**Men Are From Blue, Women Are From Pink:**

**A Diachronic Study of the Differences in Association of Colours and Emotions In Literary Works of Male and Female Authors**

# Mareta Masaeva

University of Antwerp

# Mohammadreza Barati

University of Antwerp

# Shima Rahimi

University of Antwerp

# Introduction

All across languages, cultures, genders and time, colour terms have been used not only to convey the way light can take different forms, but also to express meanings. Colour, in other words, plays an important role in how we perceive the world, and at the same time, in how we express ourselves   
 to the world. We recognize objects, interpret atmospheres and facial expressions, and express how we feel, through colour. Additionally, to different people, say an artist versus a politician, a certain colour can mean something different. Even more so, combinations and placements of colours can mean different things to different people; especially across time. The symbol of a rainbow, for example, holds special meaning to people of the 21st century, while in the Medieval Ages, a rainbow was just a weather phenomenon.

In short, colour terms are used differently by different people. This depends on who these people are and in what period of history they live. We believe two of the most crucial factors affecting use of colour are gender and time. Thus, in this paper, we will attempt to identify possibly different types of meanings, specifically emotions and moods, expressed by authors using colour terms in their literary works. Differences in gender and the period these authors lived in will be taken into account; by this, we aim to find potential distinctions in how male or female authors, or authors from different centuries, use colour in their prose.

In the first section of this paper, we will go over the literature key to the subject of colour symbol-

ism in general and applied to the distinction we aim to make. Subsequently, we will describe the methods employed to perform our analyses. Then, we will share our results and discuss them. Lastly, we will summarise what we learned from our re- search.

# Literature

In psychology, many studies have been performed to find what exactly the correlation is between colours and the emotions they evoke in people. However, to find these correlations and distinc- tions, one must first study the way that humans perceive the world in general. According to Lakoff and Johnson (1999), human concepts do not just derive from the external reality of the world we live in, but are greatly influenced by our own bod- ies. Especially our brains, and our senses, play a large role in this. Different colours are, theo- retically, manifestations of how different objects reflect light. But that is not all; colours are per- ceived by our visual sensory system. The majority of scientists believe that colour only exists through our visual sensory system, and that colour is not a property of physical objects (Johnson and Wright, 2006). Lakoff and Johnson agree with the majority:

*“What could be simpler or more obvious than colors? The sky is blue. Fresh grass is green. Blood is red. The sun and moon are yellow. We see colors as inhering in things. Blue is in the sky, green in the grass, red in the blood, yellow in the sun. We see color, and yet it is false, as false as another thing we see, the moving sun rising past the edge of the stationary earth. Just as astron- omy tells us that the earth moves around the sun, not the sun around a stationary earth, so cogni- tive science tells us that colors do not exist in the external world. Given the world, our bodies and*

*brains have evolved to create color.”* (Lakoff and Johnson, 1999, p. 31)

The way we perceive colours is also impacted by the surrounding light conditions, the (in)correct workings of the three colour cones in our retinas, and so on. Colour is, thus, not a static physical property of an object. It is, in other words, depen- dent on context. Colours are, at least partly, sub- jective. Not only colours, but the world in general is experienced by humans subjectively, depend- ing on numerous factors, such as the language we speak. This hypothesis was first proposed by Sapir (1929) and Whorf and Caroll (1973), and is also referred to as the Sapir-Whorf Hypothesis of Lin- guistic Relativity. In essence, the Sapir-Whorf the- ory holds that our thoughts and experience of real- ity is shaped by our native language. Speakers of different native languages, thus, think about the world differently. When applied to the concept of colour, this would mean that colours are experienced differently depending on what prop- erties they are assigned in someone’s native lan- guage. When a native speaker of English thinks of the colour blue, they might envision a colour that differs from when a native speaker of Russian thinks of blue; the Russian language having not one, but two terms for the colour blue. Different languages divide the colour spectrum differently, which, according to the Sapir-Whorf hypothesis, means people with different native languages see and experience these colours differently; just like they would experience concepts described by words in their na tive languages differently.

