

# 自由記述データに着目した限界自治体における生活質評価に関する分析 - 群馬県南牧村を対象として -

## AN EVALUATION ON THE QUALITY OF LIFE IN MARGINAL LOCAL COMMUNITY CONSIDERING FREE ANSWER

- A case study on Nanmoku-mura, Gunma -

塚田伸也\*, 森田哲夫\*\*, 西尾敏和\*\*\*, 湯沢 昭\*\*\*\*

*Shinya TSUKADA, Tetsuo MORITA, Toshikazu NISHIO  
and Akira YUZAWA*

Currently, the decrease in population and the aging have become important problem in Japan. The purpose of this study is analyzing quality of life in the marginal local community, considering free answer. The data of the questionnaire survey in Nanmoku-Mura Gunma was used for all analysis. Nanmoku-Mura is one of the typical marginal local communities. This study aims to discuss the characteristics and analysis method of free answers in questionnaire survey in comparison with pre-code data and free answers, by one of the language processing techniques for text data acquired from free answers. We clarified that the quality of life and residential attitude are composed of individual attributes and district characteristics.

**Keywords :** *Free answer, Marginal local community, Quality of life, Text mining*

自由記述, 限界自治体, 生活質, テキストマイニング

### 1. はじめに

過疎化, 高齢化が急速に進行している山間地域では, 近年, 高齢化率 50%を超える限界自治体<sup>1)</sup>が出現している。現在は少数の自治体ではあるが, 近い将来において増加が見込まれる。これら自治体内の集落の多くは限界集落であり, 日常生活や社会基盤の維持が問題である。このような地域の実態把握や地域計画への基礎調査として, 住民を対象としたアンケート調査が多く用いられる。アンケート調査は, 選択肢を予め設定した調査票から得られるデータ(以下, プリコードデータと称す)と自由記述データがある。住民を対象としたアンケート調査の調査票には, 自由記述欄が設けられている場合が多いが, 定量的に十分な分析がなされているとはいえない。

一方, 最近では自然言語処理分野で研究が進められているテキストマイニングにより自由記述アンケートなどを定量的に分析できるようになってきた。このため, 地域計画分野においても, 自然言語処理分野の成果を用いた研究がみられるようになってきている。本研究では, 対象地域を山間部の限界自治体とした住民アンケート調査の自由記述データに着目する。この理由は, 都市部と異なり, 山間部の魅力の発掘と合わせて, 深刻で切実な生活状況といった住民の生活質の意識が自由記述データの内容に顕れ, 生活質評価の構造をより具体的に検討できるのではないかと考えたためである。本研究で

は, テキストマイニングを用い, 限界自治体における生活質評価に関するアンケート調査票の自由記述欄に記入されたテキストデータ(以下, 自由記述データと称す)の特性を把握する。また, 自由記述データの有無及び内容の記載がプリコードデータで構築した生活質評価のモデルに, どのような影響を与えるのか基礎的な知見を得ることを目的とする。

### 2. 本研究の位置づけと研究の流れ

#### (1) 既存研究と本研究の位置づけ

本研究が着目する, テキストマイニングを適用した研究, 山間自治体・限界自治体の生活質評価・居住意向を分析した研究の2点から既存研究を整理した上で本研究の位置づけを示す。

都市及び地域計画分野におけるテキストマイニングを適用した研究は, 意見の分類・分析に自然言語処理技術を用いた研究がみられるようになった<sup>2), 3)</sup>。佐々木ら<sup>4), 5)</sup>によるワークショップや塚田ら<sup>6)</sup>による委員会の討議内容の分析, 佐々木ら<sup>7)</sup>によるブログマイニングによる行動データの抽出とアンケート調査との比較がある。長ら<sup>8)</sup>は, 自由回答のインタビュー調査データを用い, 対象者の経験知識を抽出する方法を検討している。さまざまな場面で得られたデータを, マイニング手法により定量的に分析する研究が進められてい

\* 前橋市建設部公園緑地課 係長・博士(工学)

\*\* 東北工業大学工学部都市マネジメント学科  
教授・博士(工学)

\*\*\* 前橋工科大学大学院環境・生命工学専攻  
大学院生・修士(工学)

\*\*\*\* 前橋工科大学工学部社会環境工学科 教授・工博

Construction Department, Maebashi City Office, Dr. Eng.  
Prof., Tohoku Institute of Technology, Dr. Eng.

Graduate Student, Maebashi Institute of Technology, M. Eng.

Prof., Maebashi Institute of Technology, Dr. Eng.

る。森田・入澤<sup>9)</sup>は、群馬県前橋市の居住者に対するアンケート調査データを用い、自由記述データとプリコードデータの関係をコーレスポネンダ分析により把握している。プリコードデータと自由記述欄の記載事項を比較した研究としては、永野らの路面電車の利用意識を扱った研究<sup>10)</sup>があるが、山間自治体や限界自治体を扱った研究はみられない。山間自治体や限界自治体の生活質評価・居住意向を分析した研究は、交通サービス水準や交通特性に着目したものが多い。例えば、高齢化社会におけるモビリティに着目した溝上らの研究<sup>11)</sup>、過疎地域の公共交通サービス水準に関する森山<sup>12)</sup>の研究、交通サービス水準と生活の質の関連に関する宮崎<sup>13)</sup>の研究がみられる。谷本<sup>14)</sup>が、定住意向に影響を及ぼしている社会生活環境の要素を分析している。

