

# テキストマイニングによる短期海外研修の自由記述の分析

飯塚雄一 ケイン・エレナ 小玉容子 松本亥智江

1. 目的
2. 研究方法
  - (1) 対象
  - (2) 方法
3. 結果
4. 考察

## 1. 目的

近年、学生の短期海外研修への応募が減少傾向である。高い研修費用をどうするかという検討も重要であるが、どうすれば学生の短期海外研修へ興味や関心を高めてやるかの方策を検討する必要がある。1996年から出雲キャンパスで始まった語学・看護学海外研修及び浜田キャンパス、松江キャンパスでの研修に対する学生の感想文（自由記述）を資料として検討する。

古くから異文化接触の記録はマルコポーロ、クック、小泉八雲などをはじめとして多い。また短期間の物理的移動と適応の関係をみる自由記述の主観的分析も多い。本研究では、短期海外研修の自由記述の報告書を、最近の新しい手法であるテキストマイニングによって検討した。複数年度にわたる大量な自由記述感想文の中から有用な情報を選別し取り出すために、テキストマイニングによって大量な発言内容をコンピュータにより処理すれば、代表的な感想や傾向を把握することができる。また処理された言語は数量化されているので様々な分析にかけることができる。これを活用することで、学生たちの希望や不満などを把握することもできる。また学生たちがどのような収穫を得たかを探ることもできる。

そこでこれまでに浜田、松江、出雲の各キャンパスで海外研修へ参加した学生たちがどのような感想を述べているかを調べる。つまり出雲キャンパス、松江キャンパス、浜田キャンパスの既存の海外研修報告書の自由記述感想を、使われている単語という側面からイメージを分析する。更に年度やキャンパスによる違いがあるかなども調べる。海外情報がマスコミなどであふれている現状で、実際に海外へ行くことから得られるものが何かを探り、これを学生への動機づけの一助として活用できないか検討する。

## 2. 研究方法

### (1) 対象

島根県立大学、短大部出雲キャンパス、松江キャンパスで実施された短期海外研修の報告書を対象とする。松江キャンパス 1991 年度(ワシントン州エレンズバーグとシアトル)、島根県立大学は 2007 年度(カリフォルニア州モントレイ)、出雲キャンパスは 1996, 1998, 2000, 2003, 2006, 2007 の各年度(ワシントン州ワナチとシアトル)の報告書を取りあげた。

## (2) 方法

質的データ解析はテキストマイニングのソフトウェアである TMStudio3.0 (数理システム) を用い分析した。感想文から得られたテキスト型データはまず分かち書きされ単語 (構成要素) に分けられ、様々な分析を行った。

## 3. 結果

### 1) 頻度分析

#### (1) 単語頻度解析

テキスト情報の品詞別出現回数で頻出順に、名詞、動詞、副詞、形容詞、などで感動詞も多かった。

まず、報告書にどのような単語が多く出現しているかをみるために、単語頻度解析を行った (図1)。すべての報告書を合わせて最も頻度の多い単語は、「アメリカ」、「日本」、「人」、「行く」、「自分」、「英語」などの単語が多く出ていることがわかる。