To express this idea differently, colour is a vi- sual language. It can convey meanings, emotions, moods, and so on. One example of people that use colour in this way are, of course, artists. Artists of all media know which combinations of hues, vibrancy, luminance and saturation to use to con- vey a specific mood in their work. The infinite amounts of books, articles, tutorials and semi- nars about colour studies confirm this. One artist known for the intense emotions in their paintings is Edvard Munch. Munch himself described the colour red as one that evokes a screaming emo- tion: *‘I was walking along a road one evening... The sun went down – the clouds were stained red, as if with blood. [...] I painted that picture, painting the clouds like real blood. The colours screamed.’* (Stang, 1979). Vincent Van Gogh is another painter that expresses emotion through

colour, and ardently wrote about it: *‘I have tried to express the terrible passions of humanity by means of red and green’,* and *‘this combination of red ochre, of green gloomed over by gray, the black streaks surrounding the contours, produces some- thing of the sensation of anguish [...]*’. Both of these artists, it appears, linked particular colours and colour combinations to moods and emotional states (da Pos and Green-Armytage, 2007). These colours, however, are visual. We perceive these colours as they are with our eyes. An art where colours are perceived in a different way, is the art of writing. Since we are students of Digital Text Analysis, this is the art medium we will be analysing in this paper.

When researchers compared the emotions aroused in people by colour, often, studies have made the distinction between colours as the rep- resented colour or the colour term. Jonauskaite et al. (2020) studied the emotional reactions when colours were presented either as patches or as terms. For most colours, they found that the emotions associated between the patch or the term were similar. In this, lightness and saturation played a role as well; for example, focal red (named red) was typically associated with both positive and negative emotions (love and anger), while light red (named pink) was only associated with positive emotions (love, joy and pleasure). The colour purple did arouse different emotion as- sociations between patches and terms, and more importantly, carried too many different associa- tions that not one stood out. Furthermore, for the colour and the term black, differences were ob- served as well. The researchers concluded that this can be explained by the fact that black is associ- ated with more negative emotions and is experi- enced with a higher intensity when it is presented as a term.

This idea of which emotions colours are asso- ciated with, depending on whether one sees the colour in reality or reads the colour term in one’s own native language, can be linked to Lakoff’s Contemporary Theory of Metaphor (1993). In metaphors, which are of course frequently used in literature, colours often pop up: ‘green with jeal- ousy’, ‘a brown study’, ‘being blacklisted’, and so on. Being a continuation of Lakoff and John- sons ongoing research, the Contemporary The- ory of Metaphor states that metaphor is a mat- ter of thought, not of language. Our place and

view in the world is core to our use of metaphors. This is thus also language-dependent; we do not use the same metaphors across all languages (and even more so, translating metaphors to other lan- guages is incredibly difficult), and do not assign the same emotions to the same colours across lan- guages (Soriano and Valenzuela, 2009 and Big- nozzi, 2021) cultures (Ashliman, 2004), and time (Feisner and Reed, 2020). This research of emo- tional symbolism of colour belongs to the field of colour symbolism; a vast field of study that refers to the use of colour across cultures, time, gender, and so on. Many colour symbolism studies have been performed on fairy tales, folklore, fables, and other culture-specific mythologies. Ashliman (2004) writes that *‘white, the color of new snow and bleached cloth, often symbolizes purity, as does colorless transparency, as in Cinderella’s glass slippers’* and *‘in Western cultures black rep- resents mourning, a role played by white in the Orient. In Norse mythology blue symbolizes death and mourning.’* In British folklore, especially the primary colours yellow, red and blue possess spe- cial symbolism (Nassau, 1997). Colour symbol- ism of the colour black can also be found through- out the tales of the Brothers Grimm. In all but one of their tales, however, *‘the blackness is not whitened by the tale’s conclusion; and, not sur- prisingly in the view of black color symbolism, it is thus also the only one in which there is no happy ending.’* (Schmiesing, 2016) In American writer Herman Melville’s works, the colour white is prominent and is associated with heaven, heavy winds over vast oceans, reefs, deity, death and sterility (Creeger, 1960). In Melville’s Moby Dick, white is the symbol of God and death, which results in Melvilles conlusion that God is dead.

There have been little to no studies on the dif- ferences between how male and female authors use colours in their prose. A study by Mostovaia (2009) did show that, when it comes to Russian authors and Russian literature, female author used significantly more colour terms compared to male authors. However, when landscapes are de- scribed, male authors make more use of colours. When there is an emotional connotation to colours used to describe body parts, clothes and changes in the body, women use almost double the amount of colours in their description; in other words, fe- male Russian authors use more phrases like *‘red with anger’* and *‘pale out of fear’* than male Rus-

sian authors do.

How and whether English male and female au- thors differ in their use of colour, and what influ- ence time has on these differences, has yet to be investigated. Especially colour symbolism across genders is a subject that hitherto has not been stud- ied. With this paper, we would like to fill this gap and contribute to the existing literature on colour symbolism in fiction.

# Methodology

In this section, we will describe the data that was used in our analysis, as well as specific models that were employed.

