

構成要素カテゴリを用いた情報メディアの構造の分析：
言語表現に関する考察に基づく分析基準の再検討

Structure Analysis of Information Media using the Categories:
Re-examination of the Criterion Based on
Linguistic Consideration

神 門 典 子
Noriko Kando

Résumé

This paper conducts functional structure analysis of research articles, using the Categories, the author previously proposed as a framework for analysing informational content of research articles. First, reexamination on the Criterion of the Categories is conducted through consideration on linguistic feature of the sentences. A Number of lexical-clue-combination-patterns are extracted from the corpus of research articles selected from three specialties. These patterns are classified by combination of sentence-end-representations, verbs, and other lexical clues. After that, these patterns are listed and added to previous Criterion. Junction between sentences are also considered.

- I. 情報メディアの伝達内容の構造
 - A. 情報メディアの構造
 - B. 構成要素カテゴリとは
- II. 構成要素カテゴリの分析基準の問題点
- III. 言語表現にもとづくカテゴリ分析基準の再検討
 - A. 言語表現に着目した問題1の再定義
 - B. 分析基準の再検討の方針
 - C. 特徴的な言語表現パターンの抽出
 - D. 分析基準の例示
 - E. 問題2の言語表現に基づく検討
- IV. 構成要素カテゴリの再検討
 - A. 内容判断を伴うカテゴリ

神門典子：慶應義塾大学大学院博士課程，東京都港区三田 2-15-45。（日本学術振興会特別研究員）
Noriko Kando: Graduate School of Library and Information Science, Keio University, 2-15-45, Mita,
Minato-ku, Tokyo 108, Japan.

(*Research Fellow of the Japan Society for the Promotion of Science*)

1994年3月17日受付

B. 構成要素カテゴリの修正

V. まとめ

I. 情報メディアの伝達内容の構造

A. 情報メディアの構造

情報メディアは、文やデータの無秩序な寄せ集めではなく、その内容は送り手が受け手に伝えるという意図をもって組み立てられており、構造がある。ここでは、情報メディアは、人間の知的あるいは芸術的創造の成果をひとまとまりのものとして記録し、伝達するものとする。従来から「文献」または「資料」と呼ばれているものとはほぼ同義であり、紙に印刷されたものだけでなく、いかなる物理的形態も含むものとする。

図書館・情報学では、従来から、情報メディアの特性を調べるため、あるいは、抄録作成や情報検索へ応用するために情報メディアの構造の分析がなされてきたが、分析方法は分析の目的に応じて様々であった¹⁾。また、ここで扱う構造は、情報メディアの伝達内容の構造である。これは、レイアウトなどの情報メディアの物理構造や章・節・段落といった文書書式を物理的な固定から解放して扱う情報メディアの論理構造ではない。伝達内容の構造と章・節・段落といった文書書式の構造とは関連する部分もあるが、完全に一致するわけではない。

情報メディアの伝達内容の構造は送り手が内容を組み立てた結果として存在するものである。また、情報メディアは、送り手が属するコミュニティの慣習や社会的な規範の影響の下で個別に作成され、その結果として存在する情報メディアは、社会の中で、新聞記事、学術書などというように特定種類の情報メディアと認められる。

情報メディアを介した送り手から受け手への情報伝達には、送り手や受け手の内部の状態や、送り手が作成する過程や受け手が読みとる過程が関わっているが、ここではこのような主観的な作成過程を経た結果として、また、主観的な読みの対象として、一つの実体として外部に固定されて存在している情報メディアを分析対象とする。このような外部に存在している情報メディアの伝達内容を分析する手がかりとなるのは、メディアが持っている表現である。したがって、情報メディアの中で、ことばによって表現され、記録されている部分については、言語表現から伝達内容をとらえることになる。

一方、情報メディアの構造分析の可能性と意義を検討

するには、具体例を示し、実際に構造を分析し、分析の問題点や一貫性を持った処理の可能性、応用可能性とその意義について検討する必要がある。このような意図のもとに、情報メディアの伝達内容の特性を、同一種類のメディア群内で共通にとらえる基盤の具体例の一つとして、構成要素カテゴリを提案した。

