

日本語文章における照応・省略現象の基本的検討

藤澤 伸二[†] 増山 繁[†] 内藤 昭三^{††}

本論文では、日本語文章において、文の冗長度を下げることを目的としてしばしば使用される照応と省略が、実際の文章でどのように用いられているかを調査した結果、および、その結果から得られた考察を報告する。まず、科学雑誌の記事5編と天声人語70編を調査の対象として、どのような句が省略されているか、どのような指示詞、代名詞が使われているか、これらの先行詞がどのくらい前（あるいは後）の文に現れるかを調査した。また、照応・省略に関する八つの仮説を立て、これらの成立についても検証を行い、全体的に見てほぼ成立していることがわかった。さらに、この検証結果から得られた照応と省略の使われ方の違いと、これらの仮説に対する反例が生じた理由を考察した。

A Basic Study on Ellipsis and Anaphora in Japanese Sentences

SHINJI FUJISAWA,[†] SHIGERU MASUYAMA[†] and SHOZO NAITO^{††}

In this paper we investigate how ellipsis and anaphora are used in actual Japanese sentences. We study what kind of phrases are omitted, when anaphora is used, what kind of pronouns and proverbs are used, and where and how many times they appear against articles and essays. We also propose eight hypotheses about ellipsis and anaphora and examine their validity. Then we analyze the distinction between how ellipsis and anaphora are used and the reason why some examples do not meet some hypotheses.

1. はじめに

本論文では、日本語文章における照応と省略現象に焦点を絞り、それらが実際の文章でどのように用いられているかを調査した結果と、その結果に対する考察を述べる。

照応・省略は、日本語においてしばしば用いられるもので、聞き手が文脈から復元可能であると話し手が判断できる情報を、簡略に表現したり、省略することにより、文の冗長度を下げることができる¹⁾。照応とは、代名詞、指示詞などの照応語とその先行詞との組によって表される言語現象で、伝達される情報のかなりの部分が談話文脈から推定できる²⁾。一方、省略は、聞き手にとって自明の情報を省く言語現象であるが、ここでは、省略を IPAL 辞書^{3),4)} 記述中の必須格要素の省略と定義する。省略は、照応現象の一種（「ゼロ照応」とよぶ）と見なすことができる。

照応や省略に関する従来の研究の中で、日本語文を

対象とした研究としては、照応や省略の定性的評価をした報告¹¹⁾や、照応と省略を用いた意味解析や談話解析を行ったという報告^{12),13)}、さらには、省略語の補完（復元）についての報告^{15),16)}などがある。このように、照応・省略は、顕著な談話現象であり、その扱いを考察した報告は少なくない¹⁷⁾⁻²¹⁾。また、英語文を対象とした研究としては、実際の文章を分析し、得られた結果に基づき照応についての考察を行ったという報告¹⁴⁾がある。しかし、これらの研究では、照応・省略と段落との関連性や、複数の照応・省略間の関係について論じられてはいない。本研究では、段落と照応・省略との関連性、複数の照応・省略間における非交差性などの性質に注目することによって、照応・省略現象に関する言語的な制約に関する理解を深めることをねらいとしている。これらの性質は、文章の段落分け⁹⁾や照応・省略の先行詞復元、照応・省略表現の生成などの談話処理における有力な制約やヒューリスティックスとして利用できる。

本論文の具体的な内容は以下のとおりである。まず、日本語文章において、照応や省略がどのように使われているか、また照応と省略の使用法にどのような違いがあるかを知るために、統計調査を行った。調査の対象とした文章は、科学雑誌（日経サイエンス）の

[†] 豊橋技術科学大学知識情報工学系
Department of Knowledge-based Information
Engineering, Toyohashi University of Technology
^{††} NTT 基礎研究所
NTT Basic Research Laboratories

記事5編と天声人語70編である。この調査の項目は、どのような指示詞、代名詞、代動詞が使われているか、どのような句が省略されているか、これらの先行詞がどのくらい前の文に出現しているか、の3点である。さらに、照応・省略に関する八つの仮説を立て、その成立についての検証を行った。次に、これらの仮説の検証結果から明らかになった照応と省略の使われ方の違いと、これらの仮説に対する反例が生じた理由を考察した。

本論文の構成は、以下のとおりである。第2章で文タイプの分類について述べ、第3章で照応・省略に関して立てた八つの仮説を提示する。第4章で、仮説の検証について、その調査方法と調査結果を報告する。第5章で、調査結果に対する考察、および、各仮説の反例とそれらが生じた理由を述べ、最後に、第6章でまとめと今後の課題を述べる。

