UPPSALA
UNIVERSITET

Word Frequency as a Predictor of Word Intensity

Rebeca Padilla López

Uppsala University
Department of Linguistics and Philology
Master's Programme in Language Technology
Master's Thesis in Language Technology

June 12, 2017

Supervisors:
Joakim Nivre, Uppsala University
Magnus Sahlgren, SICS

Abstract

In this thesis we explore the intensity of adjectives and how it can be predicted by different word features. We investigate how to accurately determine intensity between synonymous adjectives. For this, we look at features such as word frequency, number of senses and syllable length. Our study is inspired by life satisfaction and happiness surveys and the possibility that differences in intensity in the translation of the adjectives used for the questionnaires could explain the high degree of satisfaction that some countries show. We base our hypothesis on the theories of grammaticalization and semantic bleaching and the discoveries made by other researches about the relations between these word features and word intensity. We focus on studying Danish, English and French. Our study points to a statistically significant negative correlation between word frequency and word intensity.

Contents

List of Tables **4**

Acknowledgements **5**

1 Introduction **6**

2 Background **9**

3 Methodology **12**

4 Experiments **14**
4.1 Total Frequency of English adjectives 14
4.1.1 Happiness and Satisfaction Adjectives 14
4.1.2 Quality Adjectives 16
4.2 Total Frequency of Danish adjectives 18
4.2.1 Happiness and Satisfaction adjectives 18
4.2.2 Quality Adjectives 20
4.3 Total frequency of French adjectives 21
4.3.1 Happiness and satisfaction adjectives 22
4.3.2 Kindness and meanness adjectives 24

5 Comparison between languages **26**
5.1 English, Danish and French 26
5.2 English and Danish 28
5.3 Multilingual study 29

6 Discussion **31**

7 Conclusion **33**

Bibliography **34**

List of Tables

1.1 Results from Overall satisfaction survey 7

2.1 Results from the Sharma and Bhattacharyya (2013) experiments . 10

3.1 Words from the European Values Study (2008) questionnaire . . . 12

4.1 Positive English happiness adjectives 15
4.2 Negative English happiness adjectives 15
4.3 Correlation coefficients for English happiness adjectives 15
4.4 Positive English quality adjectives 17
4.5 Negative English quality adjectives 17
4.6 Correlation coefficients for English quality adjectives 17
4.7 Positive Danish happiness adjectives 18
4.8 Negative Danish happiness adjectives 19
4.9 Correlation coefficients for Danish happiness adjectives 19
4.10 Positive Danish quality adjectives 20
4.11 Negative Danish quality adjectives 21
4.12 Correlation coefficients for Danish quality adjectives 21
4.13 Positive French happiness adjectives 22
4.14 Negative French happiness adjectives 23
4.15 Correlation coefficients for French happiness adjectives 23
4.16 Positive French kindness adjectives 24
4.17 Negative French kindness adjectives 24
4.18 Correlation coefficients for French kindness and meanness adjectives 25

5.1 Correlation comparison between English, Danish and French . . . 26
5.2 Comparison between all positive happiness adjectives 27
5.3 Comparison between all negative happiness adjectives 27
5.4 Coefficient comparison for all happiness and satisfaction adjectives 27
5.7 Coefficient comparison for all quality adjectives 28
5.5 Comparison between all positive quality adjectives 28
5.6 Comparison between all negative quality adjectives 28
5.8 Words used in the European Values Study questionnaire and their frequency 29

6.1 Correlation coefficient comparison for complete sets of adjectives in all their languages 31

Acknowledgements

First of all I want to thank my supervisors, Joakim Nivre, for guiding and helping me in the writing of this thesis and for being an excellent professor, and Magnus Sahlgren, for giving me the opportunity to be part of something bigger than just a Master's thesis.

I would also like to thank my parents for supporting me in so many ways during all these years, and my grandparents for being my biggest fans.

Finally, I want to thank Oliver for his love and support and let him know that he will always have mine.

1 Introduction

When looking at life satisfaction and happiness surveys, it is not uncommon to see that certain countries, such as Denmark, are always at the top of the list. In the first World Map of Happiness ever created, Denmark came in first, and other Scandinavian countries such as Finland and Sweden appeared in 6th and 7th place respectively.[1] In 2017, the World Happiness Report positioned Norway as the happiest country in the world, with Denmark as a close second (Helliwell et al., 2017).

It is not strange to see Denmark so high up these lists, but it can be surprising to find that it is always there, year after year. It could be argued that differences in life quality between Denmark and other similar rich welfare states are not significant enough to explain these results, and that the explanation for these could be in the questionnaire itself and the translation of its questions.

In fact, the influence of language in these surveys has been researched. Lolle and Andersen (2016) investigated whether the translation of survey questions from an English master questionnaire to a Danish one had any effect on the response pattern. In their experiments, randomly chosen groups of Danish university students responded to questionnaires in English and Danish. They found that the answers were significantly different depending on the language, and that similar students that were expected to give the same responses actually gave different answers depending on the language used in the questionnaire.

[1] http://www.le.ac.uk/ebulletin-archive/ebulletin/news/press-releases/2000-2009/2006/07/nparticle.2006-07-28.html

[2] https://www.youtube.com/watch?v=Se2gfFKp1Iw

Figure 1.1: Ten happiest countries in the world, according to the World Happiness Report[2]

	Very satisfied	Fairly satisfied	Not very satisfied	Not at all satisfied
English	41%	51%	8%	0%
Danish	52%	43%	5%	0%

Table 1.1: Results obtained in the Overall life satisfaction survey, Eurobarometer variant (Lolle and Andersen, 2016). While only 41% of the students who took the questionnaire in English answered "Very satisfied", 52% of the students who took it in Danish did. The most common English response (51%) was "Fairly satisfied."

They also concluded that the overall pattern could also be found in the real world, and they confirmed that the term "happy" and its translation "lykkelig" were not equivalent. The English "happy" is weaker in intensity than "lykkelig." Moreover, as seen on 1, the answers given on life satisfaction were higher than the English ones (Lolle and Andersen, 2016).

It is important to mention that this difference in intensity between languages might not be the only factor affecting the results of the surveys. It is possible that, due to a social desirability bias, participants may tend to give more socially acceptable answers, and framing effects have been shown to influence survey results too (Garcia et al., 2012). Bryan et al. (2011) showed in their study that there was a significant change in the answers of a voter turnout survey when the word "voting" was changed to "being a voter" (Garcia et al., 2012).

The work by Lolle and Andersen (2016) and the aforementioned life satisfaction survey results serve as an inspiration for our thesis. We argue that word intensity is a markedness phenomenon and that more intensity is indicated by more marked terminology than neutral intensity. We maintain that this markedness can be indicated by features such as frequency of occurrence, number of senses or word length. Our main hypothesis is that there is a strong negative correlation between the frequency of a word and its intensity, and that this correlation allows us to use word frequency as a predictor of word intensity. Additionally, we also argue that there is a negative correlation between intensity and number of senses and a positive correlation between intensity and word length or number of syllables. We also expect these features to be useful in the prediction of word intensity. For our study, we will focus on adjectives that indicate happiness or life satisfaction, quality, and kindness and meanness and look at the correlation between their intensity and their frequency, length and number of senses. Our hypothesis is based on semantic bleaching, the process by which, while becoming more and more functional, a morpheme loses meaning. Some other terms used to describe this process are semantic weakening, semantic fading, desemanticization and generalization or weakening of semantic content (Von Fintel, 1995).

We will analyse our results and try to determine if it is possible to obtain an accurate, fine-grained intensity classification with just this information. Annotating words with different degrees of intensity offers more information than a simple categorical classification of "positive", "neutral" and "negative". In the case of the life satisfaction and happiness questionnaires, knowing the specific intensity of the words in the different languages used in the questions

would help normalise the results of the survey.

In most emotion and sentiment work this kind of categorical annotation or coarse rating scales are the ones used to obtain annotations. This is not surprising, since fine-grained annotation usually involves manual annotation, which is costly and hard to achieve. Not only do the annotators have to make a great mental effort, but it is also very hard to ensure that the intensity scores are consistent. For example, while one annotator might assign a score of 3.2 to a word, another annotator might assign a lower score of 2.6 to this same word. Inconsistencies are also found within the scores given by the same annotator. It is common to find that the same annotator has assigned a certain score to a word at certain point in time and a different score at some other point. Last but not least, annotators can also be biased towards different parts of the scale, which is called the scale region bias. Despite this, a key question is whether humans are actually able to discriminate across very slight intensity differences, or if we can only distinguish a small number of coarse intensity levels (Mohammad, 2017).

All of these issues could be minimised if we were able to determine a word's intensity just by analysing the word's features, thus eliminating most if not all need for human intervention and manual annotation.

We must note that the intensity of a word can change depending on the context in which it is being used and, if it has multiple senses, on which sense is being referred to. It would be ideal to give a word a specific intensity score for each of its senses, however, in this thesis we are trying to establish a general score. This would be a helpful guideline for fields such as sentiment analysis, and it could serve as a baseline for future work that aims for a more specific classification.

