on April 19, 2018.

Al-Rasheed, Abdulrahman S., Humood H. Al-Sharif, Mohammed J. Thabit, Norah S. Al-Mohimeed, and Ian R.L. Davies. 2011. “Basic colour terms of Arabic.” In New directions in colour studies, edited by Carole P. Biggam, Carole Hough, Christian Kay, and David R. Simmons, 53–58. Amsterdam: John Benjamins Publishing Company. Available online at https://benjamins.com/catalog/z.167.10alr, accessed on April 19, 2018.

Ratliff, Floyd. 1976. “On the psychophysiological bases of universal color terms.” Proceedings of the American philosophical society 120 (5): 311–330. Available online at http://www.jstor.org/stable/986265, accessed on April 19, 2018.

Rätsep, Kaidi. 2006. “A comparison of Estonian and Turkish basic color terms and the underlying mechanism for cultural divisions of color.” Bachelor thesis, Tallinn: Tallinn University Estonian Institute of Humanities.

Rätsep, Kaidi. 2010. “Turkish basic and non-basic colour terms.” Master thesis, Tallinn: Tallinn University Estonian Institute of Humanities.

Rätsep, Kaidi. 2011a. “Preliminary research on turkish basic colour terms with an emphasis on blue.” In New directions in colour studies, edited by Carole P. Biggam, Carole Hough, Christian Kay, and David R. Simmons, 133–146. Amsterdam: John Benjamins Publishing Company. Available online at https://benjamins.com/catalog/z.167.16rat, accessed on April 19, 2018.

Rätsep, Kaidi. 2011b. “Türgi põhivärvinimed rõhuasetusega sinisel.” In Värvinimede Raamat, edited by Urmas Sutrop and Mari Uusküla, 351–383. Töid antropoloogilise ja etnolingvistika vallast 5. Tallinn: Eesti Keele Sihtasutus.

Rätsep, Kaidi. 2012. “Colour term ‘black’in Estonian place names.” ESUKA [Eesti ja some-ugri keeleteaduse ajakiri] – JEFUL [Journal of Estonian and Finno-Ugric Linguistics] 3 (1): 129–154. Available online at http://core.ac.uk/download/pdf/25738505.pdf, accessed on April 19, 2018.

Rätsep, Kaidi. 2012. “Sorting and naming blue: an Estonian case study.” Presented at the Progress in Colour Studies 2012 (PICS12), Glasgow, July 13. Abstract available online at http://www.gla.ac.uk/media/media_236385_en.pdf, accessed on April 19, 2018.

Rätsep, Kaidi. 2013. “Some remarks on gender differences in turkish colour vocabulary”. In Proceedings of the second Central European conference in linguistics for postgraduate students (CECIL’S 2), 219–229. Available online at http://cecils.btk.ppke.hu/wp-content/uploads/2013/02/Ratsep_Some-remarks-on-gender_REVISED_v7_fin_corr1.pdf, accessed on April 19, 2018.

Rätsep, Kaidi. 2016. “Must ‘black’ and valge ‘white’ in Estonian place names: their naming motives in folk etymology.” In Color language and color categorization, edited by Mari Uusküla, Geda Paulsen, and Jonathan Brindle. Cambridge Scholars Publishing.

Regier, T., P. Kay, and N. Khetarpal. 2007. “Color naming reflects optimal partitions of color space.” Proceedings of the national academy of sciences 104 (4): 1436–1441. Available online at https://doi.org/10.1073/pnas.0610341104, accessed on April 19, 2018.

Regier, Terry, and Paul Kay. 2009. “Language, thought, and color: Whorf was half right.” Trends in cognitive sciences 13 (10): 439–446. Available online at https://doi.org/10.1016/j.tics.2009.07.001, accessed on April 19, 2018.

Regier, Terry, Paul Kay, Aubrey L. Gilbert, and Richard B. Ivry. 2010. “Language and thought: which side are you on, anyway.” Words and the mind: how words capture human experience, 165–182. Available online at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.159.3010&rep=rep1&type=pdf, accessed on April 19, 2018.

Rjabina, Elena. 2011. Основные цветообозначения в пермских языках. Dissertationes Philologiae Uralicae Universitatis Tartuensis 11. Tartu: Tartu Ülikooli Kirjastus. Available online at http://hdl.handle.net/10062/18808, accessed on April 19, 2018.

