

語用論的条件の解釈に基づく日本語ゼロ代名詞の指示対象同定

堂 坂 浩 二^{†,*}

本稿では、日本語対話において待遇表現、受給表現、特定の文末形式を発話環境の下で適切に使用するための語用論的条件を解釈することにより、対話登場人物を指示するゼロ代名詞の指示対象を同定するモデルを提案する。このモデルでは、発話環境として、対話登場人物間の待遇関係、話者の視点、情報のなわばりに関わる言語外情報を利用する。発話環境は、環境のあり方を規定する制約の集合として表される。語用論的条件の解釈は、環境に含まれる制約を用いて語用論的条件を充足させることによって行う。この解釈プロセスによって、文に含まれるゼロ代名詞の指示対象が同定され、また、新たな制約が環境の中に導入されることもある。新たに導入された制約は、後続する文の解釈に役立てられる。このモデルは、対話ドメインに特有な制約を与えることにより、対話ドメインにおける行為やイベントの因果関係に関する膨大な量の知識に頼ることなく、対話登場人物を指示するゼロ代名詞の指示対象を同定できるという利点をもつ。

Identifying the Referents of Japanese Zero Pronouns based on Pragmatic Condition Interpretation

KOJI DOHSAKA^{†,*}

A model is proposed to identify the referents of Japanese zero pronouns referring to persons by interpreting the pragmatic conditions of the appropriate usage of honorific expressions, donatory verbs and specific types of sentence-final grammatical forms in their circumstances. This model exploits circumstances pertaining to social relationships between dialogue participants, the speaker's point of view and the territory of information. The circumstances of utterance are represented as constraints which determine what the circumstances are like. Interpreting pragmatic conditions is regarded as the process of satisfying them using the constraints constituting the circumstances. Through the interpretation process, the referents of zero pronouns in a sentence can be identified, and there are cases where new constraints are introduced into circumstances. The new constraints serve to interpret the subsequent utterances. Moreover, this model has the advantage of identifying zero pronoun referents using constraints characteristic of dialogue domains without resort to a great deal of knowledge about actions and causal relationships in dialogue domains.

1. はじめに

1.1 日本語ゼロ代名詞と発話環境

自然言語文によるコミュニケーションにおいては、話者は、聴者に対して伝達しようとする情報のすべてを言語表現上に明示する必要はない。これは、聴者が、文が使用された際の発話環境を利用することによって、言語表現として明示されていない情報を補完することが期待できるためである。発話環境とは、文が使用された際に、話者と聴者が共有している言語外情報のことをいう。日本語対話においては、伝達すべき

情報の一部が、頻繁にゼロ代名詞化される⁷⁾。特に、対話登場人物を指示する表現は、ゼロ化されることが普通であり、ゼロ代名詞の指示対象は、それが使用された発話環境に依存して決定される。

計算言語学においては、ゼロ代名詞をはじめとする参照表現の適切な使用を説明するためのモデルの構築が、重要な課題として数えられてきた。また、対話翻訳などの対話を処理しようとする試みにおいても、参照表現を扱う手法の開発が、重要な課題として認識されている⁸⁾。こういった参照表現への関心の高まりを反映して、発話環境を利用して参照表現の指示対象を同定するさまざまなモデルが提案されてきた。本稿では、まず、続く第1.2節で、日本語ゼロ代名詞をはじめとする参照表現を扱うことのできる従来の発話解釈モデルの問題点を述べ、第1.3節で、その問題点を解決できる方策が存在することを示す。

[†] (株)ATR 自動翻訳電話研究所
ATR Interpreting Telephony Research Laboratories

* 現在、NTT 基礎研究所
Presently with NTT Basic Research Laboratories

1.2 従来の発話解釈モデルの問題点

言語表現と発話環境との関係を説明する一般的な枠組としては、Grosz らが提案した言語構造、意図構造、注視状態の3つの構成要素から成る談話構造理論がある⁹⁾。また、さらに詳細なモデルとして、対話参加者のもつプラン（意図構造）に基づいて省略表現の解釈を行うモデル²⁾、¹³⁾やセンター（注視状態）の継続性に基づいて日本語ゼロ代名詞の解釈を行うモデルがある⁹⁾、¹⁷⁾。

しかし、対話参加者のもつプランの認識に基づくモデルには、対話ドメインにおける行為やイベント間の因果関係についての膨大な知識の存在を前提としているという難点がある。また、センターの継続性に基づくモデルは一見簡易に実現できそうであるが、センターが継続する部分対話を切り出すためには、対話参加者のもつプランを認識する必要がある。実際、注視状態の継続性と意図構造との間には相互関係があることが、Grosz などによって論じられている⁹⁾。例えば、次の対話 (d1) を見てみよう。

対話 (d1)

- (s1.1) 質問者: 会議に参加したいのですが。
 (s1.2) 事務局: 登録用紙に必要事項を記入する必要があります。
 (s1.3) 事務局: もうお持ちでしょうか?
 (s1.4) 質問者: 持っていません。
 (s1.5) 事務局: それではお送りします。

この対話は、会議への参加申し込みに関する質問者と事務局との間の対話である。今、文 (s1.5) について注目する。この文では、「送る」の行為者、受益者、対象がゼロ代名詞化されている。これらのゼロ代名詞の指示対象は、それぞれ事務局、質問者、登録用紙であるような個体である。このうち、行為者に対応するゼロ代名詞の指示対象である事務局は、文 (s1.5) 以前の発話内では参照されることのなかった個体であるから、センターの継続性に基づくモデルを適用してもこのゼロ代名詞を扱うことはできない。もちろん、文 (s1.5) から新たな部分対話が始まっているとして、センターの継続性に基づく手法を文 (s1.5) に適用することは、手法の適用範囲を越えていると言うこともできるが、その場合でも部分対話の切り出しには、対話参加者のもつプランを認識する必要がある。プランの認識に基づいて発話を解釈するモデルを適用すれば、文 (s1.5) に含まれるすべてのゼロ代名詞の指示対象を同定することができるが、そのために必要とな

