

2016 年熊本地震における避難に関する新聞報道のテキストマイニング

岐阜大学大学院 正会員 ○加藤宏紀
 岐阜大学 正会員 能島暢呂・小山真紀
 中部電力株式会社 非会員 永井小雪里

1. はじめに

大規模地震後、家屋被害やライフラインの途絶等の理由から多数の避難者が発生する。避難所では避難者は不自由な生活を強いられ、避難所運営は行政機関の震災関連業務等にも支障をきたす。一連の影響は公助と自助・共助を組み合わせた対策（例えば、公共施設・家屋の耐震化、水の備蓄など）により軽減できる。これらの対策を促進するためにも、まずは避難に関する実態の把握が不可欠である。これまでに筆者らは、2016 年熊本地震における避難者の推移とその要因を災害対策本部等の資料に基づいて分析し、避難者数の推移に関する予測モデルを構築した¹⁾。本研究では、モデルの高精度化に向けた基礎資料の作成を目的として、新聞報道に着目し、2016 年熊本地震における地方紙と全国紙の避難関連記事の推移を分析する。

2. 使用データ・ツール

地方紙については熊本日日新聞の Web サイト²⁾から「平成 28 年熊本地震」特集に掲載されている記事を収集した。全国紙については朝日新聞の本紙面・地方面の記事³⁾を対象とした。ここで全国紙については熊本地震の関連記事を抽出するため、5 語（熊本地震・災害・避難・ライフライン・仮設住宅）のうち 1 語以上含む場合を対象とした。分析対象期間は、地方紙および全国紙（以降、2 紙）ともに 4 月 15 日～8 月 10 日である。

表 1 のように記事をそれぞれ 6 分割した。テキストマイニングツールとして KH Coder (Ver.2.00f)⁴⁾、形態素解析器として MeCab⁵⁾を使用した。形態素解析とは自然言語を意味のある最小の単位（形態素）に分割する作業のことである。この作業にあたり KH Coder の IPA 辞書⁴⁾及び 2018 年 1 月 15 日時点での日本語版 Wikipedia 項目リスト⁶⁾から作成したユーザー辞書を用いた。

表 1 各フェーズの対象期間と記事数

フェーズ	期間	地方紙 (記事数)	全国紙 (記事数)
1	4/15～4/22	102	538
2	4/23～4/30	91	389
3	5/01～5/15	128	407
4	5/16～6/15	226	753
5	6/16～7/15	217	556
6	7/16～8/10	194	393

3. 地方紙と全国紙を対象とした避難関連記事の推移

3.1 避難に関する頻出語の時系列推移

形態素解析の結果、地方紙の上位 150 位までの頻出語のうち避難に関連する語の順位について分析対象期間（以降、全期間）と各フェーズの結果を図 1 に示す。全期間では、各語とも上位 100 位以内に出現している。各フェーズに着目すると、「避難所」「避難者」はフェーズ 2 をピークに、出現順位が低下している。一方で「仮設住宅」「復興」の順位は時間経過と共に上昇している。図示は省略するが、全国紙においても「避難所」「避難者」「復興」については地方紙とほぼ同様の傾向が見られた。

3.2 自己組織化マップによる記事群のテーマの把握

自己組織化マップとは、教師なしニューラルネットワークアルゴリズムの一種で、高次元データを 2 次元平面上へ非線形写像するデータ解析方法である。KH Coder では、文書の頻出語を用いて出現パターンが似通った言葉ほど近くに布置される言葉のマップ⁴⁾を指す。地方紙を用いた自己組織化マップを作成し、さらにクラスター分析を行った結果を図 2 に示す。各クラスターに含まれる語群から「避難」「建物被害」「仮設住宅」「余震」などのテーマが読みとれる。