2. 文の種類

文の種類は通常、単文、複文、重文に分けられる。以下に、それぞれの定義と例文を示す。この定義は、文献5)に基づき、文献6)も考慮に入れて修正を加えたものである。

複文：少なくとも一つ埋め込み構造を持つ文。複文中の節を埋め込み文という。

(例文1) 雪の降る日は寒い。

重文：最上位で複数の主述関係が互いに埋め込み文とならずに並行しているもの。

(例文2) 花は咲き、鳥は歌う。

単文：重文でなくかつ複文でない文。

(例文3) 犬がいる。

この定義は、重文かつ複文も認めている。さらに、重文でない文(つまり、最上位での主述関係が一回だけのもの)を単位文と定義する⁶⁾。上の例で、(例文1)と(例文3)が単位文である。また、複文の中の主述関係の認められる従属的部分や重文の中の主述関係の成り立っているそれぞれの部分を節という⁶⁾。ただし、重文を除いた文自身も一つの節であると定義する。埋め込み文は、次の(例文4)のように何重にも入れ子状にできる。そのとき、入れ子状の外側(内側)を上位(下位)の埋め込み(節)という。

(例文4) 太郎が来たとき次郎が言ったとき三郎は思った。

3. 照応・省略に関する仮説

検証した照応・省略に関する仮説は以下のとおりである。(仮説とともに示した例文は、(例文5)を新規に作成し、その他は日経サイエンスから引用した文である)

(仮説1) 照応は段落に跨ることがない。

(仮説2) 照応は交差することがない。

(例文5) 山に登れば、そこから町全体を見渡せ、その地形が手にとるようにわかる。

(例文6) マイブリッジは高速度撮影を純粋に楽しむために開発した。今日、その技術は、彼やスタンフォードの予想をはるかに越えてしまった。

(例文5)と(例文6)はそれぞれ照応が2回行われている。(例文5)において、そこの先行詞として「山」、そのの先行詞として「町」が考えられる。また、(例文6)において、そのの先行詞として「高速度撮影」、彼のの先行詞として「マイブリッジ」が考えられる。したがって、これらの例文において、二つの照応語と先行詞の組の間には交差が見られない。ゆえに、(例文5)、(例文6)のいずれも(仮説2)を満たす。

(仮説3) 複文および重文の先頭を除く節中の照応に対する先行詞は、その節および、その節の上位の節には含まれない。

(例文7) 分子の運動はきわめて速く、それを撮影するのに必要な時間分解能、つまりシャッタースピードは、私たちの時間の感覚をはるかに越えている。

(例文7)は重文であり、そのの先行詞は、最初節にある「分子の運動」である。それは重文の2番目の節に含まれ、その先行詞は、その節でも上位の節でもない重文の先頭の節に含まれている。したがって、この例文は(仮説3)を満たす。

(仮説4) 単文中の照応、および、重文の先頭節で、しかも埋め込み節でない節中の照応に対する先行詞は、それより前の文(段落を跨ってもよい)にある。

(例文8) 機械観測史上初期の記録は、あまり役立たなかった。これらの記録は西ヨーロッパしかカバーされていなかったからである。

(例文8)において、照応語これが使われている。これの先行詞は、「機械観測史上初期」である。こ

の(例文8)の第2文は単文であり、そこに含まれる照応語これの先行詞が、それより前の第1文にある。したがって、この例文は(仮説4)を満たす。

(仮説5) 省略は段落間を跨らない。すなわち、省略の先行詞は、省略語と同一段階中にある。

(仮説6) 省略は交差しない。すなわち、省略と先行詞の組の間には、交差が生じない。

(例文9) 日本では、「さきがけ」と「すいせい」のコンビがハレー彗星に送られた。そして、 ϕ_{1-1} それまでに磁気圏内で磨いた技術をいかし、 ϕ_{2-2} 太陽風の観測に成功した^{2*}。

(例文9)での2回の省略(ϕ_{1-1} と ϕ_{2-2})では、「日本が」が省略されており、その先行詞はどちらも最初の文の先頭の「日本では」にあり、確かに一つめの先行詞と省略語の組と二つめの先行詞と省略語の組との間には交差は生じていない。したがって、この例文は(仮説6)を満たす。

(仮説7) 最上位の節を除く各節において、省略の先行詞はその上位の節に含まれる。

(例文10) 私たちの研究グループは、 ϕ_{1-1} 南半球(全球)の温度変動を解析する¹ための10年間におよぶ研究をこのほど終了した。

(例文10)において、「解析する」の節の省略語 ϕ_{1-1} の先行詞はその節の上位の節の「私たちの研究グループ」にある。この例文において、最上位の節ではない「 ϕ_{1-1} 南半球(全球)の温度変動を解析する」の節に含まれている省略語 ϕ_{1-1} の先行詞がこの節の上位の節にある。したがって、(例文10)は(仮説7)を満たす。