2 Background

George K. Zipf was one of the first linguists to study word frequency. Zipf's most famous law states that the frequency of any word is inversely proportional to its rank in the frequency table. Zipf also states that longer words are more rarely used than short words, as a result of a principle of least effort. (Zipf, 1949). Piantadosi et al. (2017) improve on this and show that word length increases with information content, and that the latter is a better predictor of length than frequency is (Garcia et al., 2012).

The concept of grammaticalization is even older than Zipf's law. It was first defined by Meillet (1912) as the "passing of an autonomous word to the role of grammatical element" or the "attribution of grammatical character to a previously autonomous word." On the other hand, Moder (2007) argues for a different definition, "one that recognizes the crucial role of repetition in grammaticization and characterizes it as the process by which a frequently used sequence of words or morphemes becomes automated as a single processing unit."

Moder (2007) argues that frequent repetition plays an important role in changes that take place in grammaticization, such as frequency of use leading to weakening of semantic force by habituation She argues this is the mechanism behind semantic bleaching.

On the other hand, Von Fintel (1995) argues that functional meanings do not always have to be frequent, and there are some functional morphemes, such as "unless" or "ought", which are infrequent. He also claims that even though there is a link between the functional status of a word and its frequency, this does not act as a good predictor, and that one can only say that functional items are frequent because there are few of them and their presence is mandatory (Von Fintel, 1995).

In our case, we will focus on examples that show semantic bleaching without the structural change that comes with full grammaticalization, and we will not be working with completely functional words. We will test whether our hypothesis holds for adjectives related to satisfaction and happiness adjectives in Danish, English and French.

Sharma and Bhattacharyya (2013) mention that, while there is no shortage of research on grammaticalization, sentiment analysis and opinion words, there is a lack of research on intensity variation among synonyms.

In their paper, Sharma and Bhattacharyya (2013) do explore intensity among synonyms, and they conclude that relative frequency for a word in a specific sense can predict the intensity of the word in that sense. A word with a high relative frequency will be very intense for that particular sense, while a a word with low relative frequency will not be intense when used in that

Synonyms (w_1, w_2)	**Sense definition**	**Total count (w_1)**	**Total count (w_2)**	**Sense count (w_1)**	**Sense count (w_2)**
Awful, Painful	Exceptionally bad or displeasing	10	14	1	12
Proficient, Good	Having or showing knowledge and skill and attitude	4	263	3	2
लाभदायक, उपयोगी, (Beneficial, Useful)	Giving an advantage	22	35	22	15
उत्तम, अच्छा (Exquisite, Substantial)	Having or marked by unusual and impressive intelligence	181	270	181	1

Table 2.1: Four examples (two English and two Hindi) from the experiments from Sharma and Bhattacharyya (2013). In all the examples, the first word is more intense than the second word for that sense. The total frequency is lower for the most intense word in all the cases.

sense. They validate this correlation in English, Marathi and Hindi.

It is also important to mention that Sharma and Bhattacharyya (2013) show that there is a negative correlation between total frequency and intensity, even though they argue that the total frequency of a word cannot predict the intensity of a word within a particular sense. The semantic bleaching phenomenon maintains that the higher the total frequency of a word, the lower its intensity, while Sharma and Bhattacharyya (2013) focus on the intensity of a particular sense and claim that the count of a word in the sense is what can help us determine the intensity of the word in that specific sense. However, the intensity of a word in a sense could change depending on, for example, the register of the speaker or the domain of the text. This method still does not account for the different intensities that a word in a sense could have when used in different contexts.

The correlation between word frequency and intensity is also explored by Garcia et al. (2012). In their paper, they analyse German, Spanish and English texts. Their analysis suggests that there is a correlation between word intensity and frequency of use. The words that carry the most information are the most infrequent and specific and the least expected. They also argue that we can approximate the value of information in words when given the frequencies of words and word sequences.

Lewis et al. (2014) explore word frequency too, but they also look at word length. They study the bias for longer words to map to more complex meanings, and they show that this bias exists both in the lexicon and at the moment of language interpretation. They also write about theories that predict that longer utterances should be associated with more complex meanings. Horn (1984) proposed such a theory. He argued for a relationship between the cost of an utterance and the probability of meaning and suggested that the costlier the phrase, the less probable its meaning.

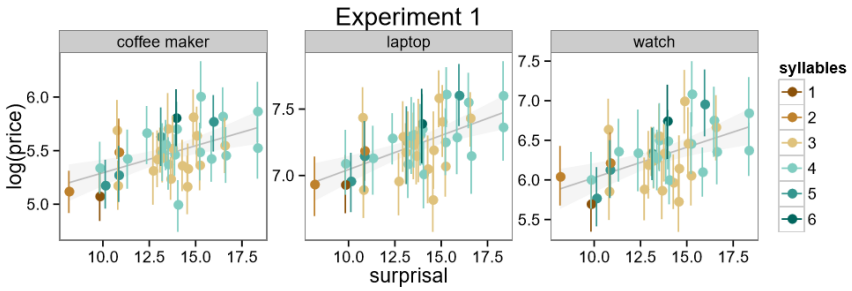

Figure 2.1: Results of Experiment 1 from Bennett and Goodman (2015). Participants were given descriptions of objects which included intensifiers and the word "expensive". They were asked to estimate the price of the objects. As the length and surprisal of the intensifier increased, so did the price responses the participants gave.

Bennett and Goodman (2015) have also taken word length into consideration in their work, in this case, with adverbs. They looked not only at the frequency of the adverbs but also at their length, and they showed that adjective phrases are more intense the longer and less frequent they are. They claim that this suggests that a word's meaning is not totally arbitrary and it can actually come from features of the word's form and distribution, such as length and frequency.

There has also been work focusing on producing detailed intensity lexicons. Mohammad (2017) has created an affect intensity lexicon with real-valued scores of association for different emotions: anger, fear, joy, sadness, disgust, trust, anticipation, and surprise. His technique, called best-worst scaling, improves annotation consistency and obtains reliable fine-grained scores. To achieve this, annotators are given several sets for each emotion. These are usually sets of 4 items, from which the annotators are asked to choose the best item and the worst item for the particular emotion. In this case, best translates to the most intense word, and worst to the least. These sets are randomly generated so that each word appears in 8 different tuples of 4, and the sets cannot have more than two items in common. The score, which goes from 0 to 1, for each word is the proportion of times it was chosen as being the best item in the set minus the proportion of times it was being the worst. This method obtains high correlations on repeat annotations, provides stable results and helps identify small differences in intensity (Mohammad, 2017).

3 Methodology

Based on the theory of semantic bleaching and on the many studies mentioned in Chapters 1 and 2, we argue that by looking at the frequency, length and number of senses of several synonymous adjectives, we can order these adjectives from least to most intense. Our study will focus on English, Danish and French. Table 3.1 contains an example of the kind of adjectives that we will be working with the most in our study. They are those used in the different translations for the questionnaires used in the European Values Study, 2008.

In our study we will focus mainly on word frequency. The expectation is that the least frequent adjectives will be the most intense ones, and vice versa. We will look at and comment on length and number of senses too. We will focus on analysing the results and the possible ways to improve on our method. We will also compare all the results we obtained from the different languages, in an attempt to show that the correlation is still maintained when working with more than one language and that we can establish if the translation of a word in one language has the same intensity as it does in a different language.

Our method is simple: first we select the adjectives we want to study and we make sure that they appear in our lists and databases so that we can obtain their intensity score. When we have their scores, we find out their frequency, their length or number of syllables, and their number of senses. Finally, we calculate the Spearman's rank correlation coefficient between the group of adjectives' intensity score and the three different features. It is important to mention that, when possible, we looked at the frequency of the adjective in all of its forms (plural, female, male...) instead of just its lemma.

In the cases of English and Danish, we will use large corpora of Internet text to obtain the word frequency. We are aware that, in their study of positive

Language	**Happiness**	**Life satisfaction**
Danish	lykkelig	tilfreds
English	happy	satisfied
Finnish	onnellinen	tyytyväinen
German	glücklich	zufrieden
Norwegian	lykkelig	tilfreds
Spanish	feliz	satisfecho
Swedish	lycklig	tillfredsställd
French	heureux	satisfait

Table 3.1: List of the words used in questions about happiness and in questions about life satisfaction by the European Values Study, 2008

words and their information content, Garcia et al. (2012) have shown that the language used on the Internet is not neutral but emotionally charged. They mention that this can affect quantitative analyses of the emotions in the text, because the "emotional reference point" is higher than zero. Nevertheless, this did not affect the conclusion of their research, which remained the same independently of the corpus they used to obtain the word frequencies. For French, we will use the Lexique corpus of film subtitles and the Lexique corpus of books (New et al., 2001). We chose to use these corpora instead of Internet text because it was faster and easier.

