

Violence in the Brothers Grimm's Fairy Tales: A Corpus-Based Approach

María Alcantud Díaz
University of Valencia
Maria.alcantud@uv.es

ABSTRACT

The purpose of this article is to carry out a corpus-based study on the presence of violence in a selection of eight tales by the Grimm's Brothers by looking at the terms which can be said to relate to the semantic field of violence. More specifically, this study will analyse a selection of eight tales in which the frequency of the words *cut*, *dead* and *blood* will be studied in detail. These words have been chosen due to their possible connection to violence after carrying out a quantitative analysis of the frequency of the whole main corpus. My initial hypothesis is that the corpus-based study of those eight tales would support my intuition regarding the high percentage of violence in the Brothers Grimm collection as opposed to the content of violence in a wider variety of texts. The study initially involved the analysis of frequencies of the lexical units in the Brothers' Grimm corpus and, secondly, a comparison of the results obtained in the frequency test to two reference corpora: the British National Corpus and the Cobuild Concordancer. The comparison and its results seem to indicate that there was a higher than average percentage of the use of words related directly or indirectly to violence.

1. Introduction

Kinder und Hausmärchen, known in English as *Children's and Household Tales* or *the Grimm's collection*) is a classic among children's literature. As children, we have all been fascinated by characters such as *Snow-White*, *Cinderella* and so on, and we have all dreamt of becoming one of them. Tatar (2004: xv), one of the worldwide renowned authorities on folklore and fairy tales, states in her preface to *The annotated Brothers Grimm* that "the fairy

tales collected in Germany almost two centuries ago by the Brothers Grimm continue to have a powerful hold on our culture. Adapted, revised, rescripted and bowdlerized, they greet us at the movies such as *Pretty Woman* or *Working Girl*, at the opera with *Hansel and Gretel* or *La Cenerentola*, and in advertisements for everything from Chanel to chocolate, and in visual media as often as in print”.

The purpose of this article is to carry out a corpus-based study on the presence of violence in a selection of eight tales by the Grimm’s Brothers, which compose the main corpus of the present study- by looking at the terms which can be said to relate to the semantic field of violence. More specifically, this study will analyse the frequency of the words *cut*, *dead* and *blood*. These words have been chosen due to their possible connection to violence after carrying out a pilot survey of the frequency of occurrence of the whole main corpus. For the present analysis, I understand violence as a verbal or non verbal action that results in physical or psychological harm to a person, object, place or animal. My initial hypothesis was that the corpus-based study of those eight tales would support my intuition regarding the high percentage of violence in the Brothers Grimm collection as opposed to the content of violence in a wider variety of texts. The study primarily involved the analysis of frequencies of the lexical units in the Brothers’ Grimm corpus and, secondly, a comparison of the results obtained in the frequency test with two reference corpora: British National Corpus and Cobuild Concordancer. The comparison and its results support the presence of higher- than-average percentage of the use of words related directly or indirectly to violence.

If my hypothesis was to be confirmed, the implications of my findings would be, that yet again, children are being exposed to violence through products which, supposedly, are products aimed at them.

Although this has been often affirmed and argued in numerous studies (i.e. Tatar 1987, 1992, 2004; Haase 2008; Zipes 1991;) My contribution may somehow be justified by bringing in a new view on the content of the brothers Grimm’s tales. That is, although many studies have paid attention to the violent content of the brothers Grimm’s tales, none, to my knowledge have supported their argument empirically. This is therefore the aim of this study, to fill an existing gap in the already numerous studies dedicated to the brothers Grimm’s tales. A corpus-based approach will provide a better understanding of such tales and make it possible for educators to deal with the tales more adequately. Labelling those readings more empirically and classifying in detail the lexical units belonging to the semantic field of violence can help understand the type of violence we are dealing with and, in addition, can help deal with it more effectively.

More specifically, this study will focus on the study of the frequency of occurrence of some of the lexical units referred to cruelty and violence found in eight out of the 210 tales which compose the brothers Grimm’s collection. These eight tales will be analyzed using Concapp, in order to prove that the content of violence in these tales is higher than average. The analysis will provide me with a frequency list of the lexical items which compose the Grimm’s corpus so as to highlight the percentages of some previously selected words related to cruelty and violence found in the tales. After that I will collate these results and compare them to other two corpora: the British National Corpus and the Cobuild Concordancer for comparison purpose. These results will be displayed in different graphics to make data understanding easier and visually clearer.

2. Outline of the present study

2.1 Statement of hypothesis

My main hypothesis here is that a corpus-based approach will contribute to a better understanding of such tales and will make it possible for educators to deal with the tales more adequately. Labelling the readings more empirically and classifying in detail the lexical units belonging to the semantic field of violence can help understand the type of violence we are dealing with and help deal with it more effectively.