(仮説8) 単文中の省略、および、重文の先頭節で、しかも埋め込み節でない節中の省略に対する先行詞は、それより前の文中にある。

(例文11) 温度計が熱を持った近くに設置されたかを知ることはほとんど不可能である。このような場所に ϕ_{1-1} 設置されて¹おれば、温度計の読みは高くなってしまう。

(例文11)において、重文である2番目の文の最初の節において、主格が省略されている。その省略語の先行詞は、最初の文の「温度計が」である。したがって、この例文は(仮説8)を満たす。

これらの仮説は、書き手は「協調性の原則」に従

う²⁾、すなわち、できるだけ読み手にとってわかりやすい文章を書くとの仮定の下に立てたものである。

「協調性の原則」などの談話の制約は、必ずしも完全に成立することを要求するものではなく、「原則」という表現にも現れているように、デフォルト制約的な要因と考えるのが妥当である。ここで提案した仮説も、このような性質を持ったものである。特に、(仮説2)と(仮説6)は一般によく知られている「1文内における係り受けの非交差性」から類推して立てたものである。たとえば、(仮説1)、(仮説2)は完全に成立するというのではなく、ほぼ成立するものであり、その成立の度合いは、(仮説5)、(仮説6)に比べて低いことが予想される。その理由の一つとしては、省略の場合と違って、照応の場合は先行詞に関する程度の情報を照応語が持つことができるということが挙げられる。

4. 仮説の検証

4.1 調査方法

今回の調査は、以下のように人間の視察によって行った。

4.1.1 省略語の先行詞の決定方法

Step 1 (述語の復元)

もし、述語が存在しなければ、以下の処理を行う。述語の省略には、並列表現による省略、「ダ」パターンによる省略³⁾、慣用的な表現による省略の3種類があるので、並列表現による省略の場合には対応する述語を選び、他の場合は、その述語の含まれる文の要素から予想される述語を選びだし、その中から適切なものを選ぶ。ここで、適切なものとは、文に矛盾が生じず、かつ、述語が省略されている文にその述語を当てはめた文が日本語文として適格であると判断される述語である。

Step 2 (省略格の認定)

文中の要素に、その述語のIPAL辞書記述中の必須格が一つでも欠けているならば、その欠けている要素を省略されている格(省略語)とする。

Step 3 (先行詞の決定方法)

その省略されている格に、以前出現した文の要素やそれらと関連のある要素をあてはめ、適切なものを省略語の先行詞の候補として選ぶ。ここで、適切なものとは、文に矛盾が生じず、かつ、省略されている文にその要素を当てはめた文が日本語文として適格であると判断される語である。その中で、省略語の位置から

* 省略を示す記号 ϕ_{i-1} のサフィックス1-1の最初の数は、省略の番号を表し、2番目の数は、(番号が付されている)どの述部に対する省略かを表している。この例は、述部「いかし」に対する省略格である。

最も近くにある語をその省略語の先行詞とする。

4.1.2 照応語の先行詞の決定方法

照応の先行詞の決定方法は、以下の通りである。まず、以前出現した文の要素や照応語と関連のある要素(句、節、文、段落)を、照応語のある位置にあてはめ、適切なものを照応語の先行詞の候補として選ぶ。ここで、適切なものとは、文に矛盾が生じず、かつ、照応語をその先行詞に置き換えた文が日本語文として適格であると判断される語である。その中で、照応語の位置から最も近くにある語をその照応語の先行詞とする。

4.2 調査結果

調査結果を表1から表12に示す。ただし、表3では、照応語の先行詞が文に跨る場合(出現数はサイエンスで26、天声人語で30)を除いており、また、表8、表9では、省略の全出現から「不特定(多数)者、筆者、読者」(以下、不特定者等、とよぶ)を除いたもの(出現数はサイエンスで308、天声人語で125)を対象とした。また、表4、表10では、照応や省略に交差が生じていない回数の決定が困難であるため、1文中で2回以上照応語や省略語が現れた文のみを対象として、照応や省略に交差が生じている回数、生じていない回数を調査した。

表1 照応語の出現回数
Table 1 The number of occurrences of anaphoras.

語句	出現回数	
	サイエンス	天声人語
指示詞	420	281
これ	70	32
この	191	56
こんな	0	8
ここ	1	4
こう(副詞)	0	3
それ	58	56
その	85	79
そんな	0	15
そこ	15	19
そう(副詞)	0	4
あの	0	3
あれ	0	2
代名詞	15	27
わたし	1	2
彼	14	20
その他	0	5
代動詞	17	29
こう(動詞)	12	16
そう(動詞)	5	13

今回調査を行った科学雑誌と天声人語において、段落数と文数はそれぞれ、科学雑誌の記事5編で317、1087、天声人語70編で436、1426であった。1段落あたりの文数の平均は、科学雑誌で3.4、天声人語で

表2 照応の先行詞の含まれる文(文それ自身を指す場合も含む)

Table 2 The position of the sentence in which the antecedent of the anaphora is included. (The antecedent may be in the same sentence in which it is included.)