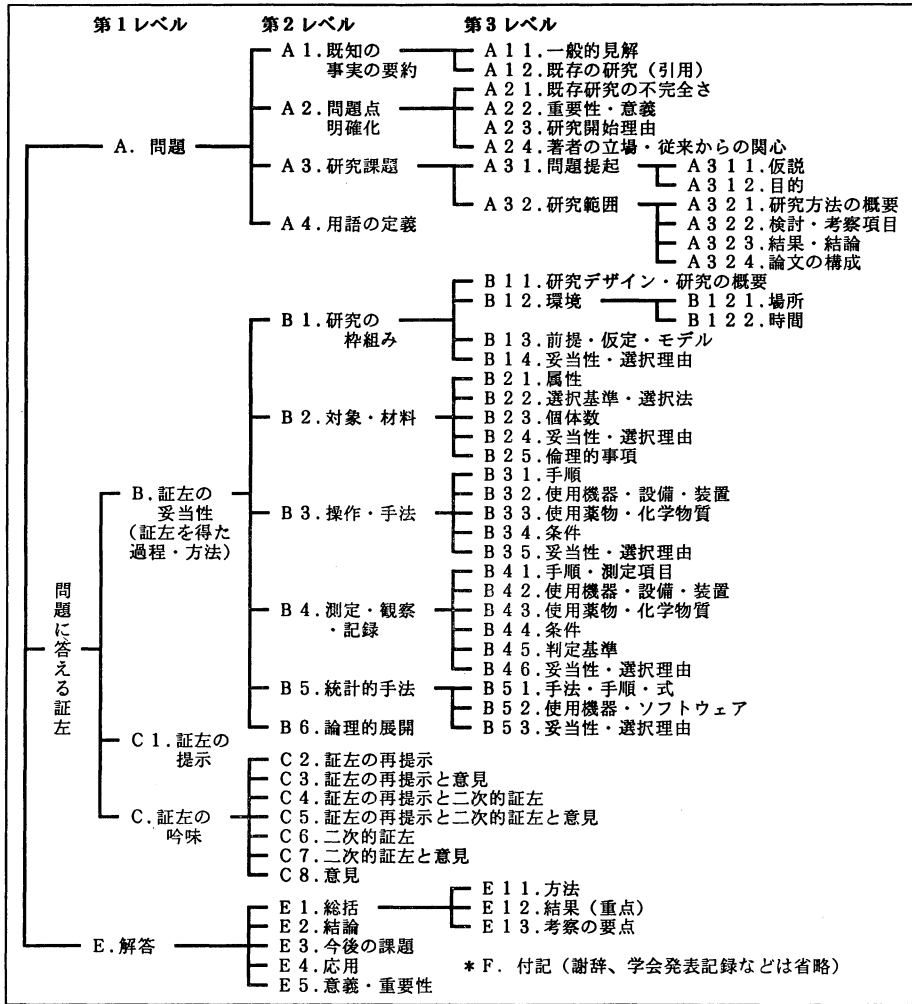
B. 構成要素カテゴリとは

構成要素カテゴリとは、情報メディアの伝達内容の特性を、領域に関わらず、メディアの種類ごとにも共通にとらえる基盤として筆者が提案したものである²⁾。各カテゴリは、情報メディア内の部分が果たしている役割や機能を表わす。また、一つの情報メディア内でのカテゴリの有無や出現順序によって、情報メディアの内部構造を記述することができる。これは情報メディアの伝達内容を機能という側面からとらえた構造である。

第1図に示した一連の構成要素カテゴリは、情報メディアの一種である原著論文を分析するために各種の論文執筆マニュアルと日本語で書かれた原著論文の内容の分析を通じて、カテゴリの追加やカテゴリの定義および分析基準の修正を繰り返し、全体の調整をとって設定した²⁾。一連の構成要素カテゴリは完全に固定されたものではなく、今後、分析対象の範囲が拡大するにつれて、カテゴリの追加、分析基準やカテゴリ定義の修正を繰り返しながら適用範囲を拡大していく性質のものである。

一連の構成要素カテゴリには階層がある。カテゴリの付与は、もっとも詳細なレベル（第1図ではもっとも右側）を用い、分析基準に従って手作業で付与する。付与の単位は原則として文であり、全ての文にカテゴリを付与する。必要に応じて1文内の部分にも付与でき、図表の参照や注にも付与する。全体から見た役割と局所的な役割とを二重に付与する「入れ子構造」も認めている。一方、カテゴリの出現型は、第2図の右側に示したように、適宜より上位レベルにまとめてとらえる。

構成要素カテゴリの基本的な考え方は、情報メディアの伝達内容を構成している部分部分に着目し、各部分と当該メディア内部の関係、および、同じ種類の情報メディア群内でメディア間に共通な関係を保持しながら、部分を単位として扱うことである。この特性をいかして、情報メディアの蓄積、検索、加工へのさまざまな応用が



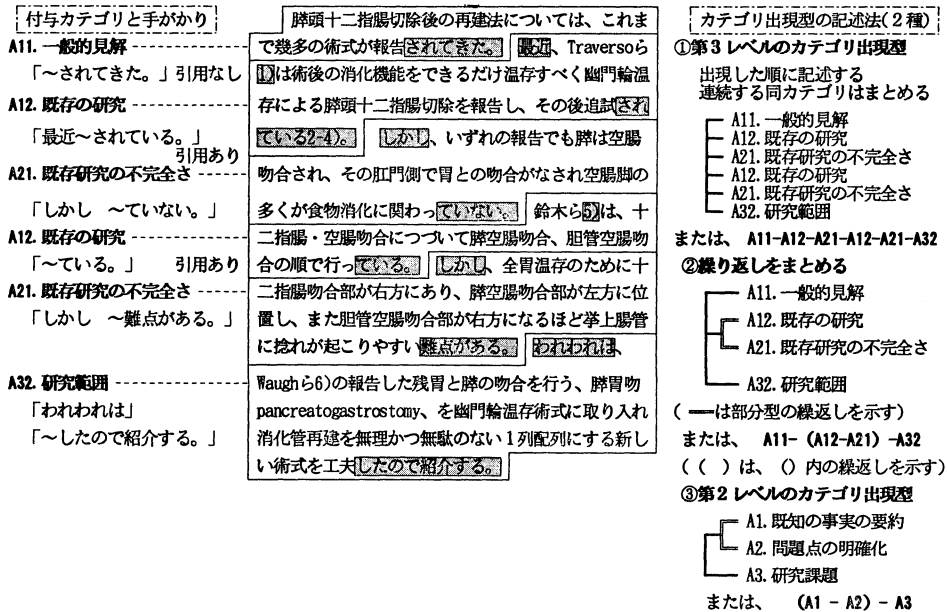
第1図 原著論文を分析するために設定した構成要素カテゴリ

考えられる²⁾。特に、全文データベースでは、全文にカテゴリが付与されたものと仮定した場合、様々な柔軟な利用法が考えられる。また、今後情報メディアに新たな物理的な形態が出現した際にも、情報メディアの内容に適した物理的な形態を考える上で有用であろう。