限界集落を対象とした「中山間地域のモビリティの確保と地域再編戦略に関する研究」<sup>15)</sup>が進められ、住民の居住・移住意向には地域への愛着が大きく影響し、移動利便性が低いことにより直ちに移住意向につながらないことを明らかにした。森田<sup>16)</sup>は、山間地域を対象に居住意向を分析し、生活に不便さを感じていながらも定住意向が高く、高齢者ほどこの傾向が高いことを明らかにしている。塚井<sup>17)</sup>は、中山間地域の住民の移住意向と移住要件との関係を分析し、移住に係る費用等の条件が整えば移住意向を示す世帯が存在することを明らかにした。また、森田<sup>18)</sup>は、限界自治体である群馬県南牧村を対象に生活質評価と居住意向の関係を分析しているが、統計的にはさらなる分析が必要であるとしている。

過疎・高齢地域においては、人口減少と高齢化は免れ得ない状況である。この問題に対し定住促進や人口流出対策を施さないと自治体自体の存続が危ぶまれる。このため、本研究では森田<sup>18)</sup>らの既存研究における限界自治体の生活質の評価に関する研究をさらに深めることを着眼点とし、プリコードデータの活用と合せて自由記述データとの比較を試みるところを特徴とする。

## (2) 本研究の流れ

本研究は、次の流れに沿って進める。3章では、限界自治体である対象地域を設定し、住民の生活質の評価・居住意向に関するプリコードデータ、自由記述データを収集・整理する。4章では、得られた自由記述データについて、テキストマイニングを用いた分析を行うことにより、プリコードデータの評価項目で捉えられない記述データの特徴を把握する。また、出現語間の関係からより具体的な特徴を把握する。5章では、自由記述の有無の属性に着目し、プリコードデータを活用した生活質の評価との関係について構造的に把握する。6章では、前章までの分析結果を踏まえ、本研究で得られた結果をまとめ、自由記述データに着目することで得られた生活質評価に関連する情報について整理する。

本研究では、自然言語処理の要素技術を用いたテキストマイニングのためのフリーウェアKH Coder<sup>19), 20)</sup>を利用する。KH Coderは、操作性が高く、多くの分野の研究論文で利用されている。形態素解析器として「茶筌」<sup>21)</sup>を用いており、精度の高い単語抽出を行うことができる。自由記述有無の属性に着目し、プリコードデータ活用した生活質評価モデルの分析においては、山下<sup>19)</sup>と同様に共分散構造分析 (SPSS Amos17) を用いる。

## 3. 対象地域の設定と収集データの概要

表1 全国の限界自治体 (2010年)

	自治体名	高齢者数	高齢化率 %
1	南牧村 (群馬県)	1,387	57.2
2	金山町 (福島県)	1,356	55.1
3	天龍村 (長野県)	896	54.1
4	大豊町 (高知県)	2,549	54.0
5	昭和村 (福島県)	798	53.2
6	上勝町 (徳島県)	935	52.4
7	神流町 (群馬県)	1,231	52.3
8	大鹿村 (長野県)	598	51.6
9	川上村 (奈良県)	833	50.7
10	北山村 (和歌山県)	245	50.4
11	仁淀川町 (高知県)	3,267	50.3

資料：2010年国勢調査



図1 群馬県南牧村の位置

### (1) 対象地域の設定

限界自治体、限界集落の概念を提唱した大野<sup>1)</sup>は、限界自治体を「65歳以上の高齢者が自治体総人口の半数を超え、税収入の減少と老人福祉・高齢者医療関連の支出増という状況の中で財政維持が困難な状態におかれている自治体」とし、全国の山間地域で限界自治体が発生すること、集落においては自治機能が低下し、高齢者の生価モデルを作成し、自由記述データの記入属性による評価間の影響活状態の悪化し、独居世帯が残されると指摘している。

2005年国勢調査によると限界自治体 (高齢化率50%以上) は全国に6町村存在した。5年後の2010年国勢調査では、限界自治体は11町村となり (表1)、大野が予見したように、限界自治体は着実に増加している。最も高齢化率が高いのは、群馬県南牧村であり、2005年の53.4%から2010年の57.2%に上昇した。0～14才人口は4.3%、15～64才は38.5%である。人口推移をみると、1955年に10,573人であった人口が、2010年に2,423人 (国勢調査) と減少している。南牧村は、急傾斜地が多く災害危険性の高く、2007年9月の台風9号、2014年2月の大雪では、道路の寸断、孤立集落の発生など大きな被害を経験した。本研究では南牧村を対象とする。その理由は、1) 高齢化率が最も高い自治体であるが、将来的に限界自治体の増加が見込まれることか

表2 アンケート調査の概要

調査日	配布：2010年11月1日 回収：2010年11月21日（郵送投函期限）
対象地域	群馬県甘楽郡南牧村全域
対象者	全1,117戸の世帯主あるいは代表者（1戸に複数の世帯が居住している場合があるため、住民基本台帳の世帯数とは異なる）
調査方法	配布：分区長による戸別配布（60分区、本研究では「集落」と称す） 回収：郵送回収
調査内容	1)世帯属性（世帯主年齢、世帯人数、居住形態、自動車保有台数、居住年数） 2)災害による被害経験（2007年9月台風9号による被害、それ以前の被害） 3)生活質評価（23項目、総合評価） 4)居住意向（定住/転居意向、転居意向理由） 5)自由記述欄 設問文「南牧村のイメージを、短い言葉や文で、自由に書いてください。」
回収数	配布数：1,117票 回収数：637票 有効回収数：631票(有効回収率：56.5%)
調査主体	群馬工業高等専門学校 環境都市工学科 群馬県 県土整備部 都市計画課

