

科学文献のテキスト分析に基づく

科学知識の形成過程

: 医学文献の分析を例として

武者小路澄子

図書館情報大学

新しい科学的発見が発表されてから、専門分野の知識へと統合されるまでの過程を「科学知識の成過程」と捉えた。「科学知識の形成過程」は、多様な科学文献を含む科学のテキストがいかに作成されるかへの探索なしには、研究することができない。事例として基礎医学分野の<エンドセリンの発見>に関連する6種類の科学のテキストを取り上げ、<発見>の記述や<発見>を構築するテキストの構成を分析した。別種のテキストや後続するテキストでは、提示される<発見>が変わり、医学知識としての位置づけも展開していた。<発見>を言い換えていくテキストの連鎖の過程こそが、「科学知識の形成過程」を示すものである。

' THE PROCESS OF CONSTRUCTING
SCIENTIFIC KNOWLEDGE'
BASED ON TEXT ANALYSIS
OF SCIENTIFIC LITERATURE
: THE CASE OF <DISCOVERY OF ENDOTHELIN>

Sumiko Mushakoji

University of Library and Information Science

1-2 Kasuga, Tsukuba-shi, Ibaraki-ken, 305 Japan

'The process of constructing scientific knowledge' is defined as a process in which a new scientific finding is transferred as 'scientific information' until it is finally integrated into the 'expert knowledge'. This process cannot be studied without exploring how different kinds of scientific texts (about the same finding) are written in each stage of the process. Comparing six kinds of texts about the <Discovery of Endothelin>, how the <Discovery> is described was examined, and how the features of textual organization construct the <Discovery> itself was analyzed. The <Discovery> changes in the different kinds of / successive texts and their relevance to the 'knowledge' was found to be progressive. In the sequence of these scientific texts, 'the process of constructing scientific knowledge' is achieved.

1 「科学知識の形成過程」

情報学には、科学情報の流れに関するいくつかの理論的モデルが存在する。その大半は、研究者が新しい科学的発見やアイデアを発表すると、それが多様な伝達経路を経て伝達されていくことを仮定している。そうした伝達の過程では、科学情報は繰り返し伝達され、更に様々な形に加工されて、最終的には特定の専門分野の知識へと統合されると言われる。このようなモデルを、本研究では、「科学知識の形成過程」と呼ぶことにする。

こうしたモデルの中で重要な役割を果たすと考えられるのが、様々な機能を持つとされる科学文献である。例えばSubramanyamは、科学情報の流れを解説する中で、情報の流れの各段階に属する科学文献を挙げている¹⁾。科学文献は、一方では続く科学文献の中で繰り返し引用され、他方ではそこから重要な情報だけを抽出して有効利用するために新たな文献へと加工されていく。前者の側面に対し、引用行動によって科学を研究することが試みられている²⁾。後者のためには、科学文献の分析や統合、評価、あるいは最近では知識ベースのための知識獲得といった領域で盛んに研究が行なわれている。

以上のように、科学文献がいかにかに作成されるかを捉えることは「科学知識の形成過程」を研究する上で不可欠である。しかしながらそれを、具体的な科学文献を取り上げて、そこで科学者や我々の実践の中で「科学情報の流れ」や「知識」がいかにかに構成されるか、を探ろうとする試みは多くない。本研究では、科学文献に関わる概念として「科学のテキスト」に焦点を当て、特定の研究活動に基づいて作成された複数の科学のテキストを分析することから、「科学知識の形成過程」を捉え直すことを試みる。

2 <エンドセリンの発見>を例としたテキスト分析

「科学情報の伝達」を行為として捉えた場合、それは人が情報メディアに<何か>を表現し、その表現によって相手がその<何か>を伝えられた、とみなされることによって成立する³⁾。こうした場合に、人々によってひとまとまりの表現とみなされるものを「テキスト」と捉えることにする。

前研究³⁾に続き、「科学情報」は「伝達」行為の中で科学のテキストを対象化して説明し直した

ものであると考え、そのような伝達行為の中にある科学のテキストを取り上げることにした。「科学情報」として挙げられる主要なもの一つに、研究活動から生み出された新しい発見やアイデアがある。そこで、特定の<科学の発見>が提示されているとされる場合を取り上げ、その<発見>を伝達する科学のテキストに着目した。そして、「科学知識の形成過程」の別の段階に属する複数の科学のテキストを比較して分析した。

