

# 計量テキスト分析による被災者の自由記述内容に関する構造分析 —被災2年後・2013年調査を事例として—

A Study on the Structure of Victim's opinion using the Quantitative Text Analysis  
 A Case Study of 2013 Survey After two years of the Great East Japan Earthquake

坪井 塑太郎<sup>1</sup>, 中林 一樹<sup>2</sup>, 小田切 利栄<sup>3</sup>, 土屋 依子<sup>4</sup>  
 Sotaro TSUBOI<sup>1</sup>, Itsuki NAKABAYASHI<sup>2</sup>, Rie OTAGIRI<sup>3</sup>, Yoriko TSUCHIYA<sup>3</sup>

<sup>1</sup>公益財団法人ひょうご震災記念21世紀研究機構 人と防災未来センター  
 Disaster Reduction and Human Renovation Institution

<sup>2</sup>明治大学大学院 政治経済学研究科  
 Graduate school of Political Science and Economics, Meiji University

<sup>3</sup>明治大学 研究・知財戦略機構・株式会社オリエンタルコンサルタンツ  
 The organization for the Strategic Coordination of Research and Intelligence Properties, Meiji University

<sup>4</sup>明治大学 研究・知財戦略機構・国土交通省国土交通政策研究所  
 The organization for the Strategic Coordination of Research and Intelligence Properties, Meiji University

The purpose of this research is to clarify the support, requests, and dissatisfaction contents at the initial stage of recovery from the description contents in the questionnaire survey. In this research, structural analysis was carried out using quantitative text analysis on the content of free Descriptive answer in 2013 after two years of disaster. In this analysis, considering the relation between the terms and the grammatical structure of Japanese Language. As a result of this research, discontent appeared in "work (employment)", "housing", "health", "transportation".

**Keywords** : Free descriptive answer, Quantitative Text Analysis, Grammatical structure, Life reconstruction Evaluation, The Great East Japan Earthquake

## 1. 研究背景と目的

復興初期段階における被災者の支援・要望の傾向を構造的に明らかにすることは、効果的な生活再建施策のあり方を検討するうえで重要な課題であると考えられる。本研究は、岩手県大船渡市・宮城県気仙沼市・福島県新地町を対象に、東日本大震災による被災から一年後の2012年3月より、被災者の復興感の変遷の考究を目的として、毎年継続して郵送式質問紙調査を実施しており、本稿では、被災から2年後の2013年調査の結果をもとに論考を行う。

質問紙調査に設定される自由記述欄（自由回答）は、これまで、尺度評価の結果を補完説明するための素材として多くは用いられてきた経緯を持つが、近年では優れた言語形態素解析エンジンの開発により、記述内容（テキスト）を主体的かつ計量的に分析をする方法が一般化してきている。しかし、自由記述欄の多くは、特定の項目に対する記述を促すものは少なく、その結果、多様な回答が混在することから、分析と解釈に高い負担が要されている。また、日本語のもつ特有の文法構造を考慮した解釈に際しての採用品詞の議論は必ずしも充分ではないことが課題となっている。本研究ではこれらの問題意識のもと、一括で取得された自由記述内容から被災者の支援・要望、不満に関する意見を構造的に明らかにすることを目的とする。表1に2013年調査の概要を示す。分析に際しては、「記述あり」の回答者のうち属性欠損値を除去した1,296件を対象とする。本調査では調査票回答者のうち対象3地域全てにおいて50%以上の自由回答記述がみられたことが特徴として挙げられる。

表1 調査概要（被災2年後・2013年調査）分析対象

項目	岩手県 大船渡市		宮城県 気仙沼市		福島県 新地町	
	人数	割合	人数	割合	人数	割合
基本項目	発送数	3141	4538	516		
	回収数	857	1324	168		
	回収率	27.3%	29.2%	32.6%		
分析対象 自由回答	記入者数	467	729	100		
	記入率	54.5%	55.1%	59.5%		