ら、速やかに住民の意向を把握することは将来における限界自治体の政策検討のために重要な資料となると考えたためである。また、2)今後、社会基盤整備・維持のための予算は減少すると予想され、災害危険性の高い自治体を対象とすることは、今後の社会基盤整備・維持に関する政策を検討する上で有用な知見を得られると考えたためである。南牧村位置を図1に示した。南牧村は群馬県南西部の県境、高崎市から約40kmに位置する。集落別にみると、役場付近の中心部は高齢化率が低い、中心部から離れた地区には高齢化率の高い集落が分布する。2007年台風9号による河川の決壊地点の奥部にも集落があり、それら集落の高齢化率は高い。また、2014年2月の群馬県の各地へ甚大な被害を及ぼした集中豪雪においては、孤立した集落の救助・支援が問題となった。

#### (2) 収集データの概要

南牧村の全世帯を対象として分区長による戸別配布、郵送回収によりアンケート調査を実施した。南牧村には60の分区(本研究では、以下「集落」と称す)が存在し、全ての集落・世帯に配布・回収した。調査内容は、世帯属性、災害による被害経験、生活質評価、居住意向、自由記述である(表2)。

全体から世帯属性を捉えると、世帯主年齢は60歳代以上が80.8%を占めた。世帯構成は、独居世帯が24%、2人世帯が46%であった。また、92%が持家であり、居住年数20年以上が92%を占めた。「自動車を保有していない」が52%であった。

台風9号による被害の経験は、「自宅・建物の被害(21%)」、「自宅の孤立(21%)」、「家族の避難(15%)」であった。生活質の評価では、買物、通勤・通学、病院・福祉施設の利便性に関わる事項、公共交通、自転車の移動性に関わる評価の項目が総じて低い。騒音・振動、身近な緑、川・水辺に関する評価が高く、スポーツやレクリエーションの場所については評価が低い。また、危険性に関連する水害に関する安全性(台風や大風)は評価が低い。居住意向では、「現在地に住み続けると思う(75%)」であり、「村内の別の場所に転居すると

表3 自由記述欄の世帯属性別の平均記入語数

世帯主の年齢階層	65才未満	65才以上
サンプル数	109	188
語数の合計	1918	4131
平均語数	17.6	22.0

世帯人数	独居世帯	2人以上
サンプル数	54	240
語数の合計	1067	4940
平均語数	19.8	20.6

思う(1%)」、「村外に転居すると思う(7%)」と村内外に転居意向を持つ人は少ない。転居意向の理由は、「病院などへの通院が不便であること(26%)」、「公共交通が不便であること(16%)」、「災害が心配(12%)」であった。

本調査は、森田らの研究<sup>17)</sup>に用いた調査と同一であり、新たに自由記述データを分析データとして加える。自由記述欄の設問文は「南牧村のイメージを、短い言葉や文で、自由に書いてください。」であり、よいイメージ、わるいイメージの両面を捉えるものとした。

#### 4. 自由記述データの分析

##### (1) 出現する語の分析

自由記述の記載された回収票は298票あり、有効回収数の47.2%を占めた。本項では、自由記述欄の出現語を分析することにより、記入者の意図を検討する。語は、形態素解析器として「茶筌」<sup>19)</sup>を用いて抽出した。表3に世帯属性別の自由記述データの平均記入語数を整理した。語数は、記入者の回答に対する熱心さを表すと考えた。世帯主の年齢では、65才以上の人のほうが多く語数を記入した。また、世帯人数では複数人の世帯のほうが多くの語数を記入した。

表4に自由記述データから抽出した頻出50語を整理した。プリコードデータで評価の高かった緑や水辺については、10位までの体言に「自然」「水」等の南牧村の水緑環境として語が抽出され、20位までに「空気」「山」といったプリコードデータで設定しなかった関連する語が確認された。また、20位までには「高齢者」「不便」「生活」等の過疎の進む地域の居住環境に関する語が確認された。用言をみると、「ない」「多い」等が現れている。プリコードデータで設定した評価の低い、買物、通勤・通学、病院、福祉施設、公共交通などの利便性に関する語を確認することができなかった。65歳代以上と65歳代未満の階層別で出現語を分析する(表5)。出現率に着目すると、「多い」について65歳代以上が26%に対し65歳代未満が31%であった。「良い」について65歳代以上が32%に対し65歳代未満が13%であった。「きれい」について65歳代以上19%に対し65歳代未満が9%であった。水緑環境に関する語では、「自然」について65歳以上が18%に対し65歳未満が28%であった。「水」について65歳以上が17%に対し65歳未満が13%であった。65歳以上と65歳未満ともに、水緑環境に関する頻出数に大きな差異がなく、「多い」まで認識している。しかし、65歳代以上は「多い」の認識から、さらに「良い」「きれい」といった肯定的な評価の記述が多いという傾向が窺われた。