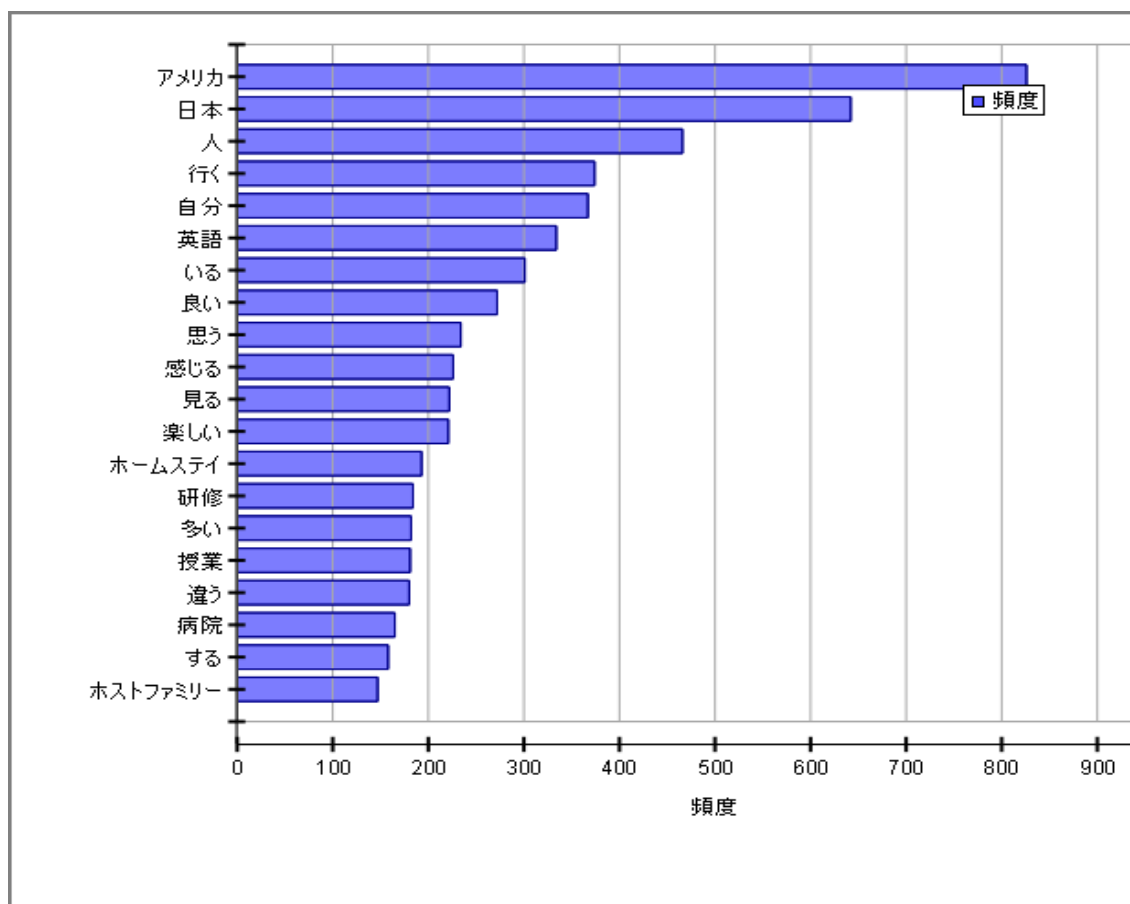


図1. 単語出現頻度

#### (2) 係り受け頻度解析

単語頻度解析は1語を見ているが、さらに詳細に見るため2語一係り・受け関係一における単語同士の関係を見る。図2は感想文の中に表れている、係り受け表現について、係

り元単語と、係り先単語の頻度を求め、対応分析によりデータを2次元上に配置してみた結果である。対応分析では関連のあるものは、近い点に配置される。「参加は良いもので、喜びも大きく、出会いが大切である」、そして、「別れは辛い」、「英語が苦手」というメッセージが読み取れる。

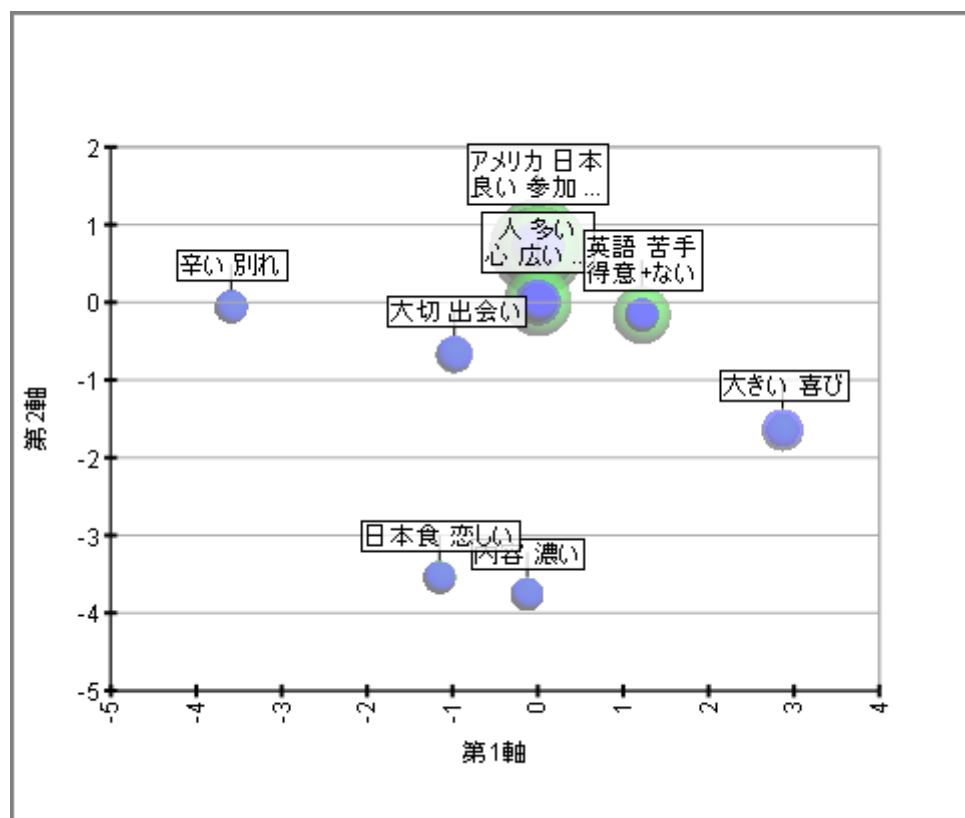


図2. 係り-受けの関係

### (3) 文章分類

次に、感想文の中のすべての文を言葉の使われ方が、似ているもの同士を5つに分類してみた。文章分類とは、使われている単語に応じて、テキストの行もしくは文章をクラスターに分類し、クラスターごとの特徴単語を表示することである。分類1は「自分が行って見るのは良い」というクラスター（図3）。分類2は、「アメリカへ自分が行く」というクラスター（図4）、分類3は、「英語の授業やホームステイは楽しい」というクラスター（図5）、分類4は、「人」などの単語を多く含む文章の集まり（図6）、分類5は「日本とアメリカの病院は違う」というクラスター（図7）があった。