2.1 分析対象

事例として<エンドセリンの発見>に関する科学のテキストを分析の対象とした。

「エンドセリン」(endothelin)は、1988年に筑波大学基礎医学系の研究グループにより最初にNature誌に発表された、血管収縮作用に関与する新しいペプチドである。この発表はそれ以後、エンドセリンに関する様々な研究を誘発し、現在エンドセリンの研究は一つの研究領域としての地位を確立している⁴⁾。医学分野では最も利用頻度が高い医学のデータベース、MEDLINEではエンドセリンに関する文献は数千件にも上り、Nature誌に発表された最初の原著論文は、1991年以降毎年700件を上回り続けている。

取り上げた科学のテキストは、次の6種類である(詳細は付表)：

- 1) Nature誌に最初に発表された原著論文の本文 (以後、1988 原著 本文と呼ぶ)
- 2) 同 見出し (1988 原著 見出し)
- 3) 同 標題 (1988 原著 標題)
- 4) 同一著者が4年後に執筆した総説の本文 (1992 総説 本文)
- 5) 同 標題 (1992 総説 標題)
- 6) 同一著者へインタビューを行なった録音の転記 (transcription) (1994 インタビュー)

2.2 分析手順

前研究³⁾において行なった1988 原著と見出しに対する命題分析、言語学的分析(モダリティ、ダイクシスの分析)、スタイルの分析の結果をふまえて、1992 総説に対して同様に分析して比較した。その際、1992 総説 本文では、a)全体のテキスト、b)1988 原著を直接引用している部分のテキスト、の2通りに対しての比較を行なった。

著者である基礎医学の研究者に、「構造化され

ていない (unstructured) 」インタビューを反復的に行なうことを依頼した。インタビューでは、特に<エンドセリンの発見>に関わる研究活動 (発表活動を含む) を話題に展開するよう、会話を進めた。会話を録音する許可を得て、転記した。

転記したテキスト中で<エンドセリンの発見>に関連する箇所を、モダリティ、ダイクシスの点から分析し、他のテキストと比較した。

更に、各テキストの組織のされ方が、「科学知識の形成過程」におけるどのような段階の<発見>を構築しているかという点から、全てのテキストを比較した。その際、エスノメソドロジーの枠組みに立って科学のテキストを分析したWoolgarが提示した、<発見>というクラスに属する現象を表現するための規則や手続きを与える、テキストの組織化の諸特性⁶⁾を参考にした。そして、各々のテキストが、<エンドセリンの発見>の現象としての認識、信頼性、重要性、医学知識の中での位置づけの点で、どのように構成されているかをまとめた。

2.3 結果

分析対象において、<エンドセリンの発見>がテキストとして直接的に提示されている部分の記述を、例示的に第1表に抜き出した。ただし、これらの記述は、全体のテキストにおかれて初めて、<エンドセリンの発見>についてであるという理解が達成されるものである。

読み手が一番最初に目にすると思われるのが、原著 1988 標題である。標題には、直接的な「発見された」という記述も、発見者や現場、発見の方法についての記述もない。“novel (新規な)”という語は、発見に関連し、実体化された<もの>(即ち、“peptide”)>を修飾する。比較して、原著 1988 見出しでは、<もの>が「単離され、〜と示された」という行為が記述されている。続いて、実験手順によってそれがどのようなものであり、そのことからどのような仮説が示唆されるかが記述される。こうした<もの>に施される行為の記述ではなく、原著 1988 本文では、“We (我々) ”、即ち科学者が、<もの>を「単離する」主体となる。言い換えれば、“we”が働きかけることによって<エンドセリンの発見>が起こった、という記述になっている。同時に、<発見>を導く過程がはるかに詳細に記述され、<もの>の単離、精

製、DNA配列決定、生合成、活性の測定、DNA構造の決定、などに関する方法と手順を一つのプロセスとして記述する中に<発見>が位置づけられている。

一方、総説 1992では、<エンドセリン>が最初から主題化されている点で原著 1988 本文のテキストと異なる。言い換えれば、それ以前に存在しなかった<エンドセリンの発見>が導入されるのではなく、既に発見済みの、既知の存在としての<エンドセリン>を紹介する記述に変化している。総説 1992 本文は、原著 1988を引用してはいるが、これ自体は特定の主体が、特定の現場で、一つのプロセスとして達成する<発見>ではなく、予め知識の中に位置づけられた<エンドセリンの発見>を示すものである。

