

## 【研究ノート】

# マクロ・ソーシャル・マーケティング研究における テキストデータ分析に関する一考察 —新聞言説を用いた KH Coder による実証分析を通して—

石澤 泉\*

## Abstract

The purpose of this paper is to examine the possibility of text data analysis as an analysis method with macro-social marketing research focused on the transition of social phenomena as a clue. Therefore, in recent years, the author focus on text mining, which has drawn attention as a method of qualitative analysis, and confirm the characteristics of each analytical approach. After that, the author empirically study based on newspaper article data on smoking cessation and consider how to do qualitative research. While this attempt is exploratory, it will present concrete ways to deal with the problem of lack of research methods which is cited as the reason why macro-social marketing research is not progressing.

## 1. 問題意識と研究目的

今日、直面している社会問題、例えば、禁煙やアルコール、過食などによる健康問題、さらには環境問題など多くは、ミクロな意思決定の集合というよりは、マクロな社会現象として捉えることができるように見える。もちろん、マクロとミクロは密接なループ関係にあるが（坂田, 2005; Layton, 2015）、少なくともこうした問題を研究対象とする、ソーシャル・マーケティング（Kotler, 1971）研究では、マクロな社会現象があまり注目されてこなかった。

その理由として、一つには、もともとソーシャル・マーケティングは直接的な個人の態度変革や社会変革をかなり実践的な点から取り扱ってきた（French & Gordon, 2015）ということが指摘できる。関連する研究として、メガ・マーケティング（Kotler, 1986）の視点からのカジノのイメージ変化を捉える考察（Humphreys, 2010a, 2010b）や、ポトックスの受容プロセスを例に新技術への社会からの拒否反応とそれを克服するマーケティング間の相互作用の考察（Giesler, 2012）を挙げることができるが、その数自体もまだ多くはない。

\* 首都大学東京 大学院社会科学研究科 経営学専攻 博士後期課程

それからもう一点、マクロな社会現象やその歴史の変遷を具体的に捉える方法があまりはつきりとしていなかったという点も指摘できるかもしれない。例えば、先述の、Giesler (2012) では、時代ごとのボトックスの対立的なブランドイメージを膨大な資料やインタビューから明らかにしているが、これらはあくまで質的な分析であり、解釈的である。これに対して、Humphreys (2010a, 2010b) では新聞記事の統計的な処理と分析が試みられているが、特定フレームの登場頻度の推移を考察しており、カジノのイメージ変化する現象の全体像を捉えているわけではない。

そこで、本研究では、質的研究の方法として近年注目されてきたテキストマイニングを取り上げ、各々の分析アプローチの特徴を確認した上で、禁煙に関する新聞記事データをもとに実証的に質的研究のあり方について考察することを目的とする。これは、マクロ・ソーシャル・マーケティング研究の具体的な方法を提示することでもある。

次節では、社会現象を捉えようとする質的研究の現状を簡単に概観する。その上で、3節で、二次データによるテキストマイニングを用いた先行研究を確認する。4節で、テキストマイニング分析の手法について三つの手法（多次元尺度構成法・クラスター分析・共起ネットワーク）を挙げて各々の特徴を確認する。5節で、具体的事例として禁煙に関する新聞記事データを基に、KH Coder を用いて多次元尺度構成法、クラスター分析及び共起ネットワークの分析方法について述べる。6節で、実証分析の結果を示し、各々の特徴を確認する。7節で、先行研究が指摘する分析視点での各々の当てはまりを考察し、そして最後の8節で帰結を述べる。

## 2. 社会現象を捉えようとする分析の現状

社会現象を実証的に説明する研究は、定量的なデータを扱う量的研究と、定性的なデータを扱う質的研究に分けることができる。量的研究を行うためには、研究対象とするトピックに関連した先行研究の中で、すでに概念が抽出されている必要があり、まだそのレベルまで研究成果が蓄積されていない場合には質的研究を用いることになる（戈木クレイグヒル, 2014）。

戈木クレイグヒル (2014) は、1990年から2010年までの日本における質的研究の論文数、研究法の内訳などを調査し、保健医療系雑誌への発表に限られた結果であるものの、同分野が国内で情報量が一番多いものであるとして、日本における質的研究の一般的な傾向を明らかにしている。質的研究は増加し、その中で、研究法別では、KJ法 (618本, 19.3%)、グランデッド・セオリー・アプローチ (432本, 19.3%)、内容分析 (390本, 17.4%)、ナラティブ (337本, 15.0%)、ライフストーリー/ライフヒストリー (223本, 10.0%) の五つが10%以上の論文で使われ、それらに続いて、現象学、アクション

リサーチ、フィールドワーク、エスノグラフィー、エスノメソドロジー・会話分析、談話分析・ディスコース分析が使われていたという（戈木クレイグヒル，2014,p.31）。社会現象を捉えようとする研究が対象とする事象は、複雑、多様なものであり、それぞれの社会的文脈において捉える必要がある。そのため戈木クレイグヒル（2014）の分析にあるように、KJ法やグランデッド・セオリー・アプローチ（GTA）、事例研究など質的研究が積極的に行われてきた（日和，2013）。一方で、質的分析には、研究者や実務担当者の解釈が本質的な部分を占めることになり、量的方法と比べて信頼性と妥当性が問われることが少なくない。そのため、一つの研究において複数の方法を用いたり、異なる視点で解釈が取束する点を探索したりするトライアンギュレーションによって研究の妥当性が高められる（Jick, 1979）とされてきた。

しかしながら、近年では、質的データ分析の支援を目的とした、QDA（Qualitative Data Analysis）ソフト、あるいは、CAQDAS（Computer Assisted Qualitative Data Analysis Software）と呼ばれるツールが増えてきている。それら分析ツールの多くは、研究者がテキストデータを解釈し、それらに対してコード化・カテゴリー化を行い、そこから重要な概念を見つけ出すという、GTAに基づくデータ分析作業の支援を目的にデザインされている（稲葉・抱井，2011）。そのことが表裏の関係になって研究方法としてGTAを採用する研究が多くなっている一つの理由かもしれない。

