

発話産出における相互作用性

羽尻公一郎

ATR 知能映像通信研究所

1 まえがき

本稿では人間の発話産出過程の解明に向けた試みの一端として、筆者の所属する研究グループでの研究内容を紹介する。発話産出の研究のなかでも、いわゆる計算モデルの構築の歴史は比較的浅く、まだ20年に満たない程度である。しかし、AIや心理学などとの関係による、多様な観点からのアプローチが試みられており、次第にその全貌を明らかにしつつある。特に、「発話意図の形成→心的辞書からの単語の検索→統語的操作による発話文の構築と音韻的操作によるプロソディの形成→調音活動」という一連の処理段階の存在と、それぞれの段階が自立したモジュールとして機能していること、それらのモジュールの中では一定単位の局所的な情報のやりとりがあり、モジュール間に相互作用があることなどの指摘は、近年の研究の大きな成果である¹。

こうしたモジュールの集合として発話産出機構を見做し、いわゆる分散並列プロセスとして記述する試みは、各所で行なわれている。特に近年は、こうした並列/分散パラダイムとモジュールの相互作用性を裏付ける現象として、漸次的発話産出という発話様式が注目されている。漸次的発話産出とは、発話意図の想起が不完全なままに、発話文の形成を始め、文章のような大きな単位の形成が終らぬうちに、調音活動をし、発話してしまふような、いわゆる“考えながら話し、話しながら考える”といった、即興性の高い発話様式である。この漸次的発話産出は、上記のモジュールの相互作用として記述される発話産出の計算モデルによって、初めて記述可能となった発話様式である。

筆者の属する研究グループは、漸次的発話産出のような即興性の高い発話の分析を進めてきた。近年では音声対話コーパスの分析などを通じて、特に既発話（直前に発した言葉）と発話産出機構の動作に綿密な関係があることを見出し、これを漸次的精緻化と名付けた。漸次的精緻化のような既発話と発話産出機構の密な関係を計算モデルに反映させることで、日常的な発話において頻出する、即時的な省略や倒置の発生、さらに言い間違いなどのシミュレーションを行なうことは、近年の我々のサブゴールのひとつである。また我々は計算モデルの構築とエンピリカルな言語分析を常に並行して行なうことで、計算モデルへのエンピリカルデータの供給と、言語分析における新たな指標の獲得との循環を重要視し、実際に心理言語実験や音声対話コーパスの分析も行なっている。

本稿ではこのような一連の試みの一端として、マルチエージェントシステムを用いた同的選択というメカニズムによる発話産出モデルの概要と、モデルへ与えるパラメータや制約規則などを検討するために行なった心理言語実験の様子を紹介する。

2 漸次的発話産出の計算モデル

2.1 マルチエージェントシステムによる動的選択

本研究はモデルのドライビング・エンジンとして、Maesによって提案された Agent Network Architecture (ANA)[4]を拡張したモデルを用いている。ANAは各エージェントに動的な知識が存在せず、STRIPSライクなオペレータ記述（条件リスト、追加リスト、削除リスト）によって一意に他のエージェントとの関係とそれに沿った活性エネルギーの流れる方向（関係と活性エネルギーについては後述）が産出されるのに対して、本研究の動的選択モデルでは活性エネルギーの流れる方向を各エージェントが知識を用いて自律的

¹こうした近年の発話産出の計算モデルの発展は、文献[3]に詳しい。

に決定する機構が備わっている点が拡張されている。本研究の動的選択モデルでは各エージェントが環境の状態や他の要素との関係を参照し、知識を用いて活性エネルギーの流れる方向を自律的に制御する。この自律的に決定する活性エネルギーの流れの方向と量をエージェントの戦略、その決定に用いる知識を戦略選定知識と呼ぶ。