## Data

The data used in this study consists of the Corpus of Late Modern English Texts (CLMET3.0) pro- vided to us by Professor Hendrik De Smet of the KULeuven. The CLMET contains about ten mil- lion words of texts, divided over three 70-year sub- periods: 1710-1780, 1780-1850 and 1850-1920.

In our selected subset, the first time period con- sists of 29 texts, the second of 35 texts, and the third of 51 texts.

All texts were written by British authors who are native speakers of English, and vary in genre. The main genre is narrative fiction, but the cor- pus also includes drama, personal letters, scientific texts, and treatise. For our analysis, we decided to only include texts of the narrative fiction and drama genre. This resulted in 201 texts, 127 from the narrative fiction genre and 74 from the drama genre.

The authors of the texts are both male and fe- male and are of varying social classes, however, inevitably, the emphasis is on men belonging to higher classes. In the unedited corpus, 158 texts were written by men, while only 45 texts were written by women. Unfortunately, we were un- able to find a corpus with a better balance between male and female writers. The corpus also includes three texts written by two authors, two of which were written by a man and woman together. Of one text, the author’s gender was unknown. These four texts were excluded for our analysis; we only included the texts with one author that was either male or female. After deleting texts from the excluded genres, our selected subset consisted of 78 male authors and 37 female authors.

As said above, we only included narrative and

drama texts written by either a male or a female author. Of the narrative fiction texts, 85 were writ- ten by men and 40 were written by women. Of the drama texts, 59 were written by men and 15 were written by women. Table 1 shows the word counts per gender per period in the selected subset. The gender imbalance was, as said, inevitable.

|  |  |  |
| --- | --- | --- |
|  | Male | Female |
| 1st period | 5 014 723 | 915 400 |
| 2nd period | 3 245 829 | 2 859 915 |
| 3rd period | 5 996 609 | 2 278 919 |

Table 1: Tokens per gender per period

Instead of choosing merely primary, secondary

and tertiary colours, we decided to use an ex- tensive list of colours in our analysis. This list consisted of the 256 HEX colours, however, we slightly adapted this list to fit our data. Colours like ‘French’, ‘Indian’ and so on were excluded, as they most often refer to nationalities instead of colours. ’Gold’, for example, was excluded as well, since many times this colour was not used as an adjective, but as a material. ’Silver’, on the other hand, was used more as a colour term than as a material, thus we did include it in the analysis. Colours that could not be put into one clear category of a primary, secondary or tertiary colour, were excluded as well. This left us with 34 colours. In the Table 2, you can find all the colours along with the colour family they were mapped to in our analysis.

|  |  |
| --- | --- |
| Red | Red, umber, scarlet, crimson, ruby,  maroon |
| Orange | Orange, amber |
| Yellow | Yellow, ecru, flaxen |
| Green | Green, emerald |
| Blue | Blue, azure, turquoise |
| Purple | Purple, violet, indigo, mauve |
| Pink | Pink, carnation, magenta, cerise |
| White | White, creamy |
| Black | Black |
| Grey | Grey, gray, silver |
| Brown | Brown, tawny, khaki, russet |

Table 2: Colours with their colour family

## Analysis

After thoroughly analyzing the CLMET3.0 cor- pus, our final version of the data to use in the anal-

ysis consisted of a raw text column, the Part-of- Speech tag of the tokens and the lemmas for each word. Eventually, only the POS-tag and tokenized texts were used, as they returned better results.

As said, most tokens in our data were from male authors. To balance this out, we decided to shave off parts of texts from male authors. In this way, the number of tokens for male authors comes closer to the number of tokens for female authors. This shaving off was done from the beginning and end of each file from a male author. This eventually resulted in a more balanced number of tokens for each gender in each time period: around 914 500 tokens per gender in the first time period, 2 857 400 tokens per gender in the second time pe riod and 2 276 500 tokens per gender in the third time period.

As predictors, we used three models: the Distilbert-base-uncased-emotion-, XLM-EMO- and Text2Emotion classifiers. Eventually, we would find out by looking at annotated data, which of these models performed the best. The data was converted to a nested dictionary, with an identifier as key and as values the file itself, the match (the actual colour word found), the colour family (the colour mapped to a primary, secondary or tertiary colour). Along with these is some metadata, like the title of the work, gender of the author, time period, birth year of the author, and so on.

We decided to pass multiple versions of the matched strings, more specifically with five differ- ent window sizes, through the emotion classifiers. Deciding which window size would be best was difficult, as there is no real literature on the best size for such a task and the different sizes returned different results. Because of this, we figured in- cluding multiple window sizes must be a reason- able approach.

