

話し言葉による説明発話の特徴の分析と具体化について -格要素の語順に関して-

古田 陽 島津 明

北陸先端科学技術大学院大学 情報科学研究科
e-mail: {a-furuta, shimazu}@jaist.ac.jp

1 はじめに

近年、人間と機械とのより自然なインタラクションを求め、対話処理に関する様々な研究が行われている。音声対話では書き言葉ではなく話し言葉によるユーザからの発話を受けて、システムが出力する発話についても人が聞いたときに自然な発話となっていることが、有用なインタフェースとして望まれる。従って、対話処理において自然な対話を扱うことが、音声での対話処理を行う際に重要となる。

本研究は、人が物事を説明するときの発話に関して、人と人との対話を書き起こしたコーパスに見られる人の発話の特徴を抽出、規則化することを目的としている。人と人との対話を書き起こしたコーパスより規則を抽出することで、その規則に従い発話を生成することでより人が説明を行うときの発話に近い発話が生成でき、それは人に対して聞きやすい自然な発話であることが期待できる。なお、ここでの発話の特徴とは、表層的な文字情報に見られる特徴を指すのではなく、深層的な格構造レベルおよび発話の間の修辞構造などの特徴を指す。また、対象ドメインとして主に交通経路案内を用いて、そこに見られる特徴の調査と具体化について調べる。

2 関連研究

対話処理システムの研究の中には、人にとって自然に聞こえる発話を生成するための研究も広く行われていて、[5][6]などが挙げられる。しかし、これらの研究では発話生成そのものに対して、話し言葉の特徴の分析に基づいた表現を用いることは考慮されていない。

話し言葉の表現を用いた発話生成に関する研究として、堂坂らの研究 [1] がある。堂坂らは交通経路案内の対話コーパス 15 対話を見ることにより修辞構造による談話関係の頻度分布を調べ、路線図における交通経路の問題を入力とし、問題解決により立案されたドメインプ

ランを階層的プランニング技法により、談話関係を用いて表層的な情報伝達行為まで詳細化し、漸次的に発話を生成している。この研究では漸次的発話生成に対して、人の発話で多く見られた修辞関係をプランスキーマにより表現し、言語運用制約を用いて話し言葉を生成している。しかし、プランスキーマ化された修辞関係について、話し言葉の表現を表すのに十分であるかは不明である。また、それらのプランスキーマも明確でない部分があることから、それらの修辞関係による人の発話の特徴をより明確化する必要があると言える。

本研究では堂坂らの研究では考慮されていない特徴のひとつであり、発話を生成するさいの基本ともいえる、動詞に対する格要素の発話語順に関して調査し、話し言葉による発話生成の規則として明確化した。

3 調査

3.1 使用コーパス

特徴の分析に関して、本研究ではテキストに書き起こされた交通経路案内コーパスを利用した。41 対話について、コーパス中の導入部分の挨拶、自己紹介、交通経路を案内した後の会話部分、および聞き手側の発話すべてを除いた、説明者が交通経路を案内しているときの 3011 発話を調査の対象とした。書き起こされたコーパスの一部を図 1 に示す。

3.2 調査した情報

本研究では、動詞およびその動詞に対する格要素や注視対象に対して、それらが発話される語順に関する特徴を取り上げる。日本語は語順の制約が比較的少ない言語であり、動詞に対する格要素の発話順序というものは一意には決まるとは言えない。しかし、格要素の新旧性や注視対象といった面から、人が発話を行うときの発話順序の特徴を見いだすことは可能と考え、それを明確にす

(説明者:KiFu, 問合者:ToFu)

KiFu: で、えーと、多摩センターから小田急線で、
<はい>えーと、新百合ヶ丘までまず、出て
いただくことになりますね。

ToFu: はい。

KiFu: で、そこからですね、えーと、<はい>小田
原行きか、<はい>えーと、箱根湯元行き、
<はい>に乗っていただいて、えーと、愛甲
石田まで、<はい>ということになります。

