

# 質的データ分析におけるグラウンデッドな 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 ——

稲葉 光行・抱井 尚子

## 要旨

本研究では、コンピュータを用いることで、解釈者としての研究者の感受性を活かしつつ、同時に客観性を確保した形で分析プロセスと結果の提示を行うための、新しい質的データ分析手法の確立を目指している。本稿ではまず、質的研究にコンピュータを用いることについての様々な議論を整理する。次に、これらの議論を踏まえて、研究者自身による分析とコンピュータによる分析の両方の利点を生かした新しい方法として、「グラウンデッドなテキストマイニング・アプローチ」手法（GTMA）の枠組みを提案する。そして、がん告知の可否をめぐるフォーカスグループの逐語録を題材とし、研究者自身が行った分析と、GTMAによって分析した結果を比較検討する。最後に、質的研究に関する評価規準の視点から、GTMAの特徴と限界について議論する。

## Keywords

Qualitative Data Analysis, CAQDAS, Grounded Theory Approach, Text Mining

## I. 背景

近年の情報通信技術の普及に伴い、研究活動のあらゆる側面でコンピュータが活用されはじめている。この傾向は、自然科学や工学の領域において顕著であるが、最近では、人文・社会科学の領域でも、コンピュータ技術を全く用いない研究活動の実践はほとんどあり得ないと言える状態になってきている。さらに、人間としての研究者自身の主観性や感受性が特に重視される質的研究においても、様々なコンピュータ技術が用いられるようになってきた。そしてそれに伴い、コンピュータを質的研究に用いることについての利点と欠点について様々な議論がなされている。ここではまず、質的研究へのコンピュータ利用に関して、これらの2つの対立する立場での議論を整理する。

## I. 1 質的研究におけるコンピュータ利用の利点

Miles & Huberman (1994) は、質的研究の実践において、コンピュータが役立つ作業として、以下の点を挙げている。

1. フィールドノーツをとる
2. フィールドノーツを転記・清書する
3. 編集：フィールドノーツを修正・詳述・改訂する
4. コード化：テキスト・画像・音声・動画のセグメント（意味的なまとまりを持った一部分）にキーワードやタグをつけて、あとから検索しやすいようにする
5. 貯蔵：組織化したデータベースにテキストを保存する
6. 検索：テキストの関連するセグメントを突き止めて閲覧ができるようにする
7. データのリンク：関連するデータのセグメントを互いに結合し、情報のカテゴリー、クラスター、ネットワークを作成する
8. メモ書き：より深い分析への基礎として、データ、理論、方法のいくつかの面について反省的なコメントを書き込む
9. 内容分析：単語や語句の出現頻度、連続の程度、位置を数え上げる
10. データ表示：選択・縮約したデータを行列やネットワークといった凝縮され組織された形で表示し、視認しやすくする
11. 結論の導出と確証：表示したデータの解釈を助け、発見したことを検証・確証する
12. 理論構築：発見した内容に対する体系的で概念としてのまとまりを持った説明を発展させ、また仮説を検証する
13. 可視化：発見した事実や理論を表現する図表を作成する
14. レポート作成：中間の報告書、最終報告書を書く

注：訳出は『質的研究ハンドブック』第3巻・第7章「ソフトウェアと質的研究」（黒沢学訳）を参照。

なお、ここでは「text」を「テキスト」と訳しているが、本稿では「テキスト」という訳語を用いる。

さらに近年は、特に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。QDA ソフト、あるいは CAQDAS 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、グラウンデッド・セオリー・アプローチ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。これらのソフトウェアは、テキストデータ中の選択された部分に簡単にコードを付けるための機能、作成済のコードを効率良く検索・再利用するための機能、コードとカテゴリーとのリンク関係を管理する機能、あるいは、コードやカテゴリーの関係性を視覚的に提示する機能、などを備えている。Weitzman (2000) は、このようなソフトウェアを使った質的データ分析において真に期待できることとして、consistency (整合性)、speed (速さ)、representation (表

現)、consolidation (統合性)、という4点を挙げている。

このようなコンピュータを用いた質的データ分析の特長は、最近、マーケットの動向調査やCMC (Computer Mediated Communication) に関するテキスト分析において普及しつつある、「テキストマイニング」(text mining) と呼ばれる分析手法によって、さらに効率化できる可能性が見え始めている。テキストマイニングは、コンピュータを利用することで、テキストデータの集合から新しい知識を見出そうとする分析手法である (Hearst, 1999)。一般的なテキストマイニングでは、最初に、テキストに含まれる品詞を識別する形態素解析 (分かち書き) や、語句の間の修飾関係を分析する構文解析 (係り受け解析) などの言語処理が行われる。その後、多変量解析による可視化や、辞書に基づくパタンマッチングによって、テキストデータの中から、何らかの意味のある情報を取り出す過程が支援される。

これまで、コンピュータを用いて日本語を高速に解析する処理はコストがかかるものであったが、コンピュータの高速化と低価格化によって、日本語のテキストを解析し、形態素解析や構文解析を行うソフトウェアを安価に利用できるようになってきた。そして、フリーソフトウェアを用いて、質的分析の視点からテキストマイニングを実践するための入門書も発売されている (藤井美和・李政元・小杉考司, 2005; 松村真宏・三浦麻子, 2009)。

## 1. 2 質的研究におけるコンピュータ利用の問題点

このように、コンピュータを用いた質的データ分析が普及する環境が整って来ているが、一方で、ワープロやスプレッドシート、データベースといった汎用的なソフトウェアを超えて、質的データ分析のためにコンピュータを用いる例はまだ多くない。

その理由としては、まず、研究者のコンピュータスキルやソフトウェア導入の予算といった、方法論と直接関係のない課題があると考えられる。またそれに加えて、質的研究という方法論特有の課題があると考えられる。つまり、どのようなツールを使っても、質的研究の本質的な部分は変わらないという立場を取る質的研究者が多いことが、コンピュータを用いたデータ分析の普及を阻害していると考えられる。例えば Flick (1995) は、QDA ソフトを使ってコード化を行う研究者は、「ワープロを使って執筆を行う著作者のようなものと考えたほうがよい」(小田他訳, p.309) と述べている。

QDA ソフトや CAQDAS と比較して、より高度な言語処理や可視化機能を備えたテキストマイニング用ソフトウェアを利用したデータ分析に関しても、様々な問題点が指摘されている。藤井ら (2005) は、テキストマイニングという手法自体の課題として、テキストデータが持つ曖昧さをコンピュータで処理することが困難であること、テキストデータが回答者の意図をそのままの形で反映する保証がないこと、そして、テキストデータの真の意味は研究者が解釈する必要があること、を挙げている。

さらに、質的データ分析にコンピュータを用いることについての、根源的な危険性を指摘する研究者もいる。例えば Seidel (1991) は、研究者が、コンピュータを質的データ分析に用いる際に、研究上の視点よりも、扱うことができるデータの量に魅力を感じてしまい、その結果、

データと研究者を遠ざけてしまう可能性を指摘している。また Becker (1993) は、データに根ざした理論構築をめざすグラウンデッド・セオリー・アプローチの実践と、CAQDAS を利用して研究することとの間には距離があることを指摘している。そして、特にグラウンデッド・セオリー・アプローチを実践しようとする初学者に対しては、少量のテキストデータを、自分自身の手で深く分析することを推奨している。

### I. 3 コンピュータを用いたグラウンデッドなテキスト分析の可能性

このような背景から、筆者らは、研究者自身とデータの間の密接さを確保した「グラウンデッドな」分析と、コンピュータを用いることによる、整合性、速さ、表現、統合性、という Weitzman (2000) が述べた利点を生かした、新しい質的データ分析の手法の確立に取り組んでいる。言い換えれば、Willig (2001) が述べている、「小文字の q」(small q)、つまりテキストデータを出現頻度などに基づいて数量化することで質的な視点からの考察を試みるだけでなく、「大文字の Q」(big Q)、つまりテキストで語られている現象自体への深い理解を目指すための分析手法の確立を目指している。