る行為やイベントの因果関係に関する知識は膨大な量になることが予想される。

1.3 語用論的条件の解釈に基づくモデル

上で見た従来のモデルで使われた発話環境が、日本語ゼロ代名詞の使用に関わる発話環境のすべてを尽くしているわけではない。日本語対話においては、待遇表現¹⁵⁾や受給表現¹²⁾が頻繁に使用され、また、同じ命題内容に言及するためにさまざまな文末形式を用いることができる。これらの言語表現が適切に使用されるためには、それぞれ対話参加者間の待遇関係¹⁵⁾、話者の視点¹²⁾、情報のなわばり¹⁰⁾に関する発話環境においてある条件が成立していなければならない。こういった言語表現の使用の適切性を保証する条件を語用論的条件と呼ぶ。また、発話環境は、対話参加者間の待遇関係などに関するどのような事実が発話環境の中で成立すべきかを規定した制約の集まりであると考えられる。次に例示するように、こういった発話環境の下で文の使用に際する語用論的条件を解釈することにより、対話登場人物を指示するゼロ代名詞の指示対象が同定できることがある。

例えば、第1.2節で示した対話 (d1) の文 (s1.5) では、待遇表現「お送りする」が使用されている。これは、いわゆる謙譲表現であって、その使用に際する語用論的条件は、『話者（事務局）は、「送る」の行為者より受益者を上位に待遇する』というものである。発話環境に含まれる制約としては、『話者は話者自身より聴者を上位に待遇する』といったものを考えることができる。この発話環境の下で、先の語用論的条件を解釈することにより、ゼロ化された行為者は話者の事務局であり、受益者は聴者の質問者であると同定できる。第1.2節で述べたセンターの継続性に基づくモデルのみを適用した場合の問題点は、上に述べた語用論的条件の解釈によって「送る」の行為者は事務局と同定できるから、回避することができる。また、待遇関係などの特定の語用論的条件を解釈する際に用いる制約を記述することは、対話ドメインの行為やイベントの因果関係に関する知識を記述するよりも、負担が少ないことが予想される。

そこで、本稿では、日本語対話において、待遇表現、受給表現、特定の文末形式をもつ文を使用する際に課せられる語用論的条件を解釈することによって、対話登場人物を指示するゼロ代名詞の指示対象を同定するためのモデルを示す。また、特に、モデルを適用する対話ドメインとして、会議への参加申し込みに関する

対話⁷⁾を取り上げる。まず、次の第2章で、このモデルの概観を示す。

2. モデルの概観

ここでは、語用論的条件の解釈に基づいて日本語ゼロ代名詞の指示対象を同定するモデルの概観を一般的な発話解釈の枠組の中で位置付けながら示していく。

発話内容と発話環境との間の相互関係を捉える一般的な理論としては、状況理論¹⁾がある。状況理論では、発話の情報内容と発話環境との相互関係は、情報内容の間の制約という考え方をを用いて捉えられる。実際に、そのような考え方に基づいて、対話参加者がもつ共有意図という発話環境と発話内行為¹⁶⁾の内容との間の制約を用いて、発話内行為の内容を解釈する枠組がある⁴⁾。発話内行為とは、文を使用する際に、言語慣習によって遂行される依頼や約束といった行為であり、共有意図とは、対話参加者が互いに協調しながらあるイベントを成立させようとする意図をいう。また、橋田は、状況理論においても利用された制約という考え方をさらに追求し、言語にまつわる情報の部分性を扱うために、依存伝搬という制約プログラミングの方法を提示している⁶⁾。このように、多様な発話環境と発話内容との相互関係をモデル化するために、状況理論の枠組や制約という考え方は有望である。Grosz らの談話構造の理論⁵⁾は、発話解釈のための一般的な枠組と見なせるが、本稿で利用する待遇関係などに関する発話環境をその中でどのように扱えばよいかという事は明らかでない。

そこで、本稿では、状況理論が提供する枠組に基づいて、モデルを構築していく。特に、発話環境を構成する個々の情報を表現するために状況理論における事態を用い、さらに、発話環境と発話内容との相互関係を捉えるために制約という考え方をを用いる。

一般的な発話解釈の枠組の中におけるゼロ代名詞の指示対象同定モデルの概観は図1のようなになる。図では、対話(d1)の文(s1.5)を解釈する際のモデルの様子を、形式的な記述法を用いずに示してある。モデルの形式的な記述は、第3章以降で行う。また、本稿で扱う語用論的条件の解釈に基づくモデルは、図中で曲線によって囲

まれている部分である。

まず、文からは、その文の使用によって遂行される発話内行為とともに、その文を適切に使用するための語用論的条件が抽出される。本稿では、発話内行為や語用論的条件を文の語彙・統語論の特徴から取り出す解析器¹¹⁾の存在を仮定する。

図中では、文(s1.5)から、その発話内行為の情報内容として『*zero1が*zero2に*zero3を送る』ことを約束する』という表現が取り出されている。*zero1, *zero2, *zero3は、それぞれゼロ代名詞化された「送る」の行為者、受益者、対象に対応している。また、第1.3節で議論したように、発話内行為の内容の抽出と同時に、謙譲表現「お送りする」の使用から、語用論的条件として、『話し手*sp1は、「送る」の受益者*zero2を行為者*zero1より上位に待遇する』という内容をもつ表現が取り出される。

環境は、制約の集合としてモデル化される。本稿で用いる制約としては、発話や話者や聴者は誰かといった事実、対話参加者間の待遇関係や話者の視点、情報のなわばりに関する制約がある。また、本稿では扱わないが、一般の発話解釈モデルでは、対話参加者が共に認識している共有意図⁴⁾やプランといった発話環境も考えることができる。そういった発話環境は、発話内行為の内容を直接解釈するために重要である。

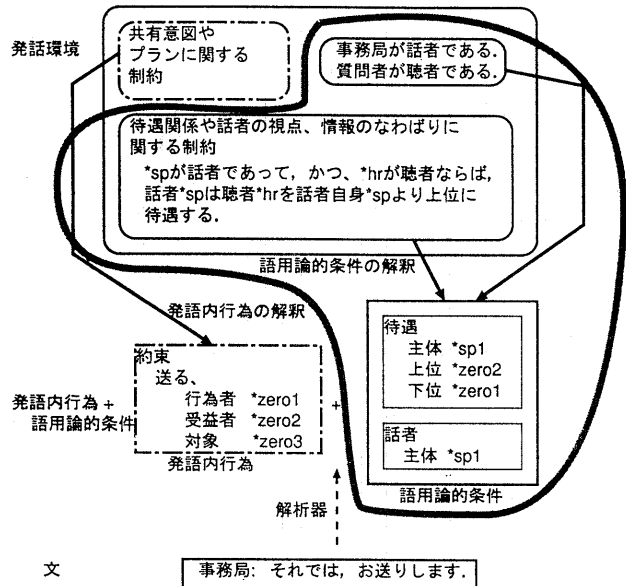


図1 ゼロ代名詞同定モデルの概観
Fig. 1 An overview of a model to identify the referents of Japanese zero pronouns.

