Fairy tales and moral values: a corpus-based approach
Contos de fadas e valores morais: um estudo baseado em corpus

Rafael Silveira Silva

Resumo: *O objetivo deste artigo é realizar um estudo de base em corpus envolvendo a presença de valores morais em uma seleção de dez contos resumidos em versão online (A Branca de Neve e os sete anões, Os três porquinhos, Cinderela, A Bela adormecida, As aventuras de Aladdin, A Bela e a Fera, A formiga e a cigarra, O príncipe Omar e a princesa Scheherazade, Chapeuzinho Vermelho, O patinho feio) dos Irmãos Grimm verificando os termos que podem estar relacionados com o campo semântico mencionado. Mais especificamente, este estudo irá investigar a frequência e o uso das palavras "love", "litlle" e "girl". Essas palavras foram escolhidas devido à possível conexão com os valores morais, após verificar-se a lista de frequência de palavras do corpus. Este estudo inicialmente envolveu a análise de frequência das palavras no corpus dos Irmãos Grimm e, em seguida, uma comparação dos resultados obtidos através do The Contemporary American English Corpus (COCA). O software utilizado para a análise dos dados no corpus dos Grimm foi o AntConc 3.2.4.*
Palavras-chave: *valores morais, contos de fadas, corpus, COCA.*

Abstract: *The purpose of this article is to carry out a corpus-based study on the presence of moral values in a selection of ten summarized online versions of the tales (Snow-white and the seven dwarfs, The three little pigs, Cinderella, The sleeping princess, The Adventures of Aladdin, Beauty and the beast, The ant and the cricket, Prince Omar and Princess Scheherazade, Little Red Riding Hood, The Ugly Duck) by the Grimm's Brothers by looking at the terms which can be said to relate to the above mentioned semantic field. More specifically, this study investigated the frequency and usage of the words" love"," little" and "girl". These words have been chosen due to their possible connection to moral values after checking the word frequency list of the whole corpus. The study initially involved the analysis of frequencies of words in the Brothers' Grimm corpus and, secondly, a comparison of the results with The Contemporary American English Corpus (Coca). The software used to analyze the data in Grimm's corpus was AntConc 3.2.4.*
Keywords: *moral values, fairy tales, corpus, COCA.*

1. Introduction

Household Tales and *the Grimm's collection* are classics among children's literature. As children we have been fascinated with Snow-white, Cinderella, The Sleep Princess and many others, we believed in a fairy who transforms the cruel reality into something as beautiful as the end of stories *And they lived happily ever after*. Fairy tales are defined as short narratives that can be told from person to person. These tales were called folk stories. The Grimm's brothers adapted the stories which were firstly invented to adults and then directed to children.

> *Fairy tales are narratives that have been shaped over centuries of retelling and that have achieved a basic narrative form that is a distillation of human experience. Their popularity is a confirmation not only of their aesthetic appeal, but also of their ability to speak to human heart* (Swann Jones, p.05, 2002).

Although fairy tales are a distinct genre within the larger category of folktales, the definition that marks a text as a fairy tale is a source of considerable dispute. There are some researches, such as Propp (1928), who supported that "fairy tales" tell the adventures of man in a world of fantasy. They say that fairy tales do not require fairies. The term *fairy tale* came with the translation of Madame D'Aulnoy's *conte de fées*, first used in her collection in 1697.

According to Zipes (1991, p.18*) fairy tales operate ideologically to indoctrinate children so that they will conform to dominant social standards which are not necessarily established on their behalf.* He believes that the stories carry more than just fun; for him, the relationship of fairy tales in the life of a child largely contributes to its education and to the understanding of the world we live in. The ideology, according to Zipes (1991), is disguised in the safety of tales which suffered a process of censorship to keep the hard cost of preserving the moral values of society.

