

『就実論叢』第42号 抜刷

就実大学・就実短期大学 2013年2月28日 発行

保育士を目指す大学生の、肢体不自由、病弱・虚弱、 発達障害、知的障害に関する意識調査

**A Survey of Nursery Teacher Training Students on Disabilities – Physical Disabilities,
Health Impairment, Developmental Disabilities, Mental Retardation.**

岡 田 信 吾
下 山 真 衣
石 山 貴 章

保育士を目指す大学生の、肢体不自由、病弱・虚弱、 発達障害、知的障害に関する意識調査

A Survey of Nursery Teacher Training Students on Disabilities – Physical Disabilities,
Health Impairment, Developmental Disabilities, Mental Retardation.

就実大学教育学部

岡田信吾

下山真衣

石山貴章

Shujitsu University Faculty of Education

OKDADA Shingo

Shimoyama Mae

Ishiyama Takaaki

抄録

大学の教員養成系学部において保育士資格を希望する学生に、肢体不自由、病弱・虚弱、発達障害、知的障害に関する意識調査を行った。自由記述による質問紙調査を実施し、テキストマイニングを利用して分析した。その結果、ほとんどの学生がこれらの障害実態に関する知識を有せず、これまでに接触した経験も少ないことが明らかとなった。また、自由記述において使用された語と障害種との対応分析の結果から、肢体不自由と病弱・虚弱とは別の障害として認識されているが、知的障害と発達障害とはほぼ同じような障害として認識されている可能性が示唆された。

キーワード；保育士、大学生、障害意識調査、テキストマイニング

I はじめに

保育士資格者は、保育所における保育のみならず、児童養護施設や障害児施設など就学前のすべての子どもに関わる専門職である。そのため、子どもに関わる一般的な知識が求められると同時に、障害のある子どもへの対応についても広く知る必要がある [1]。しかしながら、学生が大学入学前に生活や学習の中で障害のある子どもについてどのような関わりを持ち、また知識を有しているのか明確ではない。特に、病弱・虚弱、肢体不自由に関しては小学校、中学校の教員であっても知識不足が指摘されている [2] [3]。そこで、保育士資格を希望する学生の、これら障害に関する知識とこれまでの接触経験とを調査し、把握する必要

を感じた。今回の研究の第一の目的はこの確認である。

次に、研究方法についてであるが、このような調査においてはあらかじめ設定された設問に対する回答を統計的手法によって検討する方法が多く用いられている [4]。これに対して形式化されていない自然な文章（以下テキストとする）を単語やフレーズなどに分割し、出現頻度や共起関係などを集計し、統計的手法を用いて解析するテキストマイニングと呼ばれる手法が注目を集めている [5]。従来、テキストは質的研究の対象とされてきた。質的研究の方法は、研究事例の集積もされており信頼に値する研究方法である。しかし、データが膨大になった場合、その解析には多くの時間と労力がかかり、効率的な活用は容易ではない。また、解析を行う個人の経験や知識によって結果や解釈などが異なる場合もある [5]。一方テキストマイニングにおいては、解析を定式化しコンピューターを利用することによりこれらの問題点にある程度対応する事ができると考えている。今回の研究の第二の目的は、テキストマイニングを利用して自由記述表記された質問紙の分析を行い、その解析手法としての可能性を検討することとした。

II 方法

1 調査対象

20XX 年後期に、A 大学教育学部に在籍する 1 年生 64 人（女子 53 人、男子 11 人）を対象として質問紙調査を実施した。質問紙に回答した学生は、全員保育士資格取得を希望し、障害児保育に関する講義を受講した学生であった。なお、事前に研究の目的を口頭説明した。

2 質問紙について

質問紙は、肢体不自由、病弱・虚弱、知的障害、発達障害のそれぞれについて、(1)障害名を聞いたことがあるか、(2)その障害のある子どもは、どのような障害のある子どもだと考えているか、(3)その障害のある子どもは、どのような事に困り、どのような支援が必要だと考えているのかの 3 要素で構成し、(1)以外はすべて、自由記述とした。さらに、障害のある人とのこれまでの接触経験についても記述させるようにした。

