



Title	社会空間アプローチにおける計量テキスト分析の統合可能性に関する一検討
Author(s)	田中, 祐児
Citation	大阪大学教育学年報. 2026, 31, p. 29-39
Version Type	VoR
URL	https://doi.org/10.18910/104320
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

社会空間アプローチにおける計量テキスト分析の 統合可能性に関する一検討

田 中 祐 児

要旨

社会空間アプローチでは近年、質問紙調査とインタビュー調査を組み合わせる試みが盛んに行われている。本稿は、インタビューデータに計量テキスト分析を適用することが、社会空間アプローチのさらなる発展に貢献しうると提案するものである。

本稿では質問紙調査とインタビュー調査からなるデータを用い、社会空間を構築した。その後インタビューデータに計量テキスト分析を適用することによって、社会空間と計量テキスト分析とを架橋した。具体的には第1に、語彙の使用有無をサプリメンタリー変数として空間に投影することで、教育・生活スタイルを表す語彙が社会空間と相同性をもつ可能性を可視化した。第2に投影に用いた語彙についてtf-idfを算出し、同一語彙でも回答者によって前景化の程度と意味づけが異なること、さらにそれが空間上の位置と関連して分化しうることを示した。

語彙の投影とtf-idfの併用は、事例の選定過程を明示しつつ、空間上の位置と語りとを再検討可能な形で接続することを可能にする。その点で計量テキスト分析は質的分析の代替ではなく、社会空間アプローチにおける解釈や論証の足場を増やす補助的な手段として位置づけられる。

1. 問題設定

本稿の目的は、社会空間アプローチの文脈で収集されたインタビューデータの分析に、計量テキスト分析をどのように統合しうのかを探索的に検討するとともに、その方法論的意義を提案することにある。社会空間アプローチとは、「さまざまな地位変数が織りなす関連を多次元空間として把握し、そこでの客観的な位置づけをもとに人の行動や意識の態様を説明しようとするもの」(近藤 2012, 103-4頁)であり、その理論的着想はピエール・ブルデューの『ディスタンクシオン』(Bourdieu 訳書2020)に求められる。

この社会空間アプローチに基づく先行研究の多くは、地位を表す変数によって構築された社会空間と、日常生活における慣習的行動や嗜好を表現する空間との相同性を検討してきた(近藤 2011)。具体的には、多重対応分析によって構築された社会空間(あるいはライフスタイル空間)に対して、ライフスタイル(あるいは地位)に関する変数をサプリメンタリー変数として投入し、両者の対応関係を吟味する、という手順が採用されてきた。

これらの研究成果が有用であり続けているのはいうまでもない。一方で近年では、多重対応分析の結果を、インタビューデータの分析と接続しようとする試みも蓄積されつつある(Bennett et al. 2021, 知念編 2025, Iso et al. 2025, Jarness 2018, 田中 2024)。これらの研究は、まず地位変数に多重対応分析を適用することで社会空間を構築する点において、従来の社会空間アプローチと共通している。しかしその後のプロセスにおいて、社会空間上における個々の調査対象者の位置を積極的に参照し、その位置を媒介として、社

会空間とインタビューで語られる内容とを結びつけようとする点に特徴がある。

このように、空間上の個人の位置を手がかりとして量的分析と質的分析を接続することで、社会空間アプローチは新たな展開可能性を獲得する。すなわち、計量的に構築された社会空間と、インタビューにおける語りとを対応づけて解釈することにより、両者を「数学的にも社会学的にも首尾一貫した形で捉えることができる」(磯 2020, 287頁) ようになる。

この点は、「それぞれのインタビュー対象者がどこに位置しているのかを突き止め、量的調査の回答と口頭でしゃべったこととを結びつけることができる」(Bennett et al. 訳書2017, 92頁) という利点をもたらす。量的データに基づく対象者の位置と、質的データから得られる語りの内容とを相互参照的に解釈することで、両者を有機的に統合した分析が可能になるのである。

以上のように、社会空間アプローチに依拠した研究群は近年、インタビューデータを積極的に用いることで展開してきた。本稿はこの流れをふまえ、インタビューデータの分析に計量テキスト分析を導入することが、社会空間上の個人の位置とインタビューで語られる内容との対応関係を吟味するうえで、いかなる意義をもちうるのかを検討する。