表4 自由記述データの頻出50語

順位	出現語	出現数	順位	出現語	出現数
1	ない	107	26	できる	18
2	村	104		過疎	
3	多い	82		不安	
4	良い	76	29	人口	17
5	自然	64	30	川	15
6	する	60	31	人情	14
7	人	47		道路	
8	水	46	33	場所	13
9	きれい	45		働く	
10	思う	42		おいしい	
11	空気	40		悪い	
12	住む	38	37	出る	12
13	高齢者	33	38	恵まれる	11
14	なる	32	39	とても	10
	南牧村				
16	山	31		環境	
17	ある	27	心		
18	不便	26	43	近所	9
19	生活	25		進む	
20	少ない	24		日本一	
	いる		役場		
22	若い	23	47	今	8
23	高齢化	22		春	
	緑			冬	
25	子供	19		大変(な)	

表5 年齢階層別の頻出語

順位	出現語	65才未満 N=109		65才以上 N=188	
		出現数	出現率	出現数	出現率
1	ない	33	30%	73	39%
2	村	41	38%	63	34%
3	多い	34	31%	48	26%
4	良い	14	13%	61	32%
5	自然	30	28%	34	18%
6	する	18	17%	41	22%
7	人	11	10%	35	19%
8	水	14	13%	32	17%
9	きれい	10	9%	35	19%
10	思う	16	15%	25	13%
11	空気	10	9%	30	16%
12	住む	10	9%	27	14%
13	高齢者	8	7%	25	13%
14	なる	12	11%	20	11%
15	南牧村	5	5%	27	14%

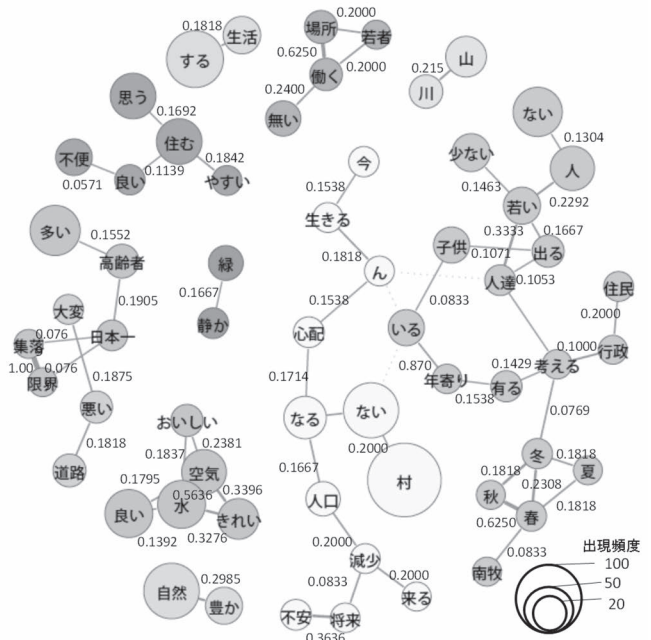


図2 出現語の共起関係(共起ネットワーク)

(2) 出現語間の関係

本節では、出現語間の関係について分析することによって、出現語間の関連性や語がどのような文脈で出現しているかといった語と語のつながりのパターンに着目した分析を行う。出現語間の関係については、自然言語処理で一般的に用いられる語と語の共起関係を分析する方法を用いる。

本研究では、共起関係を確率的に表す Jaccard 係数<sup>21)</sup>を用いることとした。本研究で用いている KH Coder は、Jaccard 係数による分析に対応しており、共起関係を視覚的に出力できる。Jaccard 係数とは、語  $\omega$  と語  $\omega'$  の共起関係を示し、大きいほどつながりが強いことを示す。 $p(\omega \cap \omega')$  は語  $\omega$  と語  $\omega'$  の共起確率を表し、 $p(\omega \cup \omega')$  は語  $\omega$  と語  $\omega'$  のいずれかが出現する確率を表す。語  $\omega$  と語  $\omega'$  の Jaccard 係数は式(1)で与えられる。

$$Jaccard(\omega, \omega') = \frac{p(\omega \cap \omega')}{p(\omega \cup \omega')} \quad (1)$$

Jaccard 係数は、0 から 1 の値をとる。Jaccard 係数の適用方法については、自然言語処理分野において研究が続けられて、図中の語の配置は意味をもたない。作成した共起ネットワークを図2に示す。前節の頻出語にもあった「空気・水・きれい・おいしい・良い」「自然・豊か」「山・川」などの水緑環境に関連するよいイメージの語の結び付きが具体的に示された。また、頻出語には「高齢者」「不便」「生活」で留まっていた居住環境について、「若い・少ない・出る・人達」などのコミュニティに関連するイメージ、「住む・やすい・良い・不安」や「道路・悪い・大変」、「若者・働く・場所・無い」などわるいイメージの語の結び付きが具体的に示された。中でも「村・人口・将来・不安・心配」という時間軸を含んだイメージなど、ブリーコードデータには設定されていない語の結び付きを見ることができた。また、「高齢者・日本一・多い」「限界・集落」など、限界自治体である南牧村ならではの語の結び付きも見られ。一方、災害や防災など「安全性」をイメージする語の結び付きは本分析において見られなかった。



表 6 21 個の評価項目

変数	評価項目	SV11	集落の防犯
SV1	買い物の便利さ	SV12	交通事故の危険が少ない
SV2	通勤・通学の便利さ	SV13	地域の活動(祭, イベントなど)
SV3	公共交通の便利さ	SV14	趣味やスポーツ活動
SV4	郵便局や銀行の便利	SV15	日頃の近所付き合い
SV5	自転車の乗りやすさ	SV16	ゴミや排水などの衛生状況
SV6	歩きやすさ	SV17	住宅、庭のゆとり
SV7	スポーツ施設が身近にある	SV18	騒音・振動が少ない
SV8	まちなみや家なみのよさ	SV19	日当たりや風通し
SV9	地震、火災に関する安全性	SV20	身近な川・水辺に恵まれる
SV10	水害の安全性(台風・大雨)	SV21	身近な緑に恵まれる