QDAソフトを活用することによる分析作業の時間的、質的、データ管理の簡素化など大きな利点がある（稲葉・抱井，2011）が、テキストの解釈は研究者に委ねられている。一方、樋口（2014）は、大量のデータに対して量的方法を適用するためにはコンピュータ利用の必要性を強調し、その上で、むしろ、量的方法によってデータの全体像を得ることで偏った、不完全な、そして非常に選択的な印象の形式を避けられる、と言う。また、計量的分析手法をテキスト型データに適用（テキストマイニング）することで信頼性・客観性の向上とデータの探索・要約の容易性というメリットがあると指摘している。

### 3. 二次データを用いたテキストマイニングによる先行研究

樋口（2014）が指摘するように、このような手法を使って、近年では過去の新聞記事などのデータを分析対象とし、より客観性を高めた上でマクロな社会現象を捉えることが可能になりつつある。新聞は、記事のデータベース化・利用サービス化が進み、日経新聞の「日経テレコン」、読売新聞の「ヨミダス」、朝日新聞の「聞蔵」など全国紙の蓄積されたデータをもとにして、過去の意識を探り得る、そして過去と現代を比較し得るという、一般的な社会調査法にない利点があり、ステイクホルダーの関係性も含む複雑な関係性や、包括的に社会的文脈形成プロセスを経時的に追跡するときには有用な資料となる（松井，2013；

樋口,2014)。同様に、雑誌でも「大宅壮一文庫」や「Magazineplus」などの記事データベースから、雑誌ならではの世相やトレンドを調べる上で有用な資料を入手することができる(松井, 2013)。

こうした新聞記事をデータとして使用することの信頼性について、樋口(2014)は、全国紙の内容分析から社会意識の探索の有効性を確認する実証分析を行い、新聞紙面に多く現れる主語や用語ほど、人々の念頭に浮かびやすいということ、その意味で社会意識においても高い顕出性を示すという関連が見出され、部分的にはあるが、新聞報道と社会意識の類似性・相関関係を確認できることを明らかにしている。経営学領域においても、新聞記事の言説を基にした先行研究として、松井(2013)、水越・コールバッハ(2015)、勝又・西本(2016)、高井(2017)などを挙げるができる。同様に雑誌記事の言説を分析対象にした織田(2017)や公文書議事録、政策評価書等公文書の言説を用いた谷口(2017)も挙げられる。

その中でも、松井(2013)は、共起ネットワーク分析により、「癒し」言説と消費の関係を明らかにしようとしている。同様に、織田(2017)は、共起ネットワーク分析により、スティグマ分析の視点から「婚活」言説の意味の変容を考察し、谷口(2017)は、「バイオマス予算」の獲得と予算配分の失敗の関係を明らかにしている。一方、高井(2017)は、コレスポネンス分析により、黎明期のオンライン証券企業間の戦略ポジショニングを考察している。これらは先述のHumphreys(2010a, 2010b)やGiesler(2012)の考察を分析手法という点でうまく補完すると思われる。

#### 4. テキストデータの量的分析

テキストマイニングとは、テキストデータを様々な計量的手法によって分析し、形式化されていない膨大なテキストデータという鉱脈の中から言葉(キーワード)どうしに見られるパターンや規則性を見つけ、役に立ちそうな知識・情報を取り出そうとする手法・技術(藤井, 2005)である。論文検索サイトCiNiiで「テキストマイニング」「社会」で検索すると318件の論文があり、そのうち福祉、医療分野が多い。内容を概観すると同分野では、関係者の自由回答やインタビューが一般的な調査方法になっていることからかもしれない。同様に「テキストマイニング」「マーケティング」では56件確認できる<sup>1</sup>。

近年では、こうした一次データに限らず、マスメディアや公文書などのテキストデータをデータとして用いて社会現象を探索する研究が増えてきている。どのような社会的行為が生成されていくのか、社会現象を、外的基準をもたず探索的に、言説間の共起関係を分析し可視化するために大別して、多次元尺度構成法やコレスポネンス分析のように言説間の全体的な関係を俯瞰的な布置による表現する分析方法と、共起ネットワーク分析や

クラスター分析のように言説間の局所的な共起関係に注目する分析方法がある。(与謝野, 2006)。ここでは、主として三つの方法について各々の特徴を確認することにする。

多次元尺度構成法 (MDS) は多変量解析の手法の中で、いわゆる外的基準をもたない手法に位置づけられる。連続的な多次元空間を用いて、データに潜んでいる情報を目で見てわかりやすいように幾何学的に表現する。対象間の距離や非類似度の大・小を、次元空間上における遠・近として表すことを目的とし、出現パターンの似通った語の組み合わせにはどんなものがあつたのかを探索するのに適していると言われる。言説の距離が近いほど言説間の関係性が強い類似性の高い言説が近くに布置される (岡太, 2002)。

クラスター分析は、対象を類似しているか類似していないのかによって分類する方法の総称であり、大別して階層的な手法と非階層的な手法に分けられる。階層的クラスター分析は、分類しようとしている対象 (n 個) 一つ一つがその対象自身のみからなるクラスターの状態 (n 個のクラスター) から、対象全体が一つのクラスターになる階層的な構造を作っていく方法である。一方、非階層的クラスター分析は、あらかじめクラスターの数を決めておくか、分析の過程で決めるなど、与えられた何らかの基準を最適化するようにクラスターにまとめ、分類する方法である (与謝野他, 2006)。

共起ネットワーク分析は、共起の程度が強い語を線で結ぶことで言説間のネットワーク構造を表現する。MDS とは異なり、単に言説がお互いに近くに布置されているだけでは、それらの言説の間に強い共起関係があることを意味しない。重要なのは線で結ばれているかどうかであって、近くに布置されているだけで、線で結ばれていなければ、強い共起関係にはない。言説間の距離には意味がなく、線で結ばれているか否かが重要になる。ネットワークを見やすく配置するための手法である (樋口, 2014)。コンピュータの解析処理の結果を示すことで、データの全体像を簡潔な形で示すことができる。また、線で結ばれているか否かで、分析者の解釈が、どのようなプロセスで出されたのか、第三者の研究者の理解を容易にすること可能になる (稲葉・抱井, 2011)。そもそも、共起性とは、同じようなネットワークに属し、同じような他者との関係を持つ同値性の高いノードは、アクセスできる情報、外部環境についての状況認知や思考パターンが似てくるため、社会的な行為に至る過程でも、外部からの同じような刺激に対して同じような意思決定のプロセスを辿り、その判断や趣味・嗜好に類似性が生まれるという、ノードの社会心理を表した概念である (中島, 2013)。

