LUNDS UNIVERSITET

Colexification and semantic change in colour terms in Sino-Tibetan and Indo-European languages

Kajsa Söderqvist

Bachelor Thesis in General Linguistics
Supervisor: Gerd Carling
Autumn semester 2016/2017
University of Lund, Centre for Languages and Literature

Abstract

Colour terms is a highly interesting field when investigating linguistic universals and how language vary cross-linguistically. Colour semantics, the investigation of the meaning of colour, consists in largely of two opposing sides: the universalists, proposing that colour terms are universal (Berlin & Kay 1969) and the relativists claiming a variation in meaning cross-linguistically (Wierzbicka 2008).

The highly changeable field lexical semantic change is defined as the change in meaning in concepts connected to a lexical item and a typical pattern of change is words becoming polysemous (Durkin 2009). To gain an expanded picture and understanding of a term, a historical investigation and etymological research of its derived concepts is a useful resource. Biggam (2012) points out that specifically colour terms are less stable and that historical colour terms tend to have broader coverage than the modern terms, which makes them an interesting object of investigation.

The focus of this thesis is consequently to investigate and contrast the synchronic colexifications and diachronic derivations of ten colour terms in ten Sino-Tibetan and ten Indo-European languages. A dataset in DiACL (Carling 2017) has been constructed to gather the collected lexemes, followed by a manual extraction to semantic networks for a visual representation (Felbaum 2012). The lexical meanings have then been grouped into semantic classifications (Haspelmath & Tadmor 2009) for further analyze.

The results showed very small overlap of colexified lexical meanings for each colour term in the diachronic perspective, but showed a conformity of semantic categories between the families. The type of change that occurred most frequently was narrowing and the direction of the semantic change went most frequently from more abstract to more concrete. When changes in the opposite direction occurred, it was almost exclusively in the Indo-European languages, not consistent with previous studies (Campbell 2004, Warth-Szczyglowska 2014).

Preface and acknowledgements

A big thank you to Gerd Carling for the guidance and support this winter (and spring), for providing me with the DiACL database and steering me in the right direction when I was out and about. Thank you Chundra Cathcart for giving me the great advice to use the STEDT database and giving me instructions and tips on how to use it. And last, but not least, a big thank you to my rock, my everything Damian Luna for being there night and day, supporting me through good and bad. Couldn't have done it without you.

Table of contents

1. Introduction

Colours are a curious thing, on the one hand we are all biologically constructed in the same way with the same biological features (except those with some sort of vision deficiency), hence we should all be able to see the same thing. On the other hand, we are all a world of our own, with our own specific experiences and personalities, raised in different cultures with its specific colour spectra and with our own specific mindset that could most certainly affect the way we see and label the world. Colours are and have always been a natural part of the world we live in, as a mean of successfully communicating the dazzling sunset, the ripe juicy apple or the deepest of ocean.

In many contemporary cultures, colours convey not only a descriptive function but also important vital messages in everything from warning signs and traffic lights to national boundaries and political ideologies. These concepts and the way in which we use them are today taken for granted and something that everyone is expected to master in order to successfully convey a picture or an idea to members of the same speech community. Not only are we to correctly identify and label all colours around us but we are also supposed to grasp associations and the meanings of each colour term that we are being exposed to daily. A green car really has nothing to do with the actual colour of the vehicle, a blue person doesn't likely refer to a smurf and a black soul hasn't got anything to do with the visual colour of it.

Colours are also an interesting phenomenon since they rarely are being experienced in isolation, detached from something else. They are naturally almost always an attribute or a part of another object that we, in the Western world, have differentiated and thus created the concept of colour, but that doesn't mean that one can assume that all societies around the world has done the same. Societies develop, language changes and so does words and colour terms. When looking at the history of a linguistic element, it is obvious that more and less drastic changes occur and has occurred and those can be explained by lexical semantic change. Biggam (2012) points out that colour terms are *less* stable which makes them an interesting object of investigation. A stage of polysemy, where a word takes on an additional meaning, is a common indicator of a lexical semantic change in the making (Durkin 2009). With colour being such a special field in linguistics, do they follow the typical patterns of semantic change?

Since colour is a field that contains both biological and linguistic features, it is a well addressed and studied field when discussing the link between thought and language. Colour semantics is thus a field that gives an opportunity to further investigate the constant question of whether or not our language shapes or constrains our thought or if they are shaped by universals of human cognition (Berlin & Kay 1969).

Etymology is a field that deals with the history of a linguistic element, one could say that etymology provides a more or less detailed picture of the history of a word. In order to get a deeper understanding of each colour term, it is hence crucial to look at its history. There are plenty of etymological data available concerning colour terms and my idea was to gather them for the sake of investigating the differences, connections and changes. I will investigate ten colour terms since it is troublesome to consider a word's semantic development in isolation from other words in the same field, and then I aim to contrast two language families against each other (Durkin 2009).

2. Theoretical background

2.1 Colour semantics

Colour semantics concerns the part of semantics (the study of meaning) that deals with investigation of the meaning of colour (Biggam 2012). The list of research conducted in the field is long and one of the central, still unanswered, questions is whether we humans, independently of the language that we speak, categorize colours in the same way. If our concept of colour is shaped by universals of human cognition, or whether our language shapes and determines the way we express it.

Several studies has been made on this topic, some pointing towards a universalist view as for example Berlin & Kay and their famous book from 1969: *Basic colour terms: their universality and evolution.* There, Berlin & Kay present a theory based on basic colour terms (BCT) and their existence or non-existence in a language being directly linked to the development of its culture. They claim that there are universal basic colour terms,

more precisely eleven of them and that they are not randomly chosen. Although languages differ in the number of BCT's they lexicalize, the theory suggests that the order in which they do so, is *fixed* and *universal*. So, even though different languages might differ in the form of a colour term: *green* (English), *grön* (Swedish), *verde* (Spanish), they still denote the same category GREEN[1].

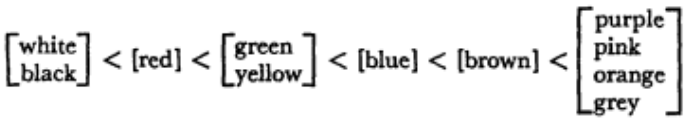

Figure 1: Proposed universal evolutionary sequence by Berlin & Kay (1969).

Although Berlin & Kay's theory very much formed the landscape of colour semantics and it indeed is an appealing theory in its neatness and simplicity, it has been criticized and challenged (Wierzbicka 2008). The linguistic relativity theory on the other side, has had a huge influence on the colour debate during the whole second half of the century and even more after the 1990s, claiming in opposition to the universalists that each language encodes their specific parts of the colour spectra in a completely arbitrary manner (Regier et al. 2010). If this turn out to be true, colour terms would instead be influenced or formed by the language and the way we speak about them. Studies that conclude that the categorization of colour terms might not be as simple as Berlin & Kay suggests, has been made by amongst other Vejdemo et al. (2014) discovering two colour categories of *pink* in a few of the Germanic languages in the study.

One could interpret the fact that not all languages has a word for *colour*, as an evidence for the phenomena being a culture-specific artefact, hence being of a complicated internal structure (Wierzbicka 2008). One suggestion would be to apply a wider perspective when investigating this matter and hence focusing on the universal concept of SEEING instead of the questionable non-universal concept of colour. Biggam (2012) suggests that there exist prehistoric colour category prototypes deriving from our need of conceptualize opposing pairs. That the earliest human beings had concepts for SEE

[1] I will from here on use SMALL CAPS when referring to colour categories and concepts and *italics* for lexical meanings.

and NOT SEE is very probable, which could be translated into LIGHT and DARK, colour category prototypes for WHITE and BLACK. Further, Wierzbicka (2008) suggest that FIRE is the colour category prototype for RED since that was a crucial part of human survival after the obvious contrast of day and night. Accordingly, only these two (BLACK and WHITE), or possibly three (RED), categories can be considered universal according to Biggam (2012).

2.2 Lexical semantic change

Lexical semantic change concerns the change of meaning in concepts associated with a lexical item in a language through time (Urban 2015). Lexical semantics is defined as "the study of meaning as conveyed by words and phrases" by Biggam (2012 p. 9). In opposition to other fields in linguistics as for example grammar, semantic change is a field that shows less resistance to change which makes it an especially interesting field (Biggam 2012).

The theory of semantic change goes back to Aristotle's analyses of metaphors where he explains how one can displace or relocate the meaning of a word. The field was then more seen as an art form and not until the mid-eighteenth/early nineteenth century did linguistic meaning become a field of study by researchers such as Bréal and Reisig (Sjöström 2001) and from there on, the interest in semantic change became a widespread field of research. It is a field that developed as a fascination for the correct description of the historical development of meanings and words. Diachronic semantics is concerned with the classification of mechanisms of semantic change, an activity that links lexicography with historical linguistics (Geeraerts 2010). When a concept associated with a word or a meaning of an individual word changes, what we have is then a lexical semantic change (Campbell 2004).

Lexical semantics investigates the hidden mechanism and flexibility of polysemy. When looking at diachronic development of a word it is likely that one will observe that the peripheral meaning of the word may develop to become the prototypical meaning (Durkin 2009). There is a great unpredictability in semantic change which can result in much greater challenges for etymological research (Durkin 2009). Sociocultural historical facts are often relevant - therefore some say it is useless to seek

generalizations to explain semantic change, but most admit that there are some general statements about how and why meanings change, even though they aren't regular nor predictable (Campbell 2004).

Although some claim that since we can't exactly explain semantic change, we can't either predict it. Others claim that we, one day, will be able to understand how the many complicated factors interact with each other. However, many seems to agree upon the explanation that language generally, is going towards the direction of protecting its functional needs (Campbell 2004). Another typical trend in semantic change is that the direction is going from a more concrete to a more abstract sense (Campbell 2004).
The four most common changes when it comes to semantic change are the following:

- Broadening, the restrictions associated with the word are lost, the meaning hence goes from a more concrete to a more abstract sense and the meaning of the word increases, becoming less specific (Campbell 2004).
- Narrowing, the meaning goes from a more abstract to a more concrete sense hence becoming more restricted and specific in its use (Campbell 2004).
- Metaphor, when a word extends its meaning and keep a semantic similarity or connection with both the original sense and the new sense. The new sense of the word gets put in a new sphere but there is still a connection to the original sense of the word (Campbell 2004).
- Metonym, the meaning increases its senses by including closely associated senses very near to its original meaning (Campbell 2004).

Semantic loans are another type of semantic change and refers to when a word broadens its meanings as a consequence of association with a meaning of a similar word in a different language. The two terms might be historically related or similar to each other, it is therefore difficult to establish whether it is a coincidental semantic development or a semantic borrowing.

Extensive studies investigating the semantic change, synchronically and diachronically, concerning *one* colour term in two languages has been done concluding that the most usual semantic shift is from more concrete to more abstract (Warth-Szczyglowska

2014) and metaphor and metonymy to be important in semantic change concerning colour terms (Xing 2008, Gieron-Czepczor 2013, Hamilton 2016).

Biggam (2012) claims that colour terms are less stable, which makes an investigation of the semantic change particularly interesting. A semantic shift can occur when a meaning changes from one cognitive category to another, it can cumulate to more drastic shifts and take place over a few generations or even only happen in the languages of a few native speakers (Biggam 2012).

2.3 Polysemy and colexification

When researching semantic change and etymology, another important and crucial factor to mention and explain is the field of polysemy. Polysemy is defined as one single word form associated with two or more related senses (Urban 2015). A clear example of that is the highly polysemous English word *line* that means both to read a *line*, to wait in *line* and to draw a *line* (Falkum, Vicente 2015). A typical pattern for semantic change hence, is that words become polysemous (Durkin 2009).

