

3 - 4 東日本大震災前後の変化

東日本大震災前後で経済構造や景気循環の変化が起こったどうかを考察するに際し、本項では以下の2つの観点からのアプローチを試みた。

分析E：震災前後で構造的な変化、すなわち時間が経過しても元には（少なくとも容易には）戻らない変化が起こったかどうか

分析F：震災直後に変化が起こったものの、時間の経過とともに縮小、あるいは解消の方向に向かったいわば一時的変化がどう起こったか

分析Eでは、東日本大震災の前後で時期を2分割した上で、景気ウォッチャー調査のコメントで言及された景気に関連するテキスト特徴量が前後において差があったかを分析した。

また分析Fでは、景気ウォッチャー調査のコメントにおいて景気に関連する特徴表現の出現回数が震災直後からどのように変化していったかを分析した。

表 3 - 4 - 1 3 - 4 における分析の概要

	分析	条件設定	概要・内訳
東日本大震災前後の変化	E	分析内容 学習期間 予測期間 説明変数 目的変数	震災前後のテキスト特徴量の変化に関する分析 各地域・各月の集計データ（2010年1月～2013年12月の48か月分×全国を含む12地域総データ数576個） なし テキスト特徴量(下は4分析における特徴量の内訳と区分) E-1-1：現状判断、地域別月別(データ数=576) E-1-2：先行き判断、地域別月別(データ数=576) E-2-1：現状判断、個人別(震災前/後・各2500件) E-2-2：先行き判断、個人別(震災前/後・各2500件) 震災前を0、震災後を1として震災の前か後かを判別
	F	分析内容 分析方法	震災前後の特徴表現の出現回数分析 東日本大震災に関係が深い特徴表現13語 「大震災」「東日本大震災」「東日本」「大震災の影響」「東日本大震災の影響」「大震災の影響で」「東日本大震災の影響で」「影響」「福島」「第一原子力発電所」「キャンセル」「停電」が2011年1月から2013年12月までの現状判断コメントに使われている頻度を月次でカウント、推移を比較

3 - 4 - 1 分析E：震災前後のテキスト特徴量に関する分析

分析Eでは、分析を分析E - 1、分析E - 2の2種類に分け、それぞれに対して震災前と震災後を判別するための学習を異種混合学習エンジンに実施させている。分析内容と学習データは以下の記載通りである。

分析E - 1：

目的変数：震災前を0、震災後を1として震災前後を判別

説明変数：各地域・各月ごとのテキスト特徴量の集計データ および 地域コード
(3 - 4 - 1 - 1として現状判断、3 - 4 - 1 - 2として先行き判断のテキスト特徴量の分析を行った)

学習データ：各地域・各月の集計データ (2010/1~2013/12の48カ月分×全国を含む12地域 = データ数576個)

予測データ：なし

分析E - 2：

目的変数：震災前を0、震災後を1として震災前後を判別

説明変数：各個人のテキスト特徴 および 地域コード
(3 - 4 - 1 - 3として現状判断、3 - 4 - 1 - 4として先行き判断のテキスト特徴量の分析を行った)

学習データ：各個人のコメントテキスト 2010/1~2013/12から、震災前2,500件・震災後2,500件(データ数計5,000個)

予測データ：なし

今回の分析では、震災前のデータ数が限られるということと、主目的が震災前後のテキスト特徴量の傾向差を見るためであり予測することが目的ではないことから、予測データをなしとしている。