これらの分析手法に対して、これら何れの手法も、分析対象とする単語の数を増やしていくと、クラスターやネットワークが複雑化し過ぎ、意味のまとまりを解釈するのが困難になるという問題や、単語の顕在的な共起関係のみに基づいて分析を行うため、少数の単語からなる周辺の話題は比較的良好に分離できるが、多くの単語が複雑に共起する中心的な話題は大きなかたまりとなって内容を把握するのが難しいという問題を抱えている



(Matsukawa et al., 2015) として、確率モデルのトピックモデル (岩田, 2015) を用いたテキストデータの自動分類による研究方法への取組みが提案され、その有用性が確認されている (岩田, 2015; 松河他, 2017)。

トピックモデルは、文書が生成される過程を、確率を用いてモデル化した確率モデルで、確率を用いることによって不確実性を扱うことができ、ノイズを含むデータの中から本質的な情報を抽出できるようになる (岩田, 2015) もので、トピックモデルを用いれば、テキストデータ全体の中にどのような単語の集合から構成されるトピックが含まれているのかという概要把握が可能になるのに加え、各文書がどのようなトピックから構成されているかということも推測が可能になる。トピックモデルでは、データが隠れた構造を持つことを仮定し、潜在的な単語間の関係も考慮されるため、従来の手法では難しかった、文脈によって同じ単語が違う意味を表すケースなどにも対応が可能であり、データから隠れた構造を発見し、高い分類精度が期待できる (Matsukawa et al., 2015) と言われている。

## 5. KH Coder を用いたテキストマイニングによる実証分析

### 5.1 実証分析の視点

以下では、実際に質的データを用いて分析を行う。具体的には、二次データである「禁煙」言説の歴史的変遷のデータを基に、記事出現数の推移とともに、その内容の推移について、テキストマイニング用ソフトウェア「KH Coder」(樋口, 2014) を用いて日本における禁煙意識の広がりを分析対象とすることにする。このソフトウェアの特徴は、無償でウェブサイトから入手でき、インストールから実際の分析まで簡単な操作で使用できる。また、優れたデータ抽出機能を備えており、テキストから様々な情報を簡単に取り出すことができ、コンピュータ・スキルをそれほど問わない利点がある。本研究では、具体的な二次データを基に、KH Coder を用いて MDS とクラスター分析及び共起ネットワークの各々の方法で分析し、KH Coder が示す結果について各々の特徴を実証確認することにする。なお、トピックモデルはその分析において異なるソフトウェアを使用することになるため、本研究では、分析手法の対象としない。

分析に当たっては、これまで、マクロ・ソーシャル・マーケティング研究において社会変革や行動変革を分析するために指摘されてきた注目点を方法選択判断の視点として利用する。

#### ● 分析における社会変革や行動変革の注目点

- a) ソーシャル・マーケティングでは、ステイクホルダーは予め顕在化しているのではなく、むしろステイクホルダーの顕在化と獲得が重要な活動になる。しかも、彼ら

の利害関係は関係性次第で動的に変化するため、ステイクホルダーの多様性を考慮する上で関係性概念に注目することが重要である（水越・日高，2014）。

- b) 社会的変革の考察に関して、マーケティング・システム全体がどのように機能しているのかを捉えることが重要である。このさい、マイクロレベルでの個人の相互作用やマイクロレベルでの個人の主観的な認知が、どのようにしてマクロレベルでの制度や文化などの社会構造と結びついているのかというマクロ・マイクロリンクが重要である（Layton, 2015）。
- c) 市場の創造に関して、ダイアドの関係だけではなく、市場が創造される背景に存在する多数の主体間の複雑な相互作用を明らかにするための社会的、文化的、規制的な制度的特徴を捉えることが重要である（Humphreys, 2010a）。

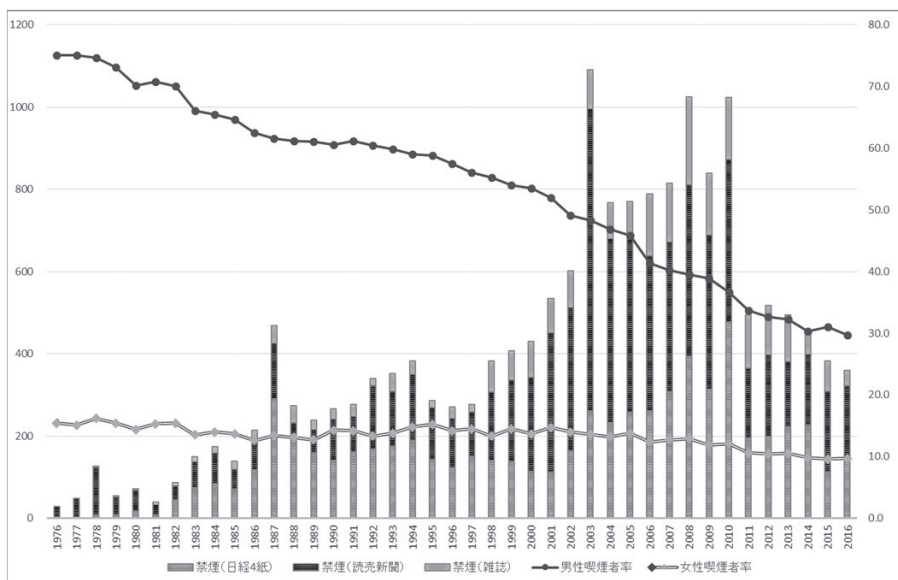
## 5.2 方法

### 5.2.1 データの取得方法

JTが毎年発表している「全国喫煙者率調査」から1976年から2016年までの40年間の「喫煙者率」データに合わせ同期間とし、記事検索では、40年間のデータ取得が可能な「日経テレコン21」から日経四紙、「ヨミダス」から読売新聞を検索した。また、補完的に、雑誌記事からも同様に、「Magazineplus」を用いてデータを収集、解釈した。多次元尺度構成法（MDS）とクラスター分析及び共起ネットワーク分析のためのデータの取得は、日経テレコンから「禁煙」言説で検索し、出現した記事見出しと記事をExcelに貼り付け、CSVに変換してデータセットを行った。