文から取り出された語用論的条件は、上に述べた環境に含まれる制約を用いて解釈される。第1.3節で議論したように、この解釈によって、ゼロ代名詞の指示対象が同定される。図中では、語用論的条件に含まれる *zero1 と *zero2 の指示対象がそれぞれ話者の事務局、聴者の質問者と同定され、結果として、発語内行為中の *zero1 と *zero2 の指示対象が定まる。このモデルでは、*zero3 に対応するゼロ代名詞の指示対象は同定できない。

以下において、第3章で、モデルの形式的な記述の枠組を説明し、第4章で、言語表現と語用論的条件の関係について述べ、第5章で、環境を構成する制約を示す。第6章では、語用論的条件の解釈プロセスを提案し、続く第7章で、その解釈プロセスの振舞いを対話例を用いて示す。第8章で、モデルの評価を行い、最後に、第9章で、モデルの利点と今後の課題を述べて結ぶ。

3. 情報の記述の枠組

このモデルでは、語用論的条件や環境を構成する制約を表現する際に、状況理論における事態を用いる。事態は次の形をしている。

$$\langle\langle R, o_1, \dots, o_n; p \rangle\rangle. \quad (1)$$

ただし、式(1)で、 R は関係、 o_1, \dots, o_n は対象の列、 p は極性を表す。式(1)は、極性が1のとき、対象 o_1, \dots, o_n の上に関係 R が成立することを示し、極性が0のときその関係が成立しないことを示す。関係 R の引数 o_i ($i=1, \dots, n$) には、定数かパラメタが割り当てられる。パラメタは不定項を表し、「*」で始まる記号として書く。解釈プロセスの中でいまだ同定されていないゼロ代名詞の指示対象は、パラメタとして書かれる。このモデルで用いる事態を次に示す。

$$\langle\langle \text{Speaker}, x; 1 \rangle\rangle: \quad (2)$$

人物 x は話者である。

$$\langle\langle \text{Hearer}, x; 1 \rangle\rangle: \quad (3)$$

人物 x は聴者である。

$$\langle\langle \text{Person}, x; 1 \rangle\rangle: \quad (4)$$

対象 x は人物である。

$$\langle\langle \text{Respect}, x_1, x_2, x_3; 1 \rangle\rangle: \quad (5)$$

人物 x_1 は人物 x_2 を人物 x_3 より上位に待遇する。

$$\langle\langle \text{Social-rel}, x_1, x_2; 1 \rangle\rangle: \quad (6)$$

人物 x_1 は人物 x_2 より上位に待遇される。

$$\langle\langle \text{Attrib}, x_1, x_2; 1 \rangle\rangle: \quad (7)$$

人物 x_1 は人物 x_2 に帰属する。

$$\langle\langle \text{Empathy}, x_1, x_2, x_3; 1 \rangle\rangle: \quad (8)$$

人物 x_1 は人物 x_3 より人物 x_2 寄りの視点をとる。

$$\langle\langle \text{Territory}, x, \text{infor}; 1 \rangle\rangle: \quad (9)$$

事態 infor は人物 x のなわばりに入る。

$$\langle\langle \text{Agent}, x, \text{infor}; 1 \rangle\rangle: \quad (10)$$

人物 x は行為 infor の行為者である。

$$\langle\langle \text{Neq}, x_1, x_2; 1 \rangle\rangle: \quad (11)$$

x と y は異なる人物である。

上の事態の中で、(5)と(6)は、対話登場人物間の待遇関係を示す。(5)は、ある対話参加者 x_1 から見たときに、人物 x_2 が人物 x_3 より上位に待遇されていることを示すために用いる。一方、(6)は、対話参加者間で共通に認識されている待遇関係を表すために用いる。

事態(7)において、人物 x_1 が人物 x_2 に帰属するとは、 x_1 が x_2 の身近な人物であることを意味する。例えば、ある人物にとって、その妻や子はその人物に帰属する。こういった帰属関係は、待遇関係や情報のなわばりに関する制約を定義する際に重要である。

事態(8)は、よく知られた話者の視点(共感度)¹²⁾を表すために用いる。事態(9)で使われた情報のなわばり¹⁰⁾という概念については、第4.3節で説明する。

最後に、事態(10)と(11)はモデルの中で特別な扱いを受ける。まず、事態(10)に関しては、モデルは、行為 infor の行為者に対応する引数に割り当てられた対象が、実際に、関係 Agent の第一引数に割り当てられている対象と一致するかどうかを判定する手続きをもっていると仮定する。また、(11)の形の事態が現れるときには、必ず、両引数に具体的な引数が割り当てられており、モデルは両引数に割り当てられた対象が異なるかどうかを判定できると仮定する。

4. 言語表現の使用と語用論的条件

ここでは、待遇表現、受給表現、特定の文末形式の使用と語用論的条件との関係について述べる。語用論的条件は、パラメタを含んだ事態として表現される。

4.1 待遇表現の使用と待遇関係

日本語対話においては、待遇表現が頻繁に用いられる。待遇表現としては、動詞句の尊敬型(例:「お～になる」,「お～なさる」)や謙譲型(例:「お～する」,「お～致す」)などがある。これらの待遇表現の使用に際しては、以下に示すような語用論的条件が課せられ

る。

(p1) 動詞句の尊敬型の使用に際しては、話者は主語の対象を非主語の対象より上位に待遇する。

(p2) 動詞句の謙譲型の使用に際しては、話者は非主語の対象を主語の対象より上位に待遇する。

例えば、対話 (d1) の文 (s1.5) では、謙譲表現「お送りします」が使用されているので、(p2) より、次の語用論的条件が抽出される。ただし、「送る」のゼロ代名詞化された行為者と受益者を表すパラメタをそれぞれ $*zero_1$ と $*zero_2$ で表す。

《Speaker, $*sp_1$; 1》. (12)

《Respect, $*sp_1$, $*zero_2$, $*zero_1$; 1》. (13)

