

テキストマイニングによる 若者の観光需要の推定 ～大学生のライフスタイルに着目して～

田中 良典[†] 井出 明^{††}

本研究は、大学生の観光ニーズを推定することを目的としている。実験として、大学生に対し、希望する旅行についての自由記述アンケートとライフスタイル調査を実施した上で、得られたデータを元に、テキストマイニング分析をおこなった。その結果、属性ごとで特徴的な語や対応分析から、女性は歴史好きな層を中心に世界遺産をゆっくり巡る旅行、男性は理系が一人旅、文系は個人ツアーを好むなど具体的な観光ニーズを推定することが出来た。

Estimating Travel Demands of Youths and Students by Text Mining ---Focusing on Lifestyle of University Students---

Yoshinori Tanaka[†] Akira Ide^{††}

The purpose of this research is to seek for needs of tourism among university students. We firstly conducted a free writing questionnaire and survey about the lifestyle of university students and then conducted text-mining analysis. From the analysis, we have discovered some interesting features that depended on personal attribution. For example, "Women, who love history, are fond of slow travels to see the world heritages," "Male science course students like to travel alone," and "Arts students choose personal tours" were some of the interesting inclinations obtained from the analysis.

1. 序論

1.1 研究の背景と目的

近年、若者の旅行離れが問題視されている。観光庁では、特に大学生の成長性が高いと考え、大学生の観光推進を重点分野として設定している¹。

若者の旅行離れに関する先行研究では、そのほとんどが若者はなぜ旅行しなくなったのかという要因分析が多い²。一方で、どのような旅行だったら行くかという観光ニーズについての研究は少なく、これらの先行研究から得られた阻害要因を観光ニーズとして捉えることも出来るが、若者の潜在的観光ニーズとは言い難い。

一般的に、仮説検証研究から得たデータは、当該の理論が「まことらしい」かどうかを判断する資料にはなるが、新しい理論の土台となることは少ない³。選択式アンケート調査において、人々の潜在需要を掘り起こすには限界がある。それは、アンケートを受けた本人にさえ自分のニーズに気づいていないことが原因であり、その人の意識に頼るような研究方法では取り出せないことが多い。しかし、自由記述を基礎データとすることで、選択式アンケートでは取り出せない潜在的な観光ニーズを推定することが可能であり、その手法にテキストマイニングがある。

テキストマイニング (text mining) あるいはテキストデータマイニング (text data mining) は、データマイニングの一種で、莫大なテキストデータを形態素解析^aすることで単語に分解し、そこから情報や知識を取り出す技術である。Hearst (1999) によれば、データベース検索や情報検索で得られる既存情報ではなく、そこからでは得られない新しい知識をテキストデータから発見することができるとしている⁴。よって、今まで得られていない若者の観光ニーズを発見できるのではないかと考えた。

これらを踏まえ、本研究では新知の大学生観光ニーズを明らかにすることを目的とする。

ここで、本研究での「若者」は「大学生」と定義した。本節で用いられている「若者」という言葉の多くは20歳代と設定されており、社会人も含まれている。しかし、志向について大学生に通ずる部分も多いと考え、1.2で述べる観光ニーズについても本研究の大学生観光ニーズとして比較することとした。

電通国際情報サービス
Information Service International-Dentsu, Ltd.
追手門学院大学
Otemon Gakuin University

^a形態素解析は、文章を構成する語を形態素(Morpheme)に分解する作業で、言語で意味をもつ最小単位である品詞を判別する作業である。

1.2 研究方法

本研究では、大学生 218 名を対象にライフスタイル調査と競合消費項目調査、観光ニーズに関する自由記述アンケート調査を実施した。

研究手法はテキストマイニング、及び、計量テキスト分析で、分析ソフトには樋口耕一氏（立命館大学）の KH Coder を用いることにした。KH Coder は、テキストマイニング分析で得られた結果で終わらせるのではなく、その結果から分析者自身の仮説を立て、語をいくつかのカテゴリに分類し、再度分析を試みて解釈を深める計量テキスト分析のためのコーディングと呼ばれる機能がある。よって、分析手順は樋口氏の示す分析手順に従うことにし、図 1 に分析手順を示す。

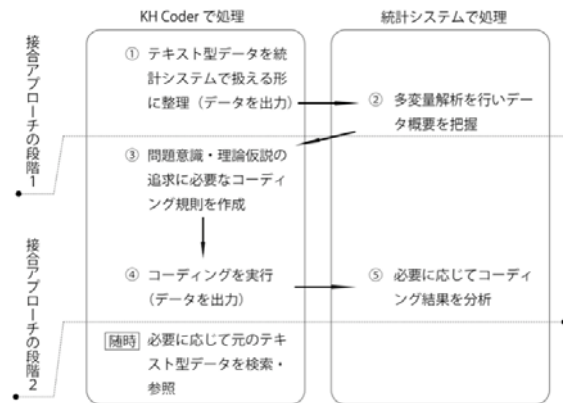


図 1 分析手順
(樋口耕一 2005 より筆者作成)