3.3 主要テーマの出現率の推移

各テーマを表す語群を以下のように設定し、その出現率（＝語群を 1 語以上含む記事数 / 各フェーズの記事数）の推移を図 3 に示す。

- ・建物被害：建物被害、全壊、半壊、一部損壊、建物 + （倒壊、被災、損傷、危険）

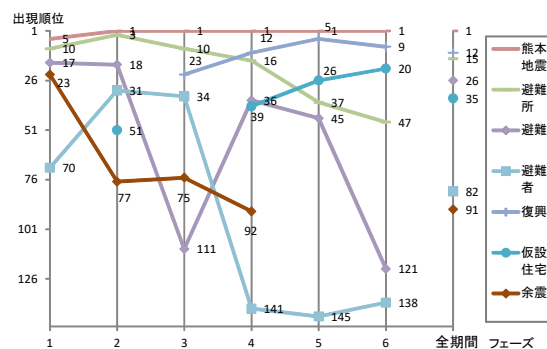


図 1 避難に関する頻出語の順位の推移（地方紙）

- ・震度：余震，震度，震源
- ・ライフライン被害：停電，断水，ガス+（漏れ，破損），停止+（電気，水，ガス），ライフライン+（被害，不安定，停止）
- ・避難：避難，避難者，避難所，車中泊，テント村，避難+（体育館，学校，テント）
- ・仮設住宅：仮設住宅，公営，仮設+（みなし，団地），住宅+（借り上げ）

（「，」印：or 条件，「+」印：and 条件を表す）

図3を見ると「震度」「ライフライン被害」「仮設住宅等」の出現率の増減は2紙とも類似している。一方、「避難」「建物被害」の出現率はフェーズ3以降で異なる。「避難」については東日本大震災に言及している記事が地方紙より全国紙で相対的に多いためであり、「建物被害」については地方紙で建物の復興を報じる内容が増加したためである。また，全期間を対象として2紙の各テーマの出現数においてカイ二乗検定⁷⁾を行ったところ，「避難」「建物被害」に関しては1%水準で有意な差が見られた（表2）。

3.4 主要テーマと関連語の推移

共起ネットワークを用いて，前節で設定したテーマと特定の語との共起関係の推移を考察する。これにより，各語の共起関係および文脈が推察できる。

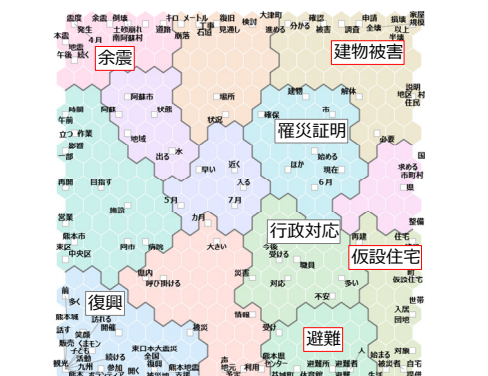


図2 頻出語を用いた自己組織化マップ（地方紙，四角枠は読み取れたテーマを表す）

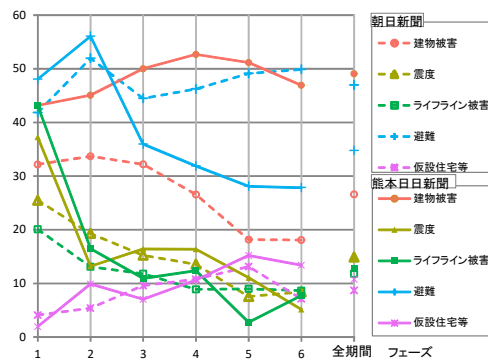


図3 2紙のテーマ出現率の推移

表2 2紙のテーマ出現数・出現率の比較
(n.s.: 非有意, *: $p < 0.05$, **: $p < 0.01$)

コード	熊本日日新聞	熊本日日新聞	熊本日日新聞	熊本日日新聞	熊本日日新聞	熊本日日新聞	熊本日日新聞	熊本日日新聞	熊本日日新聞
	出現数	非出現数	出現率	出現数	非出現数	出現率	カイ2乗値	p値	判定
建物被害	470	488	49.1%	1324	4503	26.6%	293.45	2.2E-16	**
震度	142	816	14.8%	672	5155	14.9%	8.44	3.7E-03	n.s.
ライフライン被害	122	836	12.7%	670	5157	11.8%	1.22	2.7E-01	n.s.
避難	333	625	34.8%	2786	3041	46.9%	56.43	5.8E-14	**
仮設住宅	103	855	10.8%	524	5303	8.7%	3.04	8.1E-02	n.s.