2. 社会空間アプローチにおけるインタビューデータの分析

社会空間アプローチにもとづく研究において、インタビューデータは、社会空間上の位置がどのように主観的に経験され、意味づけられているのかを明らかにするための重要な資料として位置づけられてきた。とりわけ、社会空間上の位置と生活様式との連関を具体的に描き出すうえで、インタビューデータの質的な解釈は重要な役割を果たしてきた (Bennett et al. 訳書2017)。

このような質的な解釈によって、調査対象者の語りに即して、実践や嗜好、価値判断がどのような文脈のもとで形成されているのかを詳細に把握することが可能になる。そのため、社会空間アプローチにおいては、量的に構築された空間構造を、インタビューによって厚みのある記述として補完するという研究戦略が採用されてきた。

しかし同時に、質的な分析は、いくつかの技術的な制約を伴う。第1に、複数の対象者を横断した比較が困難になりやすい点である。第2に、社会空間上の位置と語りとの対応関係が、個別事例の記述にとどまり、空間全体の構造との関係として明示化されにくいという問題がある。これらは言うまでもなく、質的な分析の価値を損なうものではない。しかし同時に、社会空間アプローチの射程を考えるうえでは看過できない論点である。

こうした課題を踏まえると、インタビューデータの分析において、質的手法を前提としつつも、それを補助する分析手法の導入可能性を検討する余地がある。その1つが計量テキスト分析である。計量テキスト分析は、インタビューにおける語りを数量化し、複数の対象者間で比較可能な形に変換することを可能にする(樋口 2020)。この点で、計量テキスト分析は社会空間分析と親和的であると考えられる。

重要なのは、計量テキスト分析を質的読解の代替として位置づけるのではなく、社会空間上の個人位置と語りとの対応関係を検討するための補助線として用いる点にある。すなわち、計量テキスト分析を通じて、語りの分布や特徴を空間上の位置と結びつけて把握することで、質的な分析では捉えにくかった構造的対応関係を可視化することが可能となる (Grimmer et al. 2022)。このような視点から、次節以降では、本稿で用いるデータと分析手続きについて説明する。

3. データと方法

本稿で用いるデータと分析方法について説明する。ただし、データは田中（2024）および知念編（2025）と同様であるほか、分析方法のうち多重対応分析についても田中（2024）と同様であるため、本節では詳細を割愛する。詳しくは上記の文献を参照されたい⁽¹⁾。

3.1. データの概要

質問紙調査とインタビュー調査を組み合わせたデータを用いる。質問紙調査は東京都内のある自治体に居住し、質問紙調査実施時点（2021年1月）で高校1年生の学年に相当する子どもとその保護者のうち、無作為に抽出された1,000ペアを対象に実施された。有効回答ケースは300ペアであった。インタビュー調査はこれら300ペアのうち、調査に同意した23ペアを対象に実施した。本データは保護者とその子によるペアデータとなっているが、本稿では紙幅を鑑み、保護者によるデータのみを分析する。

インタビューは保護者と子どもとで別々に実施した。保護者に対しては本人の生活史を中心としつつ、子どもの生育や学校経験についても聞き取った。インタビューに要した時間は概ね、1時間30分から2時間程度である。インタビューの実施期間は2022年3月から2023年3月までである。

3.2. 方法

本項では分析の手順を説明する。本稿は、社会空間アプローチにおけるインタビューデータ分析に、計量テキスト分析をいかに統合しうるかを検討することを目的とするものであった。この目的を果たすため、質問紙調査データに多重対応分析を適用して社会空間を構築し、各回答者を空間上に配置する。これにより、資本量および資本構成の差異が反映された社会空間を得る。

その後、インタビューデータに計量テキスト分析を適用し、社会空間上の個人の位置と、インタビューにおいて語られる内容との対応関係を検討する。ただし本稿では、計量テキスト分析を2つの段階に分けて行う。