すなわち、話者 $*sp_1$ は人物 $*zero_2$ を $*zero_1$ より上位に待遇していなければならない。

4.2 受給表現の使用と話者の視点

日本語対話において受給表現の使用は、話者の対話登場人物に対する視点 (共感度) と関係があることが知られている¹²⁾。例えば、受給表現「～てもらう」の使用に際しては、次の話者の視点に関する語用論的条件が課せられる。

(p3) 「～てもらう」の使用に際しては、話者は間接目的語の対象より主語の対象寄りの視点をとる。

話者の視点に関する語用論的条件は、第3章で示した表現 (8) を用いて記述する。

4.3 文末形式と情報のなわばり

情報のなわばり¹⁰⁾とは、発話内で言及された事態全体を話者がどのように眺めているかということを考察することから神尾によって考案された概念である。神尾は、ある情報が話者のなわばりに属するか否かを決定する条件として、次のものをあげている¹⁰⁾。

- (a) 情報が話者自身の行動予定などに関するものであること。
- (b) 話者が直接体験して得た情報であること。
- (c) 話者および話者の特に身近な人物に関する情報であること。
- (d) 話者の専門領域に関する情報であること。

さらに、神尾は、話者のなわばりに入る情報のみが直接形によって言及されるという一般化を行っている。本稿では、文末が断定的な確言 (例:「お話になります」) で終わっている場合、その文末は直接形であると考え、また、文末が伝聞・推量の形 (例:「～そうですが」) で終わっている場合、その文末は間接形であるとする。そこで、本稿では、特定の文末形式の使用には次の語用論的条件が課せられるとする。

(p4) 直接形の文末形式を用いて言及された事態は、話者のなわばりに入る。

(p5) 間接形の文末形式を用いて言及された事態は、話者のなわばりに入らない。

例えば、次の対話 (d2) について見てみよう。

対話 (d2)

(s2.1) 質問者: 田中先生を会議にお招きするという
ことだそうですが。

(s2.2) 事務局: はい。

(s2.3) 事務局: 基調講演でお話しになります。

上の文 (s2.1) では、「田中先生を招く」という事態が「～ということだそうですが」という間接形を用いて言及されているので、(p5) より、その事態は話者のなわばりに入らないという語用論的条件が課せられる。また、文 (s2.3) では、直接形「お話しになります」が使用されているので、(p4) より、「基調講演でお話しになります」という事態は話者のなわばりに入るという語用論的条件が抽出される。これらの語用論的条件は、第3章の(9)の形の事態で記述される。

5. 発話環境

発話環境は制約の集合として表される。ここでは、制約の記述形式と、会議参加申し込みという対話ドメインにおいて有効な制約を示す。

$$R_1(\vec{*x}_1) \wedge \dots \wedge R_{n-1}(\vec{*x}_{n-1}) \Rightarrow R_n(\vec{*x}_n). \quad (14)$$

上の式 (14) で、 $\vec{*x}_i$ ($i=1, \dots, n$) はパラメタの列である。 $R_i(\vec{*x}_i)$ ($i=1, \dots, n$) は、パラメタ化された事態である。上の制約は、左辺の条件 $R_i(\vec{*x}_i)$ ($i=1, \dots, n-1$) がすべて充足するなら、右辺の条件 $R_n(\vec{*x}_n)$ も充足することを示している。

制約 (14) で、左辺の条件が必要ない場合は、制約は次の形で書かれる。

$$R(\vec{*x}). \quad (15)$$

上の形の制約 (15) が、もしパラメタを含まないなら、その制約は事実と呼ばれる。

会議参加申し込みに関する対話ドメインにおいて有効な制約を次に示す。

《Speaker, $*sp$; 1》 \wedge (16)

《Hearer, $*hr$; 1》 \wedge

《Person, $*less$; 1》 \wedge

《Person, $*more$; 1》 \wedge

《Neq, $*less$, $*more$; 1》 \wedge

《Attrib, $*less$, $*sp$; 1》 \wedge

《Attrib, $*more$, $*hr$; 1》

- ⇒<<Respect, *sp, *more, *less; 1>>. (17)
- <<Speaker, *sp; 1>> (17)
- <<Person, *less; 1>>∧
- <<Person, *more; 1>>∧
- <<Neq, *less, *more; 1>>∧
- <<Social-rel, *more, *less; 1>>
- ⇒<<Respect, *sp, *more, *less; 1>>.
- <<Speaker, *sp; 1>>∧ (18)
- <<Hearer, *hr; 1>>∧
- <<Person, *less; 1>>∧
- <<Person, *more; 1>>∧
- <<Neq, *less, *more; 1>>∧
- <<Attrib, *more, *sp; 1>>∧
- <<Attrib, *less, *hr; 1>>
- ⇒<<Empathy, *sp, *more, *less; 1>>.
- <<Attrib, *x, *x; 1>>. (19)
- <<Agent, *agent, *soa; 1>> (20)
- ⇒<<Territory, *agent, *soa; 1>>.
- <<Speaker, *sp; 1>>∧ (21)
- <<Hearer, *hr; 1>>∧
- <<Agent, *hr, *infor; 1>>
- ⇒<<Territory, *sp, *infor; 0>>.
- <<Attrib, *agent 1, *agent 2; 1>>∧ (22)
- <<Neq, *agent 1, *agent 2; 1>>∧
- <<Territory, *agent 1, *soa; 1>>
- ⇒<<Territory, *agent 2, *soa; 1>>.

制約(16)は、話者は聴者に帰属する人物を話者に帰属する人物より上位に待遇することを示す。制約(17)は、人物 *more が人物 *less より上位に待遇されることが既知なら、話者 *sp は *more を *less より上位に待遇することを示している。制約(19)からは、質問者は質問者自身に、事務局は事務局自身に帰属することがわかる。

制約(18)は、話者は聴者に帰属する人物よりも話者に帰属する人物寄りの視点をとることを示している¹²⁾。

制約(20), (21), (22)は、情報のなわばりに関する制約であり、第4.3節で示した神尾の提案(a)~(d)をこの対話ドメインに適合するように近似したものである。まず、制約(20)は、(a)の近似であり、事態 *soa の行為者が人物 *agent なら、その事態は、その人物の行動予定を示すものであり、その人物の情報のなわばりに入ると考えている。また、制約(21)は、聴者が行為者であるような事態は、話者のなわばりには入らな