時系列データの分析において重要になるのは、期間の区分である（松井，2013）。図1が示すとおり、喫煙者率は、データ取得初年の1976年から2016年に至るまで一貫して減少している。1976年には男性75.1%、女性15.4%だったのが、2016年には、それぞれ29.7%と9.7%と大幅に減少している。特に男性は45.4ポイントも減少した。一方で、「禁煙」記事は増加しており、「禁煙」記事出現数と男性喫煙者率の相関係数は $-0.6667$ で両者は「負の相関関係」<sup>2</sup>にある。マクロな社会現象として、喫煙率は下がり、一方で「禁煙」記事の数は増えてきたということになる。その一方で、「禁煙」記事には突発的な増加のタイミングがみられる。そこで本研究では、記事出現の推移から、最初に記事が突発的に増加する1987年から次に突発的に記事が増加する2003年の前年の2002年の記事出現のデータを基に分析する。なお、2003年以降についても、2011年を前後して出現頻度に違いがみられる。これらは社会の主テーマが「震災」に代わったためと思われる<sup>3</sup>。2003年以降の分析では震災前と震災後に留意する必要があると思われる。

図1 JT全国喫煙者率推移と「禁煙」記事出現数変遷



JT全国喫煙者率調査及び日経テレコン・ヨミダス・Magazineplusのデータから筆者作成

### 5.2.2 分析言説の抽出方法及び分析方法

日経テレコンで「禁煙」で検索した1987年から2002年の記事を元データ（使用文数10,920文）に、その中から、出現回数50回以上の語のうち、名詞、名詞B、名詞C、固有名詞、サ変名詞、未知語を抽出し分析対象として用いて（表1）、多次元尺度構成法（MDS）、クラスター分析及び共起ネットワークで分析を行った。

表1 禁煙記事 頻出語抽出リスト

| 順位 | 抽出語 | 回数   | 順位 | 抽出語  | 回数 | 順位 | 抽出語  | 回数 |
|----|-----|------|----|------|----|----|------|----|
| 1  | 禁煙  | 3447 | 19 | 企業   | 78 | 37 | 情報   | 59 |
| 2  | 喫煙  | 450  | 20 | 拡大   | 77 | 38 | 全国   | 59 |
| 3  | たばこ | 387  | 21 | オフィス | 74 | 39 | 指定   | 59 |
| 4  | 社長  | 212  | 22 | タイム  | 74 | 40 | 市場   | 57 |
| 5  | 全面  | 202  | 23 | 職場   | 73 | 41 | 場所   | 57 |
| 6  | 運動  | 185  | 24 | ホテル  | 72 | 42 | バス   | 56 |
| 7  | 席   | 178  | 25 | 車両   | 71 | 43 | 開発   | 56 |
| 8  | 世界  | 145  | 26 | タバコ  | 69 | 44 | 導入   | 55 |
| 9  | 分煙  | 144  | 27 | 会社   | 68 | 45 | 対策   | 54 |
| 10 | 人   | 135  | 28 | 煙    | 68 | 46 | 調査   | 54 |
| 11 | 補助  | 115  | 29 | 愛煙   | 64 | 47 | コーナー | 53 |
| 12 | 駅   | 113  | 30 | 公共   | 64 | 48 | 店内   | 53 |
| 13 | 女性  | 100  | 31 | 教育   | 64 | 49 | 生活   | 53 |
| 14 | 全席  | 98   | 32 | 効果   | 62 | 50 | トップ  | 52 |
| 15 | 嫌煙  | 92   | 33 | 商品   | 62 | 51 | サービス | 52 |
| 16 | 実施  | 87   | 34 | 路線   | 60 | 52 | パイポ  | 50 |
| 17 | 条例  | 83   | 35 | タクシー | 59 |    |      |    |
| 18 | 販売  | 81   | 36 | 航空   | 59 |    |      |    |

「禁煙」検索記事を元にKH Coderで筆者作成



## 6. 結果

### 6.1 多次元尺度構成法 (MDS)

抽出語を用いて、非計量的多次元尺度構成法 (Kruskal) による分析を行った。KH Coderでは語と語の類似度を見るために語句同士の関連を示す Jaccard 係数を用いている。Jaccard 係数は、集合 X と Y の共通要素数を少なくとも一方にある要素の総数で除したもので、語が共起しているかどうかを重視する係数であり、語と語の共起をカウントする (樋口, 2014, p.155)。分析対象の禁煙の言説間の二次元空間上の距離の単調関係のズレ (岡太・今泉, 1994) を表す不適合度 (stress) は、0.342 で当てはまりがよくない (図 2. I)。そこで、三次元空間 (図 2. II) で距離の不適合度を見てみると、stress は 0.270 で 0.072 当てはまりが改善された。対象数や布置の次元数も考慮する必要があるが、およそ 0.05 から 0.5 程度までの stress であればその次元数の結果を解として採用することができるといわれている (中山・横山, 2010) ことから、一定の評価はできることになる。しかしながら、分析の布置からどのような意味を持つのか、解釈するのは難しそうに見える。

それでも、MDS 二次元分析図 (図 2. I) を観察すると、次のようなことを読み取ることができる。「禁煙」を中心に、次元 1 が正で次元 2 の 0 の方向に、「車両」、「路線」、「全席」、「航空」、「バス」など交通機関のグループを示している。同様に、次元 1 が負で次元 2 が正の中頃に「嫌煙」、「運動」、「職場」、「条例」などタバコを抑制しようとするグループを示している。次元 1 が 0、次元 2 が 10 の方向に「場所」、「コーナー」、「駅」、「分煙」、「実施」、「タイム」など喫煙と禁煙を時間や空間で棲み分けしようとするグループを示している。次元 1 が 0、次元 2 がマイナス 10 の方向に「商品」、「開発」、「販売」、「補助」、「市場」など商品開発をしようとするグループを示している。これらから、軸を解釈すると次元 1 は、「場所」を示す軸であり、次元 2 は、「主体」を示す軸であると推定できるように見える。