The purpose of this article is to report on a corpus-based study on the presence of moral values in a selection of ten summarized online versions of the tales by the Grimm's Brothers by looking at the terms which can be said to relate to the semantic field of moral values (Snow-white and the seven dwarfs, The three little pigs, Cinderella, The sleeping princess, The Adventures of Aladdin, Beauty and the beast, The ant and the cricket, Prince Omar and Princess Scheherazade, Little Red Riding Hood, The Ugly Duck) in which the frequency of the words *love, little and girl* will be studied. These words have been chosen due to their possible connection to moral values after carrying out a quantitative analysis of the word frequency in the whole corpus. For the present analysis, I understand that moral values are part of social construction of human beings. According Charles Perrault (Ashliman, 2004: 04):

the moral of the story" usually is not difficult to ascertain. In traditional fairy tales, morals typically center around the preservation of existing values and the maintenance of social stability. There is much to praise in the messages of ancient tales. Many of the values that then held families and communities together in a hostile world, enabling them not only to survive but sometimes to thrive and prosper, seem timeless in their merit: diligence, honesty, generosity, dependability, perseverance, courage, and a unique balance of self-reliance and selflessness.

Although moral values in fairy tales have been discussed before (Seifert 1960), my contribution may be justified by bringing a new view on the content of the Grimm's tales. That is, some studies have given attention to the topic in the Brothers Grimm' tales; none, to my knowledge, have supported their argument empirically. So, this is the aim of this study: to supply the existing gap of some studies dedicated to the brothers Grimm's tales.

A corpus-based approach to unveiling moral values, probably, leads to greater understanding of the world we live in and how we can construct and deconstruct our values. In addition, the corpus may help many educators to work properly with these values.

More clearly, this research focused on the study of the frequency of occurrence of some words that refer or relate to moral values found in ten out of the 210 tales which compose the brothers Grimm's collection. These ten tales were analyzed using AntConc software which calculates the frequency of words in a corpus showing the occurrences and the percentage.

The analysis of frequency list highlighted the percentages of some previously selected words related to the semantic field of moral values in the tales. After that, I compared the words to The Corpus of Contemporary English (COCA).

2. What are fairy tales?

According to Zipes (2002: 160), *A fairy- story is one which touches on or uses Faerie[1], whatever its main purpose may be: satire, adventure, morality, fantasy. Faerie itself may perhaps most nearly be translated by Magic – but is a magic of peculiar mood and power, at the furthest pole from the vulgar devices of the laborious, scientific magician.*

In this study I define fairy tales according to previous analysis after reading the theories published about the topic. Fairy tales are psychological plots involving a central character that needs to go through adversity, and even struggle, to find the strength and determination to succeed. More than a beautiful story for children and adapted for cinemas, the tales at a young age can have the function of transmitting a message that causes reflection, even though its real understanding arises at maturity. The stories, in spite of their imperfections and disguised

---

[1] A *fairy* (also *faery, faerie, fay, fae*; euphemistically *wee folk, good folk, people of peace, fair folk*, etc.) is a type of mythical being or legendary creature, a form of spirit, often described as metaphysical, supernatural or preternatural.

ideologies, can help children discover that the world around them is not innocent; children can open their eyes to also be able to understand the fine line between what is a good character and a bad character. Thus, one can say that it is from the fairy tales that the first relationships are established with the construction of moral values, because when the child comes into contact with the universe of the fairy tales they rediscover their own world and pre-establish relationships. According to Ashliman (2004, p.197) *the central function of many folktales, especially fables and parables, is to deliver a moral teaching. Such tales often end with a proverb- like restatement of tale's moral conclusion.* Seifert (2002) supports the same idea, but in another perspective *the moral invites a rereading and reformulation –a different understanding – of the tale's moral value. The new understanding can take the form of nostalgic and/ or utopian (re)readings of the marvelous, which, I would argue, afford a pleasure of their own.* The Brothers indirectly construct the formation of the children through their stories.

3. Methodology

In this section I describe the design of my study. I define the computational tool used to carry out my study and indicate the steps which have been followed. AntConc 3.2.4w is the program which has been selected for this task. As Anthony's web page argues, the author of the software, (it) "A freeware concordance program for Windows, Macintosh OS X, and Linux." AntConc has been used in order to find the frequency lists as well as the concordance lists in my Grimm's corpus.