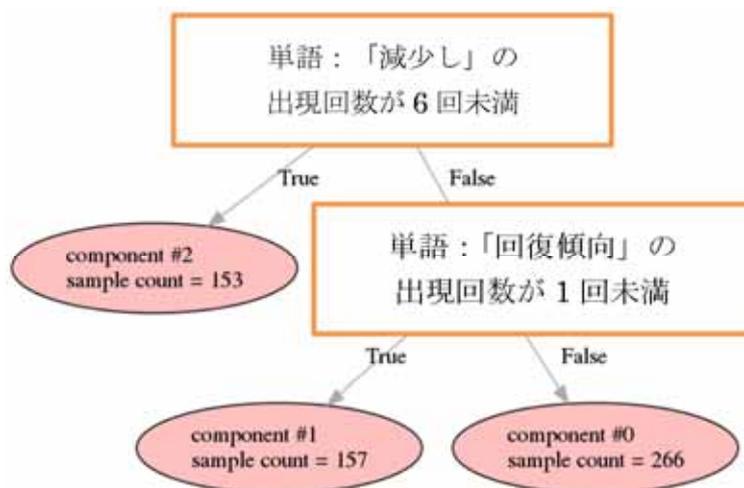
3 - 4 - 1 - 1 分析 E - 1 の結果：現状判断を使った結果

表 3 - 4 - 2 現状判断・地域別月別の判別結果

(数値は月別・地域別コメントのサンプル数)		予測値		
		震災後(P)	震災前(N)	
実績値	震災後(P)	402	6	408
	震災前(N)	103	65	168
	計	505	71	576

適合率(P)=79.6%、F(P)=88.1%、適合率(N)=91.5%、F(N)=54.4%

学習結果の判別モデルは以下のとおりである。



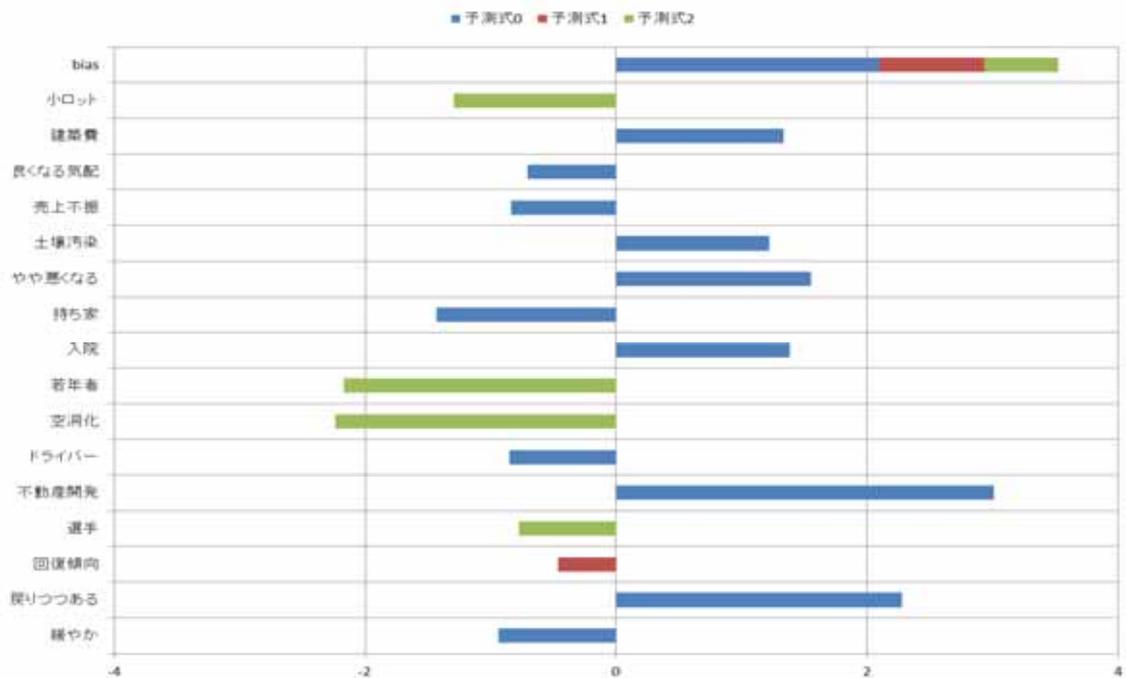


図 3 - 4 - 1 . 抽出表現の係数の重み (3 - 4 - 1 - 1)

「減少し」の出現回数が 6 回以上、「回復傾向」の出現回数が 1 回以上となる地域では、震災前に「売上不振」「良くなる気配」、震災後に「不動産開発」「土壌汚染」「建築費」等という表現が多く使用された。「減少し」の出現回数が 6 回未満となる地域では、「空洞化」「若年者」「小ロット」という表現が多く使用され、全て係数はマイナスであった。