段階 1 では、データ中から語を自動的に取り出して、その結果を集計・解析する。これによって、分析者の予断をなるべく交えずに、データの特徴を探ったり、データを要約できる。具体的な作業は、自然言語処理と統計処理である。

段階 2 では、分析自身が語をグルーピングすることで、コードを作成し、データ中からコンセプトを取り出す。その結果の集計・解析は段階 1 と同様である。

本研究では、テキストマイニング分析（図 1 の①②）と計量テキスト分析（図 1 の③④⑤）を行うこととした。

また、新知の大学生観光ニーズを求める為に、既知の観光ニーズを定義する。ここでは、若者の旅行離れの先行研究から得られる観光ニーズを既知ニーズとして定義した。

阻害要因は要因を満たすあるいは、要因を補える旅行と解釈することで観光ニーズとし、例えば「お金がかかる」を「安い旅行」と変換して求める旅行とした。表 1 のように先行研究から得られた観光ニーズをまとめ、既知とする観光ニーズを表 2 に示す。これらと得られた観光ニーズを比較することで新知の観光ニーズを明らかにする。

表 1 既知とする観光ニーズ

男性	女性	全般
趣味のための旅行	リフレッシュ旅行	安い旅行
友人との旅行	家族旅行	精神的・肉体的に負担のかからない旅行
一人旅	カップル旅行	旅行の価値を明確にした旅行
		事前学習旅行
		安心・安全な旅行
		リフレッシュ旅行
		新しいフロンティア発見旅行
		新知開拓旅行
		リラックスできる旅行
		のんびりできる旅行
		伝統文化を巡る旅行
		知識・教養を高める旅行

2. ライフスタイル分析から見る若者の観光観 -属性に着目して-

2.1 調査概要

まず具体的調査内容について述べておきたい。本調査では、大学生の観光ニーズを推定するために、平成 22 年度後期において首都大学東京で開講された“ツーリズム産業論”を受講していた大学生 218 名を対象に①観光②ライフスタイル③競合消費項目調査の 3 つのカテゴリでデータの収集を行った。対象者の性別と学年を図 2、図 3 に、質問内容を表 2 にそれぞれ示す。観光に関する調査とライフスタイルに関する調査は 3.3 以降に、競合消費項目調査は次節でテキストマイニングをおこなう。

b 当該講義の内容は、毎回一線級の実務家を招き、学生に観光産業の最前線について語っていただくという内容であった。



図2 性別

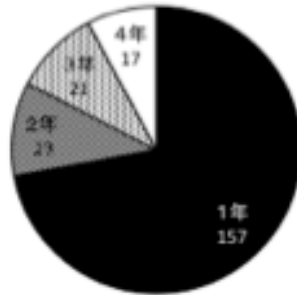


図3 学年

表2 質問内容

調査内容	質問内容	回答形式
観光に関する調査 10月上旬実施	観光に対する意識 1. 観光するのは好きである 2. どちらかというと観光するのは好きである 3. どちらかというと観光するのは嫌いである 4. 観光するのは嫌いである	4段階尺度選択
	なぜ観光するのが嫌いか (観光が嫌いと答えた人のみ)	自由記述
	どうすれば観光するようになるか (観光が嫌いと答えた人のみ)	自由記述
	行きたい旅行とはどのようなものですか	自由記述
ライフスタイルに関する調査 12月上旬実施	あなたは理系だと思えますか?	5段階尺度選択
	あなたはインドア派だと思えますか?	5段階尺度選択
	あなたは社交的だと思えますか?	5段階尺度選択
	歴史は好きですか?	5段階尺度選択
	年間旅行回数	自由記述
	1日あたりのTVの視聴時間	7段階選択
	1日あたりのインターネット利用時間	7段階選択
	1か月あたりの読書量	7段階選択
現在の旅行に対する意向	4段階尺度選択	
競合消費項目調査 12月上旬実施	3万円の旅行券があったら 1. 旅行に使う 2. 換金して自由に使う ※旅行券は手数料なしで換金できるものとする	2択
	3万円の範囲内でどんな旅行をするか (旅行に使うを選択した人のみ)	自由記述
	3万円で何を買うか (自由に使うを選択した人のみ)	自由記述

2.2 競合消費項目調査結果

2010年の国内旅行の平均消費額 31,500円 (JTB 広報室 2009) という推計から、3万円の旅行券が入手可能であると仮定して、その3万円の旅行券を旅行券に使うか、換金して自由に使うか、自由に使う場合は何に優先して使うかについての回答をまとめたものが図4である。自由に使うと答えた学生は、ファッションや趣味に使いたいという層が多かった。ここで、旅行券を現金に換金し、3万円かからないディズニーランドやスキーなどと食事、お土産などを組み合わせて書いていた学生を自由旅行とした。

