

学生の大学生生活、授業に対する満足・ 不満点に関する自由記述の分析

木村 清*・水田 恵三**

Qualitative Analysis of Students' Description of
College Life Satisfaction and Class Satisfaction

Kiyoshi Kimura, Keizo Mizuta

要 旨

学生による大学生生活および授業に対する満足・不満点に関する自由記述文を、テキストマイニングの手法を用いて分析した。その結果、個々の自由記述文を引用することなく、抽出されたキーワードの分析結果から、ある程度の傾向を把握することができた。

キーワード 学生 授業評価 自由記述 テキストマイニング 対応分析

1. はじめに

大学のユニバーサル化が現実のものとなり、学生を一定の目的、傾向を持つ集団として捉えにくくなっている。また、学生の意識や意見も多様化しており、従来型のアンケート調査（学生生活関連や学生による授業評価など）で有益な情報が得られるとは必ずしもいえない状況になっている。

この状況は、社会における消費者のニーズの多様化と企業の対応に重ね合わせて考えることができるであろう。たとえば、インターネット上ではWeb2.0と称されるサービスの普及がめざましい。サービスを提供するサイトには利用者の行動（クリック）や発言（書き込み＝テキストデータ）に関する膨大な情報が蓄積されている。そして、それらの情報から消費者のニーズを抽出し、マーケティングに活用する例が増えている。このとき使われる探索的な分析手法としてテキスト

マイニングと呼ばれる手法が注目されている^{1) 2)}。

そこで筆者らは、学生の意識やニーズを探る際にも、自由記述文をテキストマイニングの手法で分析することで、従来型のアンケート調査では得られにくい情報が得られるのではないかと考えた。

大学生生活や授業に関連して学生の自由記述文を分析した例として西川和夫他（2004）³⁾、井上 仁他（2007）⁴⁾などが散見されるが、テキストマイニングの適用例が充分にあるとはいえない。Webサイトなどで用いられる実際のテキストマイニングのシステムは非常に高価であり、幅広い研究分野で気軽に使えるソフトウェアが少ないことも一因であろう。

本研究では以下のように、比較的安価に手に入るツールを組み合わせることで分析を行うことにした。

テキストマイニングの要素技術の一つに、日本語文章を品詞に分解する形態素解析があ

* 総合人間科学部 現代社会学科

** 総合人間科学部 人間心理学科

る。最近では一般的なパソコンでも形態素解析ができる無償のソフトウェアがいくつか公開されている。本研究ではKH Coder⁵⁾を使用した。KH Coderは、形態素解析エンジンに「茶筌」⁶⁾を使用し、日本語文章を品詞に分解し、単語の頻度集計やコンコダンス検索といった簡単なテキストマイニングが行えるものである。

一方、一般的なテキストマイニングを手軽にできるソフトウェアも商品化されている。そのうち本研究では日本電子計算のWordMinerを使用した。WordMinerでは不定形文から抽出された構成要素変数を基に、クラスター分析や対応分析を探索的に行うことができる。

本研究では後述のようにテキストマイニングの適用可能性についての知見を得る目論見もあり、敢えて事前のカテゴリー化をしないようにして、学生の自由記述文を採取する方針を取った。

本稿に先立つ予備的調査では、まったくカテゴリー化していない回答用紙に、授業に関する満足・不満点を自由記述させた。しかし、たとえば「理解できる」と「理解できない」という記述があった場合に、理解(名詞)、できる(動詞)、ない(助動詞)に分解されるが、単純な処理ではポジティブな記述なのかネガティブな記述なのかの区別がつきにくく、満足な分析ができないことが分かった。

そこで最低限のカテゴリーとして、大学生活について、授業について、それぞれ満足(Positive)と不満(Negative)の2つのカテゴリーに分け、自由記述を求めた。その上で、キーワードの頻度分析や対応分析を行った。その結果、個々の記述を具体的に引用しなくとも、ある程度の傾向や特徴を見いだすことができたので、本稿ではそれを報告する。

