

# 災害報道からの個人属性別の状況の抽出 ～東日本大震災後の全国紙と地方紙のテキストマイニングから～

## A Study of Circumstances of Disaster Victims According to Multiple Attributes Using Text Mining Method for Newspaper Articles Related to The 2011 Great East Japan Earthquake

藤森 崇浩<sup>1</sup>, 小山 真紀<sup>2</sup>, 清野 純史<sup>3</sup>

Takahiro FUJIMORI<sup>1</sup>, Maki KOYAMA<sup>2</sup> and Junji KIYONO<sup>3</sup>

<sup>1</sup>京都大学 工学研究科 都市社会工学専攻

Department of Urban Management, Graduate School of Engineering, Kyoto University

<sup>2</sup>京都大学 医学研究科 人間健康科学系専攻

Human Health Sciences, Graduate School of Medicine, Kyoto University

<sup>3</sup>京都大学 地球環境学堂

Graduate School of Global Environmental Studies, Kyoto University

When a natural disaster occurs, victims' information needs vary with their personal attributes. (e.g. age, gender, occupation and so on), and the needs also vary with the time passed from the disaster. Thus, it is required to make a system that enables victims to easily satisfy their information needs considering their personal attributes and time phase. As a fundamental study before making such a system, we did time series analysis of circumstances of victims by applying text mining method for newspaper articles about the 2011 Great East Japan Earthquake. As a result, we clarified that it is possible to find the time series variation of correlation between some major personal attributes and provided information from newspaper articles.

**Keywords:** The 2011 Great East Japan Earthquake, Personal Attributes, Text Mining, Newspaper Articles

### 1. はじめに

災害が起こった際の被災者の置かれる状況は、災害が起こってからのフェーズや、個人属性によって様々である。例えば、災害直後は災害の規模や被害状況、ライフラインの復旧度合などの情報ニーズがあるが、ある人にとってはそれらの情報のうち自分の居住地域を対象とした情報のみが必要となり、また別の人では家族がその時滞在している場所の情報も必要になるなどである。さらに被災から時間が経過するにつれて置かれる状況にも幅が出てくる。慢性疾患を抱えている人や高齢者にとっては病院や薬のある場所の情報、仕事を失った人にとっては失業対策の情報、子供のいる家庭にとっては学校の再開の情報がそれぞれに必要ななど、一口に「被災者」といっても様々な状況に置かれていることが、アンケート調査から明らかになっている<sup>1)</sup>。

しかしながら、現状の災害情報提供は被災者一般を対象としたものが多く、自分にとって必要な情報を入手するにはそれらの情報を網羅したうえで自ら取捨選択する必要がある。このような状況を踏まえると、個人属性や被災からのタイムフェーズを考慮し、被災者個人が必要

とする情報に容易に到達できるようにするための支援システムが求められると考えられる。

本研究では、そのための基礎的研究として、震災時どのような属性の人物がどのような状況に置かれていたかを明らかにすることを目的として、2011年東日本大震災後から2011年末までの朝日新聞記事<sup>2)</sup>および河北新報(宮城県地方紙)<sup>3)</sup>のテキストマイニングを通じて、種々の個人属性ごとに置かれた状況の時系列評価を行った。

今回、新聞記事をデータとして用いた理由としては、扱う情報の種類が広範であり、多様な個人属性と情報との相関を分析することが可能であること、また時系列変化を分析することが容易であることが挙げられる。このような理由から、新聞記事のテキストマイニングからニーズを抽出する研究が行われた事例<sup>4)</sup>があるほか、災害報道についても、特徴的な単語に注目した時系列展開<sup>5)</sup>、地域比較<sup>7)</sup>が行われてきたが、個人属性ごとの分析を行う試みはなされていない。また、本研究では全国紙と地方紙における報道内容の違いを分析することも狙いの一つとする。

## 2. データと手法

2011年3月12日から12月31日までの朝日新聞の本紙記事(地方版を除く)および河北新報の全文記事の中から、「地震」、「震災」、「津波」、「被災」、「避難」、「原発」、「原子力発電所」、「計画停電」、「節電」、「帰宅困難者」および「帰宅難民」の語を一つでも含む記事を分析対象として抽出した。

その上で、被災地の復興フェーズは災害後の初期ほど短期間で変化していくこと<sup>8)</sup>を考慮し、表1に示すようなフェーズにより新聞記事を区分し、各フェーズにおける記事内容の分析を行うこととした。各フェーズで得られた震災関連記事数を表2に示す。括弧書きは全記事に対する、抽出した震災関連記事の割合である。

次に、新聞記事を記事内容によって分類するため、「コーディングルール」を作成して各記事にコードを付与した。例えば「高齢者」というコードは「高齢」、「高齢者」、「お年寄り」、「老人」などの語を含む記事に付与される。個人属性に関するコードと、情報・物資に関するコードの両方を用意し(表3)、それらの出現率の変化や相関の分析を行った。

個人属性については、「年齢」、「職業」、および「要援護者」の観点からコードを作成した。「性別」や「家族構成」といった個人属性も重要な要素であると考えられるが、新聞記事からの抽出は困難であるためコード設定は行わなかった。職業については、沿岸部の被害が大きかったため、「農家」および「漁師」を設定し、要援護者は内閣府による災害時要援護者の区分<sup>9)</sup>を参考にして設定した。「一人暮らし」は独居老人や一人暮らしの学生を包括するコードとして設定している。

情報・物資コードに関しては、既往のアンケート調査項目に、筆者らが被災地にて行った聞き取り調査で被災者が話題にしていた項目(「心のケア」および「風呂」)を加えて設定した。無形のものとしての「情報」、有形のものとしての「物資」という括りでのコード設定となっている。

コーディングルールの作成時には、暫定的に作成したコーディングルールで記事にコードを付与した上で、各コードが付与された記事内での頻出語を分析するといった手法で、類義語の見落としが無いように配慮した。一つの記事中に、複数のコードに該当する語が出現した場合は、一つの記事に複数のコードが付与されるものとする。

テキストマイニングツールにはKH Coder(Ver2.b.30b)<sup>10)</sup>、形態素解析にはMeCab(Ver.0.996)<sup>11)</sup>を用いた。形態素解析のための辞書としてはIPA辞書と2012年10月21日時点での日本語版Wikipedia<sup>12)</sup>の項目リストから作成したユーザー辞書を用いた。

表1 区切りに用いたタイムフェーズ

フェーズ	期間	経過時間
1	3/12~3/14	~3日後
2	3/15~3/18	~1週間後
3	3/19~3/31	~3月末
4	4/1~4/11	~1ヶ月後
5	4/12~5/11	~2ヶ月後
6	5/12~8/11	~5ヶ月後
7	8/12~12/31	~2011年末

表2 震災関連記事数

フェーズ	朝日新聞	河北新報
1	244 (68.3%)	277 (86.0%)
2	559 (79.9%)	566 (91.3%)
3	1,575 (64.0%)	1,710 (86.0%)
4	1,291 (59.5%)	1,633 (81.0%)
5	2,821 (51.8%)	3,970 (70.8%)
6	6,755 (37.2%)	11,185 (55.5%)
7	6,667 (24.4%)	13,833 (44.6%)

表3 用いたコード

選定基準	個人属性コード
年齢	乳幼児, 子ども, 学生, 高齢者
職業	会社員, 主婦, 農家, 漁師
要援護者	傷病者, 障害者, 妊婦, 一人暮らし, 外国人

