

「コ」「ソ」系指示語の用法に関する仮説とその検証

2K-4

三輪 倫子† 山本 和英† 増山 繁† 内藤 昭三†

{miwa,yamamoto,masuyama}@smlab.tutkie.tut.ac.jp † naito@slab.ntt.jp†

豊橋技術科学大学† NTTソフトウェア研究所†

1 はじめに

自然言語の研究分野として談話解析に関する研究が盛んに行なわれている。本稿では談話構造を明らかにするために必要不可欠な指示語の指示対象同定に着目し、それを自動的に行なうための基礎として二つの仮説を立て、実際の日本語文章を用いてその検証を行なった。日本語の指示語の体系が「コソア」を共通項として持つことはよく知られているが、今回の仮説は特に科学雑誌における「コ」系「ソ」系の指示語の用法の違いに着目して立てている。「ア」系指示語に関しては、科学雑誌での出現は稀であるため、本稿では考察対象から外した。

2 対象とする指示語

表1に本稿で対象とした指示語を示す。

	コ系	ソ系
代名詞 (単数)	これ	それ
代名詞 (複数)	これら	それら
代名詞 (場所)	ここ	そこ
代名詞 (方角)	こちら	そちら
代名詞 (方角)	こっち	そっち
連体詞	この	その
形容動詞	こんな	そんな
副詞	こう	そう

指示語の指示形式の分類方法は文献によって様々であるが、今回は大きく文脈指示と非文脈指示の二つに分類した。それぞれの定義を以下に示す。

- 文脈指示とはテキスト中に先行詞が存在する場合である。
- 非文脈指示とは指示対象がテキスト中に先行詞として存在しない場合である。

ここで、先行詞とは「指示語と同一の指示対象を指示する言語形式（単語、句、節、文、段落）、または指示対象を結果として与える言語形式」を言う。これは文献[1]での定義を拡張したもので、このようにすることにより指示対象が言語形式として直接テキスト中に現れない以下のような場合も文脈指示の範疇とできる。（同様の説明を文献[1]では談話記憶によって行なっているが、本稿ではテキストとの関連を重視し、指示語と先行詞と

の関連により説明した。）

(例) 卵と牛乳1カップを小麦粉に入れてかき混ぜよ。それを熱したフライパンに注げ。[1]

本研究の目的は、文章内に出現した指示語の指示範囲を検出するアルゴリズム設計のための基礎データを得ることであり、指示対象が文脈中に先行詞として存在する場合を対象としている。よって、以降は文脈指示を行なう「コ」「ソ」系指示語について考察を行なう。

3 仮説とその検証

3.1 仮説

今回立てた仮説は以下の2つである。

仮説1 文中の先行文脈指示を行なう「コ」系の指示語の先行詞は、その指示語と同一の文内に存在することはない。

仮説2 文中の先行文脈指示を行なう「ソ」系の指示語は、その先行詞が同一文内に存在する。

ここで、“文中”とは、“文頭以外、あるいは文頭の接続詞や副詞の直後以外の場所”のことをいう。

これらの仮説は国語学の方面からすでに研究されている「コ」系「ソ」系の二種の指示語に関する深層的な用法の違い[1, 2, 3]をもとに立てたものである。それぞれの仮説を導いた指示語の深層的な用法の違いを以下に示す。

「コ」系の性質 (仮説1)

「コ」系指示語による指示対象は書き手、読み手の双方にとって、その指示対象の正体が明確になっている。よって、叙述の完了しない段階にあるものは指示できない[1, 2]。仮説1は「句点が一つの叙述の完了」とみなした上で立てた。

「ソ」系の性質 (仮説2)

「ソ」系指示語は叙述の完了しない段階にあるものでも指示できる。また、「コ」系指示語が指示対象をイメージ化して書き手、読み手の意識の中心に指示対象を“持ってくる”働きがあるのに対し、「ソ」系指示語は指示対象をその位置に据え

Hypotheses on the Usage of 'ko-' and 'so-' demonstratives and its Verification.

