

統計解析ソフト R によるテキストマイニングの学校現場での応用可能性

－授業改善のためのアクション・リサーチ報告書の分析を例に

Using the Statistical Software R for Text Mining in School Settings

- A trial analysis of a potential new tool with English class action research reports

村越 亮治

MURAKOSHI Ryoji

1. はじめに

現在学校現場では、学校評価(文部科学省, 2007)の制度や学校評議員(文部科学省, 2000)の設置に伴い、生徒や保護者を対象とした、授業アンケートや学校満足度アンケートなどを実施することが普通になっている。また、授業改善の意識が高い教師は、自分の担当クラスで実施しているだろうし、教科全体としてアンケートをまとめている場合もあるだろう。特に、研究指定を受けたような場合は、報告事項の一つとしてアンケート調査の結果は欠かせないものである。言うまでもなくアンケート調査は多くの場合質問紙を用いて行われ、その質問の多くは、多肢選択式のものや、程度・段階を答えるリカート・スケールによって回答するものであるが、ほとんどの質問紙では、回答者の自由な意見を引き出すための自由記述欄が設けられる。選択式の回答については、それぞれの選択肢のパーセンテージを示したり、グラフによって視覚化したりするなどの処理がなされ、さらには、あまり学校現場で行われることはないかもしれないが、カイ二乗検定やノンパラメトリック検定、相関係数の算出などの統計分析が可能な場合もある。それに対して、自由記述はどのように処理されているだろうか。ドルニエイ(2006)では、記述のなかのキーワードをまとめてグループ化し、数値処理することなどが提案されている。加藤他(2007)にも、同様に、記述内容を見ながらカテゴリー化して変数にする「アフターコード」という方法が示されている。ある程度焦点化した質問ならば、回答をグループ化・カテゴリー化できるかもしれないが、「最後に授業の感想・要望などを書いてください」というような、まったくのフリーコメントの場合は難しい。しかしこのような自由記述のなかにこそ、選択式の質問では出しきれない、回答者の本音が隠れていることがある。

本稿では、テキストデータから情報を掘り起こす方法の一つとして、WEB 上に無料で公開されているコンピュータ・ソフトウェアを使ったテキストマイニングを取り上げる。本稿の主たる目的は、分析した結果から何かしらの知見を示すことではなく、石田(2008)の手法に基づき、現場教師による授業改善報告書のテキストデータを使って、準備から分析までの手順の一例を示し、学校現場、特に学校で行われているアンケート調査への応用可能性について考察することである。

2. 学校現場で実施されているアンケートとその処理方法の実際

2.1 実施されているアンケートと自由記述欄

公立高等学校の英語科教員 33 名に、校内で行われているアンケート調査について、その種類や体裁などをたずねたところ、すべての学校で、学校全体として生徒対象の授業アンケートが実施されており、そのなかの 29 校 (87.9%) が質問紙に自由記述欄を設けていることがわかった。それに加えて、英語科として別の生徒向けアンケートも実施しているという学校は 12 校 (36.4%) で、そのうちの 10 校 (83.3%) が自由記述欄を設定していた。また、保護者対象の学校満足度アンケートについては 12 校 (36.4%) で実施されており、自由記述欄を設けている学校はそのうちの 9 校 (75.0%) であった。

2.2 自由記述のデータ処理方法

自由記述欄のコメントの処理方法について、「主なコメントを列挙する」「主なものを肯定的な意見・否定的な意見に分けて示す」「キーワードを列挙する」「キーワードの頻度などを分析する」「特に何もしない」「その他」の 6 つの選択肢でたずねた。自由記述欄を設けているすべてのアンケート 49 件 (学校アンケート、教科アンケート、保護者アンケート) についての回答を見てみると、「主なコメントを列挙する」が一番多く、29 件 (59.2%) であった。次に多かったものは「特に何もしない (11 件:22.4%)」で、「その他 (7 件:14.3%)」「肯定的な意見・否定的な意見に分ける (2 件:4.1%)」が続く。「その他」の内訳は、「すべてのコメントを列挙する (4 件:全体の 8.2%)」「扱いは個人に任せる (2 件)」「不明 (1 件)」であった。