いことを示している。これは、神尾が与えている条件(a)をさらに強めたものになっている。さらに、制約(22)は(c)に対応しており、ある人物 *agent 2 のなわばりは、その人物に帰属する人物 *agent 1 のなわばり内の情報を含むことを示している。なお、条件(b)や(d)に対応する制約は与えない。そういった制約を与えるためには、対話ドメインを構成する行為やイベントに関する知識が必要とされると予測されるからである。

また、各発話ごとに、話者と聴者が誰であるかという事態が環境に含まれているとする。すなわち、次に示す事態のうち、(23)と(24)か、(25)と(26)かのどちらか2つの事態が環境に含まれる。

<<Speaker, 事務局; 1>>. (23)

<<Hearer, 質問者; 1>>. (24)

<<Speaker, 質問者; 1>>. (25)

<<Hearer, 事務局; 1>>. (26)

さらに、次の事実が環境に含まれているとする。

<<Attrib, 事務局, 質問者; 0>>. (27)

<<Attrib, 質問者, 事務局; 0>>. (28)

<<Person, 質問者; 1>>. (29)

<<Person, 事務局; 1>>. (30)

すなわち、この対話ドメインでは、事務局と質問者が互いに他に帰属することはないことを仮定する。また、事務局を人物として扱うことも仮定されている。

6. 語用論的条件の解釈

本章では、文から抽出された語用論的条件を制約を用いて解釈するプロセスについて説明する。以下において、まず、第6.1節で事態の単一化について説明した後で、第6.2節で、解釈プロセスの手順を示す。さらに、第6.3節で、文から抽出された語用論的条件の解釈によってゼロ代名詞の指示対象が一意に定まらないときの方策について述べる。

6.1 単一化

パラメタを含まない2つの事態が与えられるとき、各事態の関係名と極性が同じで、引数位置に対応する対象が同じなら、その2つの事態は単一化可能である。単一化した結果の事態は、元の2つの事態に含まれるすべての引数位置に対応する対象をもつ。また、極性のみが異なる2つの事態は互いに矛盾するという。

次に、パラメタを含む事態の間の単一化について説明する。パラメタへの対象の割当てをアンカーと呼

ぶ。パラメタを含む2つの事態が与えられるとき、各事態に含まれるパラメタに対象を割り当てることによって、2つの事態が単一化可能になるような最小のアンカーを見つける。そのようなアンカー F が存在するならば、パラメタを含む2つの事態はアンカー F の下に単一化可能であるという。

6.2 解釈プロセス

本節では、文から抽出された語用論的条件を環境の下に解釈するプロセスについて述べる。語用論的条件には、ゼロ代名詞の指示対象を表すパラメタが含まれている。文の使用が適切であるためには、語用論的条件が、パラメタに指示対象を割り当てるようなアンカーの下に、環境を構成する制約の集合から論理的に導かれるものであればよい。このためには、線形導出法¹⁴⁾を適用して、与えられた語用論的条件が制約の集合から導出できることを判定する手続きを実行すればよいと考えられる。しかし、文の使用に際する語用論的条件のすべてが、その時点の環境を構成する制約から導き出せるものではなく、文によって新たな制約が環境に導入されることもありうる。そういった場合を扱うためには、環境に含まれていない制約の存在を仮定しながら解釈を進めていく必要がある。仮定された制約は、解釈が終わった後環境に導入され、後続する文の解釈に役立てられる。

今、文から抽出された語用論的条件の列 \vec{C} が与えられたとする。さらに、発話環境を構成する制約の集合 Σ 、 \vec{C} 内の語用論的条件に含まれるすべてのパラメタに対するアンカーを F 、解釈の結果新たに導入される新制約の集合を Δ とする。語用論的条件の解釈は、次の手順を繰り返すことを行う。

- (1) \vec{C} の先頭から語用論的条件 C_1 を取り出す。
- (2) If C_1 がパラメタを含まず、かつ、
 C_1 と矛盾する事実が Σ から導ける。
 Then この解釈プロセスは失敗する。
- (3) Else If C_1 がパラメタを含まず、かつ、
 C_1 と矛盾する事実が Σ から導けない。
 Then C_1 を \vec{C} から消去し、
 C_1 を Δ に付け加える。
- (4) Else If 制約 $(R_1 \wedge \dots \wedge R_{n-1} \Rightarrow R_n)$ が Σ に含まれ、かつ、 C_1 と R_n がアンカー G の下に単一化可能。
 Then \vec{C} から C_1 を消去し、
 R_1, \dots, R_{n-1} を \vec{C} に付け加え、
 F に G の割当て情報を付け加える。

(5) Else この解釈プロセスは失敗する。

(4)は、制約によって語用論的条件が書き換えられるプロセスであり、一般の線形導出法と同等のものである。(4)で適用される制約が、左辺の条件をもたない式(15)の形の場合は、 \vec{C} には何も付け加えない。このとき、 C_1 は充足されたといわれる。(2)と(3)において、 C_1 と極性だけが異なるような事実が環境から導き出せるかどうか判定することは、一般の線形導出法に基づいて行う。(5)は、 C_1 にパラメタが含まれ、適用可能な制約が存在しない場合で、このとき解釈プロセスは失敗する。

上の手順で、(3)の場合が、一般の線形導出法の拡張であり、新制約を仮定することによって推論を進め、仮定された制約は解釈プロセス全体が成功したとき環境に付け加えられる。新たに環境に導入される制約は、パラメタを含まない事実型の制約のみである。

解釈プロセスは、上の手順を繰り返すことによる。適用可能な制約が複数ある場合には、解釈プロセスを分岐させる。なお、アンカー F の初期値は空であり、新制約の集合 Δ の初期値は空集合である。解釈プロセスは列 \vec{C} が空列となるなら停止する。

6.3 解の選択

解釈プロセスによって、ゼロ代名詞の指示対象は、文から抽出された語用論的条件に含まれるパラメタに対象を割り当てるアンカー F によって与えられる。すべての解釈プロセスが失敗したなら、その文は解釈不能であって、その文に含まれるゼロ代名詞は同定不能とする。さらに、成功に終わる複数の解釈プロセスがあるときは、次の基準に基づいて解釈を選択する。

