

2021年度 修士論文

訪日中国人の旅行ブログの
テキストマイニングで探る
訪日中国人の温泉体験への
関心と評価

東京都立大学院 都市環境科学研究科 観光科学域

ZHU YIWEN(朱逸雯) 20854507
指導教員 倉田陽平

要旨

日本のインバウンド観光振興にとって、中国は規模と成長性を持つ有力な市場である。そして、パンデミック終息後の観光インバウンド復興には中国人観光客が重要な役割を持つ。訪日中国人観光客は 2015 年から「爆買い」により注目され、今までショッピングを中心とした観光活動の印象が強いが、JNTO などからの調査では近年の観光動向は体験型観光へ変わる傾向がある。本研究で筆者は日本の代表的体験型観光である温泉旅行に注目する。

本研究の研究目的は、訪日中国人が日本での温泉旅行に対する印象と活動実態を把握することである。

社会情勢とサンプリング数を考慮し、本研究の研究方法はテキストマイニングである。ただし、従来のテキストマイニング研究対象はクチコミがほとんどであるが、クチコミの内容はホテルと特定の観光スポットに限り、温泉旅行のような温泉地到着から離れるまでの過程をカバーできない。従って、本研究はより詳細な記述を獲得できる旅行ブログを研究対象とする。

そうしたデータにはテキスト量が膨大すぎるというデメリットもあるため、より視点を絞った分析方法を提案する必要がある。また、ブログからのデータはクチコミと異なり、星付けなどの評価機能がないため、ユーザーの書き込みの態度が見えにくく、ほかのアプローチでブロガーの感情を把握する必要がある。本研究はコーディングによる特定のテーマ関連記述の抽出と感情分析を併用することで、ブログ分析における問題点を解決した。

使用する分析ソフトはテキストマイニング分析ソフト KH Coder である。分析対象は中国大手オンライン旅行予約サイト ctrip で選定した 14 個の温泉地の訪日温泉旅行ブログで、スクレイピングでデータを取得した。コーディングによるテーマ関連記述の抽出と感情スコアを算出することで、中国人が温泉旅行に対する注目点とその評価を得た。結果、中国人は「精緻」という語を頻繁にかつポジティブに温泉地で体験を描写していることが判明した。また、温泉地記述と訪日旅行全体の感情分析結果を比較したことで、温泉地での記述は旅行全体の記述よりポジティブであることがわかった。

さらに、温泉旅行全般からの視点と違ったエリア別の温泉旅行の特徴を見出すため、活動エリア別にブログを分類し、分析を実施し、中国人観光客がエリア別の温泉旅行での体験の特徴をまとめた。

前の分析結果での頻出語「精緻」の使用場面を究明するため、「精緻」のコーディングルールを作成し、料理関連、建築関連、工芸品関連、三つの使用場面にまとめた。

温泉地でのコード間の比較により、交通に関連する記述にポジティブな描写が少なく、感情を含まない記録的な記述が多いことから、ブログから見た中国人観光客の交通体験は相対的に印象を残らなかったことがわかり、体験向上させるため、「精緻」の要素を取り込みやすい駅弁に力を入れることを提言した。

最後に、コーディングによる再訪意向の影響要素を確かめ、「見残し・やり残し」による再訪、今回満足したので次も体験したい、異なる時期にまた来たい、主にこの三つの要素により再訪意欲が生じることが分かった。

本研究の実学的価値は、中国人観光客は温泉旅行で何に関心を持ち、さらに、どう評価していたのかを分析により示したことである。こうした分析を通して、パンデミックが終息に向かっている中で、パンデミック後のインバウンド観光誘致に中国人観光客の評価に合う誘致策を提言でき、また、中国人観光客が観光中に情報不足による観光地の魅力の見逃しなども観光地に示唆を与えることができる。中国人観光客に向けて、どうプロモーションするかの参考になりうると考えられる。また、再訪についての影響要素を整理したことで、観光地にとって、どうやって中国人観光客の再訪意欲を上昇させるかの参考材料にもなりうる。

本研究の学術的価値は、従来のクチコミやレビューなどの短い文書のユーザーのフィードバックとは異なった視点で、ブログのようなより多くの情報を持つテキストデータの分析方法に提示できたことである。

結論としては、まず、訪日旅行において温泉地での滞在経験が日本旅行にプラスの評価をもたらし、日本旅行のイメージ向上につながることがわかった。次に、中国人観光客は温泉旅行、特に会席料理を始めた食事関連の面では「精緻」という評価語を使い、高評価を与える。将来的に「精緻」のさらなる使用場面を研究し、中国人観光客が持つ「精緻」のイメージが如何なるものかをより明確にする必要がある。そのうえで、「精緻」の要素を取り込んだ観光客向けの商品開発も提案できる。

目次

第1章	序論	1
1-1.	問題意識	1
1-2.	研究目的	1
1-3.	本論の構成	2
第2章	研究背景	2
2-1.	COVID-19	2
2-2.	訪日中国人	3
2-2-1.	訪日中国人観光客の増加	3
2-2-2.	訪日旅行に関する期待内容	5
2-3.	温泉旅行	5
第3章	先行研究	6
3-1.	訪日中国人に対する先行研究	6
3-2.	テキスト分析での先行研究	7
3-3.	ブログとクチコミの比較	8
第4章	研究方法	8
4-1.	テキストマイニング	8
4-1-1.	形態素解析	8
4-1-2.	構文解析	9
4-1-3.	頻度集計	9
4-1-4.	共起ネットワーク	9
4-2.	自然言語処理ツール	9
4-2-1.	KH Coder	9
4-2-2.	中国語の自然言語処理	10
4-3.	コーディングルール定義	10
4-4.	データ収集	11
4-4-1.	スクレイピング	11
4-4-2.	データ処理	14
4-5.	テキスト分析	16
4-5-1.	内容分析	16
4-5-2.	計量テキスト分析	17
4-6.	感情分析	17
第5章	訪日温泉旅行ブログテキスト分析	19
5-1.	温泉地記述と旅行全体記述との対照	19
5-1.	目的地活動エリア別の比較	23
5-2.	再訪意向	25

第6章	分析結果	25
6-1.	温泉地全般	25
6-1-1.	温泉地と旅行全体との対照	26
6-2.	目的地活動エリア別の比較	39
6-2-1.	交通	40
6-2-2.	ホテル	43
6-2-3.	日本	46
	温泉	47
6-2-4.		47
6-2-5.	食事	49
6-3.	精緻の概念	50
6-4.	再訪意向	54
第7章	考察	55
7-1.	温泉地全般について考察	55
7-2.	精緻の考察	55
第8章	結論	56
8-1.	制約と展望	57
8-1-1.	感情分析のデメリット	57
8-1-2.	データの偏り	57
引用・参考文献		58
謝辞		60

表目次

表 2-1 JNTO2020 世界各国への中国人訪問者数の推移 (人)	3
表 2-2 「訪日外国人消費動向調査 (2011 年～ 2019 年)」に基づき日本政府観光局 (JNTO) が作成	4
表 5-1 コーディングルール一覧 (日本語翻訳)	20
表 5-2 コーディングルール一覧 (中国語原文)	20
表 5-3 各コードの数と温泉地記述における割合	21
表 5-4 各コーディングの類似度行列 (Jaccard 係数)	21
表 5-5 空港別のブログ数	24
表 6-1 交通文ポジ感情の特徴名詞	27
表 6-2 交通文ポジ感情の特徴形容詞	27
表 6-3 交通文ネガティブ感情の特徴名詞	28
表 6-4 ホテル文ポジティブ感情の特徴名詞	29
表 6-5 ホテル文ポジ感情の特徴形容詞	30
表 6-6 日本文ネガ感情の特徴名詞	31
表 6-7 日本文ネガティブ感情の特徴形容詞	32
表 6-8 日本文における全体と温泉地記述の特徴名詞の比較	33
表 6-9 日本文における全体と温泉地記述の特徴形容詞の比較	34
表 6-10 温泉文ポジ感情の特徴名詞	35
表 6-11 温泉文ポジ感情の特徴形容詞	36
表 6-12 食事文における全体と温泉地記述の特徴形容詞の比較	37
表 6-13 食事文における全体と温泉地記述の特徴名詞の比較	38
表 6-14 エリア別のネガティブ感情文	40
表 6-15 ホテルネガティブ感情文	43
表 6-16 日本ネガティブ感情文	46
表 6-17 食事ネガティブ感情文	49
表 6-18 精緻コードの関連名詞	51
表 6-19 精緻の使用例	52
表 6-20 各温泉地ブログの精緻使用状況	53
表 6-21 各コードと精緻コードとの類似度行列 (jaccard 係数)	54
表 6-22 再訪記述の具体例	54

目次

図 2-1 「訪日外国人消費動向調査（2011 年～ 2019 年）」に基づき日本政府観光局（JNTO）が作成	4
図 2-2 2012 年-2016 年中国国内での温泉産業収入および収入成長率	6
図 4-1 ctrip のブログ検索ページ	11
図 4-2 ctrip のブログ画面	13
図 4-3 Data Miner のレシピ作成画面	14
図 4-4 クリーニング後の各地ブログ数	16
図 5-1 KH Coder 公式説明スライド	23
図 6-1 温泉地記述での各感情の割合	25
図 6-2 交通文における全体と温泉地記述の対照	26
図 6-3 ホテル文における全体と温泉地記述の対照	28
図 6-4 日本文における全体と温泉地記述の対照	31
図 6-5 食事文における全体と温泉地記述の対照	37
図 6-6 ポジティブ感情交通文 x 外部変数：名詞・共起ネットワーク	41
図 6-7 ポジティブ感情交通文 x 外部変数：形容詞・共起ネットワーク	42
図 6-8 ポジティブ感情ホテル文 x 外部変数：形容詞・共起ネットワーク	43
図 6-9 ポジティブ感情ホテル文 x 外部変数：名詞・共起ネットワーク	45
図 6-10 ポジティブ温泉文 x 外部変数：形容詞・共起ネットワーク	47
図 6-11 ポジティブ温泉文 x 外部変数：名詞・共起ネットワーク	48
図 6-12 精緻コードの感情分布	50

第1章 序論

1-1. 問題意識

日本のインバウンド観光振興にとって、中国は規模と成長性を持つ有力な市場である。ポスト COVID-19 のインバウンド復興で、中国人観光客も重要な役割を持つ。しかし、先行研究から見て、中国人がある特定の観光活動に関する研究が重要であるにも関わらず、いくつかの研究上の難しさにより十分に研究されてこなかった。一つ目は中国人観光客が急増したのは近年であり、2015年以前では団体ツアーによる観光がメジャーで、観光活動のバリエーションが少ない。二つ目、訪日中国人観光客の観光活動を調査するには言語の壁がある。三つ目、従来の訪日中国人観光客は主に都会で活動しており、ショッピングがメインである旅行も多いが、近年では中国の旅行クチコミサイト「馬蜂窝 (mafengwo)」が実施した調査報告「2019 中国出境自由行大数据報告」と JNTO が訪日中国人の訪日旅行に関する期待内容 (2015-2019) から、中国人は近年では体験型旅行に切り替える傾向が見えた。そのため、体験型旅行への研究がまだ十分でないと考えられる。

また、オンラインユーザーデータを対象にしたテキスト分析ではクチコミが分析対象であるのがほとんどである。ただし、クチコミから温泉旅行行動を分析するには課題点がある。温泉地に関する評価はクチコミではホテルか、特定の観光スポットのクチコミしかなく、温泉地全体に関する記述と評価はクチコミがないことが多いためである。なので、温泉旅行のような特定形態の観光活動についての研究にはより詳しく温泉地での活動記録を探す必要がある。

そういうことから、筆者が目をつけたのが旅行ブログである。旅行ブログは、出発から帰宅まで記録したものがほとんどである。そこからクチコミより詳細な温泉地に関わる記述が抽出できる。ただし、そうしたデータにはテキスト量が膨大すぎるという研究上の難点がある。そこで、本研究ではコーディング方法の導入により視点を絞った分析方法を提案した。そうした複合的旅行体験のテキスト分析の難点を克服し、温泉旅行のような宿泊を含めた観光活動についての研究を行った。

さらに、ブログからのデータはクチコミと異なり、星付けなどの機能がないため、観光客が記録した際の態度が見えにくく、ほかのアプローチでブロガーの感情を把握する必要がある。本研究では感情分析の導入により、膨大なテキストデータから中国人観光客の感情を動かす要素を探し出すことができた。

1-2. 研究目的

本研究の研究目的は、日本の温泉地を訪れた中国人観光客の旅行ブログを分析することで、訪日中国人が日本での温泉旅行に対する印象と活動実態を把握することである。

本研究の実学的価値は、中国人観光客は温泉旅行で何に関心を持ち、さらに、どう評価していたのかを究明することである。こうした分析を通して、パンデミックが終息に向かっている背景の中で、パンデミック後のインバウンド観光誘致策を提言し、再訪意

欲上昇，サービス改善といった復興の手がかりを探る。

本研究の学術的価値は，従来のクチコミやレビューなどの短い文書のユーザーのフィードバックとは異なった視点で，ブログのようなより多くの情報を持つテキストデータの分析方法を提示できる。

1-3. 本論の構成

本研究は8つの章から構成されている。

第2章では，2020年1月を境目に，パンデミック以前の中国人観光客の観光実態，再訪観光客数，並びに今回と次回旅行での期待内容との比較により，温泉旅行を研究対照に決めた。そして，中国国内での温泉旅行現状と日本温泉の中国での事業進出を事例にあげ，温泉旅行が中国での人気を述べた。

第3章では今までの訪日中国人観光客に対しての先行研究を振り返り，先行研究の結果と不足を指摘した。また，ユーザーを対象としたテキスト分析の先行研究をまとめた。先行研究が主にホテルと観光スポットのクチコミとレビューを対象として分析するが，旅行の全体図からの視点がないため，本研究は旅行ブログを対象に研究することとなった，そして，クチコミとブログを比較した。

本研究の中心内容は，第4章から第7章にわたって論じる。第4章では，旅行ブログの分析方法と分析対象を選定する。本研究に使用する分析方法はテキストマイニングである。テキストマイニング定義の共通点は，文書の集合体，文書情報から，頻出する言葉や潜在する関連性，顕著なパターン，データの規則性，類似性など重要な情報や知識を得ることである。さらに大量のテキストデータを処理し，分析したテキスト情報を可視化し，明瞭な解釈を行う分析手法でもある。用いる分析ソフトはテキストマイニング分析ソフト KH Coder である。まず，分析方法、データ抽出作業などを説明し，分析対象である中国大手オンライン旅行予約サイト ctrip での訪日温泉旅行ブログを14個の温泉地に選定し，それぞれの概況を示す。

第5章と第6章では，テキストマイニング分析によって，得られた研究結果を示す。コーディングによるテーマ関連記述の抽出と感情スコアを算出することで，中国人が温泉旅行に対する注目点とその評価を得る。さらに，活動エリア別にブログを分類し，日本でエリア別の温泉旅行の特徴を見出すのも狙いである。そして，「再訪」に関するコードを作り，再訪意向を示す文書を定性的に分析し，再訪に関連する要素を究明する。

結章では，本研究の理論的意義を述べ，研究結果によって日本の観光政策にいくつかの実学的提案をするとともに今後に残されている研究課題を述べる。

第2章 研究背景

2-1. COVID-19

新型コロナウイルス COVID-19 の感染拡大の影響によって，2020年の世界全体の実質

GDP 成長率は-3.3%となり、全世界での国際観光客数は前年に比して 73.1% 減少し、約 4 億人となった。日本では 2009 年以来、11 年ぶりの-4.7% の実質マイナス成長率であった。また、世界的な渡航制限、人的交流の抑制が行われたことから、訪日外国人旅行者（インバウンド）数は前年比 87.1% 減の約 400 万人にとどまった。これは「観光立国」が宣言される前の 1998 年の水準と同等である。

中国人観光客に関しては、2020 年 2 月から本稿を執筆している 2021 年 12 月の現在でも中国人観光ビザによる入国が認められない状況である。さらに、中国国内での厳しい水際対策による隔離政策があるため、日本在住の中国人を除き、日本国内では中国からの中国人観光客がほぼいないのが現状である。

グローバル的なワクチン接種普及と対 COVID-19 特効薬が次々と承認された今では、パンデミックの収束の明かりが少し見えてきた。国際交流が COVID-19 以前の水準に戻り、インバウンド復興、さらにインバウンドのさらなる拡大まで発展したいなら、外国人観光客が訪日観光において何に関心を持ち、どう評価しているのかを知る必要がある。

外国人観光客の「現在」に対する研究ができないなら、外国人観光客、特に中国人観光客が今までの観光体験では何に関心を持ち、どう評価したのかを詳しく分析し、ポスト COVID-19 の観光回復に役立つ研究が求められる（示唆を得ることが期待される。）

2-2. 訪日中国人

2-2-1. 訪日中国人観光客の増加

日本政府観光局の統計調査によると、2019 年の訪日中国人観光客数は 959 万 4394 人に達し、訪日外国人全体の 30% を占めた。訪日中国人観光客数は 2007 年では、94 万 2439 人で 100 万人に届かない水準だったが、そこから 12 年で 10 倍に増えたことに至った。

JNTO の訪日旅行データハンドブック 2020 の世界各国への中国人訪問者数の推移（2014 年-2018 年）では、日本は中国人が選ぶ海外旅行目的地で、タイに次ぎ人気第二位の目的地である。

順位	訪問先	2014 年	2015 年	2016 年	2017 年	2018 年
1 位	タイ	4,636,298	7,936,795	8,757,646	9,806,260	10,535,241
2 位	日本	2,409,158	4,994,689	6,373,564	7,355,818	8,380,034
3 位	ベトナム	1,947,236	1,780,918	2,696,848	4,008,253	4,966,468

表 2-1 JNTO2020 世界各国への中国人訪問者数の推移（人）

日本のインバウンド観光振興にとって、中国は規模と成長性を持つ有力な市場である。ポスト COVID-19 にインバウンド復興には中国人観光客も重要な役割を持つ。また、近年における各国の中国人対象の観光ビザの緩和によって中国から観光旅行できる地域が拡大しており、グローバルで中国人客争奪戦も激しさを増している。そのため、今後の中国人観光客誘致については、中国人観光客の関心や行動の特徴を理解した上で、戦略

的にすすめていく必要があると筆者は考える。

また、日本政府観光局（JNTO）が作成した観光庁「訪日外国人消費動向調査（2011年～2019年）」から、訪日中国人観光客が年々増加しているとともに、日本への再訪客つまりリピーターの割合も増加していることが読み取れる。そのため、リピーターが次の訪日旅行に対して何に対して期待するかを知るのも重要である。

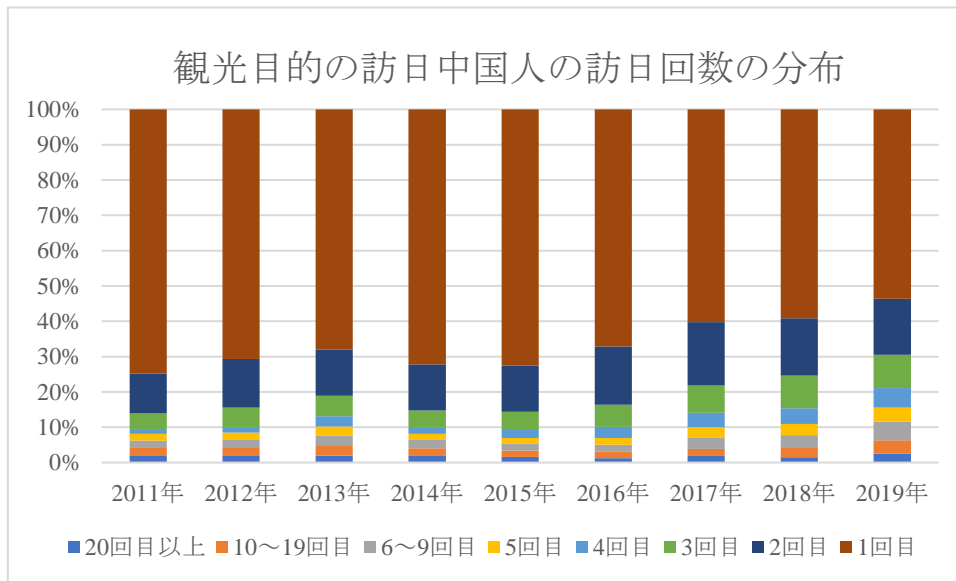


図 2-1 「訪日外国人消費動向調査（2011年～2019年）」に基づき日本政府観光局（JNTO）が作成

	2011年	2012年	2013年	2014年	2015年	2016年	2017年	2018年	2019年
20回目以上	1.9	1.9	2.0	1.9	1.6	1.2	1.9	1.5	2.5
10～19回目	2.2	2.2	2.7	2.1	1.8	1.8	2.0	2.6	3.8
6～9回目	2.1	2.5	2.9	2.4	2.0	2.0	3.2	3.7	5.3
5回目	1.9	1.9	2.6	1.7	1.5	1.9	2.9	3.1	4.0
4回目	1.4	1.5	2.9	1.9	2.3	3.2	4.1	4.5	5.5
3回目	4.4	5.5	5.8	4.7	5.2	6.2	7.7	9.2	9.4
2回目	11.3	13.8	13.1	13.1	13.0	16.4	18.0	16.2	15.9
1回目	74.8	70.6	68.0	72.1	72.5	66.7	60.2	59.1	53.6