3 分析方法

分析は、テキストを統計的な手法で分析するテキストマイニングによって実施した。

テキストマイニングに先んじて、質問紙に記述された文を、すべて電子データとして入力し、前処理を行った。前処理は、①文中の記号や特殊記号を外す、②明らかな誤字や書き間違いと思われる語を修正する、③「子ども」と「子供」、「足」と「脚」など漢字・仮名表記等の統一、④「同級生」と「クラスメイト」など、意味は同一と考えられるが表現が異なる語の統一であった。これらの処理を行ったテキストを、CSV（カンマ区切り）形式で出力し、テキストマイニング用のデータを作成した。出力フォーマットは、1 行目が項目名で、2 行目以降各セルに 1 項目のデータを入力した。

処理に当たっての変数名として、PD1、PD2、HD1、HD2、MR1、MR2、DD1、DD2 を使用した。

これは、それぞれの障害種の略称に、「(2) その障害のある子どもはどのような障害のある子どもだと考えているか」の設問の回答に1を、「(3) その障害のある子どもは、どのような事に困り、どのような支援が必要だと考えているのか」の回答に2を付け加えて指定した物である。

1) 辞書の整備

テキストマイニングに用いる辞書の調整を行った。新たに登録した形態素（文中において意味の分かる最小単位）は、アスペルガー障害など障害名に関する物、特別支援学校などの学校・学級種に関する物など32語であった。辞書の単語登録の手続きは以下の情報を参照した（確認日2012年1月）。

<http://mecab.googlecode.com/svn/trunk/mecab/doc/dic.html>

2) 分析

分析は、R_2.14.1 [6] と MeCab_0.99 [7]、RMeCab_0.99 [8] を使用した。あわせて、データ整形に Excel 2010 [9]、ネットワーク分析の図示に Cytoscape_2.8.2 [10] を使用した。

(1) 形態素分析

形態素分析は、文章を特定の辞書を用いて、意味の分かる最小単位（形態素）に分割する手続きである。日本語において形態素分析を行った場合、活用形の問題が存在するが、今回の分析においては、すべて終止形で出力し処理した。

実際の分析にあたっては、以下の手続きによって、CSV 形式で保存されたデータを一括して読み込み形態素分析を実施した。なお、手続き中の XX はそれぞれの変数名とする。

```
library ("RMeCab")          # R にパッケージ RMeCab を読み込む
data <- read.csv ("data.csv", header=TRUE)      #CSV ファイル "data.csv" を読み込み、
変数 data に格納する
XX1 <- RMeCabDF (data, "XX", 1)      #data の項目名 (XX) を対象としてデータ毎に形態
素分析を実施し、終止形で出力し、変数 (XX) に格納する
```

(2) データ長とヒストグラム

それぞれのデータの形態素長を計数し、ヒストグラムに表した。手続きを以下に表す。

```
for (i in 1:length(XX)) {if (any (is.na(XX[[i]]))) {XXLen[i] <- NA}
  else {XXLen[i] <- length(X[[i]])}}
hist(XXLen)
```

(3) 品詞毎の出現度数

個々のデータに使用される名詞、動詞、形容詞の出現度数を計数し、CSV 形式で出力した。その手続きを以下に示す。名詞を例として取り上げる。

```

XXFDNOUN <- unlist(XX) # リスト表示
XXFDNOUN <- XXFDNOUN[names(XXFDNOUN) == "名詞 "] # 名詞を抽出
XXFDNOUN <- table(XXFDNOUN) # 度数分布
XXFDNOUN <- XXFDNOUN[rev(order(XXFDNOUN))][1:20] # 並び替えて上位20を出力
Write.csv(XXFDNOUN, "XXFDNOUN.csv") # 度数分布を CSV ファイルとして出力