第1に、語彙の使用状況を社会空間に投影することにより、社会空間と使用語彙との対応関係を検討する。第4節では教育・生活スタイルに関する語彙に着目することによって、それらと社会空間との相同性を把握する。第2に、第4節で確認された語彙と社会空間との対応関係をふまえつつ、それらの語彙が語りのなかで果たす役割を回答者間で比較検討する。

換言すれば、本稿では、語彙の分布と社会空間との対応関係を把握する分析（第4節）と、個々の語りにおいて、それらの語彙がどのような位置づけのもとで使用されるのかを検討する分析（第5節）を行う。これにより、社会空間上の位置が、ある語彙を使用するか否かという論点にとどまらず、その語彙をどのように使用するか、いかなる意味づけを付与するのかという論点をも射程に含むことを示す。

なお本稿の計量テキスト分析では、その対象を回答者の語りに限定し、調査者による発言は分析から除外する。一般に、調査者の発言が強く構造化され、すべてのインタビューにおいて同一であれば、それらは全回答者に共通する要素として作用し、計量テキスト分析に含めたとしても分析結果に影響を及ぼすことはない。しかし、本稿で用いるインタビューデータは複数の調査者によって収集されており、生活史としての性質上、調査者の発言は回答者ごとに柔軟に変化している。そのため、調査者の発言を分析対象に含めた場合、その差異が分析結果に反映される可能性がある。本稿の目的は、社会空間上の回答者の位置と語りの内容との対応関係を検討することにあるため、そのような差異が結果に現れることは適切ではない。以上の理由か

ら、本稿では回答者の語りのみを分析対象とする。

3.3. 計量テキスト分析の手続

本項では、本稿で用いた計量テキスト分析の具体的な実装手続について説明する。まず、インタビューデータのテキストファイルに対して形態素解析を行い、語彙単位へと分割した。形態素解析器としてMeCab(0.996)を用い、辞書にはIPA辞書を使用した。IPA辞書には含まれていないものの、分析上重要な意味をもつ語彙については辞書登録を行った(「指定」と「校」を合わせて「指定校」など)。語彙単位には表層形ではなく原形を採用した。解析後、品詞情報にもとづき、名詞・形容詞・動詞のみを分析対象とした。また数字を含む語彙や、個人や地域が特定されうる固有名詞、頻度が一定程度高いものの、分析の目的に照らして意味をなさない判断できた語彙についてはストップワードとして除外した⁽²⁾。

以上の前処理を経て、各語彙の出現頻度表(インタビュー回答者別の頻度も含む)を作成した。この頻度表をもとに、第4節および第5節の作業内容に応じて集計方法を変更した。

第4節の分析では、語彙の使用状況に関するデータセットを作成した。そこでは行に回答者、列に語彙を配置し、各セルには語彙の使用有無を格納した。これによって得た語彙の使用状況に関する情報を、多重対応分析におけるサプリメンタリー変数として投入した⁽³⁾。

第5節の分析では、語彙の使用頻度にもとづきtf-idfを算出した。tf(term frequency)の算出にあたっては、回答者ごとの発話量の差を考慮するため、各語彙の生起回数を当該回答者の総語数で除し、相対頻度として正規化した。idf(inverse document frequency)は、文書単位を回答者数(N=23)とし、当該語彙が1回以上出現した回答者数(document frequency, df)にもとづいて $\log(N/df)$ の形式で算出した。これらの積をtf-idfとし、各回答者における特徴語の計算を行った。

4. 語彙の使用からみた教育・生活スタイルと社会空間の相同性

本節では、地位変数にもとづいて構築した社会空間と、インタビュー回答者が使用した語彙との相同性について検討する。図1は、社会空間上に語彙の使用状況をサプリメンタリー変数として投影するとともに、各インタビュー回答者の位置を示したものである。

なお図の簡潔さに鑑み、分析対象としたすべての語彙ではなく、教育・生活スタイルに関する語彙に限定したうえで、原点からの距離が十分に大きい語彙のみを表示した。加えて、それらの語彙を使用したことを示すカテゴリーのみを投影した。図1においてグレーで示したラベルは地位変数カテゴリーであり、これらによって社会空間は構築されている。黒字が語彙の使用を表すカテゴリーである。白抜きのアールファベットは回答者の位置とそのラベルを表している。