- (a) できるだけ多くのパラメタに対象を割り当てることのできる解釈を優先する。
- (b) 新制約の集合がより小さい解釈を優先する。

(b)は、発話を解釈するために新たに導入される仮説をできるだけ少なくするという方策であり、一種の節約原則¹⁵⁾と見なすことができる。この基準によって、解釈が一意に定まるなら、新制約の集合 Δ に含まれる新たな事実を環境 Σ に付け加え、続く文の解釈に進む。また、この基準によっても解釈を一意に定められない場合には、文内のゼロ代名詞の指示対象は不定となり、この場合への対処はこのモデルの責任ではない。

7. ゼロ代名詞同定例

本章では、今までに述べたモデルに従って、日本語ゼロ代名詞がどのように同定されるかを対話 (d2) の

文 (s 2.1) と (s 2.3) とを用いて説明する。

まず、文 (s 2.1) において、ゼロ代名詞化された「招く」の行為者の同定を行う。このゼロ代名詞の適切な指示対象は事務局である。この文が発話されるとき、環境には話者と聴者が誰かを示す事実(25)と(26)が含まれている。また、次の語用論的条件(31)~(34)が抽出される。なお、文 (s 2.1) の「招く」のゼロ代名詞化された行為者をパラメタ *zero₄ で表す。

《Speaker, *sp₂; 1》。 (31)

《Person, 田中; 1》。 (32)

《Respect, *sp₂, 田中, *zero₄; 1》。 (33)

《Territory, *sp₂, () (34)

《招く, *zero₄, 田中, 会議; 1》; 0》。

語用論的条件(以下、条件と略す)(33)は、(p 2)より、謙讓表現「お招きする」の使用から抽出される。条件(34)は、第4.3節で議論したように、文 (s 2.1) では、間接形「~ということだそうですか」が使用されていることから抽出される。

まず、環境に含まれる事実(25)によって、条件(31)は充足し、パラメタ *sp₂ の値は質問者となる。また、条件(32)と矛盾する事実は環境から導出できないから、条件(32)は新制約の集合に付け加えられる。

さらに、パラメタ *sp₂ の値が決まったことから、条件(33)は、次の(35)に書き換えられている。

《Respect, 質問者, 田中, *zero₄; 1》。 (35)

条件(35)には制約(16)と(17)が適用できる。まず、制約(16)が適用されるとすると、結局、次の条件が導き出される(話者の聴者が誰かという条件の充足に関しては、以下省略する)。

《Person, *zero₄; 1》 (36)

《Person, 田中; 1》 (37)

《Neg, *zero₄, 田中; 1》 (38)

《Attrib, *zero₄, 質問者; 1》。 (39)

《Attrib, 田中, 事務局; 1》。 (40)

このうち、上の条件(36)は、事実(29)か(30)と単一化可能である。まず、事実(29)と単一化されたとすると、パラメタ *zero₄ は質問者となる。このとき、条件(38)は充足され、条件(39)は次のように書き換えられている。

《Attrib, 質問者, 質問者; 1》。 (41)

条件(41)は制約(19)によって充足する。また、他の条件(37)や(40)に関して、これらと矛盾する事実は導けない。

一方、条件(36)が事実(30)と単一化されたときのことを考えてみる。このとき、パラメタ *zero₄ の値は事務局となり、条件(39)は、次のように書き換えられている。

《Attrib, 事務局, 質問者; 1》。 (42)

上の条件(42)は、事実(27)と矛盾する。

結局、条件(35)に制約(16)を適用した結果、パラメタ *zero₄ の値は質問者でなければならないことがわかった。

次に、条件(34)の充足を考える。今、パラメタ *zero₄ が質問者と同定されているので、条件(34)は次のように書き換えられている。

《Territory, 質問者, () (43)

《招く, 質問者, 田中, 会議; 1》; 0》。

しかし、この条件(43)と矛盾する事実が、制約(20)を用いて導けるので、条件(35)に制約(16)を適用した結果生じた解釈プロセスは失敗する。しかし、条件(35)には制約(17)も適用できる。

今、条件(35)に制約(17)を適用すると、次の条件が得られる。

《Person, *zero₄; 1》 (44)

《Person, 田中; 1》 (45)

《Neg, *zero₄, 田中; 1》 (46)

《Social-rel, 田中, *zero₄; 1》。 (47)

上の条件(44)は、事実(29)か(30)と単一化可能である。まず、事実(29)と単一化されたとすると、パラメタ *zero₄ は質問者となる。しかし、パラメタ *zero₄ を質問者と同定して解釈プロセスを進めても、すでに議論したように、条件(34)が環境と矛盾してしまう。よって、条件(44)が事実(30)と単一化した場合のみを考えればよい。このとき、パラメタ *zero₄ は事務局と同定され、条件(46)は充足され、条件(47)は次のように書き換えられている。

《Social-rel, 田中, 事務局; 1》。 (48)

条件(48)と矛盾する事実は環境から導けないから、この条件は新制約の集合に付け加えられる。このとき、条件(34)は、次のように書き換えられている。

《Territory, 質問者, () (49)

《招く, 事務局, 田中, 会議; 1》; 0》。

条件(49)と矛盾する事実は環境から導き出せないのので、この条件は新制約の集合に付け加えられる。なお、この条件は、制約(21)を使えば充足させることができるが、このモデルではパラメタを含まない条件で環境に矛盾しないものは新制約として見なされる。

結局、文 (s 2.1) から抽出された語用論的条件を解釈することによって、パラメタ *zero₄ の値が適切に事務局と同定された。さらに、新たな事実(32)と(48)、(49)が環境に導入される。

次に、文 (s 2.3) のゼロ代名詞化された「お話しになる」の行為者の同定について考える。このゼロ代名詞の適切な指示対象は、文 (s 2.1) で導入された人物「田中」である。この文が発話される時、環境には、事実(23)と(24)が含まれている。また、次の語用論的条件が抽出される。ただし、「お話しになる」のゼロ化された行為者をパラメタ *zero₅ で表す。

《Speaker, *sp₃; 1》. (50)

《Respect, *sp₃, *zero₅, *sp₃; 1》. (51)

《Territory, *sp₃, (52)

《話す, *zero₅, 基調講演; 1》; 1》.