To summarize, each classifier thus predicted an emotion from five different lengths of strings. This means, for each matched colour word, 15 emo- tions were predicted, five by each classifier. These results were stored in three JSON files, one for each time period. Eventually, we compared our results to annotated data to decide which classi- fier returned the best results. This turned out to be the BERT model.

# Results

# 4. 1 Best predictor

# To determine which of the three predictors would return the best results, we decided to feed them annotated data. This literary data, consisting of 1205 sentences, is annotated for five emotions: angry/disgusted, fearful, happy, sad and surprised.

# The BERT model includes six emotions; it includes love in addition to the above emotions that the data was annotated with. Love was included in the happy category of the original dataset. The BERT model returned an accuracy of 66.3%. The confusion matrix of the BERT model can be found below.

# Chart Description automatically generated

Figure 1: Confusion matrix BERT

# The second predictor was Text2Emotion, which predicts emotions that totally match the ones from the dataset. This model performed the least accurately, and seemed to be overfit on the surprise emotion. The accuracy of this model was 38.8%, and the confusion matrix can be found below.

# Graphical user interface, chart, treemap chart Description automatically generated

Figure 2: Confusion matrix Text2Emotion

The last predictor, XLM-EMO, predicts only four emotions. It ignores surprise. This means that it naturally returns the highest performance of the three, as it misses one label. The accuracy of this model was 71.5%, and the confusion matrix can be found below.

Graphical user interface, chart

Description automatically generated

Figure 3: Confusion matrix XLM-EMO

We decided to use the BERT model for our final classification, because, even though XLM-EMO has a higher accuracy, BERT includes two more emotions and returned quite a high accuracy, even after merging the love with the happy emotion. Thus, the results of our analysis will have come from this BERT model.

**4.2 Results from BERT model**

As was explained above, we decided to run our data through the model with multiple window sizes: 5, 10, 15, 20 and 25 tokens before and after the matched colour term. The results we will report on here are from the data with a window size of 20 tokens on either side of the matched word or 40 in total, however, results from data with other window sizes can be found in our code. We choose 20 here because we believe it provides the classifiers with just enough context. A window size of 25 (50 in total) might be too high, thus we settle for 20 here. Also, we would like to add, charts that are not shared in this paper can be found in the code. That includes charts with either raw or normalized counts, overall and for each time period separately.

**4.2.1 Gender and colour**

When looking at how male and female authors use colour, disregarding the time periods they lived in, we see that male authors make more use of colours than female authors. There is, of course, an imbalance between the amount of male and female authors, which must be taken into account. The normalized use of colours by male and female authors in all time periods can be found in Figure 4 below.

Chart, bar chart

Description automatically generated

Figure 4 Gender and colour use (overall, raw)

As the figure shows, the colour used the most by both genders, is red; red is used 158 times by male authors and 108 times by female authors. For most colours, the difference between male and female use is quite large, except for the ones that are used less frequently: orange (a difference of 1), pink (a difference of 2) and purple (a difference of 8). When we look at the normalized counts from the second time period, where the amount of texts by male and female authors is the same number (18), we see the trend of male authors using more colours continuing (Figure 5).

Chart, bar chart

Description automatically generated

Figure 5: Gender and colour use (2nd period, normalized)

As Figure 5 shows, male authors tend to use more colour words in their writing compared to female authors. This difference is slightly smaller for red, white and yellow, even smaller for black, and inexistent for orange. Only the colour pink is used more by female authors, by large. When we look at the raw counts, we see that pink is used 10 times by female authors and 7 times by male authors. This is the only instance, across all three time periods, where female authors use a colour more than male authors.

Comparing the normalized counts of the three time periods, a couple of remarks must be made. In the first time period, three colours were not used by the female authors: orange, pink and purple. Because of this, the male group receives a 100% normalized count for these colours, but they were rarely used. Orange and purple were used twice by male authors in the first period, pink was used once. The colours used most frequently in this time period by male authors, looking at the raw counts, were black, red and grey as the top three, and white and blue following. For female authors, the top colours were red and white, followed by black, grey and brown.

Chart, bar chart

Description automatically generated

Figure 6: Gender and colour use (1st period, normalized)

In the third time period, red and grey were the colours used the most by both genders. For the male group, black and white tie at number three. For female authors, blue was the third most used colour. In this time period, female authors used the colour orange more than male authors did, and they used purple in the same amount. All of these raw counts can be found in the notebook.

Chart, bar chart

Description automatically generated

Figure 7: Gender and colour use (3rd period, normalized)

**4.2.2 Colour and emotion**

Here, we will show the results of the correlations between colours and emotions, disregarding gender of the author.