5. 自由記述データに着目した生活質評価の分析

(1) 自由記述欄の記入有無別の生活質評価の特性

森田らは<sup>17)</sup>、南牧村を対象に、客観的変数である世帯属性、地区特性、主観的変数の生活質評価・居住意向、生活質の潜在変数の関係を統合的に検証した。本研究では、プリコードデータで構築した生活質評価のモデルに自由記述データの有無が影響を与えると考えた。表6は、生活質評価を行うために設定した21個の生活質評価の項目である。先行研究<sup>17)</sup>を参考としながら、図3に示すモデルを仮定した。モデルは、21個の評価を最下位の観測変数として配置し、先行研究<sup>17)</sup>において行った因子分析の結果を踏まえ、観測変数の上位に「利便性」「安全性」「コミュニティ」「居住環境」「水緑環境」の5つの潜在変数を配置した。さらに、5つの潜在変数をまとめる

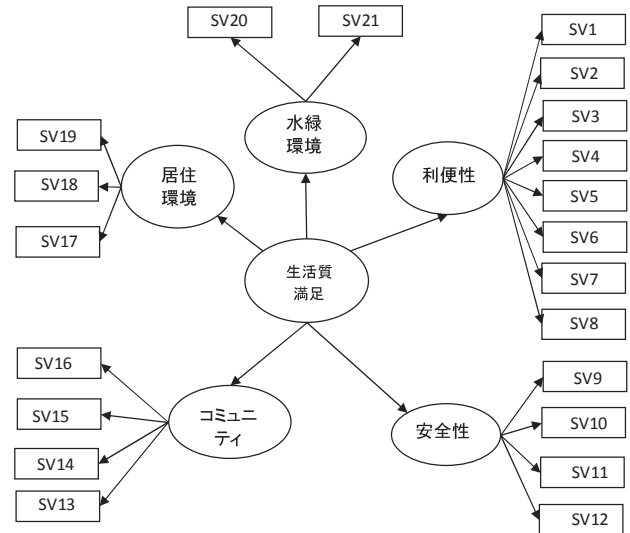


図3 生活質評価モデルの体系

潜在変数として最上位に「生活質満足」を配置した3階層のモデルで共分散構造分析を行った。自由記述を「記入あり」と「記入なし」に分類し分析した結果を表7に示す。最上層の「生活質満足」と5つの潜在変数との関係では、「記入あり」と「記入なし」の双方で、標準化係数の値より「コミュニティ」が最も影響を与えていた。

「記入あり」と「記入なし」において、標準化係数の差の大きさに着目すると、「生活質満足」と「居住環境」の標準化係数の差は、「記

表 7 生活質評価の共分散構造分析の結果

変数	標準化係数	t 値	検定	(記入あり)			(記入なし)			標準係数の差
				標準化係数	t 値	検定	標準化係数	t 値	検定	
利便性 <- 生活質満足	0.726	—	—	0.750	—	—	0.750	—	—	-0.024
安全性 <- 生活質満足	0.711	7.042	**	0.795	7.457	**	0.795	7.457	**	-0.084
コミュニティ <- 生活質満足	0.801	7.435	**	0.850	7.187	**	0.850	7.187	**	-0.049
居住環境 <- 生活質満足	0.629	6.076	**	0.762	6.853	**	0.762	6.853	**	-0.133
水緑環境 <- 生活質満足	0.520	5.481	**	0.532	5.278	**	0.532	5.278	**	-0.012
SV1 <- 利便性	0.721	—	—	0.692	—	—	0.692	—	—	0.029
SV2 <- 利便性	0.627	8.993	**	0.630	8.605	**	0.630	8.605	**	-0.003
SV3 <- 利便性	0.647	9.272	**	0.749	10.067	**	0.749	10.067	**	-0.102
SV4 <- 利便性	0.550	7.932	**	0.644	8.790	**	0.644	8.790	**	-0.094
SV5 <- 利便性	0.582	8.387	**	0.694	9.416	**	0.694	9.416	**	-0.112
SV6 <- 利便性	0.560	8.082	**	0.673	9.149	**	0.673	9.149	**	-0.113
SV7 <- 利便性	0.558	8.052	**	0.566	7.783	**	0.566	7.783	**	-0.008
SV8 <- 利便性	0.554	7.990	**	0.576	7.922	**	0.576	7.922	**	-0.022
SV9 <- 安全性	0.814	—	—	0.831	—	—	0.831	—	—	-0.017
SV10 <- 安全性	0.694	10.417	**	0.800	12.643	**	0.800	12.643	**	-0.106
SV11 <- 安全性	0.684	10.272	**	0.682	10.566	**	0.682	10.566	**	0.002
SV12 <- 安全性	0.578	8.656	**	0.593	8.977	**	0.593	8.977	**	-0.015
SV13 <- コミュニティ	0.800	—	—	0.739	—	—	0.739	—	—	0.061
SV14 <- コミュニティ	0.724	10.901	**	0.669	8.810	**	0.669	8.810	**	0.055
SV15 <- コミュニティ	0.640	9.643	**	0.594	7.918	**	0.594	7.918	**	0.046
SV16 <- コミュニティ	0.597	8.966	**	0.565	7.547	**	0.565	7.547	**	0.032
SV17 <- 居住環境	0.743	—	—	0.768	—	—	0.768	—	—	-0.025
SV18 <- 居住環境	0.614	7.355	**	0.534	6.878	**	0.534	6.878	**	0.080
SV19 <- 居住環境	0.618	7.377	**	0.682	8.350	**	0.682	8.350	**	-0.064
SV20 <- 水緑環境	0.821	—	—	0.793	—	—	0.793	—	—	0.028
SV21 <- 水緑環境	0.853	7.960	**	0.702	5.725	**	0.702	5.725	**	0.151
適合度指数	GFI	AGFI		GFI	AGFI		GFI	AGFI		
	0.822	0.779		0.834	0.794		0.834	0.794		

