

TM-0678

談話文脈構造に関する考察

浮山輝彦, 小野顕司, 住田一男,
木村和広, 木下 聡, 天野真家(東芝)

February, 1989

©1989, ICOT

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

談話文脈構造に関する考察^{*}

浮田輝彦，小野顕司，住田一男，木村和広，木下聡，天野真家

株式会社 総合研究所

1. はじめに

テキストや会話の文章の内容全体を理解するには，文章が持っている文脈の構造を考える必要がある．会話やテキストの文章には，文の内容そのものを第一として，そのほかにも種々の言語的，非言語的な情報が含まれている．

本稿では，言語学的な情報である文間の接続表現や話題提示表現の性質を検討し，対象分野に関する世界知識を考えずに，文と文との関係を記述する文脈構造を考察する．

2. 談話の文脈構造

文そのものの関係に着目して文章の構造を記述する試みがある．文と文との関係に関して，Hobbs[79]は3種類の接続関係を整理しているが，文章全体の構造を体系付けるものではなく，さらに関係の接続関係の判定には，知識を使った推論が必要になっている．Groszら[85]は，言語学的な構造，注視状態(attentional state)，意図(intention)，の3種類を統合した構造を提案しているが，その抽出方法は必ずしも明確になっていない．Mannら[87]は，文間の接続関係を重視した文章の修辭的(rhetoric)な構造を提案している．また，辻井[88]は，新聞社説を対象として，文のタイプを区別し，内容的な，主張／認識／事実といった観点から構造化を試みている．河口ら[88]は，会話の意味構造を記述する中間言語を提案し，実データの記述を試みている．高松ら[88]は，文章構造のタイプをあらかじめ分類し，トップダウンに構造を判別し文章の内容を抽出する方法を検討しているが，方法の柔軟性に問題が残る．また，Schaら[88]は文脈の構造に関する各種の知見をまとめ，文脈自由文法で表現し処理する枠組みを提案している．

一方，談話に関連した概念として，(局所的な)話題と(大局的な)談話主題がある．本稿でいう話題は，比較的局所的なものであり，2～3の文で言及されるものである(いわゆるfocusも含めておく)．話題は，文(あるいは文章)が述べている対象の“ものごと”を示しており，文の内容そのものではない．従来，代名詞や指示詞の処理を，話題と

^{*} 本研究は，ICOIからの委託により第5世代コンピュータプロジェクトの一環として行っている．

の関わりあいでも議論することが多かった（例えば、[Sidner 83]）。また話題に対応して、“文脈空間”を設定し、A T Nにより解析する方法も提案されている[Reichman 84]が、実際の解析は行われなかった。また、本稿での談話主題とは文章のある集まり全体で述べられていることがらであり、いわゆるゴール [例えばWilensky 83]に相当するレベルのものである。これは、一連の動作全体が意味するものである。例えば、ソート処理を行うプログラムでは、プログラム（文）の集まりがソートという談話主題である目的を達成するものためのといえる。やはり、これも文章全体の内容そのもののかかわりを示すものではない。これらの話題や談話主題は、文章が述べている対象を示すものであり、いわばメタ的な観点から内容を表わしたもののといえる。

図1に本稿で想定している談話処理に関連する主要な要素を示す。処理対象の文章から、文間の接続関係を記述する文脈構造と話題の流れを抽出する。次に、話題の流れと文脈構造から談話主題が抽出され、同時に文章全体の意味構造を抽出することが一つの目標となる。例えば、テキストから書かれている内容を理解するには、まず第一段階として、談話の文脈構造として、文間の関係を明らかにし、局所的な部分で話題となっている対象を抽出することが必要である。そして次の段階として、文章全体の目的などの談話主題や、全体の意味構造が取りだされる必要がある。これらの処理には、文の接続表現などに関する言語学的な知見を最も基本的な要素として、話し手（書き手）・聞き手（読み手）・“場”の状況を含めた発話状況、内容を理解するさいに前提とされる世界知識と推論処理機能、そして、これらの要素をまとめて処理するときの原則や戦略が必要になる。この原則・戦略の例としては、文理解において一貫性・最適性の観点から、知識ベースにおける最大情報量を持つ解釈を探す方法[木下88；住田88]や、文章生成において説明のための各種戦略を用いる方法[田淵88]が知られている。このような処理により取りだされる文章全体の意味構造と談話主題を使って、例えばテキスト処理では要約や表現の変換を行うことになる。