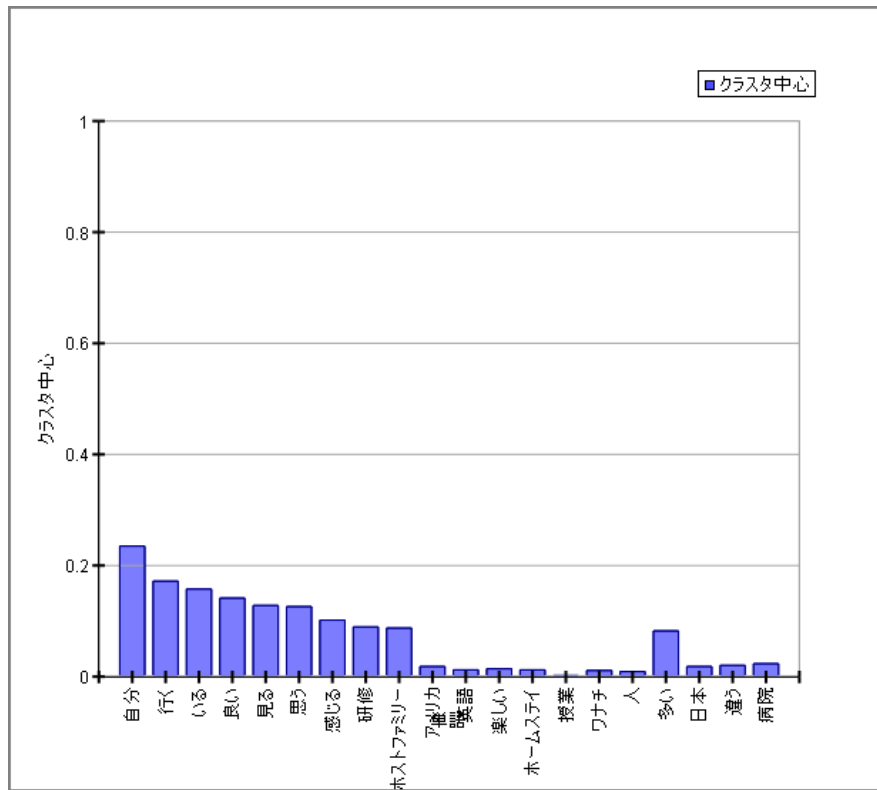


図3. 文章分類 1—自分

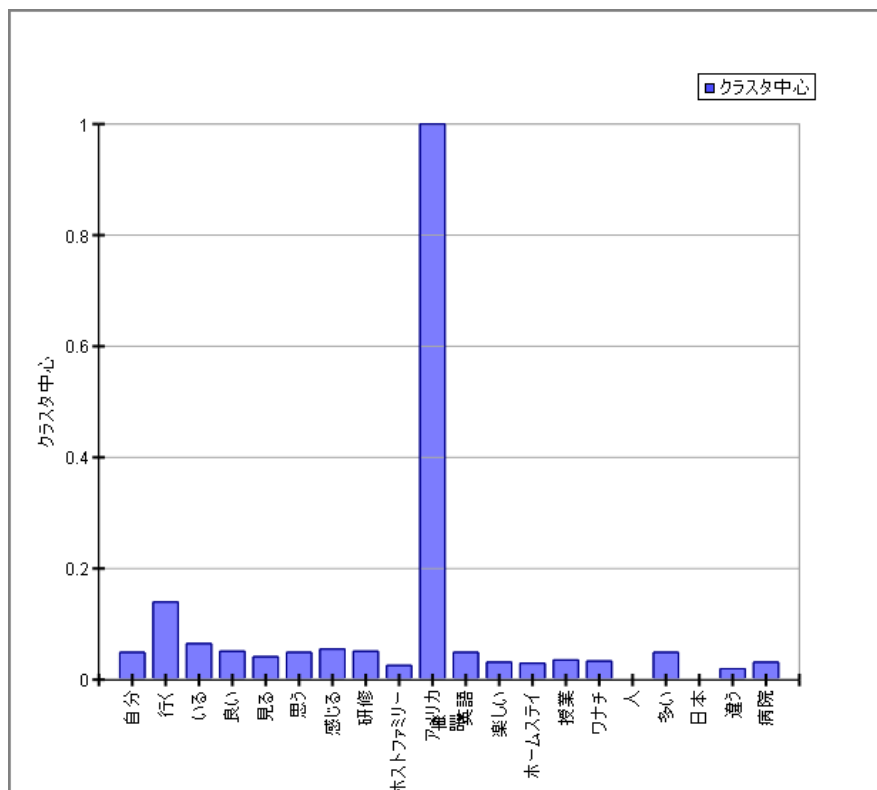


図4. 文章分類 2—アメリカ

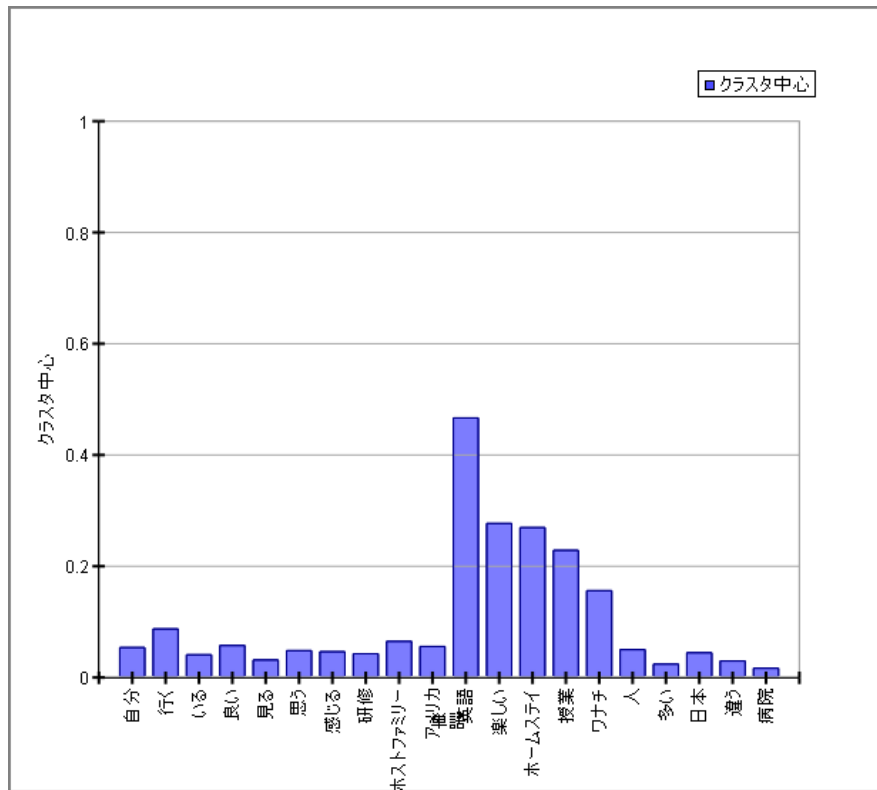


図5. 文章分類 3—英語

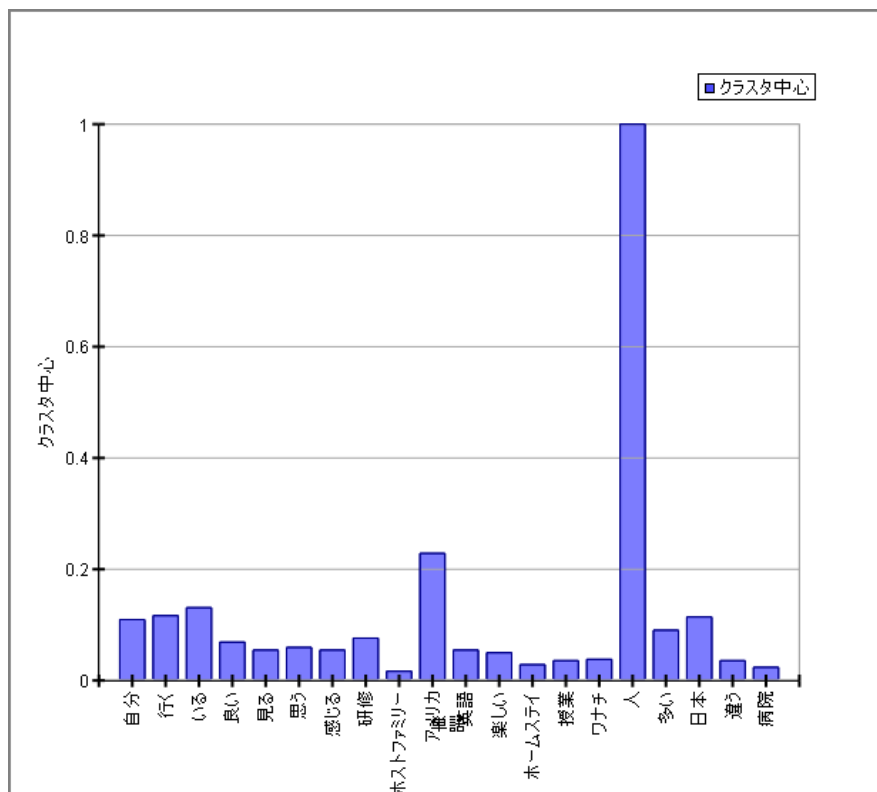


図6. 文章分類 4—一人

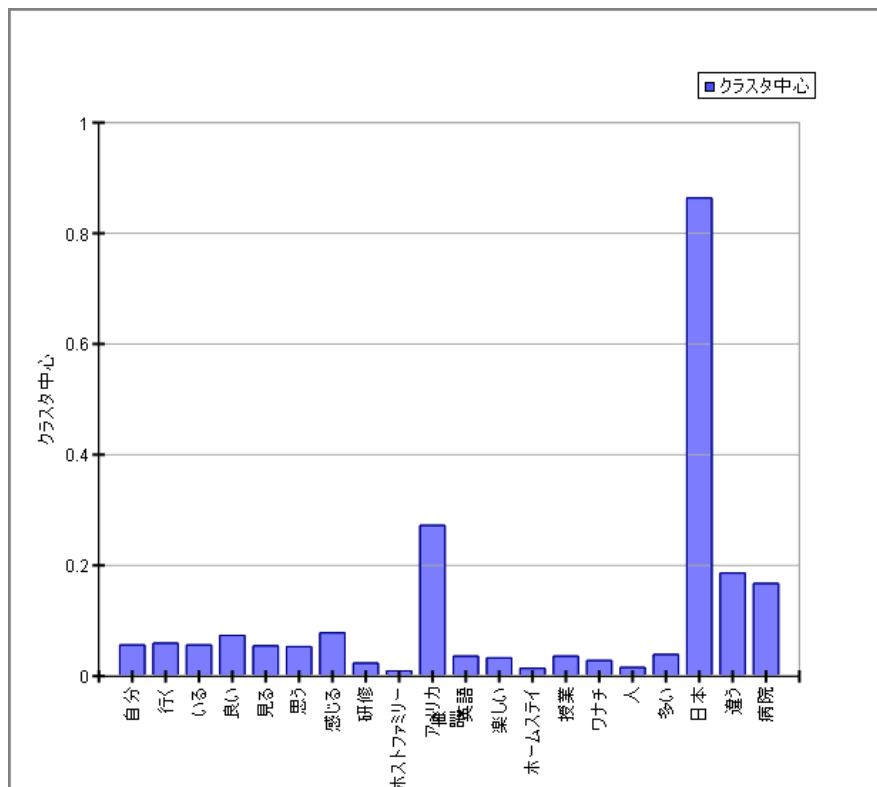


図7. 文章分類 5—日本

(4) 評判分析

次に、良いイメージで語られている単語、悪いイメージで語られている単語を調べてみた(評判抽出)。これは、単語に対して好意的な表現や非好意的な表現それぞれ語られた回数をカウントし、それをもとに、好評語・不評語のランキングを作成する。つまり単語に対して、好意的な表現、否定的な表現、それぞれで語られた回数を数え、それをもとに単語の、肯定度、否定度の順位を表した(図8, 9)。人、体験、思い出、アメリカ、などが肯定的にみられ、別れ、英語、英語力等の語は否定的にみられている。