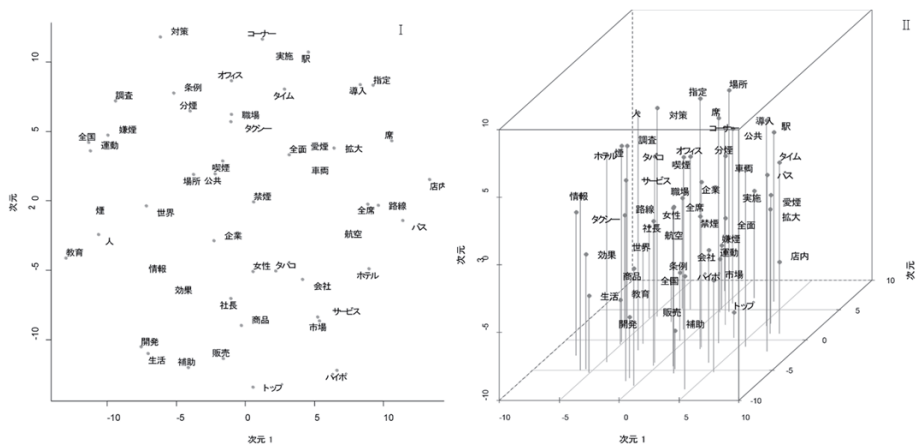
このように MDS の分析においては、一般的に分析者によって分析次元軸の特徴を解釈する。例えば、ある食品のアンケートの自由回答の MDS の結果から、次元軸が味と価格に価値軸があり、商品がそれぞれに分類され、商品の分類グループを確認することができる。しかしながら、本研究のこうした軸の解釈は、分析者の恣意性、解釈性が強くなり過ぎる懸念が生じるかもしれない。特に、本事例で用いた新聞記事の特徴とも言えるが、二次データを用いる社会現象の観察では、多様なデータの適合度を考慮しつつ解釈の容易性を求めることも必要になる (岡太・今泉, 1994) と思われる。

そこで、三次元 MDS による考察も試みた (図 2. II)。二次元 MDS よりも幾分直感的によりわかりやすい図を示している。禁煙言説がどこかに集中して布置されているわけではなく、全体的に広がっている。これは、現実においては相互の影響を無視することはできないが、直接的ではなく、多くの場合、それは間接的に、相手を横目で見ながらであった

ことが推定される。軸を解釈すると、次元1は、「場所」を示し、次元2は、「主体」であり、次元3は、「行為」を示す軸であると推定できるように見える。

岡太・今泉（1994）は、適合度と解釈の容易さをどのように調和させて最適な次元数の布置を解として選ぶかは MDS 利用する際の焦点の一つであり、布置を解釈する際には、次元がどのような意味をもっているか、布置における対象間の位置関係にどのような関係が認められるのか検討する必要であると指摘しているが、本研究の実証分析の結果ではその難しさが明らかになった。

図2 多次元尺度構成法 (MDS)



最小頻出 50 回以上の頻出語を基に KH Coder で筆者作成

## 6.2 クラスタ分析

抽出語が多く、結果として布置の解釈が難しくなる場合、分析者の解釈を軽減する方法として、KH Coder は二次元 MDS 分析で布置された似通った語を、「バブル」と「色」で 2 から 12 の任意のクラスターに分類する機能がある。本研究では、以下に、六つのクラスター分類を求めることを前提に、ユークリッド距離による Ward 法による非階層的クラスター分類の分析結果を示す (図 3. I)。本稿では、モノクロでも分かるように各々のクラスターを枠で括って描記した。

中心に枠 1 のバブルの「禁煙」、「喫煙」、「公共」、「場所」のクラスター、上方に枠 2 のバブルの「対策」、「条例」、「分煙」、「オフィス」、「タクシー」、「職場」、「コーナー」、「駅」、「タイム」、「実施」のクラスター、右方向に枠 3 のバブルの「全面」、「禁煙」、「車両」、「拡大」、「全席」、「航空」、「路線」、「バス」、「店内」、「席」のクラスター、左方向に枠 4 のバブルの「嫌煙」、「運動」、「全国」、「調査」、「世界」、「煙」、「人」、「教育」のクラスター、下方右方向の枠 5 のバブルの「タバコ」、「女性」、「社長」、「会社」、「商品」、「サービス」、「市場」、「パイポ」

のクラスター、下方左方向に枠6のバブルの「情報」、「効果」、「開発」、「生活」、「補助」、「販売」、「トップ」のクラスター、に各々クラスター分けできる。

こうしたクラスター分析は、MDSでの分析者による解釈と比較すると、コンピュータの解析でクラスター分類してくれるため、データ間のカテゴリー構造の探索的解釈の補助的効果に役立つと思われる。

### 6.3 共起ネットワーク分析

MDSと同じ語を用いて共起ネットワークで分析を行った(図3. II)。共起ネットワークでは、次元の位置や語の間の遠・近には意味がなく、線で結ばれているか否かが重要である。描画されている語数(node)は45、共起関係の数(edge=線)は41、密度(density)は.041であった。ここで言う密度(density)とは、実際に描かれている共起関係の数を、存在している共起関係(edge=線)の数で除したものである(樋口, 2014)。Min. Jaccard 係数は、36であった。共起関係の媒介性に基づくサブグラフ検出を行った結果6分類が示された。図3. IIでは、結果が解釈し易いように、Primの方法で共起の強さを考慮した最小スパニング・ツリーのみを描画、出現数が多いほどバブルを大きく、共起度が強いほど太い線で描画した。その示された六つの共起関係を以下に各々観察することにする。

「禁煙」は、「運動」、「世界」、「全面」と直接的につながっている。また、「喫煙」と弱い共起関係で直接的につながっている。「禁煙」の共起関係を見ると、「運動」を媒介にして「嫌煙」、「全国」、「条例」、「教育」と禁煙運動と共起し、一方「全面」を媒介にして「航空」、「路線」、「全席」、「店内」、「バス」、「世界」と今日までつながる公共交通機関や飲食店での禁煙の広がりを示している。