動的選択は目標記述と状況記述、および基本的な行為が記述された複数のエージェントの集合の3つの要素で構成される。目標記述と状況記述はともにエージェントの集合に影響を与える環境である。そこで、目標記述と状況記述の両方をあわせて環境記述と呼ぶ。エージェントには、行為の実行のための前提条件や行為の内容、近傍の他のエージェントとの協調方法などに関する知識などが記述されている。環境記述とエージェント、およびエージェント同士には単一化をベースにした局所的な関係の求め方が定義されており、それに基づいて活性エネルギーの授受が行なわれる。環境記述からエージェントへは、活性エネルギーが常に送られる。エージェント同士では個々のエージェントの条件、すなわち結ばれている関係の数や種類、発話の制限時間などに応じて自律的に戦略、つまり関係を結んだ先へ活性エネルギーを送るか受けとるか、そしてその量などを決定する。この自律性を実現するため、個々のエージェントには、条件を判断材料に戦略を決定する戦略選定知識、および戦略知識が記述されている。

あるエージェントの行為の条件がある状況において満たされており、かつそのエージェントの持つ活性エネルギーの保有量が、閾値を越えている全てのエージェントのなかで最も大きい場合、そのエージェントに記述されている行為が実行される。各エージェントには発話の最小単位として設定している語が記述されている。エージェントが実行されることによりその記述されている語が発話される。以後、エージェントの実行条件である活性エネルギー保有量に対する閾値を t と呼ぶ。エージェントは実行されるとその内部に記述されている行為内容に従って発話産出行為を行ない（具体的には前述の発話の最小単位として設定された語を発話する）、それによって状況記述が更新される。更新された新しい状況において再び局所的な関係の計算と活性エネルギーの授受が行なわれ、次のエージェントが実行される。このように局所的な計算による関係の導出と活性エネルギーの授受を繰り返し、実行されるエージェントが選択される。選択されたエージェントは実行され、それぞれの行為内容の記述にしたがって行為を行なう。その結果として環境が変化し、次のエージェントの選択に影響を与える。この環境とエージェントの選択の循環のなかから結果として、実行されるエージェントの系列、そして個々のエージェントが実行されることで起こる行為の系列が産まれてくる。

動的選択の大まかなながれを以下に示す。

1. 環境とエージェントとの関係の計算
2. エージェント同士の関係の計算
3. 個々のエージェントの戦略の選定、活性値の入出力の計算
4. 実行するエージェントの選定
5. エージェントに記述された行為の実行と環境の更新
6. 全ての目標が達成されれば終了／達成されていなければ 1. に戻る

この動的選択モデルの特徴は、環境の記述とエージェントの関係を求める方法を局所的に設定すれば、多様な環境記述を無数に追加できる点である。また、各関係に重みづけをするパラメータを設定し、それらの組合せを調整することで、行為の組織化の傾向を変化させることができる。図1に動的選択の概念図を示す。

2.2 エージェントと環境の記述

2.2.1 目標記述

目標記述には、発話目標が記述される。発話目標は Goal Description Unit(GDU) という単位を基本的に記述される。GDU の構成を以下に示す。

[Proposition : Modality : ActivationLevel]

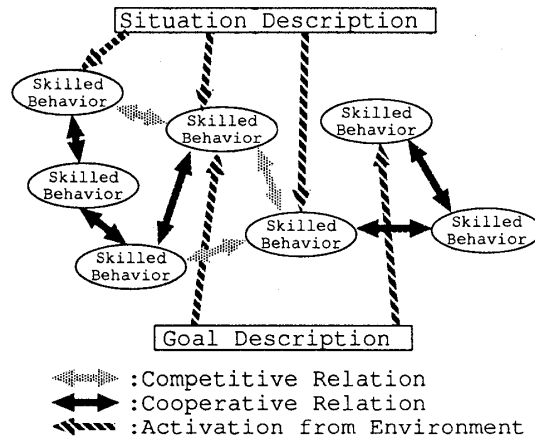


図 1: 動的選択の概念図

ここで、*Proposition* は命題情報、*Modality* にはモダリティ情報、*ActivationLevel* は送出する活性エネルギーの量である。命題情報とは、物、出来事、人などを意味する発話の核となる情報であり、モダリティ情報とは、様相や発話者の態度などを表現する情報である。目標記述には複数の GDU を記述できる。