In Figure 8, you can see the overall correlations. The colour most associated with joy, which in every time period receives very high correlation scores, is green, closely followed by orange. Another emotion that is strongly associated with certain colours is anger, and got the highest score overall with the colour brown, followed by black. The colour mostly associated with fear is black as well. Love is strongly associated with pink, orange and purple, sadness with yellow, and surprise with yellow.

Chart, bar chart

Description automatically generated

Figure 8: Colour and emotion (overall)

In the first time period, some colours are totally disassociated with some emotions, as you can see in Figure 9. Pink and purple are only associated with joy, and orange is only associated with anger. These were, however, very rarely used colours. Black, one of the top colours, is mostly associated with joy, along with green and white. Red, another top colour, is mostly associated with anger, along with yellow.

Chart, histogram

Description automatically generated

Figure 9: Colour and emotion (1st period)

In the second time period, joy is mostly associated with green and orange. Anger has a high correlation with brown and black, and love with pink. Fear has a high score with white and black, sadness with purple and yellow and surprise with yellow. All of this can be seen in Figure 10.

Chart, bar chart

Description automatically generated

Figure 10: Colour and emotion (2nd period)

The last time period shows trends similar to those of the second time period, which can be seen in Figure

11. Some differences include the high correlation of anger with purple, and fear with yellow and purple.

Chart, bar chart

Description automatically generated

Figure 11: Colour and emotion (3rd period)

**4.2.3 Colour, gender and emotion**

In this last part of Section 4, we will share the results of our analysis of the correlation between colours and emotions across the two genders. As this entails many different results, we will not be able to share all of them here. Charts and numbers can all be found in the code, and here, we will share the results we believe to be the most important.

As said above, the colours used most overall were red, grey, black, white and blue. Red, the most frequently used colour, was mostly associated with anger and joy. For female authors, red was used once more with an anger emotion than with a joy emotion, while male authors used it eight times more with a joy emotion. Female authors associated red more with fear and love than male authors did. In Figure 12 and 13, you can see more clearly which emotions male and female authors associated red with.

Chart, pie chart

Description automatically generated

Figure 12: Red associations (male, overall)

Chart, pie chart

Description automatically generated

Figure 13: Red associations (female, overall)

Black, the second most frequently used colour, was mostly associated with anger. Male authors associated black with joy in the same amount as they associated it with anger. They also associated it with fear. Female authors associated black with joy and fear as well, but less frequently. This difference can be seen in Figure 14 and 15.

Chart, pie chart

Description automatically generated

Figure 14: Black associations (male, overall)

Chart, pie chart

Description automatically generated

Figure 15: Black associations (female, overall)

All association counts, for all colours and emotions, can be found in Table 3 below.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Anger | Fear | Joy | Love | Sadness | Surprise |
| Black | M | 39 | 20 | 39 | 2 | 8 | NaN |
| F | 21 | 7 | 18 | 1 | 6 | NaN |
| Blue | M | 22 | 11 | 41 | 5 | 6 | 2 |
| F | 26 | 1 | 25 | 4 | 7 | NaN |
| Brown | M | 35 | 6 | 26 | 1 | 8 | 1 |
| F | 16 | 4 | 19 | 3 | 4 | NaN |
| Green | M | 20 | 9 | 51 | 2 | 6 | 1 |
| F | 6 | 5 | 30 | 3 | 1 | NaN |
| Grey | M | 21 | 15 | 44 | 7 | 13 | 3 |
| F | 37 | 5 | 31 | 2 | 5 | NaN |
| Orange | M | 3 | NaN | 11 | 4 | 1 | NaN |
| F | 2 | 4 | 11 | NaN | 1 | NaN |
| Pink | M | 9 | 3 | 10 | 3 | 3 | 1 |
| F | 7 | NaN | 15 | 5 | NaN | NaN |
| Purple | M | 12 | 4 | 15 | 5 | 2 | NaN |
| F | 10 | 6 | 10 | 2 | 2 | NaN |
| Red | M | 54 | 13 | 62 | 8 | 16 | 5 |
| F | 38 | 16 | 37 | 6 | 9 | 2 |
| White | M | 29 | 18 | 45 | 4 | 4 | 2 |
| F | 15 | 5 | 27 | 6 | 3 | NaN |
| Yellow | M | 20 | 10 | 25 | 2 | 5 | 1 |
| F | 13 | 3 | 18 | NaN | 8 | 2 |

Table 4: Colours, emotions and genders (overall)

For the first time period, some emotions are correlated with very few emotions. Only anger and joy are associated with more than half of the colour families. Male authors use mostly red to convey anger, along with black and grey. Female authors use these colours as well, but less frequently for anger. Female authors in this time period did use grey, black and white frequently to convey joy. Male authors mostly used black and white for this. In this period, surprise was conveyed once by the male group with the colour grey, and sadness was conveyed with black, green and yellow by the male group, and with red and yellow by the female group. Love was conveyed with blue, grey and red by males, and with white by females. All of these numbers can be found in Table 4 below.