出現位置	出現回数	
	サイエンス	天声人語
	426	307
一つ後	0	2
同一文中	121	82
一つ前	280	165
二つ前	19	27
三つ前	4	18
四つ前	2	8
その他	0	5

表3 照応の先行詞の含まれる段落(段落それ自身を指す場合も含む)

Table 3 The position of the paragraph in which the antecedent of the anaphora is included. (The antecedent may be in the same paragraph in which it is included.)

出現段落	出現回数	
	サイエンス	天声人語
	452	337
同一段落中	411	272
一つ前	41	62
二つ前	0	3

表4 仮説2の成立状況(条件付)

Table 4 The number of instances which satisfy (or do not satisfy) Hypothesis 2.

	出現回数	
	サイエンス	天声人語
成立	31	25
不成立	0	1

表5 仮説3の成立状況

Table 5 The number of instances which satisfy (or do not satisfy) Hypothesis 3.

	出現回数	
	サイエンス	天声人語
成立	416	322
不成立	36	15

表 6 仮説 4 の成立状況
Table 6 The number of instances which satisfy (or do not satisfy) Hypothesis 4.

出現位置	出現回数	
	サイエンス	天声人語
一つ後	0	2
同一文中	33	15
一つ前	272	170
二つ前	18	28
三つ前	3	14
四つ前	2	6
その他の文	0	3
前の部分	26	30

表 7 省略語の出現回数
Table 7 The number of occurrences of ellipses.

語句	出現回数	
	サイエンス	天声人語
名詞句	636	473
一は	407	406
一が	210	51
一に	10	9
一を	7	7
一と	2	0
修飾詞句	9	8
一の	7	8
一な	2	0
動詞句	18	38
並列	9	3
その他	9	35

表 8 省略の先行詞の含まれる文
Table 8 The position of the sentence in which the antecedent of the ellipsis is included.

出現位置	出現回数	
	サイエンス	天声人語
文	330	348
同一文中	264	147
一つ前	62	121
二つ前	3	48
三つ前	1	12
その他	0	20

3.2 と大差ない。今回は、ほぼ同量の文章について調査したので、天声人語のほうが1文あたりの文字数が少ないことになる。

調査結果をまとめると、以下のようになる。

1. 照応に関して

- 照応語のほとんどは指示詞であった (科学雑誌で

表 9 省略の先行詞の含まれる段落
Table 9 The position of the paragraph in which the antecedent of ellipsis is included.

出現段落	出現回数	
	サイエンス	天声人語
	330	348
同一段落中	328	295
一つ前	2	46
二つ前	0	7

表 10 仮説 6 の成立状況 (条件付)
Table 10 The number of instances which satisfy (or do not satisfy) Hypothesis 6 (under the condition stated in Section 4.2).

	出現回数	
	サイエンス	天声人語
成立	30	70
不成立	0	0

93%, 天声人語で 83%) (表 1 参照).

- 科学雑誌では、「こ」で始まる指示詞が多かった (58%) のに対し、天声人語では、「そ」で始まる指示詞が多かった (51%) (表 1 参照).
- どちらも「あ」で始まる指示詞は非常に少なかった (天声人語で 5 回のみ) (表 1 参照).
- 科学雑誌で人称代名詞は「彼」「わたし」しか用いられなかったが、天声人語ではそれらに加え「だんな」や「氏」も人称代名詞的として使われていた (表 1 参照).
- 先行詞は、照応語を含む文の直前の文にあることが多かった (科学雑誌で 62%, 天声人語で 49%) (表 2 参照).
- 天声人語には、後方照応も使用されていた (2 回のみ) (表 2 参照).

2. 省略に関して

- 省略格のほぼ 9 割が「が格」であった (科学雑誌で 93%, 天声人語で 88%) (表 7 参照).
- 先行詞は、科学雑誌では省略語を含む文に含まれる場合が多い (80%) のに対し、天声人語では、その文に含まれる場合 (42%) と、一つ前の文に含まれる場合 (35%) がほぼ同じであった (表 8 参照).