\*\*1%有意

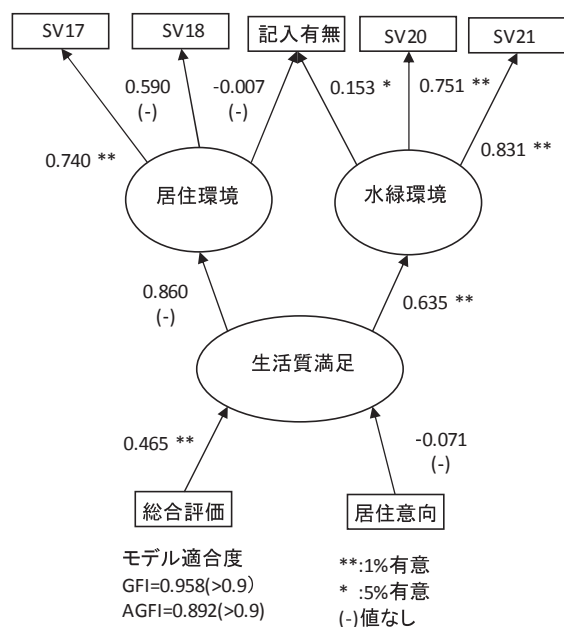


図4 生活質評価と自由記述の関係モデル

入あり (0.629)」に対し、「記入なし (0.762)」であり、配置した5つの潜在変数の中で差が-0.133と最も大きい値であった。また、「生活質満足」と「水緑環境」の標準化係数の差は、「記入あり (0.520)」に対し、「記入なし (0.532)」であり、配置した5つの潜在変数の中で差が-0.012と最も小さな値であった。潜在変数と観測変数の関係では、「利便性」と「SV3:公共交通の便利さ(係数の差-0.102)」「SV5:自転車の乗りやすさ(係数の差-0.112)」、「安全性」と「SV10:水害に関する安全性(係数の差-0.106)」が大きい値であった。

この結果から、「記入あり」と「記入なし」の属性において、生活質の評価は、「居住環境」の評価に大きな影響の差があり、「水緑環境」の評価で小さな影響の差となった。

## (2) 生活質評価と自由記述データの関係

本節では、生活質評価と自由記述データの内容について関係を分析する。自由記述データの内容は、前節において、「記入あり」と「記入なし」で影響の差に大きな相違があり、かつ頻出語として抽出された「居住環境」と「水緑環境」に関する単語とし、記入ありを「1」、記入なしを「0」としてダミー変数化した観測変数「記入有無」で扱った。図4に示す生活質評価と自由記述データの記入の関係を表すモデルを仮定し、共分散構造分析による分析を行った(図中の値は標準化係数)。

モデル全体の適合度は、GFI=0.958、AGFI=0.892となり適合度は高い結果であった。観測変数である「SV17:住宅、庭のゆとり」「SV18:騒音・振動が少ない」の回答が「居住環境」に与える影響、また「SV20:身近な川・水辺に恵まれている」、「SV21:身近な緑に恵まれている」の回答が「水緑環境」に与える影響についてはいずれも1%有意であった。

「記入有無」は影響について、「水緑環境」においては5%有意水準を満たしたが、「居住環境」との間については有意差がみられなかった。標準化係数の符号に着目すると、「記入有無」から「水緑環境」への標準化係数はプラス値をとっているのに対して、「居住環境」へ

の標準化係数は、マイナス値となっている。すなわち、有意性は示されなかったが、「居住環境」に関しての記述データは、「生活質満足」に肯定的に作用されているのではなく、むしろ否定的な内容を含み、否定的に作用する結果であると推察される。

以上から、プリコードデータである「SV17:住宅、庭のゆとり」「SV18:騒音・振動が少ない」「SV13:地域の活動」「SV14:趣味やスポーツ活動」の回答は、いずれも「生活質満足」にプラスに作用した。一方、限界自治体である南牧村を事例としたイメージの自由記述データは、「水緑環境」はプラスの値に作用されるものの、「居住環境」にはマイナスの値に作用される傾向が推察された。

## 6. おわりに

### (1) 研究のまとめ

本稿では、定量的な分析の観点からアンケート調査の自由記述データを活用するために、テキストマイニングを用いた定量的な分析を行った。対象地域を限界自治体である群馬県南牧村とし、プリコードでは把握できない情報の特性について検討した結果、以下のような知見が得られた。