例としてフェーズ3，6における「避難」の結果を示す（図4）。このうち避難所と他の語との共起関係に着目すると，地方紙ではフェーズ3で複数の語群や語がみられ，フェーズ6で被災者や仮設住宅との関連が強く表れている。一方，全国紙ではフェーズ3で被災地や支援といった語との強い関連が見られたものの，フェーズ6では共起関係にある語は見られない。これらの結果から，同じテーマであっても共起関係にある語がフェーズによって異なることと，地方紙では被災地のニーズに関する具体的な内容が詳細に報道される傾向にあることが確認できた。一方，全国紙では被災地への支援や他の地域への情報発信に重きを置いて報道されていたと推察される。

謝辞：本研究では，(国研)防災科学技術研究所「首都圏を中心としたレジリエンス総合力向上プロジェクト (a) 首都圏を中心としたレジリエンス総合力向上に資するデータ活用に向けた連携体制の構築」の補助を得た。記して謝意を表す。

参考文献 1) 繁田健嗣，能島暢呂，永井小雪里，加藤宏紀：2016年熊本地震における避難者の発生・解消過程に関する基礎的考察，土木学会論文集 A1（構造・地震工学），Vol.74，No.4，[特]地震工学論文集，Vol.37，2018（登載決定）。 2) くまにちコム：熊本のニュース，「平成28年熊本地震」特集，http://kumanichi.com/feature/kumamoto_jishin/ 3) 朝日新聞：朝日新聞記事データDVD「2016年版」（学術・研究用），4) 樋口耕一：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，237p.，株式会社ナカニシヤ出版，2014。 5) 工藤拓：MeCab: Yet another Part-of-Speech and Morphological Analyzer，<http://taku910.github.io/mecab/> 6) ウィキペディア財団：日本語版 Wikipedia，<https://dumps.wikimedia.org/jawiki/latest/> 7) 筑波大学：RCRANミラーサイト，<http://cran.md.tsukuba.ac.jp/>

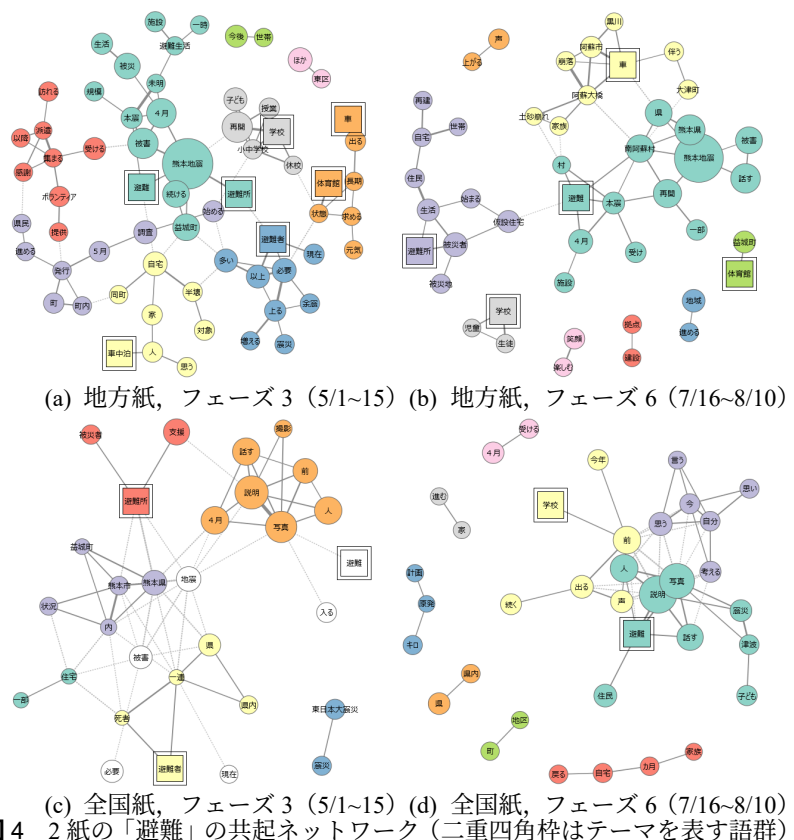


図4 2紙の「避難」の共起ネットワーク（二重四角枠はテーマを表す語群）