When it comes to the intensity scores, we will obtain the English scores from AFINN (Nielsen, 2011), a list of English words rated in intensity with an integer between -5 (negative) and +5 (positive). For Danish, we will use the Danish version of this list. For French, we will use the Quadflieg database, which includes intensity scores from 1 (very moderate) to 7 (very intense) (Quadflieg et al., 2014). While the Quadflieg database obtained its scores from the average of all scores given by 8 groups of 20 participants, the AFINN list was created by one single person. This means the quality and reliability of the AFINN list might not as high as it is for the Quadflieg database, and it could contain more biased scores.

To provide the translation of non-English terms we will use on-line dictionaries.[1] These translations are not totally equivalent to the English words and are only meant to help the reader. We will also look at on-line dictionaries to establish the number of senses the adjectives have. Since a word's number of senses is subjective, these numbers could be different to the ones we present in this thesis if looked up on another dictionary.

[1] http://www.parlor.dk/ and https://www.dict.com/

4 Experiments

4.1 Total Frequency of English adjectives

For our first analysis, we wanted to compare word frequency with word intensity for English adjectives. We focused on i) 10 positive adjectives that express satisfaction or happiness, ii) 10 negative adjectives that express sadness or discontent, iii) 10 positive adjectives that refer to quality and iv) 10 negative adjectives that refer to quality.

For the word frequencies we used Google Web 1T 5-Gram Database, a collection of frequent 5-grams extracted from approximately 1 trillion words of Web text.[1]

The intensity scores were obtained from AFINN (Nielsen, 2011), a list of 2477 English words, all of them rated in intensity with an integer between -5 (negative) and +5 (positive). In all but one case we chose adjectives that were present in the word list. The exception was "torturous", which did not appear in the word list. It was chosen due to the lack of intense negative quality adjectives in the list. We used the rating given to the noun "torture." All the definitions of English words and the number of senses have been taken from the English Oxford Dictionary. [2]

4.1.1 Happiness and Satisfaction Adjectives

In Table 4.1 we present the positive adjectives, ordered from most to least frequent, their frequencies, and their AFINN scores. Our expectation was to find that the most frequent words were the least intense ones, which was reflected in the results, though not perfectly. There is a clear increase of the intensity with the decrease of frequency, with the only exception being "happy", which is by far the most frequent word, but not the least intense. This could be due to the fact that "happy" is the word with the most senses out of the whole list. While happy has 3 different senses, the rest of words only have 1 (except for ecstatic, which has 2).

Table 4.2 shows the negative adjectives, their frequencies and scores. In this case the scores are generally lower. This might be a consequence of the fact that words with negative emotions are less often used and carry more information than positive words (Garcia et al., 2012). In some cases, we agree more with the frequency order than with the actual scores. For example, "sorrowful" ("feeling or showing grief") has a score of -2, while "apathetic" ("showing or

[1] http://nlp5.csie.org/~tzhuan/web1t5-easy/Web1T5_freq.perl?

[2] https://en.oxforddictionaries.com/

Word	Freq	Score
Happy	63507173	3
Satisfied	14589394	2
Glad	14587046	3
Pleased	13215862	3
Merry	6373348	3
Delighted	3812221	3
Joyful	1201131	3
Ecstatic	662928	4
Euphoric	255181	4
Overjoyed	248815	4

Table 4.1: Positive English happiness and satisfaction adjectives, their frequencies from most to least frequent and their AFINN score

Word	Frequency	Score
Affected	24301370	-1
Sad	13017058	-2
Depressed	3358586	-2
Unhappy	2737752	-2
Dissatisfied	1477991	-2
Apathetic	302358	-3
Sorrowful	272439	-2
Heartbroken	199986	-3
Woeful	157299	-3
Anguished	135874	-3

Table 4.2: Negative English happiness and satisfaction adjectives, their frequencies from most to least frequent and their AFINN score

Feature	Positive	Negative
Length	0.54	-0.04
Senses	0.10	-0.56
Freq.	-0.89	-0.84

Table 4.3: Spearman's correlation coefficients for syllable length/intensity, number of senses/intensity and word frequency/intensity the English happiness and satisfaction adjectives.

feeling no interest, enthusiasm, or concern") has a score of -3. A lack of interest shows less negative intensity than a state of grief. As opposed to the positive adjectives, here the most frequent word is actually the least intense one and there is no one word with a higher number of senses than all the rest. The three most frequent words have the same number of senses, 3, while the rest of the words have 1 or 2.

The correlation coefficients in Table 4.3 show that in this case, word frequency has the strongest correlation to intensity, with a coefficient of -0.89 for the positive adjectives and a coefficient of -0.84 for the negative ones. The correlation between word length and intensity is not strong enough, and is not even consistent between the positive and negative words. While it is positive for the positive words, it is slightly negative for the negative ones. In the case of number of senses, the results are similar, but in this case the negative correlation is stronger than the positive one.

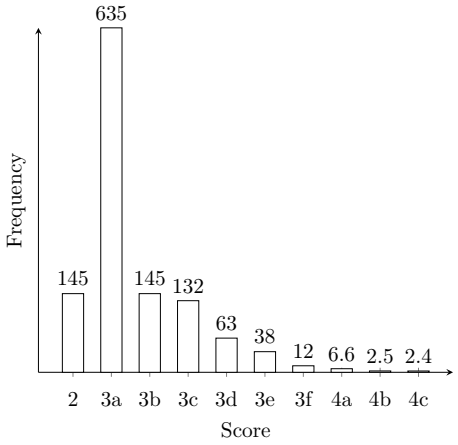

Figure 4.1: Frequency to score correlation for the positive English satisfaction and happiness adjectives (frequencies have been simplified).

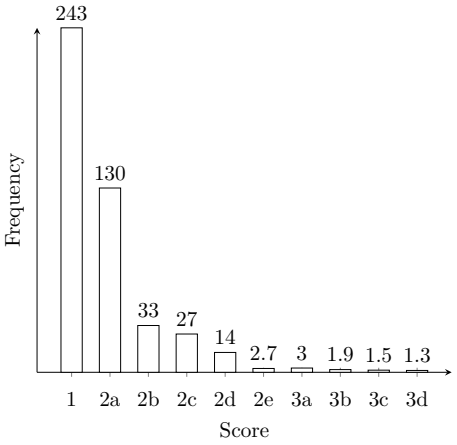

Figure 4.2: Frequency to score correlation for the negative English satisfaction and happiness adjectives (frequencies have been simplified).

4.1.2 Quality Adjectives

In this case, we obtained slightly better results with the positive adjectives. The order of the positive adjectives is almost perfect, with only "terrific" having an unexpected position in the list. If we look at their definitions we see that "terrific" means "extremely good; excellent", while "superb" means "very good; excellent". There is only a small difference, but it points to "terrific" being slightly more intense, which is reflected in their word frequency.

It is important to point out that even if "good" ("having the required qualities; of a high standard"), "great" ("of ability, quality, or eminence considerably above average") and "excellent" ("extremely good; outstanding") have the same score, it is evident to any speaker of English that "excellent" is more intense than "great" and "great" is more intense than "good". This is also reflected in their word frequency. Similarly, "bad" ("or poor quality or a low standard") has the same score as "worst" ("of the poorest quality or the lowest standard; least good or desirable"), but is not as intense.

For the negative adjectives, we see that the most frequent word is not the least intense. Once again, we argue this is due to the large amount of senses that the most frequent word has, compared to the rest. While "bad" has 8 different senses, none of the other words in the list has more than 3. In this group there is also another pair of words with an unexpected order in the list, "catastrophic" and "dreadful." While "catastrophic" has an intensity score of 4 and "dreaful" a score of 3, "dreadful" is less frequent. If we look at their definitions, we can see that they are really similar. "Catastrophic" means "involving or causing sudden great damage or suffering", while "dreadful" means "causing or involving great suffering, fear, or unhappiness; extremely bad or serious." Because of this and their similar frequencies, we would argue that both words should have the same score.

In this case, word frequency is once again the feature with the strongest correlation, with coefficients of -0.87 and -0.68. Word length and number of

senses perform better than they did for the happiness and satisfaction adjectives, but the correlation is still not as strong as it is for word frequency.

Word	Frequency	Score
Good	365822878	3
Great	301508460	3
Excellent	60325866	3
Wonderful	28800872	4
Fantastic	17379474	4
Awesome	15363751	4
Brilliant	10303324	4
Superb	7895134	5
Terrific	3652629	4
Breathtaking	2124735	5

Table 4.4: Positive English quality adjectives, their frequencies from most to least frequent and their AFINN score.

Word	Frequency	Score
Bad	95105339	-3
Poor	48053547	-2
Worst	16976651	-3
Terrible	8610277	-3
Nasty	7439073	-3
Awful	5217620	-3
Dire	2304282	-3
Catastrophic	1750256	-4
Dreadful	1240347	-3
Torturous	102719	-4*

Table 4.5: Negative English quality adjectives, their frequencies and their AFINN score. *(From "torture")

Feature	Positive	Negative
Length	0.36	0.71
Senses	-0.38	-0.49
Freq.	-0.87	-0.68

Table 4.6: Spearman's correlation coefficients for syllable length/intensity, number of senses/intensity and word frequency/intensity the English quality adjectives.