We can trace back the beginning of polysemy and that the French semantician Michel Bréal coined the term as far back as 1887 (Campbell 2004). The theory proposes that a word doesn't suddenly change meaning completely overnight, but change gradually through polysemy (Urban 2015). A word thus, starts out with one meaning, then acquires additional, multiple meanings and over time the original meaning is lost (Campbell 2004). Another way of defining polysemy is the thought that a word has a core meaning and possibly various peripheral senses and that one of those peripheral meanings becomes more central. The core then moves away from the central sense or altogether disappears (Campbell 2004).

From a synchronic point of view, polysemy is thus a rather difficult concept: very close meanings may simply show different conventional contextual uses of a single core meaning, while it is difficult to be sure that distant meanings are perceived by speakers as having more in common than the meanings of unrelated homonyms (Durkin 2009). Since the recognition of polysemy in natural languages is very constant it suggests that hearers and speakers might find it easier to extend already existing words to new

functions rather than to invent new words for each sense (Falkum & Vicente 2015). The interaction between the senses of a word demands the same model for variation in linguistic change that we encounter at various points in our examination of change in word form: A> A~B >B. The word first has one meaning, then it goes through an intermediate period in which it has both or more meanings 'A' and 'B', to later lose the original meaning and completely adapt the new meaning (Durkin 2009). One could then look upon polysemy as the synchronic side of lexical semantic change. As a still photograph taken in the exact moment when the first runner hands over the baton to the following runner in a relay race.

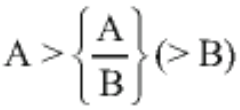

Figure 2: Linguistic change going through a stage of polysemy (Urban 2015).

An excellent example of that is the Spanish word *alcalde* which originally was borrowed from Arabic *qāḍī*, meaning 'judge (in Islamic law)' (stage A), a word that later was broadened to mean 'an official who is magistrate and mayor' *(*stage A~B*)* to eventually lose its original meaning and today only meaning 'mayor' *(*stage B*)*. It is important to understand that even though our etymological record only gives us information about a phase A and a phase B of a word, it is still highly likely to have been a phase A~B in between (Durkin 2009).

Colexification can be described as "the capacity, for two senses, to be lexified by the same lexeme in synchrony" and "a given language is said to colexify two functionally distinct senses if, and only if, it can associate them with the same lexical form" (François 2008 p. 170-171). I chose to use the term colexification in this thesis, since it corresponds to both polysemy (form is associated with two or more related meanings), homonymy (form is associated with two or several unrelated meanings) and semantic vagueness (List, Mayer, Terhalle & Urban 2014). As François further mentions: while focusing on the most exotic exceptions when looking at semantic universals, one tend to

disregard the many similarities of lexical polysemy's that can be found worldwide (François 2008).

2.4 Etymology

We can roughly say that a word's etymology is a report of its detailed story (Durkin 2009). Typically, historical investigation starts with an analysis of a single lexeme, followed by a reconstruction of a proto-form, to construct an etymology. Prehistoric unattested languages that has been reconstructed with methods of principles established through analysis and comparison of languages are in this thesis, as commonly, marked with an asterix (*) (Biggam 2012). It is hence a field that connects one chronological stage to a later one, and deals with the origin and development of a linguistic elements (Mailhammer 2015). Etymology is a field of research conducted in order to get a clear understanding and a coherent picture of the history of a single individual word. That is done through application of different methods and insights from various fields of historical linguistics (Durkin 2009). Historical colour terms are more likely to have broader coverage in comparison to modern colour terms according to Biggam (2012). A possible explanation to that could be the 'taxonomic -abstracting type' of reconstructing semantics, a method that reconstructs the semantics of a word by assigning the meanings from its cognates, creating a (in many cases) broad proto meaning. Semantic reconstructions are thus problematic and the possibility that a reconstructed word actually meant the reconstructed proto meaning is hard to tell (Urban 2015). The etymology of a specific word can hence provide valuable and useful information and a broader understanding of the history of a language. By tracing the history of a word's form and sound and by looking for regularities, the meaning of the word broadens (Durkin 2009).

3. Aim & Research questions

In the light of the previous chapter, the aim of this thesis is to combine these fields and look closer at the colexification and semantic change of ten colour terms in the Sino-Tibetan and Indo-European languages. My aim is to through colexification and

diachronic change, map each colour term's semantic derivations and more specifically answer the following questions:

- Which meanings colexify with colour terms?
- Which classifications of lexical meanings colexify with colour terms?
- What type of semantic change occurs?
- In what respect do the lexical meanings and classifications differ when looking at the colexification and semantic change?
- In which respect do the results differ between the families?

4. Method & Data

4.1 Colour terms and languages

The colour terms that I investigated are the following: *red, green, blue, yellow, grey, brown, purple, orange, black* and *white.* I chose not to include the colour term *pink* since it is a term that, at least in Europe, didn't exist until the 17th or 18th century, and is not relevant since I am interested in doing a historical, etymological study (Vejdemo et al. 2015). Thus, the colour terms that I have chosen correspond to the eleven BCT's (Berlin & Kay 1969), minus *pink*. As Durkin (2009) mentions, it is dangerous to try to consider a word's semantic development in isolation from other words in the same field, which is why I investigated as many colour terms as ten in many various languages and families.

The colour naming of western societies tends to be based upon hues such as *red, blue* and yellow. I therefore chose to exclude the accompanying words that refers to saturation, tone or brightness, for example has the meaning *pale* in *pale red* and *bright* in *bright yellow* been excluded in the networks, since it is merely a term for saturation and not hue (Biggam 2012).

The language families that I chose were the Indo-European and the Sino-Tibetan, since they are two of the largest and most important families with many speakers. With 2.6 billion speakers, the Indo-European language family is the largest language family in the word, the family has the widest distribution around the world and includes most of the

languages spoken in Europe (Thompson 2016). The Sino-Tibetan language family is also one of the largest language families in the world and consists of two main branches, the Sinitic part with the Chinese languages (or dialects) and the Tibeto-Burman branch with the remaining languages. Chinese is one of the world's largest speech communities with more than 1.1 billion speakers and is considered to be one of the most important languages in the world with regards to its cultural significance, numbers of speakers and influence on other languages. The Tibeto-Burman branch consists of hundreds of greatly differential languages from a large area amongst other in India, Vietnam, Himalayan region and China. These two branches together form a language family of tremendous range in terms of complexity and time-depth (Matisoff ed. 2016). From the two families, I selected ten languages from a range of different branches within the families, leaving me twenty languages to investigate, including its ancestor languages. Those languages were: Spanish, English, Swedish, Irish, Greek, Albanian, Russian, Latin, German, Polish, Classic Tibetan, Burmese, Bodo, Pa'o, Jingpho, Southern Qiang (Mianchi), Apatani, Newari, Mandarin Chinese and Lepcha.

Since the available information on colour terms was rather constricted for the Sino-Tibetan languages, the intention was to include languages from as many different branches as possible while still take into consideration the amount of accessible data. Unfortunately, the chosen languages were to a certain degree controlled and regulated by the two major sources available, which is why the study included both Latin and Spanish and not all branches in neither of the families. Concerning the chosen languages from the Indo-European family, the objective was still to cover as many branches as possible, but was again restricted by the limitations of the etymological resources.

4.2 Etymologies

Since I choose to do a descriptive research on colour terms in various languages, my material was already existing data from different etymological databases and lexicons. For the Sino-Tibetan colour terms, STEDTS (Sino-Tibetan Etymological Database and Thesaurus) was my main and only source since it contained a vast range of synchronic and diachronic entries that would be difficult for me to find elsewhere looking at specific languages (Matisoff ed. 2016). Many branches and languages stays rather

unexplored in the Sino-Tibetan family, and less so when it comes to colour terms, which is why my choice of languages was restricted by the etymological data available (Matisoff ed. 2016).

My main source for the Indo-European colour terms was Brill Dictionary Online and additional etymological dictionaries for the specific languages. There was a lot more available etymological information on these well-studied languages, which is why concerning some colour terms, this dataset is bigger than the Sino-Tibetan.

In the cases where I found etyma on, for instance, a colour term's Proto-Indo-European root, its form in its old variety, but lacked the obvious connection to its synchronic term, (since in many cases that is not the interest of etymological dictionaries that focuses on the very old reconstructions), I filled that link in by myself but only if the connection itself couldn't possibly be questioned.

I did not select a specific time range, instead I chose to investigate as far back as I could find data available for when doing my etymological research. The furthest back I found data, was from the Proto-Indo-European era around 3500 BC and the Proto-Tibeto-Burman and Proto-Chinese era around 4000 BC (Matisoff ed. 2016).

4.3 Semantic classifications of lexical meaning

Once my dataset was completed, I needed a way to group my lexical meanings in order to make my data more manageable and my analysis of the results simpler. The classifications are merely a way of simplifying the comparison by providing abstract groups, and is not to be seen as a way of diminishing the actual lexical meanings.

I therefor used The Loanword Typology (LWT) meaning list (Haspelmath & Tadmor 2009) which is based on the Intercontinental Dictionary Series List (Kay & Bernard eds. 2015) which in turn is an adaptation of Buck's list from *Dictionary of Selected Synonyms in the Principal Indo-European Languages* from 1949. The list consists of 24 semantic fields of which I, based on the selection of lexical meanings in my data, used the following eight classifications: The Physical World, The Body, Sense Perception, Food and Drink, Animals, Basic Actions and Technology, Agriculture and Vegetation and

Colour. The eighth classification Colour, I added myself to differentiate between colour terms and other sense perceptions.

These categories are fairly wide and the classification of the lexical meanings are not always obvious and clear. A word could for example possibly go under both Food and Drink and Agriculture and Vegetation. Nevertheless, these classifications are useful for a first orientation and should be seen as a rough way of collecting the lexical meanings to be able to get a comprehensive overview. From here on, I will use the term 'semantic fields' or 'semantic classifications' when referring to these groupings, in order to differentiate them from the *colour categories.*

4.4 Semantic networks

The fascination for the human semantic memory and how it is able to store and retrieve facts about thousands of concepts, is the thought behind the dictionary Wordnet. It is a huge dictionary that organizes words into semantic networks in order to represent the lexicalization patterns of languages and show the conceptual density of the vocabulary. The dictionary maps the lexicon to concepts to disambiguate word sense by representing the various relations such as synonymy, meronymy etc. It enables discovery of alternative expressions in a language and expansion of words to extract semantically close or related words (Fellbaum 2012).

When investigating colour terms one can't help but be fascinated by the vast amount of associations and connections of concepts that they have, which is why I chose to represent these complicated lexicalization patterns visually. Semantic networks can more clearly show the semantic relations and expand the word into further understanding of it.

To get a more comprehensive, visual representation of my data, I created four semantic networks that I completed with Google Fusion Tables. The idea came from CLICS (Database of Cross-Linguistic Colexifications) that uses visual networks in order to show areal patterns of colexification. A list with the number of semantic links to a certain concept is available together with the word expressing the specific concept (List, Mayer, Terhalle & Urban 2014).