3. 分析対象データ

テキストマイニングの対象は、『平成 23 年度英語教育アドヴァンスト研修授業改善プロジェクト報告書』(神奈川県立国際言語文化アカデミア, 2012) である。これは、現職の県立高等学校英語教員研修のなかで、個々の教師がそれぞれの授業の課題解決に取り組んだ、アクション・リサーチの報告書である。英語授業改善のためのアクション・リサーチについては、佐野 (2000; 2005) に詳しいが、おおむね、次のような手順を踏むのが普通である。

- 1) 問題の発見: 直面している事態から扱う問題を発見する。
- 2) 事前調査: 選んだ問題点に関する実態を調査する。
- 3) リサーチ・クエスチョンの設定: 調査結果から研究を方向づける
- 4) 仮説の設定: 方向性に沿って、具体的な問題解決の対策を立てる。
- 5) 計画の実践: 対策を実践し、経過を記録する。
- 6) 結果の検証: 対策の効果を検証し、必要なら対策を変更する。
- 7) 報告: 実践を振り返り、一応の結論を出して報告する。佐野 (2005)

いわゆる実証研究とは異なり、それぞれの実践者がそれぞれの状況のなかで、省察的に実践を繰り返す。

返しながら、みずからの授業を改善し、目の前の生徒の力を伸ばすことを目的としており、必ずしも理論の一般化を目指すものではない。したがって、その『報告書』には、実践者の意識や価値観が率直に表れやすいと考えられる。本稿で注目したいのは、『報告書』の記述に表れる、このような実践者の「思い」である。実践者のリサーチ・テーマは「学習意欲」「語彙」「リーディング」「ライティング」「教室での英語使用」の5つに大別することができ、報告書の共通する項目立ては、「クラスの特徴」「解決すべき課題」「事前の現状把握」「改善の目標」「改善のための手立て」「生徒の変化」「教師の変化」「今後の課題」「まとめ・感想」となっている。

『報告書』のなかの「クラスの特徴」から「改善の目標」までの部分は、アクション・リサーチの「問題の発見」から「リサーチ・クエスチョンの設定」の段階にあたり、特にここまでの記述には、それぞれのリサーチ・テーマに関する実践者の主観や意識が表現されていると思われる。一方、「改善のための手立て」「生徒の変化」(アクション・リサーチの「仮説の設定」から「結果の検証」の段階)については、事実に基づく分析、数値的データなどが示され、比較的客観的記述が多いと推測される。それに対して、「教師の変化」から「まとめ・感想」までの部分には、実践者の振り返りのなかに、それぞれの状況での授業観が表れやすいのではないかと予想される。

4. 基本的な分析の手順:形態素の頻度比較

4.1 ソフトウェアの準備

まず、無料の統計分析ソフトである「R」をインストールする(本稿執筆時点での最新版は R version 2.15.2 である)。R では、起動後に、必要に応じて機能拡張のために「パッケージ」を読み込む必要がある。次に、日本語形態素分析器である「MeCab(和布蕪)」をインストールし、さらに、石田基弘氏が開発した、これらの橋渡しをするパッケージである「RMeCab」をダウンロードして、R からインストールする。MeCab、RMeCab も無料で公開されている。山田他(2008)には、初心者にも比較的わかりやすく、R の基本操作・統計分析の手法が書かれている。

4.2 データの準備

スプレッドシートに、それぞれの実践者の ID、生徒の英語能力(基礎レベル:1, 標準レベル:2, 発展レベル:3)、リサーチ・テーマ(学習意欲:M, 語彙:V, リーディング:R, ライティング:W, 英語使用:E)、「クラスの特徴」から「改善の目標」までの『報告書』の記述(これをまとめて“THEME”とする)、「教師の変化」から「まとめ・感想」までの記述(まとめて“REFLECTION”とする)をそれぞれ入力する。生徒の英語能力については、実践者の自己申告に基づいている。シートは R で読み込むために、csv ファイル(カンマ区切りのテキストファイル)で保存しておく。

4.3 ソフトウェアの起動～形態素分析

テキストマイニングの基本的な手順を示すための試行として、『報告書』の THEME の部分にどのような語がどれくらい特徴的に記述されているかを、リサーチ・テーマ別に掘り出して比較してみることにする。

R を立ち上げると、設定にもよるが、図1のような画面があらわれる。次に日本語形態素分析器 MeCab の利用を可能にするために、「パッケージの読み込み」から RMeCab を読み込む。