以下調査の目的、方法を述べた後、今回行った集計・分析の方法と結果を記し、そこから得られたことがらについて述べる。最後にテ

キストマイニングの適用に関して今回得られた知見をまとめておく。

2. 目的

本調査では以下の2点を目的とした。

- 1) 従来的なアンケート調査のようにあらかじめ選択肢を設けることなく、自由に記述させた内容を分析し、学生の観点を探る。
- 2) 分析に際してはテキストマイニングの手法を適用しその手法の適用についての知見を得る。

3. 調査時期・方法

調査時期：2007年7月、授業終了時に調査票を配布、その場で回答させた。

設問項目：以下の4つのカテゴリーそれぞれに対し自由記述で回答させた。

| カテゴリー | | 図中での表記 | |
|-------|-------|--------|-----|
| 1 | 大学生生活 | 満足な点 | 生活P |
| 2 | | 不満な点 | 生活N |
| 3 | 授 業 | 満足な点 | 授業P |
| 4 | | 不満な点 | 授業N |

回答数：128

4. 集計・分析方法

4.1 前処理

回収した自由記述部分を1文ごとに分け、それぞれに回答者の学科、性別、記述内容のカテゴリー情報を付して1レコードとした。

次に、回答の表記上の誤り、入力ミス、語のゆらぎについては、以下のようにして入力データを直接修正した。

この作業には、KH Coderを使い、複合語や固有名詞、キーワードとして抽出すべきものについてはそれらを指定して、語として抽出するように設定した。また、品詞別の抽出

結果を観察し、表記のゆらぎその他について検討し、必要に応じて入力データを修正した。

また、この作業と並行して、意味的に一つのキーワードにまとめるべきものについてWordMinerでの置換編集作業用のデータを作成しておいた。

4.2 頻出語の集計

カテゴリー（大学生生活満足、大学生生活不満、授業満足、授業不満）ごとに頻出語を集計した。集計にはKH Coderを使用した。

4.3 対応分析

キーワード同士、およびカテゴリーとの対応関係を観察するための対応分析をWordMinerを用いて行った。この際、以下の2つの方法を用いた。

(1) 方法1

KH Coderで得られたキーワードのうち、意味的にまとめられるものをWordMinerの構成要素変数の置換編集機能を用いてまとめた。たとえば、「カフェテリア」「学食」はどちらも「食堂」に置換した。また、対応分析の対象とする語（構成要素変数）を、頻度が3以上のものに限定してから分析を行った。

(2) 方法2

方法1よりも更に意味的なまとめを進めるために、KH Coderのコーディング機能を用いた。

KH Coderのコーディング機能を使うと、たとえば、一つのデータ（文）が以下のいずれかの条件に当てはまるときは、「通学不便」というコード（キーワード）を与えることができる。

- ・「バス」と「少ない」を含む
- ・「バス」と「ない」を含む
- ・「バス」または「交通」または「通学」を含み、かつ「不便」を含む
- ・「学校」または「家」を含み、かつ「遠い」を含む

今回はコーディングにより68個のコードを付与した。

5. 結果

表1に、性別まで加えたカテゴリーと、各カテゴリーの単純集計結果を示した。ここで1サンプルあたりの平均構成要素数は、一つの文の中から抽出された語（構成要素）の個数の平均値である。大学生生活、授業両者とも、満足よりも不満カテゴリーの1サンプルあたりの平均構成要素数が多いのは、「ない」「にくい」など否定を表す語が付加されている文が多いためと思われる。

