

条件表現の語用論的性質によるゼロ主語同定

松尾 衛 龍野 弘幸 森 辰則 中川 裕志

横浜国立大学 工学部

{mamoru,tatsuno,mori}@forest.dnj.ynu.ac.jp, nakagawa@naklab.dnj.ynu.ac.jp

1 はじめに

今日、あらゆる製品にマニュアルが附属しているがこれらの機器には複雑な操作手順を要するものも多い。これを曖昧性なく記述することがマニュアルには求められている。また、海外向けのマニュアルのために複雑な手順を適切に翻訳することも困難である。そこで、我々はこのような問題の解決の基礎となる、マニュアル文を計算機で理解する手法について検討している。

一般に表層表現から直接得られる情報には曖昧性が含まれる。特に日本語の場合は主語が頻繁に省略されるためゼロ代名詞の適切な指示対象を同定することが必要になる。本研究では個別の領域知識にほとんど依存しない言語表現自身が持つ語用論的制約を用いて、言語表現から知識表現に射像する過程で曖昧性を減らしている。本稿ではゼロ代名詞の指示対象同定問題に対して、条件表現の性質を利用することを提案し、その手法の計算機処理での有効性を示す。

本稿で考える「動作主」は動作の主格になりうるものだけではなく、状態などの対象も含まれる。そこで、以下では[仁田93]のいう「主(ぬし)」とする。‘主’はより広い分類であり「対象に変化を与える主体」、「知覚、認知、思考などの主体」、「事象発生の起因的な引き起こし手」、「発生物、現象」、「属性、性質の持ち主」を含む。

2 条件表現の主に関する制約

まず、条件表現の主に関する制約の概要を説明する。なお、詳細は[森95]を参照して頂きたい。

2.1 マニュアル文における基本的制約

マニュアルを構成する最も基本的なオブジェクトとその言語的役割は、話し手がメーカー(マニュアルライター)、聞き手が利用者、第三者がシステムの全体もしくはその一部ということになる。これらを考慮するとマニュアル文中での人称は一人称がメーカー、二人称がユーザ、三人称がシステムとなる。

なお、以下の考察ではマニュアル文の主要部である操作手順の説明に現れる場合を考える。この、操作手順中の説明ではメーカーの動作はすでに完了していると考えることができる。

2.2 「と」文の主制約

[久野73]や[益岡93]による接続助詞「と」の分析から「と」文の後件には事実叙述のみが後件に現れると考えられる。つぎに、意志用法の動詞が後件で使われる場合を考える。後件には事実叙述のみが現れるため依頼、勧誘表現は現れない。そのため、動詞の「

形」が主に用いられる。「る形」で主が聞き手の場合は実質的に依頼表現となるため「と」文の後件になり得ない。また「る形」で主が話し手の場合は意志を表すが、「と」文の後件では意志を表すことができないためこの場合も後件になり得ない。主が第三者の場合は、「る形」は依頼、意志を表さないため「と」文の後件に現れることができる。以上の考察と人称に関する制約より、第三者であるシステムが後件の主となる。ただし、無意志用法の場合は「る形」が意志、命令、依頼等を表さないでこの限りではない。このように条件表現では主の意志的動作か否かが重要となる。またマニュアル文中では動作以外にも状態記述があり、これらは主の意志とは無関係である。そのため以下では意志的な動作を表すものとして「意志動作を表す動詞」、「依頼表現」、「～てみる」、「～ておく」を「動作」と表し、それ以外の「無意志動作を表す動詞(「ある」、「なる」¹⁾)」、「可能表現」、「否定」、「～ている」、「コピュラ」、「形容詞、形容動詞」を「非動作」と表す。この分類のもとで「と」の制約は次のようになる。