具体的な応用の1例として、C型肝炎論からなる実験的な全文データベースから、カテゴリを用いて、領域の現状を記述した部分を検索した。検索された文は、データベース中の複数の論文から抽出されているのでレビュー論文のようにひとまとまりの文章ではなく、専門家にとっては既知の事実であるが、容易に読むことができる

程度の文数であり、初学者や一般人がその領域の現状を把握するには有効だと思われた³⁾。また、複数の情報メディアの内容を横断的に抽出することにより、個別の文献ではわからない新たな関係が見いだされることもある。

その一方で、構成要素カテゴリの付与は時間がかかる作業なので、実用化するには付与の自動化が必要である。日本語のC型肝炎論文を対象として、「A. 問題」の下位カテゴリに関して予備的に試みたところ、95.2%の精度で自動付与に成功し、その実現可能性が示された³⁾。複数領域の論文を対象とした自動付与も試みている⁴⁾。



第2図 カテゴリの付与とカテゴリ出現型の記述法

さらに、川原村⁵⁾は、後述するように、構成要素カテゴリを用いた分析の妥当性を調べるために、妥当性検証実験を行なっている。その結果、従来の分析基準にはわかりにくい点があることを指摘している。

そこで、本報では、構成要素カテゴリを用いた分析の一貫性を高めるために、従来の分析基準の問題点を明らかにし、言語表現という面から分析基準を再検討して、精緻化した。

II. 構成要素カテゴリの分析基準の問題点

川原村⁵⁾は、対人認知、会社法、高分子化学の3領域を対象として、手作業で構成要素カテゴリをもちいた分析を行ない、分析の妥当性検証実験として、上記の3領域からそれぞれ4論文を選択し、2名の分析者が同一の論文に構成要素カテゴリを付与した結果を比較している。

その結果、2名の分析者が付与したカテゴリが一致しなかった文では、両者が付与したカテゴリの組合せに一定の傾向が見られた。特に、一方の分析者が「A11. 一般的見解」を付与した文に、もう一方の分析者が「A24. 著者の立場」を付与した例が多く見られた。

また、対人認知領域のサンプルに関しては、2名の分

析者がカテゴリ付与した後、筆者も加わり、3者で話し合って最終的なカテゴリを決定した。従来のカテゴリ付与のための分析基準は、各カテゴリの基本的な役割の説明と、少数の手がかり語の例示からなっており、例示された語が文中に出現しない場合、付与カテゴリは分析者の判断に依存する。しかし、話し合いを通じて、いくつかのカテゴリに関して、カテゴリの基本的な役割の範囲について分析者によって考え方が異なっていたことがわかり、その結果として付与カテゴリに相違が生じたと考えられた。このことから、分析の一貫性を高めるには、カテゴリの基本的役割に基づくカテゴリ付与より、手がかり語の例示を強化する方が有効であると考えられる。

さらに、作業時の問題点として、分析基準では手がかり語の例示が少なく、分析対象の文中に手がかり語が出現しない場合にはどのカテゴリを付与するか判断がつかない場合が多いことが指摘され、特に、以下の2点がわかりにくいと指摘されている。

問題1: 「A11. 一般的見解」と「A24. 著者の立場」の区別

問題2: 「A12. 既存の研究」で引用している文献の内容紹介が続いている範囲の認定

そこで、この二つの問題を解決するために、言語表現

という面から検討を加え、それに基づいて実際に複数領域の論文中の文から分析の手がかりとなる具体的な言語表現を抽出して整理し、それを従来の分析基準に追加することによって分析基準を精緻化した。

III. 言語表現にもとづくカテゴリ分析基準の再検討

A. 言語表現に着目した問題1の再定義

言語表現という側面から問題1, すなわち, 「A11. 一般的見解」と「A24. 著者の立場」の区別について検討した。

まず, 「A11. 一般的見解」「A12. 既存の研究」という「A1. 既知の事実の要約」の下位カテゴリは、いずれも著者が現状を「事実」として記述したものである。「A21. 既存研究の不完全さ」「A22. 重要性・意義」「A24. 著者の立場」という「A2. 問題点の明確化」の下位カテゴリは、著者の主観的な考えを表わしたもので、すなわち「意見」である。したがって, 「A1. 既知の事実の要約」と「A2. 問題点の明確化」との区別, あるいは, 「A11. 一般的見解」と「A24. 著者の立場」の区別は, 論文における事実の記述と意見との区別であるといいかえられる。

一方, 構成要素カテゴリの基盤には「論文は批判的な議論である」という考えがある。批判的な議論では, 伝達されている内容だけでなく, その妥当性の吟味が重要である。妥当性を吟味するには, 何を拠り所として主張を組立て, その拠り所が十分であるかを判断することが重要であり, そのため, 事実と意見との区別が重視される。したがって, 事実の記述と意見との区別は, 「A11. 一般的見解」と「A24. 著者の立場」の区別だけでなく, 構成要素カテゴリ全体に関わる問題である。

また, 木下⁶⁾によれば, 論文における事実の記述とは「事実として書いた」ことで, 真偽を客観的に判定できるものであり, 引用の記述もこれに含まれる。論文における意見には, 著者による推論, 判断, 意見, 仮説が相当する⁶⁾。

一方, 木戸⁷⁾によれば, 「文章表現においては事実も意見も書き手の主観的判断によるものである」。すなわち, 事実の記述においても, どの事実を取上げ, どのような表現で, どの話題の提示や意見の根拠として用いるかに書き手の主観的判断が関わっていると述べている。そして, 論説文中の文の機能として, 書き手の主観的判断の関わる度合いが弱い順に第1表のような6種の機能を設定している。