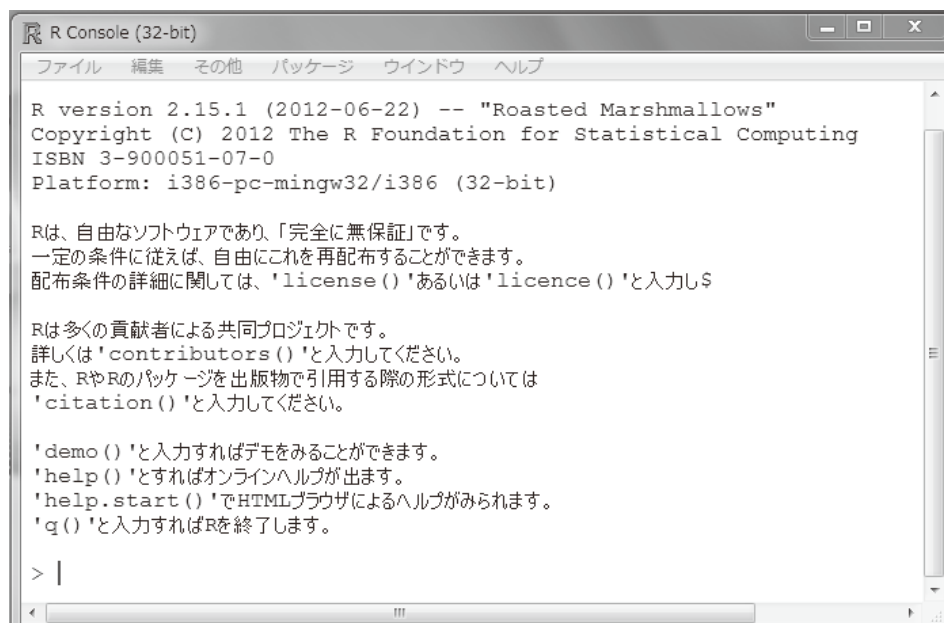


図 1 R の初期画面

次に、4. 2 で準備した csv ファイルを読み込み、名前をつけた変数=オブジェクトに代入する。このデータは表のように並んだデータフレームになっているので、RMeCabDF () という関数を使って形態素分析を行う。RMeCabDF の引数は、順に csv ファイルを代入したオブジェクト名 “theme”、分析したいテキストデータが入っている列番号 “4”、形態素の原形を返すための番号 “1” とし、その結果を新たに “theme1” と名づけたオブジェクトに代入する。さらに、内容語のなかでも実践者の意識を特徴づける可能性のある品詞として、名詞と形容詞のみを取り出し、別のオブジェクト “theme2” に収める(図 2)。

```
theme<-read.csv("AP_Report2011.csv")
# csv ファイルの読み込み
theme1<-RMeCabDF(theme,4,1)
# データフレームを MeCab に引き渡し形態素解析
theme2<-list()
for (i in 1:length(theme1)){theme2[[i]]<-theme1[[i]][names(theme1[[i]])=="名詞"
| names(theme1[[i]])=="形容詞"]}
# 名詞・形容詞のみ取り出し
```

図 2

さらに、ここで得られた、元のファイルの4列目に入っていたすべてのテキストデータ THEME に含ま

れるすべての名詞と形容詞を、「学習意欲」「語彙」「リーディング」「ライティング」「英語使用」の研究・テーマ別に取り出す。テーマを区別する記号は元のファイルの“TS”という列に入れている。共通して頻出することが予想される語として「授業」「英語」をあらかじめ除外し、残った語の頻度を求めて上位 30 位までを抽出する。ここでは「学習意欲:M」の場合のプログラムを示す(図 3)。

```

theme.M <- list()
for (i in 1:length(theme2)){if(theme$TS[i] == "M") theme.M[[i]] <- theme2[[i]}
# 列名"TS"が"M"となっているデータのみを取り出し

theme.M2 <- unlist(theme.M)[unlist(theme.M) != "授業" & unlist(theme.M) != "英語"]
# 「授業」「英語」を除外

theme.M.t <- table(theme.M2)
# 頻度表作成

theme.M.t[rev(order(theme.M.t))][1:30]
# 上位 30 位までを抽出