```

(4) クロス集計表の作成

この後の処理においては、各変数毎の形態素の出現度数を一覧にしたクロス集計表が必要となる。そのため、クロス集計表の作成手順について概説する。

① 処理用データの作成

クロス集計表の作成は、`docMatrix` を用いるが、このコマンドは、任意のフォルダ内にあるファイルを対象に、ファイル名を列名として、それぞれに存在する形態素の度数を計数するコマンドである。そのため、変数名をファイル名とする形式にデータを再構成しなければならない。そのための手続きを以下に示す。

```

data <- read.csv("data.csv", header=TRUE)
XX <- (data[, "XX"]) # data から列名 XX を抽出
write.csv(XX, "XX.csv")

```

分割したファイルは、障害の実態像に関する内容が記述されたデータをフォルダ名 `hdata1`、必要な支援に関する内容が記述されたデータをフォルダ名 `hdata2` としてそれぞれ別フォルダに保存した。

② `docMatrix` の適用

上記の手続きで処理したファイルを格納しているフォルダを対象に、`docMatrix` を適用し、列に変数名、行に使用されている形態素のクロス集計を作成する。その手続きを以下に示す。

```

library("RMeCab") # ライブラリー RMeCab の読み込み
dM1 <- docMatrix("hdata1") # フォルダ hdata1 に docMatrix を適用
write.csv(dM1, "dM1.csv") # CSV ファイルとして出力

```

(5) 形態素と障害種の対応分析

対応分析はコレスポンデンス分析とも呼ばれ、テキストを分類する手法である [11]。この分析手法は、クロス集計表で表現されるデータを2次元空間上に配置し、要素間の関連や類似点を距離で表現し直感的な把握が出来るようにする手法である [11]。この方法の特徴として、原点付近にはどの要素についても関連ある要素が集まるため、比較的特徴が薄いと解釈できる。この手続きを以下に示す。この分析においては、形態素を上位10%に絞り実施した。

```

library(MASS)           # ライブラリーの読み込み
DA <- read.csv("dM1sumsort.csv", header=TRUE, row.names=1) # データの読み込み
DA.coa<-corresp(DA[1:29,1:4], nf=4)      # コレスポネンス分析の実施
biplot(DA.coa,main="障害に関する意識    # バイプロット図の出力 ")
# 累積寄与率を求める
EV<-DA.coa$cor^2
CR<-(100*EV/sum(EV))
CCR<-sum(CR[1:2])      # 第2成分までの累積寄与率

```

(6) N-gram によるネットワーク分析

① N-gram 分析

N-gram とは、連続する要素の中から n 個の一連のデータを切り出す手続きの事である。例えば、bi-gram であれば連続する 2 つの語を抽出し、tri-gram であれば連続する 3 つの語を抽出する。今回の研究では、名詞、形容詞、動詞を対象として bi-gram 分析を行った。この手続きを以下に示す。この手続きによって出力されるのは、bi-gram の組み合わせとその出現度数である。なお、NgramDF 中の、(type) は 1 が形態素対象、2 が品詞対象である。

```

library("RMeCab")      # ライブラリーの読み込み
chanN2 <- NgramDF("XX1.csv", type =1, N=2,pos = c("名詞", "動詞", "形容詞"))
write.csv(chanN3, "XX1chan.csv")

```

② ノードの度数

ノードとは、ネットワークの枝と枝がつながる頂点の事である。Cytoscape [10] による画像出力において、ノードの出現度数が必要であった。ノードの出現度数の計数手続きを以下に示す。

```

foo<- read.csv("XXchan.csv",header=TRUE) # 対象とする N-gram データの読み込み
foaname <- c("ID","degree")             # データフレームの変数名の準備
foo1<-foo[,2]                            # 読み込んだ CSV ファイルのデータを列ごとに読み込み
foo2<-foo[,3]
foo3<-foo[,4]
foo13 <- data.frame(foo1,foo3)            #1列目と3列目からデータフレームを作成
names(foo13) <- foaname                  # 列名をつける
foo23 <- data.frame(foo2,foo3)            #2列目と3列目からデータフレームを作成
names(foo23) <- foaname                  # 列名をつける
foo123 <- rbind(foo13,foo23)              # それぞれの列を統合する
foodegree <- tapply(foo123$degree,foo123$ID,sum) # 単語毎に集計する
write.csv(foodegree,"XX1chandegree.csv") # CSV ファイルに出力する