2.2. Violence in the Brothers Grimm's Fairy Tales

As said before, the aim of the present study is to investigate the presence of violence in eight tales of the brothers Grimm's fairy tales collection using a corpus-based approach – in order to achieve an objective and empirical classification.

In my study, I primarily reviewed the different approaches which have been taken to research the presence of violence in the brothers Grimm's fairy tales, discovering that some of the most worldwide-recognised scholars who have researched the fairy tales, have provided many examples which confirm the high presence of violence in the Grimm's tales. Tatar, for instance confirms that “in fairy tales, nearly every character—from the most hardened criminal to the Virgin Mary- is capable of cruel behaviour” (1987:3-4), referring to the Grimm's tales. Interestingly, tales like *Hansel and Gretel*, *The boy who went forth to learn what fear was* and *The Juniper tree*, which all belong to the Grimm's collection, were included by Warner (1998:4), in a book described by the author herself as a book about fear. Not only violence, but the fact that the brothers Grimm tales were amongst the recommended titles by Nazis had a historical consequence: the Allied forces, after World War II, thought that the Grimm's fairy tales had contributed to Nazi atrocities and savagery. In fact, as Haase argues, the Nazis promoted German folk education and saw the folktales as a means to their racial and political ends (2008:407-408). For these reasons, their books were forbidden in England and America since, according to Haase, it was confirmed that these fairy tales were “profoundly repressive, fuelled prejudices and xenophobia, and glorified cruelty and militarism” (2008:408). Violence in the Grimm's tales is not only about killing or hurting, but also about indoctrinating male-children to learn fixed roles which have many factors in common with psychological or physical ill treatments of women by means of sexist and racist attitudes (Zipes 1991:47).

All the data mentioned above makes us question how it is possible to find so many scenes of violence in a collection of tales aimed at children. Tatar provides us with the answer to this question: the brothers Grimm had other scholars in mind when they published the tales (2004:xvi) as it was born as a philological research focused on collecting old stories from oral tradition—aimed at adults- to preserve German identity and not to achieve a collection of tales for children, in fact, the Grimm's collection became tales aimed at children due, at least in part, to marketing reasons (Alcantud 2009).

2.3 A corpus-based study of violence in the Brothers Grimm's Fairy Tales

Corpus linguistics focuses on the importance of studying patterns of real language research. More concretely, this discipline studies language through large collections of authentic –written and spoken- texts. These texts are called corpora and are used “to derive empirical knowledge about language” (Koteyko 2006:144-145). The above authors have highlighted the particular significance of *violence* in the brothers Grimm's collection from different approaches: socio-political, psychological and so on. Very little, however, has been investigated about the role played by violence within fairy tales from the point of view of a corpus-based approach. To fill this gap, I take as a starting point the work by Stubbs (1996), who explains how computer assisted analyses may provide a substantial and well documented alternative to the use of intuitive data as well as a new understanding of form-meaning relations. Stubbs establishes nine principles for this kind of analysis:

Linguistics is essentially a social science and an applied science; language should be studied in actual, attested, authentic instances of use, not as intuitive, invented, isolated sentences; the unit of study must be the whole texts; texts and texts types must be studied comparatively across text corpora; linguistics is concerned with the study of meaning: form and meaning are inseparable; there is no boundary between lexis and grammar: lexis and grammar are interdependent; much language use is routine; language in use transmits the culture; saussurian dualisms are misconceived (1996: 24-44).

The point of departure of my study will be Stubb's second and fourth principles, that is to say, that text must be studied comparatively across text corpora using authentic samples. As argued by Stubbs (1996: 45), when talking about Sinclairs' work, this

shows in a precise and concrete way how a large corpus and an associated technology create a viewpoint which can lead to innovations in linguistic description and theory. The essential vision underlying corpus linguistics is that computer-assisted analysis of language gives access to data which were previously unobservable, but which can now profoundly change our understanding of language". Thus, using a computer assisted analysis in my analysis of the Grimm's Brother tales will enable me to prove my hypothesis in an innovative and objective way.

The present study will consider the frequency of use of some words related to cruelty and violence. Following Stubb's proposal (1996), I will take into account two factors in the study of the tales: the first one is the study of collocations of those words, extracted after doing the frequency test, which could be suspected of having some relation to cruel or violent situations. The second one will be the comparison of my findings to the same words in the reference corpora in order to study the percentages of use of those words in violent and cruel situations. As said before, a quantitative and empirical investigation of the Grimm's collection will be carried out using three distinct corpora:

Firstly, the 8 tales which compose a corpus of 17,416 words. An electronic version, collated to a printed version (Grimm, 2009), has been used to make corpora research by computer easier to use. These tales are:

9. *The Twelve Brothers*. Twelve brothers are obliged to leave the castle where they live and their little sister had to find them after suffering a fatal spell which made her remain mute.
11. *Little Brother and Little Sister*. Two children who were seriously ill treated by their step mother had to escape from their house. The boy suffers a terrible spell.
15. *Hansel and Grethel*. Two children, brother and sister, are abandoned by their father and kidnapped by a mean witch who tries to eat them.
16. *The Three Snake-Leaves*. A man who is willing to do anything to bring back his dead wife, and the hard price he must pay.
21. *Cinderella*. A girl who has no mother suffers with her father's new marriage to a widow with two other daughters. She is the object of serious ill treatments.
40. *The Robber Bridegroom*. A young bride is married to a murderer.
46. *Fitcher's Bird*. A wizard takes the role of a beggar and catches pretty girls.
53. *Little Snow-White*. A girl, whose mother died, sees how her step mother tries to kill her on several occasions simply due to envy.

Secondly, I have used the Cobuild Concordancer¹ corpus. In order to compare the percentages of frequency of some words related to violence which belonged to my main corpus (Grimm's tales), I typed in some simple queries (the words: cut, dead and blood) and I got a display of 50 concordance lines chosen at random from the corpus and which have been used for comparison purposes.

Thirdly I used British National Corpus.² I typed in the same queries as in Cobuild Concordancer, (cut, dead and blood) and the search result showed the total frequency in the corpus and up to 50 examples, used for comparison purposes as well.

I first carried out the quantitative analysis of the frequency by primarily identifying in a corpus all occurrences of a node word (word form or lemma) and its raw frequency. After that, I kept a record of collocates of this node which occurs in a window of defined size (i.e. four words to left and right). Then I counted the frequency of joint occurrence of node and each collocate; in this case the frequency of each collocation when related to violence and its estimation in percentages. Finally I collated these results with the ones achieved by the same process in the two reference corpora.

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. ConcApp Concordancing program³ is the program which has been selected for this task. As ConcApp web page argues, (it) "is a free and user-friendly text analysis program. It offers concordances, collocations and word frequency statistics. It can also be used to edit text files". ConcApp has been used in order to find the frequencies lists as well as the concordances lists in my tales corpus.

The corpus selected for the present study has already been presented as the brothers Grimm's tales. The method used for this analysis is based on computational corpus linguistics and comparative analysis. Comparative analysis implies carrying out an empirical analysis

of the main corpus, which involves a computational analysis of frequencies of the lexical units in the Brothers' Grimm corpus. For that reason I have generated a list of all the words in it, ordered in alphabetical order and by percentage of frequency, in which it is possible to observe the lexical units which are used in a higher percentage in the tales. Analyzing this kind of information provided me with: firstly a study of the most frequently used words- not taking into account function words. It enabled me to check the concordances of some words which, having a high percentage of frequency of use in the Grimm's corpus, are related to violent and cruel situations. In other words, it allowed me to find out when these words have been used in the Grimm's corpus relating to violence or cruelty.

Secondly, a comparative study of the same selected words. That is to say, after finding the percentages of frequency of use in violent or cruel situations in the Grimm's corpus, I collated them to the percentages of use in the same situations achieved when typing the same words in the other two reference corpora, the Cobuild Concordancer and the British National Corpus. These two corpora provided me with a random sample of 50 examples in which the queries asked were included. So it was possible to achieve the percentages of use of the selected words in violent situations. By doing this, and after collating all the results, it was possible to prove my hypothesis.

4. Results

My first step was to generate a list of the most frequently used words, not taking into account function words. By doing this I was able to check those which were directly or indirectly related to violent and cruel situations. I worked with just the first 200 words of the frequency list of the 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 all 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 (cf. Pérez Paredes 2002).

Thus after having removed all the function words, the list was composed of 89 words (see table 1)

| WORD | FREQ | % | | | |
|--------|------|----------|-----------|----|----------|
| Said | 172 | 0,9862 % | Queen | 43 | 0,2465 % |
| Little | 113 | 0,6479 % | Saw | 43 | 0,2465 % |
| Came | 76 | 0,4358 % | Come | 41 | 0,2351 % |
| Went | 76 | 0,4358 % | Sister | 41 | 0,2351 % |
| Woman | 52 | 0,2981 % | Old | 39 | 0,2236 % |
| Have | 51 | 0,2924 % | Great | 37 | 0,2121 % |
| King | 48 | 0,2752 % | Cried | 35 | 0,2007 % |
| Go | 46 | 0,2637 % | Maiden | 34 | 0,1949 % |
| Forest | 44 | 0,2523 % | Looked | 33 | 0,1892 % |
| House | 44 | 0,2523 % | Door | 32 | 0,1835 % |
| Took | 44 | 0,2523 % | Home | 32 | 0,1835 % |
| | | | Let | 32 | 0,1835 % |
| | | | Beautiful | 30 | 0,1720 % |