表 2-2 「訪日外国人消費動向調査（2011年～2019年）」に基づき日本政府観光局（JNTO）が作成

2-2-2. 訪日旅行に関する期待内容

中国人観光客の特徴としては、「爆買い」に代表される買物好きであることが指摘されることが多い。実際に 2019 年訪日外国人客費用別旅行支出によると、中国人観光客の「買物代」が 10 万円を超え、旅行支出総額の 51.1%を占めている。訪日客数の多いアジアの主な国・地域と比較しても、支出額・支出割合ともに突出して高くなっている。

こうした背景の中で JNTO が訪日中国人に対して、訪日旅行に関する期待内容（2015-2019）を調査した。結果、日本に訪れた回数の増加につれて、訪日旅行に対する期待内容も変わる。

一般的な観光活動、例えば、食事、「爆買い」で知られるショッピング、街歩きなどは、次回の訪日旅行での期待度が下がったことが分かった。

その一方、温泉入浴をはじめ、自然体験や文化体験といった項目が次回の訪日旅行期待度が上昇した。また、自然体験も温泉旅行の一環として成り立つので、温泉旅行は中国人リピーターの期待に沿う観光活動であると考えられる。

もう一方、中国側の調査では旅行クチコミサイト「馬蜂窩 (mafengwo)」が実施した調査報告「2019 中国出境自由行大数据報告」では、中国人の海外個別手配旅行について、目的地選択の「マイノリティ化 (小衆化)」を指摘している。中国人観光客の海外観光地に関する検索データを見ると、「博物館」の検索回数は前年に比べ 41%増え、一方「買い物」の検索回数は 11%減った。「文化観光」、すなわち観光地での深い体験を目指す旅行が、中国人観光客の従来の買物を中心とする定番的な旅行に取って代わりつつあり、さらに「マイノリティ (小衆)」というキーワードが検索された回数は前年に比べ 132%増と大幅に上回った。特に若者たちは個性を求め、定番化していない「マイノリティ (小衆)」の観光地、個性のある観光地を選択する傾向が多く見られた。こうした情報から、これから中国人観光客は文化体験に注目し、人気の高い都会からより地方の個性のある観光地に赴くと推測できる。温泉旅行もそうした個性と文化体験を持ち、中国人観光客の選択対象になると考えられる。

2-3. 温泉旅行

温泉は中国国内での人気も上昇している。

2016 年までのデータしかないが、図 2-2 でわかるように、中国国内で、温泉産業収入も年々増長していた。

中国温泉と日本温泉には文化の差がある。中国の地熱資源が日本ほど開発されていないため、温浴施設の数が少なく、また、泉質より娯楽を重視する傾向がある。都会から離れた温浴施設は高級ホテルと併設し、東京ドームほどの敷地を持ち、室外で数十個の湯舟を持つのが一般的である。水着を着用しながら温泉を楽しむので、開放感があり、より長時間の滞在が狙える。では、日本の温泉文化は中国で人気を集められるか。

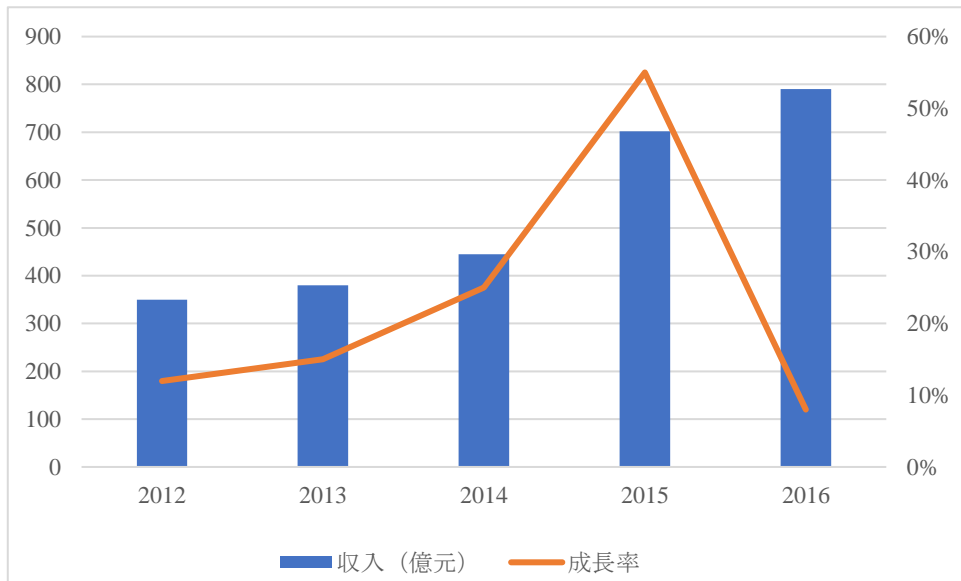


図 2-2 2012 年-2016 年中国国内での温泉産業収入および収入成長率

日本の大手スーパー銭湯チェーン「極楽湯」は、十年前から上海を拠点に事業展開をしており、2020年でも事業拡大のペースも緩めることはなかった。中国大手新聞社「第一財經」の2021年一月の報道によると、2021年の中国の長期休み春節では、消費者は海外旅行もできなく、国内旅行や帰省など省を跨ぐ移動も COVID-19 対策で相当制限されていた中で、極楽湯を含む現地の温泉施設が人気上昇した傾向が見られた。温泉旅行や温浴施設の利用も現在、中国人娯楽のトレンドにもなっている。

さらに、2021年秋に「極楽湯」は中国上海のディズニーリゾートエリア内で「百万石温泉街」という名の大型温泉ホテルをオープンした。スーパー銭湯だけではなく、本格的に日本の温泉旅行文化を中国に輸入しようとする志が見える。

「極楽湯」を筆頭とした日本の温浴施設企業が中国で長年事業展開及び事業拡大できたことから、日本の温泉文化は中国本土でもマーケットを持ち、中国の消費者たちの人気を集めていた。

第3章 先行研究

3-1. 訪日中国人に対する先行研究

呉(2019) は中国発のオンライン・トラベル・エージェント (OTA) サイトである Ctrip.com を取り上げ、京都の 30 の観光スポットを対象に、ctrip でのクチコミとガイドテキストをトピックの特徴抽出と乖離度の計測を行った。外国人観光客に対しての研究はクチコミの点数を見るだけでなく、トピックモデルによる特徴抽出し、より定性的な分析が必要だと提言した。ただし、トピックだけではポジティブかネガティブかが見えないため、極性分析による顧客の感情を把握したうえでのさらなる意味解析が必要だと補足した。

脇本,姜,大西(2018)は、沖縄を訪れる中国人観光客に対して、質問紙調査とインタビ

ューを実施して、回答者のプロフィール、最も使用する SNS 情報、沖縄のイメージを分析データとして用いて、因子分析を行うことにより中国人観光客の消費行動と SNS の関係性について分析した。分析の結果、旅行先情報と 10 行動指標に SNS の影響力が高いことが推察され、沖縄における今後の観光戦略の手法に CtoC マーケティングの重要性が示唆された。

3-2. テキスト分析での先行研究

鄔(2016) は本研究と同様、中国語テキストを対象に KH Coder のコーディング機能を使用した研究である。鄔(2016)は中国各旅行会社の訪日旅行商品での紹介文を対象に分析し、旅行会社の視点から旅行会社の影響により訪日中国人の日本観光地イメージを研究した。鄔(2016)の研究では、当時の KH Coder には中国語分析に対応しておらず、他に適切な分析ソフトが見つからなかったため、分析対象を日本語に翻訳してから分析した。しかし、翻訳する際に意味合いが変わり、翻訳者の主観的影響もあるため、原文の言語で分析することが最適である。よって本研究では中国語のまま分析を行った。さらに、鄔らはブログやクチコミなどユーザーからの直接的なデータを使用しなかったため、訪日中国人の直接的な観光体験と観光地イメージが得られなかったのも改善できる箇所である。

その他本研究と同様に、中国人訪日ブログを対象とした研究は、張章(2021) があげられる。張らは高松市、広島市、熊本市を目的地とする旅行記（旅行ブログと同義）を対象に中国人の訪日観光行動の特徴を分析した。分析対象のブログは mafengwo から抽出したもので、mafengwo は中国では ctrip に次ぐ二番目の市場シェアを持つオンライン・トラベル・エージェント（OTA）サイトで、ctrip に比べより旅行経験のシェアに力を入れているが、データ量では ctrip がより多くのデータを持つので、本研究では ctrip を研究対象に選定した。ブログを形態素解析して、頻出語を手作業で、「飲食関連」、「買物関連」、「宿泊関連」、「自然関連」、「歴史関連」、「文化関連」、「県内地域」、「県外地域」別に分類し、観光行動の特徴を分析した。本研究もブログ内容を「飲食」、「ホテル」、「交通」などのテーマに分類をしたが、頻出語の手作業による分類とは異なり。コーディングを使用することで、より多くのデータを処理でき、主観的な操作を介さない手法でもある。張らの分析結果から見ると、「マイノリティ（小衆）」の観光地の中国人観光客は、買物に対してそれほど関心がなく、旅行記に買物に関する内容を多く紹介することも少ないようである。「マイノリティ（小衆）」の観光地を訪れる中国人観光客は、特定の趣味や目的を持って特定の地域や訪問スポットを選択し、さらに食事の選択も、美味しさより特定の目的を実現するため選択している傾向がある。

また、張らの研究で抽出したデータは高松市、広島市、熊本市に関する旅行記それぞれ三十件であり、サンプリング数がやや不足していると考えられる。さらに、高松市、広島市、熊本市対象にしたのは人気の中位水準であり、マイノリティ（小衆）のある目

的地とされているが、「マイノリティ（小衆）」の定義が曖昧で、日本における立ち位置がややわかりづらいという点もある。高松市、広島市、熊本市以外、どんなどころが「マイノリティ（小衆）」に位置付けできるかを示せば、街づくりにより多くのヒントを与えることが出来ると考えられる。

3-3. ブログとクチコミの比較

前節で述べたように、現在、観光分野でのテキスト分析はクチコミとレビューを中心に展開されている。しかしクチコミなどのコメントは一つの観光スポット、ホテル、レストランに対してのそれぞれの評価なので、旅行全体の視点からの分析は難しい。

一方本研究は中国人観光客の訪日ブログを対象として分析し、旅行の一過程、特に外国人観光客にとって、宿泊は避けられないので宿泊や交通、食事、体験したサービス、風景の感想を込めて分析することもできる。さらに、分析を通して温泉地での観光活動が訪日旅行全体における立ち位置がわかる。

西平（2019）は旅行クチコミサイト 4 travel.jp での山形県の旅行クチコミを対象として、テキスト分析を行った。主に頻出語分析と共起ネットワーク分析で、山形県の人気のある観光地をまとめた。ただし、頻出語だけでの分析には定性的な視点が欠けており、感情分析のようなポジネガ判定もないため、頻出語だけでは観光地が人気であるというのは言い難いと考えられる。本研究はこれらを踏まえて、コーディングによる話題別の分析を行い、より視点を絞った定性的な分析を可能とした。

また、テキスト分析において越智（2018）のような年代別、男女別の視点からのクチコミ分析も多くあり、観光客を細分化し、より細かい観光提供が実現できる。本研究は温泉地に訪れた中国人ブロガーというすでに細分化された視点から切り込み、分析し、温泉地の潜在来訪者により良い観光サービス提供への提言を目指す。

第4章 研究方法

4-1. テキストマイニング

テキストマイニングとは、膨大なテキスト（文書）情報の中から注目すべき情報を掘り出す（マイニング）ことで、定性的データであるテキストデータを、定量的データに整理し、データマイニングの手法を用いながら、相関関係などの定量分析を行う手法である。（斎藤 2012）

4-1-1. 形態素解析

形態素とは、「言語学で、意味を持った最小の音型。ヤマ（山）のように形態素一つで単語が構成される場合もあれば、ヤマカゼ（山風）のように複数の形態素が単語を構成する場合もある（大辞泉）」とされる。文章から形態素を探し出し、その形態素単位に分割することを形態素解析と呼ぶ。（斎藤 2012）日本語の形態素解析ソフトウェアは MeCab と Chasen が主流である。中国語の形態素解析のツールとしては、スタンフォード大学の The Stanford NLP Group が提供している Stanford Word Segmenter などがある。

本研究では KH-coder 上で Stanford Word Segmenter を使用した。

4-1-2. 構文解析

形態素に分割された文章は集計の際には有効だが、文章の意味にまで踏み込んで定性的に分析を行う際には不十分である。こういった場合には、係り受け関係など構文について検討する必要がある。(斎藤 2012) 本研究で使用するテキスト分析フリーソフトウェア KH Coder には構文解析機能を用いており、形態素(抽出語)の係り受け関係を見ることが出来る。

4-1-3. 頻度集計

テキスト分析の第一段階は、語の頻度の集計である。集計方法は、大きく分けて二通りある。一つは、文章の中で一つの語が出現した個数を集計する方法である。この方法では、一つの文章である語が複数回出現した場合、それぞれを出現回数としてカウントする。もう一つは、文章の中で一つの語が出現したか否かを集計する方法である。この方法では、一つの文章の中である単語が何回出現したとしても、一回としてカウントする。異なるファイルでの語の出現有無から、ファイル全体での語の出現割合を把握できる。語の集計により、分析対象となる文章の特徴を大まかに把握することが出来る。

(斎藤 2012)

4-1-4. 共起ネットワーク

共起ネットワーク分析は、出現パターンが似通った語、すなわち、共起の程度が強い語を線で結んだネットワークを描くことである。線でつながっていることがそのまま語と語の共起を示すので、多くの人にとって理解しやすい可視化の一つである。もちろん線につながった 2 語の組み合わせに限らず、出現パターンの似通った語のグループを探せば、そこからデータ中に多く表れたテーマないしはトピックを読み取れる。このように、ネットワークを描いて探索する方法は内容分析の分野では早く用いられており (Osgood 1959)、ネットワーク分析の指標を活用することも提案されている (Danowski 1993)。

テキスト全体における共起を探るだけでなく、KH Coder ではテキストの部分ごとの特徴を見つけることもできる。語と語との共起だけでなく、語と外部変数による関係も描ける。語と外部変数の共起ネットワーク図を描けば、それぞれの外部変数の特徴語を読み取れる。(2017 樋口)

4-2. 自然言語処理ツール

4-2-1. KH Coder

KH Coder とは、テキスト型データを統計的に分析するためのフリーソフトウェアである。

樋口 (2014) はテキスト分析の手法を 2 つに分け、客観的なデータを作成して多変量解析を使って分析する Correlational アプローチと、主観的な基準(コーディングルール)で、言葉や文章を選択して分析する Dictionary-based アプローチに分けている。

Correlational アプローチは分析者の理論や問題意識の影響を受けない分析手法であり、Dictionary-based アプローチは研究者が主観的に言葉を選び、理論仮説の検証や問題意識の追求を行うアプローチである。

本研究で注目したコーディングによる文書の分類は Dictionary-based アプローチである。

KH Coder のほかにも MTMineR(石田・金編 2012, 金 2016), Random Forest などのテキスト分析ソフトがあるが、既存のソフトウェアではコーディングルールを作成することで、複雑な文書を分割してより細やかに分析できるソフトウェアが極めて限られている。(樋口 2017)

筆者は(鈴木 朱 2021) では KH Coder を使用し HIS のオンラインツアーユーザーレビューを分析した、HIS のユーザーレビューは日本語であるが、KH Coder は中国語も対応するため、中国人のブログ分析にも対応できる。

4-2-2. 中国語の自然言語処理

中国での自然言語処理分野は python のパッケージを中心に使用する。ただし python を使用するには一定のプログラミング言語に慣れる必要があり、使い手に対するハードルがある。一方、KH Coder はフリーソフトウェアで、操作画面もシンプル、プログラミング言語と形態素解析などの原理を理解しなくても、より多くの研究者がテキスト分析に触れることができるソフトウェアであると考えられる。

中国では KH Coder のような自然言語処理用のソフトもあるが、ほとんどが企業向けで、学術でよく使われていた ROST-CM (武漢大学沈陽チーム開発) は 2011 年前後からアップデートされていない。KH Coder はインターフェース言語でも、言語処理でも中国語に適用できるが、中国の学界では KH Coder を使用する論文や使い方の紹介やネットでの議論がほぼないのが現状である。

ただし KH Coder で中国語分析機能は日本語分析機能より少ないというデメリットもある。

本研究を通して、より多くの研究者に中国語のテキスト分析に少なからず示唆を与えることができると思われる。

4-3. コーディングルール定義

KH Coder が扱うコーディングルールとは、例えば次のようなものである。

*過去 懐かしい or 思い出す or 恋しい or 久しぶり

この例では、「〈懐かしい〉という語が含まれるか、あるいは〈思い出す〉という語が含まれるか、あるいは〈恋しい〉などの語が含まれている」という条件を満たす文書に「*過去」というコードが与えられる。ここでの文書は「文」、「段落」、「セル (Excel

データの場合)」三つともコード与えることができる。本研究は感情分析と併用するため最小単位である「文」を注目している。

[過去]コードの取り出し例

家族でタイ旅行に行ったことがあり、懐かしい場所が所々で確認できて家族で大変盛り上がりました。

以前の旅行を思い出しながら楽しませていただきました。

COVID-19 禍で海外に行けない現在の状況で、久しぶりに韓国の風景を観て心が癒されました。

このように、「。」で文書を区切り、コーディングルールに沿って、「過去」関連の文の取り出しができる。

一旦コーディングルールを作成すれば、そこから先は KH Coder がコーディングを行う。このような仕組みによって、様々な理論仮説や問題意識に応じた柔軟なコーディングを行った場合のことを考えれば、記述されたコーディングルールも格段に短く読みやすい、理解しやすいものとなる。これによってコーディングルールを開示した際に、第三者がそれを理解し、検討・検証することが容易になる。(樋口 2017)

4-4. データ収集

本節ではデータの収集方法、データの概要、クリーニングの手法を紹介する。

4-4-1. スクレイピング



图 4-1ctrip のブログ検索ページ

まず、データを収集するのに、分析に適するデータを探す必要がある。定量分析をす

るには一定のデータ量が求められ、なので、定量分析のできるブログの数を持つ温泉地の特定が必要である。そこで、筆者が目をつけたのは江戸時代から温泉地を大相撲の番付に見立てて格付けした温泉番付である。現在、温泉番付を作成しているのは BIGLOBE 旅行で、また、観光経済新聞社が 1 年に 1 度実施する温泉ランキング「にっぽんの温泉 100 選」も日本の主要的温泉地の特定の参考材料になる。

筆者は 3 月 29 日から ctrip のブログ掲載ページで、BIGLOBE 旅行第 13 回温泉大賞日本温泉番付の順番で、例：箱根、草津など、各温泉地名を検索、ヒットしたブログが 100 件以上の温泉地を対象に、ヒットしたブログのリストページで Google のブラウザ chrome 拡張機能の Data Miner を使用し、スクレイピングで URL などの情報を収集する。まず、ブログリストページで各ブログの URL、タイトル、日付、ユーザー名などのデータを収集できるレシピ(データ抽出手順)を作成し、スクレイピング機能で自動的にページを巡ってもらい、URL のデータなどを集めた。