以下では、このような取り組みの一つとして、筆者らが提案する新しいデータ分析手法を紹介すると共に、研究者自身による分析と、この提案手法による分析の事例について、比較検討を行う。

## II. グラウンデッドなテキストマイニング・アプローチの概要

ここでは、研究者自身による分析とコンピュータによる分析の両方の利点を生かした、新しい質的データ分析の手法として、筆者らが提案する「グラウンデッドなテキストマイニング・アプローチ」手法 (GTMA) の概要について述べる。

図 1 は、この手法の枠組みと実践プロセスを述べたものである。この図で示したように、この手法は、テキストの読み込み (ステップ①)、グラウンデッド・セオリー・アプローチに基づくデータ分析 (ステップ②)、テキストマイニング手法に基づくデータ分析 (ステップ③)、GTMA に基づく統合的分析 (ステップ④) を含む、一連のサイクルである。

また、ステップ①からステップ④を繰り返す中で、テキストデータの読み込みによって研究者が作り出した印象と、それぞれの分析結果の間に、矛盾点・疑問点があるか否かという確認が行われる。これらの点が解消されていないと判断されれば、ステップ①から④のサイクルの一部、または全体のプロセスを再度繰り返す。その際、ステップ①では、テキストデータ化の過程での間違いなどを修正する作業が行われる。ステップ②では、コード化・カテゴリー化の方法を再検討し、場合によってはそれらの作業をやり直す。ステップ③では、辞書や語句パターンの再定義、あるいはパラメータを再設定した上で、クラスターや関係性を再度視覚化する作業などが行われる。これらを繰り返し、全体として矛盾点・疑問点がなくなったと判断されれば、それまでに得られた知見を新しい知識としてまとめる作業が行われる。

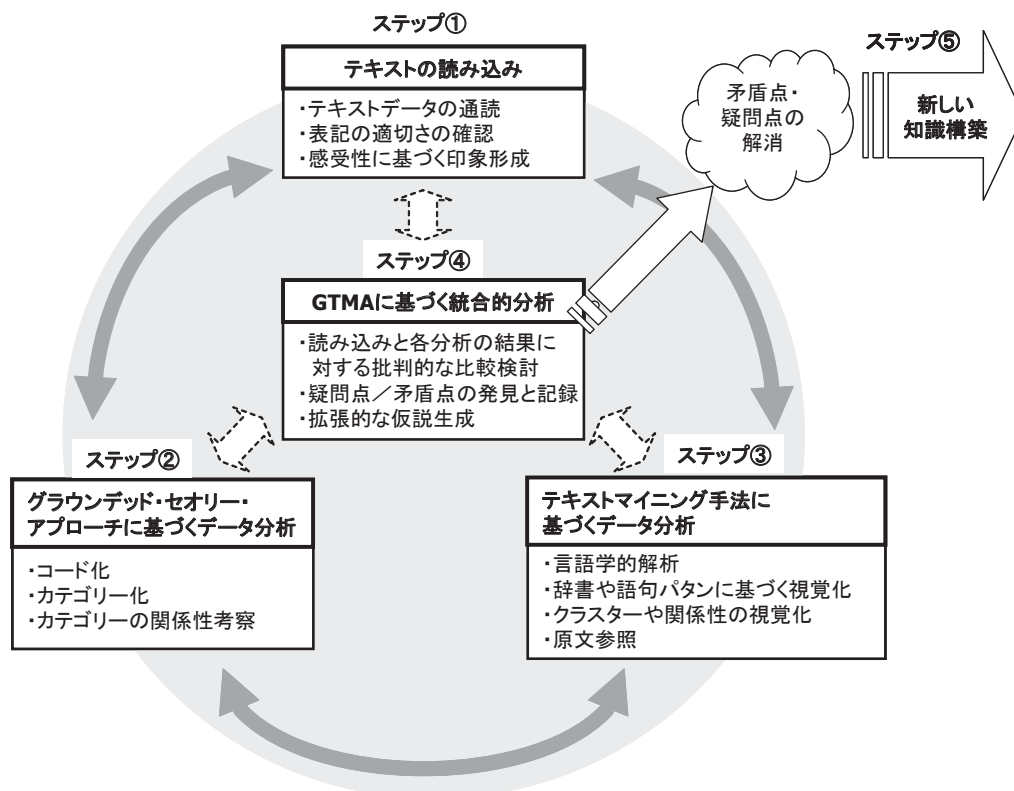


図1：グラウンデッドなテキストマイニング・アプローチ（GTMA）のプロセス

この手法のデザインの最も大きな特徴は、コンピュータによる言語処理や統計的分析による可視化と、研究者自身による解釈を併存させていることである。言い換えれば、本手法の特徴は、テキストに含まれる質的な情報を定量化（quantitization）する作業と、量化されたデータを質的な視点で解釈する定性化（qualitization）の両方が交互に行われるというもので、TeddlieとTashakkori（2009）が「変換型ミックスデザイン」（conversion mixed design）と呼ぶアプローチにあたる。これによって、研究者自身の主観性・感受性に基づく深い理解と、客観性を持った形での結果提示という、一見相反する2つの方向性を統合することが可能となる。

以下では、図1で示したそれぞれのステップの概要を述べる。

## II. 1 テキストの読み込み

研究者はまず、分析対象のテキストデータを通読することによって、自身の感受性に基づいて、テキストデータ全体に対する印象を形成する。

またこの過程で、テキストデータ中の詳細な表現や言い回しについての確認を行い、字面上の明らかな間違いがある場合は、どのように修正するのかを検討する。例えば逐語録を読み込

む場合は、ICレコーダーなどに録音された発言が適切にテキストデータ化されているのか、つまり、間違っただけの入力や、かな漢字変換のミス、あるいは変換された表現の揺れなどがないかを確認する。ネット上の掲示板やブログ、あるいは携帯メールでは、コンピュータで適切に処理できない言語表現、例えばスラングや絵文字などを、どのような形でコンピュータに処理させるのかを検討する必要がある。質問紙における自由記述欄をワープロなどでテキスト化した場合は、入力ミスや変換ミスの可能性もあるが、そもそも元の記述において間違っただけの漢字や言葉遣いが用いられている可能性もあり、それらをどのように扱うのか（そのままの形で分析するのか、あるいは明らかな間違いは修正するのか）を検討し、場合によっては、内容をそのままに、表記だけを修正する作業が必要になる場合もある。

テキストデータに対して、内容を変更しない範囲での修正を加える作業は、コンピュータを用いた言語処理を行う場合には避けられないことである。コンピュータの言語処理では、音声上は同じであっても、「書く」と「かく」は全く異なるデータとして扱われる。あるいは、「パソコン」と「PC」は、意味的には同じ単語であっても、コンピュータにはそれが同じ事物を指していることが理解できない。当然ながら、この作業は、研究倫理の視点から、相当慎重な配慮を持って進める必要がある。なぜなら、発言者が述べた、あるいは記録者が記した内容自体に関わる変更をした場合は、もはやそのテキストデータは、正当な研究活動の分析対象として用いることはできないからである。

## II. 2 グラウンデッド・セオリー・アプローチに基づくデータ分析

グラウンデッド・セオリー・アプローチは、1960年代に社会科学において主流であった実証主義／ポスト実証主義的認識論に基づく仮説演繹的な量的研究への問いかけとして誕生した、帰納的な仮説構築型研究アプローチである。この研究アプローチは、個別の事例を分析する「中範囲理論」を構築することを主張したコロンビア大学の Robert K. Merton のもとで量的研究のトレーニングを受けた Barney Glaser と、人々の主観的視点の探求を志向するシンボリック相互作用論、問題解決を一義的な研究目的とする米国プラグマティズム、フィールド研究といったシカゴ学派社会学の伝統のもとで質的研究を行う Anselm Strauss の医療社会学の共同研究から誕生した。グラウンデッド・セオリー・アプローチは、特定の社会現象に関する構造とプロセスを理解することを重視し、この手法によって生み出された理論はグラウンデッド・セオリー、つまり「データに根ざした理論」とよばれる。