```

③ ネットワークの可視化

上の手続きで得たデータを Cytoscape_2.8.2 [10] を用いてグラフ化した。Cytoscape [10]

は対話型で、直感的なデータ分析が可能であるためこれを利用した。

まず、Cytoscape にデータを読み込ませる。Cytoscape を立ち上げ、File メニューから、Import → Network from table を選択する。目標とするエクセルファイル、または CSV ファイルを読み込み、(Source interaction) に Ngram1 を、(Interaction type) に Frequency を (Target interaction) に Ngram2 をそれぞれ指定する。

次に、②で得たノードの計数データを読み込む。

File メニューから、Import → Attribute from table を選択し、上記のデータを読み込む。このデータは、CSV 形式のままでも読み込み可能ではあるが、エクセル形式 (xlsx) でファイル作成した方が、処理が容易となる。

なお、Cytoscape の初歩的な利用方法は、下記の情報を参照した (確認日2012年1月)。

<http://togotv.dbcls.jp/20110630.html?103-p01>

Ⅲ 結果と考察

1 障害者との接触経験について

表1に、障害者との接触経験について示す。

この設問に対しては無答の者が多くいた。接触経験のあった者は記述すると考え、無答は接触経験なしとして計数した。接触経験がないとした回答は、全体の4割であった。また、子どもの頃、同級生などに障害のある友達がいたとする回答とボランティアなどで接触経験のあるとした回答とはほぼ同数であった。さらに、いとこ、おじ・おばなど親族に障害のある者がいるとした回答もあった。

表1 障害者との接触経験について

接 触 経 験	人数
接触経験なし+無答	26
インターンシップ、ボランティアなどで接触あり	15
同学年、同級生に障害のある友達がいた	17
家族や親戚に障害のある者がいる	6

2 それぞれの障害に関する意識調査の結果

1) 形態素分析における基本統計量

自由記述の形態素分析の結果を示す。

4つの障害種間の自由記述文の形態素による語数の平均と分散を表2に示す。障害像について記述された文の形態素による平均語数は、困難点と支援について記述された文の平均語数と比較して一貫して少ない。

図1に自由記述された回答の障害種とその意識/必要な支援における形態素長のヒストグラムを表示する。

表 2 各障害に対する形態素の基本統計量

		肢体不自由	病弱・虚弱	知的障害	発達障害
障害像	<i>mean</i>	13.5	14.9	13.8	16.8
	<i>SD</i>	6.8	8.0	8.0	9.7
障害に起因する 困難点と支援	<i>mean</i>	24.2	24.2	25.6	27.4
	<i>SD</i>	13.5	14.7	14.7	13.8