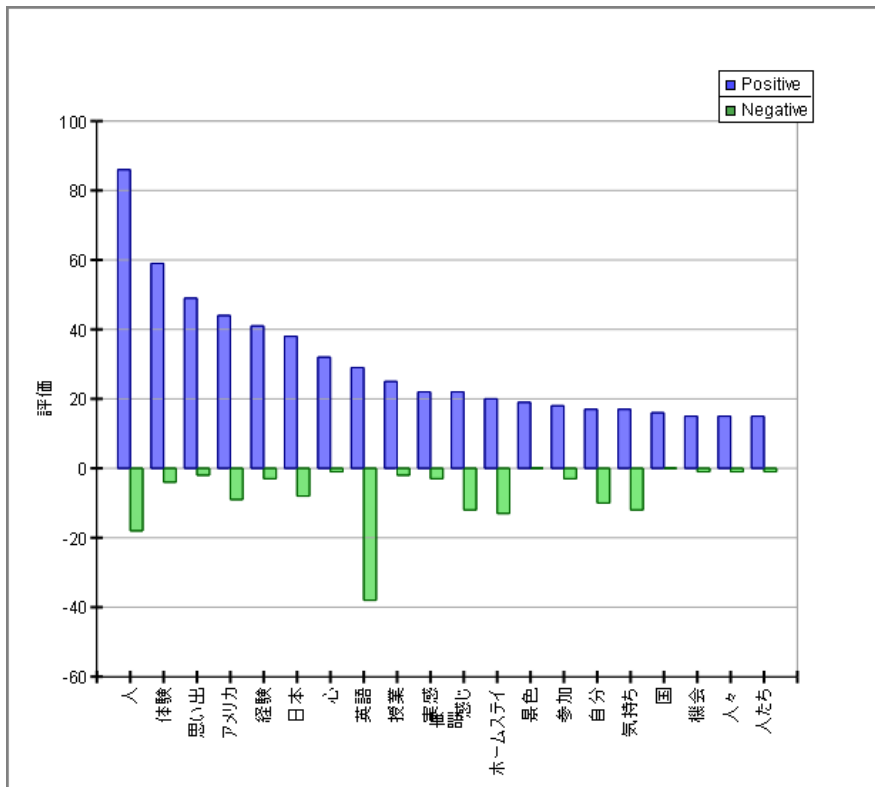


図8. 好評語ランキング

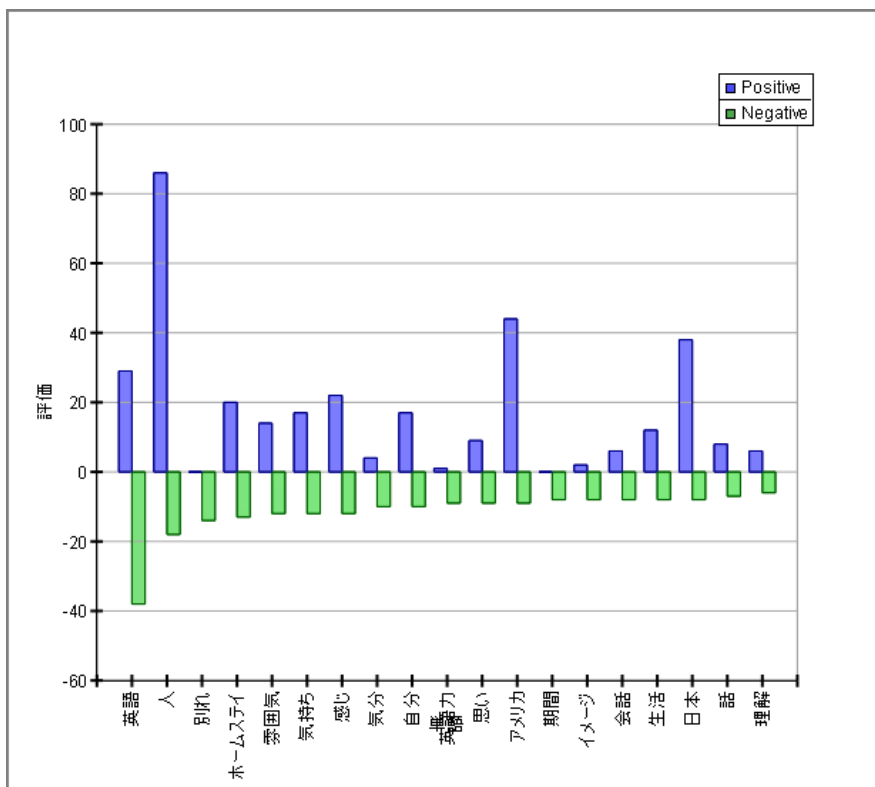


図9. 不評語ランキング

### (5) 対応分析

対応分析によって、単語による各キャンパス、年度間の関係を2次元上に配置したのが図10である。1991（松江キャンパス）と2008（浜田キャンパス；実際は2007年度だが、出雲キャンパスの2007年度と区別するため便宜上2008としてある）と出雲キャンパスの6年分のかたまりに分かれているようだった。これはプログラムの違いなどに起因するのかもしれない。どの年度でも多く出てきた「アメリカ、違う、見る」という語は、ある年度だけに特に出ているという特徴がないことから原点近くにあるといえる（図10）。