図1の社会空間は、田中(2024)と同様であるため、その詳細な解釈は当該文献を参照されたい。ここではその概略を説明する。まず縦方向(南北)に展開する第1軸は、北に高い世帯年収や義親への相談可能性などがある一方、南では世帯年収や主観的な暮らし向きが低いこと、配偶者に相談できないことなどが示されていることから、資本量の多寡が現れていると解釈できる。次に横方向(東西)に展開する第2軸では、西に両親の高い学歴や高い世帯年収などが配置され、東では両親が専門学校や短大を卒業していること、正規雇用で働いている母親が職場や専門職などに相談できることなどが示されていることから、資本構成に関する対立が表現されていると解釈できる。具体的には西側では経済資本や文化資本が優勢であり、東側では社会関係資本が優勢であると整理できる。

このような社会空間のもことで、インタビュー回答者がある語彙を使用するかどうかは、資本量および資本構成の差異と対応する形で配置されている。図1に示される語彙の配置からは、資本量の多寡および資本構成の差異に応じた教育・生活スタイルの対比を読み取ることができる。資本量の多い北側では「プール」や「ピアノ」といった教育への潤沢な投資を表す語彙が用いられるのに対し、資本量の少ない南側では「我慢」「電気」など、生活上の制約を示す語彙が現れている。実際、北に位置するBさんが「習い事は結構やってたんですね。トランポリン、プール、体操、ピアノ、公文」と語る一方、南に配置される回答者は「電気が止まったことがあって」（Vさん）や「極力、行きたいって言えば行かせてっていうか、させてはいるつもりなんですけども、たぶん子ども我慢してる部分はあるんじゃないかな。」（Wさん）と語っていた。このような語りは、図1で示される内容と整合的である。

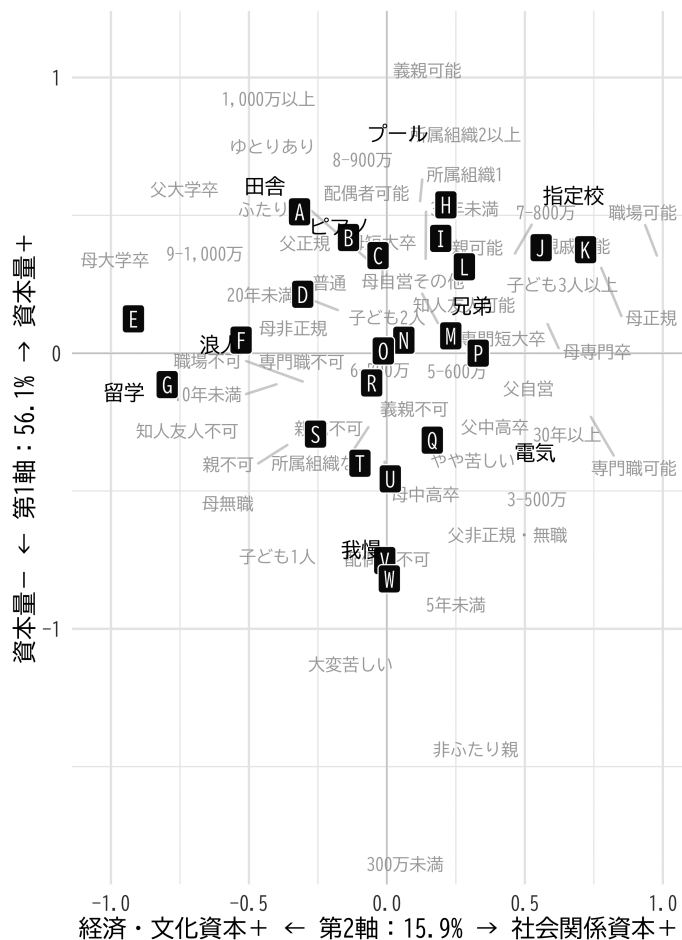


図1 社会空間における使用語彙とインタビュー回答者の配置

また、資本構成の差異に着目すると、経済資本・文化資本が優勢な西側では「留学」や「浪人」が、社会関係資本が優勢な東側では「指定校」が配置されている。これは前者では、大学受験や教育選択をめぐって時間的・金銭的コストを伴う戦略が語られるのに対し、後者では、制度を効率的に活用した進路選択が語ら