Roberson, Debi, Ian Davies & Jules Davidoff. 2000. Color categories are not universal: Replications and new evidence from a stone-age culture. Journal of Experimental Psychology: General 129(3). 369–398. doi:10.1037/0096-3445.129.3.369.

Roberson, Debi, Ian R. L. Davies, Greville G. Corbett, and Marieta Vandervyver. 2005. "Free-sorting of colors across cultures: are there universal grounds for grouping?" Journal of Cognition and Culture 5 (3): 349–386. Available online at http://epubs.surrey.ac.uk/1304/, accessed on April 19, 2018.

Romney, A. Kimball, Tom Smith, Howard E. Freeman, Jerome Kagan, and Robert E. Klein. 1979. "Concepts of success and failure." Social Science Research 8 (4): 302–326. Available online at https://doi.org/10.1016/S0049-089X(79)80002-2, accessed on April 19, 2018.

Rosch, Eleanor H. 1973. "Natural categories." Cognitive Psychology 4 (3): 328–350. Available online at https://doi.org/10.1016/0010-0285(73)90017-0, accessed on April 19, 2018.

Rosenberg, Seymour, and Moonja Park Kim. 1975. "The method of sorting as a data-gathering procedure in multivariate research." Multivariate Behavioral Research 10 (4): 489–502. Available online at https://doi.org/10.1207/s15327906mbr1004_7, accessed on April 19, 2018.

Ross, Ihaka, Paul Murrell, Kurt Hornik, Jason C. Fisher, and Achim Zeileis. 2013. Colorspace: color space manipulation (version 1.2-4.). R Package. Available online at http://CRAN.R-project.org/package=colorspace, accessed on April 19, 2018.

RStudio Team. 2016. RStudio: integrated development environment for R. Boston, MA: RStudio, Inc. Available online at http://www.rstudio.com/, accessed on April 19, 2018.

Ryabina, Elena. 2011a. "Põhivärvinimed udmurdi keeles." In Värvinimede raamat, 227–268. Töid antropoloogilise ja etnolingvistika vallast 5. Tallinn: Eesti Keele Sihtasutus.

Ryabina, Elena. 2011b. "Differences in the distribution of colour terms in colour space in the Russian, Udmurt and Komi languages." ." ESUKA [Eesti ja some-ugri keeleteaduse ajakiri] – JEFUL [Journal of Estonian and Finno-Ugric Linguistics] 2: 191–213.

Saagpakk, Paul (ed.). 1992. Eesti-inglise sõnaraamat = Estonian-English dictionary. 2. tr. Tallinn: Koolibri.

Şahin Ekici, Ebru, Cengiz Yener, and Nilgün Camgöz. 2006. "Colour naming." Optics & Laser Technology 38 (4–6): 466–485. Available online at https://doi.org/10.1016/j.optlastec.2005.06.007, accessed on April 19, 2018.

Şahin, Ebru. 1998a. Color naming. Dissertation. Ankara: Bilkent University, Department of Interior Architecture and Environmental Design and Institute of Fine Arts.

Sandford, Jodi L. 2011. "Blu, azzurro, celeste – what color is blue for Italian speakers compared to english speakers?" In Proceedings of the eighth national color conference, VIII B:281–288. Colour and colorimetry: multidisciplinary contributions. Milano: Maggioli Editore.

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

Slowikowski, Kamil. 2017. Ggrepel: repulsive text and label geoms for "Ggplot2." Available online at https://CRAN.R-project.org/package=ggrepel, accessed on April 19, 2018.

Stanulewicz, Danuta, Ewa Komorowska, and Adam Pawłowski. 2014. "Axiological aspects of Polish colour vocabulary: a study of associations." In Colour studies, edited by Wendy Anderson, Carole P. Biggam, Carole Hough, and Christian Kay, 258–272. Amsterdam: John Benjamins Publishing Company. Available online at https://doi.org/10.1075/z.191.17sta, accessed on April 19, 2018.

Stanulewicz, Danuta. 2010. "Polish terms for 'blue' in the perspective of vantage theory." Language Sciences 32 (2): 184–195. Available online at https://doi.org/10.1016/j.langsci.2009.10.004, accessed on April 19, 2018.

Starko, Vasyl'. 2013. "Ukrainian colour concepts for blue." Slovo. Journal of Slavic Languages and Literatures 54: 150–163. Available online at http://www.moderna.uu.se/digitalAssets/591/c_591534-l_1-k_14_starko.pdf, accessed on April 19, 2018.