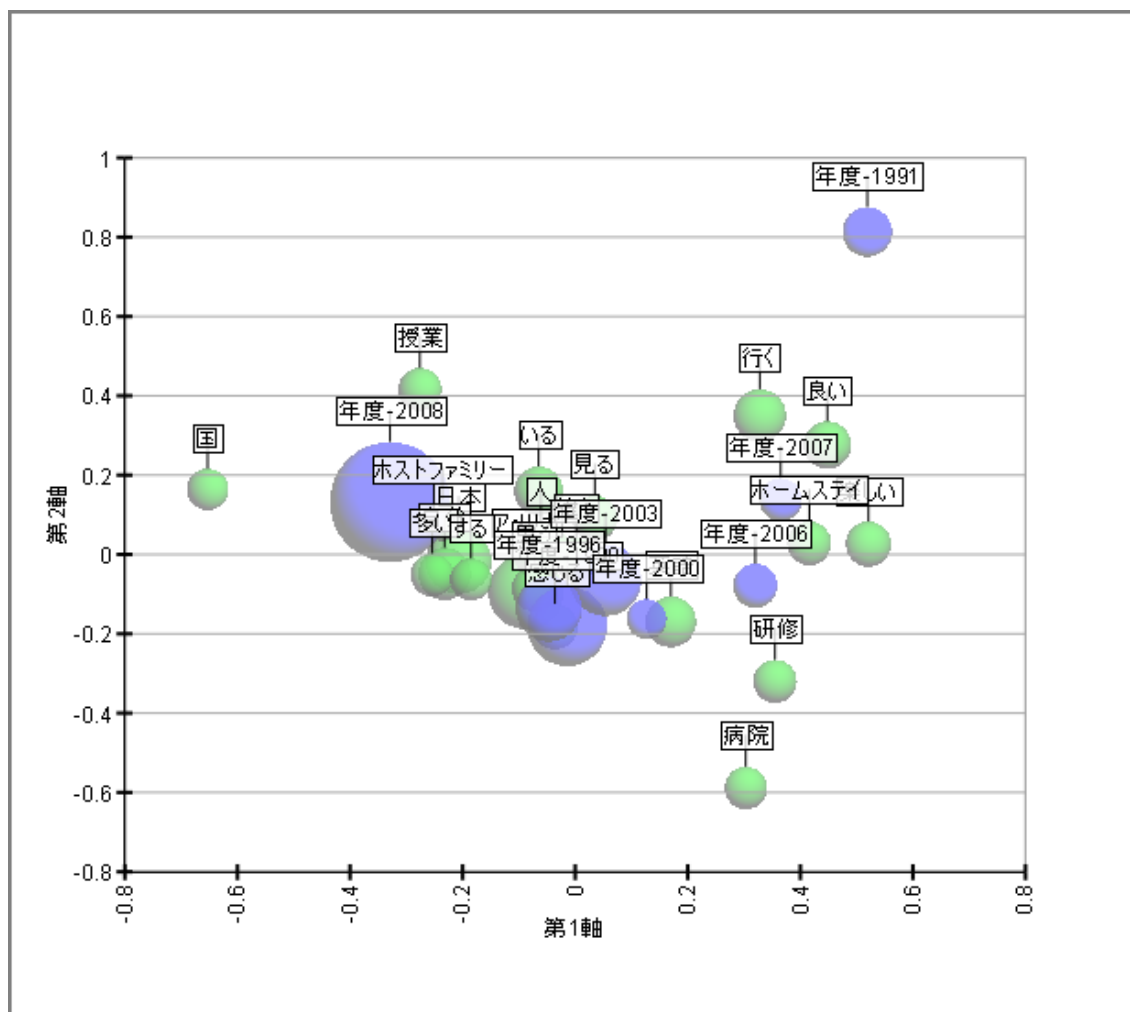


図10. 対応分析

### (6) 希望、不満の分析

報告書に表れた学生たちの希望、要望を探ってみた（図11）。これによると、行きたい、言いたい、参加したい、感謝したい、勉強したい、話したい、など積極的な意欲が示されている。また、困難や不満に感じていることは、図12のようであった。つまり、分からない、知らない、通じない、話せない、できない、などであった。