The corpus selected for the present study has already been presented as the Grimm's corpus. The method used for this analysis is based on computational corpus linguistics and comparative analysis obtained through The Contemporary American English Corpus (COCA). Comparative analysis implies carrying out an empirical analysis of the main corpus, which involves a computational analysis of frequencies of the words in the Grimm's corpus.
Next, I did a comparison with the same words by checking their occurrences and frequencies in COCA. I read almost one-hundred situations related with moral values, and then I did a selection of twenty situations involved in moral values choosing texts about fiction, because they were related with the purpose of this research.

The words *love, girl, little* were chosen, because they were always used in moral situations, every time those words occurred they were linked with situations involving an elaborate thought and perception of the problems in tales. Almost every time *little* appeared it was involved with fragility, simplicity, goodness, and dreams of new conditions of life, or to indicate the central role women/men might have in their social context. *"…he led the innocent little girl away…" (Snow- White); while Mother Duck did her best to console him. "Poor little ugly duckling!" she would say. "Why are you so different from the others?" (The Ugly Duck).*

The word *love* always occurs with the semantic escaping of any trouble. This word was able to turn the bad things that happens in life in something beautiful and magical; and give another meaning for the life of a person, such as Snow- White and the seven dwarfs -she would live if a man with a good heart gave her the real loving kiss; the real love would be the reason for new life. For example, in line 33 of Snow White, *"...he led the innocent little girl away. However, when they came to the fatal spot, the man's courage failed him and, leaving Snow White sitting beside a tree, he mumbled an excuse and ran off.*" *Love* is shown as a condition to be happy, and found a real meaning for the whole life. For example, in line 13, the sentence in Sleeping Princess *"Love," replied the fairy. "If a man of pure heart were to fall in love with her, that would bring her back to life!"*

The word *girl* was connected with adjectives like lovely, good, unfortunate, to indicate the fragility and the role women should follow, the moral are related with the usage the authors always adopt for the word, to show that a *girl* is someone who is humble, good and a victim of the cruelty of adults or someone with more experience in life. They might be indicating that a girl is sometimes fool, and she will always be. For example, in line 7 in Cinderella *"even dressed in rags with a dusty gray face from the cinders, was a lovely girl".* In The sleeping princess, the beauty of the girl is synonym of suffering because someone who has too much in life, in terms of good conditions or beauty is convicted in any way, as we see in general stories or movies. *"the most beautiful girl in the whole kingdom [...]The servant handed the girl the spindle ... and she pricked herself with it and. with a sigh, dropped to the floor."*

Finally with the results of those two corpora I designed a table comparing words ‰, by doing this I will have the percentages to highlight the moral values in the fairy tales. To generate a clearly result, the word *love* and its other forms were verified like: *loves, loved, loving.*

4. Results

My first step was to generate a list, with the help of the computer software AntConc, of the most frequently used words, leaving aside their functions. By doing this I was able to analyze those which were directly or indirectly related to moral values. I worked initially with 300 words of the frequency list of the fairy tales corpus. This list was cleared out as it is a well known fact that the words which are more frequently used in any kind of text are those with a mainly grammatical meaning. Those words are pronouns, prepositions, articles and some others. I decided to exclude them from my list mainly since they exert no influence on the final results of my research. Thus, after having removed all the function words, the list was composed of 100 words.

WORD	FREQ.	TOTAL NUMBER OF OCCURENCES %		WORD	FREQ.	TOTAL
				fairy	17	1,777684827
				merchant	17	1,777684827
				pig	17	1,777684827
		11446		came	16	1,673115131
was	132	13,80319983		do	16	1,673115131
said	49	5,123915089		house	16	1,673115131
be	36	3,764509045		told	16	1,673115131
is	34	3,555369654		wolf	16	1,673115131
little	33	3,450799958		magic	15	1,568545436
time	30	3,137090871		now	15	1,568545436
girl	28	2,92795148		very	15	1,568545436
one	28	2,92795148		eyes	14	1,46397574
could	27	2,823381784		have	14	1,46397574
castle	25	2,614242393		king	14	1,46397574
went	25	2,614242393		last	14	1,46397574
were	25	2,614242393		palace	14	1,46397574
day	23	2,405103001		princess	14	1,46397574
down	23	2,405103001		young	14	1,46397574
again	22	2,300533305		come	13	1,359406044
genie	22	2,300533305		father	13	1,359406044
did	21	2,19596361		man	13	1,359406044
lamp	20	2,091393914		began	12	1,254836348
asked	19	1,986824218		forest	12	1,254836348
daughter	19	1,986824218		found	12	1,254836348
see	19	1,986824218		good	12	1,254836348
there	19	1,986824218		life	12	1,254836348
away	18	1,882254523		old	12	1,254836348
home	18	1,882254523		replied	12	1,254836348
like	18	1,882254523		round	12	1,254836348
love	18	1,882254523		wizard	12	1,254836348
out	18	1,882254523		fell	11	1,150266653
beautiful	17	1,777684827		happy	11	1,150266653
door	17	1,777684827		left	11	1,150266653
				marry	11	1,150266653