| | | |
|---------------|----|----------|
| Answered | 29 | 0,1663 % |
| Thought | 29 | 0,1663 % |
| Dear | 28 | 0,1605 % |
| Time | 28 | 0,1605 % |
| Day | 27 | 0,1548 % |
| White | 27 | 0,1548 % |
| Children | 26 | 0,1491 % |
| Long | 26 | 0,1491 % |
| Put | 26 | 0,1491 % |
| Father | 25 | 0,1433 % |
| Heard | 25 | 0,1433 % |
| Heart | 25 | 0,1433 % |
| Son | 25 | 0,1433 % |
| Way | 25 | 0,1433 % |
| Child | 24 | 0,1376 % |
| Good | 24 | 0,1376 % |
| king's | 24 | 0,1376 % |
| Art | 23 | 0,1319 % |
| Bride | 23 | 0,1319 % |
| Dwarfs | 22 | 0,1261 % |
| looking-glass | 22 | 0,1261 % |
| Young | 22 | 0,1261 % |
| Bread | 21 | 0,1204 % |
| Gave | 21 | 0,1204 % |
| Lay | 21 | 0,1204 % |
| Like | 21 | 0,1204 % |
| Night | 21 | 0,1204 % |
| Poor | 21 | 0,1204 % |
| Pick | 21 | 0,1204 % |
| See | 21 | 0,1204 % |
| Daughter | 20 | 0,1147 % |
| Brothers | 19 | 0,1089 % |
| Brother | 19 | 0,1089 % |
| Began | 19 | 0,1089 % |
| Cut | 19 | 0,1089 % |
| Dead | 19 | 0,1089 % |
| Evening | 19 | 0,1089 % |
| Get | 19 | 0,1089 % |
| Opened | 19 | 0,1089 % |
| step-mother | 19 | 0,1089 % |
| Sat | 19 | 0,1089 % |
| Stood | 19 | 0,1089 % |
| Wife | 19 | 0,1089 % |
| Called | 18 | 0,1032 % |
| Fell | 18 | 0,1032 % |
| Girl | 18 | 0,1032 % |
| Tree | 18 | 0,1032 % |
| Blood | 17 | 0,0975 % |
| Pieces | 17 | 0,0975 % |
| Red | 17 | 0,0975 % |
| Ran | 17 | 0,0975 % |
| Fire | 16 | 0,0917 % |
| Find | 16 | 0,0917 % |
| Got | 16 | 0,0917 % |
| Has | 16 | 0,0917 % |
| Left | 16 | 0,0917 % |
| Mother | 16 | 0,0917 % |
| Morning | 16 | 0,0917 % |
| Man | 16 | 0,0917 % |
| Take | 16 | 0,0917 % |
| Asked | 15 | 0,0860 % |
| Ashes | 15 | 0,0860 % |
| Bird | 15 | 0,0860 % |
| Life | 15 | 0,0860 % |
| Pretty | 15 | 0,0860 % |
| Shoe | 15 | 0,0860 % |
| Through | 15 | 0,0860 % |

Table 1. Cleared frequency list.

Considering the results above, it can be observed that, essentially, there are a lot of verbs in past tenses, which shows it is not just a simple story but in addition a story with plenty of action in it. Besides, and more related to the present study, there are some words directly related to violence: *dead, blood*. The high frequency occurrences of these lexical units in comparison to other words thus highlight a relation of the tales to topics related to violence and cruelty. One interesting finding is that there are some words which apparently seem to have no relation to the semantic field of violence or cruelty but, if we have a closer look at their concordances within the tales corpus, there is a clear relationship with it. This is the case of words like *pieces, fire* and *heart*. They are amongst the most frequent words used in this corpus (pieces is used 17 times, heart 25 and fire 16). If we have a look at their concordances lists, we can observe the high percentage in which these words are used related to violence (fig.1 and fig. 2).

1 and then cast you into the fire, as was promised to fire, was an omen that you may not want to be in the
 2 by wild beast, but the witch was cast into fire and miserably burnt, and so was the old woman.
 3 said, "You children, lay yourselves down by fire and sleep, as I will go into the forest and cut some
 4 . But first I must have a drink, and then I will go to the fire, and then I will go to the forest and cut some
 5 of these good wood," and then she went to the fire, and when she was there, she saw a little piece of
 6 a branch, which she had delivered. They took the fire and extinguished the flames, and then they
 7 a while to go to the forest, and so they went to fire and then they went to the forest and cut some
 8 of these good wood. But as she was there they saw fire and then they went to the forest and cut some
 9 and then they went to the forest, and so they went to fire and then they went to the forest and cut some
 10 children, and so they went to fire, and then they went to the forest and cut some
 11 of these good wood, and then they went to fire and then they went to the forest and cut some
 12 of these good wood, and then they went to fire and then they went to the forest and cut some
 13 of these good wood, and then they went to fire and then they went to the forest and cut some
 14 of these good wood, and then they went to fire and then they went to the forest and cut some
 15 of these good wood, and then they went to fire and then they went to the forest and cut some
 16 of these good wood, and then they went to fire and then they went to the forest and cut some

Figure 1. Concordances of FIRE.