3 - 4 - 1 - 2 分析 E - 1 の結果：先行き判断を使った結果

表 3 - 4 - 3 先行き判断・地域別月別の判別結果

(数値は月別・地域別コメントのサンプル数)		予測値		
		震災後	震災前	
実績値	震災後	390	18	408
	震災前	88	80	168
	計	478	98	576

適合率(P)=81.6%、F(P)=88.0%、適合率(N)=81.6%、F(N)=60.2%

学習結果の判別モデルは以下のとおりである。

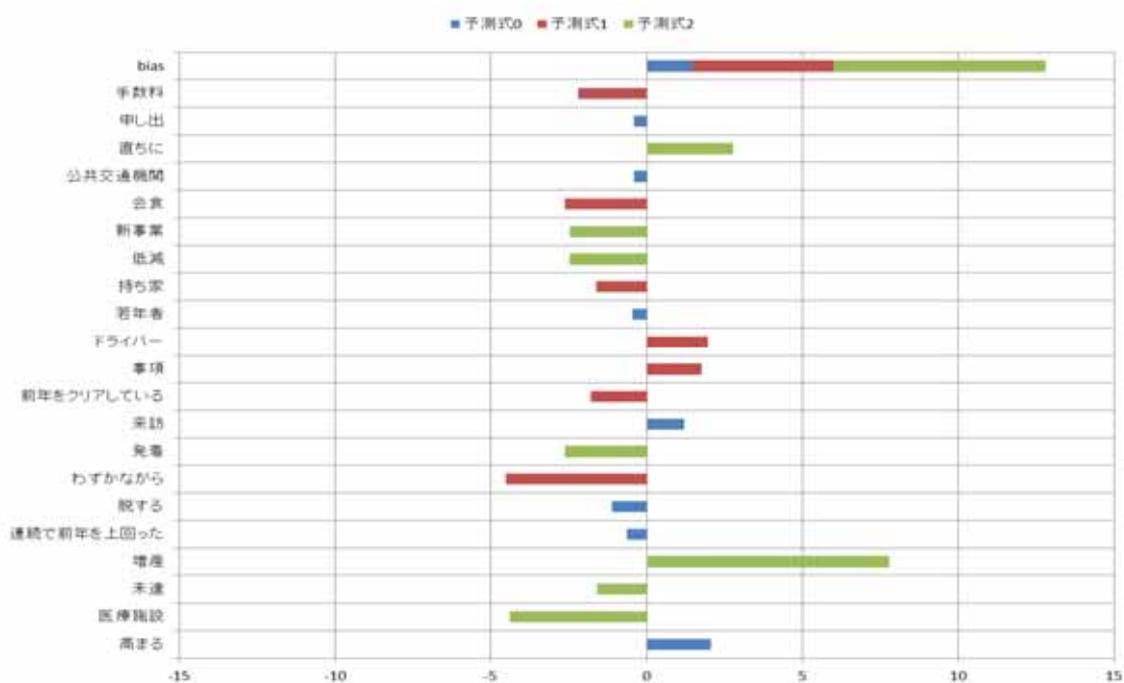
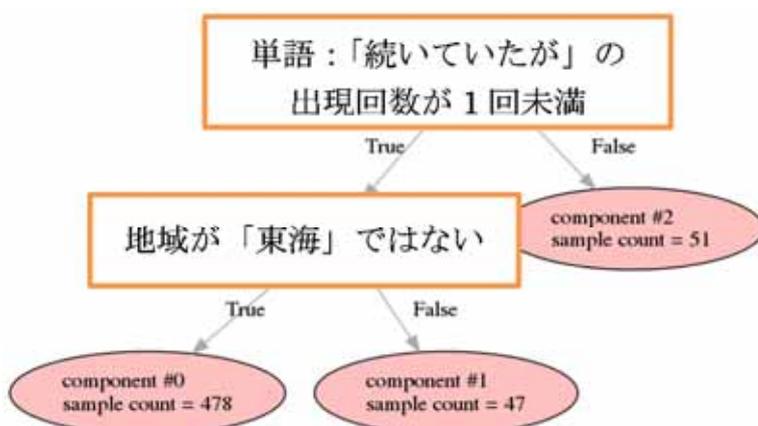


図 3 - 4 - 2 . 抽出表現の係数の重み (3 - 4 - 1 - 2)