第1表 文の機能

木戸⁷⁾より

小	事実を述べる機能
↑	報告: 事実をなるべく主観を交えずに提示する機能
主	解説: 事実を主観を交えて提示する機能 (文章中では事実の詳述あるいは一般化)
観	根拠: 判断の拠り所となった事実を提示する機能
的	
判	
度	意見を述べる機能
断	理由: 判断の拠り所となった意見を提示する機能
合	評価: ある事柄についてある判断を下す機能 (文章中では, ある事実に対する見解や感想)
い	
↓	
大	主張: 読み手に対して意見を提示する機能

ここで問題となっている「A11. 一般的見解」カテゴリは, 著者が現状の報告, すなわち, 事実として記述しているものと考えられる。しかし, 著者の見方からとらえた知識の現状であり, 論文の課題に関係すると著者が判断し選んだものであるので, 著者の主観的な判断が関わっている。また, 必ずしも真偽を客観的に判断できるとは限らない。したがって, 純粋に客観的な事実の記述ではないので, 事実の記述と意見との境界をどこにするかで「A24. 著者の立場」との区別が不明確であったと考えられる。ただし, 両者とも著者の主観的判断が関わると言っても, その程度は異なる。そこで, 木戸の枠組みを参考にして, 言語表現の面から著者の主観的判断が関わる程度と文の機能を整理することとした。

B. 分析基準の再検討の方針

1. 文末表現と述語の活用語幹とその他の要素の組合せ
言語表現の面から著者の主観的判断の関わりをとらえるために, 日本語のモダリティに関する研究⁸⁾⁻¹³⁾を概観した。モダリティとは「話者の内側にある, 主観的態度の叙述」¹²⁾である。これは, 文は客体化された世界の叙述である「命題」と主観的な態度の叙述である「モダリティ」という二つの成分から成り立つとするとらえ方によるものである。日本語学では, 研究者によって「詞」と「辞」, 「コト」と「ムード」, 「言表事態」と「言表態度」など用語は様々であるが, 同様に客観的な叙述内容と主観的な態度を表明する部分という二つの成分を認めている。

日本語では, 主観的な態度の表明に関して, 文末の助動詞と終助詞を中心とした「文末表現」が大きな働きをしている。そこで, モダリティに関する諸研究を参照し

第2表 手がかり語（群）の例

文末表現	A群	肯定判断(評価)	である。と言える。といえる。はずだ。
		否定判断	(で)ない。なかった。ず、なく、
		推量(概言)	だろう。であろう。かもしれない。
		当為	べきだ。なければならない。ねばならない。
		程度判断	多い。高い。／少ない。低い。薄い。
		可能	れる。れ うる。できる。-える。
現	B群	説明	ことになる。ところだ。わけだ。ものだ。のだ。
		過去/現在	た。／る。
現	B群	幅のある時間	ている。ていた。てきた。
		継続	つつある。
		状態	広い。狭い。固い。
述語活用	存在	(が) ある ない	
	研究	研究(する) 検討(する) 実験(する) 用い 測定(する)	
	成果	見い出(す) 経験(する) 得 確認(する) 発見(する) 明らか	
	報告	報告(する) 述べ(る) 概説(する) 説明(する) 指摘(する)	
	思考	思(う) 考え(る) 考慮(する) 配慮(する) 予測(する) 推論(する)	
	呼称	呼ぶ (と)いう 名付け(る) 意味(する)	
	問題	問題(だ) 困難(だ) 失敗 難しい	
	必要	必要(だ) 必須(だ)	
	他	重要 重要性 有用性 基本 有効性 将来性	
	最近	最近 近年 現在 ちかごろ	
	過去	従来 以前にも すでに これまで 当初	
今回	今回 そこで ここでは		
本～	本稿 本研究 本論文 本報告		
自称	我々 著者(ら) 筆者(ら) 私ども		

て、「文末表現」を整理し（第2表参照）、これを中心に著者の主観的判断が関わる程度を判定することとした。

一方、文の中心的要素は「述語」、特に文末にある主述語であり、これが文の中心的な意味を担っている¹⁰⁾。述語となっている動詞のタイプによって文の機能を類型化した研究もある¹⁴⁾。そこで文末に出現する述語にも着目し、その活用語幹を整理した（第2表参照）。

第2表において、文末表現のA群と〈必要〉グループや〈思考〉グループの述語の活用語幹は主として意見、B群は主として事実の記述に用いられるが、実際には、これらの文末表現や述語の活用語幹と、その他の語との組み合わせによって文が担う機能が異なり、対応するカテゴリも異なってくる。

また、たとえば、「目的」という語は、一見、一語だけで「A312. 目的」の明確な手がかりになると思われるが、実際には「田中^中は…を目的として研究してきた。」「彼の実験は…が目的である。」（いずれも「A12. 既存の研究」）、「…という目的以外の用途にも用いられる。」など、当該論文の目的を述べている以外の文脈でも出現する。したがって、単一の語の有無だけでなく、多くの条件を加味した基準を設定して、各文の当該論文における機能を認定する必要がある。

さらに、副詞の中にも、「おそらく…ない。」というように特定の文末表現と呼応してその意味を強める働きを

持つものもあり、分析の手がかりになる。

そこで、手がかり語は、述語の活用語幹と文末表現を中心に、前文のカテゴリ、接続表現、提題表現、副詞、その他の手がかり語、前方の文との関係などの組み合わせをパターンとして整理することとした。