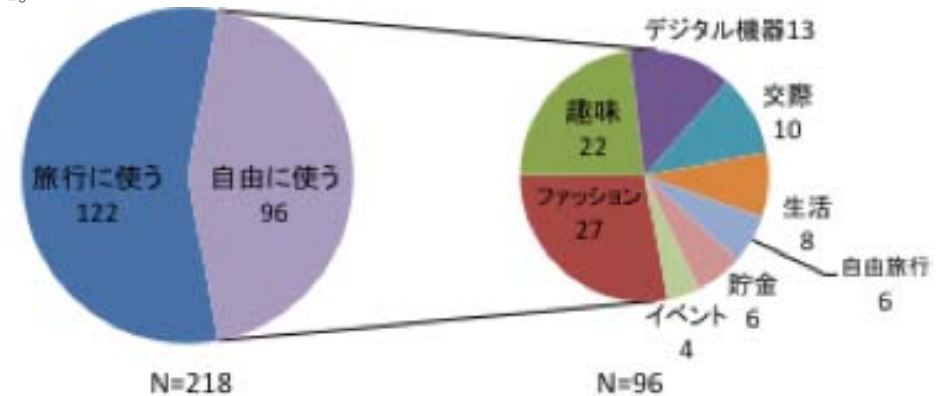


図4 3万円あったら何に使うか

また、旅行に使うと答えた学生の行きたい旅行の記述についてテキストマイニングをおこなった。ここでは地名に着目し、その頻度を降順に並べたものが表3である。

表3を見ると、一番回答の多かった地名は「北海道」「京都」であった。また、3位に「韓国」という近場の海外が上位にきており、続いて「九州」「沖縄」など東京を基準に考えると出来るだけ遠くの場所を回答する学生が多いのではないかと考える。また、「箱根」「草津」など具体的な地域名まで書かれているところは大体が温泉地だということが分かる。よって、3万円で国内旅行というと温泉を思い浮かべる傾向が出ているのではないかと考える。また、アメリカやアルゼンチンなどの海外は3万円で行けるとは考えにくい、「バリ」と答えた学生の記述からは3万を宿泊費や航空運賃の足しにしてできるだけ遠くに行きたいという記述も見られた。このことから、旅行費用に余裕が生まれれば、その分学生は旅行範囲を広くするのではないかと考える。

表3 旅行券 (3万円分) があったら行きたい場所

抽出語	出現件数	抽出語	出現件数	抽出語	出現件数
北海道	11	アメリカ	1	台湾	1
京都	11	アルゼンチン	1	大分	1
韓国	9	インドネシア	1	中国	1
日本	9	オーストラリア	1	中南米	1
九州	8	グアム	1	鳥取	1
沖縄	7	サイパン	1	徳島	1
箱根	7	ソウル	1	島根	1
広島	5	パリ	1	東北	1
青森	4	ブラジル	1	湯布院	1
石川	3	メキシコ	1	徳島	1
草津	3	茨城	1	奈良	1
アジア	2	横浜	1	熊本	1
鎌倉	2	屋久島	1	日光	1
関東	2	岩手	1	熱海	1
新潟	2	宮城	1	図書館	1
大阪	2	徳島	1	八戸	1
知床	2	近畿	1	浜松	1
長崎	2	熊本	1	福山	1
福岡	2	兼六園	1	兵庫	1
名古屋	2	弘前	1	別府	1
四国	1	香川	1	清里	1
秋田	1	讃岐	1	静岡	1
上海	1	神戸	1		
城崎	1	仙台	1		

※実際は兼六園という記述だった

2.3 男女×属性で特徴的な語

表4は、男女別に見る文系、理系ごとで特徴的な語である。Jaccard係数は語と属性の関連を示す指標で、数値が大きいほど関連が強い。女性を見ると『F文系』は「人々」「友達」など、友達と旅行先の雰囲気を楽しみたいというような語が出ているほか、「海外」なども特徴的である。一方、『F理系』は「安い」「堪能」「価格」などが上位にあることからコストパフォーマンスに優れた旅行を重視する傾向があると考えられる。また、男性を見ると『M文系』『Mやや文系』では、「個人」と「ツアー」が特徴で出ており、一人旅でなく個人ツアーなどを重視しているのではないかと考える。一方、『M理系』は「名所」「一人旅」が特徴的で、名所を一人で巡る旅行を重視するのではないかと考える。

表5は、男女別に見る歴史好き、歴史嫌いな層で特徴づけられる語である。女性を見ると、『F歴史嫌い』では、「体験」「手配」などが特徴的で、「手配」について原文を見ると、自分で手配するという使われ方がされ、その文意に自分で手配するのは不