東海地域においては、震災前に「会食」「持ち家」などの表現が出ており、「医療建設」・「新事業」なども震災前に使われていた可能性が高い。一方、「続いていたが」は震災後に多く使われていた可能性が高く、かつ「増産」「直ちに」という表現が震災後に使われていた。

3 - 4 - 1 - 3 分析 E - 2 の結果：現状判断を使った結果

表 3 - 4 - 4 現状判断・個人別の判別結果

(数値はコメントの サンプル数)		予測値		
		震災後(P)	震災前(N)	
実績値	震災後(P)	846	1,654	2,500
	震災前(N)	684	1,816	2,500
計		1,530	3,470	5,000

適合率(P)=55.3%、F(P)=42.0%、適合率(N)=52.3%、F(N)=60.8%

学習結果の判別モデルは以下のとおりである。

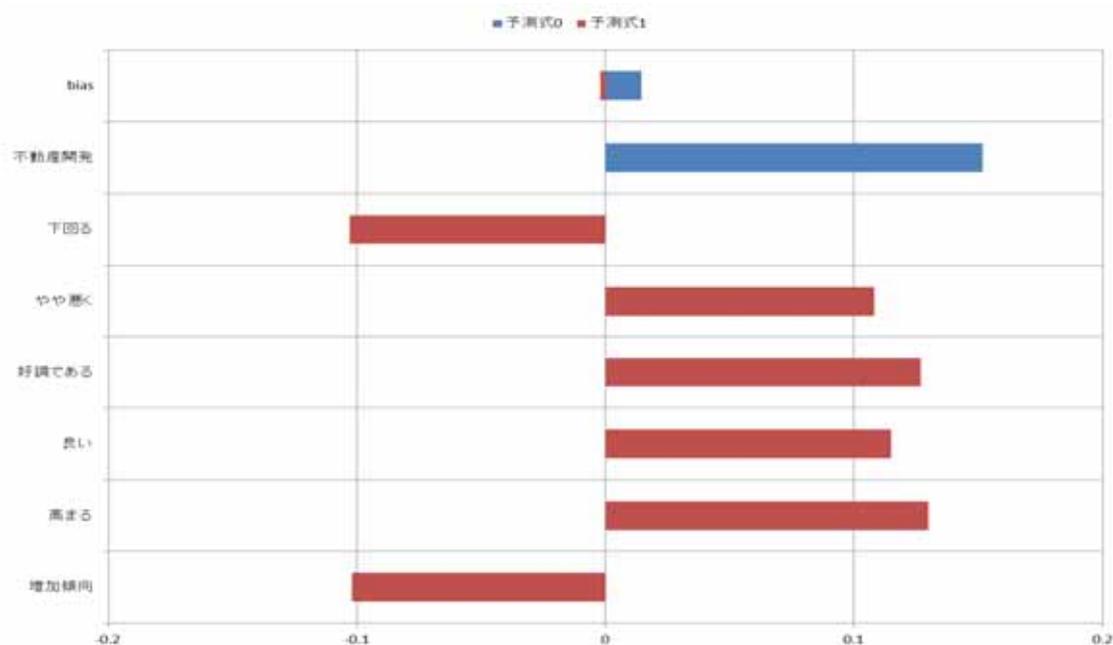
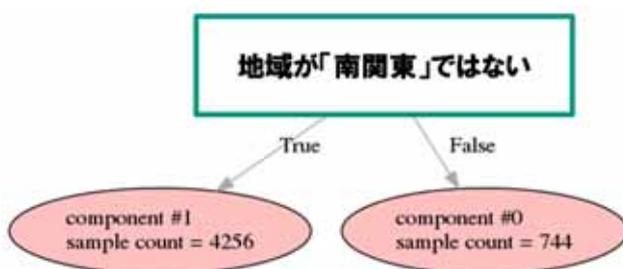


図 3 - 4 - 3 . 抽出表現の係数の重み (3 - 4 - 1 - 3)

南関東において、「不動産開発」が震災後に顕著に使われており、それ以外では、「好調」「よい」「高まる」などの語が震災後に見られるという結果となった。

3 - 4 - 1 - 4 分析 E - 2 の結果：先行き判断を使った結果

表 3 - 4 - 5 先行き判断・個人別の判別結果

(数値はコメントの サンプル数)		予測値		
		震災後(P)	震災前(N)	
実績値	震災後(P)	1,371	1,129	2,500
	震災前(N)	1,216	1,284	2,500
計		2,587	2,413	5,000