項目	人数		割合		人数		割合	
	人数	割合	人数	割合	人数	割合	人数	割合
年齢	30歳代以下	20	4.3%	17	2.3%	7	7.0%	
	40歳代	48	10.3%	88	12.1%	10	10.0%	
	50歳代	110	23.6%	164	22.5%	19	19.0%	
	60歳代	156	33.4%	226	31.0%	37	37.0%	
	70歳代	100	21.4%	177	24.3%	20	20.0%	
	80歳代以上	33	7.1%	57	7.8%	7	7.0%	
被災状況	全壊	243	52.0%	512	70.2%	63	63.0%	
	大規模半壊	34	7.3%	73	10.0%	7	7.0%	
	半壊	39	8.4%	43	5.9%	3	3.0%	
	一部破損	67	14.3%	61	8.4%	23	23.0%	
	無被害	84	18.0%	40	5.5%	4	4.0%	
現在居住	応急仮設住宅	111	23.8%	203	27.8%	40	40.0%	
	借上仮設住宅	41	8.8%	139	19.1%	3	3.0%	
	自宅/被災前同地	231	49.5%	238	32.6%	40	40.0%	
	自宅/移転新築	34	7.3%	69	9.5%	9	9.0%	
	親類宅同居	11	2.4%	22	3.0%	1	1.0%	
	賃貸住宅	19	4.1%	27	3.7%	2	2.0%	
	その他	20	4.3%	31	4.3%	5	5.0%	
生活 復興感	0~20%	69	14.8%	155	21.3%	18	18.0%	
	30~40%	105	22.5%	197	27.0%	23	23.0%	
	50~60%	120	25.7%	150	20.6%	31	31.0%	
	70~80%	91	19.5%	157	21.5%	23	23.0%	
	90~100%	82	17.6%	70	9.6%	5	5.0%	

## 2. 生活全般の主観的復興感の変化

図1に被災1年後(2012年調査)と被災2年後(2013年調査:本稿分析対象)における復興感の変化を示す。本図は、生活全般の「望ましい生活」に対する復興感について、X軸(横軸)に「復興感の比率」を、Y軸(縦軸)に「回答者の累積比率」をとり、その形状が上に凸でY軸の高位に位置するほど復興感・回復感が高い回答者が多いことを示す。

復興感の比率「70%」の割合は、2012年の22%から2013年には33%と11ポイントの上昇がみられたほか、「50%」まででは同44%から57%と13ポイントの上昇がみられた。しかし、全体の約40%は震災前の半分の水準に達していないことが看取できる。

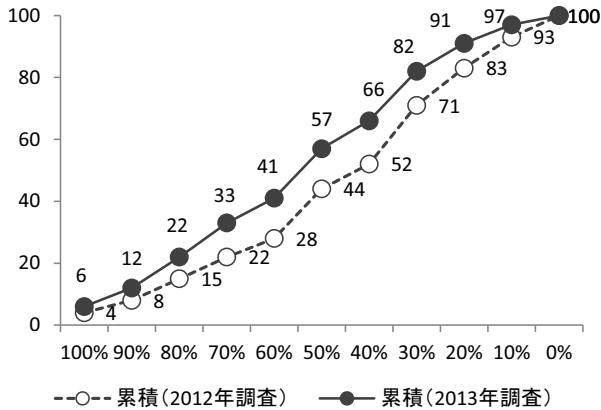


図1 生活全般の望ましい生活に対する復興感変化

## 3. 計量テキスト分析における採用品詞の検討

日本語の形態素は、品詞とその前後の接続関係をもとに分類され、分析における変数の採否は分析者の判断に委ねられることが多く、これまで必ずしも決まった方法が提示されているわけではない。そこで本研究では、全文を品詞別に分類を行った後、表2に示す形態素の採否の検討を行い、用語内容を考慮して「名詞」「形容動詞」「形容詞」とあわせ、動詞「する」に接続してサ行変格活用の動詞となる「サ変名詞」の4品詞を採用した。このほか、本研究における自由回答記述の末尾に用いられる用語の多くは要望を示す動詞「ほしい」や、個人の主張、考えを示す動詞「思う」が多く、この品詞の自立語にはキーワードとしての表現も含まれることから全文を見直した上で採否の検討を行った結果、特徴語との連結が少ないことからこれを不採用とした。

形態素の抽出に当たっては、繰り返して品詞別の分析を行い、出現用語中の「ごみ」「ゴミ」、「子供」「子ども」「こども」、「仮設」「仮設住宅」「仮設住居」など同義で表記、表現方法の異なるものについては随時関連語として登録作業を行ったほか、明らかな語句の誤りや文字の誤りについては原文をもとに前後の文脈から判断し、大意を変更させず修正を行った。表3に、地域別・品詞別出現用語と出現数一覧を示す。3地域に共通して名詞では「住宅」「土地」の用語が、サ変名詞では「支援」「仮設」「生活」「再建」「仕事」が、また形容動詞では「大変」「必要」「不安」が頻出したほか、形容詞では早期の復興要望を指す「早い」(早くの原形)や、支援への謝意を指す「有難い」「嬉しい」といった用語が特徴的な傾向として挙げられる。