選定基準	情報・物資コード
情報	ライフライン, 交通, 安否, 避難, 医療, 福祉, 原発, 心のケア, 復興, 仮設住宅, 失業・就職, 義援金
物資	食料, 防寒具, 衛生用品, ガソリン, 風呂

## 3. 個人属性コードの出現率

まず新聞記事において取り上げられやすい個人属性や、それらの全国紙と地方紙での取り上げられ方の差異を分析するため、個人属性に関する各コードの出現率(各コードが出現した記事数の、全記事数に対する割合)の分析、および2紙での各コード出現数について統計ソフトR<sup>13)</sup>を用いたカイ二乗検定を行った。結果を表4に示す。

「子ども」、「主婦」、「学生」および「高齢者」などの一般的な個人属性については出現率が高いが、「障害者」、「一人暮らし」、「妊婦」、「乳幼児」および「外国人」などの属性に関しては出現率が低く、新聞紙上ではあまり報道されていないことが明らかになった。

また、「学生」、「漁師」、「会社員」および「乳幼児」以外の個人属性に関しては、2紙の間の出現数に1%水準で有意な差が見られた。総じて朝日新聞の方が出現率が高い傾向にあるが、「農家」および「障害者」に関しては河北新報の方が出現率が高く、朝日新聞よりも重きを置いて報道していたという特徴が見てとれる。

表4 個人属性コードの出現率

個人属性コード	朝日新聞	河北新報	p	
子ども	13.7%	12.8%	3.1E-03	**
主婦	8.0%	6.1%	2.2E-16	**
学生	7.4%	7.5%	0.74	n.s.
高齢者	4.6%	3.9%	5.4E-05	**
漁師	3.9%	4.0%	0.92	n.s.
農家	2.9%	4.2%	7.9E-13	**
会社員	2.1%	2.2%	0.25	n.s.
傷病者	1.9%	1.2%	1.9E-10	**
外国人	1.9%	0.9%	2.2E-16	**
乳幼児	1.5%	1.2%	0.011	*
妊婦	0.8%	0.4%	9.2E-08	**
一人暮らし	0.7%	0.4%	8.5E-08	**
障害者	0.5%	0.9%	7.4E-07	**

n.s.:非有意, \*: p<0.05, \*\*: p<0.01

#### 4. 情報・物資コード出現率の時系列変化

次に、情報・物資に関するコードの出現率の時系列分析を行った。朝日新聞の結果を図1、河北新報の結果を図2に示す。

##### (1) フェーズ進行と共に減少傾向にある情報コード群

図1(a)および図2(a)に示すコード群は、フェーズ進行と共に出現率が減少する傾向にあるものである。

- ・ライフライン：停電、断水状況および計画停電の予定などを報じる記事に与えられる。朝日新聞、河北新報ともにフェーズ1において出現率が最大であり、フェーズ進行とともに減少していき、1ヶ月から2ヶ月後には収束する。両紙の記事内容を参照すると、朝日新聞では計画停電の話題が、河北新報では被災地のライフライン状況を報じる話題がそれぞれ多く見られた。
- ・交通：公共交通機関の運行状況を報じる記事に与えられる。朝日新聞ではライフラインと同様にフェーズ1で最大となり、その後減少していく。フェーズ4からはあまり出現率が変わらなくなる。河北新報ではピークがフェーズ2にあり、その後減少していく。これは、朝日新聞では首都圏で問題となった帰宅困難者の問題をより多く報じた後、交通インフラが復旧していったため、河北新報では鉄道の復旧していないフェーズ

2において、被災地内のバス移動のような現地の交通状況を報じていたためと見られる。

- ・安否：生存者、行方不明者および遺体安置などに関する安否情報を報じる記事に与えられる。推移の傾向としてはライフラインと同様である。津波被害の影響により安否確認が長引いたこともありフェーズ5までは出現率が10%を切ることが無かった。
- ・避難：一時避難としての避難指示と、二次避難としての避難所についての両方の記事に与えられる。朝日新聞ではフェーズ2において一旦落ち込み、避難所生活および原発事故関連の避難勧告の話題が挙がるフェーズ3で再び増加しているが、基本的には減少傾向にある。避難所生活から仮設住宅生活へと移行するフェーズ6、7において出現率が減少する。
- ・医療：受診可能な医療機関や医薬品に関する情報など、医療に関する記事に与えられる。朝日新聞ではフェーズ1、河北新報ではフェーズ2までがピークであり、その後減少していき、1ヶ月ほどで出現率は低い水準になる。
- ・福祉：老人ホーム、デイサービスおよびバリアフリーといった福祉関連の語を含む記事に与えられる。朝日新聞では漸減傾向、河北新報では漸増傾向にある。これは、地方紙では終盤のフェーズに仮設住宅での福祉サービスについて報じているためである。

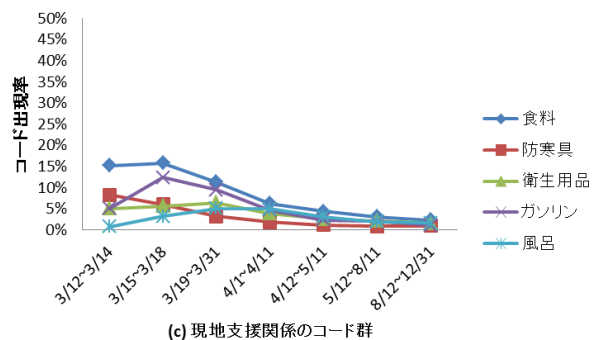
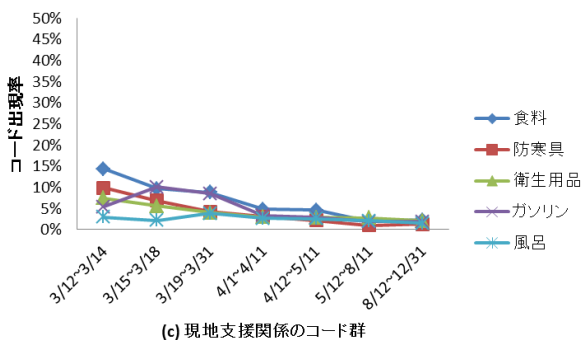
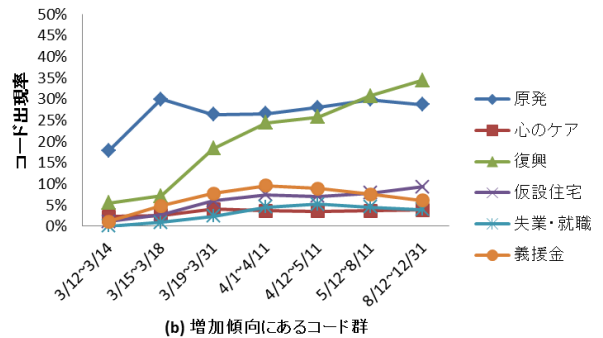
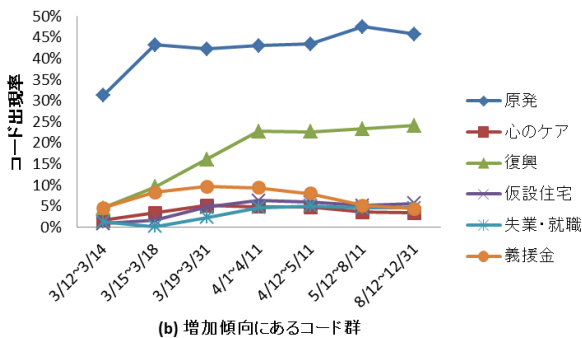
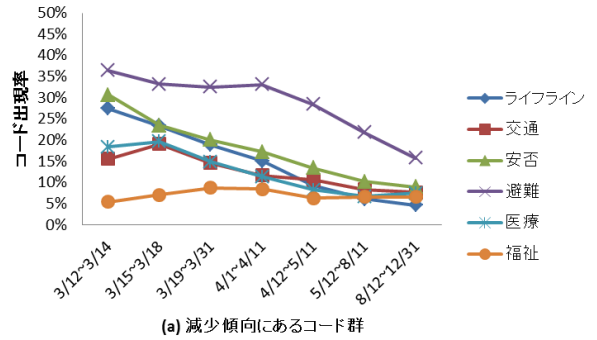
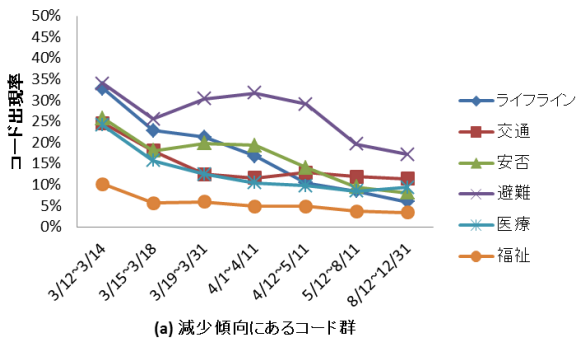