ToFu: わかりました。=

KiFu: =あの一、えーと、あっ、愛甲石田はわかり
ますか、えーと、愛するの愛に、甲乙の甲、
で、石にたんぼの田です。

図 1: 分析した交通経路案内の対話例

ることとする。

ここでは、新情報の定義として、過去に説明者および聞き手の発話で、一度も表層に現れていない情報を示す格要素が表す情報を新情報とし、一度でも表層情報として発話された情報は旧情報とした。

また、注視対象については、堂坂ら [1] および、村岡ら [2] を元に以下のような場合に遷移すると仮定する。

- 交通経路案内の対話において、ある行為を説明した場合、その行為後に聞き手が存在すると考えられる場所を注視対象とする。
- 説明者および聞き手が、何らかの情報を伝えた場合、その対象となる情報を注視対象とする。

図 1 の発話であれば注視対象は「多摩センターから小田急線で、新百合ヶ丘までまず、出ていただくことになります」の発話で新百合ヶ丘になり、「で、そこからですね、小田原行きか、箱根湯元行きに乗っていただいて」の発話で小田原行きか箱根湯元行きへ移行し、「愛甲石田まで、ということになります。」の発話で愛甲石田へと移行する。

3.3 調査結果

表 1 に、動詞、その格要素の出現順序、および注視対象の位置に関してまとめたものを示す。格要素の順序に関しては、発話として表層上に現れたもののみを調べ、ゼロ代名詞などの、省略されたと思われるものに関しては触れていない。41 対話の中で交通経路を案内するために使われた動詞は全部で 229 個存在した。

表中の格要素の順序の欄の記号は、 L_O , L_I : 場所格, O : 対象格, S : 源泉格, I : 道具格, G : 目標格, T : 時間格, V : 動詞をそれぞれ意味する。また、 L_O のように下線が引

表 1: 動詞に対する格要素の順序と注視対象の位置

格要素の順序	動詞数	主な動詞
\dot{L}_O, V	83	降りる, 出る, 着く
\dot{L}_I, V	29	乗る, 乗り換える
\dot{G}, V	24	来る, 出る, 行く
$\underline{L}_O, \dot{L}_I, V$	23	乗る
L_O, \dot{L}_I, V	14	乗る, 出る
\dot{L}_O, V	13	降りる, 着く
V, \dot{L}_O	8	降りる
$\underline{L}_I, \dot{L}_O, V$	6	降りる
\dot{L}_I, V	4	乗る
T, \dot{L}_O, V	2	降りる
T, V	2	待つ
$\underline{S}, \dot{G}, V$	2	行く
$\underline{I}, \dot{G}, V$	2	来る
$\underline{L}_I, V, \dot{L}_O$	2	降りる
T, \dot{G}, V	1	着く
\underline{L}_O, T, V	1	乗る
$\underline{L}_O, \dot{L}_I, V$	1	乗る
T, \dot{L}_O, V	1	降りる
\underline{L}_I, T, V	1	乗る
\underline{S}, T, V	1	着く
S, \dot{G}, V	1	行く
\dot{G}, V	1	行く
$\underline{S}, \underline{I}, \dot{G}, V$	1	出る
$\underline{L}_I, T, \dot{L}_O, V$	1	降りる
$\underline{L}_O, T, \dot{L}_I, V$	1	乗る
$\underline{I}, S, \dot{G}, V$	1	出る
$\underline{S}, T, \dot{G}, V$	1	来る
$\underline{S}, \dot{I}, V$	1	行く
T, V	1	乗る

かれている格はその格要素が旧情報であることを意味し、 \dot{L}_O のように格要素の上に点が付いているものは注視対象となるものである。例えば図 1 の「そこからですね、えーと、小田原行きか、えーと、箱根湯元行き、に乗っていただいて、」という発話は $\underline{S}, \dot{L}_I, V$ となる。旧情報である場所格「そこ(新百合ヶ丘)」が発話されたのち、対象格「小田原行きか箱根湯元行き」、動詞「乗る」の順に発話され、この発話の後に注視対象は場所格に移ったことを意味する。