2.2.2 状況記述

状況記述には、先に発話された言葉が記述される。このため状況記述のことを発話状況記述とも呼ぶ。発話状況記述は先に発話された言葉を記録するだけでなく、次に続く発話を実現するエージェントに対して活性エネルギーを送る。たとえば複数の発話目標が並列に存在する時、先に発話された言葉が発話を完成するため、他の発話目標を実現するエージェントより先に、続く発話を実現するエージェントを引き出す。これは自発話に対する瞬間的な推敲の発話への影響を再現する。

2.2.3 エージェントの記述内容

エージェントの記述は心的辞書の語彙に相当する。すなわち、局所的な文法情報、意味情報などである。局所的な文法情報には HPSG[5]、JPSG[2] に基づいた、下位範疇化素性情報や品詞情報などがある。また、意味情報として語彙の命題情報が記述されており、これが環境情報との関係を計算する際に参照される。さらに先にも述べたように、個々の状況に応じて戦略、すなわち近傍の関係するエージェント間での活性エネルギーの授受の方向を決定するために、戦略選定知識と戦略知識が記述されている。

図 2 にエージェントの記述例を示す。WORD: に続くのがエージェントの語彙名、HEAD: に続くのが下位範疇化素性情報以外の文法情報、SUBCAT: に続くのが下位範疇化素性情報、SEM: に続くのが意味情報である。なお、戦略選定知識および戦略知識は第 2.3 章に示す。

2.3 制約と局所的計算

2.3.1 時間的制約

発話産出行為全体に対する時間的制約として、制限時間がある。発話産出行為を組織化する過程、具体的には動的選択のながれに従ってエージェントを選択する度に時間が消費され、発話の制限時間が減少する。エージェントはシステム時計にアクセスして制限時間を知ることができる。

```

WORD: 野茂
HEAD: [POS:n,FORM:n,GR:NIL]
SUBCAT: NIL
SEM: nomo
      (a) 「野茂」

WORD: が
HEAD: [POS:p,FORM:ga,GR:_Y]
SUBCAT: [HEAD:[POS:n,FORM:n,GR:_Y]
          SUBCAT:NIL
          SEM:_X]
SEM: _X
      (b) 「が」

WORD: 獲-r
HEAD: [POS:v,FORM:root,GR:NIL]
SUBCAT: [HEAD:[POS:p,FORM:ga,GR:subj],
          SUBCAT:NIL,
          SEM:_X]
          [HEAD:[POS:p,FORM:wo,GR:obj]
          SUBCAT:NIL,
          SEM:_Y]
SEM: Get,subj:_X,obj:_Y
      (c) 「獲る」

```

図 2: エージェントの記述例

2.3.2 局所的計算の方法

目標記述とエージェントの関係

GDU においてエージェントと関係をつなぐ際に参照される項目は *Proposition* と *Modality* である。*Proposition* や *Modality* とあるエージェントの意味情報記述が単一化するとき、その GDU とエージェントは関係をつなぐ。GDU の *Proposition* と関係をつなぐエージェントを *Proposition* エージェント、*Modality* と関係をつなぐエージェントを *Modality* エージェントと呼ぶ。全ての目標記述とエージェントの間での活性エネルギーの授受の量には、重み係数 g が掛け合わされる。

エージェント同士の関係

エージェントに記述されている下位範疇化素性は、ある語とある語が主辞と補語の関係にあることを示す。図 2 の (a) と (b) において、エージェント「が」の下位範疇化素性情報とエージェント「野茂」の記述全体は単一化する。ただし、 X , Y は変数、NIL は匿名変数である。このとき両者の間には、「が」が主辞、「野茂」が補語の関係が結ばれる。