「と」の後件の主制約 接続助詞「と」による複文構造において後件の述部が動作で非過去の場合には、その主は三人称つまりシステムになる。

実際のマニュアル文でこの制約を検証した結果制約の妥当性が確認された。

2.3 「れば」文、「たら」文、「なら」文の主制約

[益岡93]によると

「れば」 時間を越えて成り立つ普遍的因果関係を表す。また、状態表現は仮定的な表現になりやすい

「たら」 1) 時間の経過にともなって実現することが予想される事態を表すものと 2) 実現するかどうかが定かではないような事態が実現したことを仮定し、それにともなってどのような事態が実現するかを表現するものがある

「なら」 後件に表現の重点があり、前件を真と仮定した想定のもとで、後件で判断や態度の表明が行われる

のようにそれらの特徴をまとめられる。「れば」は普遍的因果性を表すため後件は前件の発生にともなって必然的に生じる結果であるから原理的には話し手の態度が介入する余地がない。そのため基本的には後件に

¹「なる」には意志的動作を表すと考えられるものもあったが判断しにくいので無意志で統一した

話し手の態度が介入する余地がない。これは依頼表現は現れないことを意味する。しかし前件が状態の時は仮定的となるとあるので依頼表現が現れる可能性がある。

「たら」、「なら」については仮定的であり、仮定すること事態がすでに話し手の態度の一種であるため当然後件にも話し手の態度が現れる。よって話し手の態度である依頼が現れる可能性が高い。

以上のように「れば」、「たら」、「なら」では後件で依頼を使用することができるため「と」と同じ制約を導くことはできない。そこで、これらの使い分けを調べるためにマニュアル文から「れば」、「たら」、「なら」を含む文を集めてこれらの分類を行ったところ以下のようなデフォルト規則が導かれた。

「れば」文の後件の主制約 接続助詞「れば」による複文構造において後件は利用者の意志的動作を表さない。つまり、後件の述部が動作の場合には、その主は三人称つまりシステムになる。

「たら」「なら」の後件の主制約 接続助詞「たら」、「なら」による複文構造において、後件は利用者の動作しか表さない。つまり、後件の主は利用者である。

このように“「と」、「れば」”と“「たら」、「なら」”が相補的な分布を示していることがわかった。

3 動詞の意志性のデフォルト

前節で述べた制約のうち、「と」文および「れば」文の後件の主に関する制約を適用するためには 後件の動詞の意志性を判断する手段が必要になる。本研究ではこの意志性に関する情報を情報処理振興協会が開発された IPAL 動詞辞書を使用することにより得ている。[IPA87]では、動詞は 主によって意図的に行う動作を表す意志動詞と、主による意図的な動作を表していない無意志動詞とに分けることができるとしており、辞書項目にこの分類に関する記述がある。無意志動詞は 1,2 タイプに分類されており、意志動詞については、意志用法のみがある 3b タイプと無意志用法もあるに 3a タイプとに分類されている。この分類により動詞の意志性に関する分類を行うことができる。

しかし、IPAL 動詞辞書のこの分類法をそのまま利用すると、3a タイプの動詞については 意志用法なのか無意志用法なのかわからない という問題がある。また、IPAL の動詞辞書には多義性が存在するという問題もある。例えば「戻る」という動詞の場合は「戻る」の持つ意味により表 1 のように 7 通りに分類されていて、それぞれの意志性も異なっている。そのため

- (1) 「取消キーを押すと、文書作成画面に戻ります。」
[富士 88]

の「戻る」の主を同定する場合でも、「戻る」のみで 7 通りの解釈ができ、主もそれぞれの解釈で同定される。「戻る」には無意志用法の意味も存在するため、主は定まらなると解釈されるものも存在するのである。以上のような問題のため IPAL 動詞辞書をそのまま利用するだけでは 計算機で「と」文と「れば」文の後件の主を一意に決定することはできないのである。

表 1: IPAL 動詞辞書による「戻る」の分類

| 意味 | 意志性 |
|------------------------|-----|
| 手を離れた所有者が再び所有者のものとなる | 2 |
| ある目的で外出していた者が元いた所へ移動する | 3a |
| 元の所にいるようにする | 3b |
| それまで進んできた方向と逆の方向に進む | 3b |
| 元の状態に復帰する | 2 |
| 変化をみせていた事柄が元の状態になる | 2 |
| 一旦中断していた事を行うようにする | 3b |