4.5 Data

My tool for compiling my data was the DiACL (Diachronic Database of Comparative Linguistics) (Carling ed. 2017) and its subsection 'Lexicology' that is a tool for creating comparative lexical cognacy databases, joining information drawn from comparative methods with meeting the requirements of phylogenetic & lexicostatistical analysis. The aim of the 'Lexicology' section is to create datasets that combine lexicography and comparative linguistics, hence adding both dictionary information about the lexical meaning and etymological information. The subsection is first organized in languages that are arranged geographically with a focused macro-area, the next level is 'Word Lists' that can be defined as structures to organize lexical meanings in functional hierarchies. Under that level is 'Word List Item', where all the lexical meanings are presented together with a map of all lexemes in the macro-area for the chosen lexical meaning. The final step is the level 'Lexeme' that both attested and reconstructed languages can be given. The 'Lexeme' is the focus of the section 'Lexicology' and include information about the lexical meaning such as: transcription, transliteration, IPA, grammatical data, meaning field (synchronic polysemy accounted for here), note and source (literary/informants). 'The Etymology Controller Tool' enables any lexeme in the database to be linked to any other lexeme either as 'Descendant Lexeme' or 'Ancestor Lexeme'. From there, the internal relations of an etymology are represented as boxes (lexemes) and arrows (relations) as shown in figures 2 and 3 below (Carling ed. 2017).

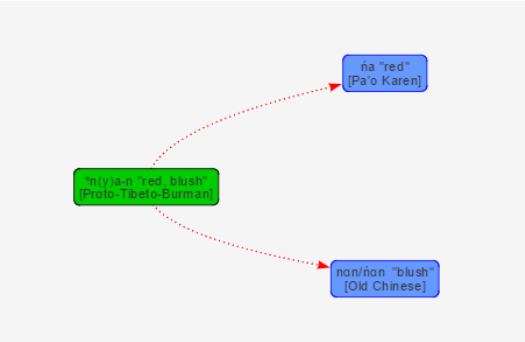

Figure 3: Etymology of the word ña *for Pa'o Karen (Sino-Tibetan) (Carling ed. 2017).*

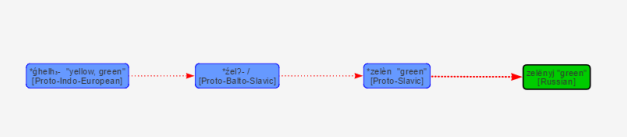

Figure 4: Etymology of the word zelënyj in Russian (Indo-European) (Carling ed. 2017).

The first step of my collection of data was to consult etymological dictionaries, in order to choose which of the Indo-European languages that I could find data on my chosen colour terms. Most of the Sino-Tibetan languages were not in the database which meant that I had to, first of all, add them into the database using Glottolog (Hammarström, Forkel, Haspelmath, Bank (ed.) 2017) as reference. Once I found my dictionaries, I started adding the terms with all the information available and first when I had entered all my terms in the database, I linked the ancestral terms with the descendant terms and left the synchronic and diachronic terms without etyma as they were. My focus in this research was how the colour term has developed semantically and hence I wanted to look at the generic meaning and trace it back, or the contrary. My main concern was to get as much data on each colour term as possible and less about getting languages with the highest possible geographical range. In the colexification section of my investigation, I excluded the reconstructed proto-forms.

My dataset is created under Lexicology → 'Word List Colour terms – Eurasia' where the ten colour terms has its own 'Word List Item' section named 'colour term black', 'colour term blue' etc. The languages represented are the 20 mentioned in section 4.1 plus its ancestral languages, including the reconstructed. Not all the lexemes had any etymological information, but they are still kept in the dataset, which is why the number is so high. An overview of the dataset is represented in the following table (table 1) and for access to the entire set, visit: https://diacl.ht.lu.se/WordListCategory/Details/11025

Table 1: Overview dataset (Carling ed. 2017)

Category	**Number**
Families	2
Languages represented from the Sino-Tibetan family	22
Languages represented from the Indo-European family	34
Colour terms	10
Lexemes in total	874

4.6 Method Discussion

By investigating languages from two families, a cross-linguistic perspective was added and made an analysis including the debate universalism vs. relativism possible. The thought behind including the Sino-Tibetan language family was mostly due to the fact that there was an easy manageable and reliable database available that included most of my colour terms. These otherwise quite unexplored languages could hence be a part of my time-limited research.

Concerning the colour term *orange*, I had some issues in both obtaining information and find out if the term I collected was a reference to the noun (fruit) or in fact the colour term. In the cases where it specifically was announced that it was exclusively a noun, I choose to exclude it from my dataset and in the case where the meaning included colour, I choose to make it a part of my research. My second issue was the colour term *purple*, that in many languages didn't have any etymological information, I also included the entries *lilac* and *violet*, here I applied the same model of consistently excluding the terms stating the meaning as a flower or noun and including the ones mentioning colour in the meaning. Still, I chose to include both terms in my research even though I realized quite quickly that the information available was restricted and in many cases unreliable.

Concerning all my colour terms, I stumbled on the issue of how to interpret the various lexical meanings, since grammatical and semantic information not always was available. While the lexical meaning *bark of tree* is not open for interpretation, *bay* or *ass* are more ambiguous and open for comprehension. In these cases, I simple proceed with the most likely interpretation, given context.

When analyzing the etymological dictionaries, I tried to be as consistent as possible and exclude any terms that expressed any doubtfulness regarding the etymological history. The same policy of consistency was practiced when extracting the etymological data, but the resources that I used were overall reliable references which simplified my research greatly.

5. Results and analysis

5.1 Results: colour by colour

RED

RED was colexified fifteen times (see Appendix A) with twelve lexical meanings. Three of the times with *fiery* and the rest with lexical meanings from various semantic fields as Sense Perception: *glowing, murderous, deadly;* Colour: *purple*, *ruddle, pink, tawny*; The Body: *blood, red-haired*; The Physical Word: *vermilion and* Animals: *monkey*.

When looking at the diachronic derivations, RED had most connections to the semantic fields Sense Perception and lexical entries as for example: *dark, hot, glowing*, Colour and lexical meanings such as: *brown, yellow, black* and *purple and* The Body and lexical entries as for example: *blood*, *wound* and *ashamed*. Added semantic fields in the diachronic change were Agriculture and Vegetation, Basic actions and Vegetation and Food and Drink. Lexical meanings that the two families had in common when looking at both the synchronic and the diachronic perspective were: *purple, vermilion (cinnabar), dark* and *blood*.

The type of semantic change that mostly occurred was narrowing (90%), as shown in the example below:

> *s-kyaŋ *red, blushing* (Proto-Tibeto-Burman) → cheng *red* (Jingpho)
> *dherg- *dark* (Proto-Indo-European) → *dergo- *red, blood-red* (Proto-Celtic) → *derg red* (Old Irish)

The most common change in the Sino-Tibetan languages, was from abstract to concrete (65%), as shown in the example below:

> *lɯŋ *red* (Proto-Tani) → ma-laŋ *rust* (Apatani)

In the Indo-European languages, there was almost as much change from concrete to abstract (46%) as abstract to concrete (54%), as shown in the examples below:

> *č'rm'n *red* (Proto-Slavic) → čeremnój *red-haired, ginger* (Russian)
> *kwr-m-i- *worm, maggot* (Proto-Indo-European) → czerwony *red* (Polish)

Table 2: RED. Table of the diachronic connections (light blue boxes represents the Indo-European languages, the dark blue boxes the Sino-Tibetan languages and the medium blue boxes, meanings that occur in both families)

Animals	Sense	Colour	Body	Agriculture	Basic Actions	Food and Drink	Physical World
worm	*murderous*	*tawny*	*red-haired*	*earth*	*hit*	*ginger*	*vermilion (cinnabar)*
maggot	*deadly*	*brown*	*colour of bruises*	*soil*	*lead*	*berry*	*gold*
rat	*dark*	*yellow*	*wound*				*rust*
monkey	*glowing*	*pink*	*blood*				
	fiery	*purple*	*ashamed*				
	hot	*black*	*blush*				

GREEN

GREEN was colexified fourteen times (see Appendix A) with eight different lexical meanings. Six of them occurred with *blue,* which wasn't a surprise since GRUE is a well-known macro-category referring to both GREEN and BLUE (Biggam 2012). The rest colexified with meanings in the semantic fields Agriculture and Vegetation: *unripe, growing well* and *leafy*, Sense Perception: *dark*, *light, Colour: yellow, blue* and Food and Drink: *vegetable*.

In the diachronic derivations, GREEN had most connections to the semantic fields Sense Perception and lexical meanings as: *bright, to glow, fair*, Agriculture and Vegetation and lexical meanings such as *grass, grow*, *blossoming*, Colour and meanings like: *blue, grey, white, yellow* and The Body and lexical meanings such as *breath*, *born* and *bile*. Added semantic fields in the diachronic derivations were The Body, The Physical World, Basic Actions and Technology and Animals. Words that the two families had in common when looking at the synchronic and the diachronic perspective were: *yellow, blue* and *grow.*

The semantic change that mostly occurred was narrowing (88%), as in the example below:

> *s-riŋ ⩙ *s-r(y)aŋ *green, live, alive, raw, give birth* (Proto-Tibeto-Burman) → *Hriŋ, hriŋ-ll *green, alive, fresh* (Proto-Kuki-Chin)
>
> glas *green, blue* (Old Irish) → glas *green* (Irish)

The semantic change went from abstract to more concrete (91%) in the Sino-Tibetan languages, as the following example:

*s-riŋ ⪤ *s-r(y)aŋ *green, live, alive, raw, give birth* (Proto-Tibeto-Burman) → *hriŋ *alive* (Proto-Tangkhulic) → haŋ-sur *breath, life* (Bodo)

In the Indo-European languages, there was exactly as much change from abstract to concrete as the reverse, as shown in the example below:

*ghroh1-ni- (Proto-Indo-European) → *grōni- *green* (Proto-Germanic) → *grōan *to grow* (Proto-Germanic)

*bhloh1/3-ro-, *bhleh3-ro- 'blossoming' (Proto-Indo-European) → blertë 'green' (Albanian)

Table 3: GREEN. Table of the diachronic connections (the light blue boxes represents the Indo-European languages, the dark blue boxes the Sino-Tibetan languages and the medium blue boxes, meanings that occurs in both families)

Animals	**Sense Perception**	**Colour**	**The Body**	**Agriculture and Veg.**	**Basic Actions**	**Food and Drink**	**The Physical World**
bear	*to gleam*	*blue-green*	*bile*	*blossoming*	*shoot*	*leek*	*lock*
	to glimmer	*grey*	*born*	*grow*	*give birth*	*uncooked*	*gold*
	to glow	*blue*	*live*	*sprout*		*raw*	
	leek-colour	*yellow*	*breath*	*new twigs*		*raw meat*	
	shining	*white*	*life*	*grass*			
	bright			*unripe*			
	fresh						
	alive						
	fair						

BLUE

BLUE was colexified twenty-three times (see Appendix A) with ten different lexical meanings. Six of the times with the colour term *green,* again not such a surprise given the well-known close connection between the two categories and the macro-category GRUE (Biggam 2012). Three of the times BLUE colexified with *dark* and the rest of the

colexifications occurred with the semantic fields Colour: *green*, *grey, black, lilac,* Sense Perception: *dark, pale, livid*, Agriculture and Vegetation: *violet* and The Body: *bloodshot.*

When looking at the diachronic derivations, BLUE had connections to the semantic fields Colour and lexical meanings such as: *yellow, purple, white*, Sense Perception and lexical meanings like: *envious, fair, whole*, Agriculture and Vegetation and meanings as: *heaven, wind, sky*, The Body and lexical meanings as for example: *grey-haired, bloodshot*, Food and Drink and lexical meanings like: *leek, mouldy*, The Physical World and the meaning *lapis lazuli* and Animals and the lexical meaning *pigeon.* Added semantic fields represented in the diachronic change were: The Physical World, The Body, Animals and Food and Drink. Words that the two families had in common when looking at both the synchronic and the diachronic perspective were: *green, grey, dark, yellow* and *black*.