図1 朝日新聞・コード出現率

図2 河北新報・コード出現率

## (2) フェーズ進行とともに増加傾向にある情報コード群

図 1(b)および図 2(b)に示すコード群は、フェーズ進行と共に増加する傾向にあるものである。

- ・原発：原発事故や、放射線量などに関する記事に与えられる。両紙ともにフェーズ 2 から一貫して高い出現率を維持している。出現率自体は朝日新聞の方が高い。
- ・心のケア：ストレスや PTSD に関する語を含む記事に与えられる。出現率自体は多くなく、フェーズ 3 にいたるまで微増し、その後はほぼ横ばいとなっている。
- ・復興：「復興」および「復興計画」といった語を含む記事に与えられる。朝日新聞、河北新報ともにフェーズ 3 から出現率が大きく伸びる。また、河北新報の方が朝日新聞よりも出現率が高いということも分かる。
- ・仮設住宅：仮設住宅に関する語を含む記事に与えられる。朝日新聞ではフェーズ 3 まで増加し、その後横ばいになる。河北新報ではフェーズ 3 以降も増加を続ける。「復興」と同様に河北新報の方が、出現率が高い結果となった。
- ・失業・就職：失業や就職など、雇用に関する語を含む記事に与えられる。朝日新聞、河北新報ともにフェーズ 4 まで出現率が増加し、その後横ばいとなる。
- ・義援金：義援金や募金に関する語を含む記事に与えられる。朝日新聞はフェーズ 3、河北新報はフェーズ 4 にそれぞれピークを迎える。各紙での初出記事の見出しを見てみると、朝日新聞が「募金活動スタート 東日本大震災の被災地支援へ」、河北新報が「東日本大震災／ボランティア自粛呼び掛け／村井宮城県知事」という義援金による現地支援を呼びかけるものとなっており、それぞれの視点の違いが表れている。

## (3) 物資など現地支援に関するコード群

図 1(c)および図 2(c)に示すコード群は、物資や風呂といった現地支援に関するものである。

- ・食料：食料の配布や給水、炊き出しといった語を含む記事に与えられる。朝日新聞、河北新報ともにフェーズ進行とともに減少していくのは同様であるが、減少が始まるのは朝日新聞ではフェーズ 2、河北新報ではフェーズ 3 となっていること、全体を通してのコード出現率は河北新報の方が高いことが相違点である。
- ・防寒具：毛布、ストーブおよびカイロといった防寒具に関する語を含む記事に与えられる。こちらは朝日新聞、河北新報ともにフェーズ 2 から減少が始まっている。なお、震災が発生した季節の特性上、避難所において「防寒具」は切迫性をもって必要とされていたが、「夏物」のニーズが発生する頃には仮設住宅への移転、買い物ができる環境の整備が進み、記事にはあまり出現しなかったため、コード設定は行わなかった。
- ・衛生用品：紙おむつ、タオル、トイレトーパーおよび生理用品といった語を含む記事に与えられる。朝日新聞では、「食料」や「防寒具」と同様の推移を示すが、河北新報ではピークがフェーズ 3 に表れている。
- ・ガソリン：ガソリンや給油に関する語を含む記事に与えられる。両紙ともにフェーズ 2 でピークを迎えるが、河北新報の方がより出現率の高いピークとなっている。
- ・風呂：風呂やシャワーに関する語を含む記事に与えられる。朝日新聞では、フェーズ 1 から 3 で相対的に多く出現する。河北新報ではフェーズ 3 および 4 にピークが見られる。

## (4) 考察

情報・物資コード出現率の時系列変化を分析することにより、新聞紙上で提供された状況の量的変化が得られた。朝日新聞では「原発」のような全国的な話題の割合が高く、河北新報では「復興」、「仮設住宅」および「食料」といった、地域住民の関心の高い話題の割合が高いことも明らかになった。また、「義援金」のコード出現率のピークが、朝日新聞においてより早く現れたのは、朝日新聞は支援を提供する側としての視点も含まれているため、義援金の募集時期からピークを迎えたのに対し、河北新報は支援を受け取る側の立場として記事が書かれているため、被災者が義援金を受け取ることができた時期にピークを迎えたという差異が表れたと考えられる。対照的に「食料」については、朝日新聞では被災地でのニーズがまだあると考えられる段階より先に出現率が低くなるという結果を示した。

以上より、一般的に認識されていた、「地方紙は全国紙よりもローカルな情報を報じている」ということを、「原発」、「復興」、「仮設住宅」および「食料」といったコードで定量的に確認することができた。また、「義援金」の例に見られるように、報道の質の面でも、支援を提供する側、される側としての差異が確認された。災害情報を適切なタイミングで提供することを考える上では、必要とされる情報種別の時間特性を考慮することが重要であり、本章で得られたような長期スパンでの量的変化を知ることは有意義であろう。

## 5. Jaccard係数によるコードの相関分析

次に個人属性に関するコードと情報・物資に関するコードの相関の分析を行った。相関の指標としては、Jaccard係数を用いた。一般にJaccard係数は共起の度合を示すが、個人属性コードと情報・物資コードが共起する(同じ記事に付与される)ということは、同じ話題の中で報道されているということになるため、2つのコードの相関が高いとみなし、相関指標としてJaccard係数を採用した。2つのコード X, Y 間のJaccard係数は式(1)により表される。

$$Jaccard_{X,Y} = \frac{|X \cap Y|}{|X \cup Y|} \quad (1)$$

ここで、 $|X \cap Y|$ はコード X および Y の両方が与えられた記事数、 $|X \cup Y|$ はコード X および Y のうちのどちらか一つでも与えられた記事数を表す。Jaccard係数が大きいほど 2 つのコードが同時に付与された記事の割合が高く、コード間に相関があるという指標になる。この指標を用いて、様々な個人属性コードと、情報・物資コードの間の相関を時系列で分析した。