この問題に対処するためマニュアル文中に使用されている動詞がどのように使用されているかを確認した。まず、マニュアル文中の動詞のみの表現について人間の判断により 動作と非動作に分類し、それらについて IPAL 動詞辞書での分類がどのようになっているかを調べてみた。ただし、「巻きもどす」のような複合動詞の場合は後の動詞「もどす」の分類で検討を行っている。その結果、次の表 2 で示すように 3a,3b タイプの意味を持つ動詞、すなわち、意志用法が 1 つでもあるような動詞は ほとんどの場合 動作側に分類しても問題がないということが分かった²。このよう

表 2: IPAL 動詞辞書の意志性と動作非動作の関係

| | 動作 | 非動作 |
|-------------|----------------|---------------|
| 3a,3b タイプあり | 255 例 78.0% | 72 例 22.0% |

に、意志用法が 1 つでもあるような動詞については 8 割り近くが動作側であると判断されていた。この結果より「意志用法を持つ動詞は動作として扱う。」というデフォルトを設定することにする。

4 制約適用システムの構成

ここで、今回使用した制約適用システムについて説明を行う。この、システムは条件接続を用いた文に、2 節で示した制約を適用することを目的に作成したシステムであり、図 1 のような構成となっている。

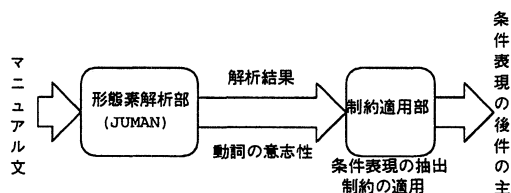


図 1: システムの構成

²ただし、「なる」に関しては無条件で非動作にするとしているため、この分類には含まれていない

4.1 形態素解析部 (JUMAN)

形態素解析システムは京都大学・奈良先端科学技術大学院大学で開発され一般に公開されている日本語形態素解析システム JUMAN を用いている。JUMANからは各形態素の意味情報も得ることができ、形態素解析と同時に動詞の意志性に関する情報を得られる。マニュアル文はまず、この JUMAN により、形態素解析が行われ、その結果得られた形態素と動詞辞書の意味情報が次の制約適用部に送られる。

4.2 制約適用部

この、制約適用部では JUMAN の形態素解析結果をもとに、条件表現が使用されている文の後件の主を決定し主の出力を行う部分である。

条件表現が含まれているかどうかということは形態素解析結果の各形態素をチェックすれば調べることができる。もし、条件表現が含まれている場合は次にシステムは制約が適用される述語を選んでくる。現在、このシステムでは制約が適用される述語として、最も文末に近い動詞および形容詞、コンピュータがその文末表現を含めて選ばれるようになっている³。この選ばれた述語に対して、「たら」文と「なら」文の場合はそのまま、「と」文と「れば」文の場合は述語の品詞、3節で述べた手法による動詞の意志性の判断および述語の文末表現により動作と非動作が判断され制約が適用され主が決められる。なお、このシステムでは条件表現の語用論的制約しか適用される制約がないため、制約もデフォルト規則も同等に扱っている。

5 制約の適用

2節で述べた「と」、「れば」、「たら」、「なら」の制約が計算機でどの程度利用できるのかということを確認するために、制約適用システムに実際のマニュアル文を入力して条件表現を含む文の後件の主を同定させてみた。今回、マニュアル文中の条件表現の制約を確認するために用いたのは、ビデオデッキ、家庭用ゲーム機、ポケットベル、日本語形態素解析システム JUMAN、構文解析システム SAX のマニュアルである。なお、これらのマニュアルは条件表現の持つ制約を導く際に利用したマニュアルとは別のマニュアルである。これらのマニュアルから 175 文の「と」文、41 文の「れば」文、11 文の「たら」文、5 文の「なら」文がシステムによって抽出された。