mother	11	1,150266653		**take**	8	0,836557566
ring	11	1,150266653		**wife**	8	0,836557566
stepmother	11	1,150266653		**work**	8	0,836557566
				years	7	1,777684827
rose	10	1,045696957		**die**	6	0,627418174
set	10	1,045696957		**handsome**	6	0,627418174
want	10	1,045696957				
find	9	0,941127261		**happened**	6	0,627418174
heart	9	0,941127261				
kind	9	0,941127261		**hope**	6	0,627418174
lovely	9	0,941127261		**terrible**	6	0,627418174
though	9	0,941127261		**widow**	6	0,627418174
three	9	0,941127261		**winter**	6	0,627418174
servant	8	0,836557566		**beauty**	5	0,522848479
sleep	8	0,836557566		**black**	5	0,522848479
splendid	8	0,836557566		**dark**	5	0,522848479

Table 1. Cleared Frequency List.

Considering the results above, we can that a great number of the words in the tales show the contrasts and differences in society, such as hate, prejudice, violence, loses. The context of some tales brings the life of a poor and beautiful person suffers in life; or someone richer who is cursed for the whole life until he/she finds real love. Maybe the selection of words in the texts try to express the idea that a person who is humble is always good, in other words, a poor person seems to be synonymous of goodness. In the frequency list we can see that there are many occurrences for some words like *love, home, life, marry etc.,* which are related with the life people desire.

The table is a clear indicator of the content of the tales corpus, but my aim is to go one step forward in the study of the most frequent words. My next resolution is to research the frequencies of three words selected from the list which are related to moral values: *love, girl and little. "…with seven neat little beds […]*
Going back to the kitchen, Snow White had an idea. "I'll make them something to eat. When they come home, they'll be glad to find a meal ready."(Snow- White).

To support the collected data I have used AntConc in order to find occurrences of the most recurrent words and then compare with the ones found in COCA. A detailed screenshot of the comparison with the occurrence lists is shown figure 4.

Total No. Of Tokens: 11446	Love/loves/loving/ loved	Girl	Little
Total of occurrences of moral situations in Fairy Corpus	24 (0, 020968‰)	21 (1, 83470‰)	15 (1, 31050‰)
Total of occurrences of moral situations in COCA	14 (1, 22313‰)	02 (0, 174733‰)	01 (0, 08736‰)

Figure 4. Screenshot of Occurrences.

These percentages of tokens (words) may be considered as high percentages bearing in mind we are working with a list composed of the 100 most frequent words in the tales corpus. This table points to the presence of moral values in these tales, since they are words connected to this semantic field, but do all these examples present moral values in the corpus? In order to answer this question, I analyzed the concordances of each word separately.

Figure 7. LOVE

If we analyze in depth the concordances of the word *love*, there are 13 out of 20 direct examples of moral situations in figure 7. For example lines 7 until 20, e.g. in line 7 (Prince Omar and Princess Sherazard), in the sentence *"I'm not in love and so I can't get married." The king, who could not bear to be crossed in such an important matter, went into a rage. He shouted for the guards and ordered them to shut the prince in an old castle in the forest."* Or in line 13 (The sleeping princess), e.g. *"Love," replied the fairy. "If a man of pure heart were to fall in love with her, that would bring her back to life!"*