ここではページ数の関係上、結果の一例として表 5、表 6 に「乳幼児」のコードと、表中に示す情報・物資コード群との間のJaccard係数の時系列変化を示す。Jaccard係数の値は、各フェーズ間の比較のため、各フェーズでの平均が 1.00 となるように正規化し、各フェーズでの値の上位 3 項目については網掛けして示している。特定のコードの記事に限定したため、個人属性コードとフェーズによっては十分な記事数が得られないこともあり、その場合は正しい相関関係を得ることが難しく、ここでは記事数の多くなるフェーズ 3 以降の結果を示している。なお、他の個人属性における分析結果は 7 章にてまとめて解説することとする。



表5 朝日新聞「乳幼児」Jaccard 係数

		タイムフェーズ				
		3	4	5	6	7
情報・物資コード	ライフライン	1.19	0.72	0.89	0.42	0.61
	交通	0.89	<b>1.47</b>	1.02	0.83	1.10
	安否	0.57	1.00	1.00	0.57	0.94
	避難	0.91	1.02	0.89	1.20	0.94
	医療	<b>1.72</b>	1.08	<b>1.57</b>	<b>1.77</b>	<b>2.04</b>
	福祉	1.15	<b>1.70</b>	1.31	1.15	<b>1.71</b>
	原発	0.78	0.86	0.38	0.68	0.66
	心のケア	0.96	1.06	0.64	1.36	<b>2.15</b>
	復興	0.37	0.49	0.47	0.37	0.44
	仮設住宅	1.21	1.06	1.46	1.30	0.77
	失業・就職	0.13	0.63	0.49	0.99	0.61
	義援金	0.56	0.78	0.86	0.52	0.44
	食料	<b>1.83</b>	<b>3.04</b>	<b>1.84</b>	<b>1.67</b>	1.32
	防寒具	<b>1.43</b>	0.53	<b>1.64</b>	0.78	1.16
	衛生用品	1.34	0.55	1.40	<b>1.93</b>	0.83
ガソリン	0.70	0.72	0.11	0.16	0.61	
風呂	1.25	0.29	1.04	1.30	0.72	

表6 河北新報「乳幼児」Jaccard 係数

		タイムフェーズ				
		3	4	5	6	7
情報・物資コード	ライフライン	0.51	0.98	0.14	0.99	0.91
	交通	0.63	1.22	0.52	0.99	0.82
	安否	0.63	0.54	0.37	0.87	0.64
	避難	0.72	1.15	0.89	1.22	1.10
	医療	1.07	0.68	1.18	1.11	<b>2.28</b>
	福祉	0.79	1.00	0.92	0.76	1.10
	原発	0.76	1.25	0.55	0.64	0.82
	心のケア	1.37	0.78	<b>1.41</b>	<b>1.75</b>	1.19
	復興	0.59	0.51	0.46	0.52	0.37
	仮設住宅	<b>1.70</b>	<b>1.98</b>	0.92	<b>1.69</b>	0.64
	失業・就職	1.11	0.44	0.32	0.17	0.27
	義援金	0.55	0.78	0.63	0.41	0.37
	食料	<b>1.72</b>	1.69	0.95	0.93	<b>2.92</b>
	防寒具	0.82	0.00	0.49	<b>2.10</b>	1.19
	衛生用品	<b>1.93</b>	<b>2.08</b>	<b>3.85</b>	1.22	0.55
ガソリン	0.58	0.17	0.00	0.35	0.18	
風呂	1.53	<b>1.74</b>	<b>3.42</b>	1.28	<b>1.65</b>	

(1)朝日新聞による結果

「医療」および「福祉」とは一貫して高い相関が得られた。「食料」、「防寒具」および「衛生用品」などのコードとはフェーズ 6 まで相関が高い状況が続いている。「心のケア」に関してはフェーズ 6 以降から相関が高くなっている。仮設住宅の入居時期にあたるフェーズ 5 から 6 では、「仮設住宅」との相関がやや高まった。

(2)河北新報による結果

朝日新聞ではそれほど高い相関が見られなかった「風呂」との高い相関が見られた。「医療」、「心のケア」「仮設住宅」、「食料」、「防寒具」および「衛生用品」などと相関の見られた点は朝日新聞と同様である。ただし、朝日新聞では「仮設住宅」はフェーズ 5 から、「心のケア」はフェーズ 6 から高い相関が表れていったのに対し、河北新報では「仮設住宅」はフェーズ 3 から、「心のケア」はフェーズ 5 からと、より早いフェーズから高い相関が表れている。

(3)考察

高い相関が得られた情報・物資コードを含む記事について、原文の内容を調査し、相関の原因を考察した。

「食料」、「防寒具」および「衛生用品」との相関が得られたのは、乳幼児にとって粉ミルク、毛布および紙おむつといった物資が必要不可欠であり、そのことが報じられているためである。「医療」に関しては、出産や医療機関による放射線被爆のチェック、感染症などに関連して相関が表れている。「心のケア」に関しては、乳幼児自体のストレスというよりは、乳幼児を持つ母親が、避難所の周囲の人への遠慮や、我が子の被爆に対する不安などからストレスを抱えている様子を報じる記事が多かったためである。「仮設住宅」に関しては、高齢者、乳幼児、妊婦および障害者のいる世帯の優先的入居がなされたためである。「風呂」に関しては、河北新報は生活関連情報として、入浴可能な施設を掲載しており、乳幼児の入浴料は無料であることや、助産婦による乳幼児の入浴サービスなどが報じられていたためである。乳幼児の被爆に関する記事があったにも関わらず「原発」との相関が見られなかった原因としては、「原発」コードが付与された記事数(全体で 5,008)に対し、「乳幼児」コードが付与された記事数(全体で 289)がかなり少なかったために高い Jaccard 係数の値が得られなかったためであると考えられる。

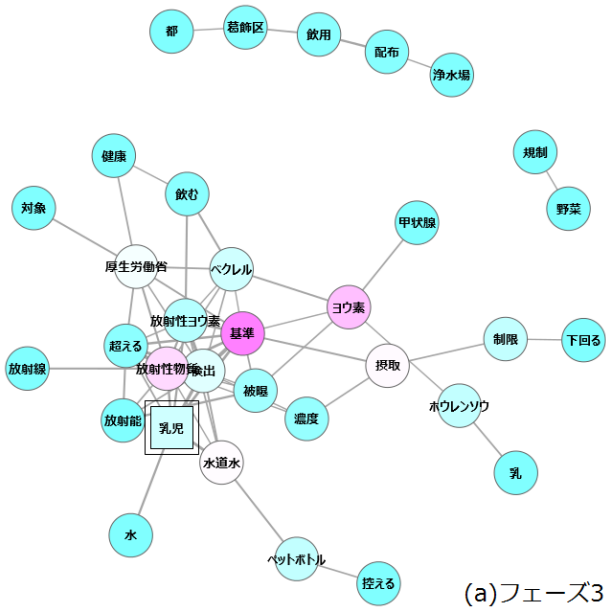
まとめとして、①河北新報は朝日新聞ではあまり取り上げなかった乳幼児の入浴サービスを取り上げるなど、被災者に密着した状況に関する報道を行っていた。②乳幼児のいる世帯の仮設住宅への優先入居を、仮設住宅の建設段階であるフェーズ 3 から取り上げていた。③乳幼児を持つ母のストレスを朝日新聞よりも早く取り上げていたなど、被災者のリアルタイムの状況をより素早く汲み取っていると考えられる。