```

図 3

図 2、図 3 に示した作業を R で実行した結果が下の図 4 である。確認すると、ここでは、研究・テーマとして「学習意欲」を掲げた実践者によって、『報告書』の「クラスの特徴」「解決すべき課題」「事前の現状把握」「改善の目標」の項に記述された文言のなかから、「授業」「英語」を除いた頻度順位 30 位までの名詞・形容詞が示されていることになる。

theme.M2									
生徒	割	的	学習	こと	クラス	問題			
43	16	14	13	13	11	10			
以上	多い	音読	よう	入試	意欲	ない			
10	9	8	8	7	7	7			
進学	全体	者	活動	リーディング	ため	7			
6	5	5	5	5	5	5			
良い	中	積極	自信	希望	基礎	関心			
4	4	4	4	4	4	4			
演習	やすい								
4	4								

図 4

他の研究・テーマの記述についても同様の分析を行い、テーマ別に頻度の高い語を比較してみると、当然のことながら、「学習意欲」では「学習」、「語彙」では「語彙」、「リーディング」では「長文」、「ライティング」では「表現」、「英語使用」では「発言」といった語が、それぞれ他のテーマより特徴的に頻度が高くなっている。

5. もう一つの分析手法: コレスポネンス分析

5.1 REFLECTION のコレスポネンス分析

コレスポンデンス分析は、クロス集計表に基づいて、2 つのカテゴリの相関関係を散布図の上に表示して視覚化したもので、対応分析などとも呼ばれる。関係性の強い要素同士は近くに、弱い要素同士は遠くに表示され、軸がクロスする原点付近に表示される要素は比較的特徴が薄いと解釈される。ここでは、『報告書』の REFLECTION データ、すなわち、「教師の変化」「今後の課題」「まとめ・感想」に現れる高頻度語と、生徒の英語能力レベル(実践者の自己申告による)の関係を見てみる。つまり、生徒の能力レベルによって、実践者である教師の振り返りのポイントが特徴的に異なるかどうかを調べてみるということである。この試行では、データ数が少ないこと、リサーチ・テーマの違いによる影響がある程度は予想されることから、結果の妥当性については主張できないが、一見して解釈が容易であることから、分析手法としての価値は高いと思われる。

5.2 分析の手順

4で行った分析と同様に、csv ファイルから REFLECTION のテキストデータを取り出し、必要な形態素分析を実行する。ここでは、生徒の英語能力を「基礎レベル:1」と判断した実践者の REFLECTION データのなかの高頻度語 20 位までを抽出する手順を示す(図 5)。

```
refl <- read.csv("AP_Report2011.csv")
# csv ファイルの読み込み(オブジェクト名:"refl"に代入)
refl1 <- RMeCabDF(refl,5,1)
# csv ファイル5列目の形態素分析
refl2 <- list()
for (i in 1:length(refl1)){refl2[[i]] <- refl1[[i]][names(refl1[[i]]) == "名詞"
 | names(refl1[[i]]) == "形容詞"]}
# 名詞・形容詞のみ取り出し

refl.b <- list()
for (i in 1:length(refl2)){if(refl$LEV[i] == "1") refl.b[[i]] <- refl2[[i]}
# 列名"LEV"が"1"(基礎レベル)となっているデータのみを取り出し
refl.b2 <- unlist(refl.b)[unlist(refl.b) != "授業" & unlist(refl.b) != "英語"]
# 「授業」「英語」を除外
refl.b.t <- table(refl.b2)
# 頻度表作成
refl.b.t <- refl.b.t[rev(order(refl.b.t))][1:20]
# 上位 20 位までを抽出し、上書き
```

図 5

「標準レベル:2」「発展レベル:3」についても同様にデータの抽出を行った後、names () 関数でそれ

それぞれのデータを取り出し、1つのまとまったベクトル(同型データの集合体)にしてから、`unique ()` 関数で重複を取り除き、新たなオブジェクト “AP.WB” に代入する。その結果、図 6 のような 34 語が残った。

```
>
> # 対応分析作業
>
> AP.WB <- unique(c(names(refl.b.t), names(refl.m.t), names(refl.a.t)))
>
> AP.WB
[1] "生徒"      "こと"      "的"        "よう"      "活動"      "指導"      "学習"
[8] "の"        "語彙"      "改善"      "必要"      "時間"      "今後"      "中"
[15] "今回"      "リーディング" "ため"      "それ"      "問題"      "文法"      "自分"
[22] "処理"      "教師"      "音読"      "英文"      "意識"      "取り組み"  "力"
[29] "評価"      "方法"      "英"        "理解"      "実践"      "向上"
> |
```