取得した URL とタイトルを目視確認で広告文か宣伝文といったブログではない文章を判断し、無効なデータとして削除した。

処理した URL をもとに、ブログページ内の情報を chrome 拡張機能の Data Miner を使い、本文、投稿時間、タイトル、ユーザー名などのデータを収集できるレシピ(事例)を作成し、クローリング方法を通して、自動的に予め設定した URL リストにアクセスし、レシピ通りにブログテキストなどのデータを収集した。

レシピを作成するには、まず図 4-2 のようなブログページ上で、自分が必要とする要素を特定する必要がある。例えば、ブログのテキスト本文を採取しようと、Data Miner



图 4-2 ctrip のブログ画面

でページ上のテキスト本文部分を選択すると、選択された部分は図 4-2 のように点線で囲まれる。

次に図 4-3 のように Data Miner で採取したデータは自分が求めるデータかどうかを確認し、一致していれば、レシピに記入する。これで、クローリング時に異なった URL でもレイアウトが図 4-2 のレシピ作成ページのレイアウトと同じな場合、同じ位置にあるブログテキスト本文のデータも自動で採取される。またブログ本文だけでなく、ユーザー名や、ブログタイトル、旅行時期など、ページ上に記載されたものなら、レシピに記

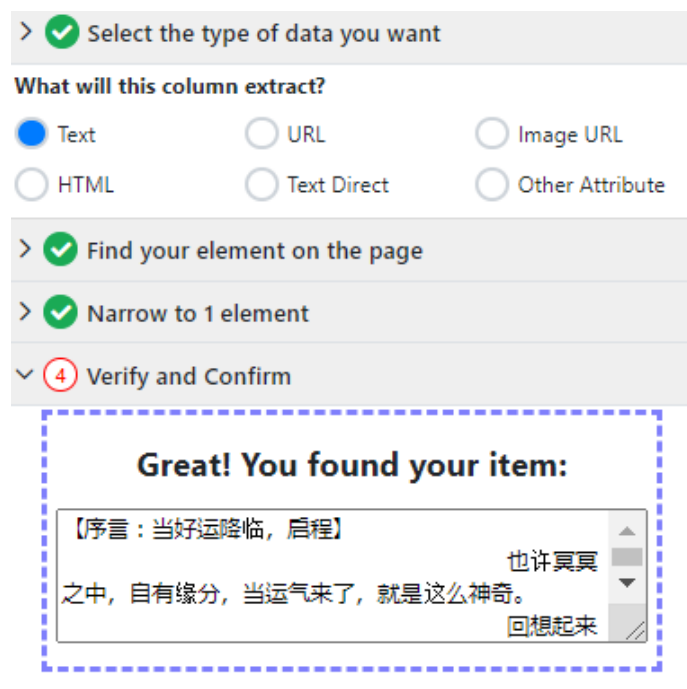


図 4-3 Data Miner のレシピ作成画面

入でき、採取できることは、スクレイピングもしくはクローリングの良さでもある。

こうしたステップを 14 個の温泉地で繰り返し、結果 14 個の温泉地に訪問経験のあるブログを収集した。

(1) データ概要

URL 個数：3996、ブログタイトル、ブログテキスト、投稿日、旅行シーズン、旅行コスト、経由地、同伴者、旅行タイプ（個人か団体）。

収集の対象期間が 2001 年 1 月から 2021 年 7 月までの間である。収集期間が非常に長いので、各ブログページのレイアウトが変化しており、収集できないデータもしくは未記入のデータもあります。

ブログの文字数は 400 文字から、40000 も超えるものもある。

4-4-2. データ処理

(1) 温泉記述

収集されたブログは温泉地に訪れていたが、ほとんど訪日旅行全体を記したブログである。そのため、温泉旅行を中心に研究するのであれば、ブログで温泉記述を旅行全体の記述から切り取る必要がある。

切り取りの方法は以下である

1. Excel の FIND 関数（FIND 関数は、指定された文字列を対象文字列で検索し、その位置を数字で返す関数である）を使用しブログ本文に関連する温泉地名、例「箱根」、あるいはキーワードである「温泉」が出現したかどうかを判断する。温泉地名と「温泉」が出現しなければ、温泉地に訪れていない可能性が大きい

ので、無効データとして廃棄する。

2. 温泉地の言及と温泉についての語の出現が確認された以上 RIGHT 関数と LEN 関数を使用する。

RIGHT 関数とは、文字列の末尾から指定された文字数の文字を返す関数である

LEN 関数は文字列の字数をカウントする関数である。

前のステップと合わせて以下の式が組み込める。

温泉言及以降文章=RIGHT(ブログ, LEN (ブログ) - 「温泉／温泉地」位置数値 + 100)		
↓	↓	↓
ブログ字数	FIND 関数結果	誤差

例 : FIND (“温泉”, a1) = 588 588 文字目に「温泉」が出現する

$$\text{温泉言及以降文章}=\text{RIGHT}(\text{a1}, \text{LEN}(\text{a1}) - 588 + 100)$$

文章の末端から「温泉」出現以前の文章を削除し、のちの目視確認のため「温泉」
以

前の 100 文字を残す。KH Coder が「。」で文書を文単位に分割し、文ごとの文字数の
平均値と中間値が 100 に近いので、この計算式の誤差を 100 字にしたのである。

この方法で大まかな温泉地旅行記述の位置を特定できる。

3. ステップ 1, 2 の作業を繰り返し、さらに、目視確認で精確な温泉地到着から離れるまでのブログ文を切り取る。

(2) 紹介文広告文の削除

変数を統一するため、温泉地で宿泊した記録のみを研究対象とする。したがって、目視確認を行う際に、まず宿泊をせず日帰りの記述を除外した。また、旅行記録ではない観光地の紹介文、自分の SNS アカウントを宣伝する広告投稿なども目視確認で判断し、研究対象から外した。



図 4-4 クリーニング後の各地ブログ数

図 4-4 クリーニング後の各地ブログ数

4-5. テキスト分析

4-5-1. 内容分析

内容分析とは「データをもとにそこから（それが組み込まれた）文脈に関して再現可能でかつ妥当な推論を行うための一つの調査技法」である。（Krippendorff 1980）

まず、コミュニケーションは人間関係や社会を成り立たせるためには不可欠である。こうしたことから、コミュニケーション内容を分析するための方法である内容分析は分野を問わず、多様な研究に応用できる。（Riffe et al.1998: 13）

次は、長期間にわたって保存・蓄積されてきた「内容」を分析対象とできることも、内容分析の特長である。（Holsti 1969: 15）本研究で使用する `ctrip` でのブログデータは 2001 年から 2021 年までの 20 年のデータである。13 年にわたっての質問紙調査やインタビュー調査は非常に難しく、効率にも問題があり、よって内容分析がより適切である。さらに、内容分析は調査対象者に気づかれにくく、実験法や質問紙調査の場合と比較し

て、内容分析ではデータを産み出す際に束縛が少ない。したがって、そうした束縛によってデータの妥当性を損なう危険性が低い (Krippendorff 1980=1989: 35-6) 。豊田ら (2003) は企業におけるマーケティング分野でのテキスト分析の活用を議論した。企業に対してブランドイメージが重要であり、ブランドイメージ通常の調査方法は質問紙調査であるが、質問紙では調査できない場合もある。例えば、新CMで視聴者が持つイメージと企業が表したいイメージが一致しているかどうかは、質問紙調査での追加刺激 (選択肢など) なしの純粹想起を取る必要がある。

本研究で使用したブログのテキストデータもそうした追加刺激がないデータである。

4-5-2. 計量テキスト分析

計量テキスト分析とはインタビューデータなど質的データをコーディングなどの手法によって数値化し、計量的分析手法を適用して、データを整理、分析、理解する方法である。(秋庭・川端 2004) 計量分析をすることによって、単なるテキストを読んでいるだけでは気づきにくいデータの「潜在的理論」を発見できる可能性がある。(川端 2003) さらに、コンピューターを使用することによって、研究者は手作業からの負担を軽減でき、より多くのデータを分析できるというメリットもある。

ただし、テキスト分析において、量的方法と質的方法は互いに不可分なものである。(Lazarsfeld & Barton 1951) 可能な限り量的分析を行ったとしても、研究の様々な段階で質的作業が必要となる。量的方法にこだわりすぎて、解釈しにくい分析結果を得るケースも少なくない。それを避けるためにも、質的方法は欠かせない。本研究において、コーディングルールの作成はそうした計量テキスト分析において欠かせない質的方法である。

4-6. 感情分析

感情分析は、指定されたテキストを調べて、そのテキストの背景にある感情的な考え方を分析する手法である。具体的には、執筆者の考え方がポジティブか、ネガティブか、ニュートラルかを判断する。

本研究では Google Cloud Platform の Natural Language API を使用して感情分析を行った。Google Cloud Platform とは、Google が提供しているクラウドコンピューティングサービスである。Natural Language API とは Google の機械学習を使用して非構造化テキストから分析情報を引き出す機能である。

感情分析の **score** は、ドキュメントの全体的な感情を示す。ドキュメントの感情分析の **magnitude** は、そのドキュメントに感情的な内容がどのくらい含まれているかを示す。この値は、ドキュメントの長さに比例する傾向がある。

例：私よりりんごが好きです。

score: 0.9

magnitude: 0.9

私はりんごが嫌いだ

score: - 0.9

magnitude: 0.9

本研究はブログを対象としたため、クチコミやレビューなどによく見られる星付けによる評価点データがなく、ブロガーの観光体験に対する直観的な評価データが得られないというデメリットがある。

それを補うため、筆者は感情分析を使用し、感情スコアから観光体験に対する評価の参考になると考えた。また、感情分析の対象はより短い文章であるほど、精度度が上がる。ブログのような何百文字さらに何万字もののデータを感情分析しても、ポジティブ感情とネガティブ感情が混合した感情スコアしか得られない。score の区間は - 0.9 から 0.9 しかないため、混合した感情でも、- 0.9 から 0.9 までの数値しか取れないというデメリットもある。また、長文章であるため、magnitude 値が 70 以上に達した場合もある。そういった結果では、さらなる解釈は難しい。つまり、Natural Language API の感情分析は、ブログ全体を分析することが難しいのである。

ここで、例文のような文単位のテキスト、つまり、「。」で区切った1センテンス単位での感情分析は magnitude 値と感情 score 値が近く、一番適切だと考えられる。

KH Coder では「。」で長い文章を文（センテンス）に区切る機能をもっている、さらにコーディングと合わせて使用すると、テーマ別の文を文章から切り取ることが可能である。

例えば、「部屋」、「チェックイン」などの語が出現した文書（この場合は文）に「*ホテル」というコードを与える。KH Coder は「*ホテル」に付与された文を大量のブログデータから切り出すことが出来る。結果、「部屋」や「チェックイン」などホテルに関連する短い文書データが得られ、感情分析に加え、ブログの「ホテル」というテーマに関する感情データが算出できる。

ただし、感情分析の結果は受動的であり、クチコミやレビューなどの評価点は能動的である。野中ら（2018）は訪日中国人観光客のオンラインホテルレビューを対象に、感情辞書による感情分析結果と評価点の関係性を分析し、結果は、感情分析の Positive 率、Negative 率ともに星の数による採点との関連性が非常に低いことを明らかにした。そのため、評価点のみ、または、感情分析のみを観光客の満足度を測る指標として利用することは適切ではないことが示唆されたが、本研究の場合、能動的な評価データがないため、受動的データであるブログの感情分析だけでも学術的価値があると考えられる。

第5章 訪日温泉旅行ブログテキスト分析

旅行ブログは観光スポットやホテルのクチコミに比べて、ユーザーのより詳細な記述が得られ、より細やかな分析ができるという利点がある。ただし、分析にあたり、テキストデータが膨大すぎるというデメリットもある。定性的性質を持つテキストデータについては、様々な視点があり、すべての視点を分析するには現実性が欠ける。したがって、本研究ではブログのテキストデータを温泉旅行の主な構成要素に基づき、テーマ別のコーディングルールを作成し、ブログテキストデータをテーマ別に分割して、分析を行った。

具体的なテーマは以下の五つの切り口である

- ホテルに関する記述
- 温泉に関する記述
- 食事に関する記述
- 交通に関する記述
- 日本に関する記述

この五つのテーマは中国人ブロガーの温泉地記述のほとんどの内容であり、各温泉地でも欠かせない内容であるため、コーディングルールはすべての温泉地に適応する。

★：本文でのグラフ、表、コーディングルールは筆者により和訳したものである。

5-1. 温泉地記述と旅行全体記述との対照

本研究では温泉地での体験が、温泉地に訪れた訪日中国人ブロガーの訪日旅行全体における位置付けを判明するため、すなわち温泉旅行体験が訪日旅行全体に比べて評価や感情に差があるかどうかを研究するため、温泉地の記述と旅行全体記述をコーディングルールのもとにブログの内容を分類し、テーマ別の感情分析を行い、温泉地と旅行全体の感情スコアと評価点を対照した。

(1) コーディングルール

本研究ではコーディングを用いて各ブログの注目すべき内容にコードを与え、またコードが付与された文書の取り出しも可能のため、ブログの内容をコード別（テーマ別）に分けることができる。

温泉地記述と旅行全体記述を比較するため、以下のコーディングルールを作成した。

表 5-1 コーディングルール一覧（日本語翻訳）

コード名	コーディングルール
ホテル	朝食, ホテル, 旅館, 部屋, 夕食, チェックイン, チェックアウト
温泉	温泉, 風呂, 浴衣, 入る (温泉に), 湯, 泉, 湯舟, 露天
食事	料理, 懐石, 会席, 朝飯, 朝食, 晩飯, 晩餐, 昼飯, 昼ごはん, レストラン, ラーメン, 茶碗, 皿, 食べる, 飲食
日本	日本, 日式, 和風, 和式, 日本人, 和服, 和室, ※東瀛
交通	バス, ケーブルカー, 駅, バス, 荷物, 電車, 下車, 新幹線, 乗る, 車

※東瀛（とうえい）とは、古代中国において、仙人の住むという東方の三神山(蓬萊・方丈)の一つを意味し、転じて日本を指す雅称である。現在でも漢民族は日本のことを東瀛とも言う。

本研究で使用したコーディングルールは筆者が頻出語をテーマによって分類し、その他二人の中国人研究者を依頼し、中国語の分類の確認に協力してもらい、日本語に訳したのが表 5-1 である。コーディングルールの原文は中国語で、KH Coder はコーディングで語を検索し、コーディングルール中に含まれた語が出現する文書にコードを付けるので、同義語なども、全部コーディングルールに入れる必要がある。しかし、本文で日本語に翻訳する場合、中国語での同義語が同一の日本語にしか訳せないため、コーディングルールが重複に見える場合もある。よって表 5-1 では重複の表現を省略したケースもある。実際に分析時に使用したのは表 5-2 の中国語原文である。「朝食」と「夕食」はホテルに出現しているのは温泉地におけるホテルに関する評価に朝食と夕食が欠かせないと考えたからである。

表 5-2 コーディングルール一覧（中国語原文）

コード名	コーディングルール
ホテル	早餐, 旅馆, 晚餐, 房间, 早饭, 晚饭, 入住, 退房, 酒店, 旅店, 民宿
温泉	温泉, 風呂, 浴衣, 泡, 汤, 泉, 池子, 露天
食事	料理, 怀石, 会席, 早饭, 早餐, 晚饭, 晚餐, 午饭, 午餐, 餐厅, 饭店, 饭馆, 拉面, 碗, 盘子, 吃, 饮食
日本	日本, 日式, 和风, 和式, 日本人, 和服, 和室, 东瀛
交通	大巴, 缆车, 站, 登山, 车站, 巴士, 站, 行李, 电车, 乘, 新干线, 列车, 火车, 车, 车子, 下车

表 5-3 各コードの数と温泉地記述における割合

コード	頻度	パーセント
*交通	2997	15.92%
*ホテル	3789	20.13%
*日本	2396	12.73%
*温泉	4368	23.20%
*食事	2086	11.08%
(文書数)	18824	83.06%(sum)

表 5-4 各コーディングの類似度行列 (Jaccard 係数)

	*交通	*ホテル	*日本	*温泉	*飲食
*交通	1	0.124	0.062	0.08	0.067
*ホテル	0.124	1	0.149	0.223	0.249
*日本	0.062	0.149	1	0.151	0.112
*温泉	0.08	0.223	0.151	1	0.099
*飲食	0.067	0.249	0.112	0.099	1

表 5-3 のように五つのコードが全文書中の割合がそれぞれで、また表 5-4 の通りコードそれぞれの類似性が低く、従ってこの五つのコードが温泉地の記述の大部分を占めている。Jaccard 係数については (3) で紹介する。

(2) 感情分析

本研究では感情分析の結果、感情スコアを、ポジティブ感情とネガティブ感情に分けて、各テーマ (コード) に関する文がポジティブ感情とネガティブ感情との共起が高い語を見る。そうした方法から、各テーマにおける、ポジティブ感情とネガティブ感情を引き起こす要素がわかる。それから、ポジティブ感情とネガティブ感情がブログではどう書かれていたのかを見ることが出来る。

本研究では、感情スコアが 0 以下の文をネガティブ感情と判定し、感情スコアが 0.5-0.9 の文をポジティブ感情と規定する。

さらに、ポジティブ感情とネガティブ感情を外部変数に設定し、KH Coder の外部変数別の特徴語という機能を使い、KH Coder は外部変数との共起が高い語から順に出力でき、ポジティブ感情とネガティブ感情に分類された文の特徴語を出力できる。

さらに、品詞別にみることも可能である。ポジティブ感情との共起が高い名詞はポジティブ感情を引き起こす対象である可能性が高い。鈴木ら（1989）とセナ（2010）は日本庭園を対象に外国人観光客の庭園像を分析するため、イメージ調査を行った。結果、外国人は日本庭園の様式や要素 に対する知識に乏しく、限定された要素、庭園の雰囲気 から庭園を評価する傾向にあることが示された。また、安原ら(2021)は旅行口コミサイト「tripadvisor」に投稿された日本語および英語のコメントを分析し、出現頻度上位100語では、日本語の方では名詞が多く、英語の方では形容詞が多かったことから、外国人は固有名詞に対する知識は乏しいものの、庭園に対する評価、感じ方、イメージを表現する形容詞が豊富であり、豊かな訪問体験が得られていることを指摘した。本研究では日本語との比較をしなかったが、名詞の分析から、中国ブロガーまたは中国人観光客が日本での温泉旅行についてどれくらい認識し、どれくらいの知識を持ったかが伺えた。中国人観光客が温泉旅行において何を体験し、体験し損ねたものも何かを知ること、観光地に示唆を与えることができると考えられる。

一方、ポジティブ感情との共起が高い形容詞は、ポジティブ感情が文書の中での表しである可能性が高く、本研究では各テーマにおいてポジティブ感情とネガティブ感情との共起が高い名詞と形容詞を分析することで、中国人ブロガーが何に対して感情があり、その感情がどう表したのかが見える。

(3) KH Coder の外部変数特徴語機能 (Jaccard 係数)

前節で述べた KH Coder の外部変数特徴語機能の原理は以下である。

外部変数の特徴語とは外部変数との関連語である、「関連語検索」という機能を使えば、特定の外部変数を与えられた文書と強く関連しているのはどんな語なのかを探索できる。この関連はデータから推測されるものであり、基本的には条件付き確率の計算によって推測が行われる。すなわち、「特定の外部変数を与えられること」といった条件を与えられたときに、そういった条件を満たす文書がまず検索され、それらの文書の中に特に高い確率で出現する語がリストアップされ、Jaccard 係数の値の降順で表示される。

Jaccard 係数は語が共起しているかどうかを重視する係数であり、一つの文書の中に語が1回出現した場合も10回出現した場合も単に「出現あり」と見なして、語と語の共起をカウントする。また Jaccard 係数には、語 A と語 B のどちらも出現していない文書が多数あったとしても、それによって、AB間の係数が大きくなるという特徴がある。語 A 語 B のどちらも含まない文書があるからといって、AB間の関連が強いとは見なさないというこの特徴は、スパースなデータの分析に適している。