The only semantic change that occurred in the Sino-Tibetan languages was narrowing, as in the example below:

*s-ŋow *blue, white, green yellow* (Proto-Tibeto-Burman) → *sŋon po *blue* (Tibetan)

The semantic change that occurred were equally divided between narrowing and broadening in the Indo-European family, as the examples below:

blāo *blue, dark, grey* (Old High German) → blau *blue* (German)

*blēwa- *blue* (Proto-Germanic) → blár *blue, livid, black* (Old Norse)

The changes were going mostly from concrete to abstract (73%), as shown in the example below:

*golǫbь *pigeon* (Old Church Slavonic) → goluboj (light) *blue* (Russian)

caelum *sky, heaven, vault of heaven* (Latin) → caerulus *blue* (Latin)

Table 4: BLUE. Table of the diachronic connections (the light blue boxes represents the Indo-European languages, the dark blue boxes the Sino-Tibetan languages and the medium blue boxes meanings that occurs in both families)

Animals	**Sense Perception**	**Colour**	**The Body**	**Agriculture and Veg.**	**Food and Drink**	**The Physical World**
pigeon	*discoloured*	*dun*	*grey-haired*	*violet*	*leek*	*lapis lazuli*
	whole	*purple*	*bloodshot*	*heaven*	*mouldy*	

	of a dull colour	*turquoise*		*earthenware*		
	envious	*lilac*		*wind*		
	colour	*grey*		*calendula oficinalis (flower)*		
	livid	*green*		*Sky*		
	dark	*yellow*				
	fair	*black*				
		white				

YELLOW

YELLOW was colexified ten times (see Appendix A) with nine different lexical meanings. Two of them occurred with *gold* and the rest with lexical meanings from the semantic fields The Physical World: *white spotted,* Sense Perception: *faded, pale, bright colour of health*, Agriculture and Vegetation: *dry as leaf, fallow* and the Body: *colour of bruises* and *blonde*.

When looking at the diachronic derivations, YELLOW was connected to Sense Perception and lexical meanings like *shine, pale, grateful*, Colour and meanings like *blue*, *green, red,* The Physical World and lexical meanings as: *dirty, lamp, gold,* Food and Drink and lexical entries like *honey, spice*, *turmeric*, Agriculture and Vegetation and meanings like: *amber, citrus tree* and Basic actions and Technology and the meaning *a float.* There were three semantic fields added in the diachronic perspective: Basic actions and Technology, Colour and Food and Drink. Words that the two families had in common when looking at both the synchronic and the diachronic perspective were: *pale, white* and *green*.

The only semantic change that occurred in the Sino-Tibetan languages was narrowing, as in the example below:

*b/s-wa *white, bright, yellow* (Proto-Tibeto-Burman) → ə-wâŋ *yellow* (Burmese)

The semantic change that mostly occurred in the Indo-European languages was broadening (70%), as in the example below:

*flāwo- *yellow* (Proto-Italic) → flāvus *yellow, blonde* (Latin)

The changes were going mostly from abstract to more concrete (80%) in the Sino-Tibetan languages as in the example below:

*s-mar *yellow, gold, butter, oil* (Proto-Tibeto-Burman) → bʑu mar *oil light/lamp* (Tibetan)

There was almost as many changes from concrete to abstract (46%) as abstract to concrete (54%) in the Indo-European languages, as shown in the example below:

*falwa- *pale* (Proto-Germanic) → falo *faded, fallow, yellow* (Old High German)

Žëlč' *bile* (Russian) → žëltyj *yellow* (Russian)

Table 5: YELLOW. Table of the diachronic connections (the light blue boxes represents the Indo-European languages, the dark blue boxes the Sino-Tibetan languages and the medium blue boxes meanings that occurs in both families)

Sense Perception	**Colour**	**The Body**	**Agriculture and Veg.**	**Food and Drink**	**Basic Actions**	**The Physical World**
bitter	*brown*	*bile*	*citrus tree*	*honey*	*a float*	*a buoy*
to gleam	*black*	*blonde*	*fallow*	*oil*		*dirty*
to glow	*red*	*colour of bruises*	*amber*	*butter*		*gold*
to glimmer	*grey*	*bright colour of health*	*bay*	*spice*		*golden*
grateful	*green*		*dry as leaf*	*turmeric*		*white spotted*
pleasing	*white*					*lamp*
faded	*blue*					
dusky						
dark						
pale						
bright						
fair						

shine						
light						

GREY

GREY was colexified sixteen times (see Appendix A) with thirteen different lexical meanings. Those were spread over the semantic fields Sense Perception: *dull, dark, pale, envious, faded, exhausted*, Colour: *green, black, tawny:* The Body: *grey-haired, grey-eyed,* The Physical World: *poor* and Food and Drink: *mouldy.*

Of the diachronic derivations, GREY was connected to mostly Sense Perception and lexical meanings like *pale, dull* and *envious*, Colour and lexical meanings as *blue, white, black*, The Body and lexical meanings like *hoary, ass*, Food and Drink and the meaning *mouldy*, Agriculture and Vegetation and the lexical meaning *fallow* and The Physical World and the meaning *poor*. The added semantic field in the diachronic derivations was only Agriculture and Vegetation. Words that the two families had in common when looking at both the synchronic and the diachronic perspective were: *black, pale* and *faded*.

The most occurring change was narrowing (75%), as shown in the examples below:

líath *grey* (Old Irish) → liath *grey, grey-haired, mouldy* (Irish)

*haswa- *grey* (Proto-Germanic) → heswe *pale, dull* (Middle High German)

The changes were mostly going from abstract to concrete (75%), as shown in the example below:

grīs *grey* (Old High German) → greis *aged* (German)

Table 6: GREY. Table of the diachronic connections (the light blue boxes represents the Indo-European languages, the dark blue boxes the Sino-Tibetan languages and the medium blue boxes meanings that occurs in both families)

Sense Perception	**Colour**	**The Body**	**Agriculture and Veg**	**Food and Drink**	**The Physical World**
bright	*tawny*	*ass*	*fallow*	*mouldy*	*poor*

clear	*brown*	*grey-haired*			
dark	*blue*	*hoary*			
envious	*white*	*grey-eyed*			
aged	*lilac*				
shiny	*green*				
pale	*black*				
faded					
dull					
exhausted					

BROWN

BROWN was colexified eleven times (see Appendix A) with eight different lexical meanings and three of them was with *dark*. Only one colexification and no diachronic derivations was gathered from my study from the Sino-Tibetan languages, and that was with *red.* The rest was with meanings from the semantic fields Sense Perception: *dusky, swarthy, sandy, dark*, Agriculture and Vegetation: *chestnut,* Colour: *red-yellow* and The Physical World: *golden*.

When looking at the diachronic derivations, BROWN was connected to the semantic fields Sense Perception and lexical meanings as: *pale, sandy, dark,* Agriculture and Vegetation and meanings like: *amber, chestnut*, The Physical World and lexical meanings as: *dirty, soot*, Food and Drink and the lexical meanings: *cinnamon, honey*, Animals and the meanings: *beaver, elk*, Colour and the meanings *red, red-yellow* and Basic actions and Technology and the lexical meaning *cut*. Three semantic fields were added in the diachronic change: Basic actions and Technology, Food and Drink and Animals. A word that the two families had in common when looking at the diachronic perspective was: *red.*

The semantic change that occurred most was broadening (72%), as seen in the examples below:

korjca *cinnamon* (Russian) → koričnevyj *brown* (Russian)

*dusno- *dark, brown* (Proto-Celtic) → donn *brown* (Old Irish)

The changes were mostly going from abstract to more concrete (73%) as in the example below:

*bhe-bhr-ú-, *bhe-bhr-o- (or *bhi-bhr-o-) *brown; brown animal, beaver* (Proto-Indo-European) → *fifro- / fefro- *beaver* (Proto-Italic)

Table 7: BROWN. Table of the diachronic connections (the light blue boxes represents the Indo-European languages, the dark blue boxes the Sino-Tibetan languages and the medium blue boxes meanings that occurs in both families)

Animals	**Sense Perception**	**Colour**	**Agriculture and Veg.**	**Food and Drink**	**The Physical World**	**Basic Actions**
beaver	*smoky*	*red-yellow*	*chestnut*	*honey*	*dirty*	*cut*
elk	*pale*	*red*	*amber*	*cinnamon*	*soot*	
	dusky		*bark of tree*		*golden*	
	swarthy					
	sandy					
	dark					

PURPLE

PURPLE was colexified six times (see Appendix A) with four different lexical meanings, two of them with *blue* and two with *red*. Only two colexifications with *red* was gathered from the Sino-Tibetan languages and no diachronic derivations. The rest of the colexifications occurred with lexical meanings in the semantic fields The Physical World: *purple dye* and The Body: *blood*.

When looking at the diachronic derivations, PURPLE was connected to the semantic fields The Physical World and the lexical meanings: *indigo, purple dye*, Agriculture and Vegetation and the meanings: *lilac, violet*, Colour and the lexical meanings: *blue, grey*, The Body and the meanings: *grey-haired, blood* and Food and Drink and the lexical meaning *mouldy*. Extended semantic fields represented in the diachronic derivations

were: Agriculture and Vegetation and Food and Drink. The two language families did not have any lexical meanings that co-occurred.

It occurred as much broadening as narrowing, as shown in the examples below:

lila *lilac* (colour and flower) (Spanish) → lila *purple* (Swedish)

ἴον violet (flower) (Classical Greek) → ἰόις *violet-colored, deep blue* (Classical Greek)

Table 8: PURPLE. Table of the diachronic connections (the light blue boxes represents the Indo-European languages, the dark blue boxes the Sino-Tibetan languages and the medium blue boxes meanings that occurs in both families)

Colour	**The Body**	**Food and Drink**	**Agriculture and Veg**	**The Physical World**
blue	*grey-haired*	*mouldy*	*lilac (flower)*	*indigo*
grey	*blood*		*violet*	*purple dye*
red				

ORANGE

ORANGE only had one diachronic derivation in the Indo-European languages, to the lexical meaning *gold* in the semantic classification The Physical World.

BLACK

BLACK was colexified thirteen times (see Appendix A) with fourteen different lexical meanings. Four of them was with *dark* and the rest were with lexical meanings in the semantic fields: Sense Perception: *gloomy, livid, lamentable*, Colour: *grey, white, blue, dun*, The Physical World: *ink, dirty, unclean* and Animals: *monkey, crow and raven.*

Diachronically, BLACK had most connections to the semantic fields Sense Perceptions and lexical meanings as: *shine, blind, deep*, Colour and meanings like: *blue, brown, white*, The Physical World and lexical meanings such as: *fireplace, gold, ink*, Animals and meanings like: *raven, monkey, cattle* and Basic actions and Technology and the lexical meaning *burn.* There was only one added field in the diachronic derivations, and that is Basic actions and Technology. Words that the two families had in common when looking at

both the synchronic and the diachronic perspective was: *dirty, grey, dark, white, raven* and *ink*.