安だという記述は見られなかった。『F旅行好き』では「世界遺産」「海外」「建造物」などが特徴的で、海外の世界遺産を巡る旅行を求める傾向にあると考える。また、男性を見ると、『M歴史嫌い』では「実際」「お金」など金銭的な面を重視する傾向にあると考えられる。一方、『M歴史好き』では「土地」「日本」「文化」「歴史」など、日本の歴史や文化に関心がある傾向にあると考える。

表4 男女×文系・理系で

Fやや文系 N=32		F文系 N=37		F理系 N=10	
海外	0.191	人々	0.196	安い	0.177
思う	0.161	いう	0.183	堪能	0.143
ない	0.155	友達	0.167	価格	0.143
ある	0.148	できる	0.165	電車	0.143
旅	0.146	感じる	0.157	世界遺産	0.136
好き	0.143	思う	0.144	魅力	0.120
できる	0.135	現地	0.136	グルメ	0.118
国	0.130	ゆっくり	0.136	食べ物	0.118
文化	0.125	人	0.130	さまざま	0.118
訪れる	0.125	なる	0.129	土地	0.113
Mやや文系 N=34		M文系 N=35		M理系 N=22	
できる	0.160	ない	0.151	名所	0.143
ない	0.152	場所	0.145	一人旅	0.133
個人	0.150	ツアー	0.118	街	0.130
体験	0.138	日本	0.115	良い	0.111
見る	0.136	感じる	0.113	する	0.110
ある	0.136	宿	0.111	ない	0.108
場所	0.119	世界	0.104	癒れ	0.107
好き	0.119	非日常	0.098	持つ	0.107
行う	0.115	考える	0.097	国	0.105
楽しむ	0.109	心	0.095	思う	0.101

※数値はJaccardの類似性測定

表5 男女×歴史の嗜好で

Fやや歴史好き N=36		F歴史嫌い N=19		F歴史好き N=28	
できる	0.176	体験	0.159	世界遺産	0.194
ゆっくり	0.167	手配	0.158	海外	0.183
思う	0.155	旅	0.136	建造物	0.162
土地	0.149	記憶	0.136	思う	0.146
ある	0.143	不安	0.130	いい	0.143
旅	0.136	いう	0.119	訪れる	0.139
好き	0.133	求める	0.115	する	0.130
各地	0.128	人	0.113	日常	0.130
リフレッシュ	0.128	見る	0.111	できる	0.130
魅力	0.125	いい	0.111	なる	0.129
Mやや歴史好き N=36		M歴史嫌い N=20		M歴史好き N=47	
目的	0.167	実際	0.161	ある	0.225
ない	0.159	お金	0.148	土地	0.185
思う	0.155	ある	0.122	できる	0.179
文化	0.146	ない	0.121	日本	0.177
楽しむ	0.125	各地	0.120	ない	0.169
現地	0.121	とる	0.120	感じる	0.167
違う	0.120	良い	0.118	見る	0.158
ない	0.117	癒れ	0.115	人	0.155
得る	0.116	場所	0.111	文化	0.154
ん	0.106	自然	0.108	歴史	0.154

※数値はJaccardの類似性測定

2.4 可視化した語と語の関係

図4は、共起ネットワーク図である。円が大きいほど出現件数が多く、語と語が結ばれている線が太い方がよい関連が強いことを示している。しかし、円の距離は関係がないため、語の島同士の関係はない。また、色は中心性を表しており、水色→白→ピンクの順で高くなる。この図で観光ニーズを示すような語と語のつながりは、「温泉」→「料理」、「移動」→「楽しめる」、「安心」→「安全」、「お金」→「安い」、「食べ物」→「景色」、「視野」→「広げる」、「非日常」→「求める」、「個人」→「ツアー」などが考えられる。

c ここでは媒介中心性の指標を用いており、他のエンティティ間をつなぐパスが多く現れるエンティティほど重要であることを示す。(フェルドマン 2010)