Tomoko MIWA†, Kazuhide YAMAMOTO†, Shigeru MASUYAMA†, Shozo NAITO†

†Toyohashi University of Technology, †NTT Software Laboratories

置いたまま、その対象を“指示する”働きがある(文献[2, 3]より考察)。“指示する”という行為は“持ってくる”という行為よりも指示対象物の存在可能な領域に強く制限を与えるため、「ソ」系指示語の指示対象は指示語の近くに存在すると考える。

3.2 調査

今回の調査は日経サイエンス(1994年1月号～1994年5月号)の10記事について人間の視察により行なった。調査結果を表2に示す。

この調査結果によれば、先行詞同一文内存在割合(ISS/MID)は「コ」系指示語で18.3%であるのに対し、「ソ」系指示語は77.4%であった。また、各記事については「ソ」系で6割以上、「コ」系で4割以下が常に保たれていた。つまり、両仮説ともつねに成立するとは限らないが、科学雑誌の文章における指示語の使用傾向を捕らえていると考えられる。

表2: 先行文脈指示における先行詞同一文内存在数

記事	こ			そ		
	NET	MID	ISS	NET	MID	ISS
1	42	18	3	31	23	14
2	49	14	5	22	14	14
3	70	22	4	44	35	27
4	43	12	1	53	35	23
5	70	26	6	37	24	21
6	69	27	2	58	40	35
7	76	29	2	70	47	37
8	54	19	4	48	27	19
9	42	19	6	25	15	12
10	61	22	5	22	14	10
合計	576	208	38	410	274	212

NET: 先行文脈指示語総出現数

MID: 先行文脈指示語文中出現数

ISS: 先行詞同一文内存在数

4 反例に対する考察

仮説1の反例文(「コ」系指示語の先行詞が指示語と同一文内に出現する)の構造を調べると次のように分類できた。

- 先行詞の存在する句が接続助詞「が」で終る。(17文)
- 先行詞の存在する句が用言の連用形で終る。(13文)
- 並列を表す助詞「と」の直後に指示語がある。(3文)
- 先行詞の存在する句が接続助詞「し」「ので」で終る。(2文)
- 先行詞と指示語が同格関係にある。(1文)

• その他(2文)

このように、反例文の多くは「～が」または用言の連用形で終る従属句¹を持つ構造のものである。これらを取り入れることにより仮説1をより成立度の高い精密な仮説にすることができると考えられる。

仮説2は「ソ」系指示語の「叙述の完了していない段階のものも指示できる」という性質のみに着目して立てたものであるが、「ソ」系指示語の性質は文献[1, 2, 3]にもその他いくつか示されており、それらを考慮に入れる必要がある。特に今回調査した記事の反例文(文中に現れた「ソ」系指示語による先行詞が前文、またはそれ以前にある)に現れた深層的特徴としては次のような点が上げられる。これらの表層への現れを検討する必要がある。

- 指示対象が仮定や空想に基づいた叙述内容である。
- 指示対象がその文において重要度の低いものである。(著者の主観によるところが大きい)
- 前文の内容に関する例示を行なう文で用いる。(例:「～はその典型である」)

5 おわりに

科学雑誌における指示語の使用に関する仮説を立て、その検証を行なった。また、仮説に対する反例文の考察から、より精密な仮説にするための有力な手がかりを得ることができた。今後は仮説を精密化するとともに、これら反例の生じる場合の表層的確定方法を見つけ、照応処理アルゴリズムに組み込める実用的な仮説にする予定である。また、今回の検証を通して、「文脈指示と非文脈指示の判別が機械的に行なえない」「ソ系指示語のなかで接続詞との区別が表層的につけられないものがある」といった、今後計算機を使って処理させる際に障壁となるであろう問題点も発見したので、これらの解決も新たな課題として検討していく。

参考文献

- [1] 吉本 啓: 日本語の指示詞コソアの体系, 日本語研究資料集 指示詞, ひつじ書房, pp.105-122 (1992).
- [2] 正保 勇: 「コソア」の体系, 日本語の指示詞, 国立国語研究所, pp.53-119 (1981).
- [3] 林 四郎: 「代名詞が指すもの, その指し方」, 運用I, 朝倉日本語新講座 5, pp.1-45 (1983).
- [4] 南 不二男: 現代日本語の構造, 大修館書店, pp.114-131 (1974).

¹接続助詞で終るものか、あるいは用言(述語部分となっている用言)の連用形で終るもの[4].