Figure 8. GIRL

On the other hand, if we study the word *girl* (figure 8) which appears 36 times, we find that almost every time this word appears, it is related to moral – 10 out of 36 examples, e.g. in line 4 (Cinderella): *"All the nice things, kind thoughts and loving touches were for her own*

daughters. And not just the kind thoughts and love, but also dresses, shoes, shawls, delicious food, comfy beds, as well as every home comfort. All this was laid on for her daughters. But, for the poor unhappy girl, there was nothing at all."; and in line 35 (Snow-White and the seven dwarfs): *"Good girl! If you promised not to open up to strangers, then of course you can't buy. You are a good girl indeed!" Then the old woman went on. "And as a reward for being good, I'm going to make you a gift of one of my apples!"*

Figure 9. LITTLE

In concordances of the word *little,* there are 10 out of 63 words related with moral values, e.g. in line 5 (Red Hiding Hood): *"the home of a pretty little girl known to everyone";* in lines 34 (Snow- White) :*"A whole world was stirring to life and the little girl was glad to see how silly her fears had been.";* and 40 (The Ugly Duck): *"Poor little ugly duckling!" she would say. "Why are you so different from the others?" And the ugly duckling felt worse than ever."*

My next step involved looking for the same words in the other corpus for comparison purpose: Grimm's and The Corpus of Contemporary American English (COCA). The motivation for this comparison was to determine if the percentages of frequency of use of these words were higher or shorter than the ones found in the reference corpora.

First a random selection of 20 occurrences was taken from the COCA corpus and the same was done in Grimm's corpus. Finally, I did the table comparisons by looking at the frequencies. The results are illustrated in tables 3 and 4 below:

	Love	Girl	Little
Samples	20	20	20
Related to moral values	**17**	**2**	**1**
Percentages	85%	10%	5%

Table 3. The Corpus of Contemporary American English (Coca)

	Love	Girl	Little
Samples	20	20	20
Related to moral values	**20**	**20**	**15**
Percentages	100%	100%	75%

Table 4. Grimm's corpus

As we can observe in table 3, these words have been used in moral situations a fewer percentage of times (love 85%, girl 10% and little 5%). The found data, in the first analyzed corpora generated less result than I expected, for example words like *girl* and *little* did not

register a great number of occurrences related to moral values. On the other hand, in Grimm's they represent a higher percentage (love 100%, girl 100% and little 75%).

5. Conclusion

The true aim of this article was to demonstrate that classical tales aimed to children may be relabelled by way of a corpus-based approach in an objective and empirical way, taking into account the evolution of society. Thus, my main idea was to demonstrate a higher percentage of situations related to moral values within Grimm's corpus empirically. Based on the results, we can say that children are exposed to a great number of words which impact their social behavior. The results obtained through the frequency list and the comparisons list proves there are words used many times in the tales in different situations. A corpus-based approach can evidence the possibility that these tales do not fit the moral values that are intended to instill in childhood according to the society we live in; because the tales carry more than a simple plot, they may carry important thoughts for the whole life in terms of moral.

I could also see that many positive and negative words can give indications of the construction of moral values, because of the high frequency found. This study was meant to show that fairy tales corpus research can be used in a more objective and empirical way, regarding developments in society.

References

ALCANTUD DÍAZ, María. *Cruelty and violence in the brothers Grimm's fairy tales collection: a corpus linguistics analysis. Revista Alicantina de Estudios Ingleses* 23: 173-185, 2010.

BRIGGS, Katharine Mary. *An encyclopedia of fairies*. New York,Pantheon Books. "Euphemistic names for fairies", 1976.

LEWIS, C. Seifert. *Fairy tales, sexuality, and gender in France 1690-1715*. Cambridge: Cambridge University Press, 2006.

ASHLIMAN, D. L. *Folk and fairy tales: a handbook*. Greenwood: Greenwood Publishing Group, 2004.

GRIMM, Jacob; GRIMM, Wilhelm. *The complete fairy tales, the Brothers Grimm*. Ware, Hertfordshire: Wordsworth Editions Limited, 2009.

GRIMM, Jacob; GRIMM, Wilhelm. *Household tales with the authors notes, translated by Margaret Hunt*, Subach,1884.

SWANN JONES, Steven. *The fairy tale: the magic mirror of the imagination*. Routledge: Routledge, 2002.

ZIPES, Jack David. *Fairy tales and the art of subversion: the classical genre for children and the process of civilization*. Routledge: Routledge, 1991.