Steingass, Francis Joseph. 1892. A Comprehensive Persian-English dictionary, including the Arabic words and phrases to be met with in Persian literature. London: Routledge & K. Paul. Available online at http://dsal.uchicago.edu/dictionaries/steingass/, accessed on April 19, 2018.

Steinvall, Anders. 2002. English colour terms in context. (Skrifter Från Moderna Språk 3). Umeå: Inst. för moderna språk, Umeå univ. Available online at https://www.diva-portal.org/smash/get/diva2:144764/FULLTEXT01.pdf, accessed on May 26, 2018.

Sutrop, Urmas. 1995. "Eesti keele põhivärvinimed." Keel ja kirjandus 12: 797–808.

Sutrop, Urmas. 2000a. "The basic colour terms of Estonian." Trames : journal of the humanities and social sciences 4 (1): 143–168.

Sutrop, Urmas. 2001. "List task and a cognitive salience index." Field Methods 13 (3): 263–276. Available online at https://doi.org/10.1177/1525822X0101300303, accessed on April 19, 2018.

Sutrop, Urmas. 2002. The vocabulary of sense perception in Estonian: structure and history. Doctoral thesis. Opuscula Fenno-Ugrica Gottingensia, Bd. 8. Frankfurt am Main: Lang.

Sutrop, Urmas. 2011a. "Towards a semiotic theory of basic colour terms and the semiotics of Juri Lotman." In New directions in colour studies, edited by Carole P. Biggam, Carole Hough, Christian Kay, and David R. Simmons, 39–48. Amsterdam: John Benjamins Publishing Company. Available online at https://doi.org/10.1075/z.167.07sut, accessed on April 19, 2018.

Sutrop, Urmas. 2011b. "Eesti keele põhivärvinimed." In Värvinimede raamat, edited by Mari Uusküla and Urmas Sutrop, 65–82. Töid antropoloogilise ja etnolingvistika vallast 5. Tallinn: Eesti Keele Instituut.

Swadesh, Morris. 1972. The origin and diversification of language. London: Routledge and K. Paul.

Thompson, E. C., and Z. Juan. 2006. "Comparative cultural salience: measures using free-list data." Field Methods 18 (4): 398–412. Available online at https://doi.org/10.1177/1525822X06293128, accessed on April 19, 2018

Truex, G. F. 1977. "Measurement of intersubject variations in categorizations." Journal of cross-cultural psychology 8 (1): 71–82. Available online at https://doi.org/10.1177/002202217781006, accessed on April 19, 2018

Tufte, Edward R. 2001. The visual display of quantitative information. 2nd edition. Cheshire, Conn: Graphics Press.

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

Uusküla, Mari, and Urmas Sutrop. 2007. "Preliminary study of basic colour terms in Modern Hungarian." Linguistica Uralica 43 (2): 102–123.

Uusküla, Mari, and Urmas Sutrop. 2010. "The puzzle of two terms for red in Hungarian." In Rara & Rarissima: documenting the fringes of linguistic diversity, edited by Jan Wohlgemuth and Michael Cysouw, 359–376. Empirical Approaches to Language Typology [EALT] 46. Berlin, New York: De Gruyter.

Uusküla, Mari. 2007. "The basic colour terms of Finnish." SKY Journal of Linguistics 20, 367–397. Available online at http://www.linguistics.fi/julkaisut/SKY2007/UUSKULAo.pdf, accessed on April 19, 2018.

Uusküla, Mari. 2008. Basic colour terms in finno-ugric and slavonic languages: myths and facts. Doctoral thesis. Tartu: Tartu University Press. Available online at http://dspace.utlib.ee/dspace/handle/10062/6401, accessed on April 19, 2018.

Uusküla, Mari. 2011. "Terms for red in Central Europe: an areal phenomenon in Hungarian and Czech." In New directions in colour studies, edited by Carole P. Biggam, Carole Hough, Christian Kay, and David R. Simmons, 147–156. Amsterdam: John Benjamins Publishing Company. Available online at https://benjamins.com/catalog/z.167.17uus, accessed on April 19, 2018.