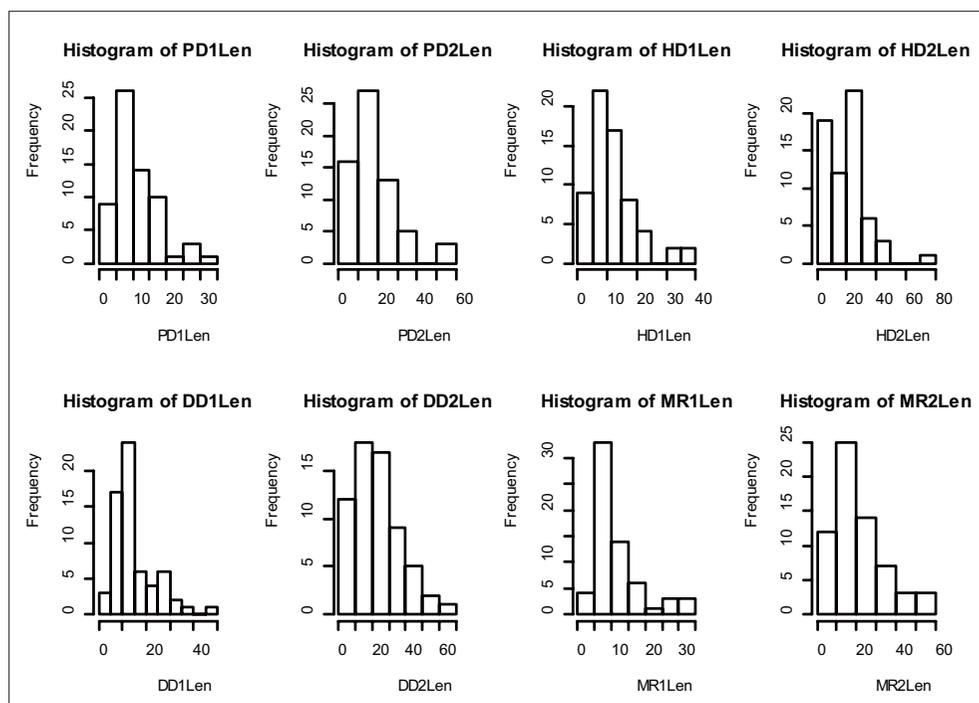


図 1 自由記述されたテキストの形態素長のヒストグラム

2) 使用された形態素 (名詞) の頻出語

表 3、表 4 にそれぞれ、「障害像」と「障害に起因する困難点と支援」に使用された形態素の名詞の上位 5 位を示す。

3) 障害種と使用された語の対応分析

障害種と使用された形態素とのクロス表を対応分析し、障害種と形態素との関連をグラフ上の距離として表現した。この対応分析においては、自由記述文から名詞、動詞、形容詞に注目して抽出し、出現度数が上位 10% にあたる形態素を対象として分析した。

表3 「障害像」、「困難さと必要とする」に使用された形態素（名詞）の上位5位

	肢体不自由		病弱		発達障害		知的障害	
	形態素	度数	形態素	度数	形態素	度数	形態素	度数
障害像	子ども	51	子ども	59	子ども	55	子ども	43
	体	31	病気	51	発達	29	こと	26
	手足	24	体	38	こと	29	人	14
	こと	20	こと	18	同級生	20	障害	14
	足	18	生まれつき	7	人	12	脳	12
困難さと必要とする支援	こと	49	こと	40	こと	53	こと	56
	よう	22	よう	27	人	24	必要	19
	人	18	学校	16	必要	22	よう	18
	日常生活	16	支援	15	よう	22	人	15
	必要	15	子ども	15	支援	18	子ども	15

(1) 障害像について

図2（左）は「障害に関する意識」を対象として処理した図である。この図の第2固有値までの累積寄与率は85.28で、十分に情報を引き出していると考えられる。なお、累積寄与率とは寄与率（各固有値が固有値全体にしめる割合）の和であり、この累積固有値の値が高いほど図に十分に情報を引き出していると考えられる指標である [12]。病弱、肢体不自由については関連する形態素が比較的分離しており、それぞれが別の概念として認識されていると考えられる。一方で、発達障害と知的障害については、関連のある形態素が密接しており、ほぼ同じ語と関連深いと考えられる。つまり、発達障害と知的障害とは対象者にとって同一概念として認識されている可能性がある。

(2) 障害に起因する困難と必要とする支援について

障害に起因する困難と必要とする支援について、障害種と使用された語の対応分析を図2（右）に示す。病弱・虚弱、肢体不自由に関しては、共通して使用された形態素が少なく、それぞれが別の概念として認識されていると考えられる。一方、発達障害と知的障害につい