(説明者:ReYa, 問合者:AkSa)
 AkSa: えーと、森の里青山行き。
 ReYa: 森の里。
 AkSa: 青山行き。
 ReYa: 青山行き。
 AkSa: か、えーと、松蔭女子短期大学行き。
 ReYa: …、はい。
 AkSa: と、いうのに乗ってください。
 ReYa: はい。
 AkSa: で、これを、通信研究所前というところで降りてください。

図 2: 旧情報より発話を行う具体例

(1) 格要素の発話順序と新旧性

語順と新旧性に関して、一般に日本語には次の事が言われている [3]。

文中の語順は、古い情報を表す要素から新しい情報を表す要素へと進み、最新の情報を表す名詞句は述語の直前に位置する。

表 1 において、動詞に対して複数の格要素が発話されている動詞の個数は 64 個存在し、そのすべてで、格要素の中に旧情報がある場合は旧情報より発話していることがわかる。よって、話し言葉による説明の発話においても、一般に日本語の語順に対して言われていることが成り立っているといえる。これらより、交通経路案内において格要素の新旧性と語順に関する規則として以下のものを用いることができる。

< 格要素の新旧性と語順に関する規則 >

動詞に対して複数の格要素を発話する場合、旧情報となる格要素より発話を行い、その後新情報となる格要素の発話を行う。

コーパスに見られた具体例を図 2 に示す。図中では、動詞「降りる」に対して、旧情報の場所格の「これ(森の里青山行きか松蔭女子短期大学行き)」を発話してから、場所格「通信研究所前」を発話しているのがわかる。

(2) 注視対象の位置

表 1 より、注視対象となる格要素についても、話し言葉において発話位置の特徴がわかる。注視対象となる格要素は話の中心となるものであり、話し手が強調して伝えたいことでもある。今回調べた説明発話に使われていた 229 個の動詞のうち、213 の動詞で注視対象は動詞の直前に発話されていた。このことは、日本語は「話し手

TiKu: で、十五分から二十分ほどで、えー、乗りまして、降りる、のが通信研究所前です。

図 3: 動詞の後に格要素の情報が発話される例

が強調したいこと(焦点)は文の陳述の部分で語られることが多い」という性質 [3] を裏付ける。

これらより、交通経路案内において、注視対象の格要素の語順に関する規則として、以下のものを用いることができる。

< 注視対象の格要素の語順に関する規則 >

注視対象にあたる格要素は動詞が発話される直前に発話する。

(3) 動詞の位置

日本語は語順の制約が緩い言語であり、それは格要素に限らず動詞にもいえる。実際に格要素や動詞に限らず、特に強く強調したい要素に関しては、他の要素より先に発話を行うことで特に強い強調を表す場合などがある。

表 1 より、交通経路案内の対話において動詞が、注視対象となる格要素や他の格要素よりも先に発話されるような、倒置の関係で強調されることは起きていなかった。動詞が発話されたのちに格要素が発話される状況は、10 個の動詞に対して見られたが、それらはすべて図 3 のような例であり、動詞を強調するために倒置を利用している発話とは考えにくい。この発話の場合、動詞「降りる」およびその場所格情報を示す「通信研究所前」の順序が、V, L₀ となっている。このような発話を話者の視点より調べたところ、発話が見られた 10 箇所のうち、8 箇所は同一の話者から発話されていた。

(4) 注視対象と格の省略の関係

表 1 より分かることとして、コーパス中に見られた 229 個の動詞のうち、165 個の動詞で表層上に現れた格要素の数が 1 個であった。これは、ゼロ代名詞やその他の要因で格要素が省略されたためと考えられる。しかし、注視対象となる格要素が存在しない動詞は、わずか 4 個のみである。これらの動詞について調べると、直前に行われた発話と比べて注視対象が移行しない場合は、注視対象である格要素が省略されることもあることが分かった。

これらより、原則として、注視対象となる格要素は旧情報であってもゼロ代名詞などとして省略されることが稀であることがあげられる。これは、実際に注視対象となる格要素は話の中心となるもので、これを省略すると情報が正確に伝わらない可能性があるからであるとみられる。しかし、例外として、直前の発話と比べ注視対象