Uusküla, Mari. 2014. "Linguistic categorization of BLUE in Standard Italian." In Colour studies, edited by Wendy Anderson, Carole P. Biggam, Carole Hough, and Christian Kay, 67–78. Amsterdam: John Benjamins Publishing Company. Available online at https://benjamins.com/catalog/z.191.04uus, accessed on April 19, 2018.

Vääri, Eduard, Richard Kleis, and Johannes Silvet. 2012. Võõrsõnade leksikon. Tallinn: Valgus.

Vejdemo, Susanne, Carsten Levisen, Cornelia van Scherpenberg, þórhalla Guðmundsdóttir Beck, Åshild Næss, Martina Zimmermann, Linnaea Stockall, and Matthew Whelpton. 2015. "Two kinds of pink: development and difference in Germanic colour semantics." Language Sciences 49 (May): 19–34. Available online at https://doi.org/10.1016/j.langsci.2014.07.007, accessed on April 19, 2018.

Wehr, Hans. 1979. A dictionary of modern written Arabic (Arabic-English). (Ed.) Milton J. Cowan. Fourth Edi. Wiesbaden: Otto Harrassowitz.

Weller, Susan C., and A. Kimball Romney. 1988. Systematic data collection. Qualitative Research Methods, v. 10. Newbury Park, Calif: Sage Publications.

Wickham, Hadley. 2007. "Reshaping data with the reshape package." Journal of Statistical Software 21 (12): 1–20. Available online at http://www.jstatsoft.org/v21/i12/, accessed on April 19, 2018.

Wickham, Hadley. 2009. Ggplot2: Elegant Graphics for Data Analysis. Available online at http://had.co.nz/ggplot2/, accessed on April 19, 2018.

Wickham, Hadley. 2011. "The split-apply-combine strategy for data analysis." Journal of Statistical Software 40 (1): 1–29. Available online at http://www.jstatsoft.org/v40/i01/, accessed on April 19, 2018.

Wickham, Hadley. 2012. Scales: scale functions for graphics. Available online at http://CRAN.R-project.org/package=scales, accessed on April 19, 2018.

Wiedemann, Ferdinand Johann. 1973. Eesti-saksa sõnaraamat = Estnisch-Deutsches Wörterbuch. 4. muutmata tr. teisest Jakob Hurda redigeeritud väljaandest. Tallinn: Valgus.

Wierzbicka, Anna. 1990. The meaning of color terms: semantics, culture, and cognition. Cognitive Linguistics 1 (1): 99–150. Available online at http://www.degruyter.com/view/j/cogl.1990.1.issue-1/cogl.1990.1.1.99/cogl.1990.1.1.99.xml?format=INT, accessed on April 8, 2018.

Wierzbicka, Anna. 1996. Semantics: primes and universals. Oxford [England]; New York: Oxford University Press.
Winawer, Jonathan, Nathan Witthoft, Michael C. Frank, Lisa Wu, Alex R. Wade, and Lera Boroditsky. 2007. "Russian blues reveal effects of language on color discrimination." Proceedings of the National Academy of Sciences 104 (19): 7780–7785. Available online at https://doi.org/10.1073/pnas.0701644104, accessed on April 19, 2018.
Woods, Anthony, Paul Fletcher, and Arthur Hughes. 2003. Statistics in language studies. Cambridge Textbooks in Linguistics. Cambridge [Cambridgeshire]; New York: Cambridge University Press.
Zabrodskaja, Anastassia. 2009. Russian-Estonian language contacts: grammatical aspects of language use and change. Tallinn University Dissertations on Humanitie. Tallinn: Tallinn University. Available online at http://www.tlulib.ee/files/arts/93/Zabro6d6de71d850daacb358353791fb9bc95.pdf, accessed on April 19, 2018.
Zimmer, Alf C. 1982. "What really is turquoise? A note on the evolution of color terms". Psychological Research 44(3), 213–230. Available online at http://link.springer.com/10.1007/BF00308421, accessed on June 1, 2018.
"City University dynamic colour vision test." 2015. City University London. Available online at https://www.city.ac.uk/health/research/centre-for-applied-vision-research/a-new-web-based-colour-vision-test, accessed on April 19, 2018.
"Metric and Nonmetric MDS." 2005. In Modern multidimensional scaling, 199–225. New York, NY: Springer New York. Available online at https://doi.org/10.1007/0-387-28981-X_9, accessed on April 19, 2018.
The Corpus of Estonian Literary Languages. 2014. Available online at http://www.cl.ut.ee/korpused/baaskorpus/, accessed on April 19, 2018
"The New Color-Aid Booklet." 2006. Color-Aid Corporation. Available online at www.coloraid.com, accessed on April 19, 2018
"Translit RU/EN – Russian transliteration and spell checker". 2007–2008. Available online at http://translit.cc/, accessed on April 19, 2018

