

漸次的発話生成における言い間違いの分析

羽尻公一郎 (ATR 知能映像通信研究所/立命館大学),
岡田美智男 (ATR 知能映像通信研究所), 小川均 (立命館大学)

1 はじめに

漸次的発話生成においては、発話目標の形成と統語的操作、調音動作と既発話の蓄積が並列に動作し、そこには時間の経過に伴う各レベル間の不整合が生じる可能性がある。例えば先に形成された部分的な発話目標を統語操作する際には、格などの情報の欠落（ただし後に補充される可能性はある）があり得る。それに対して Kempen や Levelt の言い直しの研究 [1, 2] に代表されるように、発話目標に曖昧性があるままに漸次的発話を開始し、次々と既発話が蓄積され、時間の経過とともに発話目標に曖昧性がなくなった時点で、既発話と発話目標の不整合が生じた時、言い直すというメカニズムが考えられる。このとき重要なことは、発話目標に曖昧性がある場合、統語操作などの下位のレベルである種の暫定的な仮定（発話目標に関する短期的予測、もしくはデフォルト）の存在なくしては漸次性の実現できないという点である。部分的な比較的小さい発話目標を入力とする漸次的統語操作モデルを用いて言い直しを説明する場合には、この統語操作レベルでの暫定的仮定と、発話目標と既発話の整合性という2点に留意しなければならない。

このような観点から、漸次的発話生成における言い直しの説明モデルの構築の一環として、言い直しの心理言語実験を行なった。実験においては三段階の絵の段階的・部分的な追加により、一つの絵が完成するような提示を行なう。絵の内容は男の子が女の子を傘で叩いているなど、動作主と非動作主の二要素、およびそれらの関係を示す動作から構成される。絵は最初に動作主もしくは非動作主が動作を示す手を除いて提示され、それに二つ目の要素として非動作主もしくは動作主が追加され、最後に動作を示す両者の手と道具が追加される。被験者は段階的に提示される絵の完成を待たずに、最初の要素が提示された時点で発話を開始し、絵の要素が追加されるのに合わせて描写するように指示されている。動作の要素が最後の提示で追加されるため、被験者は最初と二つ目の提示においてどちらが動作主であるか分からない。このような条件で漸次的発話をするためには、最初の二段階で提示される要素のどちらかを動作主と仮定して、発話文を組み立てなければならない。そしてその仮定が外れた時、言い直しが発生する。この実験において、提示される動作主/非動作主の順番と被験者が仮定する動作主の関係、およびその仮定に基づく発話文の語順と格助詞の使い方を調べ、漸次的発話の初期ステージにおいて発話目標に曖昧性があるとき、人間がどのようにその曖昧性を解消しつつ発話文を形成するのかがわかる。また仮定が外れて言い直しを行なう場合、言い直し発話がどのような構成になっているのかを調べることで、言い直し前と言い直し後での関係や既発話と Inner Speech の両方の資源と発話生成過程がどのように相互作用して言い直しを形成するのかなどが推測できる。

2 実験

発話の収録は防音室に行ない、被験者の発話は防音室内のマイクから外部の DAT にデジタルレコーディングされた。刺激の絵は防音室に備え付けの 80 インチのプロジェクタを用いて提示された。絵の提示は Macintosh 上で動作するソフト、PsyScope1.1 を用いて制御された。被験者が準備画面が投影されている時にマウスをクリックすると刺激の絵が漸次的に提示されはじめる。一人あたりの実験に費やした時間は約 20 分である。指示においては漸次的に絵が現れることを被験者に伝え、絵の漸次的提示に合わせて出来るだけ早く描写（発話）を開始することが指示された。また、指示において提示される全ての絵の材料が提示され、その呼び方が示された。これにより、被験者は使用可能な語彙を、動作の描写（＝動詞の使用）を除いて予め限定される。

収録された DAT の音声データは計算機上の音声データファイルに変換され、書き起こしは専用の GUI アプリケーションを用いて行なわれた。書き起こしの形式は、単語のわかし書きでひらがなのみで行なわれ、

400msec 以上のポーズを境界とする発話単位を基本として書き起こされた。録音状況 (SN 比など) は良好で、書き起こしに際して聞き取りにくい発話などはほとんどなかった。

刺激となる絵のマテリアルにはカラーのイラストで、男の子、女の子、トラック、タクシー、道具のスパナ、傘がある。動作としては人間が対象を叩く (上から道具を降り下ろしている)、押える (対象の前に道具を突き出す)、Automobile が対象にぶつかっている (衝突の衝撃を表す描写線で表現される) の 3 種類がある。マテリアルの組合せによる実験条件は、以下のようになる。

★動作主が人間の場合

- 動作主の提示順序 (先か後か)
- 対象の種別 (人間か物か: カウンターバランス, 人間の場合男か女か, 物の場合タクシーかトラックか: 相殺条件)
- 動作主の登場位置 (右側か左側か: カウンターバランス)
- 動作主の性別 (男か女か: カウンターバランス)
- 動作主の動作 (叩くか押えるか: カウンターバランス)
- 動作主の使用する道具 (スパナか傘か: 相殺条件)