If we study in depth the concordances of the word *fire*, there are 7 out of 16 direct examples of violent or cruel situations in figure 1. For example, see line 2, the sentence “by wild beast, but the witch was cast into the fire and miserably burnt [...]” or line 13 “And when she was bound fast to the stake, and the fire was licking at her clothes with its red tongue”.

On the other hand, if we study the word *pieces* (figure 2) which appears 17 times, almost every time this word appears, it is related to violence – 15 out of 17 examples.

1 and then cast you into the fire, as was promised to pieces, was an omen that you may not want to be in the
 2 by wild beast, but the witch was cast into pieces and miserably burnt, and so was the old woman.
 3 said, "You children, lay yourselves down by pieces and sleep, as I will go into the forest and cut some
 4 . But first I must have a drink, and then I will go to the pieces, and then I will go to the forest and cut some
 5 of these good wood," and then she went to the pieces, and when she was there, she saw a little piece of
 6 a branch, which she had delivered. They took the pieces and extinguished the flames, and then they
 7 a while to go to the forest, and so they went to pieces and then they went to the forest and cut some
 8 of these good wood. But as she was there they saw pieces and then they went to the forest and cut some
 9 and then they went to the forest, and so they went to pieces and then they went to the forest and cut some
 10 children, and so they went to pieces, and then they went to the forest and cut some
 11 of these good wood, and then they went to pieces and then they went to the forest and cut some
 12 of these good wood, and then they went to pieces and then they went to the forest and cut some
 13 of these good wood, and then they went to pieces and then they went to the forest and cut some
 14 of these good wood, and then they went to pieces and then they went to the forest and cut some
 15 of these good wood, and then they went to pieces and then they went to the forest and cut some
 16 of these good wood, and then they went to pieces and then they went to the forest and cut some
 17 of these good wood, and then they went to pieces and then they went to the forest and cut some

Figure 2. Concordances of PIECES.

Examples of this can be seen in line 2 “and therein lay human beings, dead and hewn to pieces, and hard by was a block of wood[...]”. Or line 16 “Have thee in their power, they will cut thee to pieces without mercy, will cook thee [...]”.

Some other examples can be seen in lexical units related to the semantic field of ‘parts of the body’. The word *heart*, for instance, provided me with interesting data: almost every time this word was used, it was related to a violent or cruel situation (figure 3).

13 out of the 25 times this word is used in the text are related to violence. For example line

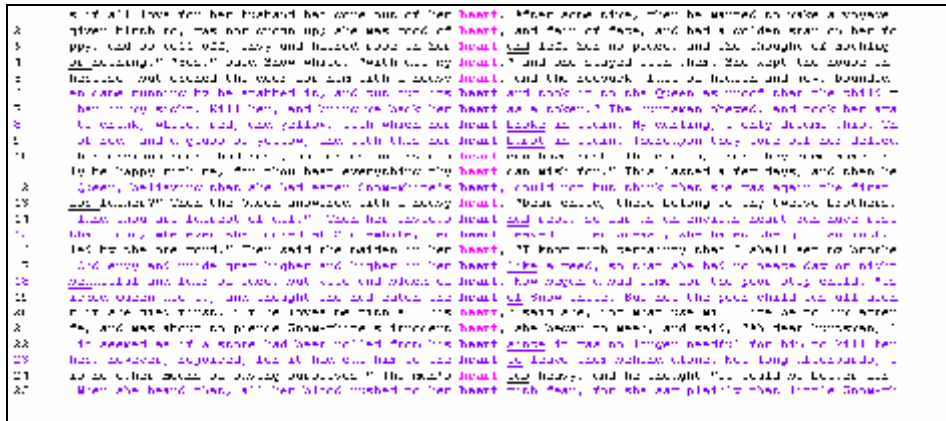


Figure 3. Concordances of HEART.

6 “Came running by he stabbed it, and cut out its heart and took it to the Queen[...] or line 12 “believing that she had eaten Snow-White’s heart, could not but thing she was the first[...]”. It has to be taken into account that this one is not a word related per se to it. These tables are a clear indicator of the content of the tales corpus, but my intention is to go one step forward in the study of the most frequent words.

My next step is to research the frequencies of three words selected from the list which, at first sight, are related to violence: *cut*, *blood*, *dead*. I have used ConcApp in order to find their frequency percentages. Figures 4, 5 and 6 show a detailed screenshot of ConcApp with the frequency percentages lists.