(Romesburg1984=1992) ただし、本研究で使用する外部変数の「特徴語」検索は外部変数と文書の中の語との共起を計算し、Jaccard 係数をえる、また KH Coder における共起関係のデフォルトセッティングが Jaccard 係数である。

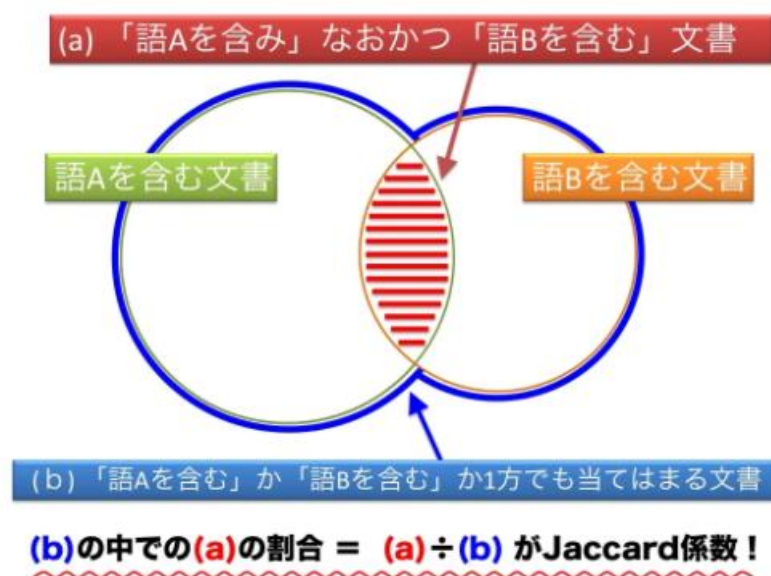


図 5-1 KH Coder 公式説明スライド

5-1. 目的地活動エリア別の比較

前節では、すべてのブログデータの温泉記述を対象に分析を行った。目視でブログデータを見ている間、筆者は箱根、熱海など関東を活動エリアとしたブログと別府、由布院、黒川温泉を活動エリアとしたブログ、並びに北海道などそれぞれ特徴があるのではないかと考えた。例えば、別府、由布院などを目的地としたブログは温泉へのこだわりが多く、箱根が代表する関東エリアが一番多くのブログの数をもち、ブログの掲載期間も一番長く、中国人が最も知るエリアでもある。その一方で、関東エリアに比べそれほどメジャーではないエリアを訪れたブログにはどういった特徴があるのかを見るため、活動エリアという視点から分析することとした。

外国人観光客にとって空港を利用することはほぼ不可欠である。また、利用空港の所在都市を拠点として活動するのが一般的であるとみられる。これに基づき、筆者は Excel を使い、国土交通省令和元年国際航空旅客動態調査での中国と直通便のある空港

名を FIND 関数で検索し、全ブログでの出現回数を統計した。統計結果が表 5-4 の通りである。

表 5-5 空港別のブログ数

温泉地／ 空港（件 数）	新千歳	福岡	佐賀	成田	羽田	関西 国際	名古屋中部
伊豆				1	2	1	
下呂		1				1	3
鬼怒川				1			
指宿		27	1				
城崎					1	1	
草津					1		
定山溪	8			4	1	1	
登別	66						
洞爺湖	29	2		8	7	3	1
白浜						1	2
箱根個人	1	4		71	32	6	
箱根ツアー	3			43	18	6	1
有馬		2		1	2	7	
由布院		30	8				
黒川温泉		11					
別府	1	62	4	4	3	1	
熱海	1	1		7	3	1	
霧島		4					
総計	109	144	13	140	70	29	7
	グループ 1	グループ 2		グループ 3			

表 5-4 から中国人ブロガーが主に利用した空港は新千歳、福岡、佐賀、成田、羽田であることがわかる。グループ 1 には定山溪、登別、洞爺湖が目的地で、利用空港は新千歳空港である。グループ 2 は由布院、黒川温泉、別府、指宿を訪問したブログ、利用空港は福岡と佐賀空港で、活動エリアは九州である。グループ 3 は箱根と熱海を訪問したブログで、利用空港は羽田と成田を中心とする、本研究では活動エリアは関東とする。また、前節の分析過程において、箱根が代表する関東エリアは中国人観光の定番ルートであることが見える。由布院、別府など中国人観光客の間ではそれほどの知名度を持たない温泉地のブログでは、湯めぐりなどの行動が見られた。北海道については、日本本州から離れており、文化も気候も本州や九州と異なり、中国人観光客の間では人気な地

域でもある。

以上のことから、本研究では北海道、九州、関東、三つの活動エリアの視点から、北海道、九州、関東を外部変数として設定し、名詞・形容詞×部変数の共起ネットワーク図を生成し分析する。共起ネットワーク図で中国人観光客のブログのそれぞれ特徴の相対的に特徴を把握することで、各温泉地、各地域の性格、あるいはそこに訪れた観光客の趣向がわかる。

5-2. 再訪意向

中国人ブロガーが訪日旅行を記録したとき、再訪についてどう語られたのかを、調査するために、以下のコーディングルールを作成した。

*再訪
(今後 and 来る) or (今後 and 日本) or (今後 and 温泉) or 次回

この、コーディングルールの意味は、文書に「今後」の次に「来る」あるいは「日本」あるいは「温泉」という語が出現した場合、或いは「次回」を含む文書に「再訪」コードを与える。こうした方法で「また日本にきたい」といった記述にたどりつけ、構文解析をすることで、どういう要素が中国人ブロガーの再訪に繋がるかを判明する。

第6章 分析結果

本章では第五章で紹介した分析方法のブログの感情別、テーマ別、活動エリア別の分析結果を説明する。

6-1. 温泉地全般

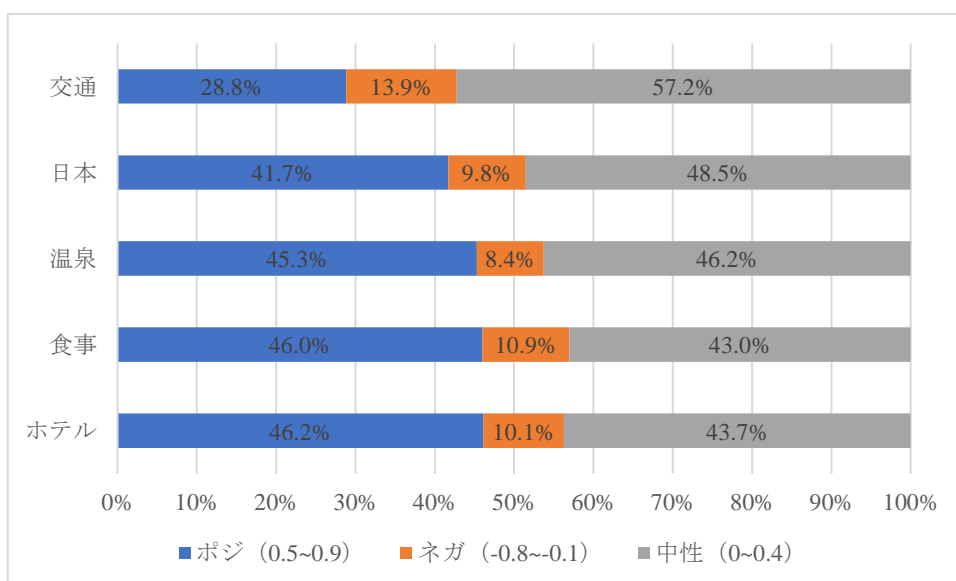


図 6-1 温泉地記述での各感情の割合

まず、五つのコードに対して、まとめて比較した結果が図 6-1 である。図 6-1 から見て、交通コードは五つのコードの中で、ポジティブな文書の割合が一番低く、ネガティブと中性的な文書の割合が両方とも一番高いことがわかる。次に温泉、食事、ホテルのコードでは三つともポジティブ文書の割合が50%に近く高い割合を持つ、その中、温泉に関する記述のネガティブ感情の割合が一番低いことがわかる。つまり、中国人ブロガーは温泉、食事、ホテルについての描写ではポジティブな描写が多く、交通面に対しては、感情を含まない単純な記録が多く、ほかのコードに比べて、ポジティブな描写も少ない。従って、訪日旅行において、どうやって交通に関する体験を観光客により深く、より良い印象を与えることが一つの課題になる。

6-1-1. 温泉地と旅行全体との対照

(1) 交通

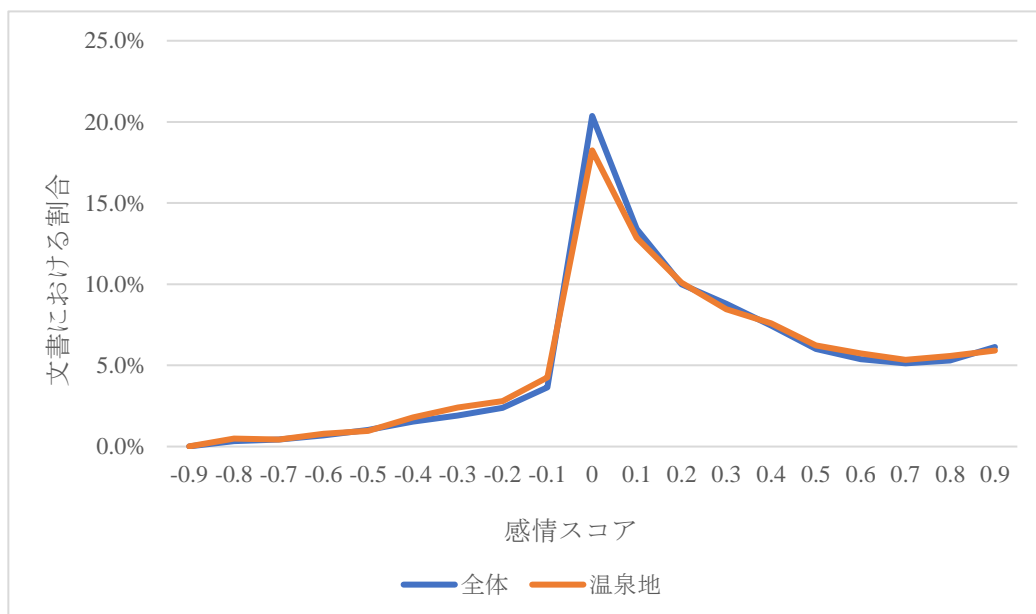


図 6-2 交通文における全体と温泉地記述の対照

図 6-2 からみると、交通に関しては、温泉地と旅行全体において目立った差がなく、中立的な感情が主に占めている。また、温泉地関連の交通が旅行全体に比べてとりわけ優れていることもなく、とりわけ改善が必要な点もないことがわかる。

次に、温泉地記述のポジネガ感情の特徴語を分析する。

表 6-1 交通文ポジ感情の特徴名詞

順位	ポジ名詞	全体	共起回数 (割合)	Jaccard
1	箱根	562 (0.246)	135 (0.262)	0.143
2	ホテル	479 (0.210)	124 (0.241)	0.143
3	温泉	345 (0.151)	100 (0.194)	0.132
4	駅	417 (0.183)	102 (0.198)	0.123
5	日本	199 (0.087)	62 (0.120)	0.095
6	ケーブルカー	289 (0.127)	69 (0.134)	0.094
7	登山	245 (0.107)	59 (0.115)	0.084
8	汽車	185 (0.081)	45 (0.087)	0.069
9	富士山	162 (0.071)	43 (0.083)	0.068
10	列車	169 (0.074)	41 (0.080)	0.064
11	風景	64 (0.028)	32 (0.062)	0.059
12	電車	157 (0.069)	36 (0.070)	0.057
13	場所	151 (0.066)	35 (0.068)	0.056
14	道	97 (0.043)	27 (0.052)	0.046
15	景色	42 (0.018)	24 (0.047)	0.045

ポジティブの関連名詞から見ると、「箱根」、「ホテル」、「温泉」など本文のなかでも頻出度の高い名詞を除いて、中国人ブロガーが評価している点は、「駅」、「ケーブルカー」、「登山列車」、「風景」などである。

表 6-2 交通文ポジ感情の特徴形容詞

順位	ポジ形容詞	全体	共起回数 (割合)	Jaccard
1	悪くない	48 (0.021)	26 (0.050)	0.0484
2	安い	29 (0.013)	16 (0.031)	0.0303
3	快適	35 (0.015)	15 (0.029)	0.028
4	美味しい	27 (0.012)	13 (0.025)	0.0246
5	近い	36 (0.016)	13 (0.025)	0.0242
6	綺麗	16 (0.007)	11 (0.021)	0.0212
7	静か	17 (0.007)	11 (0.021)	0.0211
8	きれい	23 (0.010)	11 (0.021)	0.0209
9	美しい	27 (0.012)	11 (0.021)	0.0207
10	高い	33 (0.014)	10 (0.019)	0.0186

また形容詞においては、名詞に比べて Jaccard 係数が低く、ポジティブ感情に分類された文書には形容詞が名詞に比べてあまり使用されていない、さらに、形容詞は感情スコアを大きく影響するので表 6-2 から見たポジティブ感情の割合が低いことも裏付けられる。

表 6-3 交通文ネガティブ感情の特徴名詞

順位	名詞	全体	共起回数 (割合)	Jaccard
1	車	266 (0.117)	14 (0.230)	0.0447
2	目的地	16 (0.007)	3 (0.049)	0.0405
3	霧	24 (0.011)	3 (0.049)	0.0366
4	バス	424 (0.186)	17 (0.279)	0.0363
5	荷物	230 (0.101)	10 (0.164)	0.0356
6	情報	3 (0.001)	2 (0.033)	0.0323
7	暴雨	5 (0.002)	2 (0.033)	0.0313
8	運行	6 (0.003)	2 (0.033)	0.0308
9	時間	211 (0.093)	8 (0.131)	0.0303
10	発車	8 (0.004)	2 (0.033)	0.0299

ネガティブ感情に関しては、頻出名詞と内容を確認したところ「霧」、「暴雨」のような悪天候に対しての文句があり、「情報」、「運行」は例えば大涌谷の運営停止に対して残念な気持ちを表すものが多かった、「時間」や「発車」などはバスや電車に遅れたことや、発車時間が遅れたなどのハプニングの記述である。

(2) ホテル

図 6-3 は温泉地記述と旅行全体のブログを「ホテル」コードで文単位に抽出、感情スコアを算出して作成した図である。各感情スコアの文の数を温泉地記述ホテル文の数と旅行全体ホテル文の数を割って、その割合が縦軸である。温泉地記述と旅行全体が「ホテル」テーマにおいて、ポジティブ感情とネガティブ感情の差があるかどうかを見るため、折り線グラフを作成した。

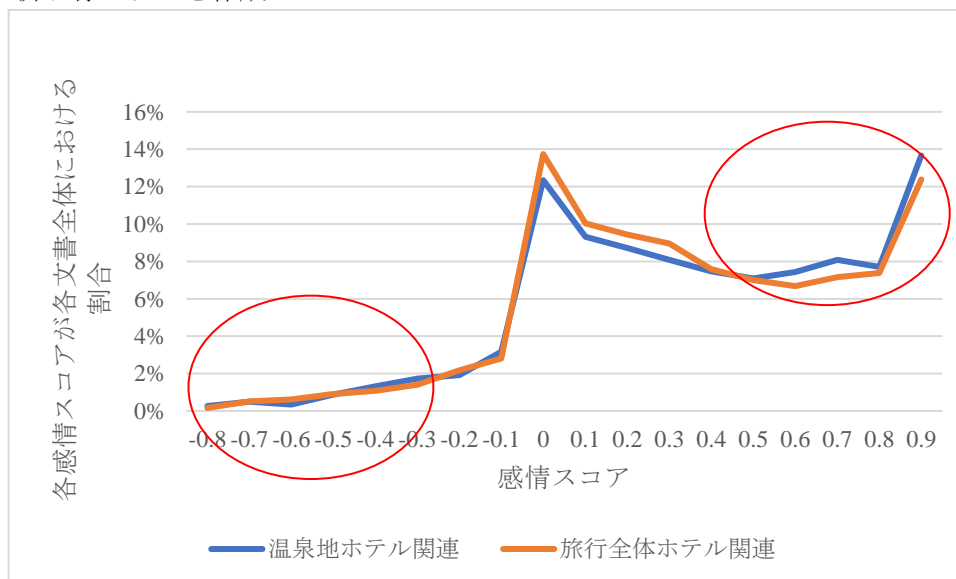


図 6-3 ホテル文における全体と温泉地記述の対照

図のように、ネガティブ感情において、温泉地と旅行全体には差がなく、ポジティブ感情では、温泉地記述の方が、ポジティブ感情の割合が少し高く見える。

つまり、温泉地でのホテル体験が旅行全体のホテル体験に比べて、より中国人ブロガーのポジティブ感情を引き起こせると推測できる。

表 6-4 ホテル文ポジティブ感情の特徴名詞

順位	名詞	全体	共起回数 (割合)	Jaccard
1	ホテル	643 (0.385)	262 (0.401)	0.2534
2	部屋	560 (0.336)	244 (0.374)	0.2518
3	温泉	569 (0.341)	238 (0.364)	0.2419
4	夕食	304 (0.182)	137 (0.210)	0.1671
5	朝食	292 (0.175)	125 (0.191)	0.1524
6	日本	187 (0.112)	78 (0.119)	0.1024
7	レストラン	76 (0.046)	34 (0.052)	0.0489
8	畳	78 (0.047)	33 (0.051)	0.0473
9	特色	48 (0.029)	31 (0.047)	0.0463
10	味	52 (0.031)	30 (0.046)	0.0444
11	サービス	58 (0.035)	25 (0.038)	0.0364
12	洞爺湖	41 (0.025)	23 (0.035)	0.0343
13	料理	43 (0.026)	23 (0.035)	0.0342
14	風景	47 (0.028)	23 (0.035)	0.034
15	バルコニー	38 (0.023)	22 (0.034)	0.0329

ホテル文のポジティブ名詞では、コーディングルールに含む語を除き「畳」、「バルコニー」などは部屋の内部にある設備に対する記述である、特に畳は日本人の角度から見れば見慣れた光景であるが、中国人にとって畳は日本文化を感じさせるものである。さらに、都会のホテルと違い、温泉地のホテルは和室が多いので、温泉地で初めて畳を経験する可能性が高い。なので、印象が強く、ポジティブな名詞の中でも目立っている。また、「サービス」との関連も強いので、中国人がサービスを評価していることが推測できる。その他「匂い」、「料理」、「夕食」などの語が大きな割合を占めていることから、中国人は食事に対しての描写が多く、しかも評価していることがわかる。

表 6-5 ホテル文ポジ感情の特徴形容詞

順位	形容詞	全体	共起回数 (割合)	Jaccard
1	悪くない	75 (0.045)	44 (0.067)	0.0643
2	快適	49 (0.029)	32 (0.049)	0.0478
3	精緻 (精巧)	44 (0.026)	27 (0.041)	0.0403
4	豊富 (食事の種類)	47 (0.028)	21 (0.032)	0.0309
5	きれい	32 (0.019)	20 (0.031)	0.0301
6	広々とした	21 (0.013)	14 (0.021)	0.0212
7	美味しい	30 (0.018)	14 (0.021)	0.0209
8	高い	25 (0.015)	13 (0.020)	0.0195
9	素晴らしい	18 (0.011)	12 (0.018)	0.0182
10	綺麗	18 (0.011)	12 (0.018)	0.0182

形容詞においては、まず「悪くない」というのは中国語では一番使われている評価語である。三番目の「精緻」は後節にて詳細に述べるが、懐石料理に使われていた食器や和風を感じるものに使う表現である。また、名詞と同じく「豊富」、「美味しい」など食事関連の評価が多い。五番目の「きれい」も、中国人がホテルの清潔さに対する評価である。さらに「広々とした」が述べていたのは、都会のホテルに比べて、温泉地のホテルの部屋が広いことから、よく言及されたと思われる。

