

テキストマイニングを利用した大学生の意識調査

Text Mining for the Attitude Survey of the University Students

河野康成（立教大学リーダーシップ研究所）

1. はじめに

大学生の意識については、様々な分野で調査が行われており、質問紙によってデータが取られていることが多い。さらに、一般的な意識調査と異なり、大学生の場合は、Web 調査特有の対象者偏向問題の影響が少ないため、Web による意識調査も増加傾向にある。

また、質問紙調査では、選択肢項目の分析が主体となっており、自由回答は補足的に用いられてきた。しかしながら、ソフトウェア技術の飛躍的な向上により、自由回答の分析も容易になり、徐々に分析が進められている。

本研究の目的は、選択肢項目と自由回答のそれぞれの長所短所を考慮し、自由回答の使い方とその分析手法について考察することにある。今回は、大学生の満足度に関する質問紙調査の分析（河野・小木, 2006a, 2006b）と授業要望に関する自由回答の分析を例として取上げる。

2. 分析

2.1. 大学生の満足度調査

まず、大学システムに関する質問項目（5段階評価）および自由回答の関連について分析について考察する。被調査者の内訳は、有効回答 131 名（留学生 61 名・日本人学生 64 名・帰国子女 6 名）、無効回答 2 名（日本語の質問紙が読めない理由で返却）であった。有効回答のうち 16 の欠損データを除いた 115 のデータに対して分析を行った。大学システムに対して、5段階評価の選択式質問（1. 大変満足、2. やや満足、3. 普通、4. やや不満、5. 大変不満）を行った結果が図 1 である。これに対して、自由回答のテキスト分析を行った結果が、図 2 である。

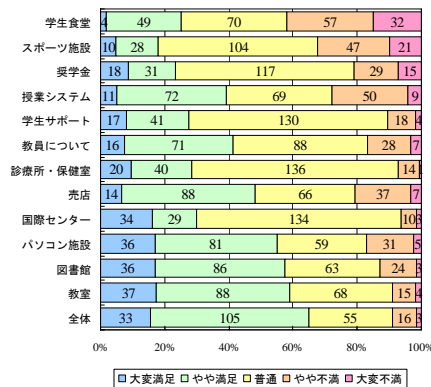


図 1. 大学に対する不満足度（選択肢）

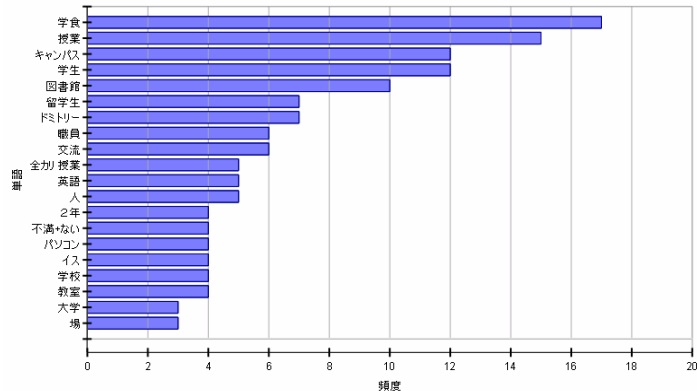


図 2. 大学に対する不満足度（自由回答）

これを見ると、選択肢と自由回答の両者共、「学食」が不満項目のトップであることは共通している。ここで注目すべきことは、選択肢で不満度が高かった「スポーツ施設」、「診療所・保健室」、「奨学金」は、自由回答に出てきていない点にある。選択肢で不満度が高かった項目が自由回答に出てこなかった理由のひとつには、最も不満が高い「学食」のような項目は自由回答にも回答されるが、2 番目、3 番目となると回答されない可能性がある。