CURRICULUM VITAE

Name: Kaidi Rätsep
Date and place of birth: 27.04.1982, Tallinn
Citizenship: Estonian
E-mail: kaidi.ratsep@gmail.com

Education

2001–2006 Tallinn University Estonian Institute of Humanities, 2006 BA Middle-Eastern and Asian Culture (Turcology)
2006–2010 Tallinn University Estonian Institute of Humanities, 2010 MA Middle-Eastern and Asian Culture (Turcology)
2010–2018 University of Tartu, PhD student (general linguistics)

Professional experience

2007–2008 Institute of the Estonian Language; Assistant (0.25)
2008–2015 Institute of the Estonian Language; Assistant (0.50)

Projects

01.01.2011–31.12.2014
SF0050037s10 "Word and concept in the development of Estonian vocabulary and place names (1.01.2010–31.12.2014)", research staff, principal investigator Urmas Sutrop, Institute of the Estonian Language.

1.01.2010–31.12.2013
ETF8168 "Areal or universal: Basic colour terms in the Baltic Sea, Central-European and Mediterranean areas (1.01.2010–31.12.2013)", senior research staff, principal investigator Mari Uusküla, Institute of the Estonian Language.

01.01.2009–31.12.2009
ETF6744 "Language family tree theory and some lexico-semantic fields (colours, the human being). (1.01.2006–31.12.2009)", senior research staff, principal investigator Urmas Sutrop, Institute of the Estonian Language.

Publications

Rätsep, Kaidi (2016). *Must* 'Black' and *valge* 'White' in Estonian Place Names: Their Naming Motives in Popular Etymology. In: Paulsen. Geda; Uusküla. Mari; Brindle. Jonathan (Ed.). Color Language and Color Categorization (333–355), Cambridge Scholars Publishing.

Rätsep, Kaidi (2012). Colour term ‘black’ in Estonian place names. Eesti and soome-ugri keeleteaduse ajakiri / Journal of Estonian and Finno-Ugric Linguistics, 3, 129–154.
Rätsep, Kaidi (2011). Preliminary research on Turkish basic colour terms with an emphasis on blue. Carole P. Biggam. Carole A. Hough. Christian J. Kay and David R. Simmons (Toim.). New Directions in Colour Studies (133–146), Amsterdam: John Benjamins.
Rätsep, Kaidi (2011). Türgi põhivärvinimed rõhuasetusega sinisel. Uusküla. Mari and Sutrop. Urmas (Toim.). Värvinimede raamat (351–383), Tallinn: Eesti Keele Sihtasutus.

ELULOOKIRJELDUS

Nimi:	Kaidi Rätsep
Sünniaeg ja-koht:	27.04.1982, Tallinn
Kodakondsus:	Eesti
E-post:	kaidi.ratsep@gmail.com

Haridus

2001–2006	Tallinna Ülikooli Eesti Humanitaarinstituut, 2006 BA Lähis-Ida and Aasia kultuurilugu (turkoloogia)
2006–2010	Tallinna Ülikooli Eesti Humanitaarinstituut, 2010 MA Lähis-Ida and Aasia kultuurilugu (turkoloogia)
2010–2018	Tartu Ülikool, doktoriõpe (üldkeeleteadus)

Töö

2007–2008	Eesti Keele Instituut; Assistent (0,25)
2008–2015	Eesti Keele Instituut; Assistent (0,50)

Projektid

01.01.2011–31.12.2014
SF0050037s10 „Sõna ja mõiste eesti sõnavara ja kohanimede kujunemisel (1.01.2010–31.12.2014)", täitja, vastutav täitja Urmas Sutrop, Eesti Keele Instituut.

1.01.2010–31.12.2013
ETF8168 „Areaalne või universaalne: põhivärvinimed Läänemere, Kesk-Euroopa ja Vahemere areaalis (1.01.2010–31.12.2013)", põhitäitja, vastutav täitja Mari Uusküla, Eesti Keele Instituut.