2. 主観的な判断が関わる程度の判定

手がかり語の組合せパターンを整理する際、言語表現の面から、どのカテゴリを付与するかという判断に迷う場合には、第1表に示した木戸のスペクトラムを参考にした。

たとえば、「～である。」は、書き手の判断を表わすので、「A24. 著者の立場」カテゴリとなる。しかし、前文が「A11. 一般的見解」や「A12. 既存の研究」など事実を記述した文で、「これは～である。」や「この+前文中の語+は～である。」など前文を詳述して解説している文は、「意見」といえるほど書き手の主観的な判断が強くないが、「事実の記述」といえるほど客観的でもなく、判断に迷う。この場合、第1表のスペクトラムから、事実を詳述し、事実を主観を交えて提示する「解説」にあたり、事実を述べる機能の1種とみなすことができる。その結果、カテゴリは前文と同じ「A11. 一般的見解」または「A12. 既存の研究」となる。

3. あいまいな表現の処理

さらに、実際の論文では、「・興味深い結果を得た。」「・画期的な手法を用いている。」などのように、文末表現は事実の記述であっても、「興味深い」「画期的な」などの主観的な語が含まれている文や、「・と考えられる。」などのように、著者の主観的な意見か事実の記述かの区別があいまいな記述が見られる。

木下⁶⁾は、この「と考えられる。」の「れる。」「られる。」は、「当然の成りゆきとしてそういう考えになる」、「自然にそう思えてくる」という意味の自発の助動詞、または「と考えることができる」という意味の可能の助動詞であり、最終的な判断を相手にゆだねて自分の意見をぼかし、「…と考えられる」が自分の考えはちがうと逃げる余地を残した表現であるので、レポートではこういうあいまいな責任回避的な表現は避けるべきであると述べている。

しかし、構成要素カテゴリを用いた原著論文の構造の分析においては、実際にある論文を分析するのであり、論文を「どう書くべきか」を扱うのではない。したがって、このようなあいまいな表現についても処理方法を決定し、一貫した方法で処理できるようにする必要がある。

4. 複数領域の実際の論文に依拠する

そのうえ、領域によって、多く出現するカテゴリに差があり²⁾、論文の構造の複雑さも異なることが報告されている²⁾、¹⁵⁾、¹⁶⁾。同一のカテゴリに対する手がかりとなる言語表現も領域によって差異がある。たとえば、「A12. 既存の研究」カテゴリについて、医学全般から選択した論文では「～ている。」という文末表現がほとんどであったのに対し、物理学論文では「～た。」もみられた²⁾という差異が見られた。

そこで、複数領域から選択した実際の原著論文にあたって、その中から、カテゴリ付与に関わる特徴的な手がかり語の組合せを抽出して整理し、それを分析基準に追加することとした。

C. 特徴的な言語表現パターンの抽出

対象領域として、C型肝炎、検索研究、対人認知を選び、その中から日本語で書かれた原著論文をそれぞれ、50件、49件、38件選択して分析対象とした。手作業でカテゴリを付与し、特徴的な言語表現の組合せパターンを抽出した。あいまいな表現を持つ文については、一貫し

た処理方針を定めて、それにしたがった。

C型肝炎論文⁴⁾と対人認知⁵⁾のサンプルはすでにカテゴリが付与されていたが、それらも見直し、あいまいな表現を持つ文については、一貫した方針に従って一部カテゴリを修正した。

本稿で問題となっている「A11. 一般的見解」と「A24. 著者の立場」を含む、「A. 問題」カテゴリ部分を中心に検討した。分析対象論文における「A. 問題」カテゴリ部分は、領域ごとの合計で、C型肝炎では293文、情報検索は914文、対人認知では1482文であった。

必要に応じて、すでにカテゴリを用いて分析した医学、物理学、経済学、国文学領域から抽出した原著論文¹⁾も参照した。

D. 分析基準の例示

分析の手がかりとなる特徴的な言語表現は、文末表現と述語の活用語幹の組合わせを中心に、提題表現、接続表現、副詞、その他の手がかり語、前文のカテゴリを考慮し、それらの組合わせで整理した。類似の機能を持つ手がかり語をグループとしてまとめて整理し、手がかり語グループの組合せパターンとした。

しかし、各カテゴリごとに手がかりとなる言語表現の組合せパターンを多数列挙することは、分析基準として実用的ではない。そこで、各カテゴリごとに、文末表現と述語の語幹を中心に、基本的に当該カテゴリとなるパターンを示し、それに、他の手がかり語との組合せによって異なるカテゴリとなるパターンを例外としてまとめて付した。第3表に例を示す。

なお、第3表および以下に示す〈 〉は、第2表の手がかり語のグループを示す。

たとえば、文末の「である。」を取りあげると、第3表に示したように、提題表現があり、文末が「である。」の文は、基本的に、著者の主観的な判断を表わす「A24. 著者の立場 (考え)」である。

ただし、「問題である (問題だ+である。)」なら「A21. 既存研究の不完全さ」となる。また、提題表現が当該論文の主題や主要な方法・成果に関わる語で、文末が「重要である。(重要だ+である。)」ならば、当該論文の課題あるいはアプローチの重要性や特徴を示す「A22. 重要性・意義」になる。また、前述のように、前文が「A11. 一般的見解」または「A12. 既知の事実」であり、提題表現が「これは」あるいは「この+前文中の語+は」など前文への指示語であれば、前文の詳述・解説となり、

第3表 手がかり語(群)の組合わせによる分析基準(例)

前文の カテゴリ	接続	提題(ハ)	その他	述語	文末	引用	付与カテゴリ
A12既存研究	○は				である。		→ A24J著者立場
	○は			<問題>	である。		→ A21. 不完全さ
		{主題語}は*1		<重要>	である。		→ A22. 重要性
		{前文語}は*2			である。		→ A12. 既存研究
		<自称>は*3		<研究>	ている。	あり	→ A12. 既存研究
		人名 ^ラ は		<研究>	てきた。	あり	→ A24J著者立場
				<研究>	てきた。	あり	→ A12. 既存研究
				<研究>	る。		→ A312目的
				<研究>	た。	あり	→ A12. 既存研究
		<自称>は*3	<過去>	<研究>	た。	あり	→ A24J著者立場
		そこで	<自称>は	<研究>	た。		→ A321研究概要
			<本～>	では	<研究>	た。	→ A321研究概要