*The Discovery of Grounded Theory* (1967) で Glaser と Strauss によって最初に提案された古典版以来、グラウンデッド・セオリー・アプローチは、Glaser 版、Strauss & Corbin 版、Charmaz による構成主義版など、複数のヴァージョンへと発展してきており、それぞれのもつ認識論的立場から独自のコード化ルールや理論の役割を提案している。中でも Charmaz (2006) により近年主張されているグラウンデッド・セオリー・アプローチは、外在する客観的現実の「発見」を志向するこれまでのグラウンデッド・セオリー・アプローチとは一線を画し、いかなる理論も研究者とデータの相互作用によって「構築」される、複数存在し得る現実の「ひとつ」

に過ぎないという前提を取る。

本稿で提案する GTMA では、特に構成主義版グラウンデッド・セオリー・アプローチ (Charmaz, 2000, 2006) のコード化プロセスを想定している。構成主義版グラウンデッド・セオリー・アプローチのコード化では、少なくとも2つの段階、つまり「初期段階のコード化 (initial coding)」と「焦点化のためのコード化 (focused coding)」が行われる。前者は、それぞれの単語、行、出来事などに名前をつける作業である。後者は、「データを分類し、総合し、統合し、整理するために最も有意義で頻出するコード」(抱井・末田監訳, p.55) を見つけ出す作業である。ここでのコード化のプロセスは、データとのやりとりという「相互行為的」なものである。つまり、研究者と研究参加者が同じ視点や世界観を共有していると思うのではなく、研究参加者の「独自の現実の見方に順応していく」(前掲書) という、構成主義的な視点が重視される。そして、このように生成される焦点化コードを、比較・分類・整理することによって、研究者は、より高次のカテゴリーをつくり出していく。

### II. 3 テキストマイニング手法に基づく分析

一般的なテキストマイニングでは、テキストデータに対して、まず形態素解析や構文解析を含む言語学的解析を行う。次に、多変量解析の手法によって、クラスターの分析や、属性と語句の関係性の視覚化を行う。さらに研究者が着目した語句については、コンコダンス機能(文脈の中で特定の単語がどう使われているのかを分かりやすく表示する機能)などを用いて原文を参照することで、テキストデータに含まれている重要な知識情報を掘り起こすことを目指す。

本稿で提案する GTMA では、構成主義版グラウンデッド・セオリー・アプローチに基づくコード化の後に、テキストマイニングの1手法である、コンピュータを用いた「計量テキスト分析」と呼ばれる作業を行うことを想定している。計量テキスト分析とは、「計量的分析手法を用いてテキスト型データを整理または分析し、内容分析 (content analysis) を行う方法」(樋口, 2006) である。つまり計量テキスト分析は、分析者が定義したパターンによってテキストを分類する内容分析と、多変量解析のような関係性の可視化という2つの手法を統合したものである。樋口 (2004) は、前者を「Dictionary-based アプローチ」、後者を「Correlational アプローチ」と呼び、計量テキスト分析は両者を統合したものであるとする。

樋口は、計量テキスト分析を支援するツールとして、KH Coder というフリーソフトを公開している。KH Coder は、有料のテキストマイニングツールの多くが備えている構文解析機能をサポートしていないが、多変量解析や、語句の共起ネットワーク作成など、優れた可視化機能を備えている。さらに KH Coder では、単語毎のコード化の結果と焦点化コードの関係を、「コーディング・ルール」として定義し、その定義に基づいて、多変量解析やネットワーク図としてコードの関係性を可視化する機能を備えているため、本アプローチの支援ツールとして適している。

### II. 4 GTMA に基づく統合的分析

前述した2つの分析の後、それらの結果に矛盾点・疑問点がなく、テキストで語られている

現象を適切に表しているかどうかを比較検討する作業が行われる。もし何らかの矛盾点や齟齬があれば、テキストマイニングにおいて、分析対象とする品詞の範囲の変更や、辞書の見直し、あるいは別の種類の可視化などを行う。さらに、問題点が解消されない場合は、構成主義版グラウンデッド・セオリー・アプローチのコード化の段階まで遡り、分析をやり直す必要がある。このようなプロセスは、研究者自身が感じる矛盾点や疑問点が解消されるまで継続される。そして、テキスト自体、およびテキストで語られている現象に対する理解の深化を目指すこのようなサイクルから、分析前には想像もしていなかった新しい知識が見出される可能性が開かれる。

なお、一般的にグラウンデッド・セオリー・アプローチでは、継続的比較法による分析が「理論的飽和」(Glaser & Strauss, 1967) —つまり、これ以上新たにデータを収集してもカテゴリーの多様性の幅が広がることはないと判断される状態—に達するまで繰り返されるが、GTMAでは、修正版グラウンデッド・セオリー・アプローチの考案者である木下(2003)が主張するように、あらかじめデータ収集の範囲を決めて分析に臨むことを前提とする。

GTMAにおいて重要な点は、たとえ限定された範囲のデータであっても、人間がテキストを理解する方法と、コンピュータがテキストを解析する方法は根源的に異なっており、両方の結果を照合することで、様々な矛盾や疑問点が次々と浮かび上がってくる可能性が高いということである。さらに、同じ人間によるテキスト理解であっても、全体を概観する場合と、詳細部分に焦点を当てる場合とでは、異なる結果が得られる可能性も十分にある。つまり、研究者がテキストを読み込むことで生成される印象、コード化・カテゴリー化の結果として見える重要なポイント、コンピュータによる可視化という3つの結果の間には、何らかの矛盾が生じる可能性が高く、そこから研究者にとっての様々な疑問点が浮かび上がってくる可能性も高い。このような矛盾点や疑問点についてはメモ書きを通して記録・考察し、それらを解消するために、改めて原文を参照し、コードやカテゴリーを再検討し、テキストデータの可視化を再度試みるという作業を繰り返すことで、テキストで語られている現象に近づくことができる可能性が出てくるのである。

### Ⅲ. 質的データ分析の事例比較

#### ～がん告知の可否をめぐるフォーカスグループの分析～

以下では、本論文における考察のための基本事例として、筆者の一人が行った、研究者自身による質的データ分析の事例(Kakai, 2009)について概説する。次に、この研究で用いられたものと同じデータ(逐語録)に対して、本稿で提案するGTMAを適用した結果について述べる。

#### Ⅲ. 1 研究者自身による質的データ分析の事例

##### Ⅲ. 1. (1) 研究目的とデータ収集

ここでは、本論文における考察のための基本事例として、筆者の一人が行った、研究者自身



による質的データ分析の例（Kakai, 2009）を紹介する。この研究全体の目的は、がん告知を題材として、日本人の倫理的意思決定が、社会文化的状況の中でどのように構築されるかを明らかにすることである。本研究では、祖父母、両親、配偶者、そして未成年の子どもに対してがん告知を行うことが正しいかどうかという判断を求められるという文脈において、以下の3つの研究設問を設定した。

- (a) 日本人はどのような倫理的な声を持ち、どのような意思決定をするのか
- (b) 日本人の中で、倫理的な声はどのように異なるのか
- (c) 個人が他者の言葉によってどのように影響を受けるのか

この研究設問に関するデータ収集のために、フォーカスグループでのインタビューを設定した。フォーカスグループへの参加者は、1997年の秋学期に、ハワイ大学の短期言語教育プログラムで英語を履修していた5人の日本人（23歳から42歳まで）である。このフォーカスグループでの議論は約2時間に渡って行われた。

フォーカスグループでは、以下の5つの状況を設定し、それぞれにおいて、医師からのがん告知を希望するかどうかという点に関する議論を行った。

状況1：

自分の祖父母が末期がんであるとしたら、そのことを本人（達）に告知すべきだと考えるか

状況2：

自分の両親が末期がんであるとしたら、そのことを本人（達）に告知すべきだと考えるか

状況3：

自分の配偶者が末期がんであるとしたら、そのことを本人に告知すべきだと考えるか

状況4：

自分の子ども（未成年）が末期がんであるとしたら、そのことを本人に告知すべきだと考えるか

状況5：

自分自身が末期がんであるとしたら、そのことを自分に告知して欲しいと考えるか

フォーカスグループの後に作成した逐語録の分析では、5つの状況に関する参加者の個々の発言を、研究者自身が解釈し、状況毎、個人毎、および相互作用という3つの視点から分析した。