★動作主が物 (Automobile) の場合

- 動作主の提示順序
- 動作主の登場位置
- 動作主の種類 (タクシーかトラックか: カウンターバランス)
- 対象の性別 (男か女か: 相殺条件)

以上、人間が動作主の場合 $2^5 = 32$ 、物が動作主の場合 $2^3 = 8$ で、合計 40 が被験者一人あたりの試行数となる。今回は動作主が人の場合のみを分析対象として想定したため、動作主が物の場合はうめぐさ (catch item) とする。よって今回注目する条件は人間が動作主の場合の動作主/被動作主の提示順序と漸次的発話の関係である。他の条件は予備実験により、語順や態の決定との相関が認められなかったため、カウンターバランスもしくは相殺条件としている。

提示順序は「動作主→被動作主→動作」もしくは「被動作主→動作主→動作」のいずれかである。先に提示される動作主/被動作主には動作に関する手がかりが欠けており (人間ならば道具を持つ手が現れない、Automobile なら衝突を示す描写線が描かれていない)、最後の動作の際に欠けている部分が現れるようになっていく。各マテリアルの提示時間は、1) 準備画面 (被験者がマウスをクリックしてスタート) → 2) 最初の動作主/被動作主の提示 (2 秒) → 3) 次の被動作主/動作主 (2 秒) → 4) 動作の提示 (絵の完成) (5 秒) → 1') 準備画面という具合になっている。この一連のサイクルが一つの試行の流れである。なお、提示時間は予備実験を参考にして決定した。

日本語の標準的文型は主語→目的語→動詞という語順であるため (1 参照)、この提示順序では最初の刺激が出た時点から漸次的発話を始めるにあたり、発話内容に曖昧性が残る。

$$\begin{aligned} & \text{Subject}(\text{NorNP}) + \text{Case} - \text{particle}(\text{PP}) \rightarrow \text{Object}(\text{NorNP}) + \text{Case} - \text{particle}(\text{PP}) \rightarrow \\ & \text{Verb}(\text{VorVP}) + \text{Voice} - \text{particle}(\text{RERUorRARERU}) \end{aligned} \quad (1)$$

↑PP として今回の実験で使われる可能性が高いものは、「が」「を」「に」である。

↑例えば上記の形式での発話例は、

能動態「(おとこのこ(N)が(PP)Sbj)(おんなのこ(N)を(PP)Obj)たたか(V)」

受動態「おんなのこ(N)が(PP)Sbj)(おとこのこ(N)に(PP)Obj)たたか(V)れる(Voice - particle)」である。

最初と 2 番目に提示された刺激のどちらが動作主であるかが分からないために、どちらかのマテリアルを暫定的に主格として発話文を組み立てなければならない。しかし最後に動作が現れて初めて、先の 2 つのマテリアルの関係が確定される。このとき、暫定的に設定した主語が被動作主である場合、言い直しか、最後の動詞の発話の際に態を即時的に決定しなければならない。日本語の標準的な語順では動詞が最後に来る。さらに受動態は動詞に接尾辞 (れる・られる) を追加することで実現される。つまり、この実験の提示順序による漸次的発話においては曖昧性を含んだ発話目標に対して暫定的に格を割り振り漸次的発話を開始し、

刺激提示が完了したときに暫定的に割り振った深層格と提示刺激の動作主／被動作主の関係に不整合が生じた時、1) 言い直す、2) 即時的に受動態に変更する、(3) 漸次的発話をせず、刺激の絵の曖昧性がなくなるまで発話しない」という対応が考えられる。

実験の被験者には奈良先端科学技術大学院大学の学生や主婦など、男女合わせて17名が集められた。被験者は4000円の謝礼を与えられた。

3 分析結果

集積された書き起こしは、全部で720の文章を含む。そのうち今回の分析対象である人間が動作主の文章は全部で544である。この半分の277ずつがそれぞれ動作主が先に出る場合、後に出る場合である。Manualによって言い直しを含む文章を抽出したところ、動作主が先に出た場合のError-repairは41、後に出た場合は70であり、Appropriate-repairや誘発によらない自発的言い直し(「スバナで。いや、傘で」など)は先に出た場合が6、後に出た場合に13発生していた。この誘発によらない言い直しは、今回の分析対象から除外した。

今回の実験では漸次的発話生成において、発話目標に曖昧性がある場合に暫定的な目標の情報の補間を行ないつつ、提示される刺激の文脈に対してどのような統語的操作が行なわれ、その結果として語順がどのように構成されるかを調べるのが目的なので、言い直し前(RM)と言い直し後(RR)の語順のパターンと受動態／能動態の使用頻度を調べた。以下にその結果を表1～6に示す。なお、表1～6の「1が2を」などの数字は、発話中の名詞や名詞句に対応する実験での提示順序の数字であり、「[]」は中の内容が省略されている場合も含む。表1～6の値の単位は全て%である。