5.1 頻出語の集計

表2にカテゴリーごとの頻出語を度数順に示した。

表1 カテゴリーと構成要素数

| カテゴリー | | | サンプル数 | 構成要素数 | 1サンプルあたりの平均構成要素数 |
|-------|----|---|-------|-------|------------------|
| 大学生生活 | 満足 | 男 | 66 | 215 | 3.26 |
| | | 女 | 93 | 288 | 3.10 |
| | 不満 | 男 | 78 | 303 | 3.88 |
| | | 女 | 143 | 561 | 3.92 |
| 授 業 | 満足 | 男 | 48 | 209 | 4.35 |
| | | 女 | 61 | 254 | 4.16 |
| | 不満 | 男 | 72 | 392 | 5.44 |
| | | 女 | 88 | 436 | 4.95 |

表 2 カテゴリー別頻出語度数

| | 大学生生活 | | | | 授 業 | | | |
|----|-------|----|------|----|-----|----|------|----|
| | 満足 | | 不満 | | 満足 | | 不満 | |
| | 抽出語 | 度数 | 抽出語 | 度数 | 抽出語 | 度数 | 抽出語 | 度数 |
| 1 | 楽しい | 23 | 少ない | 45 | 授業 | 32 | ない | 53 |
| 2 | 生協 | 19 | バス | 34 | ある | 26 | 先生 | 38 |
| 3 | 友達 | 18 | する | 33 | する | 23 | ある | 36 |
| 4 | できる | 16 | ない | 33 | やすい | 21 | 授業 | 36 |
| 5 | ある | 15 | ない | 30 | 楽しい | 18 | する | 31 |
| 6 | する | 15 | 欲しい | 19 | 分かる | 14 | いる | 24 |
| 7 | 自由 | 13 | 遠い | 15 | 先生 | 13 | 分かる | 21 |
| 8 | 良い | 13 | 本数 | 15 | 面白い | 11 | にくい | 19 |
| 9 | 多い | 12 | ある | 12 | 興味 | 9 | 多い | 16 |
| 10 | 授業 | 10 | 学食 | 12 | できる | 7 | 欲しい | 16 |
| 11 | 食堂 | 10 | 狭い | 12 | 満足 | 7 | 黒板 | 14 |
| 12 | いる | 9 | サークル | 9 | ビデオ | 6 | 字 | 14 |
| 13 | 時間 | 9 | 多い | 9 | 自分 | 6 | 書く | 13 |
| 14 | きれいな | 8 | できる | 8 | 多い | 6 | 板書 | 11 |
| 15 | 人 | 8 | なる | 8 | 内容 | 6 | ない | 10 |
| 16 | 美味しい | 8 | 学校 | 8 | 分野 | 6 | 汚い | 10 |
| 17 | ない | 7 | 食堂 | 8 | 良い | 6 | うるさい | 9 |
| 18 | 学食 | 7 | 図書館 | 8 | いる | 5 | もっと | 8 |
| 19 | 充実 | 7 | 不便 | 8 | テスト | 5 | あまり | 7 |
| 20 | 増える | 7 | もっと | 7 | 学べる | 5 | 長い | 7 |
| 21 | たくさん | 6 | 高い | 7 | 使う | 5 | できる | 6 |
| 22 | 広い | 6 | 小さい | 6 | とても | 4 | 教科 | 6 |
| 23 | 図書館 | 6 | 先生 | 6 | ない | 4 | 取る | 6 |
| 24 | サークル | 5 | 土足禁止 | 6 | 科目 | 4 | 授業中 | 6 |
| 25 | 校舎 | 5 | クーラー | 5 | 広い | 4 | 進む | 6 |

大学生生活については、“食堂”および“学食”の度数が「満足」、「不満」の両方で高い。“図書館”についても「満足」、「不満」両者にリストアップされている。その一方で“生協”は「満足」のカテゴリーにのみあり、しかも非常に頻度が高くなっている。“学食”、“図書館”については学生個人の感覚や注目点により、満足な部分と不満な部分が指摘されている一方で、“生協”（売店）については満足度の評価が非常に高いことが伺える。