### Ⅲ. 1. (2) 研究者自身によるデータ分析のプロセス

以下では、この事例で、研究者自身が行ったデータ分析のプロセスについて概説する。

まずこの研究で行われたフォーカスグループ・インタビューの中で、「祖父母への告知」という状況での、参加者の発言の例（数字は筆者）を抜粋する。

- (i) 祖母の性格から、末期癌だと知ったら、このあとはもう死ぬだけだと思って、ただ悲しい日々を過ごす、だろうと思うので、
- (ii) 残りの生活を、楽しく、過ごしてもらうために、
- (iii) ああ、伝えないで、もらいたいと思います。

<発言者：Yuka (仮名)、状況：祖父母への告知>

(i) の発言から、発言者は、自分の祖母が死の恐怖に対処できるほど強くないと考えていることがわかる。(ii) では、そのような祖母の感情に対する配慮の気持ちが示されている。(iii) では、前の発言の帰結として、医師からのがん告知を望まない、つまり真実を隠すという意味決定が示されている。

参加者への各発言に対してこのような分析を継続した結果、5名中4名の参加者に共通する態度として、患者の性格と年齢から、患者に心的負荷をかけるべきではないという心情と、その帰結として、「患者にがん告知をすべきでない」という意思決定に至るというプロセスが明らかにされた。

この研究でのデータ分析は、一般的な質的データ分析に従ったものである。つまり、研究者自身がテキストを繰り返し通読し、研究者自身の常識と感受性を働かせることで、コード化および意味の類似したコードをまとめるカテゴリー化を通して、より抽象化された上位のカテゴリーとそれらの間の関連性を明らかにしている。

さらに、結果の考察では、倫理的意思決定や社会文化的アプローチに関する研究者の知識を用いて、研究参加者らの態度と意思決定プロセスを、学術的な議論の文脈において理解することを試みている。

### Ⅲ. 1. (3) 研究の結果と考察

研究設問 (a) については、参加者毎の発言の分析から、知る権利の尊重、人生のゴールを達成しようとする他者への敬意、などを含む9つのカテゴリーが見出され、さらにそれらを解釈・整理することによって、告知という意思決定に関わる3つの概念が浮上した。1つ目は、個人の知る権利や法的枠組みに焦点を当てた「正義の声」である。2つ目は、心理的・感情的な側面に配慮する「ケアの声」、あるいは「思いやりの声」である。第3の声は、家族に対する社会的・経済的責任を重視する「実用主義的な声」と呼べるものである。

研究設問 (b) については、質問毎の議論の分析から、設定された状況によって上記の3つの声を使い分けられていることが明らかとなった。がん告知に賛成か反対かという議論の中で現れる声の多様性は、患者の年齢、性別、あるいは社会的・経済的責任などと深い関連性があった。例えば患者が、高齢者、女性、子ども、といった社会的弱者であるとき、研究参加者は、「ケアの声」を使う傾向があった。一方で、配偶者、または本人ががん患者だと仮定したとき、ほとんどの研究参加者は、「正義の声」によって、真実を伝える（または伝えてほしい）という考え

を述べた。そして、両親のように、経済的・社会的責任がある他者へのがん告知については、「実用主義的な声」を頻繁に使用していることが明らかにされた。

研究設問 (c) については、社会文化的アプローチや道徳性発達といった学術的な枠組みから解釈が行われている。例えば、フォーカスグループにおける一部の参加者の態度の変容は、日本社会全体が持つ「権威的なディスコース (authoritative discourse)」と、研究参加者自身が持つ「内的に説得力のあるディスコース (internally persuasive discourse)」(Bakhtin, 1981) との緊張関係の結果として捉えられている。

### Ⅲ. 2 GTMA によるデータ分析の事例

以下では、上で紹介した研究 (Kakai, 2009) で用いられた逐語録を、GTMA によって分析した結果を示す。なお、分析手法の比較のため、前述の研究と同じ研究課題を設定した。

表 1：構成主義版グラウンデッド・セオリー・アプローチに基づくコーディングの例（一部）

質問	名前	内容	初期コード		焦点化コード
			(単語毎)	(行毎)	
1_ 祖父母	ユカ	祖母の性格から、末期癌だと知ったら、このあとはもう死ぬだけだと思って、ただ悲しい日々を過ごす、だろうと思うので、残りの生活を、楽しく、過ごしてもらうために、あの、伝えなくて、もらいたいと思います。	性格、悲しい、楽しく、伝えない	悲しませないために、告知をしない	<u>心的要因</u> から、 <u>告知しない</u>
1_ 祖父母	タカ	僕のほうはですね。えと、僕の祖母はやっぱり同じなんですよ。気が弱そうですから、きっと聞いたら、悲しくなって、逆に悪化するだけだと思っんですね。	気が弱そう、悲しく、悪化する	悲しませないために、告知をしない	
1_ 祖父母	マキ	私は告知は望まないほうですね。うちの場合は、うちの祖父は、既にもう 90 近くです。まあ、今更、そういうショックを受けさせて、そのショックのまま死なせるよりも、すこし早まるかもしれないけども、もうそのまま、楽しいまま、死んでもらったほうが、幸せなんじゃないかなあと。	90 近く、ショック、楽しいまま、幸せ	余生を楽しいまま過ごしてもらうために、告知をしない	<u>年齢的要因</u> から、 <u>告知しない</u>

注：表中の名前は仮名である。なお、Kakai (2009) では、研究参加者の仮名をローマ字 (例：「Taka」) で記しているが、本稿では、カタカナ表記 (例：「タカ」) に変更している。

#### Ⅲ. 2. (1) 構成主義版グラウンデッド・セオリー・アプローチに基づくデータ分析

構成主義版グラウンデッド・セオリー・アプローチは、がん告知のような、複雑な意思決定に関わる課題を扱う際に特に有効である。なぜなら、研究参加者は、微妙な心情や態度を表すために、参加者毎に、あるいは状況毎に、様々な表現を用いる可能性が高く、研究者は逐語録に書かれた言葉を解釈しつつ、さらに研究参加者の独自の見方に順応しようとする姿勢が求められるからである。

表 1 は、前述したがん告知をめぐるフォーカスグループでの議論の逐語録に対して、構成主義版グラウンデッド・セオリー・アプローチに基づくコード化を行った例である。以下では、

この手法におけるコード化（「初期段階のコード化」および「焦点化のためのコード化」）の過程に焦点を当てて解説する。

### 1) 初期段階のコード化

研究者による逐語録の通読の後、初期段階のコード化を行った。ここでは、参加者の1回の発言を1行とし、その行の中の単語レベルで、本課題に関わって重要と思われる語句を抽出する「単語毎のコード化」を行った。次に、「行毎のコード化」を行った。このコード化の過程の一部を、表1に示した。左側の「質問」列は、祖父母への告知という状況を表している。「名前」列には、研究参加者の名前（仮名）が示されている。「初期コード（単語毎）」の列は、単語毎のコード化の結果が示されている。それらを元に、行毎のコード化を行った結果は、「初期コード（行毎）」の列に示されている。

### 2) 焦点化のためのコード化

焦点化のためのコード化とは、初期段階のコード化によって生成されたコードを、比較・分類・整理することである。この作業によって、分析の方向性が明確になり、データをより大きなまとまりとして統合・説明するためのコードが見出される。

Charmaz (2006) も述べているように、初期段階のコード化から焦点化のためのコード化に移るプロセスは、直線的なものではない。つまり、焦点化のためのコード化の過程で新しいアイデアや疑問点が浮上し、それらを確認する意味で、初期段階のコード化に対する再検討が行われる場合もある。