発話状況記述とエージェントの関係

発話状況記述には、すでに発話産出行為として実行、すなわち発話されたエージェントが記述される。先に発話を達成したエージェントが補語、そのエージェントに続く発話を達成しようとするエージェントが主辞で単一化が可能な場合、両者は関係をつなぐ。全ての発話状況記述とエージェントの間での活性エネルギーの授受の量には、重み係数の s が掛け合わされる。

2.3.3 戦略選定知識と戦略知識

Proposition エージェントは発話の制限時間や他の要素との関係を判断材料として、戦略選定知識を用い戦略を選定する。戦略知識は活性エネルギーの授受の方向と量を決定する。戦略は 3 種類ある。これに対して *Modality* エージェントは、常に 1 つの戦略しかとらない。すなわち、以下で説明する共同体的戦略である。この *Proposition* エージェントだけが自律的に戦略を選定するという設定は、発話産出行為にお

いては[命題+モダリティ]を発話の単位として考え、そのなかで[命題]が核となって発話産出行為を組織化する、という考え[7]にもとづく。

Proposition エージェントのもつ戦略選定知識を以下に示す。

```

if(TimeThreshold < TimeLimit){
  Community Orented Strategy
}else{
  if(ConnectionThreshold <
    EnvConnections){
    Selfish Strategy
  }else{
    Altruistic Strategy
  }
}

```

TimeLimit は発話産出行為の制限時間、EnvConnections は注目しているエージェントが環境記述と結んでいる関係の数を示す変数であり、ConnectionThreshold、TimeThreshold はそれぞれエージェントと環境記述の関係の数、制限時間の各パラメタに対する閾値である。閾値よりも制限時間が大きい、すなわち残り時間が十分にある時は共同体的戦略がとられる。これに対して十分な残り時間がない、つまり閾値よりも制限時間が小さい時は、利他的、もしくは利己的戦略がとられる。利他的戦略と利己的戦略の選択は、環境記述との関係の数が閾値を越えている場合が利己的戦略、越えていない場合が利他的戦略となる。環境記述との関係をより多く結んでいるエージェントはその状況において重要なエージェントであり、そうでないエージェントよりも優先される。それぞれの閾値は個々のエージェントで独立に設定できる。

戦略知識について以下に示す。

共同体的戦略知識

共同体的戦略をとるエージェントは、関係をもつエージェントが文法的関係にもとづいて順番に活性化されることを目標とする。そのため、文法的に見て補語の関係にあるエージェントに活性値を送る。その量は

$$\frac{\text{自分の活性エネルギー保有量}}{\text{下位範疇化素性情報の組数}} \times c \quad (1)$$

である。c は重み係数である。下位範疇化素性情報の組数で送出する活性エネルギーを割っている理由は、個々のエージェントの活性エネルギーの入出力において、エネルギーの保存則を保つためである。複数の下位範疇化素性を持つ場合には、送出エネルギーの量を下位範疇化素性の組数で割ることにより、送出される活性エネルギーの総量は送出元のエージェントの保持していた活性エネルギーと等しくなる。

利己的戦略知識

利己的戦略をとるエージェントは、他のエージェントより自分が先に活性化することを目標とする。そのため、そのエージェントから見て補語の関係にある全てのエージェントから活性エネルギーを奪う。その量は

$$\text{奪う相手の活性エネルギー保有量} \times sl \quad (2)$$

である。sl は重み係数である。

利他的戦略知識

利他的戦略をとるエージェントは、他のエージェントを自分より優先して活性化することを目標とする。そのため、そのエージェントから見て主辞の関係にある全てのエージェントに対して活性エネルギーを送出する。その量は

$$\frac{\text{自分の活性エネルギー保有量}}{\text{他のエージェントと結んでいる関係の数}} \times al \quad (3)$$

である。al は重み係数である。この式において送出するエネルギーの量を他のエージェントと結んでいる関係の数で割る理由は、式 2.3.3 の場合と同じく、個々のエージェントにおける活性エネルギーの入出力の総量を等しくするためである。