Table 3: Colours, emotions and genders (1st period)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Anger | Fear | Joy | Love | Sadness | Surprise |
| Black | M | 7 | 3 | 10 | NaN | 1 | NaN |
| F | 2 | NaN | 6 | NaN | NaN | NaN |
| Blue | M | 5 | 1 | 5 | 1 | NaN | NaN |
| F | 1 | NaN | 2 | NaN | NaN | NaN |
| Brown | M | 3 | 2 | 5 | NaN | NaN | NaN |
| F | 3 | NaN | 2 | NaN | NaN | NaN |
| Green | M | NaN | NaN | 6 | NaN | 3 | NaN |
| F | 1 | NaN | 3 | NaN | NaN | NaN |
| Grey | M | 6 | NaN | 8 | 1 | NaN | 1 |
| F | 2 | NaN | 5 | NaN | NaN | NaN |
| Orange | M | 2 | NaN | NaN | NaN | NaN | NaN |
| Pink | M | NaN | NaN | 1 | NaN | NaN | NaN |
| Purple | M | NaN | NaN | 2 | NaN | NaN | NaN |
| Red | M | 13 | 2 | 5 | 1 | NaN | NaN |
| F | 3 | 3 | 2 | NaN | 1 | NaN |
| White | M | 3 | 2 | 8 | NaN | NaN | NaN |
| F | 2 | NaN | 5 | 2 | NaN | NaN |
| Yellow | M | 3 | NaN | 1 | NaN | 1 | NaN |
| F | 3 | NaN | NaN | NaN | 1 | NaN |

In the second time period, we can see that anger is mostly conveyed with the colour red, by both genders. Joy is mostly expressed through colours like green and red, but more frequently by male authors. The colour used mostly to express joy by female authors, is red. For love, red is used as well by both genders, along with white. Male authors use green to express love as well. All of these trends can be seen in Figure 16, 17, 18 and 19.

Chart, pie chart

Description automatically generated

Figure 16: Joy conveyed by male authors (2nd period)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Anger | Fear | Joy | Love | Sadness | Surprise |
| Black | M | 21 | 12 | 21 | 2 | 5 | NaN |
| F | 9 | 2 | 6 | NaN | 3 | NaN |
| Blue | M | 12 | 6 | 20 | 2 | 3 | 2 |
| F | 7 | NaN | 15 | 2 | 3 | NaN |
| Brown | M | 21 | 4 | 13 | 1 | 4 | 1 |
| F | 6 | 2 | 13 | 1 | 4 | NaN |
| Green | M | 15 | 6 | 25 | 1 | 3 | NaN |
| F | 2 | 4 | 16 | 1 | NaN | NaN |
| Grey | M | 23 | 10 | 24 | 2 | 6 | 1 |
| F | 9 | 1 | 15 | 1 | 4 | NaN |
| Orange | M | NaN | NaN | 8 | 2 | 1 | NaN |
| F | 2 | 2 | 7 | NaN | 1 | NaN |
| Pink | M | 7 | 3 | 7 | 2 | 1 | 1 |
| F | 6 | Nan | 9 | 2 | NaN | NaN |
| Purple | M | 9 | 3 | 6 | 3 | NaN | NaN |
| F | 7 | 5 | 8 | 1 | NaN | NaN |
| Red | M | 26 | 9 | 38 | 4 | 8 | 3 |
| F | 19 | 7 | 23 | 2 | 5 | 2 |
| White | M | 20 | 9 | 26 | 2 | 3 | 1 |
| F | 5 | 2 | 15 | 1 | 2 | NaN |
| Yellow | M | 11 | 9 | 15 | 1 | 2 | NaN |
| F | 6 | 3 | 9 | NaN | 3 | 2 |

Chart, pie chart

Description automatically generated

Figure 17: Joy conveyed by female authors (2nd period)

Chart, pie chart

Description automatically generated

Figure 18: Love conveyed by male authors (2nd period)

Table 5: Colours, emotions and genders (3rd period)

Chart, pie chart

Description automatically generated

Figure 19: Love conveyed by female authors (2nd period)