「喫煙」は、「席」、「対策」、「分煙」、「調査」と直接につながり、喫煙に対する対策方法と共起関係にあることを示している。また、「分煙」を媒介にして「場所」、「公共」、「職場」、「コーナー」とつながり、多様な場所で分煙が導入されていることを示している。

「タイム」は、「実施」、「導入」、「駅」、「拡大」と直接つながり、喫煙が空間、時間的に広がっていることを示している。

「サービス」、「タクシー」、「情報」、「ホテル」など何らかの禁煙に関する現象が起きていることが推測できるが具体的内容までは理解できない。

「補助」は、「販売」、「パイポ」と直接つながり、「開発」を媒介にして「生活」、「商品」ともつながっている。禁煙の対策として、禁煙補助剤の商品開発や販売、生活習慣の改善、CMで話題になったパイポの販売など経済活動が生じていることを示している。

「人」、「社長」、「トップ」、「企業」、「タバコ」の共起関係は、企業としての禁煙の取組みを社長やトップ自らの率先している動きとして推測することができる。

本研究の結果が示す禁煙の共起関係は、共起関係の強いグループが複数(六つ)存在し、

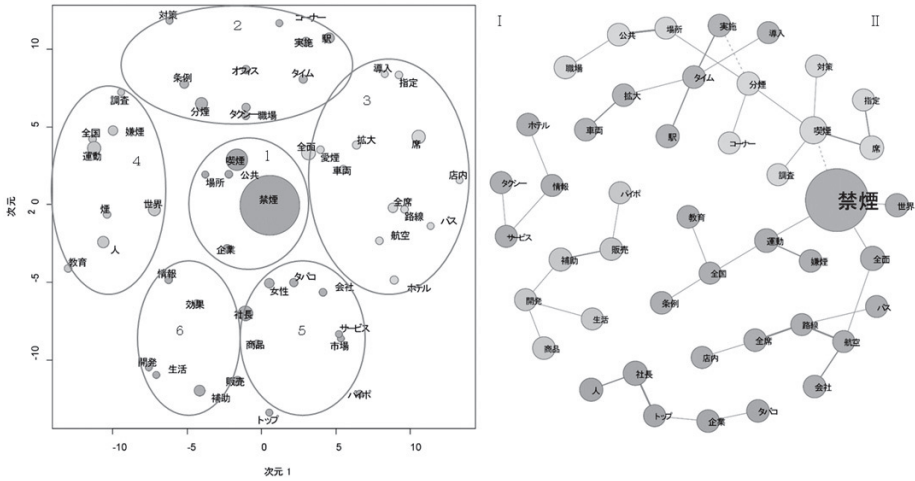
その内の一つのノードが、他の共起ネットワークの一つのノードを媒介としてつながり、直接的に、間接的に多様で複雑な共起関係が社会の現象として起きていることを明らかにしている。

共起ネットワーク分析では、上記のような示されたデータを加工せずに、ノード間の関係の有無、強弱を考察することによって全体構造の概略を把握することができる。これら言説データが示す事実から、禁煙の意識が広がる状況を文脈的に分析・解釈して次のように記述することができる。

1987年から2002年の「禁煙」の共起ネットワーク図(図3. II)のノード間の関係に注目すると、「禁煙」に結びついているのは、「席」や「店内」、「対策」といった単語である。こうした動きは、近年まで続く店内の禁煙化の傾向であり、一部の運動がより一般化し、日常の一部になってきた過程を示している。「航空」会社が「全面的」に「全席」禁煙を「導入」、「実施」している。また、鉄道の「車両」や「バス」、「駅」では「コーナー」や「タイム」など時空間の制限が「導入」「拡大」している。「店内」では、禁煙「席」が「設置」されている。「世界」の国際的な動きも観察できる。また、「嫌煙権」「運動」や禁煙「教育」、また、「条例」が「全国」に拡大している。さらには、「喫煙」の「対策」として「公共」や「場所」、「職場」で「分煙」が「導入」されている。一方で、禁煙「補助剤」や「パイポ」が「販売」され、「商品」、「開発」など経済活動も生まれている。これら言説間の共起ネットワークは、1987年から2002年の間に禁煙に関わる多様なステイクホルダーが顕在化し、また、直接的にあるいは、間接的に相互作用が起き、禁煙の社会的文脈が動態的、制度的に広がっていることを示していると考えられる。

この例示した文脈的な分析・解釈のように、共起ネットワーク分析の特徴である、線で結んで言説間の関係性の有無、及び関係性の強弱を示す表現が、分析者にとっては、その分析結果を理解しやすくさせ、分析者自身よる解釈の度合いは相当程度少なくて済む利点があるように思われる。

図3 クラスタ分析・共起ネットワーク分析



最小頻出 50 回以上の頻出語を基に KH Coder で筆者作成。図3-Iの枠は KH Coder で分析後筆者加筆

## 7. 考察

これまでに、具体的に分析した事例に基づき、先述の社会現象を研究対象とする先行研究が指摘する三つの課題を各々の分析法の当てはまりから比較確認する（表2）。

表2 三つのアプローチの比較

| 先行研究が指摘する課題                | 多次元尺度構成法 (MDS)                                       | クラスター分析  | 共起ネットワーク   |
|----------------------------|--|--|--|
| ステイクホルダーの顕在化、動態化、及び多様性、関係性 | ステイクホルダーの登場は確認できるが関係性の理解は困難                          | ステイクホルダーの顕在化及び多様性について確認できる                           | 多様なステイクホルダーの顕在化を確認できる。関係性の強弱も確認できる。ステイクホルダーを獲得したのか半自然的に生じたのかは別の方法が必要かもしれない |
| マーケティング・システム、マイクロ・マクロリンク   | マイクロ現象がインパクトを与えず新聞記事に登場する機会が少ないのかもしれない。別の方法が必要かもしれない | マイクロ現象がインパクトを与えず新聞記事に登場する機会が少ないのかもしれない。別の方法が必要かもしれない | マイクロ現象がインパクトを与えず新聞記事に登場する機会が少ないのかもしれない。別の方法が必要かもしれない                       |
| 社会的・文化的・規制など制度的特徴          | 関係性の確認が困難なので、制度的な特徴として捉えるのは困難                        | クラスター分類によってある程度は確認することができる。経時的に見ることで制度的特徴も確認できると思える  | 共起ネットワークから相互作用、動態化、制度化する制度的特徴を確認することはできる。経時的に見ることで制度的特徴を確認できると思える          |