5.1 「と」文の場合

まずはじめに述べておくが、システムによって「と」文が 175 文抽出されたのであるがその中に次の 2 の文のように格助詞の「と」が 18 文含まれていた。

- (2) 「ただし、列の先頭と末尾には、それぞれ、「文頭」および「文末」と呼ばれるダミーの形態素があると仮定する。」 [松本]

本来、このような文は排除するにしなければならないのであるがこの 2 のような文の場合は「と」を接

³そのため、現在では条件接続の後件が更に複文になっているような場合等に、かかり方によっては正しく制約が適用される述語が選ばれない場合もある

続助詞の「と」と考えるような解釈もできるので現状では排除することができない。そのため、以下に示す解析結果はこのような文も含めた結果で示している。

175 文のうち後件の述語が動作と判断されて「と」文の制約が適用されたものは 99 文存在した。この制約適用によってシステムと推定された述語の主が実際にシステムなのかどうかを人間の判断により確認した結果を次の表 3 に示す。ただし、表に成功とあるものは制約が適用された述語の主が実際にシステムだったので、失敗とあるものは人間の判断による述語の主がユーザだったものである。

表 3: 「と」文の制約適用結果

| | 成功 | 失敗 | 合計 |
|----------|---------------|---------------|--------------|
| 全体 | 83 例 83.8% | 16 例 16.2% | 99 例 100% |
| うち、主がゼロ化 | 68 例 81.0% | 16 例 19.0% | 84 例 100% |

制約が適用されて後件の主がシステムと推定されたもののうち 83.8% が正しく推定されていた。主がゼロ化されたものに限ってみても 84 文のうちの 68 文、すなわち 81.9% が正しく推定されている。このように 2 節で示した「と」の制約が計算機上でも有効に利用されることが示された。

ここで、制約適用システムによる主の推定が間違っていた 16 例の内訳について述べておく。まず、16 例のうち 7 例はこの節のはじめに述べた格助詞の「と」を接続助詞と判断してしまったためにおきた誤りであった。残りの 9 例のうち 7 例が (3) のような「と」文の後件にさらに「ので」等による接続がある文で、そのかかり方のため制約を適用させる述語の推定がうまくできなかったためにおきた誤りであった。

- (3) 「電源コードをコンセントにつないで電源を<入>にすると内部のヒーターがはたりますので十分に乾燥させて水滴の心配がなくなってからお使いください」 [三菱]

4 節でも述べたように制約適用システムは文末から制約を適用する語を探し出す。そのため、制約が適用される語として「使う」を選んでしまったのでうまく推定できなかったのである。このような文でも制約が適用される語を正しく選んでくることができれば、結果はさらに良くなるであろう。残った 2 例は (4) に示すように、「られる」が可能表現として使われていたためにおきた誤りであった。

- (4) 「+または-ボタンを押すとスロー再生の速さを変えられます」 [三菱]

この場合は可能表現のため後件の述語は非動作なのであるが受身の「られる」と区別がつかないという問題がある。この点についてはさらに検討を行う必要があるであろう。

5.2 「れば」文の場合

次に、「れば」文の場合について述べる。「れば」文 41 文のうち、後件の述語が動作と判断されて「れ

ば」文のデフォルト規則が適用されたものは11文存在した。このデフォルト規則の適用の結果、システムと推定された述語の主が実際にシステムなのかどうかを人間の判断によって確認した結果を表4に示す。表の成功と失敗は「と」文のときと同様である。

表4: 「れば」文のデフォルト規則適用結果

| | 成功 | 失敗 | 合計 |
|----------|-------------|-------------|-------------|
| 全体 | 9例 81.8% | 2例 18.2% | 11例 100% |
| うち、主がゼロ化 | 8例 80.0% | 2例 20.0% | 10例 100% |

制約が適用されて後件の主がシステムと推定されたもののうち81.8%が正しく推定されていた。主がゼロ化されたものに限ってみても10文のうちの8文、すなわち80.0%が正しく推定されている。このように例はやや少なめであるが「れば」文についても2節で示したデフォルト規則が計算機上でも有効に利用できることが示された。