表 2: 格要素の順序と注視対象の位置

格要素の順序	動詞数	主な動詞
\dot{O},V	25	クリックする, 押す
$\underline{L}_O, \dot{O}, V$	5	入れる, 持っていく
\underline{L}_O, V	5	クリックする, 離す
\dot{O}, V	4	入れる, クリックする
$\underline{S}, \dot{O}, V$	3	押す
L_O, \dot{O}, V	3	持ってくる, 入れる
O, \dot{I}, V	1	取る
S, \dot{G}, V	1	ドラッグする
S, V	1	ドラッグする
\dot{I}, V	1	閉じる
\underline{S}, V	1	クリックする
$\underline{O}, \dot{L}_O, V$	1	持ってくる

が移行しない場合は、省略しても構わない場合がある。

堂坂ら [1] では省略は「伝達済みの情報は繰り返さない」という言語運営制約を用いることで扱っている。実際の対話での省略は伝達済みの情報をすべて省略しているのでなく、伝達済みの情報でもそれが注視対象であれば省略されることが分かる。

(5) 他ドメインの格要素の発話語順

今回、用いることにした規則に関して、他のドメインの説明発話にも適用可能か調べる。利用したコーパスは、エクセルによるグラフの作成方法の説明コーパスを利用した。2 対話 (2480 ターン) の中より、説明に用いられた動詞 51 動詞と、格要素の語順のパターンに関して、表 2 に示す。これより、交通経路案内と同様の特徴があることが分かる。

4 実装

前述した規則に基づき実験システムを作成している。システムは、入力情報として与えられる意味構造を受けとり、それをプランナーによりプランニングし、発話プランとして表層的な発話へと直接変換できる形へと変換する。そしてその発話プランを実際の話し言葉の表現の発話に変換することで、発話を生成する。

入力として与える意味構造として、格構造表現、注視対象を用いる。これらは、交通経路案内ドメインのため、ある場所からある場所へ移動する、ある場所からある物に乗る、ある場所である物から降りるなどという交通経路を案内するための行為の連続になっている。

前述の規則は、プランオペレータとして記述される。

```
(:operator common-rule-os-i-fg
:parameters (?act ?s ?i ?g)
:precondition
  (and (action ?act)
        (and (old-inform ?s) (source ?s))
        (and (new-inform ?i) (instrument ?i))
        (and (focus ?g)(goal ?g)))
:effect
  (and (surface Source-> ?s Instrument-> ?i
        Goal-> ?g Action-> ?act)
        (action-fin)))
```

図 4: プランオペレータの一例

具体例としてオペレータの一部を図 4 に示す。また、プランナーには UCPOP[4] を利用している。

5 まとめ

本稿では、人の説明発話の特徴のひとつとして、基本的な特徴ともいえる、動詞に対する格要素の語順に関して調査し、その特徴を述べた。これは基本的な特徴であるが、この他にも話し言葉に見られる特徴はある。また、堂坂ら [1] の研究で確認されている特徴の中にも、明確化されていない特徴がある。今後、それらの特徴について調べ、プランオペレータとして記述し、生成された発話の評価を行っていく。

参考文献

- [1] 堂坂 浩二, 島津 明. タスク指向型対話における漸次的発話生成モデル. 情報処理学会論文誌, 37-12, pp.2190-2200, 1996.
- [2] 村岡 公一, 島津 明. 話し言葉を生成するための方略. 言語処理学会第 9 回年次大会発表論文集. pp.449-452. 2003.
- [3] 長尾 真. 自然言語処理. 1996
- [4] UCPOP. <http://www.cs.washington.edu/ai/ucpop.html>
- [5] Kazunori Komatani, Shinichi Ueno, Tatsuya Kawahara and Hiroshi G. Okuno. Flexible Guidance Generation using User Model in Spoken Dialogue Systems. 41st Annual Meeting of the ACL. pp.256-263. 2003.
- [6] 大森 久美子, 斎藤 博昭. 大語彙を対象とした音声対話インタフェースにおける自然な応答生成. 自然言語処理論文誌 V10N5. pp23-40. 2003.