3. 仮説の成立に関して

仮説の成立基準は、100% を「成立」、80% 以上~100% 未満を「ほぼ成立」、60% 以上~80% 未満を「必ずしも成立するとはいえない」、60% 未満を「成立しない」とした。

- 段落を越える照応は、科学雑誌で 452 回中 41 回 (9%)、天声人語で 337 回中 65 回 (19%) 使われていた。照応語とその先行詞が同一文中にない照応のうち、段落を越えるものの割合は、科学雑誌で 13% (305 回中 41 回)、天声人語で 29% (225 回中 65 回) であった (表 2, 表 3 参照)。したがって、(仮説 1) は科学雑誌でほぼ成立しているが、天声人語では必ずしも成立しているとはいえない。
- 照応の交差現象は、科学雑誌で 2 回、天声人語で 1 回現れていた。また、1 文中で 2 回以上照応語が現れるという条件の下では、サイエンスで 100%、天声人語で 96% の割合で成立していた (表 4 参照)。したがって、(仮説 2) はほぼ成立している。
- 複文や重文の先頭を除く各節において、照応の先行詞がその節やその上位の節には含まれない割合は、科学雑誌で 92%、天声人語で 96% であった (表 5 参照)。したがって、(仮説 3) はほぼ成立している。
- 単文、または重文の先頭の節で、埋め込み節以外の照応の先行詞がそれより前の文 (段落を跨ってもよい) にある割合は、科学雑誌で 91%、天声人語で 94% であった (表 6 参照)。したがって、(仮説 4) はほぼ成立している。
- 段落を越える省略は、科学雑誌で 330 回中 2 回 (1%)、天声人語で 348 回中 53 回 (15%) であった。(注: 省略の全回数、不特定者等の省略を除いた省略の回数) 省略語とその先行詞が同一文中にない省略のうち、段落を越えるものの割合は、科学雑誌で 3% (66 回中 2 回)、天声人語で 26% (201 回中 53 回) であった (表 8, 表 9 参照)。したがって、(仮説 5) は科学雑誌でほぼ成立しているが、天声人語では必ずしも成立しているとはいえない。
- 省略の交差現象は現れていなかった (表 10 参照)。したがって、(仮説 6) は成立している。
- 最上位の節を除く各節において、省略の先行詞がその節またはその上位の節に含まれる割合は、科学雑誌と天声人語でかなりの差があった (科学雑誌で 81%、天声人語で 64%) (表 11 参照)。したがって、(仮説 7) は科学雑誌でほぼ成立しているが、天声人語では必ずしも成立しているとはいえない。

表 11 仮説 7 の成立状況

Table 11 The number of instances which satisfy (or do not satisfy) Hypothesis 7.

	出現回数	
	サイエンス	天声人語
成立	20	14
不成立	5	7

表 12 仮説 8 の成立状況

Table 12 The number of instances which satisfy (or do not satisfy) Hypothesis 8.

出現位置	出現回数	
	サイエンス	天声人語
同一文中	1	1
一つ前	29	78
二つ前	3	31
三つ前	1	11
その他	0	16

- 単文、または重文の先頭の節で、埋め込み節でないものの省略に対する先行詞は、先行詞がそれより前の文中にある割合は極めて高かった (科学雑誌で 97%、天声人語で 99%) (表 12 参照)。したがって、(仮説 8) はほぼ成立している。

5. 考 察

5.1 照応と省略の用法の違い

今回の調査結果より、以下のような、照応と省略の用法の違いが得られる。

1. 省略では先行詞が同一文中に現れる場合が最も多いのに対し (表 8 参照)、照応では直前の文中に現れる場合が最も多い (表 2 参照)。
2. 省略句は、ほぼ 9 割が「が格」であるが (表 7 参照)、照応語の先行詞は必ずしも主格とは限らない。
3. 照応は省略ほど非交差の制約は厳しくない (表 4, 表 10 参照)。
4. 複文や重文の先頭を除く節において照応や省略が使われる場合、照応の先行詞は照応語の含まれる文より前の文に (表 2 参照)、省略の先行詞は省略語の含まれる文にある (表 8 参照)。
5. 単文、または重文の最初の節に現れる照応の割合は 9 割近くとかなり高いが (表 2, 表 6 参照)、そこに現れる省略は、全省略中の 1/4 であり (表 8, 表 12 参照)、かなりの差がある。(不特定多数者を除く省略 678 回中そこに現れる省略は 171

回で 25%、また、全照応 733 回中そこに現れる照応は 622 回で 85% である.)

