

# テキストマイニングによる中学生の自由記述データの分析

- 人や社会に対する志向性の測定を目指して -

小川 一美 齋藤 和志

(愛知淑徳大学コミュニケーション学部)

Key Words: テキストマイニング, 自由記述, 効果測定

## 【問題および目的】

近年、小、中、高等学校などで「心の教育」に関する様々な取り組みが行われるようになってきた。こうした取り組みを広く展開していこうとすると、その取り組みの効果を本人、教師、保護者、そして社会に提示していくことが必要になる。しかし、これらは児童や生徒の心理的側面に働きかけるものであるため、効果が表に現れにくかったり、効果が現れるまでに長い時間がかかったりするなど測定は容易ではない。小川ら(2001, 2002, 2004)も、中学生の「人間・社会を考える能力を刺激する」という目的の授業を実施し、中学生向け尺度を用いた縦断的調査からその効果を測定しようと試みた。しかし、尺度得点の変化などを統計的に分析するだけでは効果測定として不十分であった。そこで、本研究では、中学生による自由記述データの分析を取りあげ、質的データの分析による効果測定の可能性を探ることを目的とする。

ところで、従来、自由記述データの分析には KJ 法という手法が多く用いられてきたが、客観性および再現可能性という点が問題となる。そこで今回は、テキストデータを分解しその構造を数量的に解析する方法であるテキストマイニングという手法を用いて、より客観性と再現性のある分析を行うことを試みる。

## 【方法】

<調査対象者> 国立 A 大学附属中学校の 1 年生 80 名(男女各 40 名)、2 年生 78 名(男 38 名、女 40 名)、3 年生 81 名(男 41 名、女 40 名)の計 239 名。実施は 2002 年 6 月。

<質問紙の構成> 質問紙は、フェイスシート、社会的場面に対する自由記述、情緒的共感性尺度、パースペクティブテーキング(PT)尺度、社会考慮尺度、クリティカルシンキング(CT)志向性尺度、「社会」からの自由連想の順で構成されていた。なお、本研究では「社会的場面に対する自由記述」の「空きカンなどのゴミをポイ捨てしてはいけないのはなぜだと思いますか?」という質問に対する回答に焦点を当てて分析を行った。

## 【結果と考察】

### 1. テキストマイニングによる解析例

テキスト型データ解析ソフトウェア「WordMiner」(日本電子計算株式会社)を用いて、回答文章の分かち書き処理、キーワード抽出処理、そして多次元データ解析を行った。なお、分かち書き処理およびキーワード抽出処理には、Happiness/AiBASE(株式会社平和情報センター)が採用されている。有効サンプル数は 228 で、総構成要素数は 848、異なり構成要素数は 106 であった(「構成要素」とはデータ解析上の処理単位、「異なり構成要素数」とは同一表記の構成要素を 1 語と数えた時の語数を指す)。構成要素のクラスター化を行った結果、10 クラスターが抽出された(Table1, Figure1)。このクラスター化は構成要素の出現頻度の共変動をもとに行われることから、分析者の主観に頼る必要なく分類が可能となる。次に、人口統計的変数や個人特性との関連を検討した。性別×構成要素クラスター、学年×構成要素クラスター、さらに情緒的共感性、PT、社会考慮、CT の平均値をも

とに調査対象者を高群と低群の 2 群に分けて作成した、各群×構成要素クラスターの出現頻度のクロス表に基づき対応分析を行った。あまり深い考えをせず当然のことだからとか、「まわり」「一人一人」といった人に関する記述でまとまっているクラスター 1 は、情緒的共感性、PT、社会考慮の高い人に多いことが示された( $p < .001$  -  $p < .01$ )。具体的な記述が多いクラスター 2 は、女性や 1 年生、情緒的共感性が高く PT の低い人に多いことが示され( $p < .001$  -  $p < .05$ )、見た目や気持ちが悪いからなどといった記述でまとまっているクラスター 3 も女性に多かった( $p < .01$ )。老人や体の不自由な人といった自分とは異なる立場の他者や、まちや地球といった人間以外の対象に関する記述でまとまっているクラスター 4 は、PT が高い人に多かった( $p < .05$ )。そして、汚いからとか何となくといった浅い思考でまとまっているクラスター 6 は、3 年生、社会考慮や CT が低い人に多く( $p < .01$  -  $p < .05$ )、捨てられたゴミを食べる生き物への影響に触れたクラスター 7 は 1 年生に多かった( $p < .001$ )。環境や自然破壊といった広い視野からの記述でまとまっているクラスター 8 や 9 は男性に多いことが示された( $p < .01$ )。なお、クラスター 8 は、1 年生に少なく情緒的共感性の低い人にも多かった( $p < .001, p < .01$ )。さらに、クラスター 9 は CT の高い人に多いことも明らかとなった( $p < .001$ )。

### 2. テキストマイニングを利用する際の課題

テキストを構成要素に分解することが基本となるため、回答が単語レベルで記述されていると非常に解析しやすい。文章などの場合はその前後の文脈によって解釈が異なる単語も多数出現するため、恣意的にはなるが「分かち書き回避」という手順を踏む必要なども生じてくる。したがって、どのようなデータでもテキストマイニングを用いれば何らかの分析が出来るというわけではなく、どのレベルで自由記述をさせるのかなどについて調査前に十分検討することが重要になってくる。

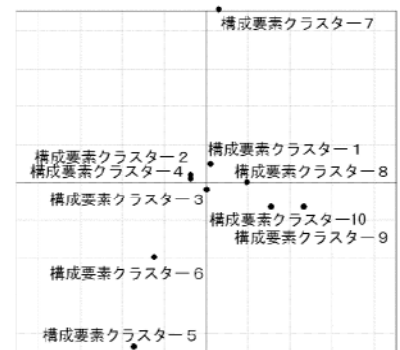


Figure1 クラスターの布置図

Table1 各クラスターの構成要素例

クラスター	サイズ	構成要素例
1	5	あたりまえ, べき, まわり, 一人一人, 考える
2	50	きれい, みんな, リサイクル, 転ぶ, 道...
3	14	ゴミ箱, 悪くなる, 汚す, 気持ち, 見た目...
4	14	けが, まち, 汚れる, 弱者, 地球...
5	2	捨わない, 臭い
6	5	もったいない, 汚い, 何となく, 見える, 残す
7	2	食べる, 生き物
8	9	ゴミ, 悪い, 汚染, 環境, 増える...
9	4	つながる, 壊す, 自然, 破壊
10	1	マナー

(OGAWA Kazumi, SAITO Kazushi)