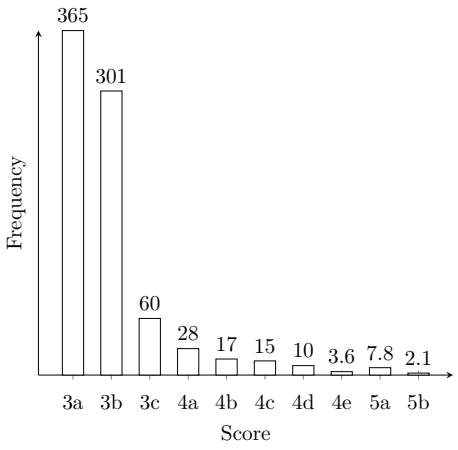

Figure 4.3: Frequency to score correlation for the positive English quality adjectives (frequencies have been simplified).

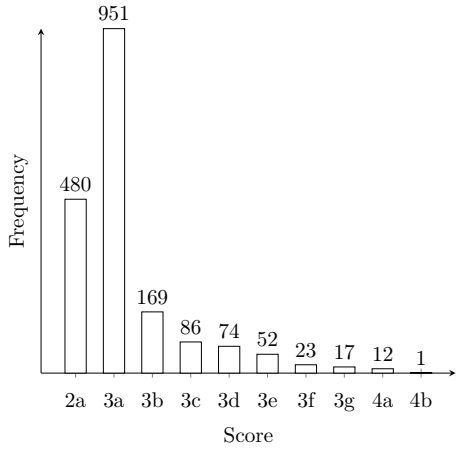

Figure 4.4: Frequency to score correlation for the negative English quality adjectives (frequencies have been simplified).

4.2 Total Frequency of Danish adjectives

Just as we did with the English adjectives, here we focused once again on i) 10 positive adjectives that express satisfaction or happiness, ii) 10 negative adjectives that express sadness or discontent, iii) 10 positive adjectives that refer to quality and iv) 10 negative adjectives that refer to quality. To find out the word frequencies, we used 4 gigabytes of tokenised Internet text in Danish and a short Python script to count the word frequency. The intensity scores were obtained from the Danish version of AFINN (Nielsen, 2011), which included 3552 words. We tried to select only adjectives that were present in the word lists (there were a couple of exceptions where the adjectives were not in the list and we used the score of similar words) and did not all receive the same rating. We obtained the number of senses from Den Danske Ordbog. [3]

4.2.1 Happiness and Satisfaction adjectives

We looked at 10 positive adjectives that express satisfaction or happiness and 10 negative adjectives that express sadness or discontent. As we did not count with any native speaker of Danish, we had to rely on dictionaries to find words with a similar meaning to the ones used in our analysis of English. Table 4.7 shows these words with their frequencies, approximate translation, and AFINN scores. Once again, the words have been ordered from most to least frequent.

Word	**Freq**	**Translation**	**Score**
Glad	122517	Glad	3
Tilfreds	46289	Satisfied	2
Glædelig	9529	Cheerful, gleeful	3
Lykkelig	9187	Happy	3
Salig	539	Blessed	3
Euforisk	477	Euphoric	4
Fornøjet	401	Pleased, glad	3*
Ekstatisk	379	Ecstatic	4
Frydefuld	86	Joyful	4*
Himmelhenrykt	33	Thrilled	4

Table 4.7: Positive Danish happiness and satisfaction adjectives, their frequencies from most to least frequent and their AFINN score. *There were no scores for these exact words, so we used the score given to a different word with the same root. In the case of "fornøjet" we used "fornøjelse" (delight, pleasure) and in the case of "frydefuld" we used "fryde" (to rejoice).

It is interesting to see that in this case, the equivalent of "happy" is not the most frequent word, but the fourth most frequent. As we mentioned in the Introduction, "lykkelig" has a stronger meaning than "happy" does, and if we compare the English and the Danish results, the word frequency reflects this. Indeed, when it comes to life satisfaction surveys, aggregate figures for Denmark usually indicate a top position. On the other hand, there is more

[3] http://ordnet.dk/ddo_en

Word	Freq	Translation	Score
Skuffet	15592	Concerned, anxious	-2
Trist	14782	Disappointed	-2
Dyster	4198	Gloomy	-2
Ulykkelig	2775	Unhappy	-2
Deprimeret	1404	Depressed	-2
Undertrykt	759	Melancholic, sad	-2
Vemodig	316	Dejected	-2
Desillusioneret	284	Saddened, dejected	-2
Bedrøvet	230	Lamentable	-2
Mismodig	59	Dispirited, discouraged	-3

Table 4.8: Negative Danish happiness and satisfaction adjectives, their frequencies from most to least frequent and their AFINN score.

Feature	Positive	Negative
Length	0.64	0.00
Senses	-0.45	-0.21
Freq.	-0.80	-0.52

Table 4.9: Spearman's correlation coefficients for syllable length/intensity, number of senses/intensity and word frequency/intensity the Danish happiness and satisfaction adjectives.

ambiguity when we talk about happiness. In the 2002 International Social Survey Programme, Denmark was ranked as number 13 among 35 countries, in contrast to the results obtained in the 2004-2008 World Values Survey, where Denmark was second among 59 countries (Lolle and Andersen, 2016).

For the negative adjectives, the intensity scores are much more similar, and once again, these adjectives are less intense than the positive ones. Once again, we find that the scores could be improved upon, since words that have the same score do not truly express the same intensity. While "trist" ("disapointed") and "deprimeret" ("depressed") share the same intensity score, they do not seem to express the same level of dissatisfaction, with "deprimeret" expressing a more intense feeling, as it is reflected in the word frequency order.

This lack of variation in the intensity scores results in much worse correlation coefficients. In Table 4.9 we can see that for the positive adjectives there is a strong correlation between length and intensity (0.64) and between word frequency and intensity (-0.80). On the other hand, when it comes to the negative adjectives, the numbers are much lower, with word frequency still having the strongest correlation (-0.52).

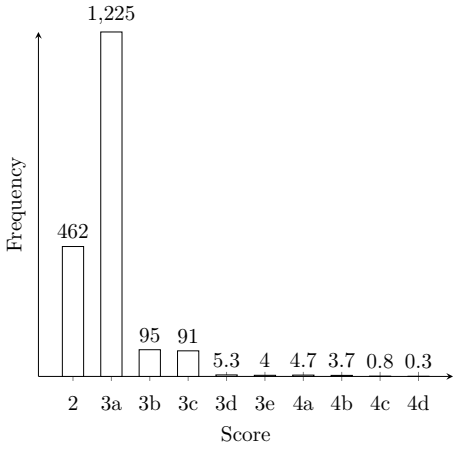

Figure 4.5: Frequency to score correlation for the positive Danish satisfaction and happiness adjectives (frequencies have been simplified).

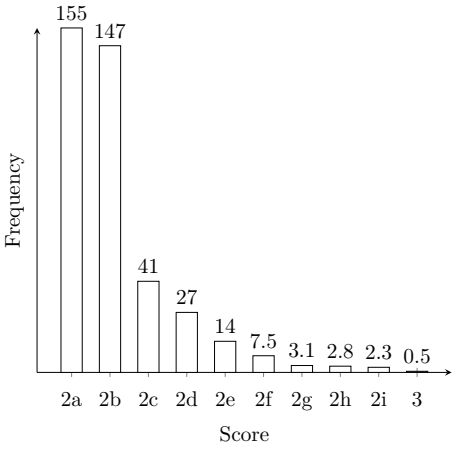

Figure 4.6: Frequency to score correlation for the negative Danish satisfaction and happiness adjectives (frequencies have been simplified).

4.2.2 Quality Adjectives

In the case of the quality adjectives, we see that the most frequent word is not the least intense one in both positive and negative words. In the case of "dårlig", it has 7 senses, while no other word in the list has more than 3. For the positive adjectives, both "god" and "fin" actually both have 6 senses. None of the other words have more than 2. Since "god" and "fin" have 6 senses, this does not explain why "gud" is more frequent. It could be the case that the score is not accurate and both words have a much more similar meaning and intensity, but we cannot analyse this like we did in the English analysis due to our lack of knowledge of Danish.