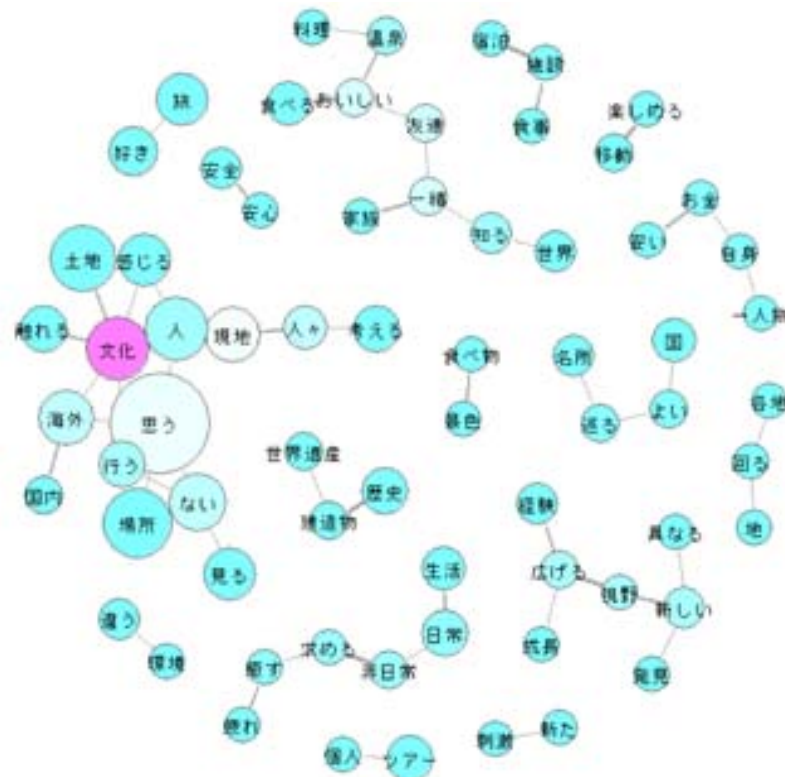


図5 共起ネットワーク

2.5 小括

本章では、観光に関する調査とライフスタイル調査、競合消費項目調査を実施し、旅行の競合や3万円あったら行きたい旅行先などを明らかにした。旅行の競合消費項目では、ファッションや趣味にお金を使う層が多かった。また、3万円あったら行きたい場所には「北海道」「京都」「韓国」が上位にあり、3万円で作れる限り遠くを求める傾向にあったが、「箱根」「草津」など関東近郊の温泉地も多かった。また、行きたい旅行とはどのようなものかという自由記述アンケートから属性に着目したテキストマイニング分析をおこない、歴史好きな女性が海外の世界遺産を巡る旅行を求めるなど具体的な観光ニーズを把握することができた。

しかし、ここで得られた特徴的な語については Jaccard 係数を基準にした為、出現

件数の少ない語も対象となり、傾向が偏ることも考えられる。よって次章では、ある程度の出現件数があり、さらに属性同士の連関を考慮した対応分析と、より特徴をとらえる為に計量テキスト分析をおこなう。

3. ライフスタイル分析から見る若者の観光観 -対応分析を中心に-

3.1 対応分析から見る語と属性の関係

図6は、歴史好きと歴史嫌いを男女別で分けた対応分析の図である。属性名にあるMは男性、Fは女性を表す。歴史好きな層の女性をみると、『F 歴史好き』『F やや歴史好き』は共に近くにあり、「世界遺産」「ゆっくり」が近くにあることから世界遺産をゆっくり見たいという傾向があると考えられる。また歴史好きな男性の層では、『M 歴史好き』『M やや歴史好き』の近くに「個人」「計画」などがあることから、個人で回りたい傾向が強く、旅行の計画を重視する傾向もあるのではないかと考える。ここで、この「計画」を調べると、「計画」は計画を立てる派と計画を立てない派の両方の意見がほぼ同数いることがわかった。

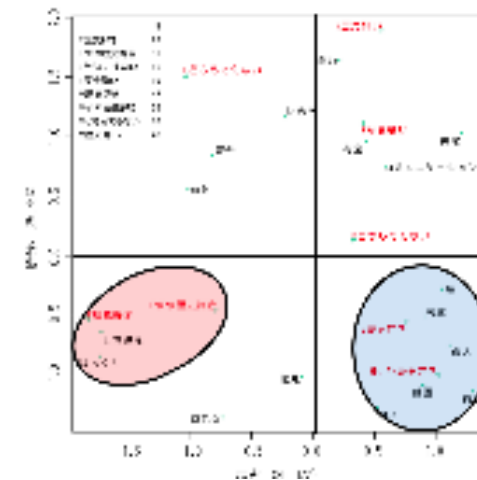


図6 対応分析に見る男女別・歴史に対する嗜好別の特徴

図7は、理系、文系の層を男女別に分けた対応分析の図である。まず、理系から見てみると、『M理系』では「思い出」「自由」が近くにきていることから自由な旅行スタイルを重視する傾向があると考えられる。また、『F理系』では突出して「安い」が近くにきており、理系女子は旅行費用をあまりかけたくない傾向が強いと考えられる。続いて、文系を見てみると『Mやや文系』では「個人」が、『F文系』では「一緒」「友達」が近くにあることから、男性は個人旅行を、女性は友達などと一緒の旅行を好む傾向にあると考えられる。また、『F文系』では「安心」「安全」などもあり、友達が一緒にいることにより安心したり、また安全な旅行を重視する傾向にあると考えられる。加えて、『M文系』『Fやや文系』の「知る」という語から、理系と比べて文系の方が旅行に対して、自己成長や非日常な要素をより求める傾向にあるのではと考える。