インタビュー 1994の場合、「これ」と言及されたものが、総説 1992と同様、最初から主題化されている。が、総説とは違い、このテキストでの<発見>は、主体、現場、方法と手順の記述をともなったプロセスとして記述されている。また、<発見>が始まってから、「今で言うエンドセリン」への命名があったことが示されている。ここでの主体は原著 1988とは一致しない(原著では9人の<発見者>がおり、インタビューでは「私」「僕」「僕ら」が主体となり、「僕ら」も場合によって2人であったり3人であったりし、同一人物を指さなかった)。より重要なことには、<発見>の過程の中にある一連のプロセスは、違ったつながり方をしていた。例えば、原著 本文では、単離・精製→生合成・活性の測定→DNA配列の決定、の直線的順序で行為が記述されるが、インタビューでは、「ペプチドを合成していくっていう、あのおう、技術もっていなかったんで、...にお願いして作ってもらったわけですね。で、その間にA君がその構造を基に、遺伝子を決めて、でこれを...」と別の連鎖をもっている。

次に、各々のテキストでの<エンドセリンの発見>の構築のされ方を、第2表に概要としてまとめた。<エンドセリン>の存在とその<発見>の信頼性は、原著 1988と比べ、総説 1992やインタビュー 1994では大きく変化を遂げた(高まっている)。一方、<発見>の重要性に関しては、原著 1988がその研究領域と直接結びつく狭い専門領域に関連するものであったのが、総説 1992やインタビュー 1994では、多様な領域の発展と関わるよう

第1表 <エンドセリンの発見>の記述の変化

原著 1988 標題	原著 1988 見出し	原著 1988 本文
A novel potent vasoconstrictor peptide produced by vascular endothelial cells	An endothelium-derived 21-residue vasoconstrictor peptide, endothelin, has been isolated, and shown to be one of the most potent vasoconstrictors known.... (発見の手順)... Expression of the endothelin gene is regulated by ..., indicating the existence of a novel cardiovascular control system. (全体で6命題)	—それ以前の研究 (Furchgott & Zawadzki, 1980)、ペプチド研究で解明された/不明なこと We have now isolated a potent vasoconstrictor peptide ..., determined its amino-acid sequence, and molecularly cloned the peptide precursor. This peptide, endothelin, does not belong to any previously known peptide family... (序の1部) 以後3節分、<発見>の順序

総説1992 標題	総説 1992 本文	インタビュー 1994
心血管系におけるエンドセリンの分子薬理 (総説 循環薬理学の新動向)	(「はじめに」第1文) 「エンドセリン (ET: endothelin) はブタ内皮細胞の培養上清中から単離、精製された21個のアミノ酸からなるペプチドであり、分子内に2個のジスフィド結合をもち、非常の持続の長い血管収縮作用を有するという、極めてユニークな内因性活性物質である ¹⁾ 」 (原著1988引用) 以後後続する研究を4章にまとめ、未解明の点や動向などを概括した「おわりに」で締めくくっている	1) 「たまたま何か面白そうなものがある、というpaper」をジャーナル・クラブで読む 2) 「それから、僕とK先生とで、いずれやってみようかと、話して」 (神経ペプチド研究者) 3) 「そしたら1年後ぐらいに、A君っていうのが、何かやりたいと。」 「我々の所で一緒にやりたいと来た」 (分子生物学的な知識を持つ研究者) 「その3人が集まって、じゃあ、これがまだやらないで残してあるから、これから始めましょう、と。これが1970、いや、1987年の、今ごろですから7年前ですかね。それで4月からもう早速始めたんです。すぐ始められる状態にあったわけですからね。それでもうその夏には、そのう今で言うエンドセリンなんですけども、そのペプチドを単離して精製して、構造も決めちゃったわけです。8月の終わりまでには全て、そこまでやっちゃっらわけです。」——以降、手順の詳細の説明 「... どころか、殆どのものが、もう本当に、もう一発で成功しちゃったっていう、これまた珍しいことです」 「例えばこういうpaperを、最初からね、頭の中に描いてこういう研究なんかできたら、これ程もう、すばらしいことはないんですけど、実際はそうじゃなくて、あ、こんなすごいもんか、こんなすごいのかっていうことで、こういうのが出来上がってくる。そういうのが現状なんですよね」