れていることがうかがえる。実際、Dさんは「交換留学とかそういうのもあったりするから、うまく利用してリーズナブルに行けるんだったら行ってくれたらいいかな」と語っていたし、Mさんも「指定校を取るためにがんばってきた」という自身の子どもの発言を参照して語っていた。

さらに、「田舎」と「兄弟」という東西に対照的に配置される語彙からは、資本構成の差異に応じた子育て経験の違いを読み取ることができる。前者は出身地の「田舎」から上京し、親族とは離れた状況のもとで子育てをしていることを示している。一方で後者は出身地である東京で子育てをしつつ「この近辺に住んでいる」(Lさん)という「兄弟」と支え合いながら子育てをするというスタイルが示されている。これらの語彙の配置は、資本構成の差異に応じて、子育て経験の前提となる地理的条件や親族関係のあり方が異なっていることを示唆している。

5. 社会空間上の位置と語彙の意味づけ

本節では、第4節で確認された社会空間と使用する語彙との対応関係を前提としつつ、特定の語彙を使用した回答者それぞれにおいて、当該語彙が語りの中でどのように重みづけられ、いかなる意味が付与されているのかを検討する。

第4節の分析は、語彙の使用の有無にもとづいて、語彙の分布と社会空間との対応関係を明らかにしたが、それだけでは、同一の語彙が個々の語りにおいてどのような役割を果たしているのかを捉えることはできない。そこで本節では、第4節で取り上げた語彙のうち、社会空間との対応関係が比較的明確に確認されたものを対象とし、各回答者の語りにおける当該語彙の相対的な重みに着目する。具体的には、各回答者を文書単位とし、語彙の出現頻度にもとづいてtf-idfを算出し、語彙ごとにその相対的な大小を手がかりとして、語彙の用いられ方を比較検討する。ここで用いるtf-idfは、語彙の使用頻度そのものを回答者間で直接比較するための指標ではなく、各回答者の語りの内部において、当該語彙がどの程度中心的な位置を占めているのかを把握するための補助的な指標として位置づけられる。

実際に着目する語彙は「留学」と「我慢」の2つである。これらは第4節において、社会空間と教育・生活スタイルとの相同性を比較的明確に示した語彙であるため、本節の分析対象として選定した。

5.1. 「留学」

本項では「留学」に注目した分析を行う。図2は図1と同一の社会空間上に、「留学」の語を実際に使用した回答者のみを抜き出して配置し、各回答者による当該語彙のtf-idf⁽⁴⁾を併記したものである⁽⁵⁾。

図1にある通り、「留学」は社会空間の西側に位置していた。これと整合的に、実際に「留学」の語を使用した5人の回答者は概ね、空間上の西に配置されている。

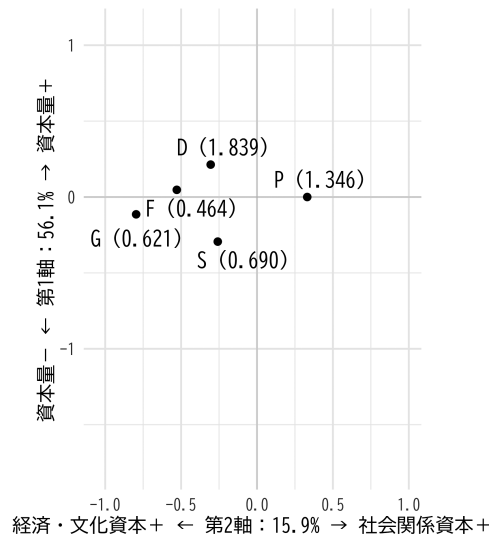


図2 「留学」使用者の空間上の位置と当該語彙のtf-idf

これら5人の回答者における「留学」のtf-idfを確認すると、相対的な高低が存在していることがわかる。tf-idfが高い事例では、当該語彙が語りの中でより前景化している可能性がある一方、低い事例では言及が限定的であると考えられる。