表2 形態素の種類と解析における採否

品詞区分	詳細	採否	品詞区分	詳細	採否		
記号	—	×					
形容詞	自立	◎	副詞	非自立	△		
	接尾	◎		一般	△		
	非自立	◎		助詞類接続	△		
	格助詞—一般	×		サ変接続	◎		
	格助詞—引用	×		一般	◎		
	格助詞—連語	×		形容動词语幹	◎		
	係助詞	×		固有名詞—組織	◎		
	接続助詞	×		数	△		
	副詞化	×		接尾—一般	△		
	副助詞	×		接尾—助数詞	△		
助動詞	副助詞—終助詞	×	名詞	接尾—助動词语幹	△		
	並立助詞	×		接尾—特殊	△		
	連体化	×		代名詞—一般	△		
	—	×		非自立—一般	△		
	—	×		非自立—副詞可能	△		
	接頭詞	×		副詞可能	△		
	動詞	自立		△	連体詞	—	△
		接尾		×	EOS	End of Sentence	×

注: ◎: 採用, △: 不採用(考慮), ×: 不採用 △については、内容に応じて採否の考慮・検討を行ったが、本研究では不採用とした。

表3 地域別・品詞別出現用語と出現数一覧

	名詞		サ変名詞		形容動詞		形容詞	
	用語	出現数	用語	出現数	用語	出現数	用語	出現数
岩手県大船渡市	住宅	135	支援	307	必要	42	早い	54
	高台	47	仮設	117	大変	38	少ない	35
	土地	47	被災	112	不安	22	多い	32
	震災	45	生活	90	不便	16	高い	16
	地域	40	再建	58	色々	13	無い	15
	津波	37	復興	48	可能	7	良い	14
	方々	35	仕事	36	健康	7	若い	11
	子供	33	移転	35	安全	6	欲しい	11
	皆様	29	お願い	33	不公平	6	有難い	10
	自分	28	感謝	31	平等	6	新しい	9
宮城県気仙沼市	住宅	236	支援	481	必要	58	早い	118
	震災	103	仮設	211	大変	42	多い	58
	土地	73	被災	192	不安	37	無い	33
	子供	71	生活	175	健康	15	良い	32
	個人	53	復興	140	十分	14	高い	27
	地域	50	再建	107	危険	11	少ない	25
	人達	49	仕事	71	無理	11	欲しい	21
	場所	48	お願い	68	いろいろ	9	悪い	17
	方々	45	移転	48	安定	9	若い	17
	自分	42	感謝	42	公的	9	新しい	13
福島県新地町	住宅	35	支援	63	大変	9	早い	12
	個人	10	生活	26	必要	9	有難い	6
	子供	8	仮設	22	不安	8	少ない	5
	震災	8	復興	19	いろいろ	3	良い	5
	ボランティア	7	被災	16	公的	3	遠い	4
	状態	7	再建	11	早急	3	多い	4
	津波	7	仕事	11	不公平	3	無い	4
	被害	7	移転	10	健康	2	嬉しい	3
	物資	7	お願い	9	不便	2	長い	3
	集団	6	避難	8	安心	1	新しい	2

被災1年後の2012年3月に実施した調査における自由回答記述の特徴は、学校施設・敷地が避難場所として長期に利用されたほか、学校自体が被災し、授業再開時期の制約等があったことから、学習支援への要望や、児童・生徒への心のケアの要望が挙げられた。また、高齢者への支援として、「買物」や「通院」などの交通・移動関連の改善要望が挙げられたほか、復興感の高い群では介護やデイサービスの改善等に関する要望が表明されていることが明らかになった。しかし、本調査(被災2年後)においては、生活基盤としての「住宅」に関する用語は継続的に見られたものの、上述の用語の出現は少なく、代わって、「仕事」の再開に関連した用語や、「健康」の不安に関する用語の出現がみられた。

