

科学研究費助成事業 研究成果報告書

平成 26 年 5 月 21 日現在

機関番号：10101

研究種目：挑戦的萌芽研究

研究期間：2011～2013

課題番号：23652067

研究課題名(和文) グリム童話集における数字使用のコンピュータ検索と、その偏差の意味考察

研究課題名(英文) Computer search of the numbers used in the Grimm's fairy tales, and the meaning consideration of its usage

研究代表者

石川 克知 (ISHIKAWA, Katsutomo)

北海道大学・メディア・コミュニケーション研究院・教授

研究者番号：30142665

交付決定額(研究期間全体)：(直接経費) 1,000,000円、(間接経費) 300,000円

研究成果の概要(和文)：グリム童話の全作品のデジタルデータベース化を実現できたことは、大変大きな成果であった。今回はその中の数字の分析を取り扱ったが、数字の内在解釈に関しても、コンピュータの処理結果と照合すれば、その説得力が明示的になったといえる。

汎用データベースが完成しているので、そのほか、グリム童話を対象とした様々なテキスト処理が可能になると言える。

研究成果の概要(英文)：It was a very great achievement that we could make a digital database of the entire work of the Grimm's fairy tales. And this time we dealt with analysis of the numbers in the Grimm's fairy tales. It can therefore be said that the result of our interpretation would be more explicit compared with the usual analog text analysis.

Since the generic database has been completed, we can also consider various usage of it.

研究分野：人文学

科研費の分科・細目：文学・ヨーロッパ文学(英文学を除く)

キーワード：グリム童話 数字 作品内在解釈 デジタルデータベース

### 1. 研究開始当初の背景

現在、電子テキストが本格的に普及しはじめ、また、それを取り扱うコンピュータの処理能力の向上や、インターネットによる汎用データベース構築環境の整備充実という、文学研究とコンピュータ処理の共同作業に向けた新しい研究体制が整い始めた。デジタルテキストの処理に関して言えば、特に言語学研究の統計処理という形では、すでに多くの研究がなされているが、文学作品の解釈に関して言えば、やはりデジタル統計処理と作品内在分析のようなアナログ処理では、ほとんど相容れない、いわば水と油の関係が続いているといえる。従って、いくら文学作品の電子テキスト化が進んでも、本格的で説得力のある作品解釈にそれが寄与することはないといえる残念な状況である。ところで、本科研の代表研究者石川克知は、長年、文学作品の研究とコンピュータ処理を組み合わせる試みを行ってきたが、現在十分な処理環境が整ったと考え、デジタル化された文学作品をただ「閲覧」するだけではなく、コンピュータにしかできない処理の対象として研究に応用することを試みることにした。研究対象としたのは、グリム童話集であるが、その選択理由は、第一に作品の分量と構造が、今回の科研期間の中で処理可能な量であると思われたこと、同時に、分担研究者の高橋吉文が、グリム童話における数字の意味を研究しており、それは、特に「数字」というコンピュータと親和性の高い対象をテーマとしていたからである。そこで、この機会に、科研費を利用し、本格的なテキスト処理が可能なグリム童話のデジタルデータベースを構築し、それを利用して、作品内在分析というアナログ処理の共同作業を試みようとした次第である。

### 2. 研究の目的

グリム作品における数字の意味に関しては、これまでも様々な研究が行われてきたが、どのような見解も、きわめて限定した作品のみに関する、主観的な解釈で、補佐する理論としては、その他、文学芸術における、一般的な数字の象徴性と関連づけられるものが多いように思われる。その理由としては、まず、グリム童話が、全体で200もの独立した内容の物語の集合体であること、また、さらには、エーレンベルク草稿という「手稿」から、第7版まで、度重なる変更が加えられており、全体を見渡すのが困難な状態にあることがあげられる。それ故、たとえば、ある童話において3という数字がシンボリックに使われていたとしても、はたしてグリム兄弟は、作品全体において3をどのように使っているかということはいえない。なぜなら、全8版に及ぶ全作品内での数字の使い方を、人間が数えていくのは、現実的な操作とはいえないからである。ところが、全作品が適切な形式のデジタルテキストデータベースにな