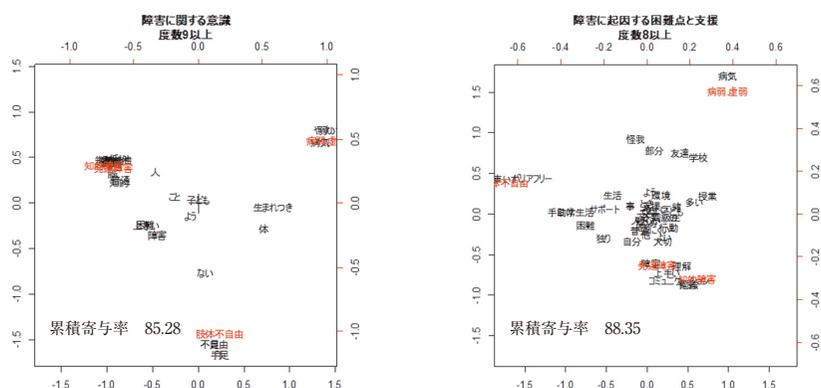


図2 対応分析の結果（左）障害種と説明に使用された形態素、（右）障害種と困難点と支援に使用された形態素

示唆されている。

肢体不自由と病弱については、対応分析においてもその分離が確認されたが、この分析においても異なった形態素で説明されている。肢体不自由に関して多く用いられているのは、手や足の不自由という形態素であり、病弱虚弱に関しては「病気-なる」「病気-かかる」などの形態素配列と「体が弱い」という形態素配列である。特に、「なる」、「かかる」といった動詞は、病弱虚弱においてのみ使用される形態素であった。

(2) 障害に起因する困難と支援について

図4に障害に起因する困難さと支援に関するネットワーク図を示す。知的障害と発達障害に関しては、その違いが分かりにくい。一方で、病弱・虚弱については「病気」、「学校」が特徴的である。また、肢体不自由については、「バリアフリー」、「日常生活」が特徴的である。これらの語が、病弱・虚弱および、肢体不自由の困難と支援を特徴付けていると考えられる。

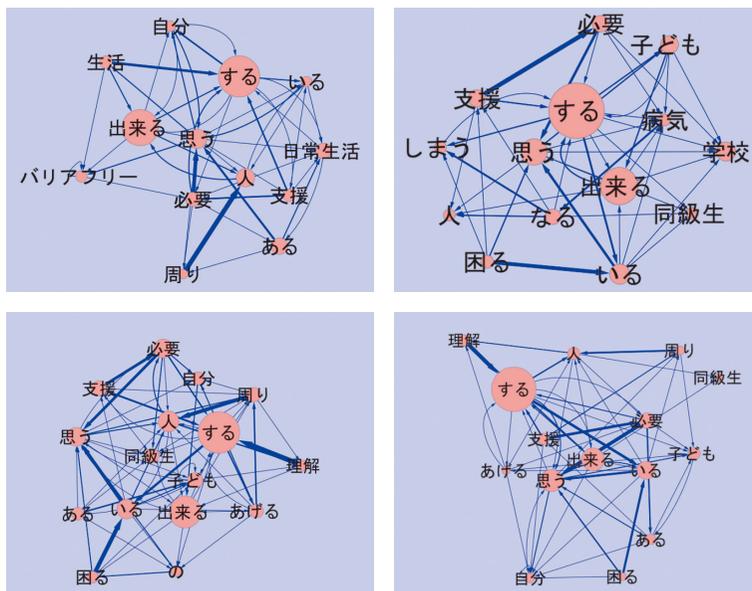


図4 障害の困難さと支援
(左上)肢体不自由、(右上)病弱虚弱、(左下)発達障害、(右下)知的障害

IV まとめ

今回の研究は、保育士希望学生の障害に関する意識を自由記述表記によって調査し、テキストマイニングによって分析した。対応分析の結果から、発達障害と知的障害の混同が示唆されたが、ネットワーク分析の結果から一応は否定された。しかし、それぞれ分析結果を詳細に検討すると、学生の障害像に関する認識はその障害を示す名称に引きずられる可能性があることが示唆された。すなわち、病弱であれば病気をイメージし、肢体不自由であれば手