The semantic change that occurred mostly was narrowing (84%), as in the example below:

*tsya(k/ŋ) *red, dark-coloured, black* (Proto-Tibeto-Burman) → chang *black* (Jingpho)

niger *black, dark* (Latin) → negro *black* (Spanish)

The changes were mostly going from abstract to concrete (89%) in the Sino-Tibetan languages, as in the example below:

*s-nak *black* (Proto-Tibeto-Burman) → nə̀-ŋá *crow* (Southern Qiang)

There was exactly as much abstract to concrete as the reverse in the Indo-European languages, as shown in the example below:

*blanka- *colourless* (Proto-Germanic) → blakkr *black, dun-coloured* (Old Norse)

*h2eh1-t(-)r- *fireplace* (Proto-Indo-European) → ātro- *black* (Proto-Italic)

Table 9: BLACK. Table of the diachronic connections (the light blue boxes represents the Indo-European languages, the dark blue boxes the Sino-Tibetan languages and the medium blue boxes meanings that occurs in both families)

Animals	**Sense Perception**	**Colour**	**Basic Actions**	**The Physical World**
raven	*blind*	*dun*	*burn*	*fireplace*
cattle	*shine*	*blue*		*first main room in roman-style house*
monkey	*lamentable*	*brown*		*unclean*
crow	*livid*	*white*		*Ink*
	dusky	*grey*		*gold*
	colourless	*red*		*dirty*
	dark			
	deep			
	gloomy			

WHITE

WHITE was colexified twenty-three times (see Appendix A) with eleven different lexical meanings. Four of the times with *bright* and the rest with lexical meanings in the semantic fields Sense Perception: *clear, airy, pale*, *brilliant*, *gleaming, shining, fair*, Basic Actions and Technology: *become*, *quick, agile*, The Physical World: *money, silver*, Colour: *black* and The Body: *fair-haired*.

When looking at the diachronic derivations, WHITE was connected to the semantic fields Sense Perception and lexical meanings like *fair, shine*, *light*; Colour and lexical meanings like *blue, green* and *yellow*, Basic actions and Technology and meanings as: *flay, quick, burn*, The Body and lexical meanings like: *skin, health*, The Physical World and the meanings: *silver, money* and Animals and the meanings *goat and swan*. Only one semantic field was added in the diachronic perspective and that was Animals. Words that the two families had in common when looking at both the synchronic and the diachronic perspective were: *clear, pale, shine, brilliant, fair, bright, silver, yellow* and *green.*

The semantic change that mostly occurred was narrowing (90%), as in the example below:

*plu[1] *white, silver, money* (Proto-Lolo-Burmese) → phlu *white* (Burmese)

*bhereĝ *to shine, white* (Proto-Indo-European) → i bardhë *white* (Albanian)

The changes mostly went from abstract to more concrete (71%) in the Sino-Tibetan languages, as in the example below:

*b/s-wa *white, bright, yellow* (Proto-Tibeto-Burman) → hu-to *light* (Apatani)

The changes were mostly going from more concrete to abstract (69%) in the Indo-European languages, as in the example below:

*blank- *to shine* (Proto-Germanic) → blancus *white* (Vulgar Latin)

Table 10: WHITE. Table of the diachronic connections (the light blue boxes represents the Indo-European languages, the dark blue boxes the Sino-Tibetan languages and the medium blue boxes meanings that occurs in both families)

Animals	**Sense Perception**	**Colour**	**The Body**	**Basic Actions**	**The Physical World**
goat	*colourless*	*lilac*	*skin*	*show*	*silver*
swan	*gleaming*	*yellow*	*faired-haired*	*quick*	*money*
	appear	*green*	*health*	*agile*	
	clear	*black*		*flay*	
	wan	*blue*		*burn*	
	airy				
	degree of cold				
	fair				
	bright				
	shine				
	brilliant				
	pale				
	light				

5.2 Semantic networks

The following networks were based on my dataset (see section 2.5) and was manually extracted into an Excel-file and then gathered in Google Fusion Tables. The networks are meant to visualize the same information as described in previous section 5.1. The first two represent the colexifications in the respective families and the last two represent the diachronic derivations (including the colexifications) in the respective families. The largest dots are the ones with the most connections and the most central terms in the networks are the ones with the most connections to the lexical meanings represented in the network.

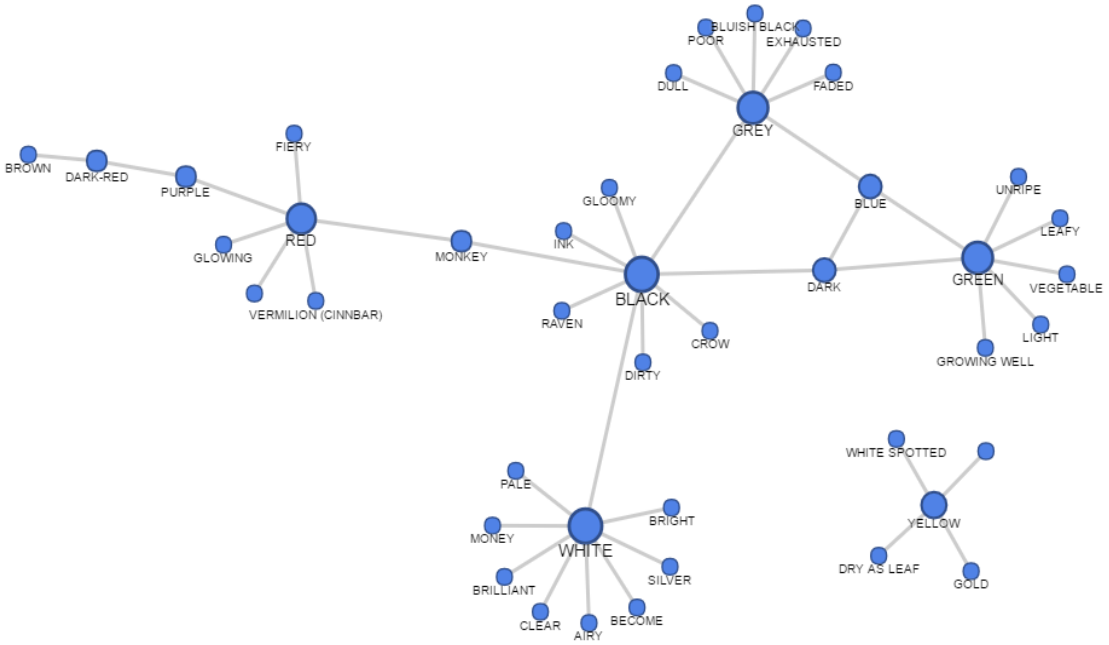

Figure 5: Synchronic colexifications in the Sino-Tibetan languages

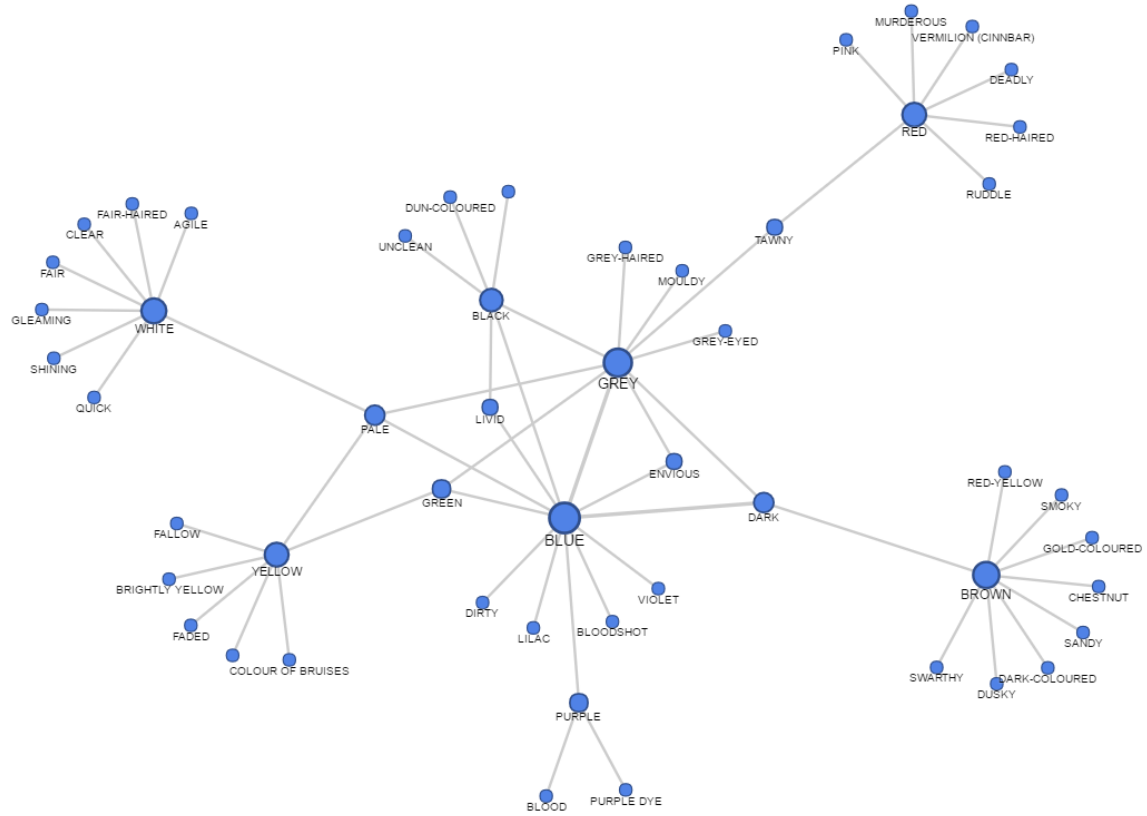

Figure 6: Synchronic colexifications in the Indo-European languages.

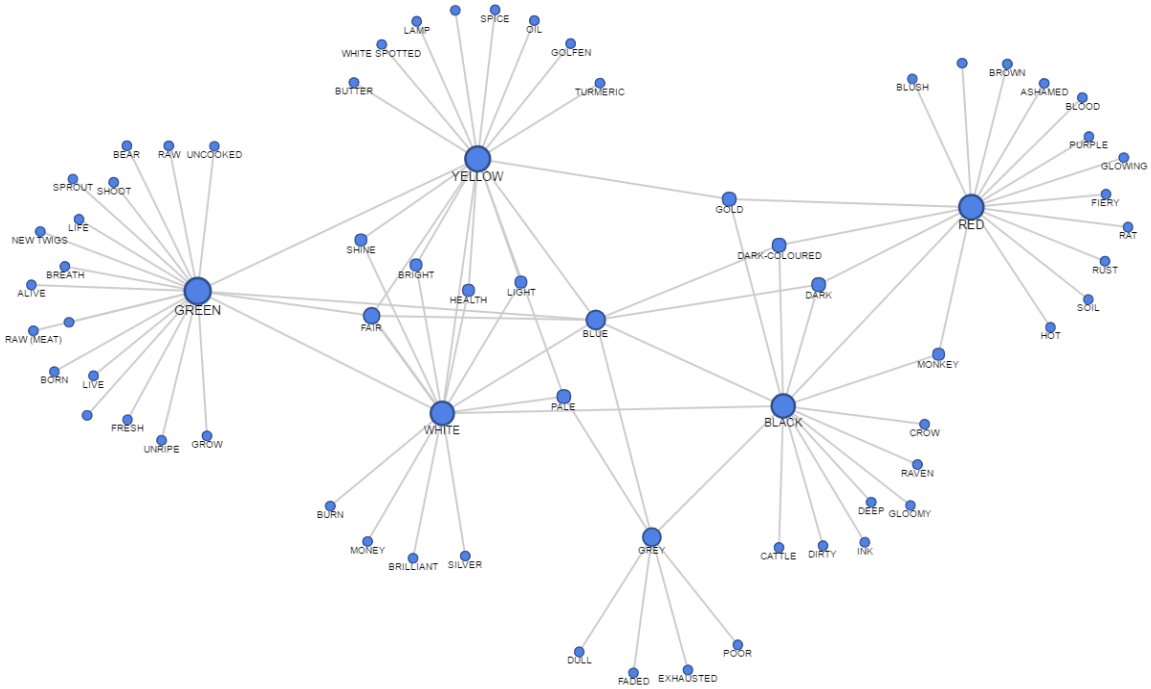

Figure 7: Diachronic connections in the Sino-Tibetan languages.