All of the association counts for the second period can be found in Table 5 below.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Anger | Fear | Joy | Love | Sadness | Surprise |
| Black | M | 11 | 5 | 8 | NaN | 2 | NaN |
| F | 10 | 5 | 6 | 1 | 3 | NaN |
| Blue | M | 5 | 4 | 16 | 2 | 3 | NaN |
| F | 8 | 1 | 8 | 2 | 4 | NaN |
| Brown | M | 11 | NaN | 8 | NaN | 4 | NaN |
| F | 7 | 2 | 4 | 2 | NaN | NaN |
| Green | M | 5 | 3 | 20 | 1 | NaN | 1 |
| F | 3 | 1 | 11 | 2 | 1 | NaN |
| Grey | M | 8 | 5 | 13 | 4 | 7 | 1 |
| F | 10 | 4 | 11 | 1 | 1 | NaN |
| Orange | M | 1 | NaN | 3 | 2 | NaN | NaN |
| F | NaN | 2 | 4 | NaN | NaN | NaN |
| Pink | M | 2 | NaN | 2 | 1 | 2 | NaN |
| F | 1 | NaN | 6 | 3 | NaN | NaN |
| Purple | M | 3 | 1 | 7 | 2 | 2 | NaN |
| F | 3 | 1 | 2 | 1 | 2 | NaN |
| Red | M | 15 | 2 | 19 | 3 | 8 | 2 |
| F | 16 | 6 | 12 | 4 | 3 | NaN |
| White | M | 6 | 7 | 11 | 2 | 1 | 1 |
| F | 8 | 3 | 7 | 3 | 1 | NaN |
| Yellow | M | 6 | 1 | 9 | 1 | 2 | 1 |
| F | 4 | NaN | 9 | NaN | 4 | NaN |

Table 6: Colours, emotions and genders (2nd period)

All association counts for the third time period can be found in Table 6.

In this table, we can see that for the last time period, the emotion with the highest colour correlation is joy, as this emotion receives the highest results. The colour most associated with joy by both genders is red. Red is also highly correlated with anger. Love is mostly conveyed with the colours pink, blue and red by female authors, and with red and purple by male authors, as you can see in Figure 20 and 21 below.

Chart, pie chart

Description automatically generated

Figure 20: Love conveyed by male authors (3rd period)

Chart, pie chart

Description automatically generated

Figure 21: Love conveyed by female authors (3rd period)

For fear, some differences between genders can be observed in Figure 22 and 23. Female authors tend to use red and purple to convey fear, while male authors tend to use black and grey the most.

Chart, pie chart

Description automatically generated

Figure 22: Fear conveyed by male authors (3rd period)

Chart, pie chart

Description automatically generated

Figure 23: Fear conveyed by female authors (3rd period)

1. **Discussion**

In this section, we will discuss the results observed and shared in the previous section.

In the first time period, for the female authors, some emotions were only associated with one or two colours; for example, sadness was associated with only red and yellow, surprise with only grey, and love with only white. As this first period had only five female authors, this could be very well explained by the lack of data for this earlier period of time. Joy seems to be expressed through similar colours when comparing genders, however, when we compare the first to the second time period, we see that joy is correlated to more colours further on in time; purple, pink and orange are not used by female authors to express joy between the 1710s and 1780s, but from 1780 to 1850, they are used. Male authors did already use pink and purple to convey joy, and start using orange as well in the 1780s. From then on, not much changes for joy.

Together with joy, anger was one of the most prevalent emotions conveyed with certain colours in our dataset. In the first time period, male and female authors use black and red to convey this emotion. Brown and yellow also seem to be important emotions to convey anger for the female authors. Throughout time, black and red continue to be important emotions for the anger emotion, along with grey slightly less frequently. Other, less important colours, stay present for this emotion, but are less dominant than black and red: blue, brown and white.

In the second time period, where there is a perfect balance between male and female authors, looking at the most frequent colours, red, grey and black, is quite interesting. Especially for fear, we can see that red is very prevalent in texts by female authors compared to texts by male authors. In texts by male authors, white is used the most to convey fear, along with black and grey. This use of white for fear by males goes down drastically in the third time period, and the use of red for fear increases. The use of grey decreases, especially for the female texts, and is replaced by the colour purple. This might be because of how purple became more associated with royalty, which might have been something to fear for people of lower classes.

As we had expected, the colour red was used very frequently by both genders to express love, along with the colour pink. Other colours used frequently for this emotion are grey, used frequently by males in the second time period, and blue, used frequently by females in the third time period. Grey especially is surprising for this emotion, as it might be seen as a very dull colour. However, as grey might also be seen as a cool and mysterious colour, it might be suited to express a kind of mysterious, secret love, or a more negative love. Blue is a very positive colour, as it is also used to express joy, and thus could easily express a more joyful type of love.