ネガティブ感情に分類された文書は 33 件しかなく、特徴語からの分析が難しいので、手作業で内容分析を行った。

具体的には、体調不良といった個人の事情によるハプニング、「宿泊料金が高い」などの文句、ほかには「朝食が豊富すぎて、食べきれない、食べすぎた」といったプラス的な記述であったものの、ネガティブとして認識された例もある。クレームと言えるのは「ベッドがなく、畳で寝るしかない」、「ホテルの売店は営業時間になっても営業開始していない」というサービス面の不満と畳文化に対する不慣れがある。

以上のことから「畳」については、マイナス的评价があるものの、ポジティブである頻度が遥かに高いので、全体から見ると、中国人観光客は畳文化について日本文化の文化体験としてプラス的に受け取っている。畳に不慣れがある観光客の体験を改善するため、ホテル予約時に部屋が和室か和洋室なのかをはっきり明記し、中国人を含めて外国人観光客にしっかり了承させる必要がある。

(3) 日本

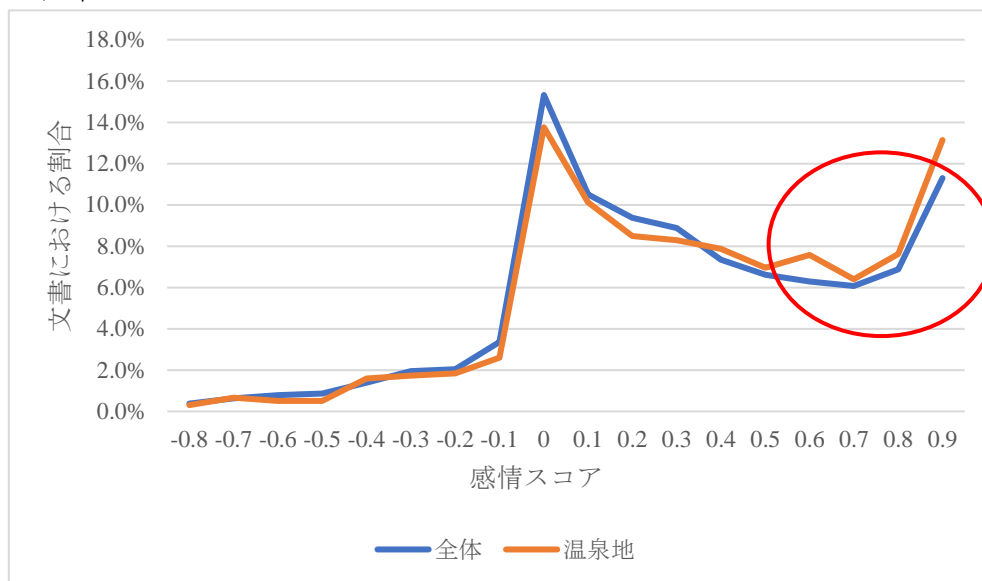


図 6-4 日本文における全体と温泉地記述の対照

前節のように、ネガティブ感情において、温泉地と旅行全体には差がなく、ネガティブ感情の割合が低い、ポジティブ感情では、温泉地記述の方が、ポジティブ感情の割合が少し高く見える。

つまり、温泉地での日本に対する描写が旅行全体のホテル体験に比べて、より中国人ブロガーのポジティブ感情を引き起こせると推測できる。

表 6-6 日本文ネガ感情の特徴名詞

順位	ネガ名詞	全体	共起	Jaccard
1	日本	1346 (0.293)	132 (0.346)	0.0827
2	日本人	203 (0.044)	35 (0.092)	0.0636
3	中国	74 (0.016)	11 (0.029)	0.0247
4	ガイド	129 (0.028)	12 (0.031)	0.024
5	時間	44 (0.010)	10 (0.026)	0.024
6	金	25 (0.005)	9 (0.024)	0.0226
7	国内	108 (0.023)	9 (0.024)	0.0187
8	時間 (単位)	21 (0.005)	7 (0.018)	0.0177
9	中国人	60 (0.013)	7 (0.018)	0.0161
10	従業員	66 (0.014)	7 (0.018)	0.0159

まず、温泉地での日本関連ネガティブ上位の名詞がほとんどコーディングルールの内容であり、主に「日本」と「日本人」の関連が一番高く、また、次に「中国」との関連

が強く、中国と日本を比較していることが見える。そしてここで初めて「ガイド」が出現する。団体ツアーにおいて、ガイドの役割が重要である。日本に関連する知識やしきたり、注意点、団体旅行のあらゆる面でもガイドの役割が大きい。ガイドのサービスに対する文句が日本との関連性も強い、ほかにも、ガイドの話聞いて、ネガティブ感情も起こりうる。例えば、「ガイドの話によると日本人は墓場を不吉だと思わない」、「ガイドの話によると神社の近くに墓場があるらしい」といったコメントがあり、墓場は中国人にとって忌む存在で、日本人はそうではないと聞き、生死観による文化の違いを感じ、ブログに表したのである。

次にネガティブ形容詞を見てみる。

表 6-7 日本文ネガティブ感情の特徴形容詞

順位	ネガ形容詞	全体	共起	Jaccard
1	価格が高い	31 (0.007)	6 (0.016)	0.0147
2	悪い	9 (0.002)	5 (0.013)	0.013
3	マナー違反	4 (0.001)	4 (0.010)	0.0105
4	頑固	2 (0.000)	2 (0.010)	0.0052
5	辛い	2 (0.000)	2(0.010)	0.0052

まずは、価格に対して、ホテル文にも出ており、中国人は価格に対する記述が多く、日本は中国より消費水準が高いので、価格はよく言及される。次に、ホテルやサービスに対してのクレームや、日本人の性格が頑固など、旅行中に日本人と接したときのフィードバックがある。「マナー違反」では、自分たちが海外の浴場でのマナーを重視して、日本でのマナーこうであるから、マナー違反しないようブログで呼びかけている。「辛い」は座敷が座りづらいとか、動きにくいなどの文句である。

ここで気になるのは、ブログでは「マナー違反」についての指摘が頻出しており、中国人が日本での「マナー」を気にしていると考えられる。感情分析ではマイナス的な感情に判定されたが、こうした気遣いは日本のマナーへの文句というより、ブログを見た人への注意を促す言い方の方が多い。かつては中国人のマナー違反が多く報道されていたが、こうしてブログのような SNS でマナーを重視する呼びかけも存在することが確認できた。

中国人観光客と日本人観光客が温泉地に訪れる際にの一番大きな違いは、中国人観光客には日本に対する評価が多いことである。中国人観光客から見た日本はどうか、温泉地から見た日本に比べてどう違うのか、それを細かく確認するため、旅行全体の日本関連記述と温泉地での日本関連記述の特徴語を比較し、分析を行った。

表 6-8 日本文における全体と温泉地記述の特徴名詞の比較

順位	温泉地ポジ名詞	全体ポジ名詞
1	日本	温泉
2	温泉	ホテル
3	ホテル	日本人
4	箱根	浴衣・着物
5	部屋	部屋
6	日本人	店
7	場所	特色
8	旅館	サービス
9	夕食	国内
10	着物	文化
11	朝食	美食
12	特色	味
13	料理	観光客
14	湯	感覚
15	味	朝食
16	レストラン	晚餐
17	感覚	歴史
18	浴衣	スポット
19	サービス	レストラン
20	畳	旅行

まず名詞から見ると、旅行全体で日本と一番関連の強いポジティブな名詞が「温泉」である。浴衣・着物も上位を占めている。そして、「日本の特色がある」などの記述も多いことから、「特色」も日本との関連が強い、旅行で感じた日本ならではの魅力を感じて書き込みに至ったと考えられる。また旅行全体において「文化」、「歴史」なども日本との関連が強く、中国人が日本旅行で「ここが日本だ」ということを実感させられるのは「文化」、「歴史」である。その他はホテルと同様、中国人観光客はホテルを拠点として活動するので、特に旅行初日ではあらゆるものが「これが日本なのね」と感じるもので、旅行全体において、ホテルと食事と日本には強い関連性を持つ。温泉地のポジティブ名詞には「料理」がある。ここの「料理」は中国語の原文でも「料理」で、中国語において「料理」は食事として使用する「日本料理」、「懐石料理」などを熟語として使用する傾向があり、温泉地に出現する「料理」は「日本料理」と「懐石料理」に指す語と考えられる。特に温泉地においては、確認したところ「懐石料理」と「会席料理」が頻出しており、普通の食事ではなく、中国人が「料理」だと思えるような料理が、日本に関連するポジティブ感情を喚起することが推測できる。

さらに、温泉地でのポジティブ名詞を見ると、温泉より日本の方が上位にあり、温泉地では中国人がポジティブに日本を想起する。その他の語はホテルと食事に集中し、浴衣と着物も温泉地で日本のイメージアップに繋がっている。

表 6-9 日本文における全体と温泉地記述の特徴形容詞の比較

順位	ポジ形容詞 温泉地	ポジ形容詞 旅行全体
1	悪くない	悪くない
2	精緻（精巧）	清潔
3	快適	旨い
4	清潔	精緻（精巧）
5	旨い	高い
6	高い	安い
7	有名	有名
8	静か	静か
9	美しい	快適
10	安い	簡単

形容詞が感情分析において比重が高い語であり、ブロガーの最も直接的な感情表現を表しているため、形容詞を見ることで、中国人ブロガーが旅行体験に対してどんな評価をしていたのかがわかる。温泉地と旅行全体の日本関連記述の関連度の高いポジティブ形容詞を比較した結果、一般的な評価「悪くない」では一致しており、「精緻（精巧）」、「清潔」、「旨い」などのプラス評価も両方ともに特徴的である。ただ、関連の強さから見ると、温泉地の方が「精緻（精巧）」との関連性が高く、つまり温泉地での記述では日本に関連する内容の中で「精緻（精巧）」という語がより使われていた。一方、全体のポジティブ日本記述の中で「清潔」との関連度がより高いである。訪日中国人は日本という話題について、旅行全体は「清潔」が代表したイメージが強く、温泉地では「精緻（精巧）」が「清潔」よりもプラス的なイメージであることが推測できる。

(4) 温泉

ネガティブ感情分析は内容を確認したところ、個人の事情、目的地別に分けており、すべての温泉地に対する一般的なマイナス評価がまとめにくいため、後節の目的地活動エリア別に定性的分析することとした。本節ではポジティブ関連の名詞と形容詞に焦点を当てる。

温泉地の温泉記述が全体での温泉に関連記述の感情分析結果がほぼ一致しており、こ

ここでは温泉地と全体の温泉記述の関連語分析結果をまとめて述べる

表 6-10 温泉文ポジ感情の特徴名詞

順位	名詞	共起回数
1	温泉	1174
2	ホテル	389
4	食事	156
5	部屋	125
6	風景	93
7	室外	76
8	特色	70
9	室内	68
13	匂い	47
14	観光客	47
16	足（足湯関連）	41
19	卵	32
20	風呂	32

まず名詞から見ると、中国人ブロガーが自然などを重視し「風景」や「新鮮（空気が）」をブログに記していた。「匂い」が浮上したのは温泉特有の硫黄の匂いが印象に残り、それをブログにし、体験をシェアしたのである。「卵」の頻数も多く、係り受けを確認したところ、箱根の黒卵と温泉卵を指していた。前節でも述べていたが、中国人は旅行中に特に食事に重視していて、中国の地熱資源も少ないことから、温泉卵の名前は聞いたことがあるが、本当の温泉で加熱した卵には魅力を感じる。また旅行に来たからには当地の名産物を食べなきゃ損という心理もあり、「卵」が多くの食品から浮上したと考えられる。

次に気になるのは「風呂」という語はもともと中国語の中にはなく、ほかにも「池（湯舟）」などの表現もあるが、中国人は中国語にある表現を使わず、あえて日本語の表現を使うのは奇妙な現象である。筆者はこれが中国人観光客が日本温泉文化に感心し、「これは日本の温泉」だから現地の言い方を使う、文化環境に合わせて使用したのではないかと考える。

表 6-11 温泉文ポジ感情の特徴形容詞

順位	形容詞	共起回数
1	気持ちいい	116
2	悪くない	85
3	きれい	67
4	綺麗	39
5	有名	33
6	旨い	24
7	精緻（精巧）	20
11	豊富 （食事の種類）	15
14	静か	22
15	安い（値段）	24
17	新鮮（空気）	11
18	趣がある	10

形容詞の方では、「気持ちいい」が一位で、中国人にとって異国にいる海外での温泉体験はリラックスを感じさせる効果があると考えられる。また、前節で述べたように、中国人は日本の「清潔」を評価しており、温泉地においてもその傾向が見える。ここにも「精緻」が出現しており、「精緻」は中国人が日本温泉に対する重要な評価で、後節で、「精緻」の概念と使用場面について詳細に分析する。

他にも、「風景」、「食事」、「匂い」、「気持ちいい」、「静か」などの語から見て、すべての五感が描写されていて、つまり、温泉旅行は中国人観光客の五感を刺激できる観光活動といえる。

これほど感覚を全般的に刺激できる観光活動はそうにないと考えられる。五感を刺激することによってより人に印象を残せ、再訪に繋がるのではないか。

ただし、温泉自体に関する描写はお湯の温度以外ほとんどなく、泉質などを気にしていたなどの記述が見当たらなかった。

(5) 食事

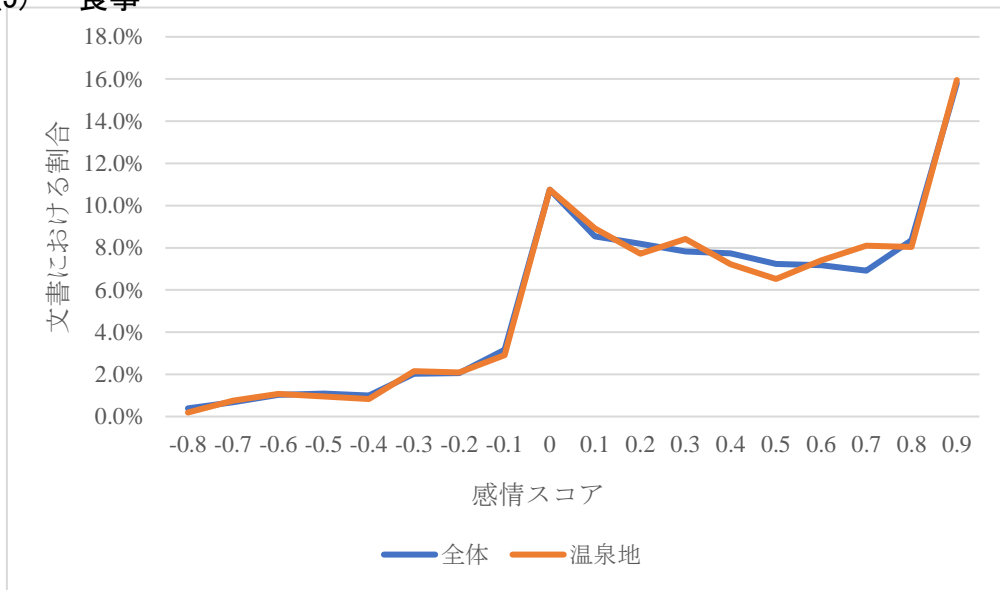


図 6-5 食事文における全体と温泉地記述の対照

図から見て、二本の折れ線がほぼ重なっており、つまり食事に関連する記述の感情スコアは旅行全体と温泉地では大きな差がないと見える。0.9 点の割合が二つともかなり高く、また、0.7-0.8 点の区間では温泉地の方が少し全体より上回る。つまり、中国人の食事に関して全般の評価が高く、温泉地は旅行全体に比べてポジティブ感情少し上回り、温泉地での食事に関する体験も、訪日旅行体験を引っ張ることが出来ると考えられる。

ネガティブ感情に関しては温泉と同様、後節で述べる。

表 6-12 食事文における全体と温泉地記述の特徴形容詞の比較

順位	温泉地ポジ形容詞	旅行全体ポジ形容詞
1	悪くない	おいしい
2	美味しい	悪くない
3	精緻 (精巧)	安い
4	豊富 (食事の種類)	精緻 (精巧)
5	高い	高い
6	きれい	有名
7	新鮮 (食材)	新鮮
8	安い	豊富
9	静か	きれい
10	美しい	シンプル
11	有名	静か
12	かわいい	かわいい
13	シンプル	きれい
14	香りがいい	美しい
15	さっぱり	こってり

まず、温泉地のポジティブ形容詞では、前節のようにやはり「精緻」が高い、和食は中国料理に比べておかずが多く、バイキングでも懐石料理でも品々の種類が多いことか

ら、「豊富」という評価が多いである。また刺身などの料理は鮮度が要なので「新鮮」も重要な評価点である。「静か」があるのは、日本人は食事中に静かで、レストラン全体の雰囲気を目指すのである。「シンプル」というのは、日本食は調理方法がシンプルであることを指す。「さっぱり」が出現するのは恐らく中国料理に比べてさっぱりであることを指す。

全体のポジティブ形容詞では「安い」、「高い」のような価格関連語が出てきやすい。理由を考えると、まず温泉地付近では飲食店が少ないことがブログでもよく指摘される。さらに、温泉地では食事つきプランを選ぶ傾向があるので、ホテルを予約する際に決済が済ませたので、食事の経験を記録する際に、価格について旅行全体のほど想起しないのである。

旅行全体では「静か」がないので、つまり温泉地の方が静かなイメージを与える。「こってり」はラーメン由来なので、温泉地でもラーメンは目立たないので、温泉地の「さっぱり」と対照的に「こってり」がある。

表 6-13 食事文における全体と温泉地記述の特徴名詞の比較

順位	温泉地ポジ名詞	全体ポジ名詞
1	ホテル	ホテル
2	温泉	レストラン
3	晚餐	晚餐
4	朝食	温泉
5	レストラン	ラーメン
6	日本	味
7	料理	料理
8	部屋	美食
9	味	美味
10	美味	ごはん
11	海鮮	部屋
12	卵	海鮮
13	美食	特色
14	ごはん	軒
15	特色	スープ
16	食材	北海道
17	懐石料理	定食
18	従業員	サービス
19	風景	スーパー
20	感覚	食材

まず、温泉地の名詞から見て、前節で述べたように、温泉地で出現する「料理」は本

格的な和食、懐石料理や会席料理を指すことが多い。温泉旅行の一つ大きな特徴は宿泊施設で会席料理が食べられることである。「料理」と「懐石料理」の両方も出現しており、しかも「料理」が上位に出現していることから、中国人観光客は温泉地での懐石料理には高く評価していることがわかる。次の「部屋」は、確認したところ部屋食を指すのがほとんどである。中国の宿泊施設では部屋食が出来ない所がほとんどで、あってもルームサービスで提供する軽食である。日本の旅館のような部屋食のためのサービスと施設がないため、部屋で本格的な料理が食べられることも文化体験の一つである。したがって、中国人は部屋で食べる懐石料理を高く評価していることがわかる。料理の「味」についてももちろん注目するが、「風景」が出ていることから、食事中に見られる風景にも高く評価しており、風景の見えるところで食事ができるのも旅行体験向上につながる。

もう一方、旅行全体の名詞を見てみると、形容詞分析にも言及していたが、温泉地以外のところでは「ラーメン」の出現頻度が大幅に上昇し、中国人は温泉地以外のところでの食事選択傾向が少し見えてきた。また「スーパー」についての描写があり、温泉地以外ではスーパーを利用して食事済ませる傾向がある。

以上のことから、温泉地での食事は懐石料理と部屋食が高く評価されており、ただし、温泉地以外の場所に比べて選択できる余地が少ないことが少し見えてくる。観光客をより満足させるには、食事の選択肢をより多く提供してもいいが、筆者は、温泉地のブランドイメージを固めるには、懐石料理や郷土料理など、現地でしか味わえないものを提供した方が、一般的な観光地との差別化でき、温泉地自身の個性を保つことができると考える。

6-2. 目的地活動エリア別の比較

本研究では北海道、九州、関東、三つの活動エリアの視点から、中国人観光客のブログを分析し、それぞれ特徴の相対的に特徴を把握することで、各温泉地、各地域の性格、あるいはそこに訪れた観光客の趣向を明らかにする。