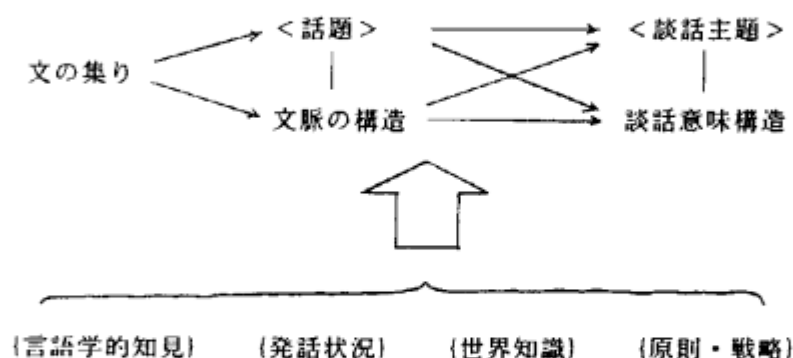


図1 談話処理に関連する主要な要素

3. 接続表現と文脈構造 [小野89]

3. 1 文脈構造

ここでは、[所 86; Mann 87] の考え方に従い、文間の修辭的な関係を文脈構造と捉え、文の接続関係を考察する。これは、「思考の流れ」しての文脈を考察することになる。特に論説文などを考えると、その論旨の展開に構造的なものを感じることができる。これは論者が自らの言説を相手に判らせるため、言説を細分化し、因果関係などでそれらを関係付けてまとめあげ、次にそれらを特定の修辭方法によって並べていくという経緯をへて論説文が書かれているからである。

修辭方法は、すべて2つの文あるいは文のまとまりの間を「関係」づけるものである。そして、どの文がどの文と、また文章のどの部分がどの部分とどう関係づけられているかを記述したものが文脈構造である。従って、まず文間の関係を定義する必要がある。いわゆる接続詞を基本に、複数のジャンル（演説文、経済記事、科学解説記事）の論説文数篇（約1000文）から抽出した接続詞相当語句約800を分類し、現在言明のレベルの関係として4種類、思考のレベルの関係として9種類定義した（表1）。この様な関係で、文章の構造を表現すると、次のようなものになる（対象とした文章を付録に示す。丸数字は文、ローマ数字がパラグラフを示す。ここで、パラグラフ内は言明のレベル、パラグラフ間を思考のレベルに相当するとしている）。

パラグラフ内：

- | | |
|-----|-----|
| I | ①→② |
| II | ③→④ |
| III | ⑤×⑥ |
| IV | ⑦ |
| V | ⑧ |

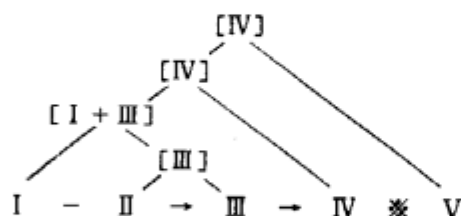
表1 文間の接続関係

(a) 言明のレベルの関係及びその修辭方法

	例	記号
①例示	“例えば”	▽
②重複	“というのは”	≡
③理由	“なぜなら”	←
④補足、婉曲	“但し”	～

パラグラフ間：

((I - (II → III)) → IV) ※ V



(b) 思考のレベルの関係およびその修辭方法

並列型：①並列	“また”	+	
②対比	“一方”	-	
直列型：③順接	“従って”	→	
④逆接	“しかし”	×	
同列型：⑤同列	“すなわち”	=	
指示型：⑥指示	“以下に…述べる”	※	
⑦参照	“図2に…示す”	※	
転換	⑧転換	“さて”	/
概括	⑨概括	“結局”	□