図 6

このうち、「こと」「的」「よう」「の」「必要」「時間」「改善」「問題」「処理」「取り組み」「今後」「中」「今回」「ため」「それ」「力」「実践」「方法」「理解」「もの」「英」「リーディング」「語彙」「音読」「英文」については、あまり特徴を示さない語、あるいは反対に特定のリサーチ・テーマにのみ直接関わりがありそうな語として除外することにする。そのためにまず、要素番号を抽出する `which ()` 関数と比較演算子 `%in%` を使って不要語のまとまり “delete” を作り、元の “AP.WB” から削除する。次に各レベル別グループのデータのなかで不要語を取り除いた “AP.WB” のデータと一致しているデータ、すなわち「基礎」「標準」「発展」の各レベルのデータのなかで共通語リストに載っているデータのみについて、その内容と頻度を取り出す。図 7 は「基礎レベル」の作業について示したものである。

```
delete <- which(AP.WB %in% c("こと", "的", "よう", "の", "必要", "時間", "改善", "問題", "処理", "取り組み", "今後", "中", "今回", "ため", "それ", "力", "実践", "方法", "理解", "もの", "英", "リーディング", "語彙", "音読", "英文"))
# 不要語を決め、まとめる
AP.WB <- AP.WB[ -delete]
# 元データから不要語のまとまりを取り除いて上書き
refl.b.t2 <- refl.b.t[names(refl.b.t) %in% AP.WB]
# 不要語を取り除いたデータと一致するグループのデータの内容・頻度を再度取り出し
```

図 7

各グループの対象データがそろったところで、コレスポンデンス分析を行うための統合データフレームを作成する。ここでははじめに、“word”(語) “id”(レベルの名称) “Freq”(頻度)の列から成る「基礎レベル」のデータフレームを作成し、そこへ `rbind ()` 関数で「標準レベル」「発展レベル」のデータフレームをつなげるという手法を使っている。そして、できあがった統合データフレームから `xtabs ()` 関数で “Freq” に基づく “word” “id” のクロス集計表を作る。ここで、それぞれの語のレベルごとの頻度合計が整

理される(図 8)。

```
AP.DF <- data.frame(word = names(refl.b.t2), id = rep("基礎",length(refl.b.t2)), Freq = refl.b.t2)
# はじめに「基礎レベル」のデータフレーム作成
AP.DF <- rbind(AP.DF, data.frame(word = names(refl.m.t2), id = rep("標準",length(refl.m.t2)),
Freq = refl.m.t2), data.frame(word = names(refl.a.t2), id = rep("発展", length(refl.a.t2)), Freq =
refl.a.t2))
# 「標準レベル」「発展レベル」のフレームを追加
AP.t <- xtabs(Freq ~ word + id, data = AP.DF)
# クロス表の作成
```

図 8

```
library(MASS)
# MASS パッケージの読み込み
AP.corr <- corresp(AP.t, nf=2)
# コレスポネンス分析の実行
biplot(AP.corr)
# グラフィックスの作成
```

図 9

最後に、図 9 のように、コレスポネンス分析を行って、図として表示させる。はじめに `library ()` 関数を使って、この分析に必要な MASS パッケージを読み込んでいるが、これは、RMeCab と同じように、R 起動後にあらかじめ済ませておいてもよい。図 10 が分析結果である。