適合率(P)=53.0%、F(P)=53.9%、適合率(N)=53.2%、F(N)=52.3%

学習結果の判別モデルは以下のとおりである。

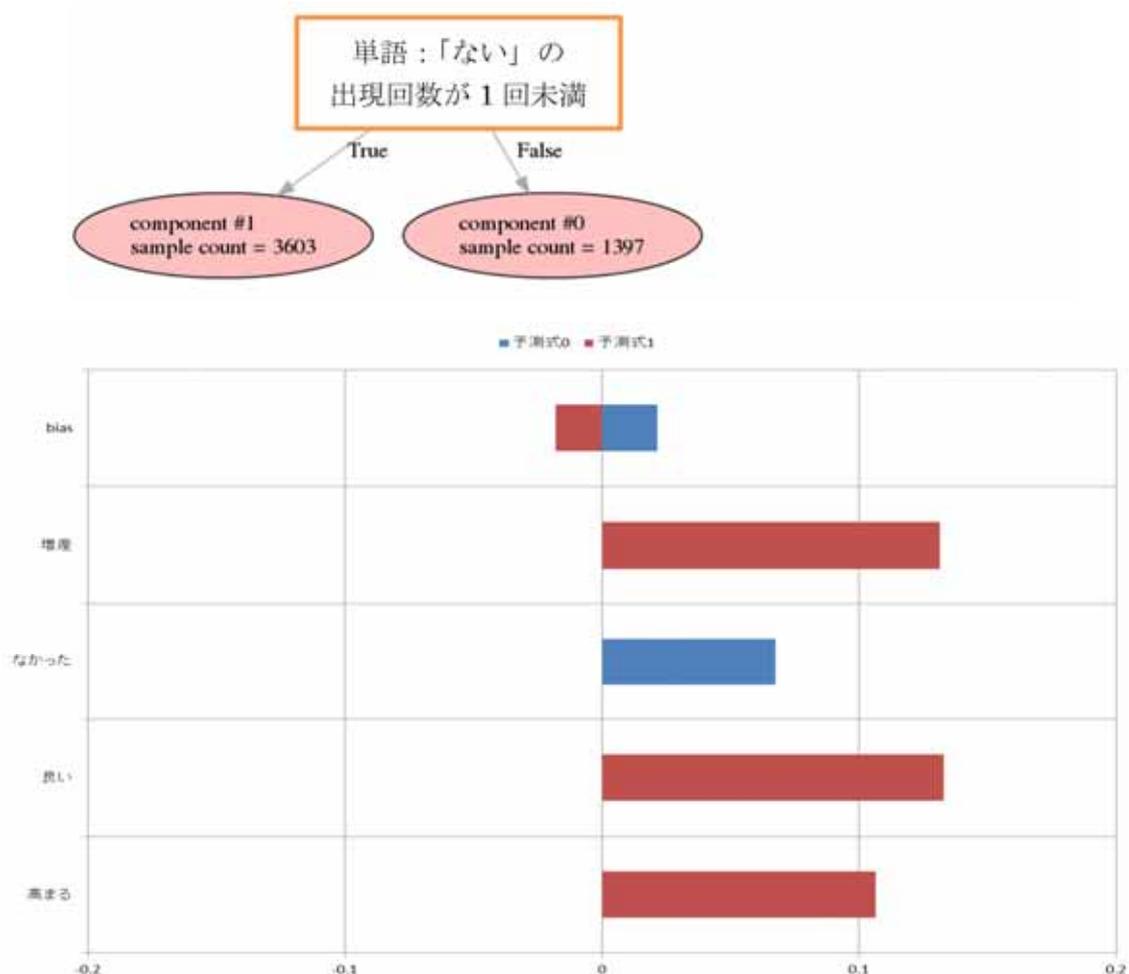


図 3 - 4 - 4 抽出表現の係数の重み (3 - 4 - 1 - 4)

震災前後の判別が明示的に連想される語句は抽出された特徴表現からはあまり出てこなかった(「増産」が震災後を示唆する語として出てきたのみ)