<p>* 人生充実                      '残りの時間' or '残りの人生' or '充実感' or 有意義                      or ( 最後 &amp; 迎える ) or ( 幸せ &amp; 死ぬ ) or ( 最後 &amp; ゴール )</p>
<p>* 自己決定・権利                      '自分が決め' or '自分で決め' or '私が決め' or 権利                      or '私が進め' or 'すべて決めたい' or '人生の選択'                      or '行きたい' or 'したい' or 'しておきたい'</p>
<p>* 社会的責任・義務                      near (社会-責任) or 財産 or 身辺                      or '人間関係' or 整理 or '残さないといけない' or '子供がいたら'                      or '家族があります' or '自分のすべき事'</p>
<p>* 年齢的要因                      年齢 or 歳 or 若い or 小さい or 'どれだけ大人か' or 老衰</p>
<p>* 心的要因                      希望 or 夢 or 性格 or '気が弱' or 恐怖 or 恐い or 悲しい or 楽しい                      or 幸せ or 残酷 or ショック</p>

図2：KH Coder のコーディング・ルールファイルの例

### Ⅲ. 2. (2) テキストマイニングに基づくデータ分析

本研究では、がん告知をめぐるフォーカスグループの逐語録を、計量テキスト分析のために設計された KH Coder を用いて分析した。具体的な分析手順は次の通りである。まず、フォーカスグループ・インタビューの逐語録に対して形態素解析を行った。そして、そこで抽出された語句の中で、出現頻度の高い語句を取り出し、語句と発言者の対応関係を、多変量解析の一種である対応分析によって可視化し解釈した。次に、単語毎のコード化によって抽出されたインビボ・コード（研究参加者自身の言葉を用いたコード）と、焦点化のためのコード化によって生成されたコードの関係を、コーディング・ルールファイルとして定義した（図2）。その後、この定義に基づいて、KH Coder のオートコーディング機能を用いて、フォーカスグループでの議論における、5つの仮定的状況と焦点化コードとの関係を可視化した（図3および4）。

### Ⅲ. 2. (3) 統合的分析

以下では、がん告知をめぐるフォーカスグループの議論に対して、GTMA に基づく統合的分析を行った例について述べる。

表2は、図2で示したコーディング・ルールを元に、フォーカスグループで議論した5つの状況（行毎の見出しとして表示されている）と、焦点化コード（列の見出しとして示される）との関係性を、KH Coder のクロス集計機能で算出した結果である。分析対象となる文章の単位（コーディング単位）は、研究参加者による一回の発言に相当する「段落」を用いた。なお、表中のケース数が示すのは、各状況について、研究参加者が発言した回数である。また、それ以外のセルの数値は、焦点化コード（列の見出しとして示される）に関する発言数の割合を表す。例えば、「1\_ 祖父母」への告知という状況に関しては、合計33回の発言があり、そのうち「告知肯定」に関する発言が、21.21%を占めている。

図5は、表2の結果を、KH Coder の Excel 連携機能によって、マトリックスとして視覚化したものである。これは Excel の等高線図機能によって作成されており、中心部の四角は、コードの出現頻度を10%単位で描画された入れ子として表現している（つまり、出現頻度が10%以下のコードは、等高線として現れない）。また、入れ子の数が多ければ多いほど、その焦点化コードの出現頻度が高いことを示している。

以下では、これらの結果に対する、3つの研究設問の視点からの考察と解釈について述べる。

表2：状況とテーマのクロス集計結果

	*告知肯定	*告知否定	*他者判断	*人生充実	*自己決定・権利	*社会的責任・義務	*年齢的要因	*心的要因	ケース数
1_ 祖父母	21.21%	24.24%	3.03%	6.06%	6.06%	3.03%	15.15%	39.39%	33
2_ 両親	64.29%	21.43%	28.57%	14.29%	0.00%	14.29%	0.00%	28.57%	14
3_ 配偶者	50.00%	16.67%	0.00%	16.67%	8.33%	8.33%	0.00%	16.67%	12
4_ 子供	44.44%	50.00%	0.00%	22.22%	0.00%	5.56%	44.44%	11.11%	18
5_ 自分	64.71%	0.00%	11.76%	5.88%	17.65%	11.76%	0.00%	23.53%	17
合計	43.62%	23.40%	7.45%	11.70%	6.38%	7.45%	13.83%	26.60%	94