Overall, making straightforward comparisons and explaining them is difficult, as there are many variables that have come into play in our analysis: gender, time, emotion and colour. This leads to many possible correlations and explanations that could be made for them. Because of this, we believe that further research into colour symbolism in literature, when comparing gender of authors across time, should probably find a way of making more clear comparisons and correlations. When many variables play a role, one should, for example, account for only primary colours, or only the very frequent colours. In our analysis, some colours were used infrequently, which might have clouded our analysis. This should be kept in mind in future research.

# Conclusion

# From our research, different conclusions can be drawn. As explained in the last part of the discussion, we believe that taking into account 11 different colour families, some of which were very infrequently used, might have clouded our analysis. Making comparisons between genders and time periods with this many different colours is difficult, and leads to doubtful conclusions. We do, however, believe that this research teaches us a lot about how female and male authors differ in their use of colour and which colours they tie to which emotions, and what effect time has on this. We hope that, in the future, more research will be performed on this topic that will lead us to more firm solutions.

1. **References**

Ashliman, D. L. (2004). Folk and Fairy tales: A handbook. Greenwood Press.

Bianchi, F., Nozza, D., amp; Hovy, D. (2022). XLM-emo: Multilingual emotion prediction in social media text. Proceedings of the 12th Workshop on Computational Approaches to Sub- jectivity, Sentiment amp; Social Media Analysis.

Bignozzi, C. (2021). The Sapir-Whorf’s Hypothesis: a comparison on colour definition and colour perception in English and Russian speakers. (dissertation).

Creeger, G. R. (1960). The Symbolism of Whiteness in Melville’s Prose Fiction. Jahrbuch Fu¨r Amerikastudien.

da Pos, O. and Green-Armytage, P. (2007). Facial Expressions, Colours and Basic Emotions. Colours: Design Creativity, 1(1), 1–20.

De Smet, H. (n.d.). The Corpus of Late Modern English Texts. KU Leu- ven. Retrieved June 17, 2022, from [https://perswww](http://www.kuleuven.be/).kuleuv[en.be/](http://www.kuleuven.be/) u0044428/clmet.htm

Feisner, E. A. and Reed, R. (2020). Color studies. Fairchild Books.

GrammarVocab Color Shades. (n.d.). Re- trieved from https://grammarvocab.com/wp- content/uploads/2021/09/Slide1-9- 1024x1536.png.webp

Johnson, K. and Wright, W. (2006). Colors as properties of the Special Sciences. Erkenntnis, 64(2), 139–168.

Jonauskaite, D., Abu-Akel, A., Dael, N., Oberfeld, D., Abdel-Khalek, A. M., Al-Rasheed,

1. S., Antonietti, J.-P., Bogushevskaya, V., Chamseddine, A., Chkonia, E., Corona, V., Fonseca-Pedrero, E., Griber, Y. A., Grimshaw, G., Hasan, A. A., Havelka, J., Hirnstein, M., Karlsson,
2. S., Laurent, E., . . . Mohr, C. (2020). Universal patterns in color-emotion associations are further shaped by linguistic and geographic proximity. Psychological Science, 31(10), 1245–1260.

Lakoff, G. (1993). The contemporary theory of metaphor. Metaphor and Thought, 202–251.

Lakoff, G. and Johnson, M. (1999). Philosophy in the flesh: The cognitive unconscious and the embodied mind: How the embodied mind creates philosophy. Basic Books.

Mostovaia, A. (2009). Color words in literary Russian: Connotations and gender differences in use. JLSE, 38(1), 1–37.

Names and hex codes of the

256 RGB (web safe) Colours. [https://www](http://www.umsiko.co.za/links/RGB-).umsik[o.co.za/links/RGB-](http://www.umsiko.co.za/links/RGB-) ColourNamesHex.pdf. (n.d.).

Nassau, K. (1997). Experimenting with color.

F. Watts.

Sapir, E. (1929). The status of linguistics as a science. Language, 5(4), 207.

Schmiesing. (2016). Blackness in the Grimms’ fairy tales. Marvels and Tales, 30(2), 210.

Soriano, C. and Valenzuela, J. (2009). Emotion and colour across languages: Implicit associations in Spanish colour terms. Social Science Informa- tion, 48(3), 421–445.

Stang, G. (1979). Edvard Munch: The man and the artist. Fraser.

Whorf, B. L. and Carroll, J. B. (1973). Lan- guage, thought, and reality selected writings.

M.I.T. Pr.