3. 2 文脈構造の抽出

文脈構造を文章データから抽出するには、その構造に関する規則性を表現する必要がある。文レベルに対する「非文」という“直観的な”基準が、文章については存在せず、単に「分かりやすい／分かりにくい」文章というとらえ方しかできないことを考えると、大局的な文章全体の構造分類は難しい。そこで局所的な規則性に着目することになる。文脈自由文法などの文法規則により、この規則性を表現しようという試みがある [辻井88; 川口88; Scha88]。しかし、原理的に人が文章を書くときの多様性を全て規則化する必要があること、すなわち論旨の進め方、内容の展開、想定している読み手の範囲などの自由度から生じる多様性を考えると、文章レベルの規則性を陽に（肯定的に）表現するには、対象分野を厳しく限定するといった条件を持ちこまない限り、現実的ではない。

ここでは、文間の関係を抽出するに当たって、局所的な規則性を制限として、関係の系列として許されない組み合わせを整理し利用することにより、文章データから文脈構造を取り出す。

前提と規則は以下の通りである。

前提 1 : 1 パラグラフ（節、段落）に対して一つのツリーが構成される。

前提 2 : ツリーは非交差である。

規則 1 : 以下の構造を含むものは、棄却する。

(..., 直列型関係 ((... (X 直列型関係 ...)))

(..., 同列型関係 ((... (X 直列型関係 ...)))

(..., 指示型関係 ((... (X 直列型関係 ...)))

(..., 指示型関係 ((... (X 並列型関係 ...)))

規則 1 中の 'X' は、その部分には 1 つの「言明」あるいは「思考」を構成する文あるいは文のまとまり全体が入ることを示している。形式的には何かの部分木（のルートノード）がそこに入ることを示している。従って、

例えば、

①	だから	②	だから	③
	順接 (→)		順接 (→)	
	直列型		直列型	

という修辞関係の列に対しては

(①だから②) ③

(① → ②) → ③

という構造は許されるが

①だから (②だから③)

① → (② → ③)

という構造は棄却される。

規則 2 : 以下の構造を含むものは、優先度を下げる。

(..., 直列型関係 ((... (X 並列型関係 ...)))

(..., 並列型関係 ((... (X 直列型関係 ...)))

(..., 同列型関係 ((... (X 並列型関係 ...)))

例えば

① さらに ② だから ③
 並列 (+) 順接 (→)
 並列型 直列型

という修辞関係の列に対しては

①さらに (②だから③)

① + (② → ③)

という構造より

(①さらに②) だから③

(① + ②) → ③

という構造が優先される。

この様な前提や規則を使うことにより、文間の全ての可能な構造から許される構造が限定できることになる。

3. 3 分析例

付録に示す文章を分析する。まず接続表現から可能な構造を抽出し、それに、話題提示表現の規則を適応する。分析結果を下に示す。これらの文章は 5 つのパラグラフに分かれており、パラグラフ単位に考えると、仮定 1, 2 を満たす構造は 14 個考えられる。それぞれの規則を適応した結果を示す。正解と考えられる構造は 9 ないし 14 に相当する。

	構造候補	接続規則	
		1	2
1	(I - (II → (III → (IV ※ V))))	X	X
2	(I - (II → ((III → IV) ※ V)))	X	X
3	(I - ((II → (III → IV)) ※ V))	X	X
4	((I - (II → (III → IV))) ※ V)	X	X
5	(I - ((II → III) → (IV ※ V)))	0	X
6	(I - (((II → III) → IV) ※ V))	0	X
7	((I - ((II → III) → IV)) ※ V)	0	X
8	((I - (II → III)) → (IV ※ V))	0	X
9	(((I - (II → III)) → IV) ※ V)	0	X
10	((I - II) → (III → (IV ※ V)))	X	0
11	((I - II) → ((III → IV) ※ V))	X	0
12	(((I - II) → (III → IV)) ※ V)	X	0
13	(((I - II) → III) → (IV ※ V))	0	0
14	((((I - II) → III) → IV) ※ V)	0	0