さらに「留学」のtf-idfが高いDさんとPさんは、「留学」の語を用いた5人の中では第2軸の得点が比較的高い⁽⁶⁾。すなわち「留学」は、使用の有無という水準で社会空間と対応するだけでなく、使用した回答者内部においても、前景化の程度（tf-idf）と社会空間上の位置とで対応している可能性が示唆される。このことは社会空間上の位置が、単なる使用の有無に関わるという論点を超えて、どのような重みで使用しているのかという論点にも関わることを意味している。

以下では、tf-idfの相対的な高低を手がかりとして実際の語りを参照し、「留学」が語りの中でどのような役割を果たし、いかなる意味づけがなされているのかを検討していく。

tf-idfの値が相対的に高い回答者においては、「留学」は教育投資における現実的な選択肢として用いられている。たとえばDさんは先に引用した通り「交換留学とかそういうのもあったりするから、うまく利用してリーズナブルに行けるんだったら行ってくれたらいいかな」と述べ、制度の活用可能性に言及していた。さらに「行けるんだったら今行っという方がいいじゃないですか。大学の間に」と語り、適切な時期を含め、「留学」を現実的なオプションとして位置づけていることがわかる。

これに対してtf-idfが低い回答者の場合、「留学」は周囲との比較のなかで、家庭の経済的制約を示す語彙として使用される。Gさんは「周りの子が留学しに行ったりとか、夏休みに沖縄に行くんだとかそういう話をして、うちは行けないんだみたい。帰省も正直厳しくって」と述べ、「留学」は「行けない」ものであると位置づけている。

このように、「留学」という同一の語彙であっても、語りの中で占める重みや意味づけは一樣ではなく、社会空間上の位置と関連しつつ分化していることがうかがえる。

5.2. 「我慢」

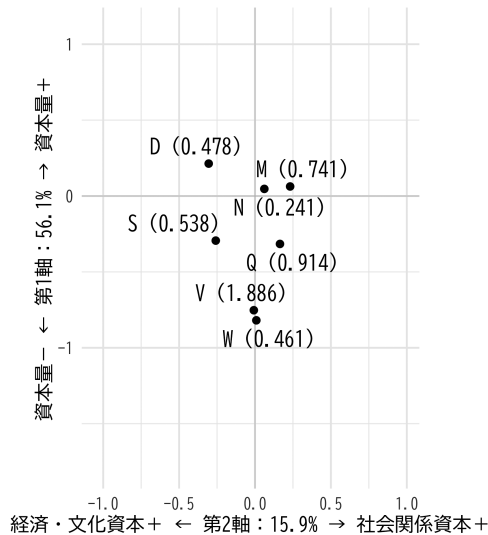


図3 「我慢」使用者の空間上の位置と当該語彙のtf-idf

次に「我慢」の語に注目する。図3は図2と同じ手法のもと、「我慢」に関する情報をまとめたものである、図1において「我慢」は南部に位置していた。一方、図3を参照すると、「我慢」の語を実際に使用した7人の回答者の位置はやや南方に集中しつつも、原点を超えて北側に位置する回答者も見られることがわかる。ただし、tf-idfの値が相対的に高いVさんやQさんは空間上の南部に位置していることから、「我慢」の語を積極的に使用することと社会空間上の位置との間には、一定程度の対応があることが示唆される。以下では前項と同様、tf-idfを手がかりに「我慢」の語を検討する。

tf-idfの値が相対的に高い回答者においては、「我慢」は生活上の制約や関係上の葛藤など、複数の場面を参照して用いられている。たとえばQさんは、「具合が悪いけど、手元にお金がないから行けないよねって、我慢しないとみたいな」と述べ、金銭的制約が受診を抑制する経験を「我慢」として語っている。またVさんも、「向こうが白って言えば白、向こうが黒って言えば黒っていう。(中略)私が、ずっと我慢してるっていう感じだった」と述べ、夫との関係性の不均衡のもとで反論を抑制してきた経験を「我慢」として位置づけている。