っていれば、この膨大な作業が瞬時に可能となる。或いは、単に数字の登場頻度表を調べても、あまり意味はないという意見もあるかもしれないが、今回の研究では、その登場頻度表だけでも、きわめて顕著でかつ重要な意味を、「客観的に」示していることを実証するのが、目的のひとつである。また、「感情のない」無機的なコンピュータ処理を、きわめて「人間的な」作品内在分析の専門家の研究と緊密に結びつけることで、一般的に水と油のように見なされている、コンピュータのデジタル処理とアナログ文学分析を有効に結びつけることが可能であることを示すのが最大の目的である。同時に、デジタル汎用データベースが完成すれば、数字以外にも、様々な応用が可能となる。つまり、今後ますます増加する文学作品のデジタルデータの、有効利用法にひとつの方向性を示すことが可能となると思われるからである。

### 3. 研究の方法

本研究では、グリム童話の、エーレンベルク草稿から第7版にいたるすべての作品を、Web 汎用文字コード UTF-8 のテキストデータファイル化して、独自の全作品データベースを構築したが、その際、各種の印刷版の書籍や公開Web データベースなどと照合し、正確を期した。分量的に大変な作業であったが、科研費と国内のグリムデータベース専門家の援助を受けて、なんとか全作品の汎用データベースを構築することができた。なお、汎用文字コードにはいくつかの種類があるが、現在は、Web で最も標準として使用される国際文字コードである、UTF-8 を用いたので、どのような環境でも、Web ブラウザーを使えば文字化けがなく利用できる。

今回の研究では、グリム作品に登場する数字すべての登場頻度表を作成することが処理内容であったが、いったん完成したデータベースの検索と内容処理操作は、デジタルテキストデータの特殊な処理方法を利用すれば、きわめて簡単な作業であった。これは、一般的に「正規表現」と言われるコンピュータ独自の文字コードの特性を利用した文字表現で、ドイツ語のように、同一語でも語尾が変化したり、文頭で大文字になる文字は、それぞれ別個に取り扱うと数が膨大になるが、正規表現を用いれば、たったひとつの表現で、複数の変化形も含めることができるのである。検索対象とした数字は、算用数字すべてと、さらに、「両者」「双方」「一対」などの数字に該当する表現も含めた。結果として作成した頻度表をみると、あまりにも顕著な(偏った)数字の使用(登場)方法に、まず驚かされた。

次に、この結果をもとに、グリム童話における数字の使用法における内在分析を行ったが、すべての作品とすべての版を網羅した結果は、個別作品のアナログ分析では予想できない傾向を示しており、内在分析の専門家

高橋吉文が、それをもとに、他の先行研究をも考慮しながら、独自の分析を行った。そして、最後に、石川と高橋の両人がその結果を検討し、学術論文として発表した。

#### 4. 研究成果

文学作品における数字のシンボリックな意味に関しては、これまで様々に論じられてきたが、本研究で行ったようなグリム童話すべての作品規模でのコンピュータ処理と、本格的な内在分析を合体させる試みは、初めてであると思われる。

まず、グリム童話においては、数字がグリム兄弟によって、意図的にシンボリックに使用されていること、具体的には作品の構造がシンボリックな数字に対応していることを、代表的で有名な作品 KHM55「ルンペルシュテイルツヒェン」(3の反復)とKHM21「灰かぶり」(2の反復)を使って確認した。その童話の選択理由は、コンピュータ検索の結果、2と3が、他の数字を大きく引き離して多いため、この二つの童話分析では、特に、2と3の数字が、内在理論的に、作品構造と密接に関連していることが証明できたと思われる。つまり、グリム童話における数字の内在論的な意味の確認である。

次に、コンピュータによる全作品対象の数字出現頻度表を詳細に分析すると、数字3の総体的減少と、それと反比例する、数字2の圧倒的に多数な状況、および改版につれての顕著な増大が確認された。このような規模における数字の特徴的な使用法は、グリム兄弟の数字に関する意図的な操作が、全作品の規模において行われていることを示唆していると思われる。ただ、コンピュータの処理結果は、単なる数字の頻度表に過ぎないので、この結果を作品内在分析の手法と関連づける必要がある。今回は、グリム童話全般にわたり、極端な使用数が確認された2という数字を対象として、具体的な内在分析を行うことにした。

グリム童話の中で、特に2に強く関連すると思われるのは、KHM60「二人兄弟」である。その理由は、まず、グリム兄弟自身がこの作品の重要性に言及しており、また、この作品が、作品量的に最も長く、内容も複雑で、解釈が難しいと見なされているからで、そのため、分析対象としては、もっとも適切であると思われたからである。