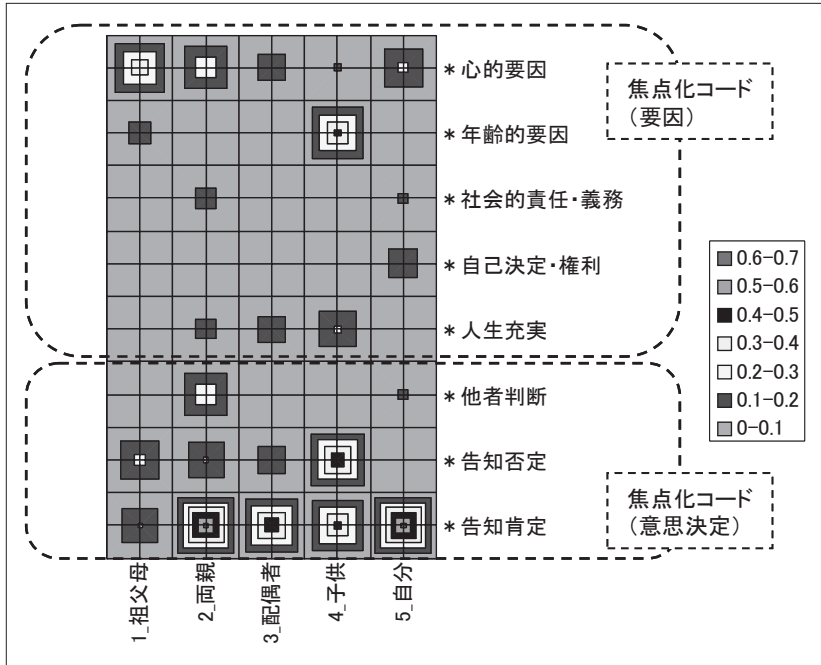


図3：5つの仮定的状況と焦点化コードのマトリックス

注：点線部分は筆者が追加

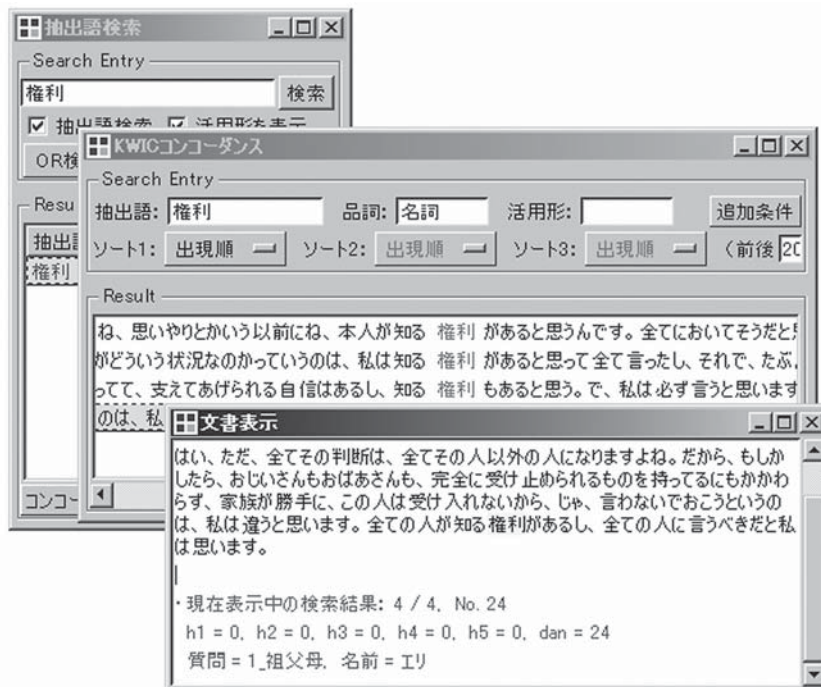


図4：コンコーダンス機能による検索例

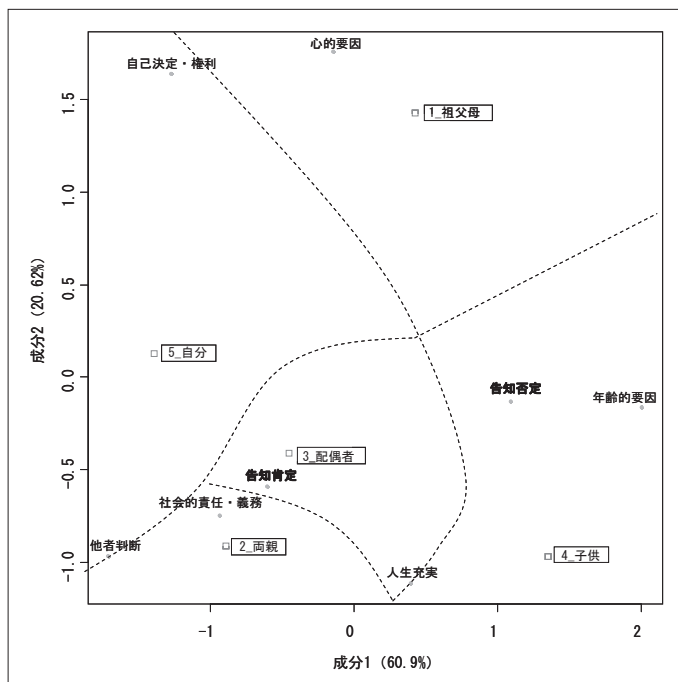


図5：仮定的状況と焦点化コードの対応関係 注：点線は筆者が追加

研究設問 (a) 「日本人はどのような倫理的な声を持ち、どのような意思決定をするのか」

図3を解釈すると、「人生充実」、「自己決定・権利」、「社会的責任・義務」、「年齢的要因」、「心的要因」といった5つの焦点化コードが、「告知肯定」、「告知否定」、「他者判断」という3つの意思決定に影響を与えていることが推察される。

ここでは、状況1（祖父母への告知）の詳細を分析した例について述べる。図3の「1\_ 祖父母」という属性の列を見ると、「告知肯定」と「告知否定」の両方の焦点化コードに関する等高線が現れており、それらの入れ子の数はほぼ同数である。そして「心的要因」と「年齢的要因」に関する言及があることがわかる。

しかし研究者は、このような等高線の表れ方について、いくつかの疑問を持つ可能性がある。例えば、「心的要因」と「年齢的要因」という、一見「告知否定」の要因になると考えられる焦点化コードにのみ等高線が現れているにもかかわらず、「告知肯定」の要因と考えられる焦点化コード（例えば「自己決定・権利」）には等高線が現れていないという疑問点が浮上したとする。そこで、マトリックスの元データである表2を参照すると、「1\_ 祖父母」と「自己決定・権利」が交差するセルには、6.06%という数値があり、10%未満であるため等高線は現れてはいないものの、「自己決定・権利」に関わる発言も出現していることがわかる。さらにその詳細を確認するため、図2のコーディング・ルールの「自己決定・権利」というコードに該当するインビボ・コードとして、「権利」に着目し、それをコンコダンス機能で探索すると、図4の「文書表示」ウィンドウに示されているような発言があることがわかる。

このように、可視化の結果から浮上した疑問点・矛盾点を元に、可視化の元になる数値、コード、原文などを確認・参照していくことで、例えば祖父母については、告知される側の心情や年齢を配慮すべきであるという理由と、知る権利や人生を充実させることを支援すべきであるという理由によって、「告知肯定」と「告知否定」の両方の意見が述べられていることがわかる。

図5は、5つの状況と焦点化コードとの関係性についてさらに詳しく調べるために、対応分析を行った結果である。この可視化によって、それぞれの状況に特徴的な焦点化コードが何かを、視覚的に理解することができる。

まず、「1\_祖父母」というラベルは、「告知否定」、および「心的要因」というコードの近くに布置されている。このことから、祖父母への告知（状況1）に関しては、「心的要因」という理由によって告知に否定的な発言をした参加者が多かったことがわかる。「2\_両親」（状況2を表す）と「3\_配偶者」（状況3を表す）については、類似した傾向が見られ、両者とも「告知肯定」に近い側に布置されている。また「社会的責任・義務」が近くにあることから、財産などの子ども達への社会的責任に関する発言が多くなされていたと推測される。「4\_子供」（状況4を表す）については、「告知否定」の近くに布置されている。その理由としては、「年齢的要因」が大きな影響を与えていると考えられる。「5\_自分」（状況5を表す）については、他者に判断を任せたいという気持ち、つまり「他者判断」という理由に加えて、「自己決定・権利」を重視する立場から、「告知肯定」の側にコードが出現していると考えられる。

#### 研究設問 (b) 「日本人の中で、倫理的な声はどのように異なるのか」

図6は、フォーカスグループで議論した5つの状況全体に対する、研究参加者と焦点化コードのマトリックスである。この図から、エリは、「心的要因」や「年齢的要因」について言及してはいるが、「自己決定・権利」および「人生充実」といった理由によって、告知については肯定的であることがわかる。ケイとタカは、「心的要因」に言及しつつ、「人生充実」あるいは「社会的責任・義務」という要因によって、告知に対しては、より肯定的な態度を取っている。マキとユカに共通する特徴は、「社会的責任・義務」と「自己決定・権利」への言及が見られないことである。そして、「告知肯定」と「告知否定」の間でアンビバレントな態度を取っているという結果が示されている。

ここで、例えばユカが、実際に、告知に対してアンビバレントな発言をしていたのか、という疑問点について調べるとする。そこで逐語録の中で、ユカの発言を参照すると、状況毎に以下のような異なる意思決定について言及していることがわかる。

「配偶者には、まず最初に伝えると思います。まあ、配偶者の人生であるわけだから…」

<発言者：ユカ（仮名）、状況：配偶者への告知>

「正直いって、自分の口からは、とても辛くて、伝えられないと思うんです。」

<発言者：ユカ（仮名）、状況：子どもへの告知>



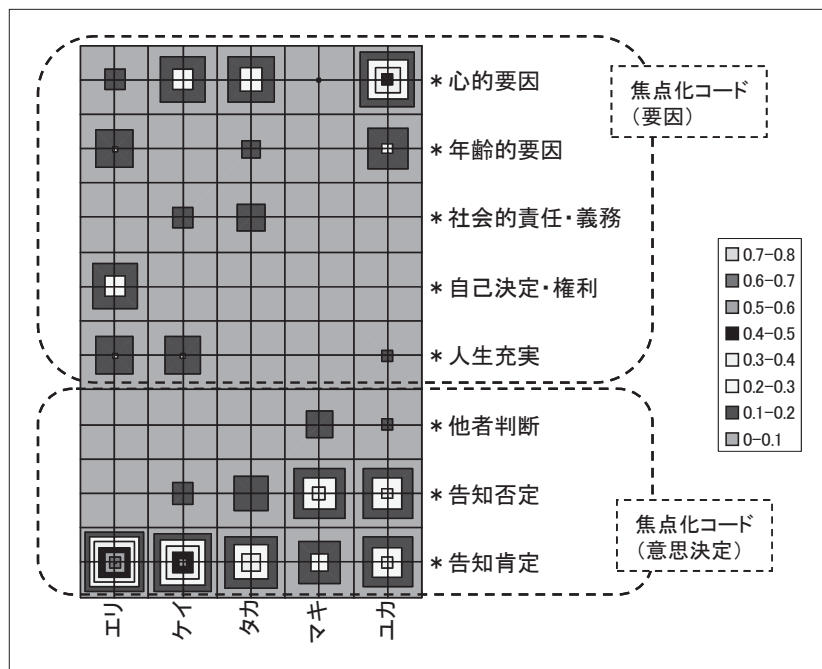


図6：研究参加者と焦点化コードのマトリックス

注：点線部分は筆者が追加

「私は知りたいです。やっぱり自分が死ぬんだって分かったら、あの、行きたい所もあるし、欲しい物もあるので、それを全部満たしたいというもあります。」

<発言者：ユカ（仮名）、状況：自分への告知>

このように、コーディング・ルールを元にしたマトリックスによる可視化と、そこから浮上した疑問点に関わる原文の参照を繰り返すことで、研究参加者の間で、告知に対する態度や意思決定がどのように異なるのかを、明瞭な形で理解することができる。そしてその結果から、参加者が持つ倫理的な声の違いについて、客観性を担保する形で、データに忠実な解釈を展開することができる。

