

## ベクトル空間モデルに則った、 近代ストア主義とメスメリズムの類似性に関する計量文体論的分析

赤間啓之

[akama@dp.hum.titech.ac.jp](mailto:akama@dp.hum.titech.ac.jp)

東京工業大学 大学院社会理工学研究科

この論考は、情報検索学の手法を計量文体論に導入し、哲学的なコーパスから文脈の自動抽出を行うことを目的とする。具体的には高頻度名詞と共起する有意味単語データを因子分析することで、18世紀後半と19世紀初頭にかけて活躍したカバニスとメスメルという、二人の思想家の予想外な類似性を計算する。

### Computational Analysis of the Similarity between the Modern Stoicism and the Mesmerism, Based on the Vector Space Model

Hiroyuki AKAMA (Tokyo Institute of Technology)

[akama@dp.hum.titech.ac.jp](mailto:akama@dp.hum.titech.ac.jp)

In this paper, it is our aim to abstract automatically the latent semantic context from philosophical corpora by introducing the IR method into the stylometrics. We apply the factor analysis to the exhaustive data of content-bearing words co-occurring with the highest frequency nouns. By way of example, the unexpected similarity between Cabanis and Mesmer, two great thinkers in the late 18<sup>th</sup> century and early in the 19<sup>th</sup> century is calculated.

#### 1. 研究の目的

本研究の目的は、計量文体論的な方法を用いて、近代初頭、晩期啓蒙主義の時代のフランスにおける、ストア主義とメスメリズムの、目に見えない思想的類似性を論証することである。両者の類似性は、従来の方法では、丹念な解読を通じて初めて直感的に把握されるものであって、単語列の一致のような客観的な根拠を元に論ぜられるものではない。いわゆる通常の文献学的研究では、証明不可能である以上、情報要約のための特別な方法論を抜きにしては、単なる印象批評のレベルで終わってしまうものである。

そもそも哲学思想の類似性研究は、これまで、部分的・局所的な照合にとどまるか、あるいはホーリスティックな印象に依拠するだけかのどちらかであった。前者の場合、特定の孤立した単語ないし単語列に、天下りの的に過度の注意が傾けられる。それがまったく有効でない事例となると、後者の場合のように、類似の印象を語る同語反復的・自己言及的な言葉を延々と紡いでいくほかはなかった。対象の全体をカバーすると同時に細部の情報を吸収できる綿密・厳密な方法の研究は、人文科学においてこれまで極端に不足していた。

そこで、本研究では、情報検索学の基本概念であるベクトル空間モデルを計量文体論に導入し、文脈の自動抽出技法を利用することで、哲学・思想の目に見えない類似性を発見的に探索したり、あるいは外的証拠の不在によって合意の得にくい類似性を承認したりする、方法論上の可能性を模索する。周知のように近年、LSA（潜在的意味分析）、Word Spaceは単語共起データの多変量解析、とくに主成分分析（あるいはその元にある特異値分解）に基づいて、文書の自動的なクラスタリングと検索のシステムを実装してきた。

特に、共起行列の概念空間を特異値分解により縮退させたWord Space（高山、Flournoy、Kaufmann、Peters）では、文書の特徴づけない不要語（以下、ノイズワードと略）を取り除いた高頻度語である内容表現語の一つの周りで、何語か以内に共起する単語を収めた「ウィンドウ」を設定している。本研究ではこの「ウィンドウ」に想を得、哲学思想テキストから単語の共起行列を出力させ、それを因子分析にかけて複雑な文脈の個別抽出を行うことを試みた。次節に見るとおり、相互交渉の証明が困難なストア主義とメスメリズム-そのいわば不可視の-「関係」は、本研究が提案する方法を用い、潜在的な共通文脈と解釈される因子をいくつか形成しようという意味で、きわめて興味深い歴史的サンプルとなりうるものである。