上の条件(51)は尊敬表現「お話しする」の使用から抽出される。また、条件(52)は、第4.3節で議論したように、「*zero₅が基調講演で話す」という事態が「お話しになる」という直接形によって言及されていることから抽出される。

ここで、条件(50)は、事実(23)によって充足され、パラメタ *sp₃ の値は事務局となる。このとき、条件(51)は、次のように書き換えられている。

《Respect, 事務局, *zero₅, 事務局; 1》. (53)

条件(53)には、制約(16)と(17)が適用できる。今、制約(16)を適用したとする。この制約適用によって、パラメタ *zero₅ の値は、聴者の質問者に帰属する人物であることがわかり、さらに、新制約を導入しない限り、それは質問者であると同定される。このとき、条件(52)は次のように書き換えられている。

《Territory, 事務局, (54)

《話す, 質問者, 基調講演; 1》; 1》.

しかし、条件(54)と矛盾する事実が、制約(21)を用いて導出されてしまう。そこで、条件(53)に制約(17)を適用する。さらに、文 (s 2.1) で導入された新事実(48)を用いることにより、パラメタ *zero₅ の値は田中となる。これはゼロ代名詞の適切な指示対象を与える。このとき、条件(52)は、次のように書き換えられている。

《Territory, 事務局, (55)

《話す, 田中, 基調講演; 1》; 1》.

上の条件(55)はパラメタを含まず、かつ、この条件と矛盾する事実は導けないので、新制約の集合に付け

加えられる。

結局、文 (s 2.3) から抽出された語用論的条件を解釈することにより、「お話しする」の行為者は「田中」であると適切に同定された。また、第4.3節の(d)で述べたように、神尾¹⁰⁾によれば、話者の専門領域に属する情報は話者のなわばりに入る。誰かが基調講演で話すという事態は事務局の専門領域に属することであるから、環境に導入される新たな事実(55)は妥当なものであるといえる。

8. モデルの評価

本モデルの有効性を評価するために、ATR 自動翻訳電話研究所で収集された「国際会議への参加申し込みに関する問い合わせ」についての対話⁷⁾を用いて、評価実験を行った。

なお、本モデルは、計算機上で実際に実現され、ATR 自動翻訳電話研究所で開発された対話翻訳実験システム NADINE の解析モジュールに組み込まれているが⁸⁾、モデル評価は、その解析モジュールに直接表層の文を与えて行うのではなく、語用論的条件を表層の語彙・統語論的特徴から取り出す解析器¹¹⁾によって、各文に付随する語用論的条件が抽出できたと仮定して、その語用論的条件にこのモデルを適用することによって行った。解析モジュールに組み込まれたものを直接用いないのは、ゼロ代名詞の処理とは直接関係のない解析誤りと、このモデル自身の能力によるゼロ代名詞の処理の失敗との混同を避けるためである。

この評価に用いた対話は、主にキーボード対話⁷⁾によって収集された10個の対話である。表1には、それぞれの対話(対話1から対話10)に含まれる対話文の個数、ゼロ代名詞の個数、対話登場人物を指示するゼロ代名詞の個数、このモデルによって指示対象同定可能なゼロ代名詞の個数を示してある。また、ここでは、日本語対話文において、表層でゼロ化されている構成要素のうち、日英対話文翻訳の際に補完される必要があるもの(主語、目的語、間接目的語など)をゼロ代名詞と認定した。

結果として、225個の対話文に含まれる対話登場人物を指示する132個のゼロ代名詞のうち、108個(約82%)が適切に指示対象を同定できることがわかった。また、対話全体に現れたゼロ代名詞のうち、25個(157個-132個)のゼロ代名詞は、対話登場人物以外の対象を指示するもので、これらはモデルの適用範囲外である。

モデルがゼロ代名詞の指示対象を同定できなかったのは、待遇表現などの特定の言語表現が使用されていなかった場合と、対話ドメインのイベントや行為に関する知識がなければ扱いきれないゼロ代名詞が現れた場合である。前者の場合の典型例は、次の対話 (d 3) であり、後者の典型例は、対話 (d 4) である。

対話 (d 3)

(s 3.1) 質問者: ホテルの手配もして頂けるのですか?

(s 3.2) 事務局: 京都ホテルは予約できます。

対話 (d 4)

(s 4.1) 質問者: 会議の間に市内観光があるそうですが。

(s 4.2) 事務局: はい, 8月5日の午後に, 清水寺, 金閣寺などを見学します。

まず, 対話 (d 3) では, 文 (s 3.2) の「予約できる」の行為者がゼロ化されている。しかし, 本モデルは, その指示対象を同定することはできない。なぜなら, この文には, 待遇表現などの特定の表現が用いられていないからである。ただし, 「予約できる」という表現のかわりに, 「ご予約することができます」といった待遇表現が使われていたとしたら, ゼロ代名詞を適切に処理できる。

次に, 対話 (d 4) では, 文 (s 4.2) の「見学する」の行為者がゼロ化されている。しかし, このゼロ代名詞の指示対象は, 対話の中にすでに導入されている対話参加者ではなく, 文 (s 4.1) で導入された「市内観光」というイベントから想起されるであろう「市内観光の参加者」という主体であると考えられる。このゼロ代名詞を扱うためには, そういった対話ドメインのイベントに関する知識が必要となる。その種の知識を用いない本モデルでは, このゼロ代名詞の指示対象について誤った予測をしてしまう。すなわち, このモデルでは, 「見学します」という直接形で言及された事態は話者のなわばりに入るとしており, かつ, 行為者が話者であるような行為は, 話者のなわばりに入っている。したがって, 「見学する」の行為者は事務局であると同定してしまう。なお, モデルが誤った指示対象を同定してしまったのはこの一例のみであった。

結局, このモデルでは, 対話ドメインに特有のイベ

表 1 ゼロ代名詞指示対象同定モデルの評価

Table 1 An evaluation of the model of identifying the referents of zero pronouns.