なお、デフォルト規則の適用による主の推定が間違っていた文のうち1文は「られる」が可能表現で使用されていたものであった。残りの1文は次の文であった。

- (5) 「リストの第1要素に形態品詞を記述し、品詞細分類があれば第2要素に品詞細分類名をリストによって記述する。」[松本]

これは、「れば」文のデフォルト規則の例外と考えられる。

5.3 「たら」文の場合

「たら」文のデフォルト規則は「後件の主はユーザになる」というものなので11文すべてに制約が適用された。この制約適用によってユーザと推定された述語の主が実際にユーザなのかどうかを人間の判断により確認した結果を表5に示す。ただし、表に成功とあるものは制約が適用された述語の主が実際にユーザであったものであり、失敗とあるものは述語の主が実際にはシステムであったものである。

表5: 「たら」文のデフォルト規則適用結果

| | 成功 | 失敗 | 合計 |
|----------|-------------|----------|-------------|
| 全体 | 11例 100% | 0例 0% | 11例 100% |
| うち、主がゼロ化 | 11例 100% | 0例 0% | 11例 100% |

このように、今回調べたマニュアルでは「たら」文の場合は後件の主がすべてゼロ化されていたのであるが、デフォルト規則に反する文は見つからず、すべて正しく主を推定することができた。このように「たら」文についても2節で示したデフォルト規則を計算機上で活用できることが示された。

5.4 「なら」文の場合

「なら」文のデフォルトは「たら」文と同じく後件の主ユーザとなるというものであったが、今回集まったデータ5文ではうまく制約がはたらかなかつた。5文の後件のうちわけはユーザの動作が2文ユーザの非動作が1文システムの動作が1文システムの非動作が1文であった。「なら」文についてはまだ、例文が少ないという問題があるためさらに検討を行う必要があるようである。ここで、[森95]の時点で集まった「なら」文も含めて後件についてのみ整理するとユーザの動作が11例、ユーザの非動作が1例、システムの動作が1例、システムの非動作が1例となる。これまでのすべてのデータを合わせた限りでは制約が成り立つようであるが今後も例文を集めて制約を検証する必要があるだろう。しかし、マニュアル文中に「なら」文がほとんどないということもまた事実である。

6 終わりに

マニュアル文に現れる条件表現「と」、「れば」、「たら」の性質より導かれた、各条件表現の後件の主についての制約およびデフォルト規則が計算機でも有効に活用できることを確認した。しかし、「なら」のデフォルト規則についてはまだ、例文が少ないためその有効性を確認しきれていない。「なら」については今後も例文を集めて検討していく必要があるだろう。また、条件表現の後件が複文になっているような場合にも存在したため、このような場合の考察も必要であろう。

参考文献

- [IPA87] 情報処理振興事業協会. 計算機用日本語基本動詞辞書 IPAL(Basic Verbs) - 解説編 -. 東京, 1987.
- [久野 73] 久野 暲. 日本文法研究. 大修館書店, 東京, 1973.
- [仁田 93] 仁田 義雄. 日本語の格を求めて. 仁田 義雄 (編), 日本語の格をめぐって, pp. 1-37. くろしお出版, 東京, 1993.
- [益岡 93] 益岡 隆志. 日本語の条件表現について. 益岡 隆志 (編), 日本語の条件表現, pp. 1-20. くろしお出版, 東京, 1993.
- [森 95] 森 辰則, 龍野 弘幸, 中川 裕志. 日本語マニュアル文における条件表現「と」「れば」「たら」「なら」から導かれる制約. 自然言語処理, pp. 3-18. 言語処理学会, 10月 1995.

マニュアル出版

- [富士 88] ハイテクノロジーコミュニケーションズ (株). OASYS Lite F・ROM 11/11 D 操作マニュアル, 1988.
- [松本] 松本 裕治, 黒橋 禎夫, 宇津 呂 武仁, 妙木 裕, 長尾 真. 日本語形態素解析システム JUMAN 使用説明書 version 1.0.
- [三菱] 三菱電機株式会社. 三菱ビデオ HV-F93 取扱説明書. など 5冊