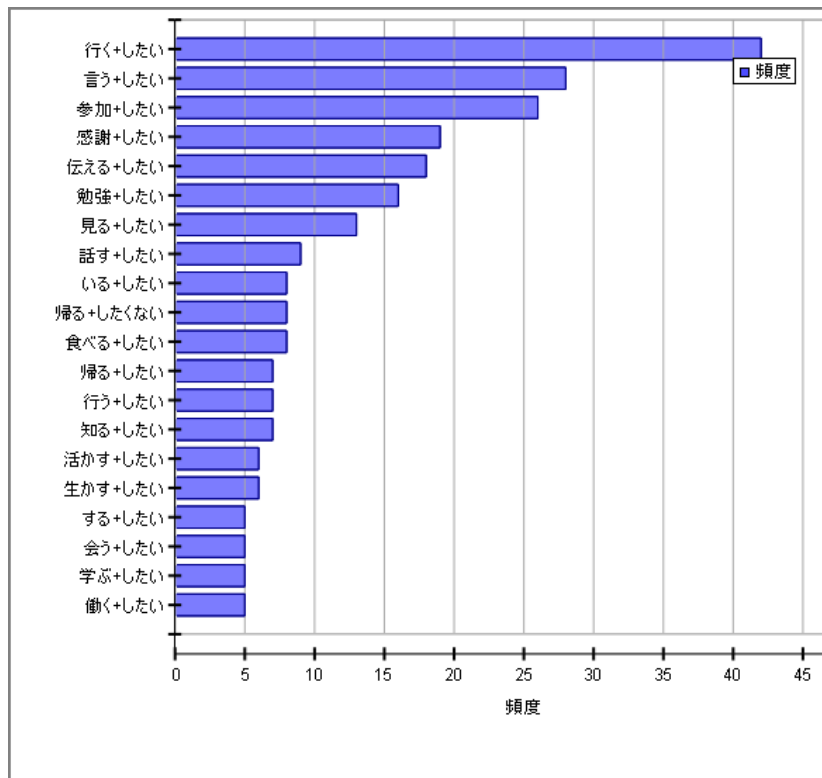


図 11. 要望の頻度

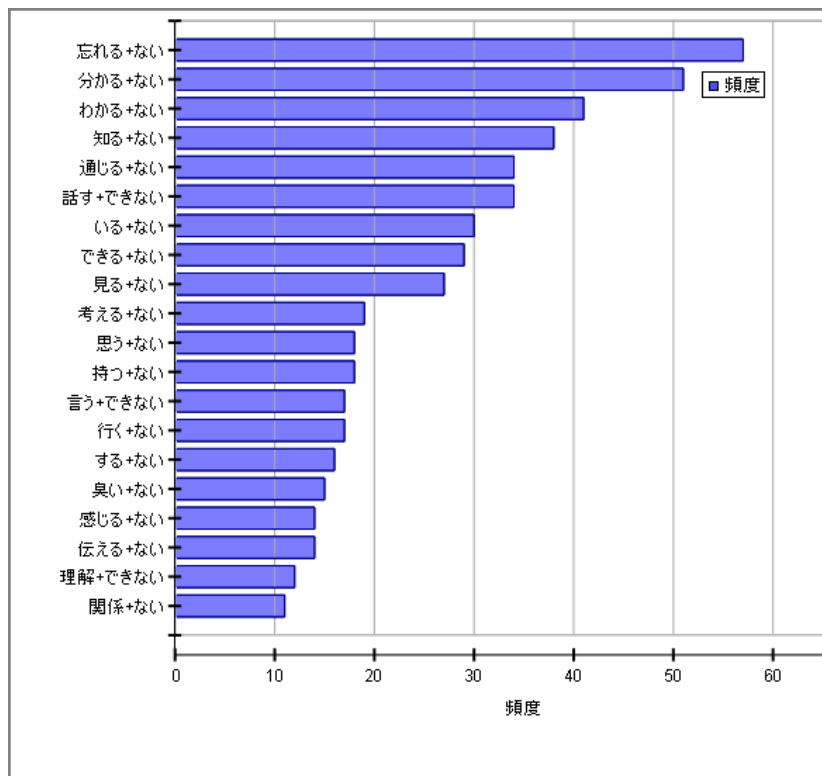


図 12. 困難、不満等の頻度

### (7) 特徴分析

次に、各年度別に、特徴的に出現していた単語を抽出してみた。まず、松江キャンパス(1991)は、「エレンズバーグ、CWU、Yakima、アメリカ人学生」(図 13)などであり、浜田キャンパス 2008 は「日本、アメリカ、自分、国、文化」(図 14)などであった。次に出雲キャンパスの 1996 は「アメリカ、日本、車椅子」など、1998 は「アメリカ、ワナチ、病院、日本など」(図 15)、2000 は「ワナチ、クライアント、講義」など、2003 は「アメリカ、日本、専攻科」など、2006 は「ワナチ、研修、楽しい」など、2007 は「笑い、良い、行く、楽しい」などが特徴的であるといえる。