### 研究設問 (c) 個人が他者の言葉によってどのように影響を受けるのか

図7は、研究設問 (c) に関して、前述と同じマトリックスを用いて可視化した例である。この課題に関する可視化の方法はさまざまな手段が考えられるが、ここでは、祖父母に対する告知という状況に関して、議論の前半部分と後半部分を分割した逐語録のデータに対して、図2で定義したコーディング・ルールを適用し分析した。前半と後半では、発言者や話題が多少異なっているため、マトリックスの現れ方も同様に異なっている。例えばケイは、前半でしか発言をしていないため、図7右側（後半に対応）ではどのコードの行にも等高線が現れていない。

同じく後半で、「社会的責任・義務」と「人生充実」というコードに関わる発言がなかったため、これらについての等高線も表示されていない。

ケイ以外の参加者について、図7中の左右の図を比較すると、エリとタカのそれぞれに関しては、「告知肯定」と「告知否定」の等高線が類似した形で現れている。しかしマキとユカに関しては、前半で「告知否定」のみに等高線が現れているのに対して、後半では、「告知肯定」と「告知否定」の両方に等高線が現れている。そこで、逐語録中のユカの発言を見ると、「祖父母への告知」という質問に対して、最初は存命中の祖母のみのことを考えて告知に否定的な発言をしていたが、議論の中で性格の違いに関する話題が出た後は、以下のように、祖父と祖母を分けて考える発言をしていた。

「私の場合は、あの、幼い頃に両方とも祖父を亡くしているので、想像つかないんですけど、私の小さい頃の、あの、思い出の中の祖父を想像すると、言ってもいいのかな、というのがあります。ていうのは、あの、祖父と祖母の性格の違いっていうか、それを聞いて受け止められるか、られないかというところで、私は判断しています。」

<発言者：ユカ（仮名）、状況：自分への告知>

つまり、他の参加者との対話から、性格の違いを意識した結果、「祖父には告知してもいいが、祖母には告知しない」という、相反する意思決定について言及するようになったことがわかる。

この例のように、GTMAによって、個々の参加者に関する特異な変化を見つけ出し、そこから参加者同士のやりとりにおける相互作用のありありとした姿を見出すことができる可能性が示唆された。

#### IV. 考察～2つのアプローチの方法論的比較

以下では、前述した2つの分析事例を元に、研究者自身が行う一般的な質的データ分析と、本稿で提案するGTMAという2つのアプローチについて比較検討する。またその際の手がかりとして、構成主義パラダイムに基づく質的研究のための評価規準（Lincoln & Guba, 1985; Guba & Lincoln 1989）を用いる。

##### IV. 1 質的研究のための評価規準

Lincoln & Guba (1985) および Guba & Lincoln (1989) は、構成主義パラダイムに基づく質的研究を評価するための規準として、authenticity（真正性）と trustworthiness（信用性）という2つの大きな概念を挙げている。前者は質的研究がもつ特有の評価規準であり、研究の公平性を評価対象とする。一方後者は、研究結果がどれだけ信用に値するものかを示す重要な研究評価規準であり、量的研究における validity（妥当性）と対比されるものである。この評価規準は、credibility（信憑性）、transferability（転用可能性）、dependability（確実性）、および

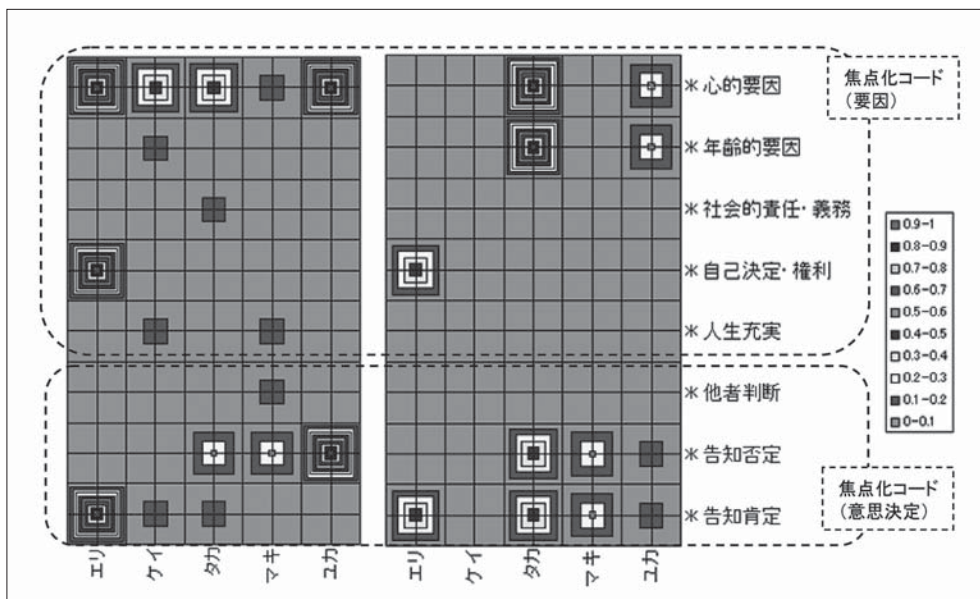


図7：祖父母への告知に関する議論 (左側：前半部、右側：後半部)

注：点線部分は筆者が追加

confirmability (確証性) という4つの概念によって支えられている。

Credibility は、研究参加者が持つリアリティと、研究者によって表現された結果が一致しているかどうかという評価規準である。これを評価するための方法としては、長期にわたる関与、継続的な観察、中立な評価、否定事例分析、メンバーチェックなどがある。Transferability は、量的研究の一般化可能性に対応する概念であり、得られた知見が類似した状況や参加者にも適用可能であるかどうかという基準である。Dependability は、量的研究における信頼性に対応する概念である。もちろん質的研究である以上、量的研究のように、研究者間でデータ分析の結果に差が出ないということは不可能であるが、別の機会に同様の状況で研究を行った場合に、同様の知見が得られるのかどうかという点は、研究の信用性を判断する上では重要な規準である。Confirmability は、量的研究における objectivity (客観性) に対応する概念である。この基準を満たすためには、どのようなデータに対してどのような解釈が行われ、そしてどのような知見が得られたのか、という研究の過程について、他の研究者が確認するための手がかり、つまり audit trail (監査証跡) が提供されている必要がある。

#### IV. 2 2つのアプローチの比較検討

上で述べた質的研究のための4つの評価規準の中で、研究者自身が行う質的データ分析と、本稿で提案する GTMA という2つのアプローチの間で、最も大きな差が出るのは、audit trail の提供が重視される confirmability である。一般的な質的研究では、分析結果を報告する際に、データや解釈のプロセスに関する「濃密な記述」が必要となるため、質的研究の論文は通常は多く

のページ数を必要とする。また分析の全体像を他の研究者が理解する場合も、量的研究に比べて時間がかかる。GTMAにおいても、研究者自身が解釈を行う必要はあるが、一般的な質的研究に比べて、データの全体像や、データ分析と解釈の過程、またその結果を、明瞭な形で示すことができる。具体的には、コンピュータによる言語処理や多変量解析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、分析に用いたコーディング・ルールやユーザ辞書などをオープンにすることで、どのような解釈や抽象化が行われ、そこから結果がどのように導きだされたのかということ、他の研究者が容易に理解することが可能である。

一般的な質的研究と比較した場合のGTMAのもう一つの長所は、dependabilityを担保することが容易であるということである。上述した通り、一般的な質的データ分析においては、異なる研究者の間で完全に同じ研究結果が得られることは通常は考えにくい。しかしGTMAでは、コンピュータによる言語解析や多変量解析といった機械的な処理の結果に基づいて分析を行う必要があるため、分析の結果に、研究者による「解釈の飛躍」が含まれる余地が少ない。