1 から 3 の理由の一つとしては、3 章で考察したように、省略では先行詞に関する情報が表層的には全く失われているのに対し、照応では照応語の種類により、ある程度先行詞に関する情報を持ち得るという相違点がある。

また、4 と 5 は省略が主として一つの文中で最上位を除く節間をつなぐのに使われ、照応が文と文との間、または、重文の節間をつなぐのに使われることを示している。この理由としては、同一文の上位の節中に先行詞がある時には、下位の節からは上位の節への情報のアクセスが極めて容易なので、省略できる場合が多く、照応を使うことが冗長になるからだと考えられる。

また、天声人語は科学雑誌に比べ、段落に跨る照応や省略が多い。この理由としては、天声人語では、科学雑誌に比べ、記事が短く、一つの記事全体で一つの主題を扱うことが多いため、その主題については段落間に跨る照応や省略の使用が多くなる。

次に、八つの仮説に対するそれぞれの反例をあげ、それらから得られた考察を以下に述べる。

5.2 仮説に対する反例

照応・省略に関する八つの仮説が実際の文章でどの程度成立するかが明らかになった。以下では、それぞれの仮説に対する反例を示す。(ここで、(仮説 1) の反例は (反例 1)、(仮説 2) の反例は (反例 2) であり、以下同様である。) なお、(反例 2)、(反例 6)、(反例 8) は新規に作成し、(反例 1)、(反例 3)、(反例 4)、(反例 5)、(反例 7) は日経サイエンスから引用した。

(反例 1) ……、この反応機構にはまた別の問題があった。つまり、光化学反応系は、一度に 1 電子の反応しか行なえないのである。

この問題を解決するために、光合成生物は特別な水の分解触媒を開発した……

(反例 1) において、照応語 この が使われている。その先行詞は、前の段落の「光化学反応系は、一度に 1 電子の反応しか行なえない」であり、照応が段落に跨っている。したがって、この例文は (仮説 1) に反する。

(反例 2) 太郎は花子の友達です。彼は彼女とよく映画を見に行きます。

この例文では、彼の 先行詞は「太郎」であり、彼女の

先行詞は、「花子」である。この 2 組の照応関係は交差しており、(仮説 2) に反している。

(反例 3) 光合成細菌の光合成複合体と緑色植物のそれとの間には多くの違いがある。

この例文は単文であり、それの先行詞として「光合成複合体」が考えられる。したがって、この例文は (仮説 3) に反する。

(反例 4) 水の酸化クロックの作用とその位置が明らかになってきた。

この例文も単文であり、その中で照応語 その が使われている。照応語 その の先行詞として、同一文(節)内にある「水の酸化クロックの」が考えられる。したがって、この例文は (仮説 4) に反する。

(反例 5) ……、重力波は、超新星爆発にともなうて、星の芯が中心に向かって崩壊する時に発生すると信じられている。

また、 ϕ_{1-1} 宇宙初期に形成されたかもしれない不思議な存在で、「宇宙のひも」と呼ばれるものに関連した高エネルギー現象からも発生するらしい、……

この例文において、省略語 ϕ_{1-1} が使われ、その先行詞は、前の段落にある「重力波は」である。したがって、この例文に段落を跨る省略が使われている。ゆえに、この例文は (仮説 5) に反する。

(反例 6) 猟師が山で鹿を追いかけたが、 ϕ_{1-1} もたついたので、 $\phi_{2-2}\phi_{3-2}$ 逃げられてしまった²。

この例文において、 ϕ_{1-1} 、 ϕ_{2-2} として「猟師が(は)」が、 ϕ_{3-2} として「鹿に」が考えられる。省略語 ϕ_{1-1} とその先行詞「猟師が」の間に、省略語 ϕ_{3-2} の先行詞である「鹿に」が存在するので、一つめの先行詞と省略語の組と二つめの先行詞と省略語の組との間に交差が生じている。したがって、この例文は (仮説 6) に反する。

(反例 7) 中心が +V だけずれた球と、 ϕ_{1-1} -V だけずれた¹ 球を考える必要がある。

(反例 7) では、埋め込み先(最上位の節ではない節)中に省略 ϕ_{1-1} が使われ、その先行詞「中心が」は、省略のある節でもその上位の節でもない節に出現している。この例文は (仮説 7) に反する。

(反例 8) ϕ_{1-1} 読書に熱中していた¹ので、太郎は呼ばれたことに気付かなかった。

この例文は重文であり、その最初の節に省略語が使われている。その省略語の先行詞は、後節にある「太郎は」であり、(仮説 8) に反している。

5.3 反例の生じた理由

本節では、前節で示したような、仮説に対する反例が生じる理由について考察する。

(仮説1) と (仮説5) に対する考察

(反例1) や (反例5) が生ずる理由のひとつとして、段落間のつながりの強度が関与していると考えられる。一般に、段落はその前の段落となんらかのつながりを持っており、特に強いつながりを持つ段落間においては、照応現象や省略現象が現れやすいと考えられる。