#### 4. 自由回答記述の全体像

本分析にあたっては任意の文章において、ある文字列と文字列が同時に出現する「共起」に着目し、この出現パターンの類似した用語を結束させることで、記述内容の全体像を示す「共起ネットワーク分析」を通して自由回答記述の全体像の検討を行う。図2に結果を示す。本図より、本調査が被災後2年目の復興初期段階にあたることから、住宅関連の用語として「仮設」や「再建」がみられるほか、生活再建のための仕事に関する用語や、同時期の高台移転に関する用語の出現がみられた。

次いで、属性と用語の関連を対応分析から検討を行う。本分析は、出現用語の行と列の相関関係が最大になるよう数量化を行い、それぞれの要素を多次元空間に表現することで、関連の強弱を各属性との位置関係において解釈を行う。図3に復興感と被害状況を外的変数として投入した結果を示す。本図より、X軸（横軸）において正方向には被害程度が大きく、復興感が低い項目が、負方向には、被害程度が小さく、復興感が高い項目が布置された。このうち、原点付近の「支援」の用語は、全カテゴリからの記述がみられたが、比較的被害が軽微で、高い復興感を持つ第2象限、第3象限においては、自身以外への配慮や、まち全体に対する要望が挙げられたことが特徴となっている。

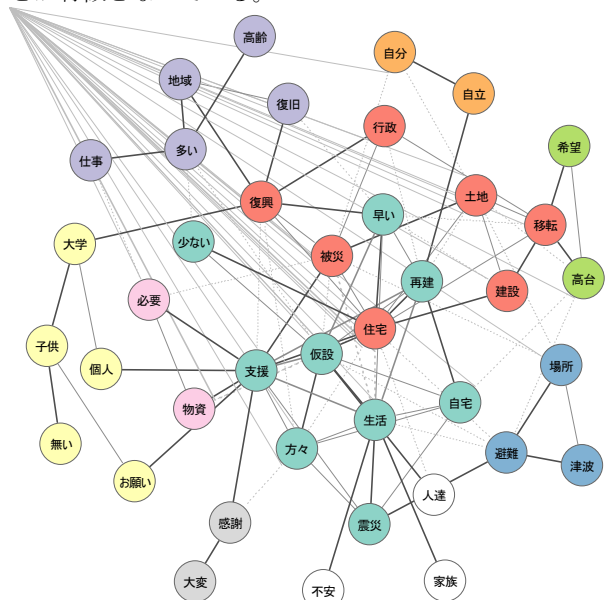


図2 共起ネットワーク分析

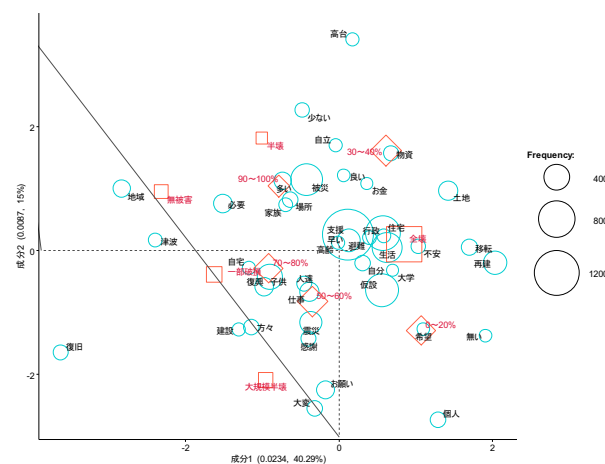


図3 対応分析 (復興感・被害状況)

#### 5. KWIC分析

本分析では、前項での全体像を踏まえ、分析対象ファイル内で、特定の用語（キーワード）がどのように用いられているのかを探的に明らかにする。分析に当たっては、KWIC (Key Word in Context) コンコーダンスによりキーワード前後の文章の抽出を行い（図4）、あわせてテキスト全体の中で出現する回数や共起強度の集計を行った。記述内容に出現する多様な要望項目のうち、本研究では、あらかじめ、Jaccard 係数により項目間の距離（非類似性）を算出し、これをもとに、Ward 法による階層クラスター分析を行ったうえで、「仕事」「住宅」「健康」「交通」の4用語に関連する語の抽出を行った。表4に抽出語と Collocation 統計表を示す。本統計は、例えば対象語「仕事」(nord word)の前後に、どのような用語が出現しているのかを全体の数量と、共起関係を併せて抽出することで、特定の語との関連を明らかにすることが可能となる。