## II. テーマの背景

ここで思想史の大きな流れを概観する形ではあるが、比較の対象となる両者の「概念図」を簡単に紹介することにしよう。まず、近代における「ストア主義」だが、これは、コンディヤック流の感覚論哲学から派生したいわゆる「イデオロジー（観念学）」において特徴的であり、これを体現する代表的な人物は、ミシェル・フーコーにより再評価された医学哲学者ピエール＝ジャン＝ジョルジュ・カパニス(1758-1808)である。一方、「メスメリズム」は、ウィーン出身でパリに出て成功した医師フランツ＝アントン・メスメール(1734-1815)が提唱した、「動物磁気」という名の一種の催眠療法であり、フロイトから遡ること百年、現在ラカン派へと引き継がれた精神分析の淵源に位置付けられている。一方は個人的禁欲の倫理、他方は集団的癒しの医学をそれぞれ指向する点、一見まったく異なるイメージを持たれがちなものだが、ともに、フーコー（認識論）ラカン（精神分析学）ら、フランス構造主義の思想を構成するいくつかの流れの源流を為す、重要な原-思潮だと言って過言でない。

にもかかわらず、両者の影響関係については、文明史関係で先行研究と呼べるものがほとんど存在していないというのが実情である。それは近代ストア主義独特の形而上学的、あるいは神秘主義的な傾向が、その母体であった啓蒙思想のもつ科学的合理性と明確な対立点を形成していないためである。理性の人カパニスと当時の秘教的科学との関係は、ようやくその研究が端緒に着いたばかりである。たとえば秘教的科学の代表である「骨相学」(ラヴァター)の分野では、最近ようやく、引用部の厳密な照合に基づく事例指向的な(レイガンの語法による)文献研究(赤間、2000)により、カパニスへの影響関係が明確に証明された。だが骨相学と並ぶ、当時の重要な秘教的科学のひとつ、動物磁気と呼ばれた催眠療法については、事例指向的研究を支える文献証拠がまったく存在していない。催眠療法(動物磁気)の提唱者メスメールがカパニスに与えた影響は、両者の基本的資質があまりに相異なることもあって、これまで論ずるに足るテーマとして論じられることがまったくなかった。

ただ、医学生であったころのカパニスには、じっさいメスメールと交わした師弟関係の契約書が今も謎めいたものとして残されている。カパニスはしかしその後、メスメールの名も著作もまったく引用することがなかった。動物磁気について暗黙のうちに言及したくんだりさえ存在していない。メスメ

ールが主な治療対象としていた神経症について、様々な医学、哲学関係のテキストを通じて執拗に論じていたにもかかわらず、である。たとえば下腹部臓器や神経系の受ける刺激印象の違いに基づく、身体的な狂気（ヒステリー、ヒポコンデリー）が、両者共通のテーマとなっていることに注目したい。その意味で両者とも、神経症の科学史においては、同じ分野で理論・実践を展開した同時代人である。

さて、本研究において比較の対象として取り上げたテキストは、Georges Cabanis, *Lettre à M.F. sur les causes premières* (ジョルジュ・カバニス、『第一原因についてのF氏への手紙』), F.-A. Mesmer, *Mémoires de F.-A. Mesmer, docteur en médecine, sur ses découvertes* (F.-A. メスメール、『医学博士F.-A.メスメールによる、彼の発見についての論文』)のふたつであり、前者は近代ストア主義、後者はメスメリズム（動物磁気）を集約した作品と言える。さらに両者を比較すると、カバニスには、思想、観念、内省に関わる「哲学」系のタームが多く、メスメールには、物体、流体、身体に関わる「医学」系のタームが多い。一見すると両グループに関わる文脈はなさそうに思える。が、当時はイデオロギーや（さらに歴史的に長い射程を有する、19世紀的な意味での）心理学とも共存して、その名も「医学哲学」という、特異な言説のジャンルが一時代を画していた。それは先述のごとく、ヒポクラテス、ガレノスの伝統でもある古典的な医学の最後の光輝であった。もしカバニス、メスメールを通底する文脈が客観的な方法で抽出できた場合、その文脈を構成する語は（因子分析をして両者に横断的な因子が取れた場合は、その因子の得点の高いものとしてクラスタリングされるサンプル語）は、「医学哲学」というジャンルを成立させる媒介の役割を果たしていたとは言えないだろうか。これがわれわれの仮説である。