それ以外の2つの規準、つまりcredibilityおよびtransferabilityは、GTMAを用いることで簡単に改善できるというものではない。しかし、「II. 4 GTMAに基づく統合的分析」で述べたとおり、人間がテキストを理解する方法と、コンピュータがテキストを解析する方法は根源的に異なっているため、両方の結果を照合し、そこでの矛盾や疑問点を解消する作業を行う過程で、テキストで語られている現象をより深く理解できる可能性が高い。つまりGTMAを用いることで、研究者自身が行う質的データ分析だけでは見えてこない矛盾や疑問点を見つけ出すことができ、またその探求によって、credibilityやtransferabilityの視点から、より充足した分析結果を導き出すことができる可能性がある。

以上、質的研究のための4つの評価規準の視点から、GTMAの利点について述べてきたが、それらと表裏一体の欠点がある。それは、研究者自身が行う質的データ分析に比べて、研究時間の短縮が保証されるわけではないということである。単純に考えれば、GTMAでは、一般的な質的データ分析とテキストマイニングの両方に取り組む必要があり、さらにその結果として、一般の質的データ分析だけでは見えてこない矛盾や疑問点まで浮上してくるため、それらを探求する作業も必要となる。つまり、最終的に研究全体に要する時間は、一般的な質的データ分析に比べて長くなる可能性は十分にある。しかし、confirmabilityとdependabilityの点では明らかに改善が見込まれ、credibilityやtransferabilityについても、より望ましい分析結果が導きだされる可能性は高い。従って、GTMAは、現在の質的研究をさらに深化させ、発展させていくための一つのヒントを提供するものであると考えられる。

## V. まとめ

本論文では、先行研究として筆者らが行った、ガン告知の可否をめぐるフォーカスグループ・インタビューの逐語録に関する質的分析を取り上げた。そして、質的データ分析支援の新しい手法として、GTMAを提案した。次に、先行研究と同じデータに対して提案手法を適用した結

果を紹介した。最後に、提案手法について、質的研究の評価基準の視点から検討を行った。

本稿で提案する GTMA のデザインは、大きく 2 つの方向性から構成される。まず 1 つは、研究者自身によるテキストの読み込みや分析によって、テキストが語っている現象についての様々な可能性を展開させ、divergent な知見を得るという方向性である。もう一つは、テキストマイニングによって、テキストの全体像や関係性などを視覚的に確認することによって、知見を収束させる convergent な方向性である。またこの 2 つの方向性での作業の繰り返しによって、テキストデータを、ミクロとマクロの両方から把握することができ、またテキストデータの表面には現れていない深いレベルでの仮説生成や、新たな知識構築ができる可能性がある。

一方で GTMA では、一般的な質的データ分析とテキストマイニングの両方に取り組む必要があり、さらにその結果として、一般の質的データ分析だけでは見えてこない矛盾や疑問点まで浮上してくるため、最終的に研究全体に要する時間は、一般的な質的データ分析に比べて長くなる可能性がある。しかし、質的研究の 4 つの評価基準の点から、GTMA によって、より望ましい分析結果が導きだされる可能性がある。従って GTMA は、現在の質的研究をさらに深化・発展させていく上での重要なヒントを提供するものである。

本稿では、がん告知をめぐるフォーカスグループ・インタビューという、限定された話題と方法で収集されたデータを扱ったが、今後は、質的研究における多様な調査手法によって集められたテキストデータを分析することで、GTMA の検証と改善に取り組んで行く予定である。さらに、本稿で提案した GTMA の実践例では、テキストマイニングによる分析を、計量テキスト分析およびその支援ツールである KH Coder を用いて紹介したが、近年この分野の技術は急速に発展しているため、今後は、このような技術的な進歩に合わせて、GTMA の高度化に取り組んでいきたい。

#### 謝辞

まず、がん告知に関するフォーカスグループでの議論に参加して頂いた方々に感謝する。KH Coder を用いた分析については、立命館大学樋口耕一先生から多くのご助言をいただいた。また本稿で提案した手法は、2008 年度多文化関係学会年次大会のプレスカンファレンスワークショップにおいて最初に紹介したが、その際参加者の方々から有益なコメントを頂いた。最後に、本稿をこのような形でまとめる機会を与えて頂いた、立命館大学 Gordon L. Ratzlaff 先生に感謝申し上げます。

#### 参考文献

- Bakhtin, M. (1981). *The dialogic imagination*. Austin, TX: University of Texas Press.
- Becker, P. H. (1993). "Common pitfalls in published grounded theory research". *Qualitative Health Research*, Vol. 3, No.2, pp. 254-260.
- Charmaz, K. (2000). "Grounded theory: Objectivist and constructivist methods". In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research*. (pp. 509-535). Thousand Oaks, CA: Sage Publications. (チャー

- マズ, K. (2006). 「グラウンデッド・セオリー: 客観主義的方法と構成主義的方法」『質的研究ハンドブック』第2巻, (平山満義, 監訳). 京都: 北大路書房.)
- Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. London: Sage Publications. (シャーマズ, K. (2008). 『グラウンデッド・セオリーの構築: 社会構成主義からの挑戦』(抱井尚子・末田清子, 監訳). 京都: ナカニシヤ出版.)
- Flick, U. (1995). *Qualitative Forschung*. Rowohlt Taschenbuch Verlag, GmbH. (フリック, U. (2002). 『質的研究入門—人間の科学>のための方法論』(小田博志・山本則子・春日常・宮地尚子, 訳) 東京: 春秋社.)
- 藤井美和・李政元・小杉考司. (2005). 『福祉・心理・看護のテキストマイニング入門』東京: 中央法規出版.
- Glaser, B.G. & Strauss, A.L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, IL: Aldine. (グレイザー B.G. & ストラウス, A.L. (1996). 『データ対話型理論の発見』(後藤隆・大出春江・水野節夫, 訳). 東京: 新曜社.)
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. Newbury Park, CA: Sage Publications.
- Hearst, M. A. (1999). "Untangling Text Data Mining". *ACL'99*, pp. 3-10.
- 樋口耕一. (2004). 「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, No. 1, pp. 101-115.
- 樋口耕一. (2006). 「内容分析から計量テキスト分析へ—継承と発展を目指して—」『大阪大学大学院人間科学研究科紀要』Vol. 32, pp. 1-27.
- Kakai, H. (2009). "Multivoicedness in Bioethical Arguments about Cancer Disclosure in Japan: A Narrative Analysis Using a Sociocultural Approach". 『青山国際政経論集』Vol. 78, pp. 179-209.
- 木下康仁. (2003). 『グラウンデッド・セオリー・アプローチの実践: 質的研究への誘い』東京: 弘文堂.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.
- 松村真宏・三浦麻子. (2009). 『人文・社会科学のためのテキストマイニング』東京: 誠信書房.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: an expanded sourcebook*. Thousand Oaks, CA: Sage Publications.
- Seidel, J. (1991). "Method and Madness in the Application of Computer Technology to Qualitative Data Analysis". In N.G. Fielding & R. M. Lee (Eds.), *Using Computers in Qualitative Research*. (pp. 107-118). London: Sage Publications.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research*. Thousand Oaks, CA: Sage Publications.
- Weitzman, E. A. (2000). "Software and Qualitative Research". In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research. 2nd ed.* (pp. 803-820). Thousand Oaks, CA: Sage Publications. (ワイツマン, E. A. (2006). 「ソフトウェアと質的研究」『質的研究ハンドブック』第3巻, (平山満義, 監訳). 京都: 北大路書房.)
- Willig, C. (2001). *Introducing Qualitative Research in Psychology: Adventures in Theory and Method*. Buckingham: Open University Press. (ウイリッグ, C. (2003). 『心理学のための質的研究法入門—創造的な探求に向けて—』(上淵寿・大家まゆみ・小松考至, 共訳). 東京: 培風館.)