As it can be seen, there are 19 examples of the word *cut* which represents 0.1089% of the

| Unique Words | Instances | Percent |
|--------------|-----------|----------|
| cut | 1 | 0.0057 % |
| cutches | 1 | 0.0057 % |
| cut | 2 | 0.0115 % |
| cuttingly | 1 | 0.0057 % |
| cuttainsly | 2 | 0.0115 % |
| cutt | 1 | 0.0057 % |
| cuttains | 1 | 0.0057 % |
| cut | 19 | 0.1089 % |
| cutting | 1 | 0.0057 % |

Figure 4. Cut.

| Unique Words | Instances | Percent |
|--------------|-----------|----------|
| daughter | 20 | 0.1147 % |
| daughter | 3 | 0.0172 % |
| daughter | 1 | 0.0057 % |
| day | 27 | 0.1548 % |
| daybreak | 1 | 0.0057 % |
| days | 4 | 0.0229 % |
| dead | 19 | 0.1089 % |
| deadly | 1 | 0.0057 % |
| Days | 28 | 0.1605 % |

Figure 5. Dead.

| Unique Words | Instances | Percent |
|--------------|-----------|----------|
| blessed | 2 | 0.0086 % |
| blessing | 2 | 0.0086 % |
| blindfold | 2 | 0.0086 % |
| blood | 2 | 0.0129 % |
| blood | 18 | 0.0774 % |
| blood-red | 2 | 0.0086 % |
| bloody | 4 | 0.0172 % |
| blowing | 2 | 0.0086 % |
| blower | 2 | 0.0086 % |

Figure 6. Blood.

whole corpus. In terms of the word *blood*, there are 18 examples (0.0774 %), which should be added to the two examples of the word *blood-red* (0.0086%) and the 4 examples of the word *bloody* (0.172%). Finally, I have found 19 examples of the word *dead* (0.1089%) plus one example of the word *deadly* (which is not representative), as illustrated in table 2 below.

| Selected words | Frequency percentage | Figure |
|----------------|----------------------|--------|
| CUT | 0.1089 % | 7 |
| BLOOD | 0.0774 % | 8 |
| DEAD | 0.1089 % | 9 |

Table 2. Frequency percentages of cut, blood and dead.

These percentages can be considered as high percentages bearing in mind we are working with a list composed of the 89 most frequent words in the tales corpus. This table confirms once again the remarkable presence of violence in these tales, since they are words related intrinsically to this semantic field, but are all these examples used violently in the corpus? In order to answer this question, my next step will be to extract percentages of how many of times these words have been used related to violence and cruelty within the tales by using a concordance analysis. In figures 7, 8, 9, 10, and 11 screenshots of the concordances related to these words in my corpus will be seen:

Let's study the times in which these words appear in violent or cruel situations in table 4:

1 "Will have to make need to go on foot." The maiden cut a bit off her hair, shaved her feet into the shoe,
 2 "Tempt Me here another space for a knife and said, "This p bit off thy hair, when thou art Queen thou wilt
 3 we off her delicate tresses, laid her on a table, cut her beautiful body in pieces and strewed hair their
 4 "When the shoe, shagged her tresses by her hair, cut her nose with the black, and bound her in golden
 5 "Will" The father, however, refused, for it had cut the hair to leave them behind when. He l
 6 he said, and then it saw the other young maid and cut in pieces, at each mark, but such case again cut
 7 and too there in the basket, mostly murdered, and cut in pieces. But she began to gather about their bag
 8 as it was hard to come off, he took at one and cut it off, but the finger sprang up in the air, and a
 9 head just then came crashing by he stumbled at, and cut, the head, and took it to the Queen as good to
 10 she dice and start, he will go into the forest and cut wood. When we have done, we will come back an
 11 and of power!" said the old woman. "Only, I like cut the apple in two pieces: you eat the one, cleave an
 12 a wedge, she got like a barrel of honey, and when cut her forehead open and pulled herself in it, and
 13 it would not come off so soon, he took at one and cut the finger off, but it sprang up in the air, away
 14 as they see that she was faced too tightly, she cut the ladder that she began to breathe a little, and
 15 will have to make need to go on foot." The maiden cut the hair off, indeed the lock into the shoe, and
 16 her. Then her mother gave her a knife and said, "This cut now off, when thou art Queen thou wilt have to
 17 cut it be (condemned)" and had at one brought and cut the tree down, but so she was in it. And when they
 18 and when they have done in their power, they will cut their hair pieces without mercy, will reach thee, and
 19 y sleep a little: we are going into the forest to cut wood, and in the evening when we are done, we will

Figure 7. CUT

1 n. But what did she see when she went in? A great bloody basin stood in the middle of the room, and therein
 2 an ounce by the red spots that she had seen in the bloody chamber. "Since thou hast gone into the room again,
 3 led away by her curiosity, opened the door of the bloody chamber, looked in, and had to stare for it such a