### III. 研究の方法

ここで、研究の方法に関して言及する。まずふたつの異なる思想について、その潜在的な類似性を論じるためには、複雑な文脈を的確に個別抽出する手段として、多変量解析、特に因子分析が有効である。ここでは、ひとつひとつの作品の内容ではなく、双方の作品が共存し複雑に絡み合っている言説空間全体を要約的に提示することを目的としている。そのため、異なる文書を、長さを規準化しつつ合併連結し、双方で最頻出するキーワード名詞を変数にし、それらと一定のウィンドウ幅内で共起するすべての語をサンプルとして出現頻度をカウントして、パリマックス回転付の因子分析を行う。合併連結するにあたり文書の長さを揃えるので、単語は出現回数1回につき単純に1個とは数えず、各文書領域の長さ（ノイズワードを取った後の総形態素数）に従って適当な重みをかけた。

データの取得には、高山らのWord Spaceにおける「ウィンドウ」に想を得、FreeBSD2.2.8上のPerl言語で書かれた自作のウィンドウ・スクリプトcoexcountを利用し、以下に述べるような条件で、哲学思想テキストから単語の共起行列を出力させた。このスクリプトは、一定の語数幅の共起ウィンドウに登場する共起語の数をカウントする（中心に置かれたキーワード自体はカウントしない）、共起ウィンドウ制御スクリプトである。引数4個を取り、第1引数はターゲットとなるテキストファイル名、第2引数はウィンドウの中心に置くキーワードのリスト、第3引数はウィンドウの幅となる語数、第4引数はウィンドウの境となるキーワード（ここを超えた位置にある単語はたとえウィンドウ幅の中にあってもカウントしない）である。最後の引数はウィンドウの幅に制約条件を設け、章や節などの切れ目で特定の範囲を超えてウィンドウが拡大しないようにする。

このcoexcountの利用にあたっては、まず当該テキストを適当に下処理した後、フランス語の統語解析・形態素解析ソフトCordial6にかけて、1行1語基本形の形に変換しておく。さらには、前もっ

て一定の基準に従いノイズワードを除去することが必要である。哲学思想の類似性判定の場合、トータルなものとしての概念体系を問題にする以上、ノイズワードはあらかじめ除去してかかるのが自然だからである。ノイズワードリスト(インターネットで参照可能)は、個人の言語経験に基づき恣意的に作ったものを鍛え上げたが、ノイズワードに単独の普通名詞を含まないという条件を設定した。

問題の二文書をそれぞれ、以降C(カバニス), M(メスメール)と略して表記するが、データを採取する対象はCとMを連結合成した、単語数にして長さ10482個のテキスト(ノイズワードは除去した)であり、これをC&Mと名づける。そしてC、Mのそれぞれ最頻出名詞上位50位まで計77語を抽出し、それらをC&M上で走らせるウィンドウの中心に置くキーワードとした。ウィンドウは、キーワード毎にひとつずつ設定されるが、C&Mの開始語から終了語に至るまで、テキスト上をくまなくスライドしていき、その中心にキーワードを発見するとストップする。そしてウィンドウ中にセンサーされた共起語をカウントした後、次のマッチングを求めて移動を再開する。この操作をすべてのキーワードについて行う。

キーワードに関しては、C、M双方に共通して50位以内のものが26個、Cのみ50位以内のものが26個、Mのみ50位以内のものが25個の名詞が抽出された。これら最頻出77語のキーワードは、ほぼ1/3ずつに等分される。

C、M双方に共通して現れる高頻度名詞(これらを共通語と呼ぶ)の多さは、テーマの類似性を示唆するが、たとえば、C、Mそれぞれから頻度数で上位20位までの名詞に限ると、およそ半分の9個の名詞が共通しているほど、双方は重なっている。(なお当該名詞のうち、それとほぼ同義の形容詞は、同条件で頻度の観点から重要と判断された場合、文脈に対する解釈上の配慮から名詞扱いとした。さらに当該名詞のうち、Mにとって最重要タームである*magnétisme animal*(動物磁気)については、ほぼ一語扱いの可能なバイグラムのコロケーションであるので、*magnétisme-animal*と連結して1変数に設定し、後続に*magnétisme*を伴わない*animal*と区別した。)