6. 個人属性コードの共起ネットワーク分析

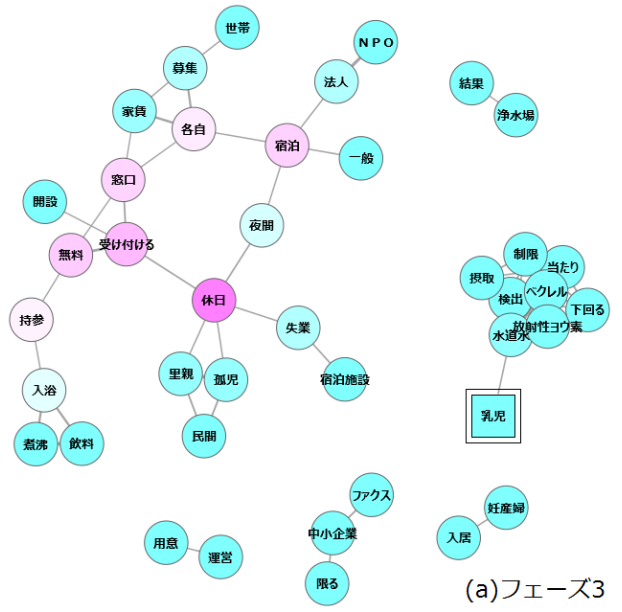
前章において表 5 および表 6 に示した、Jaccard 係数による相関分析では、2 つのコードが同じ記事に同時に付与される割合を分析しているため、記事内容までは評価することができていなかった。そこで、本章では共起ネットワークを用いて記事内容まで立ち入った検討を行うこととする。あるコードに関する共起ネットワークを描くと、そのコードが付与された記事に含まれる語の中で、関連の強い語同士を線で結んだネットワークを作成することができる。これにより単語レベルでの内容分析および文脈の類推を行うことができる。

ここでは、結果の一例として前章と同じく、「乳幼児」コードが付与された記事を例にとって紹介することにする。朝日新聞での結果を図 3、河北新報での結果を図 4 に示す。全ての共起ネットワークは、語と語の共起関係(Jaccard 係数で計算)の上位 60 位までをエッジとして表すように描いている。「乳幼児」コードに含まれる語は二重の正方形で描かれ、その他の語は円で描かれており、エッジのカラーは媒介中心性を示している。(水色・白・ピンクの順に媒介中心性が高くなる。)

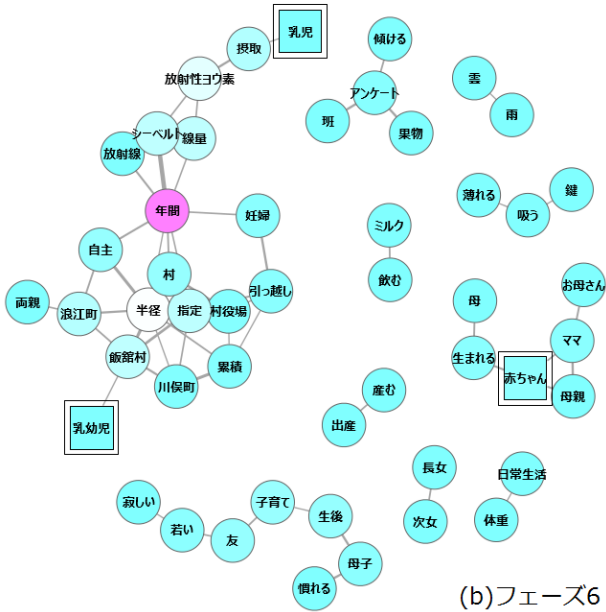
今回の共起ネットワーク分析例において、フェーズ 3、6、および 7 の 3 フェーズを用いた理由としては、この 3 フェーズでは「乳幼児」の記事数が確保でき、比較可能な形で共起ネットワークを描くことができたためである。