表 1: 分析対象全体における言い直しの発生率

	動作主(主語) 先	非動作主(目的語) 先
言い直し	14.8	25.3
非言い直し	81.2	68.2

表 2: 非言い直し文における受動態の使用頻度

	動作主(主語) 先	非動作主(目的語) 先
受動態	0.4	27.0
能動態	80.9	41.2

表 3: 分析対象の言い直し文における受動態の使用頻度

	動作主(主語) 先	非動作主(目的語) 先
受動態	0	11.4
能動態	100	88.6

表 4: 分析対象の言い直し文の訂正対象(RM)の語順

	"1が[2を等]"	それ以外
RMの語順	82.9	17.1

表 5: 非動作主先行の言い直し文の訂正部分(RR)の構成

	1が2が	2が1を	1が2に(受動態)	その他
RRのパターン	0(「Xを」だけの場合が1.6)	62.5	0(「Yに」だけの場合が10.9)	25.0

†修復対象は"1が[2を等]"のパターン

表 6: 動作主先行の言い直し文の訂正部分(RR)の構成

	1が2を	2が1が	1に2が(受動態)	2が1に(受動態)	その他
RRのパターン	10.9	0(「Yを」のみの場合が47.8)	0	0	41.3

†修復対象は"1を等[2が]"のパターン

4 考察

表1より、動作主が後に出る場合のほうが先に出る場合より、有意に言い直しが多いのがわかる。表2および表3より、動作主が後に出る場合は言い直すか、受動態を選択することによって、漸次的発話刺激提示の開始における曖昧性と、その解消に伴う既発話の不整合を修正していることがわかる。ただし、受動態を用いて不整合を解消する場合、目的語の格助詞には「を」ではなく「に」を用いなければならない。即時的に受動態への変更を行なっている被験者は、曖昧性がなくなり動作主と被動作主の関係を確認するまで2つめの格助詞の発話をしない場合と、一度能動態を仮定して「を」を発話してから「に」を言い直している場合が考えられる。これについては、表3が答えてくれる。表3において、言い直しが発生している場合にはほとんど能動態が用いられている。このことから、言い直しを回避するために受動態を用いる場合には、曖昧性が解消されるまで2つ目の格助詞を発話していないと思われる。この点は今後、収録時の発話の時間情報を参照したさらに詳しい分析で解明されると考える。また、表2より動作主が先に出た場合にも言い直しがある程度発生しているが、動作主が後に出る場合とは有意に差があると考え、例外と考える。この動作主が先に出る場合の言い直しの発生の原因としては、試行を繰り返すうちに被験者の注意が、動作主が後に出る場合に向くためと考えられる。

表4より言い直し前(RM)の語順(曖昧性のある発話目標に対する暫定的語順)は、提示された最初の刺激を主語とし、2つ目の刺激を目的語とする(「1が2を」などのパターン)場合がほとんどであり、暫定的な語順のストラテジはほぼ1つであると言える。さらに言い直し後(RR)の語順のパターンは表5からわかるように、主語と目的語を入れ換え、格助詞のデフォルトの順番(これは動詞の下位範疇化素性に基づく格助詞の結合価として日本語の動詞に関しては統計的に調べられている)を保存させている。これは刺激提示の順番よりも、語彙に備わっているデフォルトの語順が優先されることを示している。

5 今後の展開

今回の実験では17人の被験者を集め、111の言い直しの収録に成功した。今後は個人性などにも焦点を当てることを考えており、被験者の数を増やして追加実験をすることを考えている。現在の実験手法では、言い直しの誘発の成功率を完全に制御することは出来ない。今後の課題として、漸次的発話を自然に行なえる実験手順の考案、言い直しをより確実に誘発する実験条件の開発が必要である。また、今回は刺激の提示順序として動作の提示を最後に固定したが、動作が最初や2番目に提示される場合についても検討したい。

我々は現在、マルチエージェントシステムによる多自由度非線形系からの創発現象として、発話生成モデルを構築している。現在モデルは文脈からの影響や既発話の影響を反映しつつ、語彙の持つデフォルトの格助詞の順序にもとづくエージェントの相互作用として、語順を柔軟に決定する能力を持つ。このモデルを拡張することにより、漸次的発話における言い直しの生成をシミュレートし、モデルによるさらに深い理解を試みる。これに今回の実験結果や今後行なわれる追加実験の結果を積極的に参照したい。

謝辞

本実験の計画及び原稿の執筆にあたっては、奈良先端科学技術大学院大学の伝康晴氏、京都大学大学院の伝優子氏に全面的な協力を得た。ここに感謝の意を表します。

参考文献

- [1] Levelt: Speaking From Intention to Articulation, MIT Press, 1989
- [2] Kempen: A dual system for producing Self-Repairs in spontaneous speech, Cognitive Psychology, 19, pp.403-440, 1987
- [3] Kempen: Incremental Syntactic Tree Formation in Human Sentence Processing, Connection Science, Vol.1, No.3, pp.273-290, 1989
- [4] 羽尻ほか: 時間制約依存性を有する日常的な発話生成のモデル, 日本認知科学会第13回大会予稿集, pp.192-193, 1996