また、大学生生活についての記述でありながら“授業”という語が多く使われている。実際、“授業”という語の周辺を観察すると、「授業が楽しい」（4）、「授業時間の組み合わせ、時間割り関連」（4）という内容の記述が多かった（括弧内数字は度数）。大学生生活についての不満カテゴリーでは特に“バス”、“本数”、“遠い”、“少ない”、“不便”といった通

学の不便さに対する不満が非常に多いことがこれらの語の頻度の高さから伺われる。

「授業」カテゴリーについては、当然のことながら「満足」・「不満」両者に、“授業”、“先生”が高い頻度で現れている。「満足」カテゴリーに、“分かる”、“やすい”が、「不満」カテゴリーには“分かる”、“にくい”、“ない”が高い頻度で現れている。このことから、授業内容の分かりやすさ、分かりにくさが満足・不満の主な要因となることが推測される。さらに「不満」カテゴリーでは、“黒板”、“字”、“書く”、“板書”、“汚い”がいずれも高い頻度で現れている。

5.2 対応分析^{注)} (方法 1)

ここでは分析に先立ち、KH Coder で抽出されたキーワードのうち、意味的に同じものをまとめた上、3回以上出現したものを対応

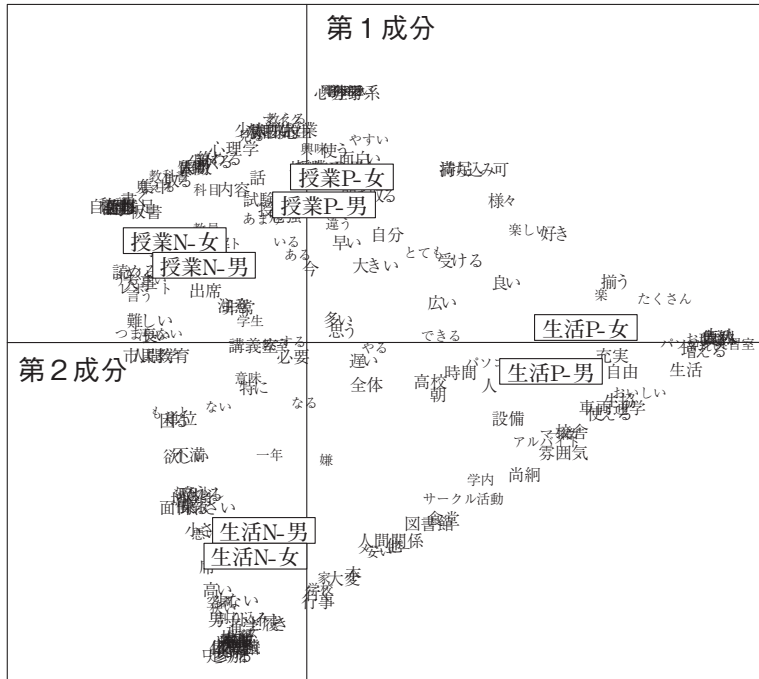


図1 キーワードとカテゴリの布置図

表3 布置図の頂点付近のキーワード

| | 横軸（第2成分） | | 横軸（第1成分） | |
|----|----------|---------|----------|-------|
| | 負 | 正 | 負 | 正 |
| | 授業不満 | 生活満足 | 生活不満 | 授業満足 |
| 1 | レポート | 友人 | 人間関係 | 興味深い |
| 2 | 板書 | 自然 | 他 | ビデオ |
| 3 | 言う | 便利 | メニュー | 分野 |
| 4 | うるさい | おばさん | 安い | 学べる |
| 5 | にくい | 環境 | 席 | 学ぶ |
| 6 | 大事 | パソコン実習室 | 本 | 心理学系 |
| 7 | 書く | きれい | 大変 | 教える |
| 8 | 読める | 増える | 家 | マイク |
| 9 | 字 | 生活 | 高い | 少人数授業 |
| 10 | 進む | たくさん | 学校 | たまに |
| 11 | 少し | おいしい | 行く | M先生 |
| 12 | 下手 | 自由 | 少ない | 詳しい |
| 13 | 自己満足 | 生協 | 行事 | 専門的 |
| 14 | 説明 | 揃う | 空調 | 見る |
| 15 | 必修 | 充実 | 狭い | やすい |