表2 文脈構造の分析例

○…正解候補がすべて含まれている △…一部棄却されている ×…すべて棄却されている

章, 節	文数*	規則1	規則2	章, 節	文数	規則1	規則2
A1 パラグラフ間	(5)	14	7 △	B1	3	2	1 ○
A2	3	2	2 ○	B2	8	429	260 ×
A3	3	2	2 ○	B3	8	429	58 ○
A4	5	14	10 ○	B4	10	4862	1358 ×
A5 パラグラフ間	(6)	42	4 ○	B5	18	129844790	解析不能
A6	4	5	5 ○	B6	9	1430	1430 ○
A7	4	5	1 ○	B7	5	14	5 ×
A8	3	2	1 ○	B8	5	14	11 ○
A9	5	14	9 ○	B9	6	42	42 ○
A10 パラグラフ間	(4)	5	2 ○	B10	8	429	359 ○
A11 パラグラフ間	(3)	2	1 ○	* 括弧内はパラグラフ数			
A12	3	2	2 ○	文書A: 東芝レビュー1988.9月号 pp. 710-714. 「火力発電設備の診断システム」 6章20段落 48パラグラフ 88文 文書B: 東芝レビュー1988.6月号 pp. 490-492. 「パーソナルワークステーションと 新しい潮流」 8段落 12パラグラフ 73文			
A13 パラグラフ間	(3)	2	1 ○				
A14	3	2	2 ○				
A15	4	5	5 ×				
A16 パラグラフ間	(4)	5	2 ○				
A17	3	2	2 ○				
A18 パラグラフ間	(3)	2	2 ○				
A19	3	2	2 ○				
A20	4	5	5 ○				
A21	5	14	1 △				
A22	5	14	14 ○				
A23 パラグラフ間	(3)	2	2 ○				
A24	3	2	2 ○				

各規則の欄でXとなっているところが、その構造候補がその規則に抵触していることを示している。この結果から、規則1では構造候補は7つにまで絞られ、正解と思われる構造はどちらもそのなかに残っている。しかし、プリファレンス規則である規則2を用いてさらに絞ると、候補は2つにまで絞れるが正解候補の1つが棄却されてしまう、ことが判る。

次に、同様の分析方法によって論説文2文書について調べた結果を述べる。このうち構造候補が複数出てくるパラグラフ、段落はそれぞれ16個、8個の計24個である。分析結果を表2に示す。

これらの結果から以下の結論を得た。

まず、各規則の妥当性（正解構造を棄却していないか）については、概ね良好な結果を得ることができた。誤判定したものを分析した結果、その原因は連体詞、代名詞の扱いによるものであることが判った。

関係抽出処理では現在、連体詞或いは代名詞で始まり（“…、この…”、“…、それらは…”等）他に接続的表現が無い箇所については直列型関係としている。しかしそれらは、他の直列型の関係表現（“従って”、“だから”等）と同じようには規則1や2に従っていないかった。

例えば

＊ ①一方②従って ③＊

という表層部分を含むものはすべて規則 2 に従っていたが、

＊ ①一方②これらは③＊

という表層部分を含むものについては、一部規則 2 に違反して、

…①－ (②→③) …

という構造になっているものも存在した。

次に各規則の候補制約力（効果）については、1/1 ～10/429と場所によっておおきなばらつきが在ることが判った。即ち、各場所での接続関係の並びに多分に依存していることが判った。平均すると、規則 1 には約 2/3、規則 2 まで用いた場合には約 1/5 にまで候補を絞る能力がある。

最終的に検出される構造の候補数は、文書の文体にも大きく依存している。文書 B は文書 A に比べて 1 文が短く、1 パラグラフ中の文数は多い。このように文数が多いと接続関係の相対的組み合わせが増え、各規則が効き易くなり、効果は大きくなる。