足の不自由さをイメージし、知的障害であれば学習上の困難や知能指数などの内容をイメージするという認識である。その中で、発達障害はその名称から直接思い浮かぶ事物がなく、知的障害と似た形態素によってその状態像を表現したと考えられた。また、病弱虚弱においてのみ、「なる」、「かかる」といった形態素が使用されていた。他の障害種は、「不自由」や「障害」、「遅れ」といったどちらかという、変化しない定常的な状態を示す語によって、説明されていた。この違いは、病弱が単に病気として解釈され、治りうる障害であるにとらえられている事を示唆している。本来、病弱虚弱とは、慢性疾患があり疾患への対応が日常的に必要とされる障害種である。次に、障害に関する支援においては、肢体不自由に関してスロープの設置やバリアフリーと言った語が見られるものの、他の障害種についての具体的な支援の方法まで示した者は少なかった。

これらの結果から考えられることは、高等学校卒業までに障害に関する具体的な知識やその支援方法についてほとんど知らないという現状である。障害者との接触経験に関しても、家族に障害者がいる者を除くと、深い関わりがあるとは言えない。この点において、幼い段階から障害のある友達と良い関係・良い関わりを保つとともに、正しく障害理解ができるよう系統的で持続的な障害理解教育の推進が重要であると考えられる。

次に、テキストマイニングという手法は、その手法を用いるだけで差の検討や弁別が出来る物では無い。結果を基に、その分野に関する専門知識を持つ者が解釈をする必要がある。しかしながら、質的な研究の方法と比較すると解釈に準拠する部分は少なく、いつ、誰がやってもある程度同じ結果の得られる手法である。その意味では今後大きな発展が期待できる。しかしながら、この手法はまだまだ開発されて日が浅く、定式化された処理の流れもない。今後、こういった研究への適応を繰り返し手法として洗練させる必要があることは確かであろう。

引用文献

1. 保育士養成校における学生の学習に対する意識調査 ―演習「障害児保育」の授業への取り組みを中心に―. 松尾寛子. 10, 2009年, 関西国際大学研究紀要, ページ: 209~216.
2. 通常の学級に於いて病弱児への教育的支援を困難と感じる理由―教師を対象とした自由記述の分析を通して―. 平賀健太郎. 2006年, 大阪教育大学障害児教育研究紀要, 第29巻, ページ: 71-78.
3. X県下小学校における肢体不自由学級の「交流教育」に関する実態調査研究. 妻木節子, 山崎由可里. 2004年, 和歌山大学教育学部教育実践総合センター紀要, 第14巻, ページ: 227-236.
4. 大学生における子ども観に関する一考察. 星野修一, 日潟淳子, 吉田圭吾. 1, 2008年, 神戸大学大学院人間発達環境学研究科研究紀要, 第2巻, ページ: 33-42.

5. 金 明哲. 統計的テキストマイニング. テキストデータの統計科学入門. 出版地不明 : 岩波書店, 2009, ページ : 1-12.
6. R Development Core Team. (オンライン) 2011年 . <http://www.R-project.org/>.
7. MeCab. (オンライン) 2011年. <http://mecab.sourceforge.net/>, MeCab: Yet Another Part-of-Speech and Morphological Analyzer.
8. 石田基広. (オンライン) 2011年 . <http://rmecab.jp/wiki/index.php>.
9. Microsoft. Excel 2010. 2010年.
10. Cytoscape:An Open Source Platform for Complex Analysis and Visualization. (オンライン) (引用日 : 2012年 1月10日 .) <http://www.cytoscape.org/>.
11. 金 明哲. テキストの特徴と話題分析. テキストデータの統計科学入門. 出版地不明 : 岩波書店, 2009, ページ : 141-158.
12. 高橋 信. 分析結果の精度. Excel で学ぶコレスポネンス分析. 出版地不明:オーム社, 2005, ページ : 125-127.
13. 金 明哲. テキストデータの統計科学入門. 出版地不明 : 岩波書店, 2009. ページ : 26.
14. 教員養成課程在籍学生の子ども観に関する一考察. 吉澤千夏, 大瀧ミドリ. 2011年, 上越教育大学研究紀要, 第30巻, ページ : 221-233.