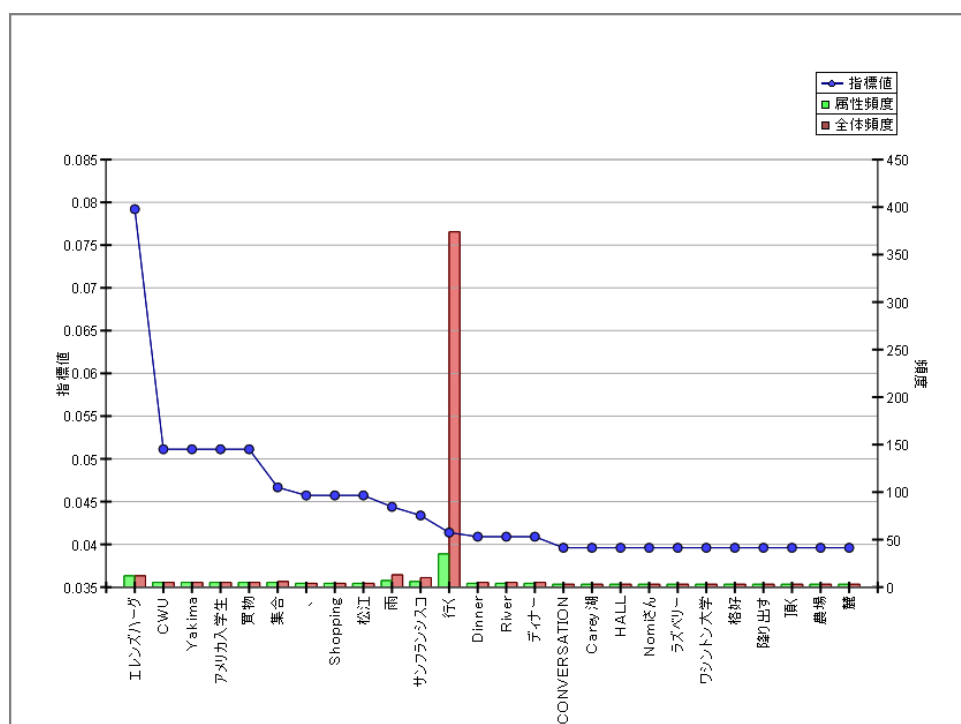


図 13. 松江キャンパス(1991)の特徴語

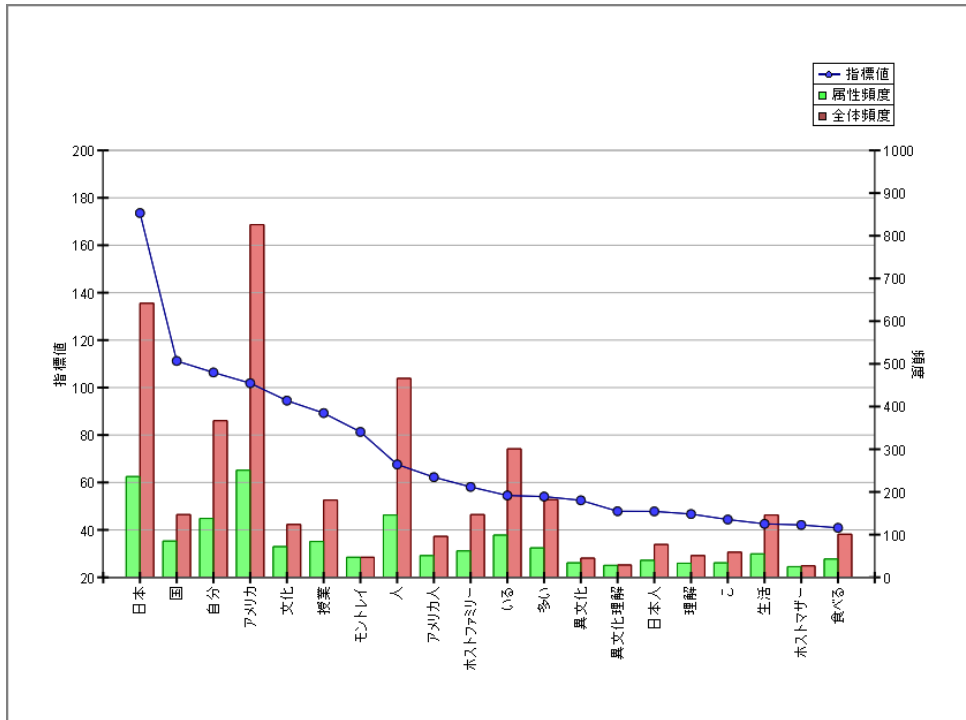


図 14. 浜田キャンパス(2007)の特徴語

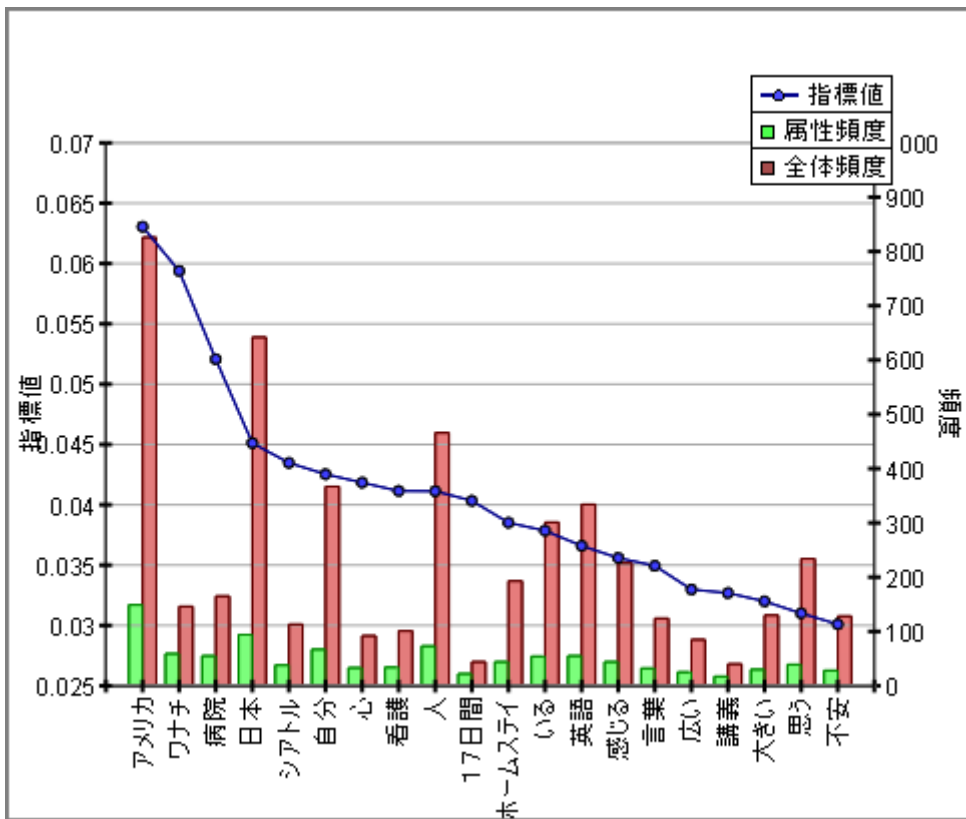


図 15. 出雲キャンパス (1998) の特徴語

#### 4. 考察

単語頻度解析によると、すべての報告書を合わせて最も頻度の多い単語は、「アメリカ」、「日本」、「人」、「英語」などであった。アメリカ研修後の報告書なので、これらの言葉が上位に来るのはうなずける。また、単語の品詞頻度解析の結果、名詞、動詞、副詞、形容詞などに次いで、感動詞の頻度の多いのが特徴的であった。これは学生が海外研修によって感動的な体験をしてきたということが推測できる。海外での実地体験をすることが、現在最も欠けていると思われる「感動を与える教育」になっているのかもしれない。このことから、「だまされたと思って、感動するからとにかくいってみなさい」などと言って動機づけてやるのがよいのかもしれない。係り受け分析からも「参加しては良かった、喜びも大きかった」、そして「別れが辛かった」などというメッセージが代表的なものであった。このように、単語分析の結果は直接的な国際交流の重要性を物語っていると思われる。

次に自由記述文に散在している好意的な印象と非好意的な印象を選別するために評判分析を行った。海外研修を終えて、多数の学生が、楽しかった、きてよかった、よい体験をした、などといった、満足感や達成感、喜びなどの肯定的感情の表現をしていた。良いイメージで語られている単語には、アメリカ、アメリカ人などがあげられている。しかし非好意的なイメージで語られている単語に、英語や英語力が挙げられていることは注目に値する(図8、図9)。アメリカやアメリカ人に、好意を持ってどんどん話してみたいのに、自分の思うように、意思伝達できない不満、いらだちが表れているようである。例えば、「ホームステイで切実に感じたのが、自分の英語力の低さだった。電子手帳と筆記用具と紙をいつも持ち歩き、わからなければ筆談したりして、わかるまでには疲労も感じたが、通じたときの喜びはとても大きかった。もう少し英語力があればもっといろいろなことが話せたのにと残念にも思う。」というような感想が多くみられた。要望、不満の分析結果(図11、12)からも、積極的な意欲が示されているにもかかわらず、短期間には英語力が伸びない不満も示されている。Ward, Bochner & Furnham (2001)も、短期海外旅行は、愉快で、楽しく、面白い経験であるが、同時にまたストレスをも伴う経験であると述べている(p. 130)。このジレンマを解決してやるのが、今後の課題であると言えよう。つまり、海外研修に行く前に、ある程度の英語力を身に付けさせていく必要があると思われる。

報告書の分析から、たとえ短期間のあっても、異文化との出会い、交流が、学生に様々な喜びや感動をもたらしていることが明らかにされた。

#### 文献

- Ward, C., Bochner, S., & Furnham, A. 2001 *The Psychology of Culture Shock* (2<sup>nd</sup> ed.)  
Routledge
- 数理システム Text Mining Studio 操作マニュアル (v. 3.0)

**キーワード**：短期海外研修、自由記述、テキストマイニング、単語頻度解析

(IIZUKA Yuichi, KANE, Eleanor Anne, KODAMA Yoko, and MATSUMOTO Ichie)

## 要約

短期海外研修に対する感想文の単語頻度解析の結果、「アメリカ」、「日本」、「とても」、「人」、「英語」などの感想が得られた。また文章分類の結果「自分」、「アメリカ」、「英語」、「人」、「日本」の5つのクラスタがみられた。全般的にみて、学生たちの短期海外研修の評価としては良好であったと思われる。特徴的な印象語であった英語や英語力に対して肯定的、否定的な相反する評価がありアンビバレントな傾向がみられ、英語に対して好意的な反応がある反面、思うように英語を使えないいらいらだちがうかがえた。