対話番号	対話文の個数	ゼロ代名詞の個数	人物を指示するゼロ代名詞の個数	同定可能なゼロ代名詞の個数
対話 1	20	12	12	11
対話 2	21	13	11	10
対話 3	16	7	5	5
対話 4	22	20	16	16
対話 5	22	20	18	16
対話 6	23	18	10	7
対話 7	21	12	10	10
対話 8	25	25	22	11
対話 9	17	8	8	5
対話 10	38	22	20	17
総計	225	157	132	108(108/132=0.82)

ントや行為に関する知識を用いていないにもかかわらず, 約 82% の対話登場人物を指示するゼロ代名詞の指示対象を同定できた。これは, かなり良好な結果であるといつてよいであろう。もちろん, 対話登場人物以外の対象を指示するゼロ代名詞や, 対話ドメインに特有のイベントや行為に関する知識を用いなければ処理できないゼロ代名詞を扱うためには, プランやセンターの継続性に基づく他の手法が必要となる。本稿の主張は, 本モデルがそういった既存の手法に置き換わるものであるということではなく, 第 1.2 節で述べたように, 既存の手法が依存していた対話ドメインのイベントや行為の因果関係に関する膨大な知識に頼らずとも, 待遇表現などの表層表現の使用に関する語用論的条件を用いることによって, ある範囲のゼロ代名詞を扱うことができるという主張である。

9. おわりに

本稿では, 日本語対話において, 待遇表現, 受給表現, 特定の文末形式の使用に際して課せられる語用論的条件を発話環境の下に解釈することにより, 対話登場人物を指示するゼロ代名詞の指示対象を同定するモデルを示した。このモデルは以下の利点をもつ。

- (1) 対話ドメインの行為やイベントの因果関係に関する膨大な量の知識に頼らなくても, 対話ドメインにおいて有効な待遇関係, 話者の視点, 情報のなわばりに関する制約を与えることにより, 対話登場人物を指示するゼロ代名詞の指示対象を同定できる。
- (2) 文の解釈により新たな制約を環境に導入し, そ

れを続く文の解釈に役立てることができる。

- (3) ここで用いられる語用論的条件は、既存の解析器¹⁾を使って表層の言語表現から直接抽出できるものである。

最後に、今後の課題としては、以下のことがある。

第一に、このモデルは、対話登場人物以外を指示するゼロ代名詞や、待遇表現などの特定の言語表現が使われていない発話には適用できない。そういった場合に対処するためには、本モデルで扱わなかった発話環境も利用できるように、このモデルを第2章で触れた発話解釈モデル全体へ統合していく必要がある。

第二に、このモデルでは、文の解釈は第6.3節で述べた戦略に従って一意に決まり、その解釈の結果が続く文の解釈によって覆されるようなことはないとしている。しかし、対話の中では、ある文の解釈のあいまいさが、後続の文の解釈との整合性によって解消される場合がある。こういったことを扱うためには、ある文の解釈があいまいであっても、それを一意に定めないうで、解釈プロセスを文間を越えて分岐させ、後続の文の解釈によって、先行する文の解釈を絞り込めるように、解釈プロセスを拡張することが必要である。

第三に、このモデルでは、ゼロ代名詞は単独の個体を指示するとして扱った。しかし、対話 (d2) の文 (s2.1) では、質問者と事務局の双方が田中先生を会議に招くという読みも存在する。すなわち、質問者と事務局の双方が田中先生を上位に待遇するという解釈である。この解釈を扱うためには、ゼロ代名詞が個体の複合体も指示することができるとし、対話参加者の双方が第三者を上位に待遇することを許すように、待遇関係に関する制約を精密化する必要がある。

参 考 文 献

- 1) Barwise, J.: *The Situation in Logic*, CSLI Lecture Notes Number 17, CSLI Publications, Stanford (1989).
- 2) Carberry, S.: A Pragmatics-based Approach to Understanding Intersentential Ellipsis, *Proceedings of the 23rd Annual Meeting for Computational Linguistics*, pp. 188-197 (1985).
- 3) Crain, S. and Steedman, M.: On Not Being Led up the Garden Path, *Natural Language Parsing*, Dowty, D.R., Karttunen, L. and Zwicky, A.M. (eds.), Cambridge University Press, pp. 320-358 (1985).
- 4) 堂坂浩二: 対話参加者の心的状態に関する制約に基づく発話解釈モデル, コンピュータソフトウェア, Vol. 4, No. 4, 日本ソフトウェア科学会, pp. 42-54 (1989).
- 5) Grosz, B. J. and Sidner, C. L.: Attention, Intentions, and the Structure of Discourse, *Computational Linguistics*, Vol. 12, pp. 175-204 (1986).
- 6) 橋田浩一: 制約と言語, コンピュータソフトウェア, Vol. 4, No. 4, 日本ソフトウェア科学会, pp. 16-29 (1989).
- 7) 飯田 仁: 自然言語対話の言語運用特性と対話処理の研究課題, 人工知能学会誌, Vol. 3, No. 4, pp. 445-452 (1988).
- 8) 飯田 仁: 対話翻訳と高度自然言語処理, 人工知能学会誌, Vol. 6, No. 3, pp. 328-337 (1991).
- 9) Kameyama, M.: A Property-sharing Constraint in Centering, *Proceedings of the 24th Annual Meetings for Computational Linguistics*, pp. 200-206 (1986).
- 10) 神尾昭雄: 談話における視点, 日本語学, Vol. 4, No. 12, pp. 10-21 (1985).
- 11) Kogure, K.: Parsing Japanese Spoken Sentences Based on HPSG, *Proceedings of International Workshop on Parsing Technologies*, pp. 132-141 (1989).
- 12) Kuno, S. and Kaburaki, E.: Empathy and Syntax, *Linguistic Inquiry*, Vol. 8, No. 4, pp. 627-672 (1977).
- 13) Litman, D. J.: Understanding Plan Ellipsis, *Proceedings of the 4th National Conference on Artificial Intelligence*, pp. 619-624 (1986).
- 14) Loveland, D. W.: A Unifying View of Some Linear Herbrand Procedures, *Journal of the Association for Computing Machinery*, Vol. 9, pp. 366-384 (1972).
- 15) 水谷静夫: 待遇表現の仕組み, 運用 I, 林四郎, 荻野綱男, 田中幸子, 樺島忠夫(編), 朝倉書店, pp. 158-178 (1983).
- 16) Searle, J.: *Expression and Meaning*, Cambridge University Press (1979).
- 17) Walker, M., Iida, M. and Cote, S.: Centering in Japanese Discourse, *Proceedings of 13th International Conference on Computational Linguistics* (supplement), pp. 10-15 (1990).
(平成4年12月17日採録)
(平成6年1月13日採録)

堂坂 浩二 (正会員)

1984年大阪大学基礎工学部卒業。

1986年同大学院基礎工学研究科博士前期課程修了。同年NTT入社。

1988年から4年間ATR自動翻訳

電話研究所に出向。現在NTT基

礎研究所。自然言語処理, 対話理解, 語用論の研究に従事。人工知能学会, 日本ソフトウェア科学会, 日本認知科学会各会員。