候補が殆ど絞れていないものについて調べてみると、それらは接続表現が乏しい部分であった。関係抽出処理では現在、接続表現が無い箇所は、並列型或いは直列型の関係として扱っている。しかしそのような箇所が多いと、各規則があまり効かなくなってしまう。文書 B ではこれが原因となって候補が絞れていないケースが大半であった。

4. 話題と文脈構造

4. 1 話題提示表現

従来、話題提示表現は、照応・省略表現との関わりで論じられることが多かった【例えば Sidner 83; Kamoyama 86; Yoshimoto 88】が、ここでは、話題提示表現の性質とその文脈構造へのかかわりを中心に考察する。

話題が設定できることにより、書き手は、話題との関わり合いにおいて簡潔な表現をすることができ、Grice の公準（量の公準 [Grice 75]）としていわれるように、過不足ない表現を行うことができる。反対に、読み手にとっては、その注意を局所化し、文章の内容を理解する助けとしての機能を持つといえる。従って、話題の流れ自体は、書き手から読み手へと、言語表現を通じて受け渡される注意の範囲であり、文章の関係を表現する文脈構造との直接的な関係はないものと考えられる。

しかしながら、話題となっているものが聞き手にとって自明でない場合には、既出の言葉を受けるという形で、その概念を話題として提示し、それがどのようなものであるかといったことを、文章として表現することが多くみられる。従って、話題の表現が文間にわたって明確に検知できる場合には、それらを含む文の間の関係を決定する手助けとして使える可能性を持っている。例えば、ある解説文の 1 パラグラフの文章を見してみる。

対話を通じて、ユーザから要求仕様を獲得するプログラム合成システムにPSIがある。PSIは、緊密に相互作用しあうプログラムモジュール(エキスパートと呼ぶ)からなる、知識依存型のシステムである。PSIでは、システム全体の動作を仕様獲得フェーズとプログラム合成フェーズに分けて考えている。仕様獲得フェーズでは、まずパーサ/インタプリタ・エキスパートが入力文を解析し、構文解析木を解釈してプログラムネットと呼ぶネットワーク型データ構造を作成する。パーサ/インタプリタ・エキスパートは約70のプログラミング概念と約175語の語彙に関する知識を持っている。プログラミング概念のなかには、データ構造、制御構造、プリミティブな演算、複雑なアルゴリズムの概念などが含まれている。

<辻井・上原：“ソフトウェア工学と自然言語処理”，情報処理，28.7(1987)，p.915より>

この文章の場合、最初の文を除いて、下線で示される話題提示表現が観察され、しかも、各文の“話題”となる表現は、その直前の文のなかに既に出現している。話題の流れとして、順に話が展開されていく様子が良くわかる。文脈構造として、文間の関係について考えると、この例における話題の展開は、“補足”という関係で文が順次結びつけられていると言える。

話題の提示表現は、永野[86]等々に示されているが、ここでは最も簡単な表現である助詞の「は」について検討する。まず、簡単な形態素解析を行った後に、助詞の「は」を持っている表現を取り出し、それらを分類する。話題表現が前方に現れるかどうかを中心に分類した結果を表3に示す。分析データは、先の分析に用いた論文2編で、161文である(ここで文は簡単のため読点(。)から読点までとしている)。話題提示表現としてピックアップしたものは、

「XXは」，「XXには」，「XXでは」，「XXに関しては」，「XXとしては」
「XXのためには」，「XXの場合には」，

という形式のものであり、XXが

“名詞”，“代名詞”，“名詞+の+名詞”，“指示詞+名詞”

の場合である。また話題提示表現の名詞が既に出現したかどうかの判定は、最小の節内(なんらかの見出しがある場合には、そこを節の始まりと考えた)の範囲で、いま検討している“話題”から前方のみをチェックしている。

表3の主題提示表現の内訳を概観する。

1. の“前方に出現しないもの”は、ほとんどのものが読み手には自明であろうと考えられるものである。従ってこれらの表現は、話題を提示するが、必ずしもテキストにおける文間の関係を提示しないと考えてよい。