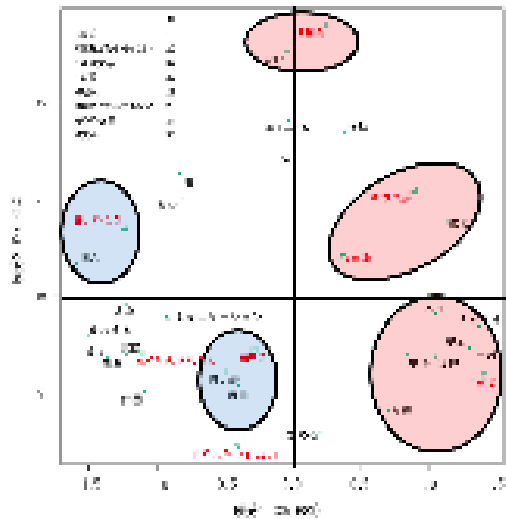


図7 対抗分析に見る男女、文系・理系別の特徴

3.2 計量テキスト分析の為の階層的クラスター分析

前節の対応分析では、属性や性別ごとに特徴的な語が出てきており、それを元に分析を試みた。本節では、計量テキスト分析の為の階層的クラスター分析をおこなう。階層的クラスター分析をおこなうことで、これまでに共起ネットワークなどから確認した概念や使われ方の似た傾向にある語をまとめる。その結果、グループの意味解釈が可能な26のクラスターに分けることができ、それぞれのクラスターのコード名と累積件数を示した表が表6である。ここで、cl.9のよく出る語について、「国内」、「海外」

「思う」「考える」などは旅行に関するアンケートとして非常によく使われる語である。このような語は分析時でもあまり傾向が出る語ではないと判断し、cl.9は分析対象から外した。よって、得られた25のクラスターについてコーディングをおこない、再度対応分析をおこなう。

表6 コード名と累積件数

コード名	語	累積件数
cl1	世界遺産(15) 歴史(28) 建造物(15)	58
cl2	文化に触れる 土地(49) 文化(58) 触れる(28)	145
cl3	現地交流 現地(38) 人(50) 人々(24)	112
cl4	おいしい料理 料理(9) 食べる(24) おいしい(18)	51
cl5	行ったことない場所 行方(34) 場所(40) ない(50)	144
cl6	名所 スポット(10) 有名(16) 名所(18)	62
cl7	コミュニケーション 特(14) 心(11) コミュニケーション(10)	90
cl8	計画 自由(14) 計画(11) しっかり(9)	34
cl9	よく出る語 言う(12) 歩く(10) 良い(18)	273
	好き(32) 考える(33) 海外(43)	
	よい(14) 国内(13) 思う(68)	
cl10	安心・安全 安心(10) 安全(17)	27
cl11	魅力 魅力(18) 感じる(44) 訪れる(13)	89
cl12	安い お金(11) 安い(10) 国(20)	41
cl13	個人ツアー ツアー(22) 個人(12)	34
cl14	思い出に残る 思い出(12) 残る(13)	25
cl15	友達と一緒に 楽しい(24) 友達(19) 一緒(11)	54
cl16	自己成長 自身(12) 成長(12) いい(12)	70
	一人旅(12) 経験(22)	
cl17	食のバラエティ 理想(11) さまざま(9) 食事(12)	63
	食(10) 価値(9) 様々(12)	
	体験(32) 景色(10) 見る(41)	
cl18	景色 多い(12) 地(10) 実際(16)	138
	感動(11)	
cl19	非日常(体験) 環境(9) 違う(20) 味わえる(11)	60
	味わう(12) 雰囲気(8)	
cl20	世界を知る 世界(18) 知る(18) 興味(24)	69
	持つ(9)	
cl21	移動を楽しむ 風景(14) 移動(12) 楽しめる(10)	36
cl22	温泉 温泉(10) ゆっくり(13) 目的(20)	119
	離れる(28) 自然(21) 楽しむ(22)	
cl23	のんびり過ごす 過ごす(10) のんびり(10)	20
cl24	非日常(体験) 求める(10) 癒す(12) 刺激(11)	111
	非日常(10) 出来る(15) 生活(29)	
	日常(24)	
cl25	視野を広げる 視野(10) 広げる(13)	23
cl26	新しい発見 異なる(10) 発見(10) 新しい(14)	34

3.3 対応分析から見るコードと属性の関係

図8は、現在の旅行に対する意向の男女別の対応分析の図である。ここで、旅行の意向について「行きたい」と答えた層が『積極派』、「行きたいが行けない」が『希望派』、「なんとなく行きたいが行かない」が『消極派』を表している。図のように線を描くと右が積極派、左上が消極派、左下が希望派に分かれている。消極派に着目すると、『F消極派』は「友達と一緒に」など複数人での旅行を重視し、『M消極派』では「非日常(体験)」が近くにあり、非日常を重視する傾向にあると考える。また希望派は、男女においてそれほど違いはないように見えるが、希望派を総じて見ると「思い出に残る」「行ったことない場所」など、旅行に行ったことに対する証としての記憶を重視する傾向があると考えられる。そして、積極派では『F積極派』が「文化遺産」「文化に触れる」など、やはり女性は文化遺産などを重視する傾向があることがわかる。逆に、『M積極派』は「名所」と布置が一致することから文化遺産などは関係なく、いわゆる名所と呼ばれる場所を重視し、新たな発見から視野を広げていける旅行を好むのではないかと考える。