本章であげた g, s, c, sl, al の各係数と前章であげたエージェントの実行条件であるエネルギー保有量に対する閾値 t は、組織化の全体的な傾向に影響を与える変数である。これらの変数をグローバル変数と呼ぶ。

2.4 動作例

システムのおおまかなながれを、図3、図4を用いて示す。図3は制限時間の制約が比較的緩いため、全ての

Proposition エージェントが共同体的戦略をとった場合、図4は制限時間の制約が厳しく、利己的戦略をとる *Proposition* エージェントと利他的戦略をとる *Proposition* エージェントが混在する場合である。図3および図4において、Goal Description は目標記述を、Spoken Words は発話状態記述を、Society of Agents はエージェントの集合を意味する。グローバル変数のバランスは組織化の傾向に影響するが、説明の簡略化のため、それらのバランスは文法的な関係、つまり「主辞」→「補語」の関係を重視する設定であるとする。具体的には、 $g = 1.0, s = 1.2 > c = 0.8 > sl = 0.6 > al = 0.4$ である。また閾値は $t = 100$ である。

全てのエージェントが共同体的戦略をとった場合の動作のながれを、図3で説明する。最初 (a) の状態では目標記述には

```

GDU1
Proposition: Nomo
Modality: ga, ne
ActivationLevel: 100

GDU2
Proposition: Get
Modality: [past], da, yo
ActivationLevel: 150

GDU3
Proposition: Swingout
Modality: wo, ne
ActivationLevel: 50

```

の3つのGDUが存在し、それぞれに対応するエージェントと関係を結んでいる。また、個々のエージェントに記述された局所的な文法情報にもとづいて「主辞」→「補語」の関係が $ne \rightarrow ga, ga \rightarrow Nomo, ne \rightarrow wo,$

$wo \rightarrow Swingout, yo \rightarrow da, da \rightarrow [past], [past] \rightarrow Get, Get \rightarrow ga, Get \rightarrow wo$ に結ばれている。これが (a) の状態である。まず、目標記述からエージェントへと活性エネルギーが送出される。活性エネルギーを受けとったエージェントは共同体的戦略にもとづいて補語の関係にあるエージェントへと、活性エネルギーを送る。その結果、矢印の方向をたどって活性エネルギーは *Swingout* と *Nomo* に集まる。ここでGDU1のほうがGDU3よりも *ActivationLevel* の値が大きいため、*Nomo* が *Swingout* よりも多くの活性エネルギーを保持するので、*Nomo* が選択され、発話される。発話された *Nomo* は状況記述に記述され、発話を完成するために続くエージェントに活性エネルギーを送る。これが (b) の状態である。そして ga, ne が次々と選択され、GDU1の目標が達成された状態が (c) である。同じように今度はGDU3の目標が達成され ((d) の状態)、最後にGDU2の目標が達成され、全ての発話が終了する ((e) の状態)。

このとき得られた発話文は、

「野茂がね、三振をね、獲ったんだよ」 (α)

である。もとの目標記述の状態では、発話目標の力関係は

$$GDU2 > GDU1 > GDU3 \quad (4)$$

であったにもかかわらず、目標記述間の文法的な関係を反映した発話になっている。

つぎに図4を用いて、利己的/利他的戦略を選定したエージェントが混在する場合の動作の流れを説明する。ここでは図3の場合と同じ内容の3つのGDUがある。エージェントの数および関係の数も同じであるが、*Get* が利己的戦略、*Nomo* と *Swingout* が利他的戦略をとっている。このためエージェント同