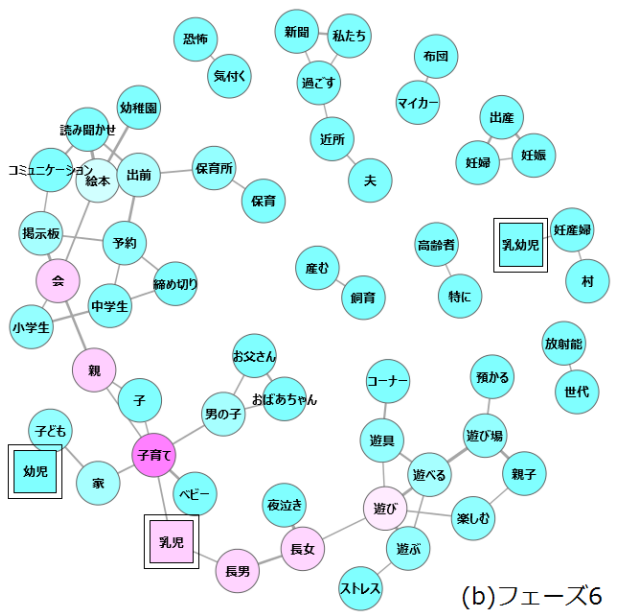
(a)フェーズ3



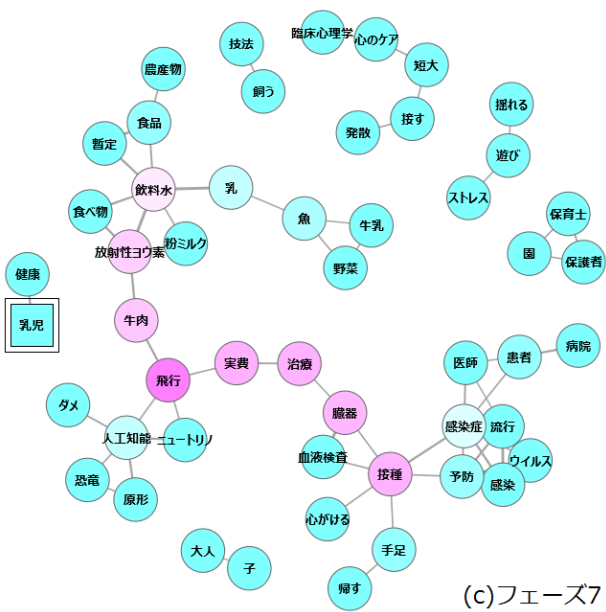
(a)フェーズ3



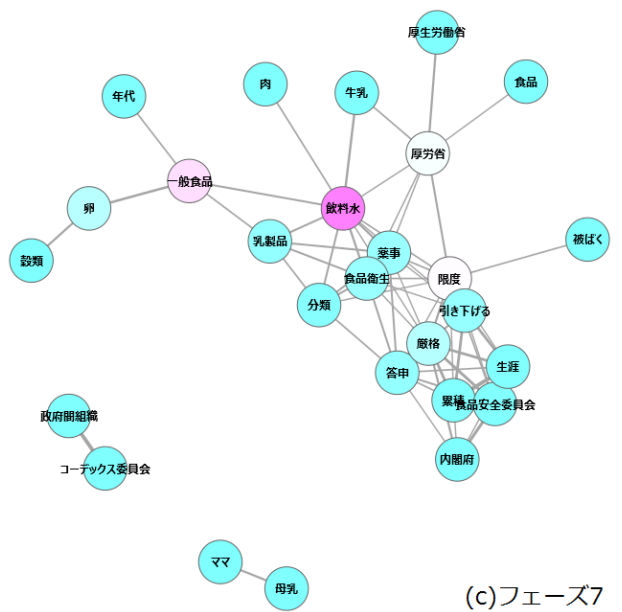
(b)フェーズ6



(b)フェーズ6



(c)フェーズ7



(c)フェーズ7


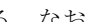
図3 朝日新聞「乳幼児」共起ネット

図4 河北新報「乳幼児」共起ネット

図 3 および図 4 より、朝日新聞、河北新報ともに「放射性ヨウ素」、「厚生労働省」および「飲料水」といった語が頻出していることが分かる。これらは、厚生労働省が食品衛生法に基き放射性物質で汚染された食品の出荷や販売を規制するため、飲料水、乳製品、野菜および肉などに含まれる放射性ヨウ素やセシウムの基準を設け、特に乳幼児用の食品は基準を厳しくしたこと<sup>14)</sup>を報じるものである。これより、前章の表 5 および表 6 には表れていなかったものの、「乳幼児」と「原発」の間には相関が強いということが明らかになった。

また、図 4 (b)では「ストレス」、「遊ぶ」、「遊具」および「コーナー」といった語が表れている。これらは、避難所で遊ぶことができずにストレスを抱えている乳幼児のために遊具コーナーを設置したという記事のものである。このような話題は河北新報でしか取り上げられておらず、ここからも地方紙においては地域に根差した話題も取り上げられているということが分かった。他の個人属性における分析結果は 5 章での結果とあわせて次章にてまとめて解説する。

## 7. 個人属性ごとの情報との相関まとめ

5 章の Jaccard 係数による相関分析について、全ての個人属性での分析結果を図 7 に示す。各フェーズにおいて Jaccard 係数の上位 3 位を、朝日新聞は (上段)、河北新報は (下段)で網掛けしている。なお、各個人属性コードが与えられた記事数が 10 に満たないフェーズについては、記事数が不十分として分析結果を省略(斜線)している。

### (1) 各個人属性ごとの結果

- ・乳幼児：出産の話題が多く報じられ、「医療」との相関が得られた。また、仮設住宅への優先入居や、避難所での乳幼児や母親のストレスの問題もあり「仮設住宅」および「心のケア」との相関も得られたが、河北新報の方が大きな相関が表れていた。また、河北新報では乳幼児の入浴可能施設・時間帯なども掲載され「風呂」との相関が得られた。また、共起ネットワークからは、内部被爆の話題に関連して「原発」との共起関係も得られた。
- ・子ども：早期は「安否」との相関が高かった。子どもを持つ家庭で子どもの安否を心配する声反映されていると考えられる。また、避難所での子どもの様子が報じられていたため、「避難」との相関が高まった。朝日新聞では、乳幼児と同様に「医療」、終盤には「原発」との相関が得られ、河北新報では子どものストレスを取り上げたため「心のケア」との相関も見られた。また 1 ヶ月後から相関が得られた「復興」は、未来の子どもたちのために復興を、という論調の記事が多かったためである。
- ・学生：「子ども」と相関の特徴が似通っており、早期は「安否」、「避難」および「心のケア」、1 ヶ月からは「復興」および、学生が行った募金活動のために「義援金」といったコードと相関が見られた。
- ・高齢者：ほぼ全フェーズ通して「医療」および「福祉」と相関が得られた。早期では「安否」および「避難」、仮設住宅の入居時期からは「仮設住宅」との相関が高く、河北新報では入浴を望むお年寄りの声を取り上げ、「風呂」との相関も得られた。
- ・会社員：初期において、帰宅困難時や通勤に関連して「交通」との相関とともに、家族の安否を心配する声から「安否」との相関も高かった。朝日新聞の中では、1 ヶ月後からは計画停電とそれに関連した鉄道ダイヤの変更を気にする会社員の声を取り上げられ「ライフライン」との相関が高まった。被災地では 1 ヶ月後も安否確認がとれないケースがあり、河北新報では「安否」との相関が長く続いた。
- ・主婦：家族の安否を心配する声や、避難所での主婦のコメントなどから「安否」および「避難」との相関が、また仮設住宅入居の時期になると「仮設住宅」との相関も高まった。朝日新聞では 1 ヶ月以降、病院や医薬品の入手場所などへの関心から「医療」との相関も得られた。
- ・農家：原発事故の風評被害への懸念や、厚生労働省が作物中に含まれる放射性物質の基準値を設けた記事から「原発」との相関が見られた。これに関連して、原発付近からの避難指示の話題も同時に掲載されたことから 1 ヶ月後から「避難」との相関も高まった。田畑に津波被害を受けた被災者が、国からの補償などの復興支援を求める記事が掲載されたため、「失業・就職」、「義援金」および「復興」との相関も得られた。
- ・漁師：津波で漁船を流されてしまった漁業従事者は、水揚げをすることすらできなかったため、「農家」のように「原発」との相関は高くなかった。失業対策や復興に向けた構想と関連した報道が多かったため、「復興」および「失業・就職」との相関が高まった。
- ・傷病者：当然ながら全フェーズを通して「医療」および「福祉」との相関が得られた。終盤になると長引く避難生活へのストレスから「心のケア」に関する話題も多くなったが、これも河北新報でより早いフェーズから相関が表れていた。共起ネットワークの分析では、初期フェーズにおいて、透析や人工呼吸器に電気が不可欠なことから「ライフライン」関係の単語との共起関係も得られた。
- ・障害者：通常の避難所では生活が困難な障害者は、福祉避難所に関連した報道が多く、「福祉」との相関が高かった。また、バリアフリータイプの仮設住宅を必要とするため、「仮設住宅」との相関も高かった。朝日新聞では障害者を支える NPO への義援金の話題が挙げられ、「義援金」との相関が得られた。また、共起ネットワーク分析からは、避難所での入浴に苦勞する様子が伺え、「風呂」との共起関係が得られた。
- ・妊婦：初期は、受診可能な産婦人科の情報が掲載され、「医療」との相関が高まった。その後、優先入居が可能な「仮設住宅」と相関が見られたほか、共起ネットワークからは、胎児への影響への懸念から「原発」との共起関係も見られるなど、「乳児」と相関の傾向はほぼ同じであった。
- ・一人暮らし：一人暮らし高齢者の記事が多く、「高齢者」の相関の特徴とやや似ている。それに加え、一人暮らし高齢者の避難所での孤立や、仮設住宅での孤独死が問題となり、「心のケア」との相関も高くなった。
- ・外国人：震災後の外国人観光客の減少を報じる記事が多く、残念ながら被災者としての外国人の状況把握は難しかった。しかしながら、共起ネットワークからは災害情報の翻訳ボランティアの助けを借りて情報収集をする様子も見取れた。