(仮説2) に対する考察

(反例2) は、日本語文として不適格な文ではない。その理由として、使われている照応語が「彼」と「彼女」という語なので、容易にそれぞれの先行詞が復元でき、文が曖昧でないからだと思われる。照応語は、指示詞、代名詞、代動詞に分けられ、さらに、素性や使われ方によって細分される。もし、照応語として、同じ品詞で同じ素性で同じ使われ方をする語が2回以上使われ、それらが交差を生じているような場合、照応語の先行詞復元は、省略語の先行詞の復元と同じ状況に陥ってしまう。つまり、照応の交差は、品詞や素性、使われ方が違い、先行詞が容易に復元できる場合に生じるのではないかと考えられる。

(仮説3), (仮説4) に対する考察

(反例3), (反例4) は、どちらも並列構造中の照応であり、どちらも先行詞は、並列構造中の対応する要素となっている。一般に、並列構造における照応に関しては、並列構造の対応関係による制約が強く作用すると考えられる^{22), 23)}。また、これらの反例は、科学雑誌 (日経サイエンス) に掲載されたもので、英文を日本語文に訳したものであり、英文原文の影響が出ている。日本語としては、以下のような、同語反復や同一要素の括り出しにより、照応語を使わない文の使用が普通である。(反例3), (反例4) の内容を表す、より一般的な文を (例文12), (例文13) として、以下に示す。

(例文12) 光合成細菌の光合成複合体と緑色植物の光合成複合体の間には、多くの違いがある。

(例文13) 水の酸化クロックの作用と位置が明らかになってきた。

(仮説6) に対する考察

(反例6) に使われている動詞 (「もたつく」「逃げる」) に対する格解析だけからは何が主語となるかは曖昧だが、事象関係に関する意味まで考慮すると、明白になる。つまり、「もたつく」の主語は「猟師」で、「逃げる」の主語は「鹿」である。この場合には、省略間の非交差よりも事象関係に関する制約が強いので、聞き手が省略格の補完に失敗することはないだろうという話し手の判断が働いていると考えられることができる。

また、(反例6) にある「もたついたので」という節は、 ϕ_{3-2} (鹿に) とともに「逃げられてしまった」という節に係る。つまり、次の (例文14) のように、省略語の位置を移動させても、意味的に変わらない。

(例文14) 猟師が山で鹿を追いかけたが、 ϕ_{1-2} (鹿に,) ϕ_{2-1} (猟師が,) もたついたので、逃げられてしまった²。

このように、日本語文において、語順はかなりの自由度を持っているので、(反例6) においても、仮説は、省略語の位置を適当に設定すると成立しているといえる。

なお、1文の中に2個以上の省略語がある文が少なかったために、(反例5) のように省略に交差が生じている文が出現しなかった可能性も排除できない。

(仮説7) に対する考察

(反例7) において、「中心が+V だけずれた」という埋め込み文と省略語 ϕ_{1-1} を含む埋め込み文「 ϕ_{1-1} -V だけずれた」は並列構造の中の対応する要素である。この場合、これらの埋め込み文の中の対応する格「中心が」は両方で共通であり、省略しても並列構造の対応関係より容易に復元可能である。

(仮説8) に対する考察

(反例6) が生じた理由としては、(仮説6) に対する反例で考察したと同様に、日本語文の語順に関する自由度の高さが考えられる。日本語としては、(仮説6) を満たす以下の語順の使用が普通である。

(例文15) 太郎は、 ϕ_{1-1} 読書に熱中していた¹ので、呼ばれたことに気付かなかった。

6. むすび

本論文では、日本語文における照応と省略の使われ方を分析するとともに、照応・省略に関する仮説を立て、その成立に関する検証を行った。これらの仮説は協調性の原則に基づき立てた仮説であるが、同時に照応や省略と段落との関連性、および、複数の照応や省略間の非交差性を考慮に入れている。これらの仮説が実際の文章で成立している程度に関して得られた結果を、

1. 計算機による日本語文章の段落分け⁹⁾

2. 照応・省略を適切に含む文の生成

等に応用することが考えられる。(仮説1)と(仮説3)は、照応・省略の先行詞が照応語や省略語と同一段落中にあると主張している。これらの仮説が実際の文章で成立している度合いは高いので、このことを段落分けのヒューリスティックスとして利用することができる。すなわち、ある文中で文に跨る照応語や省略語があれば、その先行詞を含む文と同一段落に入れる、または、先行詞の決定が困難な場合でも、一つ前の文と同一段落に入れるなどにより、段落分けに利用することができる。また、照応語や省略語の先行詞探索のヒューリスティックスとして、これらの制約を用いることができる。