- 1) 年齢属性の影響を踏まえて、自由記述の頻出語について65歳未満と65歳以上に整理したところ、65歳以上が、水緑環境に関する頻出語の出現に大きな差異はなかったが、肯定的な記述をしている傾向がうかがえた。
- 2) 共起ネットワーク図により、語間の関係について検討した。図より、水緑環境を連想するよいイメージの単語の結び付き、コミュニティを連想するイメージ、居住環境に関連するイメージの結び付きが示された。中でも時間軸を含むわるいイメージなど、プリコードデータには設定されていない内容を読み取ることができた。一方で、限界集落が課題とする地震・火災、水害、防災などの安全性に連想するイメージの結び付きは見られなかった。また、65歳以上と65歳未満ともに、水緑環境に関する頻出数に大きな差異がなく、「多い」まで認識しているが、65歳代以上は「多い」の認識から、さらに「良い」「きれい」といった肯定的な評価の記述が多いという傾向が窺われた。
- 3) プリコードデータに記述有無を加えて、共分散構造分析により生活質の評価を構造化した。「生活質満足」と5つの潜在変数との関係では、「記入あり」と「記入なし」の双方で、標準化係数の値より「コミュニティ」が最も影響を与えていた。また、「記入あり」と「記入なし」において、標準化係数の差の大きさに着目すると、「居住環境」の評価は大きな影響の差があり、「水緑環境」の評価は小さな影響の差であった。
- 4) 生活質の評価と自由記述との関係では、限界自治体である南牧村のイメージに関する自由記述に、「水緑環境」においては正の値に評価するものの、「居住環境」には、負の値に評価するという特性が得られた。このように限界自治体である南牧村の評価においてプリコードデータによる分析でなく、自由記述データを加えることにより、より実態的なモデルの検証が期待できる可能性が明らかになった。
- 5) 今回のケースでは、南牧村の生活質を改善する地域計画の施策として、定住促進と人口流出対策を検討した場合、定住促進を目指した施策としての魅力発掘においては、水緑環境面でプリコー

ドデータに加えて自由記述データを活用することにより具体的な全体像の把握に効果的な活用できるのではないかと考える。しかし、限界自治体である南牧村の課題の抽出といった面においては、今回のケースでは、居住環境の実態に十分な全体像の把握への活用が見いだせなかった。

## (2) 今後の研究課題

群馬県南牧村を事例とした本研究から、プリコードデータに加えて自由記述データを活用する方法に関する今後の課題を述べる。今回の自由記述欄の設問であった「南牧村のイメージを自由に書いてください。」の自由記述では、水緑環境を中心とする肯定的な記述が顕著にうかがわれた。今回は南牧村に限られた事例研究であるが、例えば、限界自治体の定住促進に伴う魅力づくりの創出を行うための基礎資料として、例えば景観計画の作成や地域の景観のあり方の意思決定にこれらの自由記述の活用が有効であると考えられた。

一方、高齢者が多いといったプリコードデータに設定されていない居住環境の将来不安の記述があったが、居住環境における生活質の評価モデルとの関連性が十分に得られなかった。限界自治体、限界集落の課題解決に自由記述を活用するためには、将来不安をどのように考えていくべきかといった直接的な設問内容の設定などが自由意見の抽出に必要と考えられ、これについては自由記述欄のアンケート設計も踏まえて検討したい。さらに、生活質評価をおこなうために設定した21個の生活質の評価項目に、歴史・文化に対する思いや誇りに関する項目がない。歴史・文化は生活質評価を行う上で重要な構成要素でもあると考えることから、今後の生活質評価の研究の展開に際して評価項目へ加えていくことを今後の課題とした。