Figure 9. BLOODY

1 It was, however, not the white, but the blood-red flag which announced that they were
 2 "— whosoever we find a girl, her red blood shall flow." Thereupon they went deeper
 3 that her face was hurt, she washed the blood off him, laid herbs on the wound, and as
 4 a. And hardly had he done this than the blood stirred in her veins, rose into her pale
 5 "Turn and peep, turn and peep, there's blood within the shoe, the shoe it is too small
 6 "He looked at her foot and saw how the blood was streaming from it. He turned his hor
 7 "Turn and peep, turn and peep, there's blood within the shoe The shoe it is too small
 8 looked down at her foot and saw how the blood was running out of her shoe, and how it
 9 d -- "Turn and peep, turn and peep, the blood is in the shoe. The shoe is not too small
 10 he basin. She got it out and washed the blood off, but in vain, it appeared again in a
 11 ce, and bawed her in places so that her blood ran on the ground. Then he threw her int
 12 nd as he could perceive no trace of any blood on it, he said, "Thou hast stood the tes
 13 ger with the needle, and three drops of blood fell upon the snow. And the red looked p
 14 had a child as white as snow, as red as blood, and as black as the wood of the window-
 15 who was as white as snow, and as red as blood, and her hair was as black as ebony and
 16 as she." When she heard that, all her blood rushed to her heart with fear, for she s
 17 a loud and said, "White as snow, red as blood, black as ebony-wood! this time the dea
 18 for she was as white as snow, as red as blood, and her hair was as black as ebony. It
 19 being 3 0,0172 & 112 blood-red 1 0,0057 & 113 blood
 20 blood-red 1 0,0057 & 113 blood 17 0,0976 & 114 bewitched 5 0

Figure 8. BLOOD

1 g raised. It was, however, not the white, but the blood-red flag which announced that they were all to die. Th

Figure 10. BLOOD-RED

| | DEAD | CUT | BLOOD |
|--------------------------------------|------|-------|-------|
| Total numb. of examples | 19 | 19 | 24 |
| Numb. of times in violent situations | 19 | 12 | 15 |
| Percentages | 100% | 63.1% | 62.5% |

Table 3. Percentages of frequency

As we can see in table 3, a high percentage of times (100% for *dead*) these words have been used in violent situations. These are very high percentages and they might already confirm my hypothesis which stated the presence of a high percentage of violence in the Grimm's tales. However once I have studied the corpus, my next step is to look for the same words in the two corpora proposed for comparison purpose: BNC, The British National Corpus and Cobuild Concordancer. The reason for this comparison is to determine if the percentages of frequency of use of these words are higher or shorter than the ones found in the reference corpora.

The next step in the analysis involves generating examples of the use of these three words in a larger corpus; with the intention to compare them to the corpus of the brothers Grimm's tales.

First a random selection of 50 solutions was taken from the British National corpus and the Cobuild Concordancer. The results are illustrated in tables 4 and 5 below:

| | DEAD | CUT | BLOOD |
|-----------------------------|------|-----|-------|
| Samples | 50 | 50 | 50 |
| Times in violent situations | 17 | 3 | 9 |
| Percentages | 34% | 6% | 18% |

Table 4. British National Corpus.

| | DEAD | CUT | BLOOD |
|-----------------------------|------|-----|-------|
| Samples | 50 | 50 | 50 |
| Times in violent situations | 14 | 0 | 6 |
| Percentages | 22% | 0% | 12% |

Table 5. Cobuild Concordancer.

If we combine all the percentages in frequencies related to violent situations in just one table (table 6):

| | Cut | Blood | Dead |
|----------------------|-------|-------|------|
| Grimms Tales | 63.1% | 62.5% | 100% |
| BNC | 6% | 18% | 12% |
| Cobuild Concordancer | 0% | 12% | 22% |

Table 6. Comparison British National Corpus, Cobuild Concordancer and Grimm's corpora.

And we transform these data into a graph in order to obtain visual evidence of the difference in percentages (see graphic 1), we can observe that the Brother Grimm's corpus percentages on the use of violence are largely higher than the results obtained in the two reference corpora used for comparison purposes.

5. Conclusions

In this research into the brothers Grimm's fairy tales, I have studied a corpus composed of 8 of their tales in order to make a twofold computational analysis. This study has been carried out in two steps: firstly the analysis of frequencies of the lexical units in the Brothers' Grimm corpus. Secondly, comparing it to the British National Corpus (BNC) and the Cobuild Concordancer in which I found a high percentage of use of words related directly or indirectly to violence.