3 - 4 - 1 - 5 分析結果の評価、及び考察

3 - 4 - 1 - 1 および 3 - 4 - 1 - 2 から、地域別・月別で集計した際の震災前後の判別はテキストデータからある程度できる可能性が確認された。

たとえば、3 - 4 - 1 - 1 では、「減少し」が地域内で 6 回未満であると震災前である可能性が高く（係数がマイナスに効いている特徴量が多いため）かつその場合特に「若年者」「空洞化」などが多いことが震災前に見られること、「減少し」が 6 回以上・「回復傾向」が 1 回以上の地域においては、「建築費」「土壌汚染」「不動産開発」が震災後に多くみられる（これらの特徴量は全て係数がプラスに効いている）と学習していることがわかる。他にも、3 - 4 - 1 - 2 において、震災後において「増産」「直ちに」、震災前について「医療建設」「新事業」が多かったことがわかる。

このように、震災前後での判別問題の学習結果から、震災前・震災後に顕著に変化した言葉を発見することができる可能性を確認した。

なお、震災前の「若年者」「空洞化」、震災後の「建築費」「土壌汚染」「不動産開発」という語群から経済や景気に影響を与えるトピック、事項の変化が構造的に起こったかどうかを判別するのはこの分析だけでは十分とはいえない。しかし、震災前の上記 2 語は日本経済の構造問題の所在を示唆する語であるのに対し、震災後の上記 3 語は震災後の復興対処やそれにかかる問題を示唆する語と考えられることから、震災をきっかけに国民が景気を考える上での焦点が（少なくとも震災後数年の間は）構造問題から震災復興対応に移転してきた可能性があると考えられる

一方、3 - 4 - 1 - 3 および 3 - 4 - 1 - 4 においては、震災前後の判別精度が低く、満足な学習結果が得られなかった。これは、今回のアンケートデータ内の文字数からは、震災前後の差以上に個人ごとの言葉のばらつきが非常に大きく、震災前後の判別が困難であったことがうかがえる。ただし、これは、個人ごとの回答テキストデータの文章量が増加することで改善する可能性があり、例えば今後、twitter やブログデータなど一人あたりのテキスト量が多い情報ソースの分析を行うことでよりよい結果が得られる可能性があると考えられる。

3 - 4 - 2 分析F：震災前後の特徴表現の出現回数分析

分析Fでは、景気ウォッチャー調査のコメントにおいて景気に関連する特徴表現の出現回数が震災直後からどのように変化していったかを分析した。具体的には、3 - 2 - 6の表3 - 2 - 8で取り上げた東日本大震災後の特徴表現上位リストから特に東日本大震災に関係が深いと考えられる特徴表現を取り出し、この特徴表現が2011年1月から2013年12月までの期間、景気ウォッチャー調査のコメント（現状判断）の中にどれくらいの頻度で使われているかを月次でカウントした。

抽出した特徴表現は、「大震災」「東日本大震災」「東日本」「大震災の影響」「東日本大震災の影響」「大震災の影響で」「東日本大震災の影響で」「影響」「福島」「第一原子力発電所」「福島第一原子力発電所」「キャンセル」「停電」の13語である。

いずれの特徴表現も2011年2月以前にはほとんど出現しないが、2011年3月の震災発生時に急激に増加し、その後は徐々に回数が減少している。特徴表現ごとの出現回数の差を見ると、東日本大震災を直接さし示す「大震災」、「東日本大震災」「東日本」といった特徴表現の頻度が多く、「第一原子力発電所」「キャンセル」「停電」といった東日本大震災によってもたらされた事象や課題を示す特徴表現についてはやや頻度が少ないという結果が見て取れる。

これらの特徴表現が頻繁に使われている期間は震災発生から1年程度で、それ以降の出現回数は大きく減少しており、震災が景気に与えた影響は2012年の前半までということが景気ウォッチャー調査のコメント数からは読み取れる。

個別の特徴表現をみると、「停電」については、他の特徴表現よりも早い減少が見られており、2011年3月に計画停電が行われるなど一時は緊迫した電力供給の状況が比較的速やかに改善されたことがうかがえる。一方、「影響」については3月に急増してからの減少が他の特徴表現よりも小さく、大震災を超えて幅広い用途で使われていることが推察される。