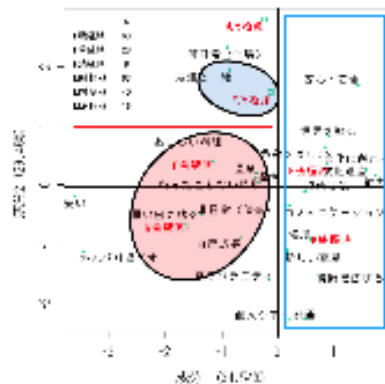


図8 対抗分析に見る旅行の意向別、男女別の特徴

3.4 得られた観光ニーズのまとめ

第2章と第3章で得られた結果についてまとめをおこない、表7で得られた観光ニーズを示した。分析全体を通して明らかになったことは以下の3点である。

1. 男女や属性で重視する点が異なる
2. 年間旅行回数や旅行の意向で重視する点が異なる
3. 性別と属性のクロス分析で違いがより明確になる

表7 得られた観光ニーズ

属性	男性	女性	全般
全般	何かがあることを重視	何かができることを重視 文化、歴史などを重視する	
体系	自由な旅行スタイルを重視 名所を一人旅	コストパフォーマンスに優れた旅行を重視 安い旅行を重視	街、ドライブを重視
文系	個人ツアーを重視 個人旅行を重視	友達と一緒に旅行するのを重視	目的地に対するニーズが多い 自己成長や非日常を重視
インドア	受動的な旅を重視	少人数でリフレッシュを目的とした旅行	のんびり過ごす 非日常を重視
アウトドア	主体的な旅を重視	非日常を重視	主体的な旅を重視
社交的	非日常を重視 自己成長を重視	リフレッシュが目的 世界遺産などの文化に触れる 非日常を重視	文化や人に触れる 自己成長を求める
非社交的	その土地ならではの文化に関心 お金をかけずに刺激を得られる一人旅を重視	精神的な負担をかけない 料理や風景を楽しむ のんびり安全な旅行 リフレッシュ目的	自己啓発を重視 リフレッシュが目的 安い個人ツアー 現地交流を重視
歴史好き	日本の歴史や文化に関心 個人で回りたい。計画を重視。 ※しっかり計画を立てると計画を立てない人が半々	海外の世界遺産(建造物)を巡る旅行 世界遺産をゆっくり見たい	文化を重視 世界遺産、建造物を重視 非日常を重視
歴史嫌い	全概面を重視		リフレッシュを重視
旅行回数が多い		旅行に自己成長を期待	非日常を重視
旅行回数が少ない	旅行に自己成長を期待	文化や人を重視 安全な旅 友達と一緒にの旅	
旅行回数が多い	ガイドブックに載っていない場所 海外志向	海外志向、文化遺産に興味	自己成長を求める
TVを見ない	コミュニケーションがとれる旅行	リフレッシュが目的	刺激を重視
TVをあまり見ない			体験を重視
TVをよく見る	リフレッシュが目的 テレビで見た所を自分の目でみたい	魅力を探す	体感を重視
インターネットをしらない			食を堪能
インターネットをあまりしらない			インターネットからは得られない非 日常体験を求める 安い旅行を重視
インターネットをよくする			
読書をしらない			
読書をあまりしらない			
読書をよくする			移動にも楽しさを求める
3万円あったら旅行に使う	個人旅行を重視 現地の人とのコミュニケーション 温泉を重視	海外や文化に関心 友達と一緒に旅行するのを重視 移動時間も楽しむ	文化や歴史に関心
3万円あったら自由に使う	安い旅行	現地の生活に関心 安心・安全が確保された旅行	
積極派	一人旅を重視 名所を重視	文化遺産を重視 複数人で安心した旅行	自己成長を求める 文化に触れる
希望派	お金を重視		お金をかけずにのんびりリフレ ッシュできる旅行を重視 思い出に残る旅行
消極派	非日常を重視	旅行に不安感あり リフレッシュが目的 友達と一緒に旅行するのを重視	非日常を重視

これらの観光ニーズを要約し、さらに出現傾向の高い順に分けたものが表 8 である。男性よりも女性の方が出現傾向の高いニーズが多かった。

表 8 出現傾向の高かった観光ニーズ

	男性	女性	全般
1	一人旅(理系)や個人ツアー(文系)など一人の旅行	世界遺産(建造物)をゆっくり巡る旅	非日常体験
2	コミュニケーションが取れる旅行	友達と一緒に旅行	自己成長・自己啓発
3	自己成長できる旅行	リフレッシュ旅行	文化に触れる
4	安い旅行	安心・安全な旅行	リフレッシュ旅行
5	その土地の文化に触れる旅行	安い旅行(特に理系)	安い旅行