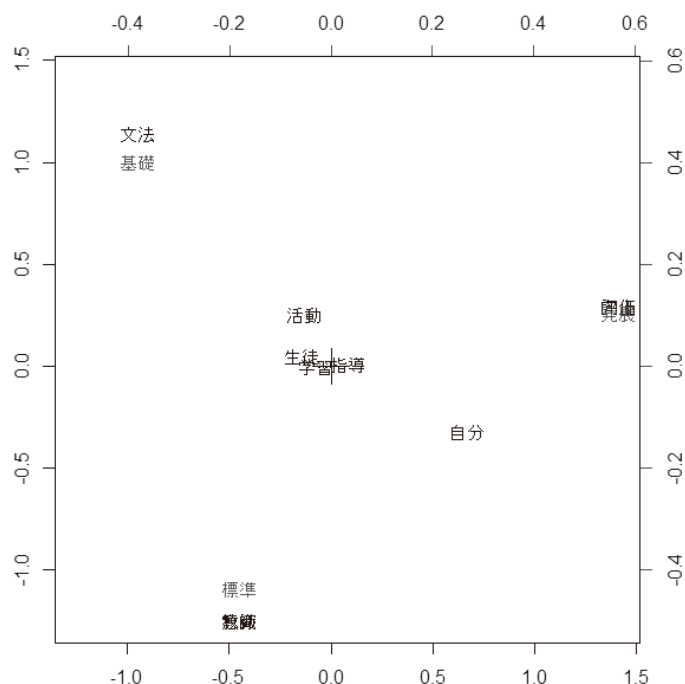


図 10

前述のとおり、原点の+印付近の語は「あまり特徴を示していない語」である。「基礎」と表示されている周辺を見ると、距離が近いところに「文法」がある。「標準」「発展」の近くの語は重なってしまっているため、Rによる作業の過程で作成されている実際の頻度表データの内容を見たところ、「標準」の近くには「意識」「教師」が、右端中ほどにある「発展」の近くには「評価」「向上」がそれぞれプロットされているということがわかった。データ数が少ないため、この結果の解釈はあまり大きな意味を持たないと思うが、文法知識の定着が大きな課題の一つである「基礎レベル」の生徒を指導する教師が、「文法」という語に多く言及しているというのは何となくうなずけるのではないだろうか。より多くのデータで同様の分析をすれば、より精緻にそれぞれの関係性を推測できるかもしれない。

6. 考察

6.1 現状のアンケート処理の課題

前述の学校現場におけるアンケートに関する調査から、多くの学校、多くのアンケートで生徒や保護者に自由に意見を記述する機会が提供されていることがわかった。しかし、その分析方法には一考の余地があるかもしれない。

まず、自由記述の回答処理として「特に何もしない」となると、回答者に対する説明責任を十分に果たせないだけでなく、生徒・保護者の本音から提供される、授業や学校をよりよいものにするための貴重なヒントを見過ごしてしまう可能性がある。アンケートとして自由に意見を述べる機会があることだけでも、ある程度回答者の満足感を得ることはできるかもしれないが、何のフィードバックもなければ、次回からは率直な意見を書いてくれなくなる恐れもある。

コメントをただ「すべて列挙する」という場合は、まず、アンケートの規模が大きくなれば、それだけいっそう、フィードバックの際にかなりのスペースが必要になる。また、読むだけでも大変な上に、そこからどのような解釈を得るかは、報告を受け取る側に任されてしまう。必ず自分が書いたコメントが載るということで、アンケート調査の誠実さを示すことはできるかもしれないが、さらにデータとしての価値を高められる可能性を残している。

「主なコメントを列挙する」という手法は、今回の調査のなかでは一番多く使われているものであったが、何をもって「主な」とするかという取捨選択が分析担当者任せになるなど、かなり主観的にならざるを得ない。意地悪な見方をすれば、都合の悪いコメントは載せないということもできてしまう。

「肯定的な意見・否定的な意見」にグループ化することは、何らかの分析が加わるという点で、データとしての価値をある程度高めるだろう。ただ、この場合は「肯定か否定か」という一つの尺度で分類していることに気をつけなければならない。どちらにも分類されなかったコメントを含め、肯定・否定の判断に直接関わることのなかった文言のなかにも、重要な視点が潜んでいるかもしれない。

6.2 学校アンケートにおけるテキストマイニング応用イメージと留意点

すべての自由記述から回答者の意識を端的に表すキーワードを抽出することができる、テキストマイニングを学校アンケートなどに応用する際には、いくつか留意すべきことが考えられる。まず、本当に必要なデータが得られるような質問紙をデザインしなければならないのは当然のこととして、回答を電子データとしてスプレッドシートに入力する時の約束事を決めておいたほうがよいだろう。例えば、次のようなコメントがあった場合には注意が必要になる。