これに対し、tf-idfが低い回答者の場合、「我慢」は状況への対処を述べる補助的な語として現れ、制約や葛藤を説明する中心概念としては前景化しにくい。たとえばSさんは、「自分に生活力がないし、なんかもう我慢するしかないっていうか」と述べ、「我慢」を生活上の受容として簡潔に言及している。

5.3. 小括

以上のように、本節では、社会空間上の位置と当該語彙のtf-idfとの関係を確認したうえで、tf-idfが高い回答者と低い回答者の語りを比較検討した。その結果、社会空間上の位置は、語彙を使用するか否かという水準にとどまらず、同一語彙が語りの中でどのような役割を果たし、いかなる意味づけが付与されるのかという水準にも対応して現れることが示唆された。

6. 結論

本稿の目的は、社会空間アプローチの文脈で収集されたインタビューデータの分析に、計量テキスト分析をどのように統合しうるのがかを探索的に検討し、その方法論的意義を提示することにあった。具体的には、質問紙調査データにもとづいて多重対応分析により社会空間を構築したうえで、第1に語彙の使用有無をサブリメンタリー変数として空間上に投影し、第2に、投影結果をもとに選定した語彙についてtf-idfを用い、個々の語りの中での前景化の程度と意味づけを検討する、という2段階の手続きを採用した。

第4節では、教育・生活スタイルに関する語彙の使用が、資本量および資本構成の差異と対応する形で社会空間上に配置されることを確認した。語彙の投影は、インタビューにおけるトピックの分布を、社会空間上の位置と接続して可視化するものである。これにより、社会空間とライフスタイルとの同定性仮説が、インタビューにおける語彙使用の水準からも検証しうることが示された。

もっとも、語彙の投影は、語彙が語りの中でどのような役割を担うのか、あるいは同一語彙がどのような文脈で用いられるのかまでは捉えない。そこで第5節では、「留学」「我慢」という語彙に着目し、語彙使用者の空間上の位置と、語りの中での前景化の程度(tf-idf)を手がかりとして実際の語りを比較検討した。その結果、同一の語彙であっても、語りの中心として用いられる場合と、限定的・補助的に言及される場合とが区別され、その意味づけは社会空間上の位置と関連しつつ分化しうることが確認された。すなわち、本稿の手続きは、「語彙を使う／使わない」という水準での対応関係にとどまらず、「同じ語彙をいかなる重みで、いかなる含意のもとで用いるのか」という水準へと分析の焦点を展開する足場を提供するものである。

以上を踏まえると、社会空間アプローチに計量テキスト分析を統合する試みには、以下の3つの方法論的意義があるといえる。第1に、語彙の使用をサブリメンタリー変数として投影することは、インタビューにおけるトピックの分布を社会空間上に配置し、複数のインタビューデータを鳥瞰的に捉えることを可能にする。第2に、tf-idfは、語彙が各事例の語りの中でどの程度前景化しているかを示す指標として、質的な分析に貢献する。第3に、語彙の投影とtf-idfを組み合わせることで、「どの語彙に注目し、どの事例に焦点化するのか」という選定過程を明示したうえで、「空間上の位置」と「語りの内容」とを接続できる。これにより、個別事例の断片的な解釈に依存しにくくなり、根拠の提示と再検討可能性を高めることが可能となる。計量テキスト分析は質的分析の代替ではなく、社会空間アプローチにおける解釈や論証の足場を増やす補助的な手段として位置づけられる。

今後の課題として、語彙の選定基準や前処理（ストップワード設定、辞書登録等）の影響を点検しつつ、対象語彙を拡張していくことが挙げられる。また、語彙の前後文脈を系統的に確認する手続き（KWIC等）を組み込むことで、計量指標と語りの解釈との往復をより強固にできよう。さらに、本稿の手続きは語彙を手がかりに分析の入口を設定する一方で、出来事の系列や転機といった——生活史では特に重要な——時間的連関を含む語りの構造を直接に捉えるものではない。したがって、文・段落・エピソードといった発話単位での集計や、複数語彙の共起関係にもとづくコードを導入することで、「語彙中心の接続」と「生活史としての全体像」との橋渡しを図る必要がある。これらの検討を通じて、社会空間アプローチにおける質的資料の活用を、より比較可能で累積的な形へと組み替える可能性が開かれると考えられる。