2.2. 授業要望に関する自由回答分析

次に、授業要望に関して、都内の 4 つの私立大学で授業に関する事前調査のうち、自由回答のテキスト分析事例を紹介する。

第一の例として、社会心理学調査実習の授業（受講生 22 名）において、「具体的にはどのようなこと

を調査したいと思いますか。」という質問に対する自由回答の単語頻度解析を行った（図3）。

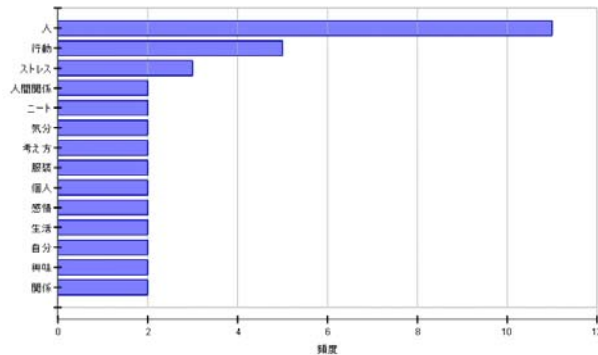


図 3. 大学に対する不満足度（自由回答）

この例では、受講生が心理学を専攻している学生であるため、対人行動を調査したいという意識が見受けられる。データ数が少ないため、具体的なところまでは把握できないが、学生の意識ははっきりと示されている。

第二の例として、情報処理初級の授業（受講生は3時限で計70名）で、「Excelで扱ってみたい項目と内容」に関する自由回答の単語頻度解析を行った結果が、図4と図5である。

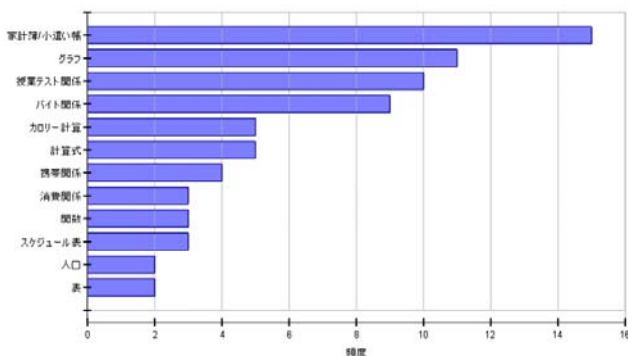


図 4. Excelで扱ってみたい項目

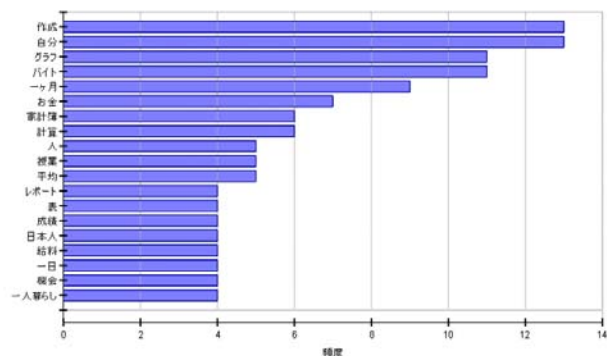


図 5. Excelで扱ってみたい内容

この例では、図4の項目については明確に回答結果が出ているが、内容については、Excelで何ができるかがよくわからないこともあり、身の回りに関することは推定できるがはっきりとは解釈しにくい結果となっている。

3. まとめ

今回の分析で、選択肢項目のとりこぼしや、不満や希望などにおける強い意識については、自由回答のテキスト分析が有効であることがわかった。しかしながら、選択肢と比較すると、欠損が多くなったり、回答結果が不明瞭になったりしやすいため、欠損にならないような対処や解釈を容易にする分析手法を考案していく必要がある。また、日本語の場合、文字に付随したコンテキストを反映できるような対応も今後の課題である。

引用文献

河野康成・小木しのぶ (2006a). 多言語を含む自由回答のテキスト分析. 日本計算機統計学会第20回大会論文集, 191-194.
 河野康成・小木しのぶ (2006b). 大学に対する満足度調査の質的・量的分析. 日本行動計量学会第34回大会発表論文抄録集, 82-83.