<>は表2に示した手がかり語のグループを表わす。^ラは文献の引用を表わす。
○は提題表現があることを表わす。
*1: 論文全体の主題、主要な方法やアプローチを表わす語
*2: 前方の文中に含まれる語
*3: <自称>グループの中でも特に「筆者(ら)は」が多い。

前文と同一カテゴリになる。

同様に「研究」グループの述語の活用語幹が出現する文は、文末表現が「ている。」ならば、基本的に「A1. 既知の事実の要約」の下位カテゴリとなり、文献引用があれば「A12. 既存の研究」である。

ただし、文末表現が「てきた。」では、「A24. 著者の立場(いままでの著者自身の研究)」または「A12. 既存の研究」になる。前者の場合、提題表現は「筆者らは」「著者らは」「われわれは」などの「自称」グループであり、その中でも特に「筆者(ら)は」のことが多い。後者の「A12. 既存の研究」の場合、提題表現は引用文献の著者名か研究の内容になる。

さらに、文末表現が「た。」の場合、文献引用があり、提題表現が「自称」でなければ「A12. 既存の研究」となる。提題表現が「自称」で、文献引用がある、もしくは、「すでに」「～年より」「これまで」「既報にて」など「過去」グループの語があれば「A24. 著者の立場(いままでの著者自身の研究)」となる。「本稿では」「本報では」「本論文では」などの「本～」グループの語が出現したら「A321. 研究方法の概要」の可能性が高い。

21群 633語の手がかり語を認定し、それらの組み合わせを文末表現と述語の活用語幹を中心に整理し、分析基準に追加した。なお、この中には、手がかり語のグループに対する処理と、グループ中の個々の語に対する処理とが含まれる。

E. 問題2の言語表現に基づく検討

川原村⁹⁾が指摘したもう一つの問題は、「A12. 既存の

a) Darley & Fazio(1980)は、期待者と被期待者間の社会的相互作用の1つのモデルとして、6段階の期待確認過程を提示している。それらは、(a)期待者の被期待者に対する期待の形成、(b)期待に一致する期待者の行動、(c)期待者の行動についての被期待者の解釈、(d)被期待者の反応、(e)被期待者の反応についての期待者の解釈、(f)被期待者の自分自身の反応についての解釈である。 A12. 既存研究

b) Dalto, Aizen & Kaplan(1979)は、開示内容の内面性と望ましさを2要因を独立に操作する実験を行なった。その結果、自己開示の望ましさは、対人魅力に正の効果を及ぼすのに対して、その内面性それ自体は何らの効果も及ぼさないことを見出した。開示の望ましさが魅力に効果を及ぼすのは、その要因が被開示者によって認知された開示者の傾性の望ましさに直接影響を及ぼすためであると説明している。 A12. 既存研究

第3A図 数文でひとまとまりの内容を持つ例

A12. 既存の研究

B型肝炎に対するインターフェロン両方の有効性⁽⁴⁾が確認され、C型肝炎に対してもIFN療法の評価⁽⁵⁾がなされつつあるが、注射剤であることより、入院あるいは頻回の通院が必要となり、患者はかなりの負担を強いられるという難点がある。 A21. 不完全さ

第3B図 2つのカテゴリを付与した文の例

研究」カテゴリが続く範囲の認定が困難であるということである。この「A12. 既存の研究」カテゴリに関わらず、実際の論文では、第3A図のように数分からなる内

第4表 文の接続に伴うカテゴリ遷移の手がかり (例)

関係		接続表現の例
順接	前文に基づく判断	したがって このことから
接	前文に基づく実施	そこで
	前文の帰結・結果	その結果 このため

同格	換言	すなわち いいかえると
	例示	たとえば 具体的には

連鎖	前文の解説 前置き 話題場面設定など	<ul style="list-style-type: none"> 提題表現の省略 前方指示・前文語の提題表現 箇条書き(列挙表現) 述語の連鎖、

面的に密接に関連したひとまとまりの部分がある。

このようなひとまとまりの部分は、通常、読み手が読みとった内容に基づいてひとまとまりの内容が続いていると認定される。そのため、分析者によって判断が一致しない場合もあるし、専門的な知識がないと内容のまとまりを判断できない場合もある。また、言語表現という面からも、このようなひとまとまりの部分の中では、繰り返し出現する語句が省略されることも多く、前述の手がかり語の組合せパターンだけでは、カテゴリを判断できない場合が多い。そこで、言語表現によってひとまとまりの範囲をとらえるために、文の接続関係に着目した。

日本語の文章論において、寺村ら¹⁷⁾は、接続関係として順接、逆接、添加、対比、同列、転換、補足、連鎖の8類型を示している。ここでは、それらの中から実際の論文に出現し、文の接続関係に伴うカテゴリの遷移に関連するものを選択した(第4表参照)。以下、接続詞などの明示的な接続表現が出現する場合と、そうでないものに分けて述べる。

1. 明示的な接続表現が出現する場合

接続詞などの明示的な接続表現は、前文のカテゴリとの関係において、カテゴリ付与の手がかりとなる場合がある。

たとえば、第4表に示した、順接の中の「前文に基づく判断」で接続している文は、前文の事実の記述(「A11. 一般的見解」「A12. 既存の研究」)を根拠とする判断(「A24. 著者の立場」)、あるいは前文の意見(「A21. 既存研究の不完全さ」「A22. 重要性・意義」「A24. 著者の