これを第 1 章で示した既知とする観光ニーズ(表 1)と比較する。既知とした男女の観光ニーズについてはそのまま比較し、全般にあるものも男女にそのまま適用したが、男性の一人旅や女性の安い旅行については、はっきりと属性傾向が出たため、そのまま残した。また、伝統文化を巡る旅行についても、世界遺産という具体的な対象が出たため、そのまま残した。しかし、全般の新知の観光ニーズである非日常体験は、今回の先行研究からは得られていないが、観光業界で一般的に言われるニーズであるため、対象から外した。以上より、新たに発見した新知の観光ニーズを表 9 に示す。

表 9 新知の観光ニーズ

男性	女性
一人旅(理系)や個人ツアー(文系)など一人の旅行	世界遺産(建造物)をゆっくり巡る旅(特に歴史好き層)
コミュニケーションが取れる旅行	友達と一緒に旅行
自己成長できる旅行	安心・安全な旅行
現地の文化に触れる旅行	安い旅行(特に理系)

4. 結論

本研究では、テキストマイニングによる大学生の観光ニーズ推定をおこなった。そこから、女性は、歴史好きな層を中心に、友達と一緒にゆっくり世界遺産を巡る旅行を求めており、男性は志向によって一人の旅行でも、理系は一人旅、文系は個人ツアーを好む傾向にあることがわかった。今後の展望として、まずテキストマイニング分析において属性の中でも特に志向(理系・文系、歴史好き・歴史嫌い)において男女に差が出たことから、さらに傾向が出やすい変数を見つけることが挙げられる。

また、計量テキスト分析でコーディングを行う際、階層的クラスタ分析よりグルーピングをおこなったが、より特徴を出せたとは言えなかった。このような語の分類

をする際、シソーラスを用いた分類が有効であると考えているd。しかし、通常のシソーラスと観光のシソーラスでは違いが存在するe。観光の消費行動を分類体系としたシソーラスが構築され、観光分野に適したコーディングができれば、もっと明確に特徴を捉えられるかもしれない。そして、観光ニーズを把握する上で、計量テキスト分析がより有効的な分析手法となりえると考えている。

現段階では、観光分野におけるテキストマイニング、計量テキスト分析に王道と言われる分析手順は存在しない。しかし、既存の研究からは得られなかった観光ニーズを知りえることができた。今後は国内旅行と海外旅行に分けて分析を試みるとともに、さらに新知の観光ニーズを増やしたい。

参考文献・資料

- 観光庁(2009). 観光庁アクションプラン 2009
- 廣岡裕一(2008). 「若者の海外旅行離れ」に関する考察. 世界旅行博 VWC タウンミーティング(2008年9月19日, 於東京ビッグサイト)発表配布資料
- 箕浦康子(1993). 「フィールドワークの技法と実際-マイクロエスノグラフィー入門-」, ミネルヴァ書房
- 日本観光協会総合研究所, 2009, 若者は本当に旅をしなくなったのか(特集 若者たちの観光), 観光とまちづくり, 496, pp. 18-23
- 恩田裕介, 澤村博, 2008-12, 大学生の旅行目的・動機に関する研究, ツーリズム学会誌, 8, pp. 134-144
- 金春姫, 鎌田裕美, 2010-03, 若者の旅行に対する意識, 成城大学経済研究, 188, pp. 177-191
- 高井典子, 中村哲, 西村幸子, 2009-04, 若者の海外旅行離れ「論」への試み, 日本観光研究学会第23回全国大会論文集, pp. 421-424
- Marti A. Hearst, (1999), Untangling Text Data Mining, ACL'99, pp.3-10
- 樋口耕一(2005), 「計量テキスト分析の方法と実践」, 大阪大学大学院人間科学研究科 平成16年度 博士論文
- JTB 広報室(2009), 2010年の旅行動向見通し, JTB ニュースリリース
- ローネン・フェルドマン著/辻井潤一監訳(2010), 「テキストマイニングハンドブック」, 東京電機大学出版局
- 守屋豊・井出明(2009), 観光情報シソーラスの構築と観光言説比較に関する研究, 日本観光学会誌第50号, pp. 41-49

d シソーラス(thesaurus)とは、語句を意味によって分類・配列した語彙集のことで、類義語集とも呼ばれる。
 e 例えば「温泉」と「卓球」というような語の関連は通常のシソーラスには出てこない。このように普段の語の使い方とは違った関連が観光に関する語には存在するため、観光情報のシソーラス開発が必要である。(守屋・井出 2009)