筆者作成

課題 a) に対しては、何れのアプローチにおいても、多様なステイクホルダーの顕在化を確認することができるが、その関係性については、既知の通り MDS に比べてクラスター分析と共起ネットワークの方が分析者の解釈の余地が少なく済む。

課題 c) に対しては、MDS では類似性をみるのが主眼であり、それぞれの相互作用に



ついてロジカルに説明することは難しいように思われる。クラスター分析では、バブルと色の分類で確認できる。また、共起ネットワークでは、線で結ばれているか否かで関係性の有無を確認することができる上、線の太さで関係の強弱が表され、多数の主体間の複雑な相互作用を確認することができる。また、松井（2013）、谷口（2017）、織田（2017）らの研究が示すように、経時的に共起ネットワークの変容を分析することで、相互作用の動態化、制度化する制度的特徴を考察することは可能になると考える。

課題 b) のマイクロ・マクロリンクの重要性については、マイクロの言説そのものがデータとして示されていない。禁煙の促進という点でもっとも効果的であると思われるのは、一般的にたばこの値上げである（Brown, 1995; Liang, Chaloupka, Nichter, & Clayton, 2003）。当然、価格政策はソーシャル・マーケティングの一要素とも成すことができる。今回調査対象とした新聞記事の特徴と見ることもできるが、水越・日高（2017）によるダウストリームとアップストリームという観点から見れば、喫煙者に直接働きかけるようなダウストリームの活動は新聞記事上にはあまり反映されない、あるいはさらに踏み込んで、社会現象としてあまり話題にならないのかもしれない。マクロな社会現象を考えると、マクロ・マイクロ・ループの重要性が指摘される（Layton, 2015）。だが、その分析に際しては、マクロな局面は新聞記事を中心にして捉えられる一方で、マイクロな局面や、さらにはそれらのループについては、また別の分析が必要になる可能性がある。

以上、課題 a)、c) の結果で示されたように、マクロ・ソーシャル・マーケティングの探索的実証分析の結果、先述で確認した MDS、クラスター分析及び共起ネットワークの各々の分析の特徴を改めて確認することができた。

## 8. 帰結

本研究では、マクロ・ソーシャル・マーケティング研究の探索的考察として、テキストマイニングの現状を概観した上で、新聞記事の「禁煙」言説を事例にして、そのテキストデータを基に、多次元尺度構成法（MDS）とクラスター分析及び共起ネットワークの分析による結果の示し方について実証分析を通じて比較確認してきた。社会現象を捉えようとする先行研究が指摘する三つの分析課題視点で見ると、多量で多様、かつ複雑な因子の分析から社会現象を分析するには共起ネットワーク分析が分析者にとって利用しやすい結果を示していることが改めて確認された。特に、ステイクホルダーの顕在化、動態化や、社会的・文化的・規制的な相互作用、制度化的特徴を一定程度コンピュータの機械的解析で結果を示してくれることが確認できた。

これまで多くの研究が、マクロな意思決定の過程についてあまり注目してこなかった。その理由の一つとして、具体的な分析方法の問題があった。ソーシャル・マーケティング

研究が近年注目するアップストリームや、あるいはより広いマーケティング・システム、マクロ・ソーシャル・マーケティングといった観点は、いずれもマクロな現象を研究対象にしようとしている。こうした課題に対して、本研究では、質的データの分析方法を確認した上で、各々の分析方法を実証的に確認する作業を行った。具体的には、新聞記事の「禁煙」言説という限定的な資料を対象としながら、テキストマイニングの支援ツールであるKH Coderを使って、多次元尺度構成法、クラスター分析及び共起ネットワークの方法で分析を行い、分析手法の結果の示す違いを実証的に再確認することができた。

一方、それでもなお、本分析で解釈したように、比較的明瞭な結果を示した共起ネットワーク分析でも、それ単独の解釈では、恣意的であるとの批判が生じる懸念の余地が残る可能性はあるかもしれない。前述の共起ネットワークによる文脈的解釈の信頼性を検討する上で、この時期の新聞記事を概観すると次のような記事を確認することができる。1992年には、国際民間航空機関が、「国際線旅客機の全面禁煙について各国政府が必要な措置をとる」との決議を採択したことが契機となって、国内線では1998年から航空3社が全席禁煙となり、1999年には、国際線も全席禁煙になった。JRは「禁煙車」表示から「喫煙車」表示に変えて分煙化を進め、また、民間企業のオフィスでもOA化の普及に伴い煙に弱いOA機器保護のためもあり、禁煙、分煙の取り組みが進んだ。2002年には、厚労省の審議会が、たばこ消費抑制を求めたり、厚労省が全国の自治体に分煙効果の徹底を求める通知を出したりしている。こうした異なる分析方法による取り組みは、テキストマイニングによる探索的考察と、記述による検証という各々の役割において相補的であると思われる。

今後、本研究で分析手法として取り上げなかったトピックモデルでの分析や、より信頼度の高い分析結果を示す可能性を高めるために、テキストマイニングと併せて、インタビューや資料からの記述等とのトライアングレーションに取り組む必要があるだろう。

## 謝 辞

本稿の審査過程では、学会委員の竹内祐介先生、及び匿名のレフェリーの先生から貴重なご示唆をいただきました。また、本稿の作成にあたっては、指導教員である水越康介先生から懇切丁寧なご指導を賜りました。ここに記して御礼申し上げます。

## 注

- 1 2018年9月17日アクセス。
- 2 男性の喫煙者率が75.1%（1976年）～29.7%（2016年）であるのに対して、女性の喫煙者率は16.2%（最高年1978年）～9.7%（最低年2015年）であるように全体的な比率で見ると、喫煙という行為は男性に典型的な行為とみることができ、女性の変化率が小さいこともあり、