6-2-1. 交通

表 6-14 エリア別のネガティブ感情文

エリア	ネガティブ感情文交通関連	スコア
北海道	唯一残念だったのは、駅に着いたら、スタッフがホテル行のシャトルバスがなくなったので、タクシーで行くしかなかった、2000 円くらいだった。	-0.2
	アナウンスでは「洞爺温泉街行のバスが間もなく発車する」を聞いたが、温泉街の位置がわからなくて、少し躊躇したらバスが行ってしまった。	-0.1
関東	路線が複雑すぎる、まず小田急線から小田原へ、それから箱根登山線で箱根湯本まで、最後にバスを乗る、バスの路線が同じ色のバスが何本も走っていて、自分は少し日本語がわかることで助かったが、さもなければ大変になるところだ。	-0.2
	天気のせいでケーブルカーが止まり、バスに変えて登るしかないのは苦痛だ。	-0.5
	東京から伊豆の路線は結構面倒で、何回も乗り換える必要がある。	-0.7
九州	由布院の森みたいな人気列車には自由席がなく、予め指定席の切符を購入する必要がある、当日で購入するなら指定席がない状態で特急にのる覚悟が必要だ。	-0.2
	時間が過ぎるのが早すぎて惜しい、あっという間に指宿駅に入った。	-0.2

各地に関する交通面のネガティブ記述をまとめると、北海道エリアではバスの乗り遅れやバス便が少ないことが挙げられる、関東では乗り換えが複雑で、路線図がわかりづらいといった点が挙げられる。九州では指定席の購入についての心配がある。また最後の一文のように、意味合い的にはプラス的な評価に解釈できるが、「惜しい」や「あっという間」があるせいで、マイナスに判定されたケースもある。

次に、ポジティブ文書をエリアで振り分け、エリアの外部変数を作成する。そして、外部変数と頻出名詞・形容詞の共起ネットワーク図を生成することで、各エリアの相対的に特徴を把握できる。共起ネットワーク図のセッティングは KHCoder のデフォルトである上位 60 個で描画する。

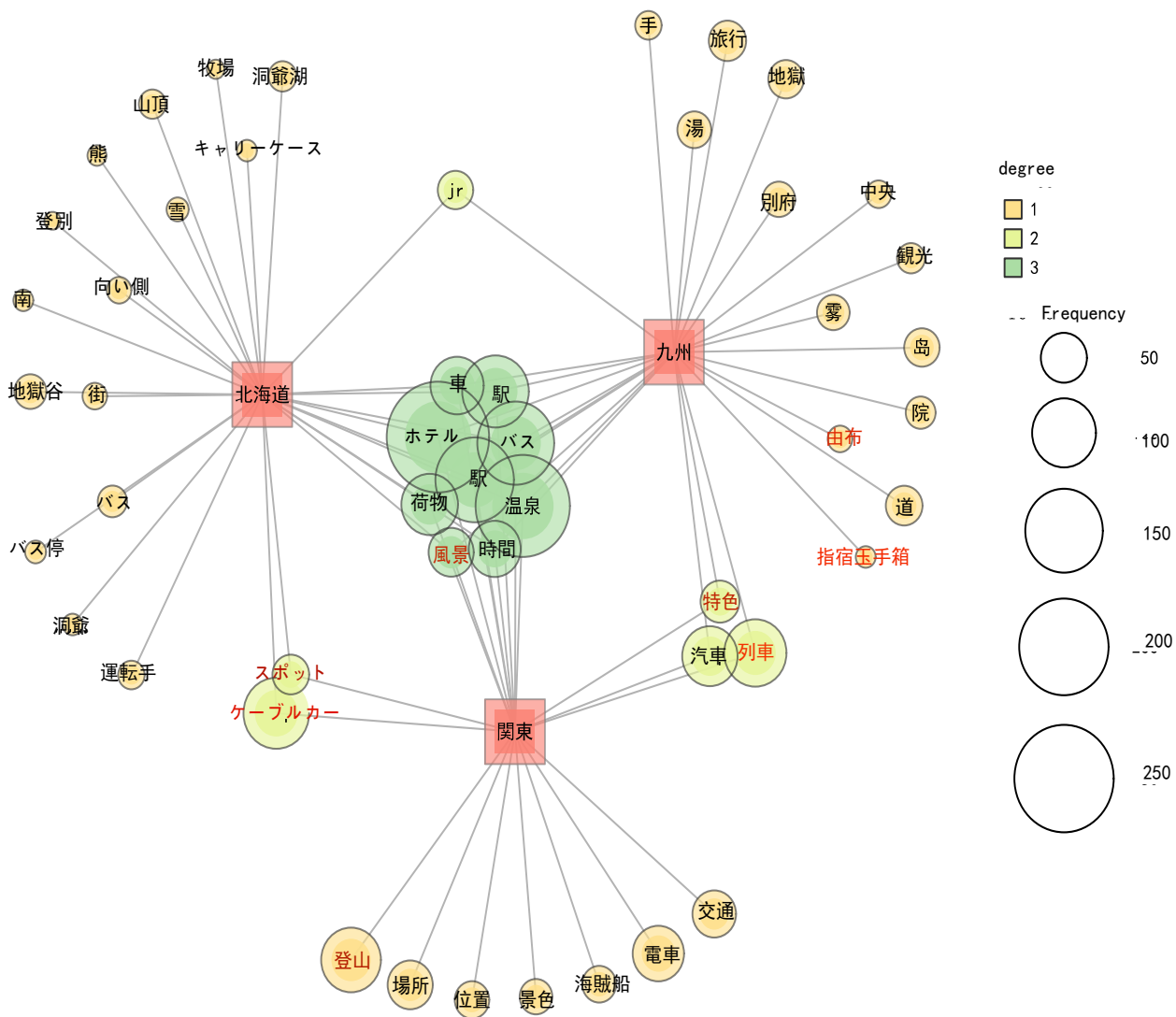


図 6-6 ポジティブ感情交通文 x 外部変数：名詞・共起ネットワーク

まず、九州と関東の共通に関連している語には「列車」がある。主に、登山列車と指宿玉手箱列車のことを指す。「特色」としては例えば「玉手箱列車が特色である」などの描写がある。九州と関東エリアの温泉地プロガーが列車をポジティブに感じ、特色のあるものを期待していると推測できる。

次に、北海道と関東の共通に関連している語には「ケーブルカー」がある。これは、大涌谷と登別のケーブルカーを指すものである。「スポット」があるのは恐らく、北海道と関東は観光客が多いので、観光スポット意識が強いのではないかと考えられる。

そして関東独自の共起特徴としては「登山」があり、「箱根登山列車」を指す。

九州独自の共起特徴には「由布院の森」と「玉手箱」があり、これも列車関連のものである。従って、中国人ブロガーが交通面では地方の特色のある列車に関心があることがわかる。

最後に、全体的に共起のある部分には「風景」が挙げられた、これは移動中の風景が評価されたと考えられる。

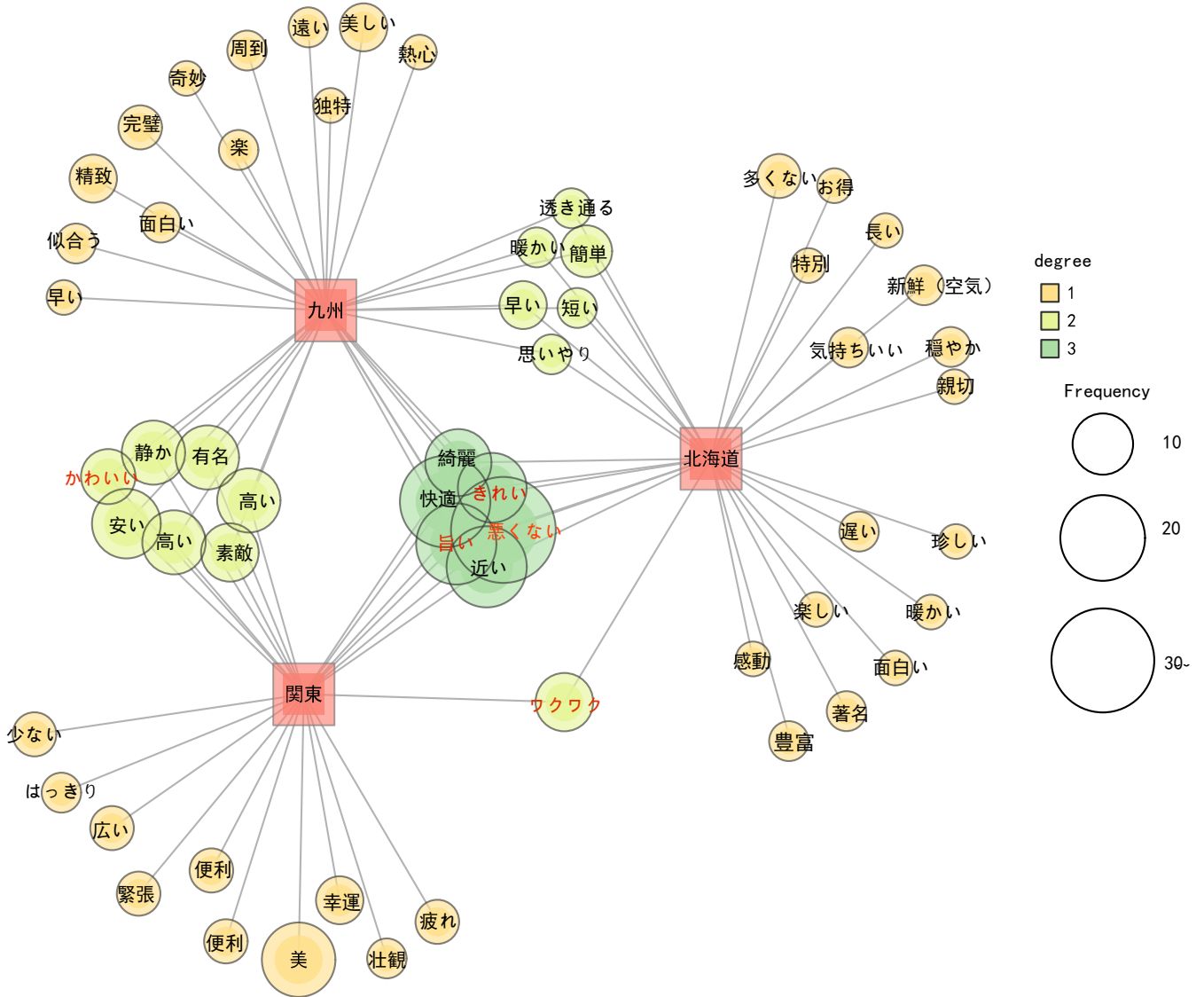


図 6-7 ポジティブ感情交通文 x 外部変数：形容詞・共起ネットワーク

まず、全体的に共起している形容詞には「旨い」があり、駅付近のグルメを指すものが多く見られる。「悪くない」は風景や駅弁、車内の環境に対する評価である。

次に関東と九州の共通する共起形容詞には「かわいい」がある、これは箱根登山鉄道、指宿玉手箱、京王バスに対してかわいいと感じたブロガーがいるからである。ここで、

写真がなかったため、どういうものに中国人ブロガーがかわいいと感じさせたことが見えないが、観光客に印象を与えるため、さらなる研究も必要だと思われる。

6-2-2. ホテル

表 6-15 ホテルネガティブ感情文

北海道	部屋の中にはバスルームがなく温泉ホテルなので、皆大浴場で風呂に入る、知らない人と一緒に風呂に入るのはどうも苦手で、家族貸し切りで入りました。	-0.1
	地震で部屋での茶器が床に落ちて、怖かった。	-0.4
関東	わざと湖が見える部屋に予約して、三泊しても離れたくない。	-0.3
	レストランを探して、道が真っ暗で少し怖い、蘆ノ湖まで歩いたが夜景を見る気分もなく、とにかく夕飯を食べたい。	-0.3
	ホテルではバスルームとトイレは分けているが、間取りが小さい。	-0.2
	いろんな温泉街に行っていたが、鬼怒川は寂れている方だ、いろんなホテルが閉まった。	-0.2
九州	スマホを持ち込んではいけない、ある子が温泉でスマホをいじったら、追い出されて、ホテル側もこのせいで通知を出した、	-0.8
	チェックインしたら、日本人のおばあちゃんが案内してくれて、荷物を代わりに持ってくれるが、ダメだと断った、老人に持たせるわけにはいかない。	-0.4
	シーズンじゃなかったら日本人はまだ多いものの、シーズンになったらホテル内でみんな中国語をしゃべっている、それだとおすすめできない。	-0.3
	日本人の老婆がまるで田舎者を見ているような目でこっちを見ている。	-0.3

ホテルに関しては、ネガティブな感情描写が以下に挙げられる、知らない人と一緒に風呂に入るのに抵抗感のある描写、日本で地震に遭遇した時の描写、ホテル付近でレストランを探すのが大変だなど、街全体に対してのネガティブ描写、温泉でのマナー注意、中国人観光客が多すぎて本物性が損ねる、日本人にたいする複雑な気持ちなどの描写もあった。特に気になるのは、中国人観光客が温泉ホテルでの高齢スタッフに荷物を持たせたくない気持ちがあることである。高齢化社会の日本では、就労年齢も高く、特に温泉地ではこういった現象が多くみられる。こうした高齢者への配慮の気持ちは、中国と日本との文化の違いによると考えられる。

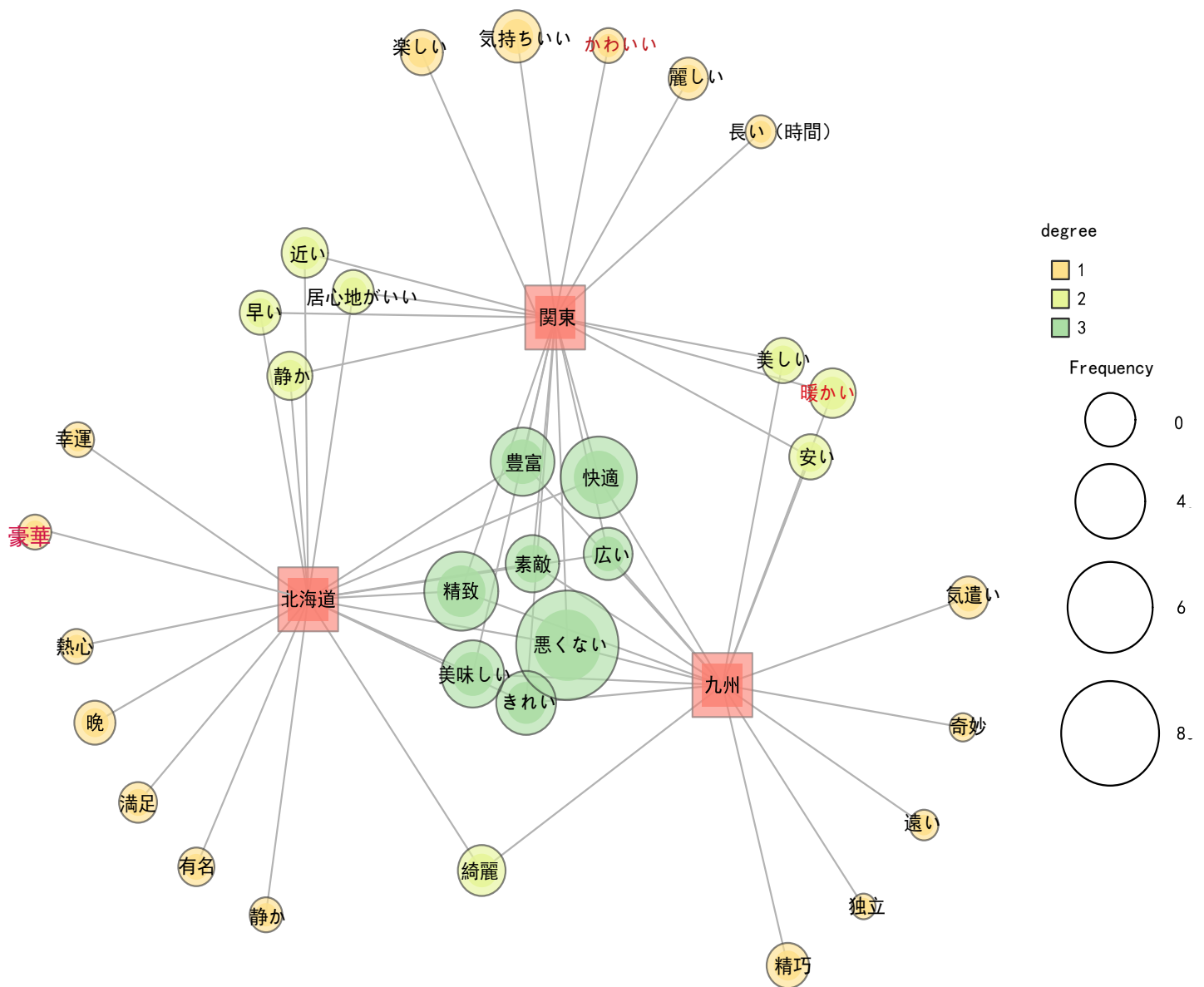


図 6-8 ポジティブ感情ホテル文 x 外部変数 : 形容詞・共起ネットワーク

まず、関東独自の特徴としては「かわいい」があげられ、ここは主に浴衣がかわいいことを指す。

北海道独自の特徴には「豪華」が入っている、ホテルが豪華だと感じた文が多かった。

九州と関東が共通した共起形容詞には「暖かい」がある。ここでの「暖かい」はサービスに暖かさを感じたことである。

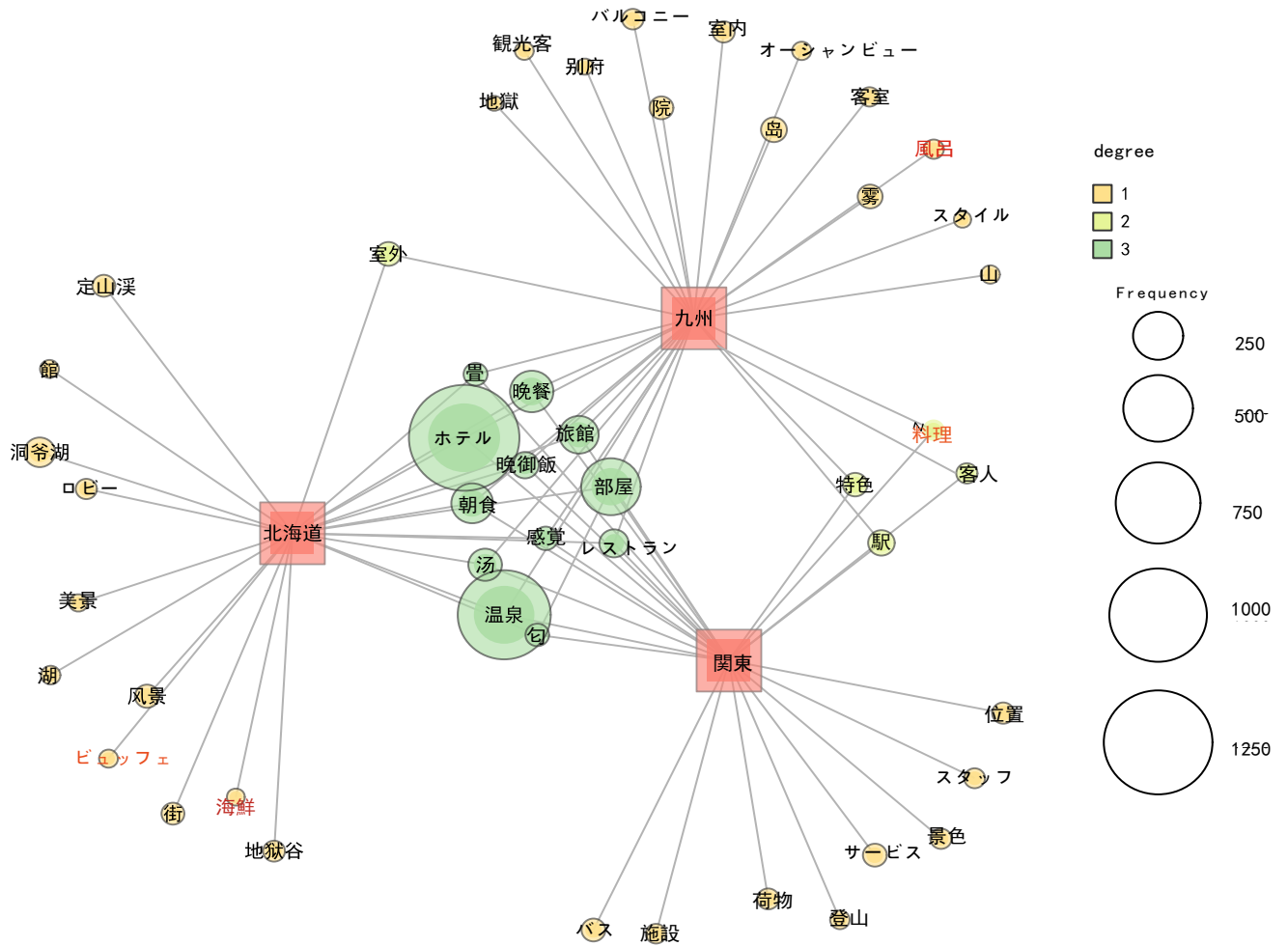


図 6-9 ポジティブ感情ホテル文 x 外部変数：名詞・共起ネットワーク

北海道ではスポット描写関連語が独自の特徴語がほとんどであるが、「ビュッフェ」と「海鮮」ではほかの二つエリアより関連性が強いことが見られる。

前節でも述べたように、「料理」という語の意味には懐石料理も含めているので、九州と関東の共通する共起語では料理も挙がっており、関東と九州は「料理」のイメージが強く、一方北海道は「ビュッフェ」のイメージが強かった。