分析の対象キーワードとした。

対象キーワードとカテゴリの布置図を図1に示した（X軸：第2成分、Y軸：第1成分）。X軸を第2成分、Y軸を第1成分としたのは、X、Yそれぞれ第1成分、第2成分

とするよりもカテゴリの配置が見やすくなるためである。

この図において、周辺部分（特に4つの頂点付近）は、その付近に布置されたカテゴリに特有の語を意味する。頻度が高くなくとも

表 4 コーディングの例

| 与えるコード (キーワード) | 条件 |
|-------------------|--------------------------------------------------------------------------------------------|
| 要望 | ほしい or 欲しい |
| 理解 | 分かる or 理解 |
| 空調 | エアコン or クーラ or クーラー or 暖房 or 暖房器 or 暖房機 or 暖房装置 or 冷房 or 冷房装置 or 暖房器具 or 冷房器具 or 冷暖房 or 空調 |
| 混雑 | 混む or 定員オーバー or 混雑 |
| 困難 | にくい or 困難 |
| 見にくい | (見える and *困難) or (読む and *困難) or (字 and 汚い) or (読む and 小さい) or (読める and ない) |

離れたところに配置されるため、そのカテゴリーを代表する語ではない。(異常値と見なせる場合もある)

参考までに表 3 に布置図の 4 つの頂点付近の語を抜き出した表を示す。

また、“なる”、“思う”、“する”、“必要”、“多い”、“やる”などは中心付近に配置されているが、これは、どのカテゴリーにも出現している語であることを意味する。

性別まで含めた 8 つのカテゴリー間の距離を見ると、いずれも男女の距離は小さいことがわかる。

次に、4 つのカテゴリー間の距離を見ると、授業に対する満足・不満は比較的近い場所に配置されているが、大学生活に対する満足・不満は互いにやや離れて配置されている。この布置図は頻度情報に基づいたクロス表の多変量解析であることから、授業に対する満足・不満は両者に同じような言葉が使われる傾向がある、つまり、評価の観点が似ていると考えられる。それと比較すると、大学生活に対する満足・不満の記述では、評価の観点が多様であることが推測できる。

5.3 対応分析 (方法 2)

KH Coder のコーディングの機能を利用し、さらに踏み込んで意味的要素を加味したキーワード付け (コードの付与) を行ってから、対応分析を試みた。表 4 にコーディング

の例を示した。

布置図を図 2 に示す (X 軸: 第 2 成分、Y 軸: 第 1 成分)。X 軸を第 2 成分、Y 軸を第 1 成分としたのは、X、Y それぞれ第 1 成分、第 2 成分とするよりもカテゴリーの配置が見やすくなるためである。

5.4 カテゴリー間の有意差検定

コーディング後のキーワードについて、各カテゴリーに特徴的に使用されているものにどんなものがあるのかを調べるため、WordMiner を使い、頻度による有意性テストを行った。有意確率が 5 % 未満のものについてまとめたものを表 5 に示す。これから以下のことがわかる。

大学生生活の満足な点については、男子が食堂や生協の施設面、利便性について、またサークル活動といった活動面を評価するのが多いのに対し、女子では友人関係の充実を指摘する傾向が見られる。

大学生生活の不満な点については、男子が記述したものが通学 (家からの距離、バスの便) に関連する不満に集中しているのに対し、女子ではそれ以外の多様な対象への不満が見られる。(上履きに対する不満、冷暖房に対する不満など)

授業に対する満足点については男女で大きな違いは見られない。一方で不満な点については、男女とも板書 (字が見にくい、汚い)