「授業の進度が少しはやいです。」

「進むのが速くてついてくのが大変です。」

ここでは、ひらがなの「はやい」と、「速くて」の形容詞原形部分「速い」は別物としてカウントされる。「進度」「進む」については品詞も異なる。文章そのものを変えるのは大変だが、せめて入力の際に、かな書きなのか漢字を使うのかを統一しておくだけでも、後の分析が楽になる。また、「が」「は」「です」など、分析に影響しない助詞などは省略することもできるだろう。

次に、毎回決まった体裁の入力フォームを使用することも、作業効率を高めるために必要である。Rのプログラムスクリプトはテキストファイルであり、コピー&ペーストで再利用が可能である。自由記述を格納する列番号が決まっていれば、スクリプトに記述されるファイル名だけを変更すれば、二度目からは簡単にデータ処理ができる。また、複数のフォームが存在すると、まとめるのが大変なばかりでなく、思わぬ誤操作・誤入力を許容してしまうことにもなりかねないため、スプレッドシートは「共有」にし、なるべく多くの人員で分担入力した方がよい。最後の分析段階のみを特定の担当者がやれば、仕事の過度の集中は避けられる。

欲しいデータの性質によって分析方法を変えることも可能である。ここで扱ってきたテキストマイニングは頻度情報によってキーワードを抽出し、さらにカテゴリ間関係性を見るという基本的なものであるが、まだまだ多くの分析方法が存在する。例えば、次のようなコメントでは分析のしかたで情報の価値が異なるものになる。

「授業の進度が少し速いです。」

「進み具合はそれほど速くないと思います。」

単語レベルで分析すると、両方のコメントから同様に「速い」が抽出される。ここでは、「両方とも『速さ』に言及している」という情報が得られるが、「速いのか速くないのか」という肯定・否定の情報は失われている。このように、1 語でなく複数語単位・フレーズ単位で抽出する必要がある場合には、Ngram () 関数や collocate () 関数などを使った分析をすることもできる。

7. おわりに

学校での教育活動だけでなく、人が関わるあらゆるプログラム評価にはアンケート調査は不可欠である。それぞれの選択肢を選んだ人数やその割合などを示す数量的データ分析は、客観的で説得力もあり、違いや成果が見えやすい。それに比べると、自由記述の回答については、実施者の受け取り方も、報告のされ方もかなり傍観的・主観的になっていないだろうか。自由記述のなかには、選択肢では拾いきれない、そのプログラムをよりよいものにするためのヒントが埋め込まれているかもしれない。そこに光を当て、より客観的なデータとして提示するための一方法として、テキストマイニングはさらなる可能性を持っていると考える。

(参考文献)

- 石田基弘. (2008).『R によるテキストマイニング入門』森北出版.
- 石田基弘.「RMeCab—R と Linux と…」<http://rmecab.jp/wiki/index.php?RMeCab> (2012 年 10 月 30 日)
- 加藤千恵子・慮志和・石村和夫. (2007).『SPSS でやさしく学ぶアンケート処理 第 2 版』東京図書.
- 佐野正之(編著). (2000).『アクション・リサーチのすすめ—新しい英語授業研究』大修館書店.
- 佐野正之(編著). (2005).『はじめてのアクション・リサーチ—英語の授業を改善するために』大修館書店.
- ドルニエイ, Z(著)八島智子・竹内理(監訳). (2006).『外国語教育のための質問紙調査入門』松柏社.
- 文部科学省. (2007).「学校評価に関する学校教育法・学校教育法施行規則の規定」
http://www.mext.go.jp/a_menu/shotou/gakko-hyoka/08021216.htm (2012 年 10 月 30 日)
- 文部科学省. (2000).「学校教育法施行規則等の一部改正について(学校評議員等関係)」
http://www.mext.go.jp/component/a_menu/education/detail/_icsFiles/afieldfile/2010/04/06/1230703_001.pdf (2012 年 10 月 30 日)
- 神奈川県立国際言語文化アカデミア. (2012).『平成 23 年度英語教育アドヴァンスト研修 授業改善プロジェクト報告書』<http://www.pref.kanagawa.jp/uploaded/attachment/440003.pdf> (2012 年 10 月 13 日)
- 山田剛史・杉澤武俊・村井潤一郎. (2008).『R によるやさしい統計学』オーム社.
- MeCab [computer software]. <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>
- R version 2.15.1. [computer software]. The R Project for Statistical Computing
<http://www.r-project.org/index.html> (2012 年 10 月 30 日)