「二人兄弟」は、6つの部分から構成されていて、それらは、前半ステージ1,前半ステージ2,中核ステージ1,中核ステージ2,後半ステージ1,後半ステージ2に区分される。それぞれの6つの構成部分では、前半ステージと後半ステージにおいて、数字の2の使用数が優勢であり、中核ステージではその他の数字が登場し、2の出現回数を弱めている。この数字の使用法を内在的に分析すると、まず物語の発端には2という数が登場し、ついで、それが分離され、離反されていくが、

最後に再び、より次元の高い統合状態となって終わる、というように読める。

「二人兄弟」における数字2の有為転変表

構成部分	意味	数字群の全体的特長
前半 1	転落	2 がやや強く羅列ぎみ
前半 2		2 がほぼ圧倒的
中核 1	冥界	諸数が混在、2 の埋没
中核 2		諸数混在の中 2 が増加
後半 1	蘇り	2 の優位性が明確化
後半 2		2 と 1 のみが支配

それではその2という数字であるが、それは、表面的には、タイトルや内容からみて、もちろん二人の兄弟であり、彼等は様々な試練に遭い離散の危機に陥るが、困難な試練を乗り越え、最終的にはより高い内容の結びつきを獲得した2へと立ち戻っている。つまり、主人公が危機的状況に陥り、比喩的に言えば冥界へと転落するが、童話に特徴的なこの「地獄巡り」を経験した後、最後にはさらなる高次のレベルの存在としてよみがえるという構造である。数字のレベルで見ると、奈落への転落が、数字の2が減少し、他の数字が台頭するという構造で表現されているのが興味深い。後半2で、1が登場するのは、高次の統合状態を暗示している。

実は、この2という数字は、グリム兄弟や、さらにはドイツロマン派にとって、重要な意味を持つ概念である。今回内在分析の対象としては、「二人兄弟」の一編だけしか取り上げなかったが、先の全作品データベース検索の結果を見ると、グリム兄弟は、2という数字の使用を、きわめて広汎に敷衍しているように思われる。

さらに、コンピュータによる頻度表で、改版が行われるたびに、3が減少傾向を見せるが、そもそもグリム兄弟をはじめ、一般的にヨーロッパの思想史においては、特に宗教的な理由から「聖なる3」という概念が最も有名である。グリム兄弟の兄ヤーコブ・グリムも、言語学研究においてはよく3の重要性を主張していた。しかし、時代が下り、兄弟の根本思想の中に、3にかわって2の重要性に注目する傾向が現れてくるが、頻度表の結果は、この二人の思想的展開を反映していると考えることができるであろう。

このように、作品内在分析の手法を用いれば、一見難解で、解釈の難しい作品も、その隠れた意図が理解できるようになる。また、その際、コンピュータを分析のツールとして利用することにより、アナログの分析では不可能な発見も可能となるのである。グリム童話全般において、2が特に著しく登場するというのは、2に対するグリム兄弟の特別な偏

愛ともいえる傾向が、数量的にも示されていると思われる。すなわち、本研究で、2と最終的にそれを統合へと向かわせる1というグリムの数字操作の意図を抽出し、実証することができたといえる。

本研究で利用した膨大な量のテキストのデジタルデータベース作成作業には2年を超える時間が必要であったが、最終年度にはデータベースが完成し、内在研究に利用できるようになったのは、大きな成果である。本研究は、一般的に関連づけが難しいコンピュータによるデジタル処理と、人間精神が行う作品内在分析を、有効に連携させる方法の一例として、価値があると思われる。本研究をとおり、ますます増加する文学作品のデジタル化を、単なる「閲覧」だけで終わらせず、コンピュータ独自の処理対象にする研究の、実際例を提示できたように思われる。

なお、このグリム童話のデータベースは、Web 汎用文字コードによる純粋なテキストデータから作成されているので、数字だけではなく、その他の様々な項目の処理にも利用できるので、存在価値は大きいと思われる。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 1 件)

石川 克知、高橋 吉文、『グリム童話集』における数字使用 コンピュータ検索と作品内在分析の連携、メディア・コミュニケーション研究、査読有、66号、2014、1-57

## 6. 研究組織

### (1) 研究代表者

石川 克知 (ISHIKAWA, Katsutomo)  
北海道大学・メディア・コミュニケーション研究院・教授  
研究者番号：30142665

### (2) 研究分担者

高橋 吉文 (TAKAHASHI, Yoshifumi)  
北海道大学・名誉教授  
研究者番号：20091473