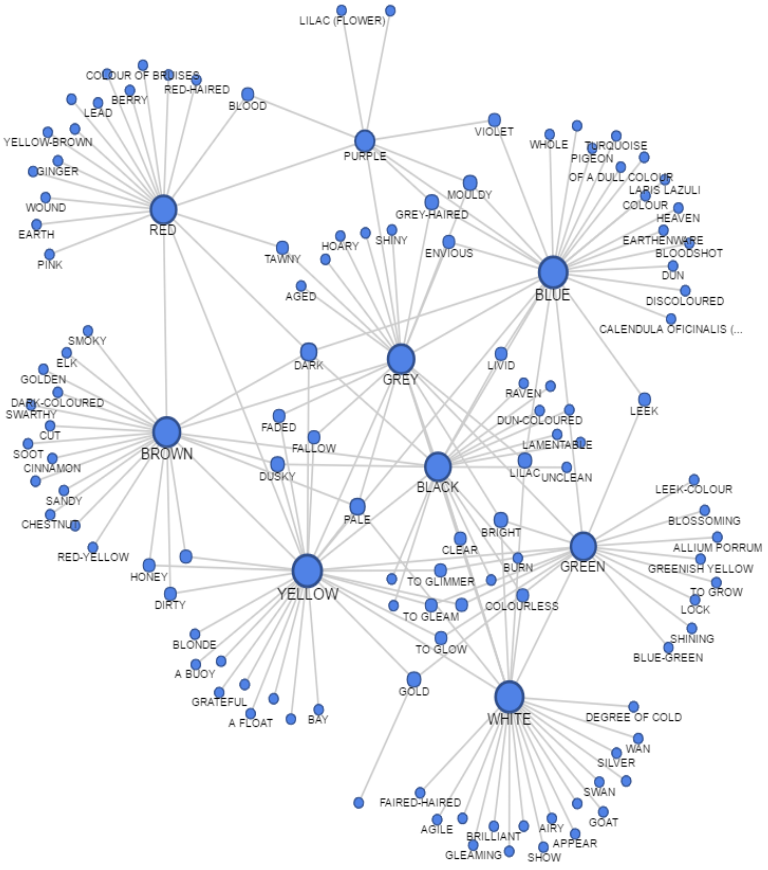

Figure 8: Diachronic connection in the Indo-European languages.

6. Conclusions

I will here present a summary of my results by going back to my initial research questions and end this chapter with some suggestions for further research in the field.

Which *lexical meanings* colexify with colour terms?

The colour terms that I investigated showed a huge range of colexified meanings. There was an overrepresentation of other colour terms, adjectives and nouns. Descriptive meanings such as *leafy, dirty* and *glowing* were recurrent, as well as words connected with the body and its physical appearance as for example *blood, blonde* and *grey-haired*. Less occurring colexifications, but still present, were words of physical objects as for example *ink, gold* and *chestnut* and words of animals such as *raven* and *monkey.*
RED had words connected with danger such as *murderous, deadly* and *blood*. GREEN colexified with words connected to nature as *unripe, leafy* and *vegetable*. BLUE colexified with different colour terms as *green,* which was to be expected since GRUE is, as already mentioned, a well-known macro-category in colour semantics (Biggam 2012). YELLOW colexified with lexical meanings like *dry, pale, faded* and words with more positive connotations like *bright colour of health* and *gold*. GREY was colexified with meanings of the more negative sense such as *dull, poor* and *exhausted* and interestingly also with both *dark* and *pale*. BLACK colexified with lexical meanings such as *unclean, dark* and *dirty*. WHITE colexified with lexical meanings connected to light as: *bright, brilliant, gleaming* and *shining* and objects as *money* and *silver*. BROWN mostly colexified with descriptive lexical meanings as *golden, swarthy, dark* and PURPLE colexified with physical objects such as *blood* and *purple dye.*

Which *classifications* of lexical meanings colexify with colour terms?

There was an extremely wide spread of derived concepts in most of the colour terms investigated with the exceptions of *brown, purple* and *orange.* These colour terms all have in common being in the last (and second last) stage in the evolutionary sequence created by Berlin & Kay (1969) (see figure 1). One of the overall trends was that all the colour terms colexified with the semantic field Sense perception. The semantic field Colour was also frequent, perhaps because they are both fairly vague groupings.

RED, GREY, PURPLE, BLACK colexified overall with lexical meanings in the semantic field The Body, The Physical World, Colours, Sense Perception and Animals. GREEN had, not so surprisingly, an overrepresentation of colexifications with lexical meanings in the Agriculture and vegetation field and BLUE, YELLOW, BROWN showed similar patterns with colexifications with lexical meanings in the classification Agriculture and vegetation, Sense perception, Colours and Food and Drink. WHITE colexified with lexical meanings in the field Sense perception, The Physical World, Basic Actions and Technology and The Body. The majority of the colour terms had a range of colexified meanings from four different semantic fields. RED, WHITE and GREY had the widest range of colexified meanings with lexical meanings from five different semantic field, while PURPLE had the least range with only three semantic fields represented.

What type of semantic change occurs?

Without doubt the most common type of semantic change was narrowing that occurred in 76% of the cases, with exception from BROWN that had more broadening and PURPLE that had half broadening and half narrowing. The colour BLUE also had half broadening and half narrowing in the Indo-European languages and the colour YELLOW had more broadening than narrowing in the Indo-European languages. Many of the reconstructed colour terms were, as Biggam (2012) stated, broader than the modern terms, which could be an explanation to narrowing being the most common change occurring.

The most common way of semantic change was the direction from more abstract to more concrete meanings which occurred in 61% of the cases, except for the colour BLUE in both language families, GREEN in the Sino-Tibetan languages and WHITE in the Indo-European languages that had more changes from more concrete to abstract. There was a few colours in the Indo-European languages that had as many changes from abstract to concrete as the reverse, those were: RED, GREEN, YELLOW and BLACK. This was not entirely in line with previous studies that stated the most common pattern in semantic change in colour terms (and generally) to be from concrete to abstract (Campbell 2004, Warth-Szczyglowska 2014). This could perhaps be due to the fact that my investigation wasn't as profound as the previous studies already mentioned and that my investigation included many more colour terms. Perhaps if my dataset would contain more data, from particular the Sino-Tibetan family, the results would be different.

In what respect do the lexical meanings and classifications differ when looking at the colexification *and* semantic change?

Clearly the list of lexical meanings expanded when including the derived senses as well as the colexified. A clear pattern showed that all colour terms had derived lexical meanings within the semantic field already represented in the colexification. Additional lexical meanings from other semantic fields were added in the diachronic perspective concerning all colour terms. GREY and BLACK added the least with only one new additional category while GREEN added the most with four new fields.

In which respect do the results differ between the families?

The semantic networks presented in section 5.2 showed the first obvious observation: that all the colour terms were connected when looking at the diachronic perspective. Most connections, as we can see from the diachronic semantic networks, was to BLACK, WHITE and GREY in both language families since they shared the most diachronic and synchronic co-occurrences of lexical meanings. BLACK and WHITE is the first step in the suggested sequence of lexified colour terms by Berlin & Key (1969) (see figure 1). This supports the theory that at least BLACK and WHITE might be universal or at least very central in most languages as also suggested by Wierzbicka (Biggam 2012).

Reoccurring semantic fields represented in the both families were Sense perception and Colour. The Indo-European languages had a lot of words connected to danger in the category RED such as: *wound, hit* and *deadly,* while the Sino-Tibetan languages instead had words like *ashamed* and *blush*. GREEN had more meanings connected with light as for example *bright* and *glow* in the Indo-European languages, than the Sino-Tibetan that had more meanings derived connected to life such as *breath, life* and *born*. I found much more data for the Indo-European languages concerning both the categories BROWN and BLUE, though the Indo-European had more abstract lexical meanings as *livid*, *whole* and meanings connected to nature like *heaven, wind* and *sky.* I found a lot more concepts connected with food and objects in the colour YELLOW in the Sino-Tibetan languages. GREY had more meanings with derogative denotations such as *dull, pale* and *poor* in the Sino-Tibetan languages than in the Indo-European, that had more words connected to the body such as *ass, grey-haired* and *grey-eyed*. Some of these artefacts might be a

consequence of the small amount of data that I had and might only occur in the chosen languages. Nonetheless, the differences and variations are many and extensive.

Of the total 89 lexical meanings colexified with the colour terms in the study, only 28 of them co-occurred in both language families. WHITE had eight co-occurring lexical meanings, BLACK and BLUE five, while PURPLE did not have any. The fact that both BLACK and WHITE are amongst the colour terms that had the most co-occurring colexifications, is another argument for them being of a more universal nature. An interesting fact was that WHITE was colexified with BLACK in the Sino-Tibetan language family and not in the Indo-European, one possible explanation could be that they are both achromatic, namely colours without hue.

There was some difference between the families regarding the type of semantic change that occurred. Narrowing was clearly the most common, while there was a tendency for the Indo-European languages to have more broadening than the Sino-Tibetan languages. There was only a small difference concerning the tendency of the direction of the change, most of the times it was going from more abstract to more concrete. Concerning some colours, the Indo-European languages had a tendency to have more changes going from concrete to abstract. The fact that the results was more varied in the Indo-European languages, could be due to the fact that I had more data.

When looking at the results at a macro level, there seemed to be a clear pattern between the families when it comes to the semantic fields colexified and derived from the colour terms. The similar expansion of semantic fields in these genetically unrelated language families, in particular concerning the colour terms BLACK and WHITE, showed universal tendencies.

However, upon closer examination, there was clear cross-cultural differences in the lexical meanings, opening up for a more relativistic view. BLACK, for example had the semantic field Animals represented in both language families, but the lexical meaning in the Indo-European was *raven* and in the Sino-Tibetan *monkey*. There was very little co-occurrence and overlapping when it comes to lexical meanings colexified and diachronic connections even though it, at a categorical level, seemed to be a lot of conformity. There

was, nonetheless, too many exceptions and variation for this study to be supporting a universal theory in my view. The study is far too narrow and limited in data to make any clear conclusions in the debate between relativism and universalism, even though the results in my opinion points towards a relativistic theory. To sum up, the results showed that the colour terms of the two families have employed somewhat similar structures for semantic change along with the semantic domain that it extends to, but when looking closer at the exact lexical meaning, they differ greatly.

For further studies, one could, extract more information from the dataset, outside of the semantic field, such as analyzing for example the derivational morphology and iconicity. Adding more languages, preferable families without any historical or geographic contact, such as for example a South American language, would also be a great next step in the research. To widen the study by adding more colour terms including other non BCT's and more data, would be a further step in the study as well.

7. References and resources

7.1 References

Berlin, Brent & Kay, Paul (1969) *Basic Colour Terms - Their Universality and Evolution.* Berkeley: California University Press

Biggam, Carole Patricia (2012) *The semantics of colour - A historical approach.* Cambridge University Press: Cambridge

Campbell, Lyle (2004) *Historical Linguistics: An Introduction*. Cambridge: MIT Press

Durkin, Philip (2009) *The Oxford Guide to Etymology*. Oxford: Oxford University Press

Falkum, Vicente (2015) *Polysemy: Current perspectives and approaches.* Retrieved 2017-01-10 from: http://ac.els-cdn.com/S0024384115000170/1-s2.0-S0024384115000170-main.pdf?_tid=842c3f92-8bb4-11e6-ab2f-00000aacb35d&acdnat=1475751991_dc82da15c5cef3ef70aafecbd9abedc7

Fellbaum , Christiane (2012) *Wordnet* In Chapelle, A. Carol Encyclopedia of applied linguistics. Blackwell Publishing Ltd.