第2表 「科学知識の形成過程」における<エンドセリンの発見>の構築のされ方の変化

	原著 1988 標題	原著 1988 見出し	原著 1988 本文	総説1992標題	総説 1992 本文	インクビュー 1994
<エンドセリン>の存在	初めて報告 :「<もの>の提 示	初めて報告 :「<もの>が発 見された」ア 命名	初めて報告—ただし 過去の発見を継ぐ :「我々が<もの> を発見した」ア 命名	総説1992標題 所与 :「<エンドセ リン>が実体 として主題化 されている	所与 :「<エンドセリン>が実体として主題化されて いる	所与 :「<発見>のプロセスの中で、<エンドセリン> と命名される
<エンドセリン>の 信頼性	雑誌名・著者・所 属機関の定評 から前提 → 本文参照	雑誌名・著者・所 属機関の定評から 前提 → 本文参照	雑誌名・著者・所 属機関の定評から前提 一貫性のある、周到 な「実験方法の記述	雑誌名・著者 ・所属機関の 定評や総説と いうメディア の性格から前 提	雑誌名・著者・所属機関の定評や総説という メディアの性格から前提 ・反復性や領域としての確立の根拠付け (117 文献が<発見>に基づいて、一領域をなす) ・総合化 (117文献が全体として<発見>の文 脈に織り込まれている)	前提: 研究者が「 <i>in vitro</i> 」に<発見>を解説するとい う文脈——紹介の段階で、既にそうした権威付け ・研究者の所属、業績、地位 ・発表後の反応: 多数の研究者が<発見>を引き継 いだ (訪問、研究材料の提供、引用回数への言及)
<エンドセリン>の 重要性	vascular endo- thelial cells やpeptide, 血 管収縮研究にレ レバント ('novel' な心 血管 制御機構を示唆 もの>)	左欄に加え、「血 管収縮作用の機構 」を採る研究にレ レバント (「 'novel' な心 血管 制御機構を示唆 もの>)	過去にない (「 This peptide, en- dothelin, does not belong to any pre- viously ...」) ・「今までで最も効 力のある哺乳類の血 管収縮ペプチド」 ・「新しい血管収縮 の機構を示唆」	総説というメ ディアの性格 からの権威付 け 総説のテーマ (循環薬理学) にレバント	総説というメディアの性格からの権威付け ・続行されかつ誘発されたものとしてまとめら れた文献の数と多様性 ・<もの>の特異性: 「極めてユニークな」 ・研究の進展における意義の言及 学術的: 「～作用は、生理的には極めて 重要な意味を持つものと思われる」 臨床的: 「<発見>に関わる病名の言及、 薬の開発への示唆	文脈で前提: 一つの研究史として語られる ・研究内容: 多様な専門家が共同で「一発で成功」 した<もの>が「<とんでもない性質を持っている」 ・原著 1988発表後の反応: ノーベル賞級の研究者 をはじめ多くの訪問/学会やツボ/リを誘発/引用 回数 ・<発見>の浸透と展開 ・学術的—遺伝子、生合成に関わる酵素、受容機構 ・社会的—薬、病気の関係
医学知識 としての 位置づけ	レレバントな 研究領域へ新し いくもの>を提 示	レレバントな研 究領域において もの発見>と命 名が示された ・<もの>を特定 した (DNA構造 を決定した) ・<もの>が新し い仮説を示唆した	過去の研究を引き 継ぎ、著者が<もの の発見>を行なうア セスを示した ・<もの>を特定し た (DNA構造を決 定した) ・<もの>の発見> 他の研究と関連付け て仮説を提示、関連 領域の研究の方向性 を指示した	領域全体の知識として<発見>を参照 ・個々の研究に知識としての一貫性を与えて関 連付ける (構造化された知識の穴埋め作業) ・個々の研究に共同な方向付けを与える : 問題や仮説、今後要請される研究の共有化 ・知識の拡張: 薬理学全体を対象読者へ ・<発見>の一般化 ・個別の研究成果や研究活動をまとめて記述 /解説の文体/「多い」「かなり」の使用 ・知識の配分 : 「ここでは詳述を避けるので、他の総説を 参照されたい ⁹⁰⁾ 。」	<発見>を一つの語りとして歴史的に位置づける ・科学文献での記述とは異なる: 現実の<発見>の 提示: 「例えばこういうpaperを頭の中に描いてこ ういう研究なんかできたら、これ程もう素晴らしい ことはないんだけれど、実際はそうじゃないで..」 ・科学論文に知識としての価値を与え: 「正にこ の論文ですよ。全体をもうずーっと絶えず、把握し ながら見ると (総説 1992に対して) 」 ・現象としての<発見>とテキストとしての論文の 切り離し ・<発見>の一般化 ・専門知識の「 <i>in vitro</i> 」への解説 ・知識の配分: 「うちがけがけにやっているのは...」	