さて、それら、二文書に共通する高頻度語彙が多い一方、共通しては存在しない(一方のみの)高頻度語彙もいくつか散見された。それらは検索語となる資格を持ってそれぞれの文書の独自性を特徴づけると考えられるので、これを独自語-検索語と規定する。たとえば、キーワード(最頻出名詞上位それぞれ50位以内)計77語のうち、一方の文書にのみ出現する独自語-検索語、あるいは基準を緩めて、他方の文書では上位100位以降にしか出現しない低頻度語(一方の文書では反対に独自語-検索語と見なされる)の個数は、以下の表のとおりとなる。

共通語	26
C独自語-検索語(Mになし)	3
C独自語-検索語(Mは50位以下)	4
C独自語-検索語(Mは100位以下)	19
M独自語-検索語(Cになし)	5
M独自語-検索語(Cは50位以下)	4
M独自語-検索語(Cは100位以下)	16
計	77個

仮説として、キーワードのうち、どちらか一方の独自語-検索語と双方の共通語が、C&Mの共起

関係をめぐって、ともにある特定の因子の構成に関与していた場合、それらは、表層的なばかりか潜在的な共通文脈を形成していると考え。言い換えれば、CとMを横断し、両テキスト全体が存している概括的・包括的な文脈を明らかにする因子において、たとえばCに1度も現れない語が、高い因子負荷量で位置付けられているならば、Mを弁別するこの最重要検索語は、まさしくCにとっては、ソシユールの意味での「アナグラム」(赤間、2000)のように、あってしかるべき場になぜか見出されない、ということである。それは逆に、両者に深いレベルで通底している、テキストの表面には現れない、包括的な思考世界を顕示することになる。

ただし、ここでは文脈抽出する上で、ウィンドウ幅nをパラメーターとして考慮しなくてはならない。以下の表は、ウィンドウ幅を適当に変化させて取得したデータが、因子分析した結果にどのような相異を示すかまとめたものである。

キーワード数	第一因子寄固有値1以上の					
	Window幅	与率	因子数	累積寄与率	回転後の累積寄与率	共起語の種類
77	3	20.18	19	57.62	43.71	1972
77	5	28.64	14	60.60	50.89	2134
77	10	40.90	11	67.68	62.18	2242
77	15	47.63	9	72.01	68.31	2255
77	20	52.10	8	74.97	72.06	2257
77	30	57.49	7	79.10	76.92	2257

ウィンドウの幅を変えた実験結果では、変数相互の相関が大きい、つまり特定の共起語がどのキーワードにもある程度は共通して高頻度で共起するため、第一因子の寄与率はかなり高く、それもウィンドウ幅の拡大と連動し、共起語の種類(異なり数)が最大限に達しても増大しつづける。また第二因子以下の単独寄与率は急に落ち込むので、そこではいろいろな相関が混在して複雑なデータになっていることが見て取れる。またウィンドウ幅がごく狭い場合、多くの因子を選択しても、累積寄与率がなかなか大きくなり、そこではさまざまな細部情報が複雑に絡まり、多くの相関が混入し、全体像をつかみづらいことがわかる。一方、それに対し、固有値1以上の因子数は減少しつづける、それに見合っただけで累積寄与率も上昇しつづける、つまり、テキストの世界が総体的に捉えられるようになり、基幹的な情報がわずかな数の因子に集約していくのがわかる。

#### IV. 研究の結果

ここでは代表例として、C&Mテキストに対するウィンドウ幅5の共起データを因子分析(主因子法)にかけ、出力された結果を取り上げてみる。変数(キーワード数)は77個、サンプル数(共起語の異なり数)は2134個。固有値1以上という条件では、15個の因子が抽出され、それらにバリマックス回転を施した。因子負荷行列から得られる因子の解釈内容は順次提示する。説明された分散は以下の表の通りで、固有値1未満の因子は割愛した。ウィンドウ幅5という広さは、共起データによる因子分析(主因子法)を行うにあたり、抽出される因子の数やそれらの命名解釈のしやすさの点で、概ね適当であると思われる。以下の表は、スペースの関係で因子1と因子2のみの因子パターンを、それも因子負荷量の大きい変数に限定して掲載したものである。