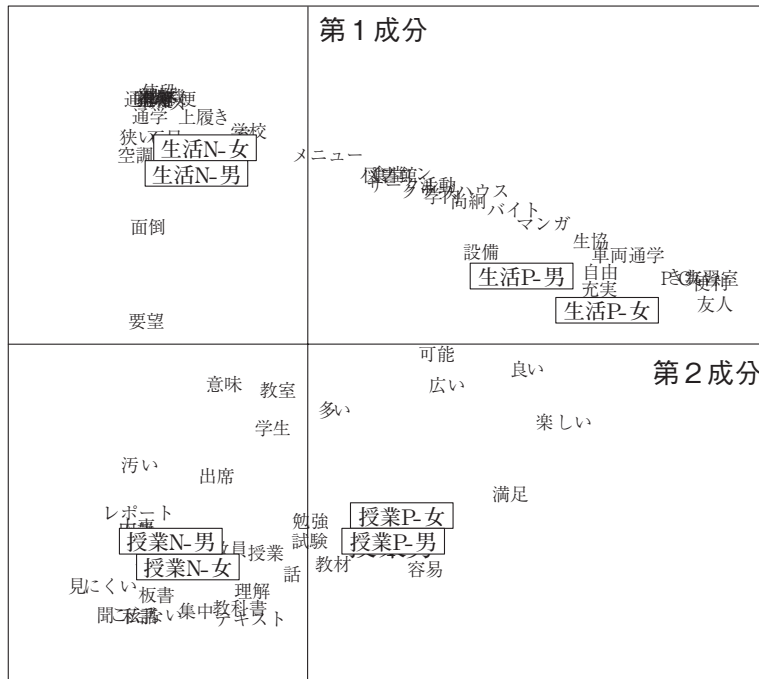


図2 コーディング後の対応分析結果

表5 頻度による有意性テスト結果

| カテゴリー | 大学生生活 | | | | 授業 | | | |
|---------|--------|------|-----|------|-----|-----|------|------|
| | 満足 | | 不満 | | 満足 | | 不満 | |
| | 男 | 女 | 男 | 女 | 男 | 女 | 男 | 女 |
| 特殊キーワード | 食堂 | 友人 | 不足 | 不足 | 容易 | 容易 | 板書 | 教員 |
| | 自由 | 生協 | バス | 通学不便 | 面白い | 教材 | 見にくい | 板書 |
| | サークル活動 | 楽しい | バス | 通学不便 | 理解 | 楽しい | 教員 | 理解 |
| | 良い | 便利 | 通学 | 空調 | 勉強 | 試験 | 教員 | 困難 |
| | 尚綱 | きれいな | 家 | 遠い | 満足 | 理解 | 汚い | 授業 |
| | 生協 | 自由 | 遠い | 上履き | | 満足 | レポート | 私語 |
| | きれいな | 可能 | | 食堂 | | 授業 | 大事 | 見にくい |
| | 広い | 充実 | | 混雑 | | 面白い | 要望 | 集中 |
| | 充実 | | | 狭い | | 良い | 話 | |
| | | | | 要望 | | | | |
| | | | 不便 | | | | | |
| | | | 自販機 | | | | | |

に対する不満、要望が非常に多い。女子ではそれに加え授業中の私語に対する不満も多いことが分かる。

以上主な分析結果について述べてきたが、それらを総合すると以下のことがいえる。

- ・ 大学生生活の満足な点：自由な雰囲気楽しく、校舎がきれいであることに満足してい

る。生協や生協の店員の方に親しみを感じている。男子ではサークル活動、女子は友人関係にポイントを置く傾向が見られる

- ・ 大学生生活の不満な点：大学が遠くにありなかつ通学（バス）の便が悪いことに不満を感じている。校舎に空調の設備が整っておらず、夏は暑く冬は寒い、校舎内は土足

禁止（上履き）であることにも不満であるが、ルールを守らない学生がいることも不満に感じている。男子は特に通学に対する不満に集中しているが、女子は男子よりも学生生活全般に関係する多様で細かい観点での不満を述べている