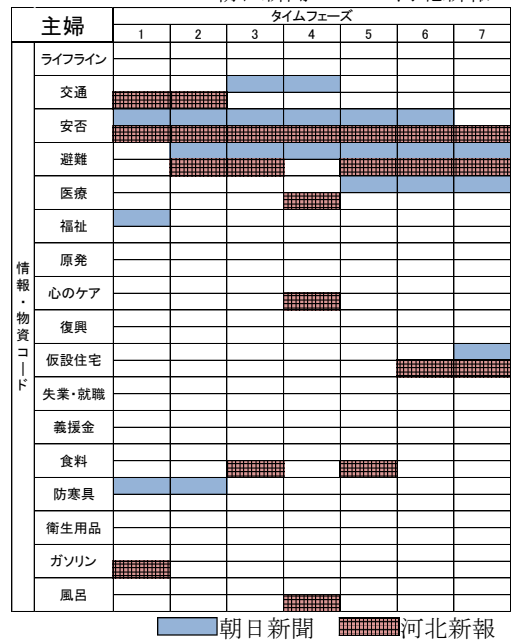
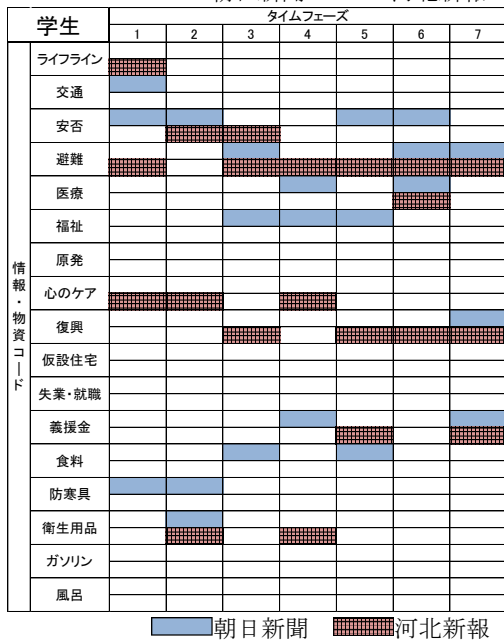
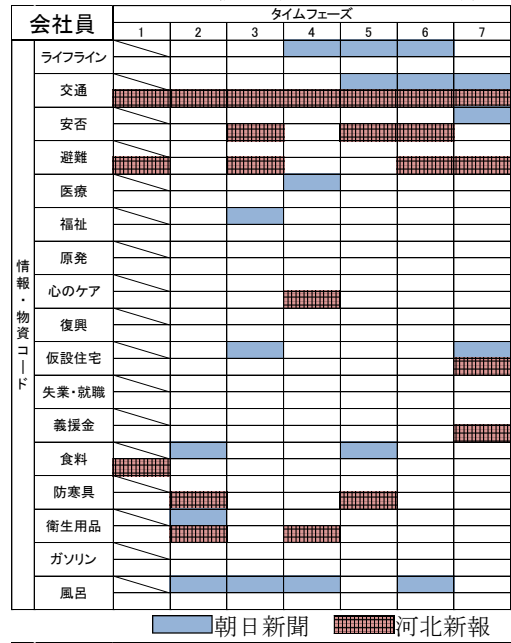
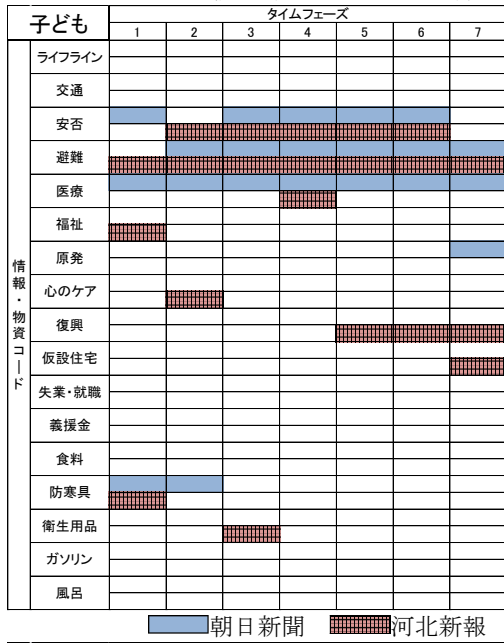
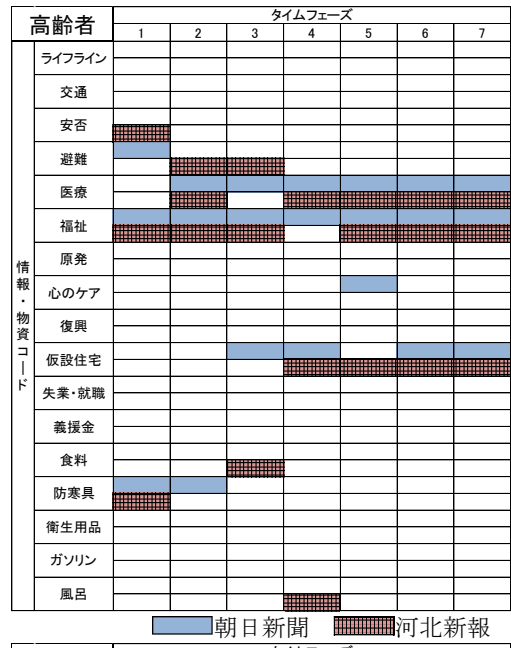
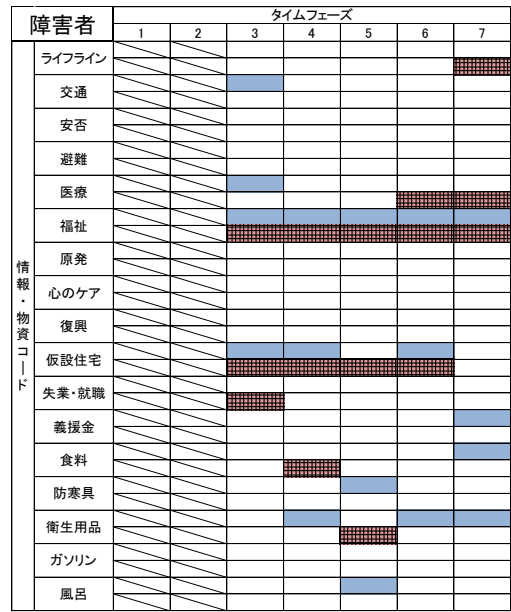