主因子法 Kaiser の正規化を伴うバリックス回転 15 回の反復で回転が収束。

CM 共通高頻度

C 高頻度

M 高頻度

大文字 片方 1 0 0 位以下

大文字・斜体 片方に無

回転後の因子行列

C での頻度 M での頻度

Window 幅 5

	C での頻度	M での頻度	因子 1	因子 2
Corps(身体)	21	38	0.681	0.201
FEU (火)	2	13	0.625	0.016
MAGNETISME (磁気)	0	19	0.611	0.130
Animal (動物)	12	29	0.602	0.250
FLUIDE (流体)	3	44	0.598	0.292
Organisation (組織)	22	15	0.587	0.220
Influence (影響)	14	14	0.576	0.102
Mouvement (運動)	31	72	0.548	0.399
Action (作用)	33	35	0.538	0.396
Propriété (特性)	22	23	0.529	0.160
MECANISME (機構)	2	17	0.486	0.190
Principe (原理)	20	30	0.476	0.130
Partie (部分)	40	26	0.442	0.217
Loi (法)	33	16	0.419	0.108
Résultat (結果)	25	11	0.412	0.235
LUMIERE (光、啓蒙)	9	15	0.397	0.052
THEORIE (理論)	1	12	0.358	0.065
Nature (自然)	56	40	0.335	0.162
RAISON (理性)	21	7	0.291	0.101
Organe (臓器)	36	36	0.243	0.723
Sens (感)	43	16	0.106	0.702
SENSATION (感覚)	3	25	0.309	0.685
NERF (神経)	3	19	0.300	0.681
Impression (印象)	19	21	0.112	0.622
SUBSTANCE (実質)	1	18	0.324	0.473
INDIVIDU (固体)	18	7	0.075	0.345

まず第一因子を見てみると、これはCテキストとMテキストを横断し、両テキスト全体が存している概括的・包括的な背景文脈を明らかにする因子であることがわかる。すなわち、CとMの共通文脈を形成する双方高頻度な共通語(corps-身体を筆頭に、organisation-組織、influence-影響、mouvement-運動、action-作用、など)が、この第一因子を形成する変数の中に認められる。これらは個人的なバイアスがそれほどかかっているとはいえない、ごく一般的な語である。

が、不思議なことに、この因子中にはCテキストに1度も現れない語もまた、高い因子負荷量で位置付けられているのである。すなわちMagnétisme-磁気という語であり、Mテキストを弁別するこの最重要な独自語-検索語は、まさしくCテキストには1度も出現せず、一種、そこにあるはずが直接は見えてこない「アナグラム」(語の下の語)のごときものになっている(Cでの頻度:0、Mでの頻度:19)。

先述のごとく、メスマール(M)は、動物磁気という一種の催眠術を提唱し、カバニス(C)は、若いころ動物磁気の使用にかんし、メスマール(M)と契約を取り交わしたが、その後、磁気についてはまったく無視して触れることがなかった。にもかかわらず、CとMの共通文脈を形成する双方高頻度な共通語に混じり、動物磁気関係のM独自語-検索語(FEU-火、MAGNETISME-磁気、FLUIDE-流体)が、この因子においては同時に高い因子負荷量を示し、最も重要な共通因子を形作っているのである。そこでこの因子を(全体潜在文脈-)磁気媒体因子と命名する。この因子の存在は、カバニスとその至ってニュートラルな側面においても、見えないところでメスマールの強い秘教的独自色を隠し持っている(継承している?)ことを暗示している。