01.01.2009–31.12.2009
ETF6744 „Keelepuuteooria ja mõned leksikaal-semantilised väljad (värvid, inimene). (1.01.2006–31.12.2009)", põhitäitja, vastutav täitja Urmas Sutrop, Eesti Keele Instituut.

Väitekirjaga seotud publikatsioonid

Rätsep, Kaidi (2016). *Must* 'black' and *valge* 'white' in Estonian place names: Their naming motives in folk etymology. Mari Uusküla, Geda Paulsen and Jonathan Brindle (Eds.). Color language and color categorization, Cambridge Scholars Publishing.

Rätsep, Kaidi (2012). Colour term 'black' in Estonian place names. Eesti and soome-ugri keeleteaduse ajakiri / Journal of Estonian and Finno-Ugric Linguistics, 3, 129–154, Rätsep, Kaidi (2011). Türgi põhivärvinimed rõhuasetusega sinisel. Uusküla, Mari and Sutrop, Urmas (Toim.). Värvinimede raamat (351–383), Tallinn: Eesti Keele Sihtasutus.

Rätsep, Kaidi (2011). Preliminary research on Turkish basic colour terms with an emphasis on blue. Carole P, Biggam, Carole A, Hough, Christian J, Kay and David R, Simmons (Eds.). New Directions in Colour Studies (133–146), Amsterdam: John Benjamins.

DISSERTATIONES LINGUISTICAE UNIVERSITATIS TARTUENSIS

1. **Anna Verschik.** Estonian yiddish and its contacts with coterritorial languages. Tartu, 2000, 196 p.
2. **Silvi Tenjes.** Nonverbal means as regulators in communication: socio-cultural perspectives. Tartu, 2001, 214 p.
3. **Ilona Tragel.** Eesti keele tuumverbid. Tartu, 2003, 196 lk.
4. **Einar Meister.** Promoting Estonian speech technology: from resources to prototypes. Tartu, 2003, 217 p.
5. **Ene Vainik.** Lexical knowledge of emotions: the structure, variability and semantics of the Estonian emotion vocabulary. Tartu, 2004, 166 p.
6. **Heili Orav.** Isiksuseomaduste sõnavara semantika eesti keeles. Tartu, 2006, 175 lk.
7. **Larissa Degel.** Intellektuaalsfäär intellektuaalseid võimeid tähistavate sõnade kasutuse põhjal eesti ja vene keeles. Tartu, 2007, 225 lk.
8. **Meelis Mihkla.** Kõne ajalise struktuuri modelleerimine eestikeelsele tekstkõne sünteesile. Modelling the temporal stucture of speech for the Estonian text-to-speech synthesis. Tartu, 2007, 176 lk.
9. **Mari Uusküla.** Basic colour terms in Finno-Ugric and Slavonic languages: myths and facts. Tartu, 2008, 207 p.
10. **Petar Kehayov.** An Areal-Typological Perspective to Evidentiality: the Cases of the Balkan and Baltic Linguistic Areas. Tartu, 2008, 201 p.
11. **Ann Veismann.** Eesti keele kaas- ja määrsõnade semantika võimalusi. Tartu, 2009, 145 lk.
12. **Erki Luuk.** The noun/verb and predicate/argument structures. Tartu, 2009, 99 p.
13. **Andriela Rääbis.** Eesti telefonivestluste sissejuhatus: struktuur ja suhtlusfunktsioonid. Tartu, 2009, 196 lk.
14. **Liivi Hollman.** Basic color terms in Estonian Sign Language. Tartu, 2010, 144 p.
15. **Jane Klavan.** Evidence in linguistics: corpus-linguistic and experimental methods for studying grammatical synonymy. Tartu, 2012, 285 p.
16. **Krista Mihkels.** Keel, keha ja kaardikepp: õpetaja algatatud parandussekventside multimodaalne analüüs. Tartu, 2013, 242 lk.
17. **Sirli Parm.** Eesti keele ajasõnade omandamine. Tartu, 2013, 190 lk.
18. **Rene Altrov.** The Creation of the Estonian Emotional Speech Corpus and the Perception of Emotions. Tartu, 2014, 145 p.
19. **Jingyi Gao.** Basic Color Terms in Chinese: Studies after the Evolutionary Theory of Basic Color Terms. Tartu, 2014, 248 p.
20. **Diana Maisla.** Eesti keele mineviku ajavormid vene emakeelega üliõpilaste kasutuses. Tartu, 2014, 149 lk.