〈注〉

- (1) 知念編（2025）においても多重対応分析が使用されているが、軸の構築に用いる変数が田中（2024）とは一部異なっている。

- (2) 具体的には「ほう、とき、所、ほか、そちら、どちら、どこ、ヶ月、か月、匹、組、冊、割、笑」をストップワードとして指定した。
- (3) 語彙に関する情報はインタビュー協力者からのみ得られているため、インタビューを実施していない質問紙調査対象者については、語彙に関する情報は全て欠損値となる。ただし多重対応分析の性質上、サブリメンタリー変数に欠損値が多いことは問題とならない。
- (4) tf-idfの値は、可読性のために1000倍にして表示している。
- (5) 整理すると、図1は「語彙を使用した」というカテゴリーの位置を投影する一方、図2は個人の位置——と、個人のtf-idf——を示すものである。
- (6) 図1において「留学」の使用カテゴリーが西側にある一方、図2において「留学」のtf-idfが相対的に高い回答者が東側に配置されていることは、前注に示した理由により、必ずしも矛盾しない。

〈引用文献〉

- Bennett, T., D. Carter, M. Gayo, M. Kelley and G. Noble 2021 Fields, Capitals, Habitus: Australian Culture, Inequalities and Social Divisions, Routledge.
- Bennett, T., M. Savage, E. B. Silva, A. Warde, M. Gayo-Cal and D. Wright 2009, 磯直樹・香川めい・森田次朗・知念渉・相澤真一訳『文化・階級・卓越化』青弓社 2017.
- Bourdieu, P. 1979, 石井洋二郎訳『ディスタンクシオン—社会的判断力批判』藤原書店 2020.
- 知念渉編 2025『東京で育つ／育てる—母子の生活史と不平等の布置』有斐閣.
- Grimmer, J., M. E. Roberts and B. M. Stewart 2022 Text as Data, Princeton University Press.
- 樋口耕一 2020『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.
- 磯直樹 2020『認識と反省性—ピエール・ブルデューの社会学的思考』法政大学出版局.
- Iso, N., H. Takashi, K. Mori and K. Fujimoto 2025 "Habitus and the Relationship between Youth Culture and Transition: A Bourdieusian Analysis of Musical Culture and Inequality in Japan", Japanese Journal of Sociology, Vol. 34, No.1, pp.54-78.
- Jarness, V. 2018 "Viewpoints and Points of View: Situating Symbolic Boundary Drawing in Social Space", European Societies, Vol. 20, No. 3, pp.503-524.
- 近藤博之 2011「社会空間の構造と相同性仮説」『理論と方法』26巻1号, 161-177頁.
- 近藤博之 2012「社会空間と学力の階層差」『教育社会学研究』90巻, 101-121頁.
- 田中祐児 2024「多重対応分析とインタビュー分析による貧困観の検討」『理論と方法』39巻2号, 152-67頁.

Exploring the Integration of Quantitative Text Analysis into Bourdieusian Social Space Analysis

TANAKA Yuji

Recent years have seen growing efforts within the social space approach to combine questionnaire surveys with interview research. This study argues that applying quantitative text analysis to interview data can further develop the social space approach. Using a dataset consisting of both survey and interview materials, we first construct a social space from the questionnaire data and then link that space to the interview narratives through quantitative text analysis. Specifically, we project the presence/absence of selected lexical items onto the social space as supplementary variables, thereby visualizing the possibility that vocabularies related to educational and everyday lifestyles exhibit homologies with positions in social space. We then compute tf-idf scores for the same lexical items and show that, even when respondents use the same word, the degree of foregrounding and the meanings attributed to it can differ among individuals, and that such differences may be differentiated in relation to their positions in social space. The combined use of lexical projection and tf-idf makes it possible to connect social positions and narratives in a way that renders case selection explicit and the linkage open to re-examination. In this sense, quantitative text analysis is not a substitute for qualitative interpretation; rather, it functions as an auxiliary means of strengthening interpretation and argumentation within the social space approach.