When it comes to word frequency, the positive adjectives have a Spearman coefficient of -0.79. The correlation is not as strong for the negative ones (-0.60), but it is still significant. For the number of senses, both correlations are too weak, and for the word length only the positive adjectives show a

Word	**Frequency**	**Translation**	**Score**
God	586351	Good	3
Fin	76732	Fine	2
Dygtig	32209	Competent	2
Sund	27512	Sound	2
Tjekket	8174	Great	2*
Vidunderlig	3391	Wonderful	4
Eminent	1680	Eminent	4
Betagende	1367	Breathtaking	5
Fabelagtig	876	Fabulous	4
Superb	730	Superb	5

Table 4.10: Positive Danish quality adjectives, from most to least frequent, with their frequencies, translations and AFINN scores. *(From "utjekket", with score -2)

Word	Frequency	Translation	Score
Dårlig	49319	Bad	-3
Forkert	39026	Wrong	-2
Værre	25070	Worse	-3
Ødelagt	13285	Broken	-2
Værst	6664	Worst	-3
Forfærdelig	4854	Apalling, Terrible	-3
Grim	4279	Horrible	-3
Frygtelig	3029	Terrible	-3
Grotesk	2652	Grotesque	-3
Katastrofal	1375	Catastrophic	-4

Table 4.11: Negative Danish quality adjectives, from most to least frequent, with their frequencies, translations an AFINN scores.

Feature	Positive	Negative
Length	0.66	0.17
Senses	-0.36	-0.27
Freq.	-0.79	-0.60

Table 4.12: Spearman's correlation coefficients for syllable length/intensity, number of senses/intensity and word frequency/intensity the Danish quality adjectives.

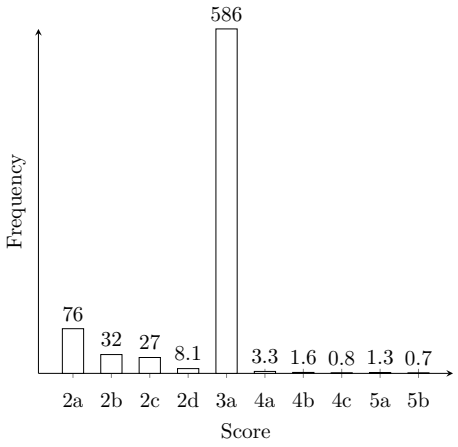

Figure 4.7: Frequency to score correlation for the positive Danish quality adjectives (frequencies have been simplified).

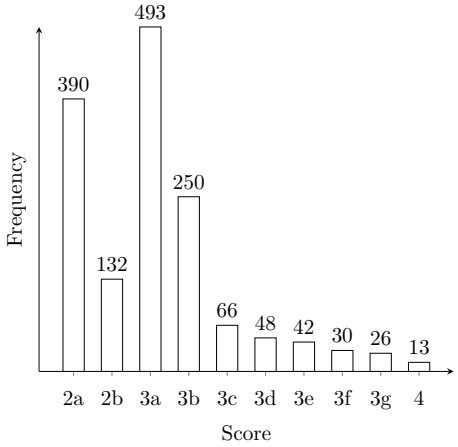

Figure 4.8: Frequency to score correlation for the negative Danish quality adjectives (frequencies have been simplified).

significant correlation (0.66).

4.3 Total frequency of French adjectives

For the analysis of French adjectives we also looked at 10 positive adjectives that express happiness and satisfaction and 10 negative adjectives that express sadness or discontent, but we could not study quality adjectives, since they

were not present in the database we used. Instead, we focused on 10 adjectives that express kindness and 10 adjectives that express meanness.

We obtained the word frequencies, number of syllables, translations and intensity scores (from 1 very moderate, to 7, very intense) from the Quadflieg database (Quadflieg et al., 2014). The senses were extracted from the Larousse dictionary.[4]

4.3.1 Happiness and satisfaction adjectives

Word	Translation	Film frq.	Book frq.	Avg. frq.	Intensity
Heureux	Happy	155.8	125.41	140.60	4.05
Content	Pleased	114.75	51.22	82.98	3.3
Joyeux	Joyous	31.81	26.69	29.25	3.85
Gai	Happy	10.74	22.64	16.69	3.85
Satisfait	Satisfied	4.42	10.41	7.41	3.05
Guilleret	Jolly	0.21	1.89	1.11	4.18
Béat	Blissfully happy	0.25	1.55	1.05	4.10
Enjoué	Jolly	0.14	2.09	0.90	4.32
Réjoui	Delighted	0.23	1.49	0.86	4.05
Extasié	Ecstatic	0.01	0.95	0.48	5.84

Table 4.13: Positive French happiness and satisfaction adjectives, their film, book and average frequency per million, and their intensity score.

Once again, "happy", here "heureux", is the most frequent word. We see similar tendencies to the ones in the analysis of English and Danish, with a general (but not perfect) correlation between frequency and intensity. Also, just like it happened in the English and Danish analysis, we can see again that the most frequent word, "heureux", is not the least intense one. It has 7 senses, while "joyeux" has 4 and the rest of the words has 3 or less.

In the analysis of English and Danish we saw that positive words were in general more intense than negative words, but in the French study we see the complete opposite, with many words having a score of over 5.

Is it worth mentioning too that the negative words present bigger differences between the film and book corpora than the positive words do, and that in both cases the film frequency shows a stronger correlation than the book frequency does. For the negative adjectives, the film frequency to score correlation coefficient is of -0.76, but for the book frequency the result is considerably lower, -0.59. On the other hand, for the positive adjectives the difference is much smaller, with the film frequency and the book frequency having Spearman coefficients of -0.69 and -0.67 respectively. We decided to work with an average of the two frequencies.

In the case of the positive words, the word frequency correlation coefficient is of -0.67, while for the negative words it is higher, -0.78. For number of senses, the positive words show a significant correlation (-0.63), but the negative words do not. There is not relevant correlation for word length in neither of the cases.

[4]http://www.larousse.fr/dictionnaires/francais/

Word	Translation	Film frq.	Book frq.	Avg. frq.	Intensity
Désolé	Desolate	193.51	6.49	100	3.47
Triste	Sad	93.03	86.89	89.96	4.00
Déçu	Disappointed	5.16	5.47	5.31	4.00
Contrarié	Annoyed	1.32	2.50	1.91	4.4
Angoissé	Anxious	0.28	2.57	1.42	5.16
Déprimé	Depressed	1.89	0.68	1.28	5.00
Désemparé	Distraught	0.20	2.23	1.21	5.42
Effaré	Aghast	0.03	1.62	0.82	5.10
Furieux	Furious	0.25	1.35	0.80	5.79
Cafardeux	Gloomy	0.07	0.34	0.20	5.00

Table 4.14: Negative French happiness and satisfaction adjectives, their film, book and average frequency per million, and their intensity score.

Feature	Positive	Negative
Length	0.12	0.31
Senses	-0.63	-0.11
Freq.	-0.67	-0.78

Table 4.15: Spearman's correlation coefficients for syllable length/intensity, number of senses/intensity and word frequency/intensity the French happiness and satisfaction adjectives.

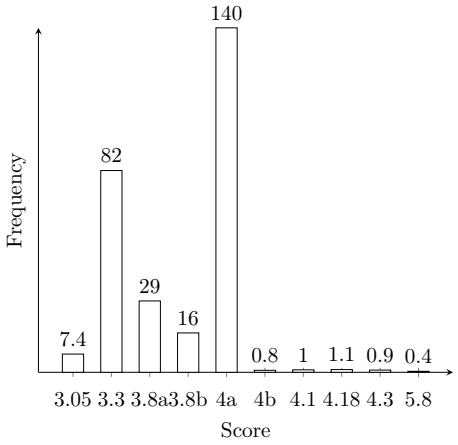

Figure 4.9: Frequency to score correlation for the positive French satisfaction and happiness adjectives (frequency and scores have been simplified).

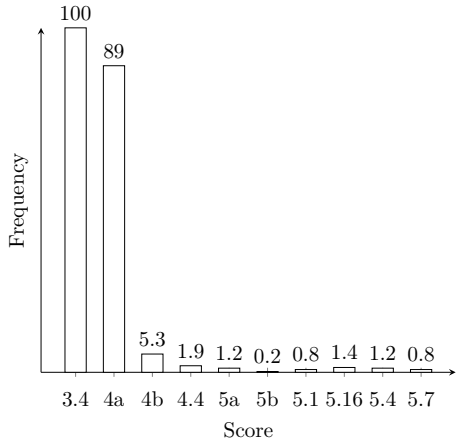

Figure 4.10: Frequency to score correlation for the negative French satisfaction and happiness adjectives (frequency and scores have been simplified).

4.3.2 Kindness and meanness adjectives

Word	Translation	Film Frq.	Book Frq.	Avg. Frq.	Score
Gentil	Kind	134.11	37.39	85.74	3.50
Joli	Nice	94.55	44.53	72.0	3.21
Agréable	Pleasant	37.71	31.35	34.53	3.30
Aimable	Kind	21.98	24.59	23.28	3.00
Amical	Friendly	4.26	11.22	7.74	3.50
Affectueux	Affectionate	3.56	7.97	5.76	3.80
Respectueux	Respectful	3.47	5.41	4.44	3.60
Débonnaire	Good-natured	0.06	2.64	1.35	4.65
Bonasse	Meek	0.08	1.62	0.85	5.11
Altruiste	Altruistic	0.52	0.00	0.26	4.74

Table 4.16: Positive French kindness adjectives, their film, book, and average frequency per million, and their intensity score.