2. の“同語反復”は、話題としての文字列が、既に出現している場合である(上記のパターンでXX = “・・・+名詞”の場合は、後ろの名詞が反復される場合)。それが前方で一ヶ所しかない場合は、その文を受けていると考えることが自然である。同語反復は、文中に現れる概念の文間に渡る対応関係が、表層レベルで明瞭に観察されるものである。従って、文間の関係度を判断する一つの基準として使うことができる。

表3 助詞「は」表現の分類

分類	個数	表現の例
1. 前方にないもの	50	「最近のエレクトロニクス技術は」「この章では」「最近では」「図3は」
2. 同語反復	25	「火力発電設備は」「エキスパートシステムとしては」
3. 部分反復	23	「火力発電所の高度情報化システムは」
4. 代名詞	7	「これは」「このためには」
5. 指示詞+名詞	10	「これらの判断は」「このような状況は」「このような際には」 「このようなエキスパートシステムでは」「このようなデータベースでは」
6. 句・文(連体修飾) + (時・場合・こと…) + は、には(24)	24	「…を成功させるためには」「…するような場合には」「…があることは」
7. 「は」なし文	161文中63文	

分析条件などは本文参照

データ：東芝レビュー「火力発電設備の診断システム」('88.9)

「パーソナルワークステーションと新しい潮流」('88.8)

分析文数161文、ただし1文中に2回以上「は」の表現を持つものがある

3. の“部分反復”は、前方の部分的に出現する表現を含む可能性が大きいと考えられるが、知識なしで、対応関係をとるのは難しい。

4. の“代名詞”は、「ここでは」等の表現を除いて、前方ないし後方と関係していることは明らかである。代名詞・指示詞も同様であるが、対応関係の同定には知識が要求されることが多く、確実性のある測度として利用するのは難しい。

5. “指示詞+名詞”は、「この章では」等を除くと、前方ないし後方と関係していることは明らかである。

6. の場合は、複合した概念や複雑な条件が述べられており、話題としての言葉の抽出自体が難しい。

161文中63文には、「は」の表現がなかった。例えば、次のような文である。

「以下に火力発電設備の診断システムに関する概況およびその具体例を紹介する。」

「図5に海水漏洩の警報に対する操作支援画面の一例を示す。」

「これが、運転支援システムの導入の必要性がとえられているゆえんである。」

これらの内、最初の2文では、論文の構成や図についての説明が行われている。これらの場合の話題は、「論文の構成」や「図」になっていると考えることができる。文章の構成についての説明を一段高いレベルから読み手に対して直接表現するものは、別途取り出し扱う必要がある。また3番目の文は、特に何についての記述であるかは表現されておらず、この文に先行する話題が継続していると考えてよい。また、箇条書きに近いような短い文の場合には、その見出しが話題となっており、文中にはほとんど「は」の表現が現れない傾向も観察できた。

「は」の表現のあるものの中で最も多く観察されたものは、1の“前方にないもの”である。次に多かったものは、同語反復や部分反復と読んでいるものである。このうち部分

反復では、「プラント運転の全自動化の普及」という表現では、その節の中で、「プラント」、「運転」は出現しているが、「全自動化」、「普及」は出現していない。この欠落している部分は、1の“前方にないもの”と同様に、読み手がその言葉を十分カバーできる（単語の意味が読手にとって自明であり、既出の単語との関係が推論でき、話題として提示される概念が形成できる）と書き手が判断していると考えることができる。従って、部分的に単語が共通に観測される文間では、その先行する文と、現在の主題表現との文の関係は深いといえる。しかし必ずしも直接的に結び付くとは断定できない。