九州ではスポット描写関連のほかにも「スタイル」と「風呂」がある。前節でも述べたようにここでの「風呂」中国語の原文でも日本語の「風呂」をそのまま使用している。つまり、九州では「温泉」を使っている記述もあるが、一方「風呂」のまま使用しているブロガーも多いことがわかる。

6-2-3. 日本

表 6-16 日本ネガティブ感情文

北海道	ガイドがずっと日本の温泉文化を強調していて、期待していたが、それほどでも...むしろ北京の温浴施設に劣るかも、施設も環境も似ているし	-0.5
関東	日本の温泉では服は着ちゃダメで、タオルもダメ	-0.3
	日本の交通費がかなり高く、バス 50 で 70 元をかかる	-0.4
	日本人はとても礼儀正しく、さよならを言う時も、ずっとお辞儀をする、自分もお辞儀するしかなかった。	-0.8
九州	別府に行くのは全くの気まぐれで、なぜなら福岡付近ではもう大体見たし、妻も日本に来たからには本物の温泉を見に行かないと残念だといった。	-0.4
	初めて日本に来たわけではなかったのなら、湯布院は来なくてもいい、ここはビジネス化しすぎて、特色を感じない。	-0.6

日本と関連するネガティブ感情文では、日本の温泉文化に関心を示せない記述もあり、日本の温泉へのマナーを気にしてネガティブに判定された文もある。また日本のサービスに不慣れを感じ、礼には礼を返さなければいけないと思っているのも見える。

九州では、湯布院のビジネス化が進みすぎて、好まない人もいるようだ。

6-2-4. 温泉

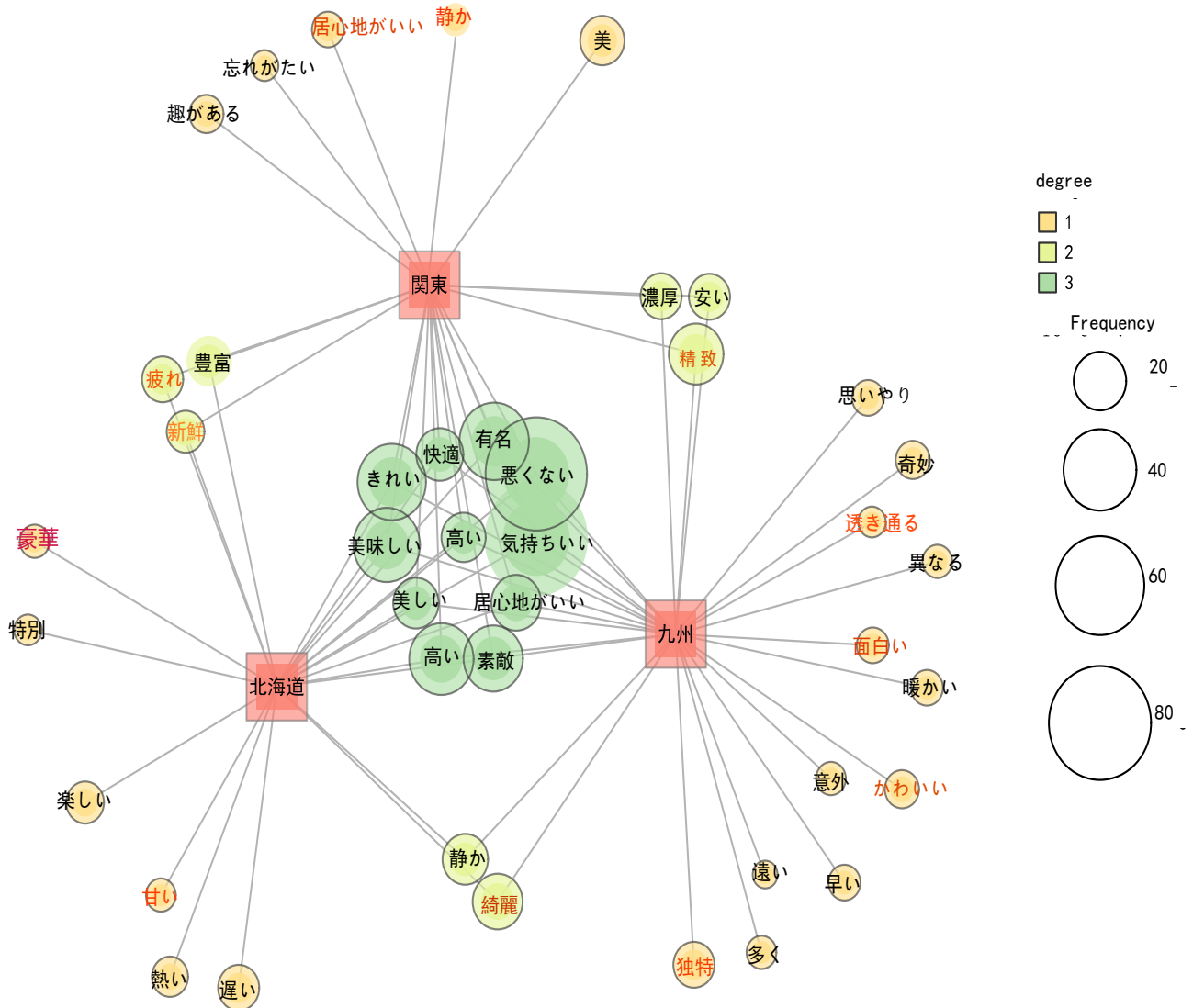


図 6-10 ポジティブ温泉文 x 外部変数：形容詞・共起ネットワーク

九州と関東が共通の共起語には「独特」があり、ここでの独特は温泉と景色を指すのである。また、「精巧（精緻）」があり、温泉に関連する語では、九州と関東が北海道より「精緻」についての描写があるとみられる。

関東と北海道が共通している共起語「疲れ」があり、確認すると、「疲れが癒された」との描写が多い。「新鮮」とは空気が新鮮を指すが多い。

北海道の独自の特徴語としては「甘い」「豪華」などがある。豪華はホテルでの分析で述べたように、主にホテルの内装を指すものである。「甘い」は甘エビとデザートなどを指すのである。

関東の独自の特徴語には「美」、「居心地が良い」、「静か」などがある。

九州の独自の特徴語は「かわいい」、「面白い」、「透き通る」などがある。中では「かわいい」は「駅がかわいい」、「ホテルの施設がかわいい」、「霧島のマスコットキャラがかわいい」、「お土産がかわいい」などがあげられる。

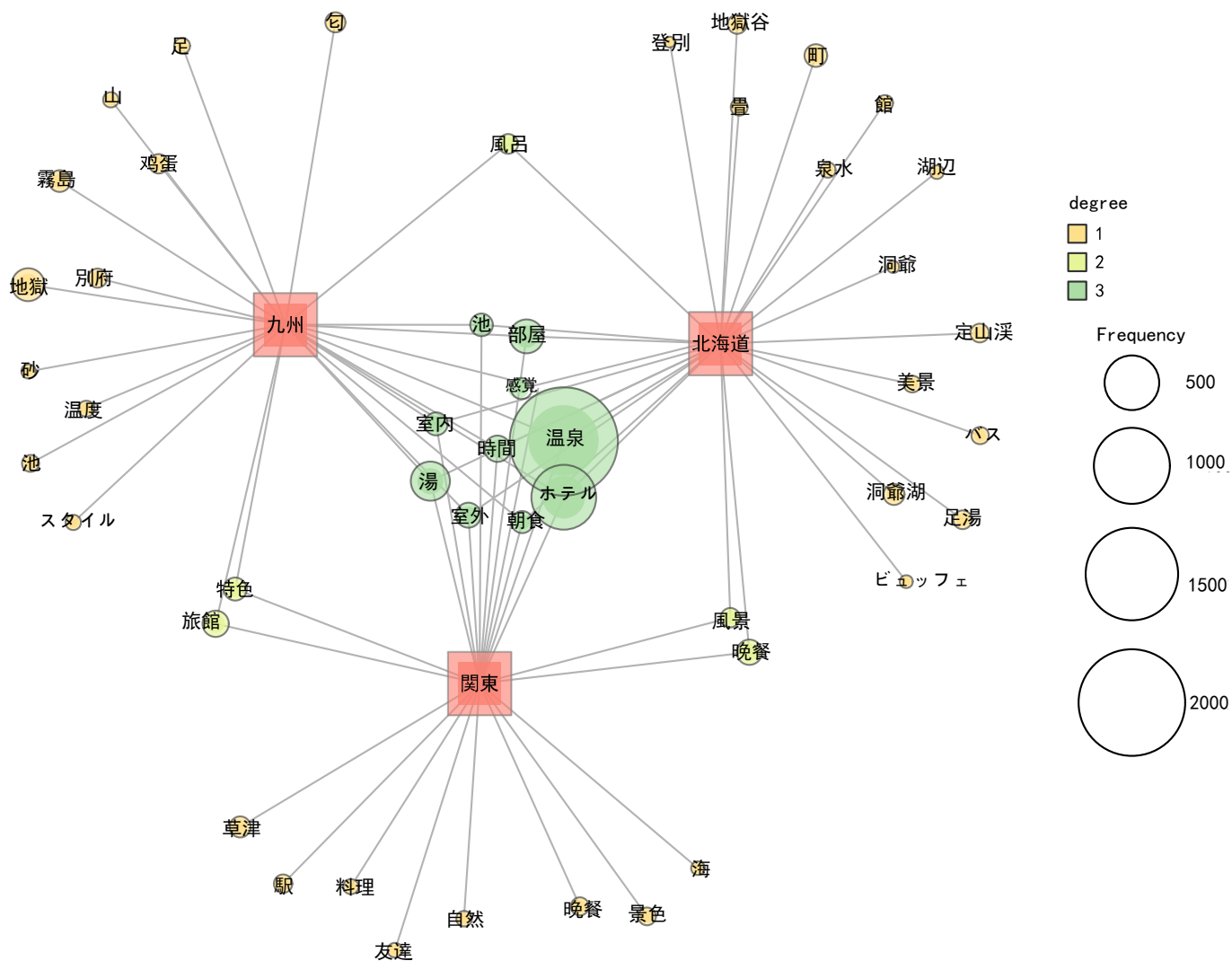


図 6-11 ポジティブ温泉文 x 外部変数：名詞・共起ネットワーク

名詞の方を見ると、九州と北海道では共通する共起語には「風呂」があり、「風呂」という特殊な言い方を使う傾向がある。

九州と関東が共通している部分には「匂」があり、硫黄の匂が気になるということである。また、関東と北海道エリアでは風景への関心が強いことが見られる。九州では「温度」との関連性がより強く、温泉の温度への関心が見えた。

6-2-5. 食事

表 6-17 食事ネガティブ感情文

北海道	懐石料理の種類が多いものの量が多すぎるという点もある，それに床に座って食べるのも少し食欲に影響するし，最後になっていっぱい食べ残した．価格の高いものを選んで食べるしかない．	-0.3
	ホテルの代金にはビュッフェの夕食と朝食が含まれている．本来なら夕食への期待は高かった，北海道三大カニ食べ放題がある．しかし，レストランが混んでいてカニも冷たいし，美味しくない．	-0.7
	ここの街はレストランがなく，観光客の夕食はホテルで済ます．	-0.4
	地図の通りレストランを探そうとしたが，営業しているレストランがどこにも見当たらない	-0.7
関東	朝食の卵は冷めているから食べてない．	-0.8
	畳に座って食べるのが辛いあつという間に疲れたポーズチェンジをしなければならない．	-0.3
	食べながら従業員に聞いてみた，この上には何が書いてあるか，彼女は小声だったので，あまり聞き取れなかった，料理長が書いた歌ではないかと思った．（品書きのことである）	-0.2
	今回はホテルで朝食と夕食を予約しなかった，失敗した．本来ならホテルの料理は美味しいと思えなかったが，外に出たら食べれるものほとんどなくて，コンビニのパンで済ますしかなかった．	-0.7
九州	湯布院の旅館で夕食を食べる，ネットではこの店での夕食の評判は高かったが期待が高すぎて普通だと思う，地鶏鍋予約していたが魚の刺身だけ印象に残った．	-0.2
	20分くらい経つと，机の上にはいろんな食べ物が置かれていた，中年の女性が日本語でどうやって食べるかを教えてくれる，色々言ったが，大体分からなかった．食べ終わってから電話するしか分からなかった．後でわかったのが，これらを全て食べ終わったらまだ料理が残って，ホテルの名物温泉で作ったプリンがあると．	-0.3
	どうも鬱なのは，湯布院のレストランはだいたい5時で閉まる，何度も壁にぶつかった後ホテルに戻り，なんとか晩御飯を提供してくれないかを聞くしかなかった．	-0.5

ポジティブ文の共起ネットワーク分析でとりわけ有意な結果は見当たらなかったが，ネガティブの記述から見て，ネガティブの評価には，冷たい食事に対する不慣れ，中国の食文化において暖かい食事を食べるのが一般的で，調理法と料理の保温などについてこだわりがある．次に，温泉地におけるレストラン探しが難しいことについてもよく言及される，中国の旅行文化にホテルでの食事は朝食のみを利用するのが一般的で，観光

地においてもレストランが林立しているので、ホテルではなく当地のレストランを探して利用することが当たり前だと思っている人も少なくない。従って、日本の温泉旅行の食事込みという文化を事前に観光客に把握してもらう必要がある。温泉旅行とは温泉に入るだけでなく、独特の食文化も温泉旅行の一環であることもきちんと紹介しなければならない。その上、外国人観光客がそうした温泉旅行の文化を受け入れられるかを調査してから、必要に応じて温泉地のレストランの増設や営業時間の調整などを視野に入れてもよい。そして目玉になる懐石料理については、スタッフによる料理の紹介も懐石料理体験の一環として不可欠だが、言語の壁があり、料理の説明を聞いてもよくわからないことは決して少なくない。筆者自身も初の懐石料理では説明を聞いても、料理を楽しむとき「なにを食べているかがわからないけどおいしい」と思っていた。しかし、料理の内容を理解した上で食べるのなら、こうしたあやふやな印象ではなく、より素敵な思い出になるはずだと考えられる。品書きを多言語に用意し、スタッフに品書きと料理を照らし合わせて説明を行う方が、外国人観光客により鮮明な印象を与えるはずである。

6-3. 精緻の概念

前節の分析において、中国人は「精緻」という語を使い日本と温泉を描写する。「精緻」は各日中辞書での意味は以下である。

白水社 中国語辞典では（庭園建築の布置や器物の製造などが）精緻（精巧）で細密である、精緻（精巧）で手が込んでいる様相を描写。精緻の対義語が粗劣である。また、では精緻の同義語には「繊巧さ」細工や技芸が繊細でたくみであることを指す。「美しい」容姿等が優雅で美しいことを指す。「巧緻」細部までうまくできていることを指す。日中中日専門用語辞典では洗練という意味合いを含むとある

*精緻

精緻 or 精美 or 精巧 or 優雅 or 雅 or 洗練 or 細やか or 繊細 or 精良 or 趣のある

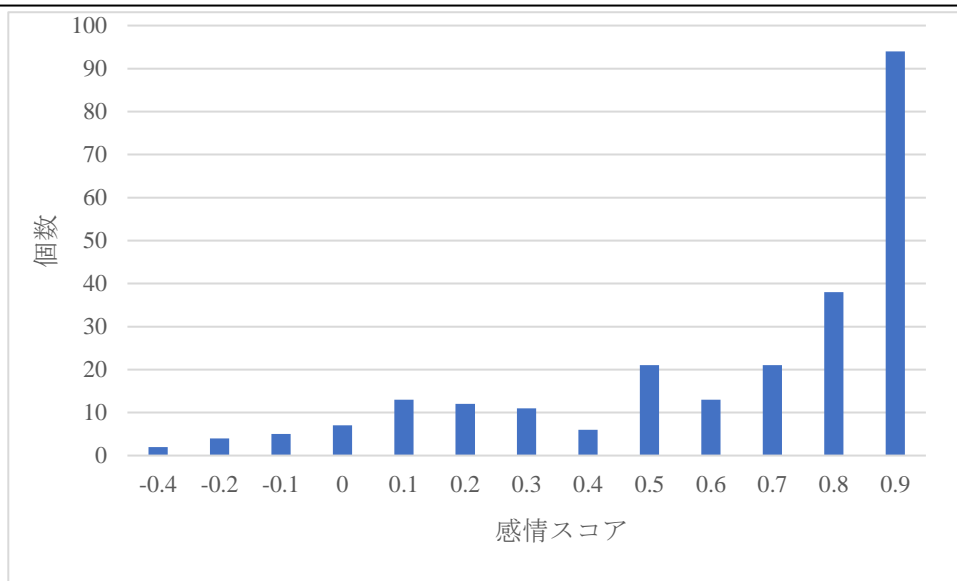


図 6-12 精緻コードの感情分布

図 6-12 からわかるように、精緻コードの 75%が 0.5 以上のポジティブ感情で、最高スコア 0.9 は 38%もあることから、精緻に関わる文書はほとんどがプラス的でかつブロガーの高評価を示している文書であることがわかる。

表 6-18 精緻コードの関連名詞

順位	関連名詞		
1	料理	11	品
2	和風	12	小さな
3	晚餐	13	デザイン
4	盛り付け	14	食べ物
5	庭園	15	飾り
6	ガラス	16	客室
7	部屋	17	晩御飯
8	懐石料理	18	美味
9	建物	19	美食
10	朝食	20	伝統

精緻コードの関連名詞には、まず料理が冒頭に立ち、料理が一番中国人に「精緻」のイメージを与えることがわかる。次に、和風と精緻の関連も強く、精緻は日本文化を描写するのに使う語であることもわかる。それから、庭園が代表する建物類に工芸品であるガラスまで幅広い範囲で使用されている。具体例をまとめると表 6-19 になる。