Word	Translation	Film Frq.	Book Frq.	Avg. Frq.	Score
Méchant	Nasty/naughty	30.23	16.01	23.12	4.85
Aigre	Sour	0.42	8.38	4.40	4.30
Mesquin	Petty	1.64	2.91	2.27	4.80
Dédaigneux	Disdainful	0.26	3.85	2.05	5.00
Toxique	Toxic	3.00	0.34	1.67	5.25
Abject	Despicable	1.40	1.69	1.54	5.15
Belliquex	Belligerent	0.41	1.28	0.84	4.95
Envahissant	Intrusive	0.33	0.88	0.60	5.05
Injurieux	Abusive	0.09	0.88	0.48	5.00
Dictatorial	Dictatorial	0.03	0.20	0.11	5.90

Table 4.17: Negative French meanness adjectives, their film, book, and average frequency per million, and their intensity score.

Once again, the negative words are considerably more intense than the positive ones. Also, in this case the book frequency shows a stronger correlation than the film frequency. For the positive adjectives, the difference is small, since the Spearman coefficient is -0.87 for book frequency and -0.82 for film. However, the negative adjectives have a -0.80 coefficient for the book frequency but a very low -0.34 for the film frequency.

We see that the pattern repeats itself and in both sets of adjectives the most frequent word is not the least intense one, but the one with the most senses. In these cases, we have the positive adjective “gentil” with 6 senses, and the negative “méchant”, with 7. For all the other positive and negative adjectives, the maximum number of senses they have is 4.

Another surprising fact is that, even though word frequency is still the feature with the strongest correlation (-0.84 for positive words and -0.69 for positive ones), the number of senses also present a strong negative correlation to intensity, with a coefficient of -0.80 for positive words and -0.62 for the negative ones. The word length feature does not have a significant correlation to intensity in this case.

Feature	Positive	Negative
Length	0.06	0.53
Senses	-0.80	-0.62
Freq.	-0.84	-0.69

Table 4.18: Spearman's correlation coefficients for syllable length/intensity, number of senses/intensity and word frequency/intensity the French kindness and meanness adjectives.

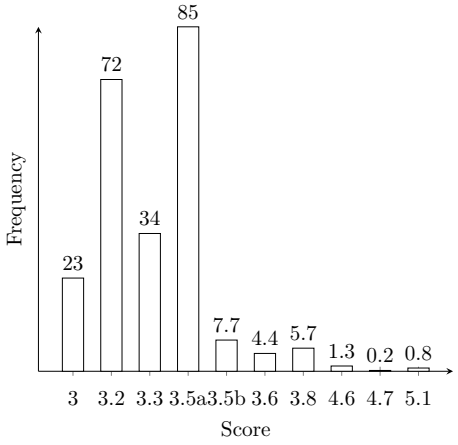

Figure 4.11: Frequency to score correlation for the positive French kindness adjectives (frequency and scores have been simplified).

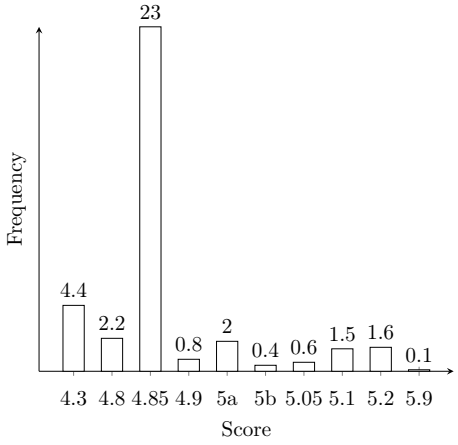

Figure 4.12: Frequency to score correlation for the negative French satisfaction and happiness adjectives (frequency and scores have been simplified).

5 Comparison between languages

When we look at the average correlation coefficients per language and feature, the results still point to word frequency as our best predictor. In Table 5.1 we can see that word frequency has a strong negative correlation to intensity in each of the languages, while number of senses and word length show a much weaker correlation, with number of senses in French having a coefficient stronger than ±0.40.

Language	Feature	Pos. H&S	Neg. H&S	Pos. Q.	Neg. Q	Pos. K	Neg. M	Average
English	**Length**	0.54	-0.04	0.36	0.71	–	–	0.39
	Senses	0.10	-0.56	-0.38	-0.49	–	–	-0.33
	Freq.	-0.89	-0.84	-0.87	-0.68	–	–	-0.82
Danish	**Length**	0.64	0.00	0.66	0.17	–	–	0.36
	Senses	-0.45	-0.21	-0.36	-0.27	–	–	-0.32
	Freq	-0.80	-0.52	-0.79	-0.60	–	–	-0.67
French	**Length**	0.12	0.31	–	–	0.06	0.53	0.25
	Senses	-0.63	-0.11	–	–	-0.80	-0.62	-0.54
	Freq.	-0.67	-0.78	–	–	-0.84	-0.69	-0.74

Table 5.1: Comparison of the Spearman correlation coefficients for syllable length/intensity, number of senses/intensity and word frequency/intensity in all the adjectives sets in English, Danish and French.

Having assessed the correlations separately, we are now going to take into consideration the English, Danish and French happiness and satisfaction adjectives together and the English and Danish quality adjectives together. We will once again look at the correlation between word intensity and word frequency, word length or number of syllables, and number of senses.

5.1 English, Danish and French

Word frequency still shows a strong negative correlation to word intensity when working with three different languages. In this case, the average coefficient was of -0.57. In general, as we can see in Tables 5.2 and 5.3 there are no big clusters of words in the same language in our two lists. There are four French words together in the positive group, “guilleret”, “béat”, “enjoué” and “réjoui”, and another four in the negative one, “déprimé”, “désemparé”, “effaré” and “furieux”, but in general the three languages seemed to be scattered all over the list, with no language focused on any extreme. We must mention that there are some issues that can affect our results. First of all, we have used two different intensity scales, with different degrees of specificity. Additionally, the different corpora and databases used to calculate word frequencies have different sources such as Internet text, books or film subtitles. Nevertheless,

	Word	Freq/ million	Inten.	Sy.	Se.		Word	Freq/ million	Inten.	Sy.	Se.
1	Glad (DA)	185.7	3.00	1	3	**16**	Joyful	1.2	3.00	2	1
2	Heureux	140.6	2.89	2	7	**17**	Guilleret	1.1	2.98	2	2
3	Content (FR)	82.9	2.35	2	3	**18**	Béat	1.0	2.92	2	2
4	Tilfreds	70.0	2.00	2	1	**19**	Enjoué	0.90	3.08	2	1
5	Happy	63.5	3.00	2	3	**20**	Réjoui	0.86	2.89	2	2
6	Joyeux	29.2	2.75	2	4	**21**	Salig	0.81	3.00	2	3
7	Gai	16.6	2.75	1	5	**22**	Euforisk	0.72	4.00	3	1
8	Satisfied	14.5	2.00	3	1	**23**	Ecstatic	0.66	4.00	3	2
9	Glad	14.5	3.00	1	1	**24**	Fornøjet	0.60	3.00	3	2
10	Glædelig	14.4	3.00	3	1	**25**	Ekstatisk	0.57	4.00	3	1
11	Lykkelig	14.0	3.00	3	2	**26**	Extasié	0.48	4.10	3	2
12	Pleased	13.2	3.00	1	1	**27**	Euphoric	0.25	4.00	3	1
13	Satisfait	7.4	2.17	3	3	**28**	Overjoyed	0.24	4.00	3	1
14	Merry	6.3	3.00	2	1	**29**	Frydefuld	0.13	4.00	3	1
15	Delighted	3.8	3.00	3	1	**30**	Himmelhenrykt	0.04	4.00	4	1

Table 5.2: Comparison of all the positive happiness and satisfaction adjectives in English, Danish and French, their frequency per million, intensity score, syllable length and number of senses. The intensity scores for French have been adapted from their 1 to 7 original rating to a 1 to 5 rating, similar to the one used for English and Danish.

	Word	Freq. per million	Intensity	Sy.	Se.		Word	Freq. per million	Intensity	Sy.	Se.
1	Désolé	100	2.47	3	3	**16**	Déprimé	1.28	3.57	3	1
2	Triste	89.9	2.80	1	6	**17**	Désemparé	1.21	3.87	4	1
3	Affected	24.3	1.00	3	3	**18**	Undertrykt	1.10	2.00	4	1
4	Skuffet	23.6	2.00	2	1	**19**	Effaré	0.82	3.64	3	1
5	Trist	22.3	2.00	2	1	**20**	Furieux	0.80	4.13	2	5
6	Sad	13.0	2.00	1	3	**21**	Vemodig	0.47	2.00	3	1
7	Dyster	6.3	2.00	2	2	**22**	Desillusioneret	0.43	2.00	3	1
8	Déçu	5.3	2.80	2	1	**23**	Bedrøvet	0.34	2.00	3	2
9	Ulykkelig	4.2	2.00	4	2	**24**	Apathetic	0.30	3.00	4	1
10	Depressed	3.3	2.00	2	3	**25**	Sorrowful	0.27	2.00	3	1
11	Unhappy	2.7	2.00	2	2	**26**	Cafardeux	0.20	3.57	3	2
12	Deprimeret:	2.1	2.00	4	1	**27**	Heartbroken	0.19	3.00	3	1
13	Contrarié	1.9	3.14	3	4	**28**	Woeful	0.15	3.00	2	2
14	Angoissé	1.4	3.68	3	3	**29**	Anguished	0.13	3.00	2	1
15	Dissatisfied	1.4	2.00	4	1	**30**	Mismodig	0.08	3.00	3	1

Table 5.3: Comparison of all the negative happiness and satisfaction adjectives in English, Danish and French, their frequency per million, intensity score, syllable length and number of senses. The intensity scores for French have been adapted from their 1 to 7 original rating to a 1 to 5 rating, similar to the one used for English and Danish.

