

日本語文章における照応・省略現象に関する仮説の考察

5 P-1 藤澤伸二* 増山 繁* 内藤昭三**

fujisawa@toki.tutkie.tut.ac.jp, masuyama@tutkie.tut.ac.jp naito@atom.ntt.jp

*豊橋技術科学大学知識情報工学系 **NTT基礎研究所

1 はじめに

照応・省略は、日本語文章においてしばしば行なわれるもので、言わなくても聞き手にとって自明の情報を省くことによって、文の冗長度を下げることが目的とする[5]。[1]では、照応や省略が、実際の文章(日経サイエンスと天声人語)でどのように使われているかという調査結果と、協調性の原則、すなわち、書き手は読み手にとって分かりやすい文章を書くとの仮定の下に、照応・省略に関して立てた8つの仮説(制約)に対する検証結果を報告した。その結果、いずれの仮説も実際の文章でほぼ成り立っていることが明らかになった。本稿では、文献[1]で紹介した照応・省略に関する8つの仮説に対する反例をあげ、それらを通じて得られた考察を報告する。

2 照応・省略に関する仮説とその反例

照応・省略に関する8つの仮説[1]を、それぞれの反例(但し、仮説1と仮説3は除く)とともに以下に示す。なお、(反例1)、(反例2)、(反例4)は新規に作成し、(反例3)、(反例5)、(反例6)は日経サイエンスから引用した。

(仮説1) 省略は段落間を跨らない。すなわち、省略の先行詞は、省略語と同一段落中にある。

(仮説2) 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。

(反例1) 猟師が山で鹿を追いかけたが、 ϕ_1 もたついたので、 $\phi_1\phi_2$ 逃げられてしまった。

この例文において、 ϕ_1 として「猟師が(は)」が、 ϕ_2 として「鹿に」が考えられる。ひとつめの先行詞と省略語の組とふたつめの先行詞と省略語の組との間に交差が生じている。

(仮説3) 照応は段落に跨ることが少ない。

(仮説4) 照応は交差することが少ない。

(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く。

この例文において、彼の先行詞として「太郎」が、彼女の先行詞として、「花子」が考えられる。従って、この例文は照応の交差現象を含んでおり、(仮説4)に反する。

(仮説5) 最上位の節を除く各節において、省略の先行詞はその節またはその上位の節に含まれることが多い。

(反例3) 中心が+Vだけずれた球と、 ϕ -Vだけずれた球を考える必要がある。

この例文において、 ϕ として「中心が」が考えられる。 ϕ の先行詞が前節にあるので、この例文は(仮説5)に反する。

(仮説6) 単文、または重文の最上位の先頭節の埋め込み文以外の節中の省略の先行詞は、それより前の文中にある。

(反例4) ϕ 読書に熱中していたので、太郎は呼ばれたことに気付かなかった。

この例文において、 ϕ として「太郎は」が考えられる。 ϕ の先行詞が後節にあるので、この例文は(仮説6)に反する。

(仮説7) 複文や重文の先頭を除く各節において、照応の先行詞は、その節やその上位の節には含まれない。

(反例5) 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。

この例文において、それの先行詞として「光合成複合体」が考えられる。従って、この例文は(仮説7)に反する。

(仮説8) 単文、または重文の先頭の節において、その埋め込み文以外の照応の先行詞は、それより前の文(いくつかの文と段落を含む)にある。

(反例6) 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。

この例文において、そのの先行詞として「水の酸化クロックの」が考えられる。従って、この例文は(仮説8)に反する。

3 考察

8つの仮説に対するそれぞれの考察を以下に述べる。

(仮説1)と(仮説3)に対する考察

文献[1]において、検証した結果、これらの仮説は、ほぼ成立することが分かったが、一部、成り立たない場合もある。その理由としては、段落は、その前の段落とある程度のつながりを持っており、強いつながりを持つ段落間においては、省略現象や照応現象が現れるためと思われる。実際の文章で、省略や照応が段落を越える場合、つまり、仮説に反する場合は、2つの段落は強く結合していた。しかも、照応が段落に跨る場合のほとんどは、段落頭に照応語が来ていた。また、文献[1]とは別に、英文において、段落は、照応の先行詞復元の最初の妥当な範囲であると文献[4]で独立に提唱されている。

(仮説2)に対する考察

(仮説2)に反する(反例1)に使われている動詞は、構文解析だけでは何が主語となるかは曖昧だが、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。

また、(反例1)にある「もたついたので」という節は、 ϕ_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1)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1) 猟師が山で鹿を追いかけたが、 ϕ_2 (鹿に、) ϕ_1 (猟師が、)もたついたので、逃げられてしまった。

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1)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1)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4)に対する考察

(反例2)は(仮説4)に反するが、日本語文として、おかしい文ではない。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、代動詞に分かれ、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と思われる。

(仮説5)に対する考察

(仮説5)に反する(反例3)を見ると、「中心が+Vだけずれた」という埋め込み文と省略語 ϕ を含む埋め込み文「 ϕ -Vだけずれた」は構造が非常に似ている。そのため、「中心が」という語がなくてもそこに省略語があることを容易に知ることができる。かえって、そこに「中心が」という語が存在すると、冗長な表現となってしまう。

(仮説6)に対する考察

(仮説6)に反する(反例4)が生じた理由は、(仮説2)に対する考察のところでも述べたように、日本語文は語順に関してかなりの自由度を持っているからだと思われる。(仮説6)を満たすように語順を変えた(例文2)を以下に示す。

(例文2) 太郎は、読書に熱中していたので、呼ばれたことに気付かなかった。

(仮説7)に対する考察

(仮説7)を満たさない(反例5)の原文は、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る。そのため、(反例5)は翻訳文独特の印象を与え、かつ、「光合成複合体」を強調しているように思われる。(反例5)の内容を表す、より日本語らしい文に書き換えたものを(例文3)として、以下に示す。

(例文3) 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。

(仮説8)に対する考察

(仮説8)に反する(反例6)は、「水の酸化クロック」を強調しているという感じがする。(反例5)の内容を表す、より一般的な文を(例文4)として、以下に示す。

(例文4) 水の酸化クロックの作用と位置が明らかになってきた。

4 今後の課題

今後の課題として、本研究で得られた知見を

1. 計算機による日本語文章の段落分け [2]
2. 照応語や省略語の先行詞復元

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ことを示している。これらの仮説が実際の文章でほぼ成立しているの、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して、実際の文章を段落わけするのに応用できることを示している。

また、8つの仮説は照応や省略語の先行詞の出現位置を限定している。そのため、照応語や省略語の先行詞復元に、これらの考察を用いることができる。

なお、(仮説2)、(仮説4)に対しては、例えば、文献[3]で提唱されているのと類似のスタック構造による談話処理メカニズムを想定することで、よりうまく説明できる可能性があり、現在考察中である。

参考文献

- [1] 藤澤、増山、内藤：日本語文章における照応・省略現象の基本的検討、情報処理学会、NL.86-6(1991)
- [2] 山本、増山、内藤：手が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、NL.84-9(1991)
- [3] B. J. Grosz, and C. L. Sidner, Attention, Intentions, and the Structure of Discourse, Computational Linguistics, Vol.12, No.3(1986)
- [4] W. Zadrozny, K. Jensen, Semantics of Paragraphs, Computational Linguistics, Vol.17, No.2(1991)
- [5] 久野：談話の文法、大修館(1978)