本研究では活用語の語幹を対象としているため、用語の前後に出現する否定語の位置関係により、「必要+ない」などの対立解釈が成立することが想定される。この点において、Collocation 統計表を用いることで、語の係り受け状況を明示でき、健康関連語においては「相談」のほか「負担」や「心配」などの不安感が表明されているほか、交通関連語は、被災1年後の2012年調査と同様、「不便」「悪い」といった不満が表明されていることが明らかになった。

図4 KWIC Concordance



表4 抽出語と Collocation 統計表

	抽出語	品詞	全体	共起	jaccard
仕事 関連 連語	再開	サ変名詞	23 (0.007)	8 (0.075)	0.0656
	収入	名詞	39 (0.011)	8 (0.075)	0.0580
	多い	形容詞	86 (0.025)	7 (0.065)	0.0376
	必要	形容動詞	103 (0.030)	7 (0.065)	0.0345
住宅 関連 連語	関係	サ変名詞	20 (0.006)	4 (0.037)	0.0325
	仮設	サ変名詞	279 (0.080)	145 (0.445)	0.3152
	再建	サ変名詞	159 (0.046)	70 (0.215)	0.1687
	支援	サ変名詞	696 (0.200)	89 (0.273)	0.0954
健康 関連 連語	公営	名詞	29 (0.008)	29 (0.089)	0.0890
	早い	形容詞	153 (0.044)	37 (0.113)	0.0837
	相談	サ変名詞	29 (0.008)	4 (0.174)	0.0833
	管理	サ変名詞	4 (0.001)	2 (0.087)	0.0800
交通 関連 連語	一人暮らし	サ変名詞	8 (0.002)	2 (0.087)	0.0690
	負担	サ変名詞	16 (0.005)	2 (0.087)	0.0541
	心配	サ変名詞	42 (0.012)	3 (0.130)	0.0484
	復旧	サ変名詞	65 (0.019)	6 (0.214)	0.0690
交通 関連 連語	不便	形容動詞	23 (0.007)	4 (0.143)	0.0851
	悪い	形容詞	21 (0.006)	4 (0.143)	0.0889
	公共	名詞	16 (0.005)	6 (0.214)	0.1579
	機関	名詞	13 (0.004)	9 (0.321)	0.2813

## 6. コーディング分析

自由回答記述には多様な内容が含まれ、語自体の関連や出現状況の分析については、共起分析、対応分析およびKWIC分析により一定程度の明示が可能である。しかし、当該の用語を含む「構成概念」をCodeとして集約することで、その出現状況と属性間の関係を明示することが可能となる。

本研究では「病院」「不安」「通院」の用語が含まれる場合には、「健康」Codeに集約するためのコーディングを行ったほか、「仕事」（雇用・収入）、「住宅」、「交通」（不便・遠い・道路・通勤・通学）についても、適宜、関連語分析と共起ネットワーク分析を連動させながらCode化を行った。

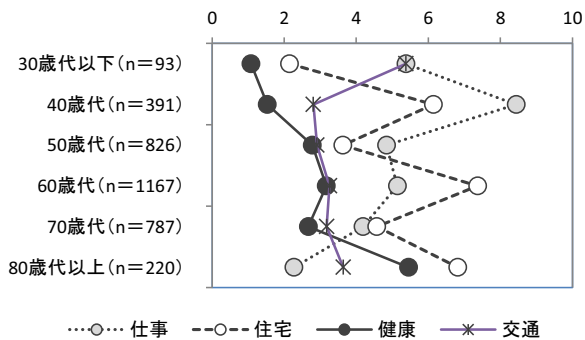


図5 年齢別・Codeの出現割合 (%)