François, Alexandre (2008) *Semantic maps and the typology of colexification - intertwining polysemous networks across languages.* Retrieved 2017-01-10 from: http://alex.francois.free.fr/data/AlexFrancois_2008_SemanticMaps-Colexification_preprint.pdf

Geeraerts, Dirk (2010) *Theories of Lexical Semantics*. Oxford: Oxford University Press

Gieron-Czepczor, Ewa (2013) *Gloomy Images of Yellow and Żółty in a Corpus-Based Cognitive Study.* Research in Language. Dec2013, Vol. 11 Issue 4, p405-421. Retrieved 2017-02-13 from:

Goddard, Cliff & Wierzbicka, Anna (2013) *Words and meanings: lexical semantics across domains, languages and cultures.* Oxford: Oxford University Press. Retrieved 2017-03-10 from: http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199668434.001.0001/acprof-9780199668434

Hamilton, Rachael Louise (2016) *Colour in English - from metonymy to metaphor* Glasgow: University of Glasgow Retrieved 2017-02-08 from: http://theses.gla.ac.uk/7353/1/2016HamiltonPhD.pdf

List, Johann-Mattis, Mayer, Thomas, Terhalle, Anselm & Urban, Matthias (2014) *CLICS: Database of Cross-Linguistic Colexifications*. Marburg: Forschungszentrum Deutscher Sprachatlas. Retrieved 2016-12-20 from: http://CLICS.lingpy.org

Mailhammer, Robert (2015) Etymology In: Bowem, Claire & Evans, Bethwyn (ed.) *The routledge handbook of historical linguistics.* (p. 423-441) London: Routledge cop

McMahon, April M. S. (1994) *Understanding Language Change*. Cambridge: Cambridge Univ. Press

Sjöström, Sören (2001) *Semantisk förändring - hur ord får nya betydelser*. Lund: Studentlitteratur

Thompson, Irene (2013) *Sino-Tibetan Language Family.* Retrieved 2016-11-08 from: http://aboutworldlanguages.com/sino-tibetan-language-family

Tufvesson, Sylvia (2011) *Analogy-making in the Semai sensory World*. Senses and society Mar2011, Vol. 6 Issue 1, p86-95. 10p. Retrieved 2017-02-23 from: http://www.tandfonline.com.ludwig.lub.lu.se/doi/pdf/10.2752/174589311X12893982233876?needAccess=true

Urban, Matthias (2015) Lexical semantic change and semantic reconstruction In Bowem, Claire & Evans, Bethwyn. P (eds.) *The routledge handbook of historical linguistics* (p. 374-392) London: Routledge cop

Vejdemo, Susanne, Levisen, Carsten, van Scherpenberg, Cornelia, Gudmunsdottir Beck, Pórhalla, Naess, Åshild, Zimmermann, Martina, Stockall, Linnaea, Whelpton, Matthew (2015) *Two kinds of pink: development and difference in Germanic colour semantics.* Language Science 49 May 2015 p. 19-34. Elsevier Ltd. Retrieved 2017-01-10 from: http://ac.els-cdn.com/S0388000114000709/1-s2.0-S0388000114000709-main.pdf?_tid=de1b964c-8fab-11e6-adba-00000aab0f26&acdnat=1476188081_ce20ddb7f03ac3995a3cc025d351bc19

Verspoor, Marjolijn, de Bie-Kerékjá rto, Ágnes (2006) *Colourful bits of experience: From bluestocking to blue movie.* English Studies, 87:1, 78-98. Retrieved 2017-02-08 from: http://www.tandfonline.com.ludwig.lub.lu.se/doi/pdf/10.1080/00138380500490819?needAccess=true

Warth, Szczyglowska, Magdalena, Malgorzata (2014) *Colour and semantic change: a corpus based comparison of English green and Polish zielony.* Retrieved 2017-01-10 from: http://theses.gla.ac.uk/5690/1/2014WarthSzczyglowskaPhD.pdf

Wierzbicka, Anna (2008) *Why there are no language universals in language and thought* The Journal of the Royal Anthropological Institute, 6/1/2008, Vol. 14, Issue 2, p. 407-425. Blackwell Publishing. Retrieve 2017-02-09 from : http://eds.b.ebscohost.com.ludwig.lub.lu.se/eds/pdfviewer/pdfviewer?sid=cf372351-e78c-4e13-bf9a-12db8fc84a6a@sessionmgr4010&vid=4&hid=4213

Xing, Janet Zhiqun (2008) *Semantics and Pragmatics of Color Terms in Chinese*. Studies of Chinese Linguistics: Functional Approaches. Hong Kong: HKUP. Retrieved 2017-03-07 from: http://myweb.facstaff.wwu.edu/zxing/XingColorSemanticPragmatic08.pdf

7.2 Resources (mainly for my dataset)

Beekes, Robert with assistant of van Beek, Lucien (1969) *Etymological Dictionary of Greek.* Brill Online. Retrieved from: http://dictionaries.brillonline.com.ludwig.lub.lu.se/greek

Benson, Ken, Strandvik, Ingemar & Santos Melero, María Esperanza (1999) *Norstedts spanska ordbok: spansk-Svensk/Svensk-Spansk: (127 000 ord och fraser)* 1st edition. Stockholm: Norstedts ordbok

Birgegård, Ulla & Sharapova Marklund, Elisabeth (ed.) (2012) *Norstedts Stora rysk-svensk/svensk-rysk ordbok (144 000 ord och fraser)* 1st edition. Stockholm: Norstedts

Carl Darling Buck (1949) *A dictionary of selected synonyms in the principal Indo-European languages*. Chicago: University of Chicago Press

Cambridge University Press (2016) *English-Russian Dictionary.* Retrieved 2016-11-11 from: http://dictionary.cambridge.org/

Carling, Gerd (ed.) (2017) *Diachronic Atlas of Comparative Linguistics Online. Lund University.* Retrieved 2017-01-10 from: https://lundic.ht.lu.se

Collins (2010) *Collins Pocket Irish Dictionary in Colour*. London: Collins Publishers Limited

Derksen, Rick (2008) *Etymological Dictionary of the Baltic Inherited Lexicon.* Brill Online. Retrieved 2017-01-10 from: http://dictionaries.brillonline.com.ludwig.lub.lu.se/baltic

Derksen, Rick (1996) *Etymological Dictionary of the Slavic Inherited Lexicon.* Brill Online Retrieved 2017-01-10 from: http://dictionaries.brillonline.com.ludwig.lub.lu.se/slavic

De Vaan, Michiel (2002) *Etymological Dictionary of Latin and other Italic Languages.* Brill Online. Retrieved 2017-01-10 from: http://dictionaries.brillonline.com.ludwig.lub.lu.se/latin

Ernby, Birgitta (2008) *Norstedts etymologiska ordbok : [20.000 uppslagsord]* Stockholm : Norstedts akademiska förlag

Fisiak, Jacek (ed.) (1996) *Collins słownik polsko-angielski / pod redakcją Jacka Fisiaka = Collins Polish-English dictionary.* Warszawa : Polska Oficyna Wydawnicza

Gòmez de Silva, Guido (1985) *Elsevier's concise Spanish etymological dictionary : containing 10000 entries, 1300 word families.* Amsterdam: Elsevier

Hammarström, Harald, Forkel, Robert, Haspelmath, Martin, Bank, Sebastian (ed.) (2017) *Glottolog 2.7*. Retrieved 2016-11-02 http://glottolog.org/

Haspelmath, Martin, & Tadmor, Uri (2009). *Loanwords in the world's languages: a comparative handbook*. Berlin: De Gruyter Mouton.

Hysa, Ramazan (2008) *Albanian Standard dictionary - Albanian-English, English-Albanian.* New York: Hippocrene

J.T. Pring (1982) *The Oxford Dictionary of Modern Greek. Greek-english/english-greek* Oxford: Clarendon Press

Key, Mary & Comrie, Bernard (eds.) (2015) *The Intercontinental Dictionary Series. Leipzig: Max Planck Institute for Evolutionary Anthropology*. Retrieved 2017-01-10 from: http://ids.clld.org

Klein, Ernest (1967) *A comprehensive etymological dictionary of the English language : dealing with the origin of words and their sense development thus illustrating the history of civilization and culture*. Amsterdam : Elsevier

Lindestam, Irmgard (2008) *Norstedts stora tyska ordbok: tysk-svensk, svensk-tysk*. 1. edition. Stockholm: Norstedts akademiska förlag cop.

Kroonen, Guus (2009) *Etymological Dictionary of Proto-Germanic.* Brill Online. Retrieved 2017-01-10 from: http://dictionaries.brillonline.com.ludwig.lub.lu.se/proto-germanic

MacLennan, Malcolm (1979) *A pronouncing and etymological dictionary of the Gaelic language : Gaelic-English, English-Gaelic.* Edinburgh : Mercat Press

Matasovic, Ranko (1995) *Etymological Dictionary of Proto-Celtic.* Brill Online. Retrieved 2017-01-10 from: http://dictionaries.brillonline.com.ludwig.lub.lu.se/proto-celtic

Matisoff, James A. (ed.) (2016) *The Sino-Tibetan Etymological Dictionary and Thesaurus (STEDT)*. Retrieved 2016-11-01 from: http://stedt.berkeley.edu/

Negre, Xavier (2017) *Lexilogos - Cyrillic to Latin Alphabet*. Retrieved 2017-01-10 from: http://www.lexilogos.com/keyboard/russian_conversion.htm

Orel, Vladimir (1998) Albanian Etymological Dictionary. Leiden: Brill

Pappas, Christos (ed.) (2005) *Norstedts ordbok: nygrekisk-svensk/svensk-nygrekisk (4000 ord och fraser)* 1st edition. Stockholm: Norstedts Akademiska förlag

Partridge, Eric (1958) *Origins: a short etymological dictionary of modern English.* London : Routledge & Kegan Paul

Philips, Janet (2006) *Oxford English-Serbian student's dictionary: englesko-srpski recnik sa srprsko-engleskim indeksom.* Oxford: Oxford University Press

Vilborg, Ebbe (2009) *Norstedts svensk-latinska ordbok: [28.000 ord och fraser]* Stockholm : Norstedts akademiska förlag

Wade, Terence (1996) *Russian Etymological Dictionary.* London: Bristol Clas Press

Yvonne, Blank (ed.) (2005) *Norstedts engelska (engelsk-svensk/svensk-engelsk: 152 000 ord och fraser)* 4th edition. Stockholm: Norstedts akademiska förlag

Zendeli, Sadulla (1994) *Lexin - Svensk-albanskt lexikon - Fjalor Suedisht-shqip.* Stockholm: Skolverket