The computational study unquestionably provided confirmation of a high percentage of violence in the brothers Grimm's tales. Thus, I have found out the frequency of use and percentages of some words related to cruelty and violence: cut + parts of the body, dead and blood found in the tales. After that I have compared these results to other two corpora, the British National Corpus and the Cobuild Concordancer, in order to discover such a high percentage of violence in the text which leads us to think of them as not aimed at children. It is widely recognized, violence can be found where you least expect it: cartoons, sports, readings, movies, etc and the presence of violence in the Grimm's collection is another key example.

The true aim of this article was to demonstrate that classical tales aimed at children can be relabelled by way of a corpus-based approach in an objective and empirical way, taking into account the evolution of society. In this sense it seemed worthwhile using the presence of violence in the brothers Grimm's fairy tales collection to detect it following Stubb's line of study on corpus analysis. Thus, my point of departure was to demonstrate a higher than average percentage of presence of actions related to violence and cruelty within the Grimm's corpus empirically. After doing it, it has become apparent that any lexical unit which might be controversial – in other words, topics which could not be total or at least in part aimed at children- might be detected.

Almost no investigation has been carried out by means of a computer-based approach to the brothers Grimm's fairy tales. My results seem to be pointing to the possibility of getting classical readings relabelled according to a standard, empirical and objective list of conventions on literature aimed at children.

The present research is just a starting point of what I hope will offer a thorough insight into the Brothers Grimm's Fairy tales. I have only covered eight out of the 209 tales taken from the original version but, as observed, the presence of violent behaviour is a characteristic of almost all of them. The results, no doubt, are interesting enough to divert the public's attention towards a more in depth and multidisciplinary study of the whole collection itself, which I hope to continue doing. I realise the limitations in the review of the literature section which I hope to extend by including the perspectives of other disciplines such as psychology and social sciences, which, I am sure will give me a richer insight into the interpretation of the tales

themselves and how they are or should be used in nowadays society.

To finish with, I want to make it clear that the historical and literary quality of the brothers Grimm Collection is not being questioned in this article. The only point here is to study the possibility that this collection does not suit the scale of values that we are trying to instil in our children according to the society in which we are living nowadays.

Notes

1. <http://www.collins.co.uk/Corpus/CorpusSearch.aspx>.
2. <http://www.natcorp.ox.ac.uk/corpus/index.xml>
3. <http://www.edict.com.hk/PUB/concapp/> and <http://www.intute.ac.uk/artsandhumanities/cgi-bin/fullrecord.pl?handle=humbul4918>

References

- Alcantud-Díaz, María (2009): *Cruelty and violence in the brothers Grimm's fairy tales collection: a corpus linguistics analysis*. Research project. Universitat de València.
- Atkins, B.T.S. and Zampolli, A.(1994): *Computational Approaches to the lexicon*. Oxford: Oxford University Press.
- Biber, D., Conrad, S. & Reppen, R. (1998): *Corpus linguistics investigating language structure and use*. Cambridge: Cambridge University Press.
- Grimm, Jacob.& Wilhelm. (2009): *The complete fairy tales, the Brothers Grimm*. Ware, Hertfordshire: Wordsworth Editions Limited.
- Grimm, Jacob & Wilhelm (trans. Margaret Hunt) (1884):*Household tales with the authors notes, translated by Margaret Hunt*.
[<http://www.surlalunefairytales.com/authors/grimms.html>].
- Haase, Donald. (2008): *The Greenwood encyclopedia of folktales & Fairy tales*. The USA: Greenwood Publishing Group.
- Koteyko, Nelya (2006): "Corpus linguistics and the study of meaning in discourse". *The linguistics Journal* 1(1): 132-157.
- Pérez Hernández, M. C. (2002): "Explotación de los córpora textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento." *Estudios de Lingüística Española (ELiEs)* 18. [<http://elies.rediris.es/elies18/>].
- Pérez Paredes, Pascual (2003): "Small corpora as assisting tools in the teaching of English news language: A preliminary tokens-based examination of Michael Swan's Practical English Usage news language wordlist". *Journal ESP world*. [http://www.esp-rld.info/articles_6/issue_6.htm].
- Rinolucrí, Mario (2003): "The challenge corpus linguistics offers to language teachers". in G. Luque Agulló, A. Bueno González & G. Tejada Molina (eds), *Las lenguas en un mundo global – Languages in a global world*. Jaén: Universidad de Jaen, 147-156.
- Stubbs, M. (1996): *Text and corpus analysis*. Oxford: Blackwell.
- Tatar, Maria (1987):*The hard facts of the Grimms' Fairy tales*. Princeton: Princeton University Press.
- _____. (1992): *Off with your heads*. Princeton: Princeton University Press.
- _____. (2004): *The annotated Brothers Grimm*. New York: W.W. Norton & Company.
- Zippe, Jack (1991): *Fairy Tales and the Art of Subversion. The Classical Genre for Children and the Process of Civilization*. New York: Routledge.