立場」)を理由とする判断(「A24. 著者の立場」「A311. 仮説」)を示す。

同じく、「前文に基づく実施」で接続している文は、前文の判断(「A24. 著者の立場」)を理由とする研究の実施、あるいは前文の事実の記述(「A11. 一般的見解」「A12. 既存の研究」)を根拠とする研究実施として「A321. 研究方法の概要」となる。「A12. 既存の研究」の文が連続している中では、前文の述語は〈思考〉グループで、当該文はそれに基づく〈研究〉グループの述語となる。

また、順接の中の「前文の帰結・結果」で接続している文も、同様に「A12. 既存の研究」の文が連続している中では、前文の述語は〈研究〉グループ、当該文は〈成果〉グループの述語となる。

さらに、同格で接続している文は、「換言」「例示」ともに、原則として前文と同一カテゴリとなる。ただし、「例示」の場合は、前文における著者の判断(「A21. 既存研究の不完全さ」「A22. 重要性・意義」「A24. 著者の立場」)の根拠となる事実の記述(「A11. 一般的見解」「A12. 既存の研究」)が示される場合もある。

2. 明示的な接続表現がない場合

連鎖は、学術的な論文では最も頻繁に見られる接続関係であるが、とくに接続詞は見られない。ここでは、提題表現、列挙構造、述語の活用語幹の連鎖に着目した。

前文とこの連鎖関係にある文では、提題表現が(1)省略される。または(2)「これは/これらは/それは/前者は/その一つは/彼は」などの指示表現となる。(3)「前(方)の文中の語+は」あるいは「この/その/このような+前(方)の文中の語+は」という指示表現となる場合が多い。しかし、実際には、直前の文だけでなく、離れた位置にある文を指示したり、前方の段落や節の内容全体を指すこともある。また、(3)の前方文中の語も、完全に同一の語が繰り返されるだけでなく、同義語、下位概念や上位概念を表わす語、複合語の一部を省略した語などで言い換える場合が多く、必ずしも前文と同一の語が繰り返されるわけではない。

また、箇条書きなどの列挙表現とその前置きの文の特徴を調べ、列挙表現のパターンを整理した。これは、例えば、「以下にその利点を述べる。①～②～」というようなものである。列挙表現には、(1)節や段落レベルのもの、(2)文レベルで項目ごとに改行されているもの、(3)1文の中に複数の項目が列記してあるものという様々なレベルがある。しかし、実際には、文レベルの

列挙構造でも、一部の項目だけには説明がついている(1)と(2)のレベルが混在しているものなど、さまざまな形式が見られた。

さらに、一部の述語グループの種類と順序を連鎖パターンとして整理できる場合があった。例を以下に示す。なお、[] は省略可能、/はいずれか一つを選択、〈 〉は第2表に示した述語の活用語幹のグループ、《 》は文末表現のグループを表わす。

「A12. 既存の研究」でみられる述語連鎖パターンの例

- ①[…は、]…〈研究〉た。[…〈研究〉た。][その結果、]…〈成果〉た。…〈報告〉ている/た。
 ②[[…は、]〈思考〉た。][そこで]…〈研究〉た。[[その結果、]…〈成果〉た。][…〈報告〉ている/た。]
 ③[…は、]…〈研究〉/〈成果〉ている/た。[そこでは/それによると/彼によると]…〈判断〉。[[また]…〈判断〉。]

「A321. 研究方法の概要」でみられる述語連鎖パターンの例

- ①…〈研究〉た。…〈成果〉た。[ので]…〈発表〉る。

これらの接続関係を考慮すると、第3A図のa)は、提題表現に「それらは」という前文中の先行語への指示表現を含む連鎖関係であることから、IIIのDで述べたように、前文の解説として、前文と同カテゴリとなる。b)では、手がかりとなる接続表現も少しあるが、述語の連鎖パターンからいずれもひとまとまりの内容であり、一文めで引用した文献の説明の続きであると判断できる。

IV. 構成要素カテゴリの再検討

A. 内容判断を伴うカテゴリ

手がかりとなる言語表現の組合せパターンだけでは充分にカテゴリを判断できないもう一つの例として、カテゴリ付与にあたって、内容判断が必要なカテゴリが挙げられる。

たとえば、当該論文の課題やアプローチの重要性や利点を示す「A22. 重要性・意義」カテゴリや、研究課題を示すときに結果や結論の概略を示す「A323. 結果・結論」カテゴリ、また今回の範囲外であるが、「C2. 証左の再提示」は、いずれも言語表現だけでなく、カテゴリの付与に当たっては、内容の判断が必要なカテゴリである。これらは、分析者による判断の差異が大きくなると考えられ、また、今後、自動付与を進めていく際にも問

題である。したがって、言語表現による認定を可能にするようにさらに検討するか、あるいは、他のカテゴリに吸収するか検討する必要があるだろう。

B. 構成要素カテゴリの修正

一連の構成要素カテゴリは、I章のBでも述べたように、それまでに分析対象としたものが矛盾なく記述できるように設定したものである。したがって、分析の範囲が拡大すれば、カテゴリの追加や修正、カテゴリの定義や分析基準の細分化や修正も生じる。

「A24. 著者の立場」カテゴリは、①著者の従来の研究を紹介することによって以前からの著者の立場を示すもの(自己引用)と、②著者の考えを意見として表明したもの(著者の考え)の2種類が混在していた。