*「周到な」は、この領域では単に物質を単離・精製してその活性を見ただけの研究発表と比べ、これが特
定のもののDNA構造まで決定しており、そのことがこの領域で「あるもの」を完全に確定する」とい
う水準に達したとみなされうる、と推察されるからである。

な意義へと広がっている。

そして、最も豊かな展開があるのが、医学知識との関連からみたテキストの構成のされ方である。〈エンドセリンの発見〉がどれほど重要なことであれ、原著 1988では、このテキスト自体を過去のテキストと結びつける指標を強調してはいても、〈発見〉は医学知識に吸収されていない。ところが、後続するテキスト、総説 1992・インタビュー 1994では、薬理学や医学という知識の権威の下に、〈エンドセリンの発見〉が秩序立てて整理されて、提示されている。知識として提示される〈エンドセリンの発見〉は、特定の研究グループが主張することではなく、科学コミュニティの（「我々の」）達成である。そればかりか、テキストは、知識の配分を決め、これを統制するものとなっている。そして、知識としての〈発見〉を他へ拡張し、一般化している。

3 科学知識の形成過程

科学文献を含む科学のテキストは、対象とする

もの（what that scientific text "is about"）、即ち研究活動や科学の発見、と1対1に対応するのではなく⁷⁾、むしろそのテキストから解釈される現象とテキストの組織のされ方は同じである⁶⁾。それゆえ医学知識の中で〈発見〉をいかに位置づけるものとしてテキストが構成されているか、を探ることこそ、我々の出発点となる。

この研究で対象とした各テキストでは、〈エンドセリンの発見〉の記述が変化していた。これは、テキスト毎に科学的発見を説明付ける（account）際のレレバンスが異なっていること、それにも関わらず、〈エンドセリンの発見〉が一つの同定可能な活動として維持されていることを意味している。〈エンドセリンの発見〉は、後続するテキストで「言い換え」³⁾られ、知識としての位置づけが展開していた。この言い換えられていくテキストの連鎖の過程こそが、「科学知識の形成過程」を表わすものである。

参考文献

- 1) Subramanyam K. Scientific literature. Encyclopedia of Library and Information Science. Vol. 26. p. 394.
- 2) 例えば、Price, Derek J. de Solla. Little Science, Big Science... and Beyond. New York, Columbia University Press, 1986. を参照。
- 3) 武者小路澄子, 野添篤毅. 複数の情報メディアにおける同一の科学情報の伝達: その同一性と差異. 1994年情報学シンポジウム講演論文集. p. 51-60.
- 4) 眞崎知生. 「序」. エンドセリン: 最新情報. 東京, 中外医学社, 1990. p. [4]-[5].
- 5) Mushakoji, Sumiko. Constructing 'identity' and 'differences' in scientific texts and their summaries. Dagstuhl-Seminar 9350: Summarizing Text for Intelligent Communication, p. 55-58.
- 6) Woolgar, Steve. Discovery: logic and sequence in a scientific text (1). The Social Process of Scientific Investigation. Dordrecht, Holland, D. Reidel, 1980. (Sociology of the Sciences. Vol. IV)
- 7) 武者小路澄子, 野添篤毅. 「専門的知識」として提示される科学情報と科学のテキスト: その非対称的関係における諸問題と解決の枠組み. 情報知識学会第1回(1993年度)研究報告会講演論文集. 1993. p. 49-54.

付表 分析対象とした原著論文及び総説

1988 原著 Masashi Yanagisawa, Hiroki Kurihara, Sadao Kimura, Yoko Tomobe, Mieko Kobayashi, Youji Mitsui, Yoshio Yazaki, Katsutoshi Goto, & Tomoh Masaki. A novel potent vasoconstrictor peptide produced by vascular endothelial cells. Nature, Vol. 332, p. 411-415.

1992 総説 後藤勝年, 桜井武, 粕谷善俊. 心血管系におけるエンドセリンの分子薬理. 日薬理誌, Vol. 100, p. 205-218. (総説 循環薬理学の新動向)