次に、話題提示表現の単語列とまったく同じものが前方の文中に観測される同語反復の場合は、文間の関係として、前方の文の展開に現在の文が位置付けられる場合が多い。また部分反復の時に、取り分け特徴的になるが、反復する語が、前方の複数の文に2か所、3か所と現れることがある。このような場合には、先行する文のうち、どれと関係があるのかは一意に決定できない。またすべての文と関係がある可能性もある。反復出現している語の数にも大小がある。反復される語数が多いものを優先するのが妥当と考えられる。

代名詞、指示詞+名詞の形式の場合は、先行詞が決められないかぎり、どの文におけるものごとを受けるとの判定はできない。単に前方のどれかの文中の概念を受けるとしか判断できない。しかし、既に提示されている話題提示表現より以前まで遡って概念を参照することはないといえる。

4. 2 話題と文脈構造

話題は、文が述べている対象を示すものである。既に前方の文で現在対象となっている概念が提出されている場合には、その文と展開や補足といった直接的な関係があるといえる。従って、文脈構造を記述する際に、これらの文間の関係が近い構造を優先する、一つの手掛かりとすることができる。

まず話題は、次のような処理原則により設定されたとする。

- (a) 話題提示表現が現れると、それによって示される話題が現在の話題となる。
 - (b) 話題提示表現が明確に示されるまで、現在の話題が継続する。
 - (c) テキストの場合、タイトル・見出しは、それ自体で話題となる。
- この話題を使って文脈構造の設定に利用する方法として次のものが基本的に整理できる。
- (1) 提示された話題が、先行する文のなかで既に同じ文字列として、反復出現している場合、現在の文は、その文と直接的な*関係を持つ。
 - (2) 提示された話題表現が「代名詞」「指示詞+名詞」の形の場合、その文は、先行する文の話題提示表現のある文以降の文のいずれかと直接的な関係を持つ。
 - (3) 話題提示表現のない文では、既に提示されている話題が保存される。したがって、現在の文はその話題を提示した文以降のものと関係づけられる。
- (* 直接的な関係とは、文脈構造の中で文と文が一つの関係（枝）で結合されることを示す。)

この様な話題呈示表現と話題となる言葉の反復出現を使って、文脈構造の曖昧さを減少できる可能性がある。

5. むすび

本稿では、日本語文章の構造情報を考察し、修辭的な文間の関係を表わす文脈構造を中心に考察した。特に言語表現の接続関係と話題提示表現の性質を検討し、それらと文脈構造の関係を考えた。文章の構成を表現する文脈構造は、文間の（文章の構成部分間の）関係を記述する必要があるとの立場から、文間の接続表現を抽出し、それにより、2分木の形式で関係を記述する方法をとった。具体的には、約800個の接続表現を13種の接続関係に分類し、それら接続関係の組み合わせについて考察し、局所的な見地から文章の構造として許されるものを規則化した。論文2編、161文を実際に分析し、正しい形態素解析を想定して、自動処理の可能性を検討した。その結果、曖昧さは残るものの、テキストが言及している対象分野の知識や発話状況の上をまったく仮定せず、言語学的な情報として接続関係の情報を使うことにより、可能な構造を効果的に削減できることを示した。また話題の提示表現を分析し、その表現で示される言葉が先行文脈に出現している場合に、文脈構造の抽出に利用できる可能性を示した。今後、より多くのデータに関して検討し、各種の規則の精密化を進めていく。

付録 分析データの例

I-①火力発電設備は大幅な負荷調整能力と運用の多様化の要求にこたえなければならないという宿命を帯びている。②このような過酷な条件下でもなお長期的な高信頼度運用を達成するためには、まず、火力発電設備を構成する各機器の安全性の確保が前提となる。

II-③一方、運転状態監視の強化や計画的予防保全などにより、定期点検間隔の延長化や、補修業務の合理化による各機器の長寿命化が試みられている。④このためには、運転中の機器の異常を早期に検出し、原因の分析や、対策に結び付く適切な情報の提供、あるいは、経年的な変化傾向を把握するなどの、きめ細かなデータの管理と分析、およびそれらの診断のための技術の確立が必要となる。