次に第二因子を見てみると、これもまたCテキストとMテキストを横断し、両テキストの類似性を表象する因子である。しかも個人の神経系、臓器系の感覚、感受性に関わる深い連想関係にあるキーワード(organe-器官、sens-感覚、SENSATION-感覚、NERF-神経、impression-印象)が高い因子負荷量で並んでいる。そこでこの因子を身体内部感覚因子と命名する。すなわち、当時の感覚論を踏まえた医学哲学に特徴的なテーマがそこに見出せる。だが、SENSATION-感覚、NERF-神経、INDIVIDU-個人の3語に特に着目すると、出現頻度には大きなばらつきがあるのがわかる。SENSATION-感覚はCテキストで3回と希少なのに対し、Mテキストで25回と高頻度で、NERF-神経はCテキストで3回と希少なのに対し、Mテキストで19回と高頻度で、INDIVIDU-個人はMテキストで7回に對し、Mテキストで18回現れている。これらの語は、当然たくさんあって然るべき場にあまり見出されなかったことになる。しかし用語の面で多少偏りがあり、その意味で双方の差異が際立ったとしても、この因子が明らかにする文脈対象-やがてはメヌ・ド・ピランらによって発展的に解消される「医学哲学」-は、潜在的には双方のテキストが同時に指向しているものである。

第三因子以下については、変数が膨大なもので、その詳細は因子負荷行列ともども省略する。なお、念のためにこれらの因子を命名解釈した結果は以下のとおりに掲載しておく。

第三因子：病的催眠因子(ほぼCM横断因子) 第四因子：主意知力因子(Cの弁別因子) 第五因子：仮説実験因子(Cの弁別因子) 第六因子：欲求関係因子(ほぼCの弁別因子) 第七因子：謬見出所因子(CM横断因子) 第八因子：天空媒体因子(Mの弁別因子) 第九因子：自我感知因子(Cの弁別因子) 第十因子：倫理道德因子(Cの弁別因子) 第十一因子：物質秩序因子(ほぼCM横断因子) 第十二因子：因果関係因子(ほぼCM横断因子) 第十三因子：努力因子(ほぼCM横断因子) 第十四因子：観察実験因子(ほぼCM横断因子) 第十五因子：観念時間因子(ほぼCM横断因子)

注目すべきは、Cを弁別する因子(弁別因子)にくらべ、Mを弁別する因子の個数が少ないということである。これは、Cに比べMが比較的単純明快、ストレートな文脈から成り立っており、反対に

Cでは様々な個別的-部分的文脈が混在して複雑な作品に仕上がっていることを示唆しているだろう。

そして第一因子すなわち(全体潜在文脈-)磁気媒体因子などに見られるとおり、Mの議論の独特でドグマ的、直截的な性格は、実はC,M共通の関心テーマ系の中でその主調(メイントーン)を決定しており、そこで我々はMのターミノロジーの孤立した珍奇さだけに迷わされてはいけない、ということになる。「催眠術の一種としての動物磁気」という、きわめて特殊な観念といえど、人間の身体組織をめぐる思想的考察という、当時の一般的なイデオロギーの中に自然にすっぽり収まるということであり、その独自の思想的バイアスは、第二因子-身体内部感覚因子-で示唆される「医学哲学」というニュートラルな言説領域において、うまく均されるのだと言ってよい。

#### V.まとめ

概括すると、ストア派のカピニス(Cテキスト)と動物磁気論者メスマール(Mテキスト)について、連結されたC&Mテキストの共起語を対象とする因子分析により、CテキストとMテキストを明確に分別する因子の他に、両者の深い類似性を示唆する複数の横断的な因子が抽出された。さらにキーワードの頻度データまで考慮すると、その類似性は同時代性の文脈と言った中間的・一般的な性格にとどまらず、独自の用語をめぐるかなり特殊な概念のレベルにまで及んでいるとすることができる。

#### 参考文献

Hiroyuki Akama(2000), *Cabanis ou le crépuscle métaphysique*, in Actes du XXVIIe Congrès ASPLF, J.Vrin, Paris, p.297~305

Gerard Salton(1975), *A theory of Indexing*, Society for Industrial and Applied Mathematics

Hinrich Shutze(1997), *Ambiguity Resolution in language learning*, CSLI Publications

高山泰博、Raymond Flournoy、Stefan Kaufmann、Stanely Peters(1999)、単語の連想関係に基づく情報検索システム InfoMAP、情報学基礎 53-1

三宅真紀、赤間啓之、佐藤研、中川正宣(2000)、因子分析による共観福音書問題の解析、統計数理、第48巻 第2号、p.327~337