いままで分析した中で医学全般から抽出した論文群²⁾やC型肝炎論文ではこのカテゴリのほとんどが①の自己引用であったが、今回、分析した対人認知論文では②の著者の考えが多かった。そこで、このカテゴリを①に相当する「A24J. 自己引用」と、②の「A24I. 著者の考え」に細分することとした。

また、「A23. 研究開始理由」は、だれかの勧めや機会が与えられたなど外的な理由を表わすが、いままで分析した論文ではほとんど出現しなかったので、今後、削除することを検討する必要があるだろう。

V. まとめ

本稿では、情報メディアの伝達内容の特性をとらえるためにその構造を分析することの可能性と意義を検討する研究の一環として、具体的な一例として提案した構成要素カテゴリの分析基準について検討した。ここでは、情報メディアは、人間が人間に伝達することを目的として社会環境の中で作成するものととらえている。また、構成要素カテゴリを用いた分析は情報メディアの伝達内容の機能的な側面からとらえた構造を分析するものであり、一つの情報メディア内の部分と部分あるいは部分と全体との関係、および、一定の情報メディア群内の複数メディア間に共通の関係を扱うものである。

構成要素カテゴリの分析基準を言語表現の面から精緻化した。このために、論文中文を、主観的判断が関わる程度に応じたスペクトラム上に位置づけ、主観的判断が関わる程度を言語表現の面から整理した。そして、従来の例示的な分析基準に、手がかり語の組合せパターンの列挙を追加した。

分析基準を言語表現の面から精緻化することは、手作業で分析する際に分析者間の一貫性を高めるとともに、今後自動付与を進めていく上でも重要なことである。

また、このような手がかり語の組合せパターンによってとらえられないものとして、教文でひとまとまりの内容を述べている部分と付与に際して内容判断が必要なカテゴリが挙げられた。前者については、接続関係に着目し、提題表現、列挙構造、述語の連鎖に着目して、言語表現を整理した。後者は今後さらに検討が必要である。

さらに、適用範囲の拡大、応用の実現とその有用性の検討、このような分析の基盤となる情報メディアの内容構成に関わる社会的要因や歴史的発展の経緯に関するより深い検討なども必要であると考えられる。

本稿をまとめるにあたり、ご指導いただきました慶應義塾大学文学部の上田修一教授に心からなる謝意を表します。

本研究は、平成5年度文部省科学研究費補助金（特別研究員奨励費）による成果の一部である。

- 1) 神門典子. 情報メディアの構造：伝達内容の分析と利用. *Library and Information Science*. No. 30, p. 1-19 (1992)
- 2) 神門典子. 構成要素カテゴリーを用いた情報メディア内部構造分析の試み：原著論文を例として. 東京, 慶應義塾大学, 227 p, 1991. 平成2年度修士論文. (分析結果の概要は, 神門典子. 構成要素カテゴリーを用いた原著論文の内部構造分析, 情報処理学会研究報告 (情報学基礎研究会 92-FI-25). Vol. 92, No. 32, p. 39-46 (1992) 参照)
- 3) 神門典子. 原著論文の機能構造の分析とその応用：C型肝炎論文を対象とした基本動向記述文の抽出とその前提としての構成要素カテゴリー自動付与の試み. 図書館学会年報. Vol. 40, No. 1, p. 11-20 (1994) 掲載予定
- 4) 神門典子. 複数領域における日本語原著論文の機能構造分析：構成要素カテゴリーの自動付与. *Library and Information Science*. No. 31, p. 25-38 掲載予定
- 5) 川原村麻里. 構成要素カテゴリーによる情報メディアの内部構造分析の妥当性と応用. 東京, 慶應義塾大学, 1992, 69 p. 卒業論文.
- 6) 木下是雄. レポートの組み立て方. 東京, 筑摩書房, 1990, 234 p.
- 7) 木戸光子. “第7章 文の機能による要約文の特徴”. 文章構造と要約文の諸相. 佐久間まゆみ編. 東京, くろしお出版, 1989. p. 112-125.
- 8) 寺村秀夫. 日本語のシンタクスと意味Ⅱ. 東京, くろしお出版, 1984, 365 p.
- 9) 益岡隆志. モダリティの文法. 東京, くろしお出版, 1991, 232 p.
- 10) 益岡隆志, 田窪行則. 基礎日本語文法. 改訂版. 東京, くろしお出版, 1992, 251 p.
- 11) 仁田義雄. “現代日本語のモダリティの体系と構造”. 日本語モダリティ. 仁田義雄ほか編. 東京, くろしお出版, 1989, p. 1-56.
- 12) 中右実. “モダリティと命題”. 英語と日本語と：林栄一教授還暦記念論文集. 東京, くろしお出版, 1979, p. 223-250.
- 13) 近藤康弘. “ムード”. 講座日本語と日本語教育第4巻, 東京, 明治書院, 1989, p. 226-246.
- 14) 森山卓郎. 日本語動詞述語文の研究. 東京, 明治書院, 1988.
- 15) Swales, John; Najjar, Hazem. The writing of research article introductions. *Written Communication*, Vol. 4, No. 2, p. 175-191 (1987)
- 16) Crookes, G. Towards a validated analysis of scientific text structure. *Applied Linguistics*, Vol. 7, No. 1, p. 57-70 (1986)
- 17) 寺村秀夫, 佐久間まゆみ, 杉戸清樹, 半澤幹一編. ケーススタディ日本語の文章・談話. 東京, 桜楓社, 1990, 189 p. (接続関係については, 市川孝. 国語教育のための文章論概説. 東京, 教育出版, 1978, 236 p. に依拠している. また, 永野賢. 文章論総説：文法論的考察. 東京, 朝倉書店, 1986, 379 p. もほぼ同様の接続関係を提示している)