以上の結果から、景気ウォッチャー調査のコメントから判断される東日本大震災の影響は2011年3月から約1年間程度であることが伺え、また個別の特徴表現に注目することで特定の事象（例：停電）が影響した期間について推察できる可能性が示されたと考える。

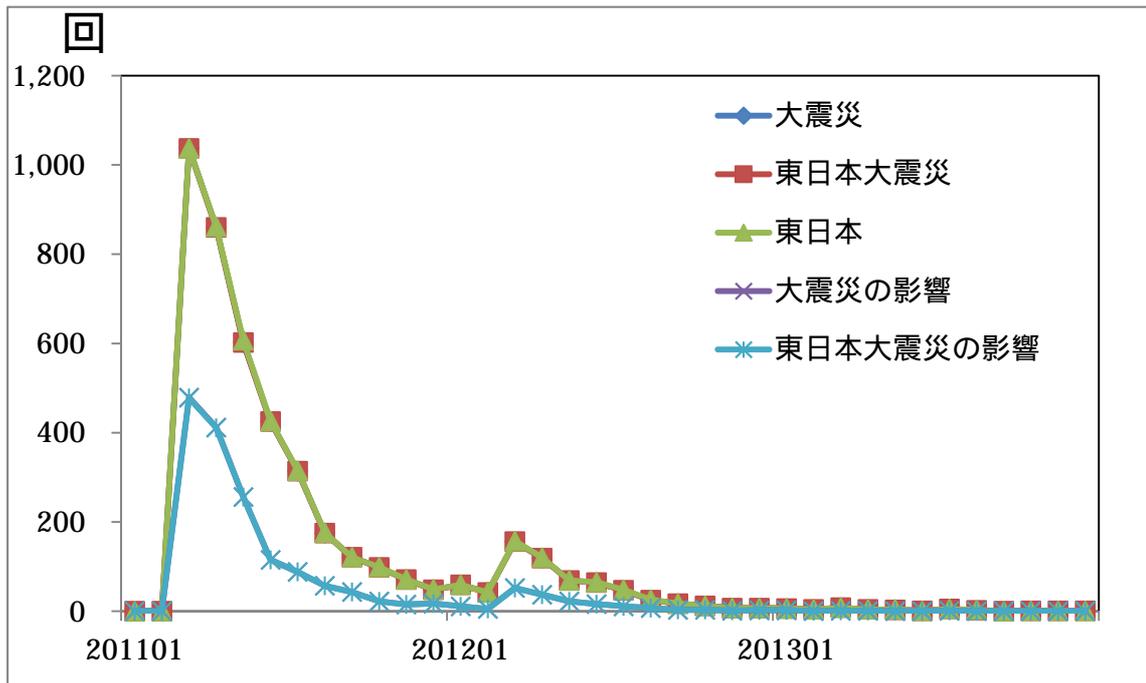