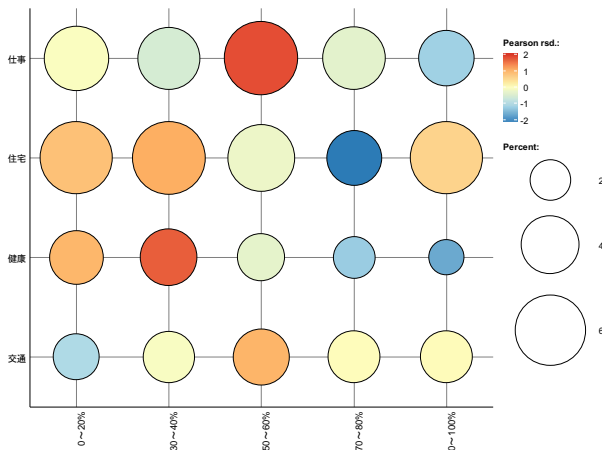


図6 復興感別・Codeの出現割合プロット

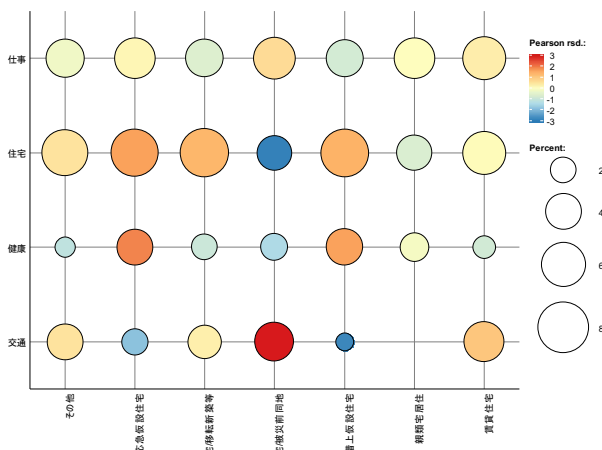


図7 現在居住別・Codeの出現割合プロット

図5に、年齢別・Codeの出現割合を示す。本図より、健康関連の語を含む「健康Code」は80歳代以上において、「仕事Code」は、40代において高い出現率があることがわかる。このほか、図6に示す復興感・Codeの出現割合からは、復興感割合が低いほど、仕事、健康に関するCodeも多く出現しており、経済的な支援と併せ、復興初期段階における公的な健康管理の増進を図っていくことの重要性が示唆された。また、図7に示す現在居住別では、修復等を行ったうえで被災前と同じ場所に居住する自宅居住者は、応急・借上仮設住宅居住者よりも高い生活復興感が見られた反面、震災前の生活道路環境の変化が要因と考えられる「交通Code」の不満の出現割合が高いことが明らかになった。これは、同時期において域内の生活用品購買店を含む巡回ルートバス等が開通していたものの、図5にみられる通り、30歳代以下の突出を除いても、全年齢層において移動障壁に係る不満の表明が特徴となっていることが明らかになった。

## 7. 結論と課題

本研究では、復興初期段階にあたる発災から2年後の2013年調査において得られた自由記述内容をもとに、分析に際しての採用品詞の議論を踏まえて、第一段階ではテキストデータ全体の概観の把握を行い、次いで、第二段階ではコーディングによる出現割合の数値解析を行った。明らかになった点は以下の通りである。

- 1) テキストの分析に際しては、あらかじめ品詞別に採否の検討を行い、本研究では、「名詞」「形容動詞」「形容詞」および「サ変名詞」の4品詞を採用した。
- 2) 自由記述内容の全体構造では、生活再建の基盤となる住宅関連の用語が頻出したほか、復興感や被害状況と関連した特徴語の出現が見られた。
- 3) 用語の係り受け状況についてKWIC分析を行い、主な支援・要望、不満の構成要素として「仕事」「住宅」「健康」「交通」の関連用語の抽出を行った。
- 4) 関連語分析において、3)で抽出した語と関連する用語をCode化し、語単体での位相関係ではなく、同語を含む「構成概念」と各属性との関連についてコーディング分析を行った。本分析により、出現数の多寡と全体における出現割合を併せて検討を行い、健康については、加齢に伴い、また、復興感が低いほどその出現割合が高いことが明らかになった。

本研究は、同一地域を対象に、本年度までに6年間の連続・継続調査を実施しており、今後においては、パネルデータを構築したうえで、記述内容の変化やその要因等についても併せて検討していくことが課題である。

## 謝辞

本研究は、科学研究費補助金(B)24300322「東日本大震災の被災者の復興感の変遷と被災地の復興過程の対応に関する研究(研究代表者:中林一樹)」および科研費(B)15H02962「津波被災者の生活再建と復興感の時系列分析から見た復興体系に関する研究(研究代表者:中林一樹)」による研究成果の一部である。記して厚く御礼を申し上げます。

## 参考文献

- 中林一樹・土屋依子・小田切利栄(2014):東日本大震災津波被災者の被災状況と被災2年後の生活復興の現状—津波被災地の生活復興に関する2013年調査その1—,日本建築学会学術講演梗概集(神戸大学)