男性のみの変化で相関関係を観察した。

- 3 2011年に「禁煙」の出現数が激減しているのは、2011年3月11日の東日本大震災の影響と思われる。ちなみに震災関連の言説の2010年と2011年の日経四紙の出現数は次の通り。「震災」531件→34794件、「津波」272件→5374件、「原発」1262件→10362件、「停電」238件→2572件。まさに桁違いの出現数の増加で新聞が震災関連記事で埋め尽くされていたことがわかる。

## 参考文献

- Andreasen, A.R.(1994). "Social Marketing: Its Definition and Domain," *Journal of Public Policy and Marketing*, 13(1), 108-114.
- Brown, A. B. (1995). "Cigarette Taxes and smoking restrictions : Impact and policy implications" *American Journal of Agricultural Economics*, 77(4), 946-951.
- 藤井美和・小林孝司・李政元(2005).『福祉・心理・看護のテキストマイニング入門』,中央法規
- French, J. and Gordon, R. (2015). *Strategic Social Marketing*, Sage.
- Giesler,Markus(2012).“How Doppelgänger Brand Images Influence the Market Creation Process:Longitudinal Insights from the Rise of Botox Cosmetic” *Journal of Marketing*, 76(6),55-68.
- 林 拓也 (2015).「社会調査データを用いたポジショニング分析の基礎」『奈良女子大学文学部紀要』, 1-68.
- 日高優一郎・水越康介(2014).「社会的企業家を捉える論理—関係性概念にもとづく公共・非営利マーケティング研究の再検討—」『流通研究』, 16(3), 69-93.
- 日和恭世 (2013).「ソーシャルワーク研究におけるテキストデータ分析に関する一考察」『評論・社会科学』 106.141-155.
- Humphreys,Ashlee(2010a),“Megamarketing: The Creation of Markets as a Social Process,” *Journal of Marketing*, 74(2),1-19.
- Humphreys,Ashlee(2010b), “Semiotic Structure and Legitimation of Consumption Practices: The Case of Casino Gambling,” *Journal of Consumer Research*, 37(3), 490-510
- 樋口耕一 (2014).『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』, ナカニシヤ出版
- 稲葉光行・抱井尚子 (2011).『「質的データ分析におけるグランデッドなテキストマイニング・アプローチの提案」—がん告知の可否をめぐるフォーカスグループでの議論の分析から—』『政策科学』, 18(3), 255-276.
- 岩田具治 (2015).『トピックモデル』, 講談社
- Jick.Todd D.(1979). Mixing Oualitative and Quantitative Methods:Triangulation in Action, “*Administrative Science Quarterly*, 24(4), 602-611
- 勝又壮太郎・西本彰宏 (2016).「市場創造と普及のダイナミクス」『組織科学』, 5(2),1-9.
- Kennedy,A.M.(2015).“Macro-socialMarketing” *Journalof Macromarketing*,36(3), 354-365.
- Kotler, Philip (1986), “Megamarketing,” *Harvard Business Review*, 64(2), 117-25
- Kotler, P. and Lee, N. (2009). *Up and out of poverty: The social marketing solution*. (塚本一郎訳『コトラーソーシャル・マーケティング』丸善, 2010)
- Kotler, P. and Roberto, E. L. (1989). *Social Marketing*, The Free Press. (井関利明監訳『ソーシャル・マーケティング—行動変革のための戦略』, ダイヤモンド社, 1995)
- 松井 剛 (2013).『言説とマーケティング,「癒し」ブームの消費社会史』, 碩学舎
- Matsukawa,H.,Arai,Y.,Iwasaki,C.,Kinjyo,Y.and Hotta,H.(2015). “Semantic Classification of Text Messages Using the Concept of Community in Social Network Analysis,” *IADIS International Conference e-society 2015, Mederia, Potugal*, 340-342.
- 松川秀哉・大山牧子・根岸千悠・新居佳子・岩崎千晶・堀田博史 (2017).「トピックモデルを用いた授業評価アンケートの自由記述の分析」『日本教育工学論文誌』 41(3), 233-244.
- 水越康介・コールパッサ・フローリアン (2015).「イクメン言説の消費」『マーケティング・ジャーナル』

34 (4) ,86-101.

- 水越康介・日高優一郎 (2017). 「ソーシャル・マーケティング研究における理論的視座の再検討」『JSMD レビュー』1(1), 33-39.
- 中野 勉 (2017). 『ソーシャル・ネットワークとイノベーション戦略』, 有斐閣
- 中山厚穂・横山暁 (2010). 「MDSを使って使って使い倒す! MDS 入門から非対称 MDS 実習まで」. 日本行動計量学会第 13 回春の合宿セミナー資料, 御殿場, 1-18.
- 岡太彬訓 (2002). 「社会学におけるクラスター分析と MDS の応用」『理論と方法』17(2),167-181.
- 岡太彬訓・今泉忠 (1994). 『パソコン多次元尺度構成法』, 共立出版
- 織田由美子 (2017). 『「消費の脱ステイグマ化」－「婚活」を事例として－』, 一橋大学博士論文.
- Layton, R.A., (2015). "Formation, Growth, and Adaptive Change in Marketing Systems" *Journal of Macromarketing*, 35(3), 302-319.
- Liang, L., Chaloupka, F., Nichter, K. & Clayton, R. (2003). Prices, policies, and youth smoking, *Addiction*, 98(1), 105-122.
- 戈木クレイグヒル慈子 (2014). 「グランデッド・セオリー・アプローチ概論」『KEIO SFC JOURNAL』,14(1), 30-43.
- 坂田隆文 (2001). 「マクロ消費の意義と展望」日本商業学会『流通研究』4(2), 49-60.
- 高井文子 (2017). 「模倣・追従の二面性:日本のオンライン証券市場黎明期における企業間競争の実証的分析」『組織科学』51(1). 46-57.
- 谷口 涼 (2017). 『シンボルを用いた資源獲得の成功による資源配分の失敗:「バイオマス・ニッポン総合戦略」の事例』『組織科学』50(4), 66-81.
- 与謝野有起ほか数理社会学会 (2006). 『社会の見方, 測り方』勁草書房