図 3 - 4 - 5 東日本大震災に関係の深い特徴表現の出現頻度

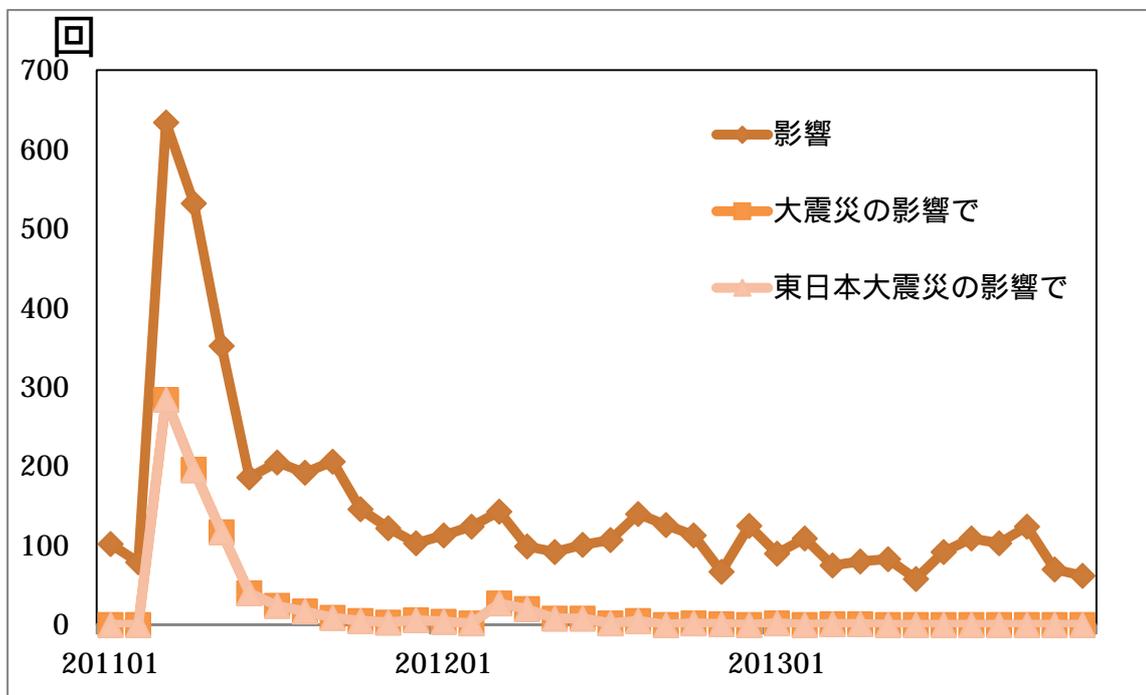


図 3 - 4 - 6 東日本大震災に関係の深い特徴表現の出現頻度

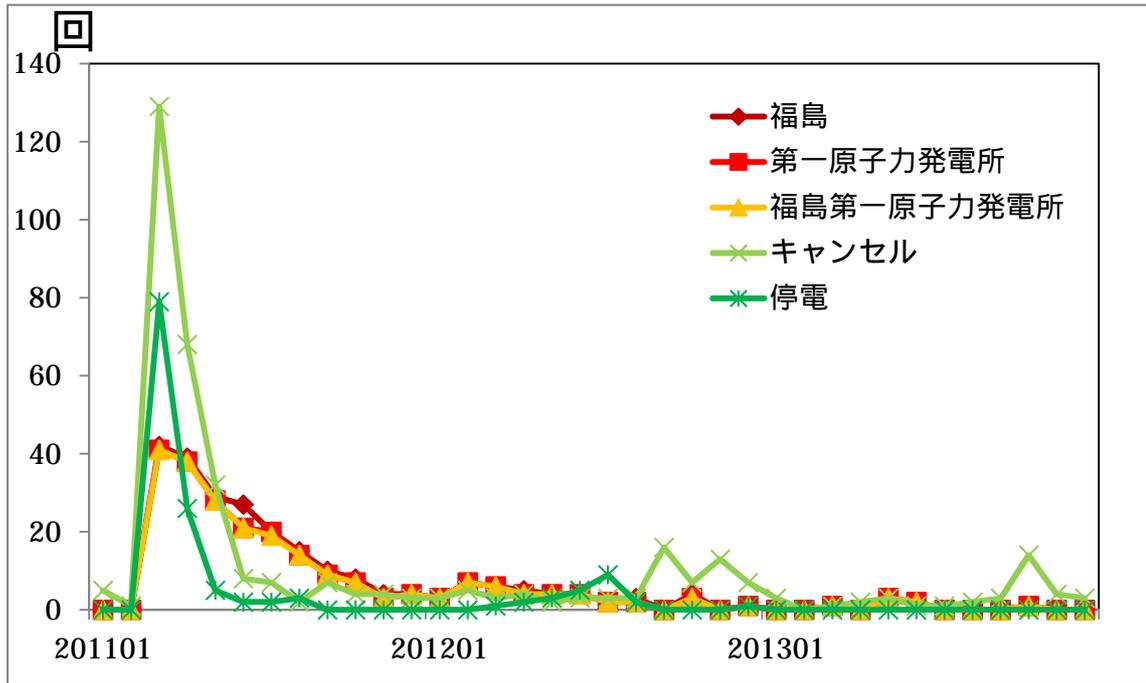


図 3 - 4 - 7 東日本大震災に関係の深い特徴表現の出現頻度