図7 個人属性別・相関の高い情報・物資コード





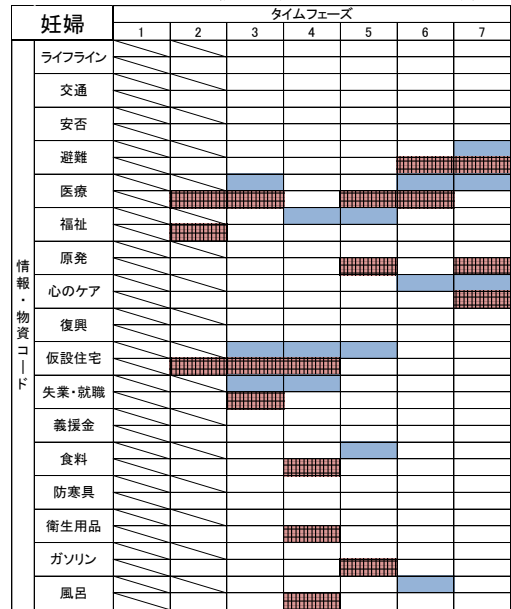
■朝日新聞 ■河北新報



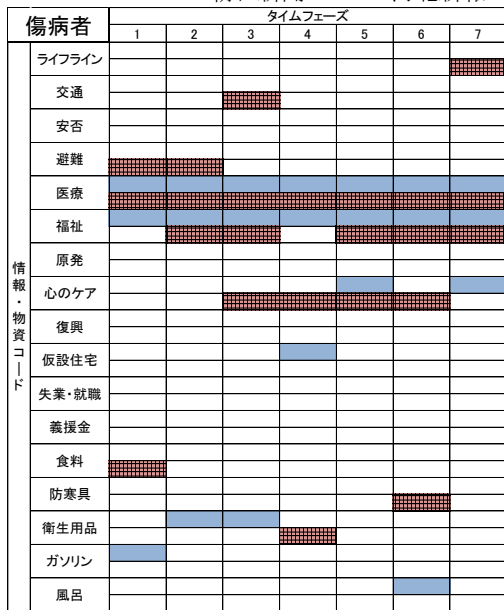
■朝日新聞 ■河北新報



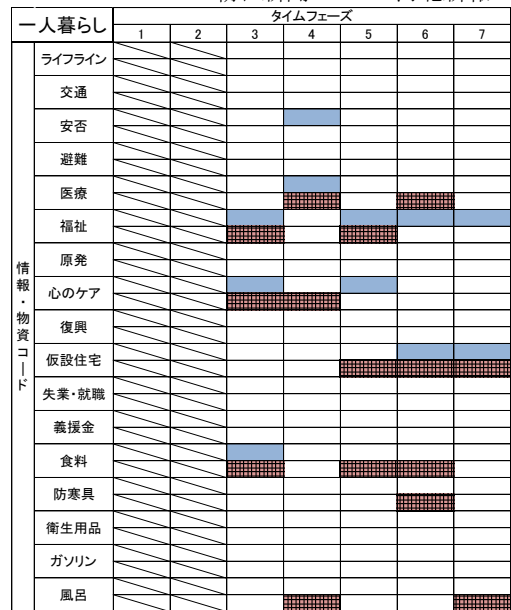
■朝日新聞 ■河北新報



■朝日新聞 ■河北新報



■朝日新聞 ■河北新報



■朝日新聞 ■河北新報

図7 個人属性別・相関の高い情報・物資コード(続き)

## (2) 両紙における差異

子どもの被曝に関連した「医療」および、計画停電に関連した「ライフライン」といった全国的な話題は朝日新聞のみで抽出されるようなケースが多かった。これに対し河北新報では、「乳幼児」や「高齢者」が「風呂」に関して困難を強いられていたことや、「心のケア」の必要な「子ども」, 「乳幼児」のいる家庭および「傷病者」がいたこと, 「農家」と「復興」との相関といった、朝日新聞からは抽出できなかった被災地内部での関心事項までも抽出することが可能であった。

## 8. 結論と今後の課題

本研究では、様々な個人属性の人々が震災後に置かれたフェーズごとの状況を新聞記事から抽出することができた。また、朝日新聞と河北新報を比較すると、下記のことことが分かった。

- ① 震災関連記事の割合は河北新報の方が高い。(表 2)
- ② 朝日新聞では、「原発」のような全国的な話題の割合が多く、河北新報では「復興」, 「仮設住宅」および「食料」といった被災地での話題の割合が多い。(図 1, 図 2)
- ③ 4 章の考察で言及したように、「義援金」の記事をとっても、朝日新聞と河北新報では、支援を提供する側とされる側という視点の違いが表れており、質的な観点からも差異が見られた。
- ④ Jaccard 係数と共起ネットワークによる相関分析の結果(7 章)からは、様々な個人属性の状況は両紙から抽出できたものの、「心のケア」, 「風呂」および「復興」といった被災地内部での関心事項の一部は地方紙である河北新報からしか抽出できない、もしくは全国紙に表れてくるまでにタイムラグがある場合があった。ただし、「子ども」と放射線医療に関連した「医療」との相関や、「会社員」の計画停電に関連した「ライフライン」との相関といった事項は朝日新聞からのみ抽出された。

①, ②より、全国紙と地方紙ではそれぞれ具体的にどのような話題に重きを置いて報道を行っていたのかということデータを上から定量的に確認することができた。また、③, ④より、両紙の視点の違いや、個人属性ごとに置かれた状況を具体的にどの程度まで抽出可能かということを確認することができた。

しかしながら、早期のフェーズでは、「農家」, 「障害者」, 「妊婦」および「一人暮らし」といった個人属性については分析に十分な記事数を確保することができず、また「外国人」被災者の状況把握も難しかったため、こうした個人属性についてはアンケート調査や市町村広報誌で提供された情報などの分析が必要になってくると考えられる。

今後は、さらに詳細な災害時情報ニーズモデルを構築した上で、その人にとって必要な情報を入手できるような仕組みの考案にも取り組んでいきたいと考えている。

## 謝辞

本研究は、JSPS 科研費(70462942)の助成を受けたものである。

## 参考文献

- 1) サーベイリサーチセンター：宮城県沿岸部における被災地アンケート, 2011, <http://www.surece.co.jp/src/research/area/20110620.html>.
- 2) 朝日新聞社：朝日新聞記事データ(学術・研究用), <http://www.asahi.com/information/cd/gakujutsu.html>
- 3) 河北新報社：河北新報データベース, <http://www.kahoku.co.jp/dbpr/tanaka/>
- 4) 杉浦政裕：テキストマイニングによる地域ニーズの分析手法の開発に関する研究—モンソーンアジア地域の水資源インフラ整備—, 日本建設情報総合センター研究助成事業成果報告会資料集, Vol.7, pp.1-55, 2009.
- 5) 佐藤翔輔, 今村文彦, 林春男：東日本大震災について報じられたウェブニュースコーパスの基礎的解析, 地域安全学会論文集, No. 15, pp. 303-311, 2011.
- 6) 沼田 宗純, 目黒 公郎：テキストデータのランニングスペクトル解析による福島民報の分析, 生産研究, Vol. 64 (2012) No. 6, pp. 927-934, 2012.
- 7) 松本誠, 鶴田和治：震災報道の“温度差”を見る 全国 4 紙の東西版紙面比較調査から, 兵庫地域研究, No.16, pp.67-80, 1999.
- 8) 小山真紀, 翠川三郎：市町村における地震時の意思決定支援に向けた災害応急対応モデル化の試み, 自然災害科学, Vol.25(1), pp.51-70, 2006.
- 9) 内閣府・災害時要援護者の避難対策に関する検討会：災害時要援護者の避難支援ガイドライン, 2008, [http://www8.cao.go.jp/shougai/suishin/kaikaku/s\\_kaigi/k\\_32/pdf/ref1.pdf](http://www8.cao.go.jp/shougai/suishin/kaikaku/s_kaigi/k_32/pdf/ref1.pdf).
- 10) 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, Vol.19(1), pp.101-115, 2004.
- 11) 工藤拓：MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>.
- 12) ウィキメディア財団：日本語版 Wikipedia, <http://ja.wikipedia.org/>.
- 13) 筑波大学：R CRAN ミラーサイト, <http://cran.md.tsukuba.ac.jp/>
- 14) 厚生労働省：食品中の放射性物質の対策と現状について, 2013, [http://www.mhlw.go.jp/shinsai\\_jouhou/dl/20131025-1.pdf](http://www.mhlw.go.jp/shinsai_jouhou/dl/20131025-1.pdf).

(原稿受付 2014.1.13)

(登載決定 2014.5.31)