APPENDIX A

GREEN

Synchronic colexification	Semantic classification	Sino-Tibetan	Diachronic change – Indo-European	Semantic classification	Diachronic change – Sino-Tibetan	Semantic classification
unripe	Agriculture and Veg.	2	yellow	Colour	Blue	Colour
growing well	Agriculture and Veg.	1	blue	Colour	white	Colour
leafy	Agriculture and Veg.	1	Blue-green	Colour	yellow	Colour
dark	Sense Perception	1	grey	Colour	uncooked	Food and Drink
light	Sense Perception	1	bile	The Body	raw	Food and Drink
vegetable	Food and Drink	1	to gleam	Sense Perception	raw (meat)	Food and Drink
blue	Colour	6	to glimmer	Sense Perception	shoot	Basic Actions and Tech.
			to glow	Sense Perception	give birth	Basic Actions and Tech.
		Indo-European	Leek-colour	Sense Perecption	sprout	Agriculture and Veg.
yellow	Colour	1	shining	Sense Perception	grow	Agriculture and Veg.
			bright	Sense Perception	new twigs	Agriculture and Veg.
			Leek	Food and Drink	grass	Agriculture and Veg.
			to grow	Agriculture and Veg.	unripe	Agriculture and Veg.
			blossoming	Agriculture and Veg.	born	The Body
			lock	The Physical World	live	THE BODY
			gold	The Physical World	breath	The Body
					life	The Body
					fresh	Sense Perception
					fair	Sense Perception
					alive	Sense Perception
					bear	Animals

RED

Synchronic colexification	Semantic classification	Sino-Tibetan	Diachronic change – Indo-European	Semantic classification	Diachronic change – Sino-Tibetan	Semantic classification
glowing	Sense Perception	1	worm	Animals	ashamed	The Body
fiery	Sense Perception	3	maggot	Animals	blood	The Body
monkey	Animals	1	tawny	Colour	blush	The Body
blood	The Body	1	purple	Colour	rat	Animals
vermilion (cinnabar)	The Physical World	1	brown	Colour	monkey	Animals
purple	Colour	1	yellow	Colour	glowing	Sense Perception
			pink	Colour	fiery	Sense Perception
		Indo-European	dark	Sense Perception	dark	Sense Perception
ruddle	Colour	1	murderous	Sense Perception	hot	Sense Perception
pink	Colour	1	deadly	Sense Perception	black	Colour
tawny	Colour	1	Red-haired	The Body	purple	Colour
Red-haired	The Body	1	blood	The Body	soil	Agriculture and Veg.
murderous	Sense Perception	1	colour of bruises	The Body	gold	The Physical World
deadly	Sense Perception	1	wound	The Body	rust	The Physical World
vermilion (cinnabar)	The Physical World	1	earth	Agriculture and Veg.	vermilion (cinnabar)	The Physical World
			hit	Basic Actions and Tech.		
			lead	Basic Actions and Tech.		
			ginger	Food and Drink		
			berry	Food and Drink		
			Vermilion (cinnabar)	The Physical World		

YELLOW

Synchronic colexification	Semantic classification	Sino-Tibetan
bright colour of health	Sense Perception	1
dry as leaf	Agriculture and Veg.	1
white spotted	The Physical World	1
gold	The Physical World	2
		Indo-European
fallow	Agriculture and Veg.	1
faded	Sense Perception	1
pale	Sense Perception	1
colour of bruises	The Body	1
blonde	THe Body	1

Diachronic change – Indo-European	Semantic classification
bitter	Sense Perception
to gleam	Sense Perception
to glimmer	Sense Perception
to glow	Sense Perception
grateful	Sense Perception
pleasing	Sense Perception
faded	Sense Perception
dusky	Sense Perception
pale	Sense Perception
dark	Sense Perception
bile	The Body
blonde	The Body
colour of bruises	The Body
gold	The Physical World
a buoy	The Physical World
dirty	The Physical World
citrus tree	Agriculture and Veg.
fallow	Agriculture and Veg.
amber	Agriculture and Veg.
bay	Agriculture and Veg.
brown	Colour
white	Colour
black	Colour
red	Colour
grey	Colour
green	Colour
a float	Basic Actions and Tech.
honey	Food and Drink

Diachronic change – Sino-Tibetan	Semantic classification
lamp	The Physical World
white spotted	The Physical World
golden	The Physical World
gold	The Physical World
bright colour of health	The Body
bright	Sense Perception
fair	Sense Perception
shine	Sense Perception
pale	Sense Perception
light	Sense Perception
dry as leaf	Agriculture and Veg.
blue	Colour
green	Colour
white	Colour
oil	Food and Drink
butter	Food and Drink
spice	Food and Drink
turmeric	Food and Drink

BLUE

Synchronic colexification	Semantic classification	Sino-Tibetan
dark	Sense Perception	1
green	Colour	6
grey	Colour	1
		Indo-European
green	Colour	2
grey	Colour	2
black	Colour	2
lilac	Colour	1
violet	Agriculture and Veg.	1
envious	Sense Perception	1
pale	Sense Perception	1
livid	Sense Perception	1
dark	Sense Perception	3
bloodshot	The Body	1

Diachronic change – Indo-European	Semantic classification
lapis lazuli	The Physical World
black	Colour
dun	Colour
purple	Colour
yellow	Colour
turquoise	Colour
green	Colour
grey	Colour
lilac	Colour
Grey-haired	The Body
bloodshot	The Body
pigeon	Animals
leek	Food and Drink
mouldy	Food and Drink
violet	Agriculture and Veg.
heaven	Agriculture and Veg.
earthenware	Agriculture and Veg.
wind	Agriculture and Veg.
calendula oficinalis (flower)	Agriculture and Veg.
sky	Agriculture and Veg.
discoloured	Sense Perception
whole	Sense Perception
of a dull colour	Sense Perception
envious	Sense Perception
colour	Sense Perception
livid	Sense Perception
dark	Sense Perception

Diachronic change – Sino-Tibetan	Semantic classification
grey	Colour
white	Colour
green	Colour
yellow	Colour
black	Colour
fair	Sense Perception
dark	Sense Perception

BROWN

Synchronic colexification	Semantic classification	Sino-Tibetan
red	Colour	1
		Indo-European
golden	The Physical World	1
Red-yellow	Colour	1
chestnut	Agriculture and Veg.	1
dusky	Sense Perception	2
swarthy	Sense Perception	1
sandy	Sense Perception	1
dark	Sense Perception	3

Diachronic change – Indo-European	Semantic classification
honey	Food and Drink
cinnamon	Food and Drink
cut	Basic Actions and Tech
dirty	The Physical World
soot	The Physical World
golden	The Physical World
red	Colour
Red-yellow	Colour
smoky	Sense Perception
pale	Sense Perception
dusky	Sense Perception
swarthy	Sense Perception
sandy	Sense Perception
dark	Sense Perception
chestnut	Agriculture and Veg.
amber	Agriculture and Veg.
bark (of tree)	Agriculture and Veg.
beaver	Animals
elk	Animals

Diachronic change – Sino-Tibetan	Semantic classification
red	Colour

GREY

Synchronic colexification	Semantic classification	Sino-Tibetan
dull	Sense Perception	1
faded	Sense Perception	1
exhausted	Sense Perception	1
poor	The Physical World	1
black	Colour	1
		Indo-European
mouldy	Food and Drink	1
Grey-haired	The Body	2
Grey-eyed	The Body	1
tawny	Colour	1
green	Colour	1
black	Colour	2
dark	Sense Perception	1
envious	Sense Perception	1
pale	Sense Perception	1

Diachronic change – Indo-European	Semantic classification
moudly	Food and Drink
ass	The Body
Grey-haired	The Body
hoary	The Body
Grey-eyed	The Body
bright	Sense Perception
clear	Sense Perception
pale	Sense Perception
faded	Sense Perception
dark	Sense Perception
envious	Sense Perception
aged	Sense Perception
shiny	Sense Perception
tawny	Colour
brown	Colour
BLACK	Colour
blue	Colour
white	Colour
lilac	Colour
green	Colour
fallow	Agriculture and Veg.

Diachronic change – Sino-Tibetan	Semantic classification
pale	Sense Perception
dull	Sense Perception
faded	Sense Perception
exhausted	Sense Perception
poor	The Physical World
black	Colour

BLACK									
Synchronic colexification	Semantic classification	Sino-Tibetan		Diachronic change – Indo-European	Semantic classification		Diachronic change – Sino-Tibetan	Semantic classification	
dark	Sense Perception	4		dark	Sense Perception		gold	The Physical World	
gloomy	Sense Perception	1		blind	Sense Perception		ink	The Physical World	
grey	Colour	2		shine	Sense Perception		dirty	The Physical World	
white	Colour	2		lamentable	Sense Perception		grey	Colour	
monkey	Animals	1		livid	Sense Perception		red	Colour	
crow	Animals	1		dusky	Sense Perception		white	Colour	
raven	Animals	3		colourless	Sense Perception		cattle	Animals	
ink	The Physical World	3		dun	Colour		monkey	Animals	
dirty	The Physical World	1		white	Colour		crow	Animals	
				grey	Colour		raven	Animals	
		Indo-European		blue	Colour		deep	Sense Perception	
unlean	The Physical World			brown	Colour		gloomy	Sense Perception	
dirty	The Physical World			burn	Basic Actions and Tech.		dark	Sense Perception	
grey	Colour			raven	Animals				
blue	Colour			fireplace	The Physical World				
dun	Colour			first main room in roman-style house	The Physical World				
livid	Sense Perception			unlean	The Physical World				
lamentable	Sense Perception								
dark	Sense Perception								

PURPLE								
Synchronic colexification	Semantic classification	Sino-Tibetan		Diachronic change – Indo-European	Semantic classification		Diachronic change – Sino-Tibetan	Semantic classification
red	Colour	2		indigo	The Physical World		red	Colour
				purple dye	The Physical World			
		Indo-European		lilac (flower)	Agriculture and Veg.			
blue	Colour	2		violet	Agriculture and Veg.			
purple dye	The Physical World	1		blue	Colour			
blood	The Body	1		Grey	Colour			
				Grey-haired	The Body			
				blood	The Body			
				mouldy	Food and Drink			

WHITE								
Synchronic colexification	**Semantic classification**	**Sino-Tibetan**		**Diachronic change – Indo-European**	**Semantic classification**		**Diachronic change - Sino-Tibetan**	**Semantic classification**
clear	Sense Perception	1		to shine	Sense Perception		health	The Body
airy	Sense Perception	1		colourless	Sense Perception		silver	The Physical World
pale	Sense Perception	1		gleaming	Sense Perception		money	The Physical World
bright	Sense Perception	4		brilliant	Sense Perception		black	Colour
brilliant	Sense Perception	1		fair	Sense Perception		blue	Colour
money	The Physical World	1		appear	Sense Perception		green	Colour
silver	The Physical World	2		bright	Sense Perception		yellow	Colour
become	Basic Actions and Tech.	1		clear	Sense Perception		fair	Sense Perception
black	Colour	2		wan	Sense Perception		bright	Sense Perception
				airy	Sense Perception		light	Sense Perception
				pale	Sense Perception		shine	Sense Perception
		Indo-European		degree of cold	Sense Perception		brilliant	Sense Perception
gleaming	Sense Perception	1		silver	The Physical World		pale	Sense Perception
clear	Sense Perception	1		skin	The Body		burn	Basic Actions and Tech.
pale	Sense Perception	1		Fair-haired	The Body			
shining	Sense Perception	2		goat	Animals			
fair	Sense Perception	1		swan	Animals			
Fair-haired	The Body	1		show	Basic Actions and Tech.			
quick	Basic Actions and Tech.	1		quick	Basic Actions and Tech.			
agile	Basic Actions and Tech.	1		agile	Basic Actions and Tech.			
				flay	Basic Actions and Tech.			
				yellow	Colour			
				green	Colour			
				lilac	Colour			

Appendix B

Dataset direct link: https://diacl.ht.lu.se/WordListCategory/Details/11025