参 考 文 献

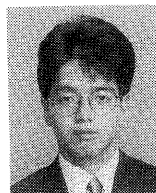
- 1) 久野 暉：談話の文法，大修館（1978）。
- 2) Grice, H. P.: *Logic and Conversation, Syntax and Semantics*, Morgan Kaufmann, pp. 41-58 (1975)。
- 3) 情報処理振興事業協会技術センター：計算機用日本語基本動詞辞書 IPAL (Basic Verbs) 解説編，辞書編（1987）。
- 4) 情報処理振興事業協会技術センター：計算機用日本語基本形容詞辞書 IPAL (Basic Adjectives) 解説編，辞書編（上，下）（1990）。
- 5) 国語学会：国語学大辞典（1991）。
- 6) 島津，内藤，野村：構造予測を用いた日本語文の意味解析法，情報処理学会論文誌，Vol. 27, No. 2, pp. 165-176 (1986)。
- 7) 山村，大西，杉江：日本語文の照応問題解決システム，電子情報通信学会論文誌，Vol. 73-D-2, No. 6, pp. 887-896 (1990)。
- 8) Zadrozny, W., Jensen, K.: *Semantics of Paragraphs, Computational Linguistics*, Vol. 17, No. 2, pp. 171-209 (1991)。
- 9) 山本，増山，内藤：手かかり語を用いた日本語文章の段落分けに関する実証的考察，情報処理学

会自然言語処理研究会研究報告，NL. 84-9 (1991)。

- 10) Sidner, C. L.: *Focusing in the comprehension of the Definite Anaphora*, Brady, M. and Berwick, R. C. (eds.), *Computational Models of Discourse*, MIT Press (1983)。
- 11) 石崎，井佐原：日本語文の複雑さの定性的・定量的特徴抽出，情報処理学会自然言語処理研究会研究報告，NL. 67-6 (1988)。
- 12) 竹木，田口，中川：質疑応答システムにおける日本語文の理解と生成，情報処理学会自然言語処理研究会研究報告，NL. 74-13 (1989)。
- 13) 平井，北橋：省略語と事象間関係から見た日本語文の意味解析と文脈解析の枠組について，電子情報通信学会自然言語とコミュニケーション研究会技術研究報告，NLC 86-8 (1986)。
- 14) 小倉，飯田，野村：英日翻訳における英文解析の照応関係処理，情報処理学会自然言語処理研究会研究報告，NL. 28-1 (1981)。
- 15) 吉本 啓：談話処理における日本語名詞の扱いについて，情報処理学会自然言語処理研究会研究報告，NL. 56-4 (1986)。
- 16) 三吉，安川，平川，向井，田中，横井：状況意味論に基づく談話理解システム DUALS，情報処理学会自然言語処理研究会研究報告，NL. 50-7 (1985)。
- 17) 池田光生：主題と焦点を用いた日本語対話文生成，情報処理学会自然言語処理研究会研究報告，NL. 78-3 (1990)。
- 18) 内藤，島津：談話意味構造の極小拡大を求めることによる照応処理について，情報処理学会自然言語処理研究会研究報告，NL. 72-4 (1989)。
- 19) 白井英俊：照応の扱いについて，電子情報通信学会技術研究報告 86-9 (1986)。
- 20) 片桐恭弘：視点表現と照応現象について，電子情報通信学会技術研究報告 89-5 (1989)。
- 21) 風斗博之：名詞句の指示と談話処理，情報処理学会自然言語処理研究会研究報告，NL. 55-3 (1986)。
- 22) 長尾，辻井，田中，石川：科学技術論文における並列句とその解析，情報処理学会自然言語処理研究会研究報告，NL. 36-4 (1983)。
- 23) 首藤，吉村，津田：日本語技術における並列構造，情報処理学会論文誌，Vol. 27, No. 2, pp. 183-190 (1986)。

(平成4年8月31日受付)

(平成5年7月8日採録)

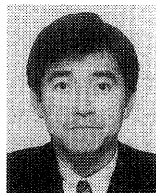


藤澤 伸二 (学生会員)

1970年生。1990年国立小山工業高等専門学校電子制御工学科卒業。1990年豊橋技術科学大学知識情報工学課程入学。現在同大学院において自然言語処理，特に，日本語文章における省略に関する研究に従事。

**増山 繁 (正会員)**

1952年生。1977年京都大学工学部数理工学科卒業。1983年同大学院博士後期課程修了。1982年日本学術振興会奨励研究員。1984年京都大学工学部数理工学科助手。1989年豊橋技術科学大学知識情報工学系講師。1990年同助教授。現在に至る。グラフ・ネットワークのアルゴリズム、組合せ最適化、並列アルゴリズム、自然言語処理等の研究に従事。工学博士。電子情報通信学会、日本オペレーションズ・リサーチ学会等会員。

**内藤 昭三 (正会員)**

1955年生。1979年京都大学工学部数理工学科修士卒業。同年 NTT 入社。基礎研究所情報科学部勤務。自然言語処理の研究に従事。計算言語学会、電子情報通信学会各会員。