## 参考文献

- 1) 大野晃：山村環境社会学序説・現代山村の限界集落化と流域共同管理，社団法人農山漁村文化協会，2005。
- 2) 福田大輔，庭田美穂，屋井鉄雄：疑問型表現自由回答データを用いた社会资本整備に対する市民の関心の抽出方法に関する基礎的研究，土木計画学研究・論文集，Vol.24，pp.139-148，2007。
- 3) 鄭蝦榮，羽鳥剛史，小林潔司，白松俊：ファセット学習モデルを用いた公的討議のプロトコル分析，土木計画学研究発表会・講演集，Vol.36，2007。
- 4) 佐々木邦明，飯島陽介，鈴木猛康，秦康範，大山勲：ワークショップ運営支援のためのテキスト分析，土木学会論文集 F4，Vol.66，No.1，pp.57-64，2010。
- 5) 佐々木邦明，丸石浩一：テキストマイニングを用いたワークショップの討議内容の特徴把握と可視化に関する研究，都市計画論文集，Vol.46，No.3，pp.1039-1044，2011。
- 6) 塚田伸也，森田哲夫，湯沢昭：委員会の発言から捉えた歩行者の交通空間整備に関する検討，第31回交通工学研究発表会論文集，pp.503-506，2011。
- 7) 佐々木邦明，紀藤舞華，山崎慧太：ログマイニングからの行動データ抽出・分析可能性とアンケート調査との比較，土木計画学研究・講演集，Vol.43，CD-ROM(145)，2011。
- 8) 長尚希，室町泰徳，板谷和也：計量的言語処理を利用した大規模交通プロジェクトに関する経験知識の抽出に関する研究，都市計画論文集，Vol.47，No.3，pp.793-798，2012。
- 9) 森田哲夫，入澤寛，長塩彩夏，野村和広，塚田伸也，大塚裕子，杉田浩：自由記述データを用いたテキストマイニングによる都市のイメージ分析，土木学会論文集 D3，Vol.68，No.5，I\_315-I\_323，2012。
- 10) 永野峻祐，小根山裕之，大口敬，鹿田成則：形態素解析を用いたアンケート調査自由記述欄の分析手法に関する研究—路面電車利用意識データを用いたケーススタディー，土木学会論文集 D3，Vol.68，No.5，I\_973-I\_981，2012。
- 11) 溝上章志，川島英敏，大森久光，永田千鶴，野尻晋一，矢口忠博：高齢化社会においてパーソナルモビリティが QOL に与える影響に関する実証調査，土木学会論文集 D3，Vol.68，No.5，I\_141-I\_153，2012。
- 12) 森山昌幸，藤原章正，杉恵頼寧：過疎地域における公共交通サービスの評価指標の提案，日本都市計画学会都市計画論文集，Vol.8，No.3，pp.475-480，2003。
- 13) 宮崎耕輔，徳永幸之，喜多秀行，谷本圭志，菊池武弘，高山純一：過疎地域におけるバス運行サービスの変化が地域住民の生活に与えた影響分析に関する研究，土木学会土木計画学研究・講演集，Vol.33，CD-ROM(78)，2006。
- 14) 谷本圭志，森健治：地方部における定住意向と社会生活環境の関係に関する考察—住民のライフステージに着目して—，環境システム研究論文集，Vol.35，pp.19-27，2007。
- 15) 日本交通政策研究会：日交研シリーズ A-473，「限界集落」を対象とした中山間地域のモビリティの確保と地域再編戦略に関する研究，2009。
- 16) 森田哲夫，塚田伸也，佐野可寸志：過疎・高齢地域における集約型居住に向けた人口動向・居留意識の分析—群馬県六合村におけるケーススタディー—，都市計画論文集，Vol.45，No.3，pp.511-516，2010。
- 17) 塚井誠人，桑野将司：中山間地域住民の移住意向と移住要件に関する分析，都市計画論文集，Vol.45，No.3，pp.277-282，2010。
- 18) 森田哲夫，木暮美仁，塚田伸也，橋本隆，杉田浩：限界自治体の生活質と居留意向に関する研究，社会技術研究論文集，Vol.10，pp.86-95，2013。
- 19) 樋口耕一：テキスト型データの計量的分析—2つのアプローチの峻別と統合—，理論と方法，No.19(1)，pp.101-115，2004。
- 20) KH Coder，<http://khc.sourceforge.net/>，2014.2.26 閲覧。
- 21) ChaSen「茶筌」形態素解析器，<http://chasen-legacy.sourceforge.jp/>，2013.5.1 閲覧。

# AN EVALUATION ON THE QUALITY OF LIFE IN MARGINAL LOCAL COMMUNITY CONSIDERING FREE ANSWER

– A case study on Nanmoku-mura, Gunma –

*Shinya TSUKADA\**, *Tetsuo MORITA\*\**, *Toshikazu NISHIO\*\*\**  
*and Akira YUZAWA\*\*\*\**

\* Construction Department, Maebashi City Office, Dr. Eng.

\*\* Prof., Tohoku Institute of Technology, Dr. Eng.

\*\*\* Graduate Student, Maebashi Institute of Technology, M. Eng.

\*\*\*\* Prof., Maebashi Institute of Technology, Dr. Eng.

## 1. Purpose of This Study

Currently, the decrease in population and the aging have become important problem in Japan. The purpose of this study is analyzing quality of life in the marginal local community, considering free answer. The data of the questionnaire survey in Nanmoku-Mura Gunma was used for all analysis. Nanmoku-Mura is one of the typical marginal local communities. This study aims to discuss the characteristics and analysis method of free answers in questionnaire survey in comparison with pre-code data and free answers, by one of the language processing techniques for text data acquired from free answers. We clarified that the quality of life and residential attitude are composed of individual attributes and district characteristics.

## 2. Method of This Study

In this study, operability is high and it is used in the thesis of many fields, KH Coder using freeware for text mining the elemental technology of the natural language processing uses a morpheme analyzer. I can extract a high word of the precision and pay my attention to an attribute of the freedom description presence. The result is obtained in the data of the quality of life evaluation model that utilized pre-cord data, using covariance structure analysis (SPSS Amos17).

## 3. Measure the Life satisfaction of inhabitants

I set an area for that is the limit local government and I collect data about a quality of life evaluation of inhabitants, the residence intention, free description data and arrange it. I analyze entry properties of the free description data in the limit local government and I grasp the characteristic of the description data which are not arrested in the end-point of pre-cord data and, furthermore,

## 4. Analysis of the quality of Life evaluation to free answer data

The coincidence network diagram considered the relation of spacing. Thereby, connection of connection of the good word of an image reminded of water green environment, the image reminded of a community, and the image relevant to living environment was shown. The contents which are not set up were able to read to a bad image about free description data by performing analysis using the text mining.

## 5. Effects of Life evaluation to free answer data

Paying attention to whether there is any description, covariance structure analysis of the nature evaluation of a life was conducted. As a result, the tendency for the influence of life satisfaction, living environment and convenience, public traffic and a car and safety, and flood damage to have a difference has been grasped.

## 6. Subjects for Next Study

I hope that this study provided based on the analysis to arrange it about the information in conjunction with the quality of life evaluation that was provided by paying its attention to freedom description data.

(2014年7月10日原稿受理, 2014年10月27日採用決定)