	Positive	Negative	Average
Length	0.46	0.07	0.26
Senses	-0.44	0.05	-0.19
Frequency	-0.74	-0.40	-0.57

Table 5.4: Spearman's rank coefficients for the positive and the negative happiness and satisfaction adjectives, and the average coefficients.

taking into account these issues, we maintain it is important to present and analyse these results.

	Positive	Negative	Average
Length	0.57	0.40	0.48
Senses	-0.35	-0.46	-0.40
Frequency	-0.77	-0.66	-0.71

Table 5.7: Spearman's rank coefficients for the positive and the negative quality adjectives, and the average coefficients.

5.2 English and Danish

We have also examined the quality adjectives. French was not included since we do not have information about quality adjectives in that language. In this case, the coefficients were stronger for all features than they were for the happiness and satisfaction adjectives. Word frequency is also the feature with the strongest correlation, -0.71. In Table 5.5 we can see that the Danish "god" is much more frequent than the English "good", which would suggest the English word has a stronger intensity. On the other hand, "superb" has a similar frequency in both languages. For the negative words, Table 5.6 shows that the English "bad" is more frequent than the Danish "dårlig", while "worst" and "værst" and "katastrofal" and "catastrophic" are much closer in frequency to each other.

	Word	Freq/ million	Inten.	Sy.	Se.		Word	Freq/ million	Inten.	Sy.	Se.
1	God	888.4	3	1	6	**11**	Tjekket	12.3	2	2	2
2	Good	365.8	3	1	7	**12**	Brilliant	10.3	4	2	3
3	Great	301.5	3	1	5	**13**	Vidunderlig	5.1	4	4	2
4	Fin	116.2	2	1	6	**14**	Eminent	2.5	4	3	1
5	Excellent	60.3	3	3	1	**15**	Betagende	2.0	5	4	2
6	Dygtig	48.8	2	2	2	**16**	Fabelagtig	1.3	4	4	1
7	Sund	41.6	2	1	1	**17**	Superb (DA)	1.1	5	2	1
8	Wonderful	28.8	4	3	1	**18**	Superb (EN)	0.7	5	2	2
9	Fantastic	17.3	4	3	2	**19**	Terrific	0.3	4	3	2
10	Awesome	15.3	4	2	1	**20**	Breathtaking	0.2	5	3	1

Table 5.5: Comparison of all the positive quality adjectives in English and Danish, their frequency per million, intensity score, syllable length and number of senses.

	Word	Freq/ million	Inten.	Sy.	Se.		Word	Freq/ million	Inten.	Sy.	Se.
1	Bad	95.1	3	1	8	**11**	Forfærdelig	7.3	3	4	2
2	Dårlig	74.7	3	2	7	**12**	Grim	6.4	3	1	3
3	Forkert	59.1	2	2	2	**13**	Awful	5.2	3	2	3
4	Poor	48.0	2	1	3	**14**	Frygtelig	4.5	3	3	2
5	Værre	37.9	3	2	2	**15**	Grotesk	4.0	3	2	1
6	Ødelagt	20.1	2	3	2	**16**	Dite	2.3	3	1	2
7	Worst	16.9	3	1	1	**17**	Katastrofal	2.0	4	4	1
8	Værst	10.0	3	1	2	**18**	Catastrophic	1.7	4	4	1
9	Terrible	8.6	3	3	2	**19**	Dreadful	1.2	3	2	2
10	Nasty	7.4	3	2	3	**20**	Torturous	0.1	4	3	1

Table 5.6: Comparison of all the negative quality adjectives in English and Danish, their frequency per million, intensity score, syllable length and number of senses.

Terms used for life satisfaction	Freq. per million	Terms used for happiness	Freq. per million
Tyytyväinen	260.38	Gelukkig	204.92
Tilfreds (DA)	70.00	Glücklich	155.82
Zufrieden	44.96	Heureux	140.60
Tevreden	42.97	Onnellinen	64.97
Satisfecho	18.24	Happy	63.50
Satisfied	14.50	Feliz	60.26
Tilfreds (NO)	14.10	счастливы	59.96
Satisfait	7.41	Lycklig	45.77
удовлетворен	6.47	Lykkelig (NO)	23.10
Tillfredsställd	1.40	Lykkelig (DA)	14.00

Table 5.8: List of the words used in questions about happiness and in questions about life satisfaction by the European Values Study (2008) ordered from most to least frequent.

5.3 Multilingual study

To study the similarities in frequency for related languages, we took the words from Table 3.1 and their frequencies per million words and compared them. We obtained this information from several different sources and corpora. As we saw in the analysis of French, the frequency can change depending on the type of corpus used, which is why using different kinds of corpora for this comparison might not be ideal, but it was the simplest and quickest method available. For French, English and Danish, we used the frequency data we have already mentioned. For the rest of languages, we used:

- SUBTLEX-NL, a database of Dutch word frequencies based on 44 million words from television and movie subtitles. (Keuleers et al., 2010)
- Matthias Buchmeier's German frequency list, generated from film and television subtitles with a total of 25,399,099 words.[1]
- The Swedish Kelly-list, a frequency-based vocabulary list generated from a web-acquired corpus of 114 million words. (Volodina and Kokkinakis, 2012)
- The Spanish CREA reference corpus. (Real Academia Española, 2017)
- Suomen sanomalehtikielen taajuussanasto, the Frequency Lexicon of the Finnish Newspaper Language, based on 43,999,826 words. (CSC - IT Center for Science, 2004)
- the Frequency lists for Norwegian spoken and written language. (The Text Laboratory, ILN, 2017)
- A Russian frequency list made up of 16,336,972 words. (Sharoff, 2002)

[1] https://en.wiktionary.org/wiki/User:Matthias_Buchmeier/German_frequency_list-1-5000

In Table 5.8 we present all the words and their frequencies. There we can see that there is a tendency for related languages to cluster together. The Dutch "gelukkig" and the German "glücklich" are the two most frequent terms in the happiness list, and for life satisfaction, "zufrieden" is the third most frequent term while "tevreden" is the fourth. These two pairs of words sound and look very similar, and these results suggest that they might also have the same intensity. Dutch and German are West Germanic languages, and even though English is West Germanic too, its frequencies are more similar to those of Spanish, a Romance language. Spanish and English are two of the most widely spoken languages in the world, with millions of non-native speakers. This and the strong influence French and Latin had on English could be a reason why they are more similar in frequencies than they are to other languages of their own family, such as English and German or Spanish and French. In the case of French, "heureux" is the third most frequent term, but "satisfait" is the seventh most frequent word, half as frequent as the English "satisfied."

In the North Germanic family we have Danish, Norwegian and Swedish. In the case of the happiness terms, these languages have the least frequent words, but for satisfaction it is different. While the Danish "tilfreds" is the second most frequent term, the Norwegian "tilfreds", with a frequency of 14.10 appearances per million, is the seventh most frequent term and the Swedish "tillfredsställd" is the least frequent term with 1.40 per million. Danish is five times more frequent than Norwegian, which is ten times more frequent than Swedish. It is interesting to see that even though both the terms for satisfaction and the terms for happiness are very similar in the three languages, their frequencies are so different for the satisfaction terms.

Finnish, an Uralic language, has a surprisingly frequent satisfaction term, "tyytyväinen", with a frequency of 260.38 per million, and the fourth most frequent term for happiness, "onnellinen." It does not seem to share any similarities with any of the other languages in the list. On the other hand, Russian, an East Slavic language, is right next to Swedish in both sets. It is important to mention that the Russian frequencies were obtained by looking not only at the frequency of that exact term, but also at the frequency of that term plus different suffixes. The frequencies of the exact terms as they appear on the list are of 2.61 for the satisfaction word, which would not change its position in the list, and 10.45 for the happiness term which would turn it into the least frequent term. Similarly, for the rest of the adjectives, we tried to find out their frequency in all of their forms. In some cases, this was beyond our control, since the corpora did not specify if the frequencies were for all forms or just the lemma.