III-⑤これらの判断は、これまで運転員あるいは補修員が監視計器などからの情報を、一般的判定基準に自己の経験を加味して行ってきたものであるが、⑥運用の多様化と介在するシステムの高度化とともに、判断を必要とする情報量も飛躍的に増加しつつあるのが現状である。

IV-⑦このような理由から、各種の情報を整理し加工して、運転や保全のための情報を提供する、計算機による診断システムの開発が急がれている。

V-⑧以下に、火力発電設備の診断システムに関する概況およびその具体例を紹介する。

（東芝レビュー1988.9月号 p710「火力発電設備の診断システム」から抜粋）

[参考文献]

- [小野 89] 小野, 浮田, 天野: “文脈構造の分析”, 情報処理学会 自然言語処理研究会資料 89-NL-70, 1989.
- [河口 88] 河口, 立花, 谷口: “会話文の文意構造記述に関する試み”, 昭和63年度人工知能学会全国大会(第2回)予稿集, 9-1, 1988.
- [木下 88] 木下, 佐野, 浮田, 住田, 天野: “文脈理解のための知識の表現と推論”, Proc. The Logic Programming Conference '88, pp.205-215.
- [住田 88] 住田, 浮田, 天野: “解釈の情報量を考慮した文脈理解”, 信学技報 NLC 88-10, 1988.
- [高松 88] 高松忍, 西田富士夫: “見出し情報を用いたテキスト解析と情報抽出”, 情報処理学会論文誌, Vol. 29, No. 8, pp.760-769, 1988.
- [田淵 88] 田淵, 辻井, 長尾: “文脈を考慮したテキスト生成システム”, 情報処理学会 自然言語処理研究会資料 88-NL-65, 1988.
- [辻井 88] 辻井 潤一: “論説文における文脈構造”, 日本学術振興会 文字言語・音声言語の知能的処理第152委員会 第7回研究会資料7-1, 1988.
- [所 86] 所 一哉: “現代文レトリック読解法”, 匠出版, 1986.
- [永野 86] 永野 賢: “文章論総説—文法論的考察—” 朝倉書店, 1986.
- [Allen 80] Allen, J.F.: “Analyzing Intention in Utterances”, Artificial Intelligence, Vol.15, pp.143-178, 1980.
- [Grice 75] Grice, H. P.: “Logic and Conversation”, Syntax and Semantics, Vol. 3, Speech Act, Seminar Press, pp. 41-58, 1975.
- [Grosz 85] Grosz, B.J. and Sidner, C.L.: “Discourse Structure and the Proper Treatment of Interruptions”, Proc. IJCAI-85, pp.832-839, 1985.
- [Hobbs 79] Hobbs, J.R.: “Coherence and Coreference”, Cognitive Science, Vol.3, pp.67-90, 1979.
- [Kameyama 86] Kameyama, M.: “A Property-sharing Constraint in Centering”, Proc. 24-th Annual Meeting of the Association for Computational Linguistics, pp.200-206, 1986.
- [Mann 87] Mann, W. and Thompson, S.: “Rhetorical Structure Theory: A Framework for the Analysis of Texts”, USC Information Science Institute Research Report, RR-87-190, 1987.
- [Scha 88] Scha, R. and Polanyi, L.: “An Augmented Context Free Grammar for Discourse”, Proc. COLING-88, pp.573-577, 1988.
- [Sidner 83] Sidner, C.L.: “Focusing in Comprehension of Definite Anaphora”, M.Brady and R.C.Berwick (Eds.), Computational Models of Discourse, MIT Press, pp.267-330, 1983.
- [Reichman 84] Reichman-Adar, R.: “Extended Person-Machine Interface”, Artificial Intelligence, Vol.22, pp. 157-218, 1984.
- [Wilensky 83] Wilensky, R.: “Planning and Understanding”, Addison-Wesley Publishing Company, 1983.
- [Yoshimoto 88] Yoshimoto, K.: “Identifying Zero Pronouns in Japanese Dialogue”, Proc. COLING-88, pp. 779-784, 1988.