- ・ 授業に対する満足な点：面白く分かりやすい授業に対しては満足度が高い。女子のほうが男子よりも具体的な満足点を記述している傾向が見られる
- ・ 授業に対する不満な点：板書が見にくい、内容が理解しにくい授業に不満を感じている。私語に対する不満の指摘は女子のほうにより多く見られる。

6. まとめ

あらかじめ大学生活に対する満足な点・不満な点、授業に対する満足・不満な点の、4つのカテゴリーに書かれた自由記述文をもとに、キーワードの抽出、コーディングを行い、頻度の集計、キーワードの対応分析を行った。その結果、個々の自由記述文を引用することなく、抽出されたキーワードの分析結果からある程度の傾向を把握することができた。

それは例えば、不満な点に限定すれば、学校生活のハード面に関しては、立地を含む校舎や教室の温度管理に関する不満、ソフト面では、規則を守らない人たちに対する不満である。また、授業に関しては、板書の見にくさや、私語への不満などである。

今回の分析を通してテキストマイニングの手法の具体については次のような知見を得ることができた。

- ・ 自由記述文から統一的な処理でキーワードを抽出し数量化を行い分析することで、ある程度客観性の高い分析が行える。
- ・ キーワードの抽出の善し悪しはその後の分析結果を左右する。たとえば WordMiner の分かち書き処理では、活用形のまま構成

要素変数として扱われてしまう。キーワード抽出については WordMiner とは独立に形態素解析ソフトを使うことができるため、今回は KH Coder でキーワード抽出を行った。いずれの場合も、同じ語の表記が違っていると別のキーワードとして抽出されてしまう。あらかじめ表記を訂正しておくか、コーディング（KH Coder の場合）を行うか、置換（WordMiner の場合）を行うかなどする必要がある。

- ・ 対応分析では対象とするキーワードの数をある程度絞ったほうがよい。そのため今回は KH Coder のコーディング機能を用いて、意味のまとまりごとにキーワードを付与（コード化）した。その結果コーディングをしない場合よりも有効な分析結果が得られた。しかし、コーディングの規則を作る際に分析者の意図が入らざるを得ない。この点で、テキストマイニングのそもそもの目的、すなわちプリコーディングされた調査票から得られない情報を得るということと矛盾しかねない。調査の目的が、統計的な傾向を把握したり、分類を行うのか、予期しなかった発見を期待するのかによって、分析の仕方も変わってくる。

注) 対応分析は、多数のカテゴリー形式の項目データの相互関連を分析する手法で、数量化 III 類や双対尺度法とも呼ばれている。

参考文献

- 1) 三室 克哉：“顧客の声マネジメント-テキストマイニングで本音を「見る」” オーム社 (2007)
- 2) 藤井美和・小杉孝司・李政元 編著：“福祉・心理・看護のテキストマイニング入門” 中央法規出版 (2005)
- 3) 西川和夫, 廣岡秀一, 中谷素之, 後藤淳子, 矢神祥代 学生は大学の授業をどう認識し、何を求めているのか：自由記述にみる率直な意見の分析 三重大学教育学部附属教育実践総合センター紀要 Vol. 24 pp39-47 (2004)
- 4) 井上 仁, 中野俊也, 白石義光, 高塚人志, 河合康明 鳥取大学医学部における学生による授業評価：3年

間の得点推移とフリーコメントの分析 医学教育
Vol.38 pp39-46 (2007)

- 5) KH Coder Index Page: (2008.9 現在)
<http://khc.sourceforge.net/>
- 6) 松本裕治 形態素解析システム「茶筌」情報処理
Vol.41 pp.1208-1214 (2000)

本研究は尚絅学院大学 2007 年度共同研究費の助成を受けて行われたものである。