6 Discussion

In this thesis we have studied the correlation between three different word features (word frequency, number of syllables and number of senses) and word intensity, in this case, focusing on adjectives. We managed to show that there is a negative correlation between word frequency and word intensity, both when working with one language and when comparing two or more. It is important to mention that the three languages we have analysed are Indo-European, the results could be much different for languages from different families.

Very often, the most frequent word was not the least intense one, which seemed to be related to its high number of senses. For example, "bad" had 8 senses, "heureux" had 7 and "dårlig" also had 7. This was only reflected in the separate studies, when we combined different languages the most frequent word was not necessarily the one with the most senses.

These two facts suggest the possibility for a more complex model than ours, perhaps a linear regression model, where both features are taken into account, with number of senses having a smaller weight than frequency. While high frequency will predict a word to be less intense, the number of senses will predict a higher intensity for the word if its number of senses is higher than the rest. Furthermore, another improvement on this study would be to find a fast way to compute relative frequency for a word in a specific sense. This would allow us to have a specific intensity score for each of the word's senses.

In general, negative words showed a lower correlation between frequency and intensity than positive words. This might be due to negative words being less intense than positive ones (Garcia et al., 2012). Interestingly, in the French analysis the negative words were more intense than the positive ones, whereas in both English and Danish the positive words were more intense. Since we are not working with very large samples, we cannot claim that this is the case for all of French adjectives, but this issue could be an interesting subject of study. Second, as we can see in Table 5.1, the number of senses seems to be a better predictor of word intensity in French than it is in Danish and English.

	Pos. H&S	**Neg. H&S**	**Pos. Q**	**Neg. Q.**	**Avg.**
Length	0.46	0.07	0.57	0.40	0.37
Senses	-0.44	0.05	-0.35	-0.46	-0.30
Frequency	-0.74	-0.40	-0.77	-0.66	-0.64

Table 6.1: Spearman's rank correlation coefficient for the complete sets of positive happiness and satisfaction adjectives (English, Danish and French), negative happiness and satisfaction adjectives (English, Danish and French), positive quality adjectives (English and Danish) and negative quality adjectives (English and Danish)

Once again, we cannot know if this applies to French as a whole or not, but it would be worthy of further analysis. Finally, we have also seen that different kinds of corpora can give us very different word frequency results. For example, "désolé" has a film frequency of 103.51 per million, but a book frequency of only 6.49 per million.

The comparison between languages shows that the negative correlation between frequency and intensity is still significant when working with two or more different languages. In Table 6.1 we can see that the average Spearman's rank coefficient for word frequency is -0.64, while number of syllables, with 0.37, and number of senses, with -0.30, have a low correlation with intensity also when it comes to two or more languages.

As we have seen, the translations of the words used in the European Values Study had very different frequencies. Since we have found a correlation between frequency and intensity, this suggests that our initial claim that the results obtained by these kinds of surveys could be affected by the translations of the questions is not entirely unfounded and that it might be possible to use word frequency to control for translation effects.

It is important to mention that there are several possible sources of error in our study. First of all, we used different intensity rating scales and corpora which were created using different text genres. These intensity ratings are also, in the case of the AFINN list, decided by only one person, which can lead to very subjective scores. Furthermore, there is no consensus on the number of senses that each word has, which means the results obtained when studying this feature could be different when using a different dictionary. Finally, we cannot be sure how many of our frequencies were obtained by only looking for the lemma, which can also affect our results.

However, our method is generally quicker, cheaper and easier than manual annotation, and though it would definitely benefit from manual evaluation, it would still be less costly.

7 Conclusion

In this thesis, we have investigated the correlation between adjective intensity and three different features: word frequency, number of senses and number of syllables or word length. Our results point to a significant negative correlation between word frequency and the intensity of the word. It was also shown that, in general, the correlation between frequency and intensity was weaker for negative adjectives. There was no proof that there was a strong correlation between numbers of senses and intensity and between number of syllables and intensity. Nevertheless, there seemed to be a pattern where the most frequent adjective was not the least intense one, but it did have more senses than the other adjectives in the list. This suggests that even though the number of senses is not enough on itself to predict word intensity, it could be used to improve on the results obtained by studying only word frequency. However, this pattern was not reflected when we looked at several languages together. Relative frequency of the word in a particular sense could also be useful for this task, and its further study would be very valuable.

In future work, it would be interesting to study the differences in correlation strength between negative and positive words, and the differences in intensity between negative and positive words in different languages, since Danish and English had more intense positive words than negative words, but in the case of French negative words were more intense than positive ones. Also, the creation of a more complex model which also takes number of senses into consideration would be a great improvement on this present work.

Bibliography

Erin Bennett and Noah D. Goodman. Extremely costly intensifiers are stronger than quite costly ones. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, 2015.

Christopher J. Bryan, Gregory M. Waltona, Todd Rogers, and Carol S. Dwecka. Motivating voter turnout by invoking the self. *Proceedings of the National Academy of Sciences*, 108(31):12653–12656, 2011. ISSN 0027-8424. doi: 10.1073/pnas.1103343108.

CSC - IT Center for Science. Frequency Lexicon of the Finnish Newspaper Language, 2004. URL `http://urn.fi/urn:nbn:fi:lb-201405272`.

David Garcia, Antonios Garas, and Frank Schweitzer. Positive words carry less information than negative words. *EPJ Data Science*, 1(1):1–12, 2012. ISSN 21931127. doi: 10.1140/epjds3.

John Helliwell, Richard Layard, and Jeffrey Sachs. World Happiness Report. *New York: Sustainable Development Solutions Network*, 2017.

Laurence R. Horn. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. *Meaning, Form and Use in Context (GURT '84)*, pages 11–42, 1984. ISSN 01867207.

Emmanuel Keuleers, Marc Brysbaert, and Boris New. SUBTLEX-NL: a new measure for Dutch word frequency based on film subtitles. *Behavior research methods*, 42(3):643–650, 2010. ISSN 1554-351X. doi: 10.3758/BRM.42.3.643.

Molly Lewis, Elise Sugarman, and Michael C. Frank. The structure of the lexicon reflects principles of communication. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*, pages 845–850, 2014.

Henrik L. Lolle and Jørgen G. Andersen. Measuring Happiness and Overall Life Satisfaction: A Danish Survey Experiment on the Impact of Language and Translation Problems. *Journal of Happiness Studies*, 17(4):1337–1350, 2016. ISSN 15737780. doi: 10.1007/s10902-015-9646-4.

Antoine Meillet. L'évolution des formes grammaticales. In *Linguistique historique et linguistique générale*, pages 130–148. Hachette, 1912.

Carol L. Moder. Mechanisms of Change in Grammaticization. In *Frequency of Use and the Organization of Language*. Oxford University Press, 2007. ISBN 9780195301571.

Saif M. Mohammad. Word Affect Intensities. *arXiv preprint arXiv:1704.08798*, 2017.

Boris New, Christophe Pallier, Ludovic Ferrand, and Rafael Matos. Une base de données lexicales du français contemporain sur internet : Lexique™//a lexical database for contemporary french : Lexique™. *L'année psychologique*, 101(3):447–462, 2001. ISSN 0003-5033. doi: 10.3406/psy.2001.1341.

Finn Å. Nielsen. Afinn, March 2011. URL `http://www2.imm.dtu.dk/pubdb/p.php?6010`.

Steven T. Piantadosi, Harry Tily, Edward Gibson, and Paul Kay. Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, 108(9):3526–3529, 2017. doi: 10.1073/pnas.1012551108/-/DCSupplemental.3526-3529.

Susanne Quadflieg, Caroline Michel, Henryk Bukowski, and Paul Kay. A database of psycholinguistic and lexical properties for French adjectives referring to human and/or nonhuman attributes. *Canadian journal of experimental psychology = Revue canadienne de psychologie expérimentale*, 68(1): 67–76, 2014. doi: 10.1037/cep0000001.

Real Academia Española. Corpus de Referencia del Español Actual - Listado de frecuencias, 2017. URL `http://corpus.rae.es/lfrecuencias.html`.

Raksha Sharma and Pushpack Bhattacharyya. High, Medium or Low? Detecting Intensity Variation Among synonyms in Sentiment WordNet. In *Proceedings of the Eighth Global WordNet Conference*, 2013. ISBN 9789730207286.

Serge Sharoff. Meaning as use: exploitation of aligned corpora for the contrastive study of lexical semantics. In *Proc. of LREC2002*, pages 447–452, May 2002.

The Text Laboratory, ILN. Frequency lists for Norwegian spoken and written language, 2017. URL `http://www.tekstlab.uio.no/frekvensordlister/`.

Elena Volodina and Sofie J. Kokkinakis. Introducing the Swedish Kelly-list, a new lexical e-resource for Swedish. In *LREC*, pages 1040–1046, 2012.

Kai Von Fintel. The Formal Semantics of Grammaticalization. In *NELS 25 Proceedings*, volume 25, pages 175–190, 1995.

George K. Zipf. *Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology*. Martino Publishing, 1949. ISBN 978-1-61427-312-7. doi: 10.2307/2226729. Reprinted in 2012.