表 6-19 精緻の使用例

料理関連	盛り付け	<p>宴会料理のすべての食事が洗練されている，食材の選びから，配膳まで，盛り付けから，さらに食べる手順まで，非常にエレガントであり，日本の究極の美しさへの追求を反映しています。</p> <p>もともとはホテルで適当に和食で済ませたかったのだが，まさか本格的な懐石料理が食べられるとは，皿の盛り付けの一つ一つが精緻（精巧）であり，食感も驚きです。丸ごと2時間食べました。</p>
	器	<p>料理が異常なほど精緻（精巧）で，当地の旬の食材を使用し，一番目立たない食後のデザートまで，最も精緻（精巧）な器で盛り付ける。</p>
建築関連	庭園	<p>部屋の小さなバルコニーからはホテルの中庭が見えます。大きくはありませんが，丁寧に配置されています。花や木の層，小さな池や鯉がレイアウトに工夫を凝らしています。日本庭園の精緻（精巧）感があふれている</p>
	和室	<p>和室の飾りはエレガントで，シンプルで，こだわりがあって，畳も綺麗いで，すべてのアメニティまで，精緻（精巧）さが見える。</p>
商品関連	寄木細工	<p>ホテルロビー1階の陳列棚には，寄木細工のコースター，箱，収納など，趣があつて精緻（精巧）に見える。</p>
		<p>寄木細工の魅力には2つがあると思います。1つは，模様のデザインが非常に精緻（精巧）であるということです。単純な幾何学的な接合でありながら，至る所に職人が細心の注意を払っていることが見える。</p>
		<p>箱根で一番有名なのは寄木細工です。事前に店も調べました。いくつかの小物を買って，絶妙な仕上がりで，しかも値段も高くない</p>
	ガラス製品	<p>箱根ガラス博物館のガラス製品が精緻（精巧）である。</p>
		<p>伊豆高原は，精緻（精巧）なガラス工芸品，人形，手工芸品，西洋絵画，東洋絵画など，さまざまな美術館が立ち並ぶ芸術的な雰囲気あふれる街です。</p>
		<p>小樽は単調ではなく，精緻（精巧）なガラス店，おとぎ話のようなオルゴール博物館，デザートショップ，寿司レストランがあります。</p>

表 6-20 各温泉地ブログの精緻使用状況

	*精緻	ケース数
下呂	3 (25.00%)	12
伊豆	9 (37.50%)	24
別府	17 (20.73%)	82
城崎	3 (21.43%)	14
定山溪	1 (3.45%)	29
指宿	9 (21.95%)	41
有馬	8 (22.86%)	35
洞爺湖	5 (7.35%)	68
熱海	11 (39.29%)	28
由布院	11 (23.91%)	46
登別	17 (18.68%)	91
白濱	2 (16.67%)	12
箱根自由行	41 (20.71%)	198
箱根跟团	20 (21.28%)	94
草津	1 (14.29%)	7
霧島	3 (20.00%)	15
鬼怒川	5 (55.56%)	9
黒川温泉	6 (26.09%)	23
合計	172 (20.77%)	828
カイ2乗値	30.752*	

表 6-20 から見て、北海道は相対的に精緻（精巧）の割合が低いことがわかる。

北海道の温泉体験は中国人に精緻（精巧）というイメージを与えることが難しいことが推測できる。

さらに、前節でコードにより分割したテーマで、どのテーマが精緻コードとの類似度が高いか、つまり関連性が強いかを見るため、また温泉地記述と全体と比較し、どちらが精緻との関連性が強いのかも、温泉旅行の特長を判明するのに重要である。

各コードと精緻コードとの類似度行列（jaccard 係数）を算出した結果が表 6-21 である。温泉地記述の各コードはブログ全体に比べて関連度が高いことがわかる。各コードを比較すると、交通が精緻との関連が一番低く、飲食が精緻との関連が一番高いことがわかる。

表 6-21 各コードと精致コードとの類似度行列（jaccard 係数）

	温泉地	全体
*交通	0.006	0.005
*ホテル	0.028	0.017
*日本	0.031	0.02
*温泉	0.017	0.014
*飲食	0.04	0.023

6-4. 再訪意向

再訪意向については、再訪に関わる文書量がそれほど多くない上で、それぞれ記述が個性的であり、より定性的に分析する必要があるので、手作業で再訪に関しての記述を確認した。

再訪意向は大きく分けて三種類ある。

一つ目は、今回の旅行ではやり残したことや見られなかったことがあったので、次の旅行で補完したい。主に箱根を目的地としたブログで見られる。

二つ目は、今回の旅行では満足したので、また同じ体験をしたい。主に別府や由布院などの九州を目的地としたブログで見られた。

三つ目は、今回の旅行体験をあまり言及せず、次の旅行で違った体験をしたい、例えば、違った季節で来たい、団体ツアーではなく個人で来たいという再訪に変化を求める記述もあった。

表 6-22 再訪記述の具体例

箱根	一日だけでは全部のアトラクションを楽しめなかったのでまた来たい
	日帰りだったので、次回は宿泊したい
	天気が悪いので、富士山が見えない、次回はしっかり見たい
	大涌谷のケーブルカーが運営停止、次回に乗りたい
別府 由布院	めちゃくちゃ満足したのでまた来たい
	お金を稼いだら、次回は三連泊にしたい
	次回また絶対に来る
	冬／夏になったらまた見てみたい
北海道	夏になったらラベンダーを見たい
	ツアーだったので、次回は個人できたい
	また同じホテルに泊まりたい

表で見た通り、箱根については、「今回は残念だったので、今度に補いたい」のよ

うな「見残し・やり残し」現象が多く存在する，日本を象徴する世界遺産富士山が綺麗に見えるかどうかは中国人観光客の多くが注目している話題である．また，箱根近辺では観光のアクティブが多いことから，やり残しが生じやすいのもひとつである．

別府・由布院に関しては，満足度が高いので，再訪したいという意向が見える．

北海道では，シーズンによって観光活動が大きく変わることもあり，再訪による新しい体験も期待できる．

「見残し・やり残し」による再訪行動についての先行研究には，赤沢・佐藤（2010）が，観光時の見残し・やり残しの存在，経験による訪問者選好の更新による新たな学びの機会への欲求、景観、移動手段、駐車場、地域の印象を満足度経路要因として示し，その影響を実証している．つまり，見残し，やり残しは観光客の満足度を下げるが，再訪意向に促すような現象でもある．

第7章 考察

7-1. 温泉地全般について考察

日本関連記述において，まず，ネガティブ感情では価格に関するものが主である．サービスや部屋などのマイナス評価もあったが数は少ない．その他「国家」「日本人」などの語から見て，日本社会に対する評価もネガティブ感情にはつながる．ほかにも交通ビザといった一般旅行に対しての評価がある．特に，中国人ブロガーがマナー違反を気にしていることがネガティブ感情に判定されたことも興味深い．

ポジティブ感情では，温泉が日本軸のポジティブ特徴語の上位にあることから，中国人ブロガーは日本に関する物事の中で温泉を高く評価していることが推測できる．従って，温泉は日本のプラス評価につながる．その他には食事や着物，歴史などの文化的事象も高評価文に出現している．つまり，中国人ブロガーが日本に対するプラス的に考えることは文化的な事柄に集中している．

以上のことから，温泉は観光客の日本のイメージ上昇につながる事がわかる．

7-2. 精緻の考察

7-1-3 の分析から見て「精緻」は日本との関連性が高く，中国人観光客にポジティブな感情をもたらす．また，7-1 の分析から，交通関連の観光体験を向上させ，より全般的に訪日温泉旅行を楽しませることに繋がる事がわかった．従って「精緻」という要素を交通体験に取り込むことで，中国人観光客に交通機関利用時の体験を向上させるのではないかと考えられる．

7-3 で分析した「精緻」との関連語に食事についての語が大多数占めていて，さらに，交通コードでの頻出語「旨い」から見て，訪日ブログでは駅弁について高く評価している事がわかる．もともと中国本土には駅弁の文化がなく，日本の新幹線に相当する高速鉄道においても，食事は持参するインスタントフードか，車内で販売する機内食のような加熱した弁当が一般的である．日本の駅弁と違い，おかずの種類が少なく，見た目

に対してそこまでこだわっていない。また、中国の食事は暖かい食べ物が一般的であるため、加熱不要な日本の駅弁は中国人観光客にとって新鮮である。そこで、駅弁に対して「旨い」だけでなく、「精緻」というイメージを中国人観光客に残せることが出来るとしたら、交通面の観光体験向上につながると筆者は考える。ここで、7-3 の分析結果を受け、いくつかの提案を行う。例えば、駅弁の中身を懐石料理風にし、オンラインとオフライン両方でプロモーションをして、駅弁＝精緻のイメージを定着させる。また、駅弁の重箱で地方の独自の要素を取り込み、地方の個性をアピールすることもできる。例えば、箱根では寄木細工要素が入った駅弁を販売することで、地方のブランドイメージ作りができるだけでなく、乗り物での体験も一層向上させることができる。

第8章 結論

本研究では、旅行ブログにおける特定の観光体験に着目し、コーディングで複合的旅行体験のテキストの特定テーマを抽出、感情分析と併用して、特定の観光体験における異なる側面でのユーザーの評価点抽出に成功した。

観光領域におけるテキストマイニングの従来の研究対象であるオンラインレビュー・クチコミと違い、本研究はより詳細の記述を持つブログを対象に分析を行った。今後のテキストマイニングにおいて、クチコミとレビューのような短い文書に限るのではなく、ブログや記事のような長い文書への分析に本研究は研究方法を提示した。本研究はこうした学術的意味を持つ。

中国人観光客を対象とした先行研究も多くあり、中国人観光客は従来のショッピング重視の観光から、体験重視の観光に変わり、そして、体験型観光で日本の一大名物とする温泉旅行が中国人観光客の視点からの分析ができたのも本研究の実学的意味でもある。

温泉地記述と旅行全体の記述の比較から、温泉地に訪れたブロガーは温泉地での体験は旅行全体の体験に比べ、よりポジティブな感情を示した。つまり、訪日旅行において温泉地での滞在経験が日本旅行にプラス的な評価をもたらす。また、日本旅行のイメージ向上にもつながる。

感情分析とコーディングにより抽出した評価点から、中国人観光客は「悪くない」、「きれい」など一般的な評価語のほかに、ある特定の評価「精緻」で温泉、食事、ホテルを描写する。本研究では「精緻」の使用場面を究明するため、「精緻」のコーディングルールを作成し、料理関連、建築関連、工芸品関連、三つのカテゴリーにまとめた。

温泉地でのコード間の比較により、交通に関連する記述にポジティブな描写が少なく、感情を含まない記録的な記述が多いことから、ブログから見た中国人観光客の交通体験は相対的に印象が残らなかったことがわかった。体験を向上させるため、「精緻」の要素を取り込みやすい駅弁に力を入れることを提言した。

最後に、コーディングによる再訪意向の影響要素を確かめ、「見残し・やり残し」による再訪、今回満足したので次も体験したい、異なる時期にまた来たい、主にこの三つ

の要素により再訪意欲が生じることが分かった。

結論としては、まず、訪日旅行において温泉地での滞在経験が日本旅行にプラス的な評価をもたらし、日本旅行のイメージ向上につながるということがわかった。次に、中国人観光客は温泉旅行、特に懐石料理を始めた食事関連の面では「精緻」という評価語を使い、高評価を与える。将来的に、「精緻」のさらなる使用場面を研究し、中国人観光客が持つ「精緻」のイメージが如何なるものかをより明確にする必要がある。そのうえで、「精緻」の要素を取り込んだ観光客向けの商品開発も提案できる。

8-1. 制約と展望

8-1-1. 感情分析のデメリット

まず感情分析のデメリットについて、本研究において感情分析はポジティブかネガティブかを判断するツールとして使用しているがどれほどポジティブかどれほどネガティブかは、今回の Google cloud platform では分析しにくいことがわかった。また、ポジティブであるにもかかわらず、ネガティブに判定されてしまった例もいくつか見られた。例えば、時間が過ぎるのが早すぎて惜しい、などはポジティブに判定されるべき例である。そして、マナー違反したらだめ、のような他人に注意を促す例が存在するため本研究のように機械を通した結果だけでなく手作業で内容を確認し内容分析を行う必要がある。さらに感情分析の結果は受動的であり、ユーザー自らの採点ではないため今後は両者の違いをより詳しく分析するためにも同じ文章を対象に感情スコアと採点の結果を比較して分析する必要がある。

8-1-2. データの偏り

今回の分析に使ったのは中国で一番市場シェアを持つ ctrip で、データ量を狙っていたが、他にも「馬蜂窝 (mafengwo)」のようなソーシャルコミュニケーションを重視する大手サイトがある。旅行ブログに特化した一面を持つので、より質の高いブログを収集できる。異なるプラットフォームでのユーザーデータを分析し比較することで、ユーザークラスターの特徴や好みを探り、将来的により効率のある広告宣伝が期待できる。

本研究の位置付けとしては、中国人から見た日本温泉という大きな目線から研究を進めたが、人気の温泉地があり、マイノリティで特徴のある温泉地もあるため、分析上、データに偏りが生じ、どうしても、ブログの数の多い温泉地に偏ってしまう。従って、マイノリティの温泉地を訪れたブログにはより細分化した分析が求められる。

さらなる観光客の特徴を探る研究は、より良い観光サービス提供につながれると信じる。

引用・参考文献

BIGLOBE 旅行第 13 回温泉大賞日本温泉番付. 参照日: 2020 年 12 月 23 日, 参照先:

<https://travel.biglobe.ne.jp/onsen/award/>

Danowski, J. A., 1993, "Network Analysis of Message Content," W. D. Richards Jr. & G. A. Barnett eds., *Progress in Communication Sciences IV*, Norwood, NJ: Ablex, 197-221

Osgood, C. E., 1959, "The Representational Model and Relevant Research Methods," I. de S. Pool ed., *Trends in Content Analysis*, Urbana, IL: University of Illinois Press, 33-88

P. Chang, M. Galley, and C. Manning. "Optimizing Chinese Word Segmentation for Machine Translation Performance": *Statmt '08 Proceedings Of The Third Workshop On Statistical Machine*, pp. 224-232, 2008.

鄔雅瓊 (2016) 訪日中国人観光客の潮流-広告コミュニケーションの分析

越智香那恵・桂信太郎・井形元彦 (2018) 「テキストマイニングによる観光地の定量的評価分析」『産学連携学会 関西・中四国支部 第 10 回研究・事例発表会&記念講演会予稿集』

大久保立樹・室町泰徳 (2014) 「旅行ガイドブックとロコミの言語解析による訪日外国人の観光地イメージに関する研究」『都市計画論文集』 Vol. 49, No. 3, pp. 573-578.

観光庁: 訪日外国人消費動向 参照日: 2021 年 9 月 12 日, 参照先:

https://www.mlit.go.jp/kankocho/topics02_000182.html,

観光庁: 訪日外国人旅行者(観光・レジャー目的)の訪日回数と消費動向の関係について, 参照日: 2021 年 6 月 12 日 参照先:

<https://www.mlit.go.jp/kankocho/siryoutoukei/content/001350782.pdf>,

牛曉雪・中村哲(2019): 中国人若者の訪日旅行前の情報探索の実態, 第 34 回日本観光学会全国大会学術論文集, pp393-396.

Google cloud Natural Language AI 参照日: 2021 年 6 月 5 日, 参照先 <https://cloud.google.com/natural-language>

呉歌・林高樹 (2019) オンライン・トラベル・エージェント上のデータ分析による訪日中国人の観光地イメージに関する研究 修士学位論文 2019 年度経営学第 3559 号

齋藤朗宏 (2012) 日本におけるテキストマイニングの応用, *The Society for Economic Studies The University of Kitakyushu Working Paper Series No. 2011-12.*

CTRIP ブログ検索ページ参照日: 2021 年 4 月 2 日, 参照先:

<https://you.ctrip.com/searchsite/Travels?query=%25e7%25ae%25b1%25e6%25a0%25b9>

スクレイピングツール Data Miner 参照日: 2020 年 11 月 12 日, 参照先:

<https://dataminer.io/>

鈴木 誠・田崎知裕・進士五十八 (1989) 外国人の日本庭園観に関する比較研究. *造園雑誌* 52 (5), 25~30

- セナ=イアネス=ラピド・下村 彰男・小野 良平 (2010) フィリピン人における日本庭園の認識及び印象評価. ランドスケープ研究 73 (5), 477~452
- 第一財經 (2021) 冬の過ごし方 参照日: 2021年3月12日 参照先: <https://www.yicai.com/news/100932094.html>
- 中国極楽湯公式ホームページ参照日: 2021年3月12日参照先: <https://www.gokurakuyu.cn/>
- 張章・山崎 隆之 中国人の訪日旅行記にみる観光行動の特徴: 高松市、広島市、熊本市を目的地とする旅行記の比較を通して日本観光研究学会全国大会学術論文集 = Proceedings of JITR annual conference 36, 139-144, 2021-12
- 西平直史 2019 「テキストマイニングによる観光クチコミ情報の分析 —山形市を対象として—」 『山形大学人文社会科学部研究年報』 16: 81-86
- 日本政府観光局: 国籍/月別訪日外客数(2003~2021), 参照日: 2021年8月22日, 参照先: https://www.jnto.go.jp/jpn/statistics/since2003_visitor_arrivals.pdf,
- 日本政府観光局: 訪日旅行データハンドブック 2020 参照日: 2021年3月20日, 参照先: <https://www.jnto.go.jp/jpn/statistics/datahandbook.html>
- 日本政府観光局: 訪日旅行データハンドブック 2021 参照日: 2022年1月4日, 参照先: <https://www.jnto.go.jp/jpn/statistics/datahandbook.html>
- 野中 尋史・平岡 透 訪日中国人観光客のオンラインホテルレビューの感情分析と 評価点の関係性分析 産業応用工学会論文誌, Vol. 6, No. 2, pp. 95-99 (Sept. 2018) The Japanese Journal of the Institute of Industrial Applications Engineers Online edition: ISSN 2187-5146, Print edition: ISSN 2189-373X, DOI: 10.12792/jjiaae.6.2.95
- 樋口耕一 (2014). 社会調査のための計量テキスト分析—内容分析の継承と発展を目指して— ナカニシヤ出版
- 樋口耕一 (2017). 社会調査のための計量テキスト分析—内容分析の継承と発展を目指して— ナカニシヤ出版
- 平井 貴幸 (2021) インバウンド観光拡大と経済成長との関連性: 構造変化を考慮した時系列分析 札幌大学研究紀要, 2021-10, 7-19
- 马蜂窝: 2019 中国出境自由行大数据报告, 参照日: 2021年8月23日, 参照先: https://www.sohu.com/a/332474882_665157
- 安原 有紗・劉 銘 (2021) ロコミの言語解析からみた日本人と外国人の浜離宮恩賜庭園に対する観光体験の相違 2021年度 環境情報科学研究発表大会 要旨集 p. 227-232
- 脇本忍・姜思義・大西隆士(2018), 沖縄におけるインバウンド市場調査—中国人観光客の消費行動と SNS の関係性—, 聖泉論叢, 26, 1-12.

謝辞

本研究に取り組み、学位論文をまとめるまでには、多くの方々のご支援とご指導を賜りました。修士論文を作成するにあたり、お世話になった皆様方に、この場をお借りして感謝の意を申し上げます。

まず、本研究を遂行し学位論文をまとめるにあたり、研究テーマの方向性や論文執筆に関するご助言・ご指導など多大なご助力をいただきました、指導教員である倉田陽平准教授に深く感謝しております。また、倉田准教授には、修士論文の完成まで、常に親心を感じる丁寧なご指導を賜り、毎週のゼミを通して修論の段取りが確認でき、論文執筆に前向きに取り組むことができるよう心強く支えてくださり、非常に感謝しております。

また、本研究の副査を務めていただき、多大なる有益なご指摘をくださった日原勝也准教授、岡村 祐准教授に深く御礼申し上げます。続きまして、TAとして学部の方々と共に課題を完成し、その際に多数の貴重なご助言・ご指摘をくださるとともに、研究のスタンスや研究者としての姿勢などをご教示いただきました小笠原悠助教に深く御礼申し上げます。そして、一年目で指導教員に当たる直井岳人准教授と直井研究室の先輩、後輩の皆様、河田浩昭さん、岡野雄気さん、羅楽さん、黒須康太さんはゼミやその他の時間を通じて、貴重なご助言・ご指摘をくださったことに深く御礼申し上げます。

次に、二年目でお世話になった倉田研究室の方々、Googleクラウドについて助言していただいた吉田伊武貴さん、修論作成にネイティブチェックをしてくださり、共に二年の大学院を送ってくださった元直井研究室で現倉田研究室の鈴木美樹さんに深く御礼申し上げます。

また、観光科学域の教員及び職員の皆様、同学年の皆様へ多大な感謝の意を表させていただきます。

最後になりましたが、筆者の修士課程への進学を応援してくれ、これまで温かい目で見守ってくれた上海の家族と友人たちには、特に心からの感謝を伝えたいと思います。多くの皆様に支えられて修士論文を終えられたことに、改めて深い感謝を覚えずにはいられません。本当にありがとうございました。

2022年2月10日 朱逸雯