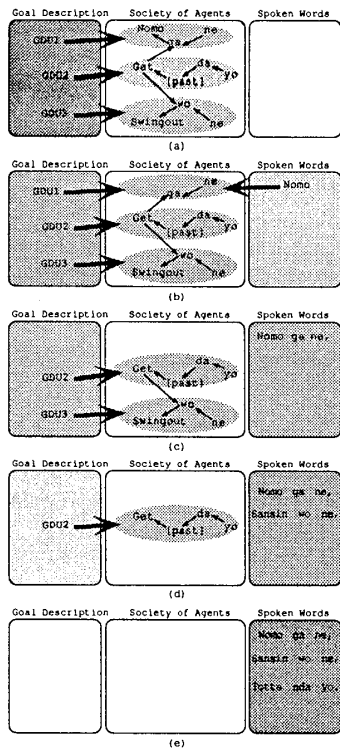


図 3: 共同体的戦略による動作の例

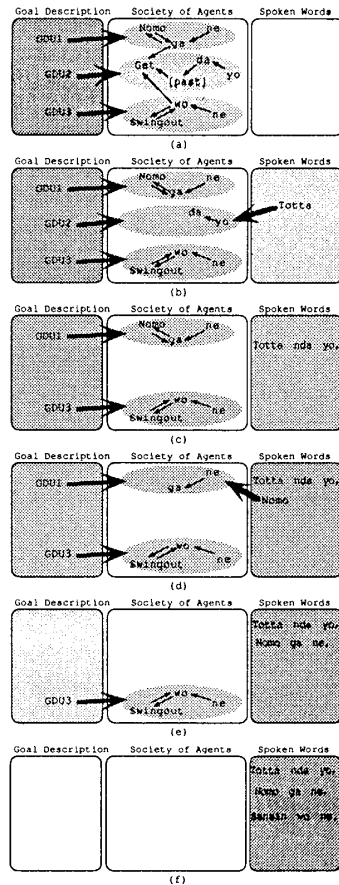


図 4: 利己的/利他的戦略が混在する場合の動作の例

士での活性エネルギーの送出方向が図3とは異なっている。ここでは関係の方向が *Get* に集中しているため *Totta* が発話され ((b) の状態), 続いて *nda, yo* が発話されている ((c) の状態)。この時点で3つの GDU を結んでいた関係は断たれるため, 残された GDU1 と GDU2 は独立に活性エネルギーの授受を行なう。ここでは GDU1 の *ActivationLevel* の値が GDU3 よりも大きいことを反映して *Nomo* が発話され ((d) の状態), 続く発話を完成させ ((e) の状態), 最後に GDU3 の目標が達成されている ((f) の状態)。この場合も発話目標の力関係は式 (4) のとおりであったにもかかわらず, 利己的戦略をとったエージェントを含んだ発話目標 GDU2 が優先して発話された結果,

「獲ったんだよ, 野茂が, 三振を」 (β)

という, (α) と比較して倒置の発話文が組織化されている。

2.5 今後の展開

前章までで示したように, 現在のモデルの発話産出能力は極めて限定されたものである。副文などの埋め込み構造や, 心的辞書の検索機構, さらにモダリティの処理能力など, 汎用の発話産出計算モデルとして備えるべき能力は, まだ備わっていない。これらの能力を実現した計算モデルは英語などの分析をもとにしたものがすでにいくつか存在しており [1], これらが日本語についても適用可能かどうかの検討を進めている。また, 漸次的発話産出における並列動作と局所的情報処理が必然的に孕む問題である計算の失敗や情報の不整合の発生に対する自己修正能力が, 現在の我々の最大の関心である。この自己修正活動は「言い直し (self repair)」という極めて日常的な発話として表層化する。この言い直しについては, 以下に

述べるように我々自身で心理言語実験を行ない、分析を進めている。この心理言語実験によって得られる様々な知見を、モデルに反映させることが、次のモデルの拡張目標である。

3 心理言語実験

漸次的発話産出においては、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 [3, 6] 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、漸次的発話産出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教示されている。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産出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。

3.1 実験

発話の収録は防音室にて行ない、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 Machintosh 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。教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）を開始することが教示された。また、教示において提示される全ての絵のマテリアルが提示され、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、単語のわかち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、道具のスパン、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）
- 動作主の登場位置（右側か左側か）
- 動作主の性別（男か女か）
- 動作主の動作（叩くか押えるか）
- 動作主の使用する道具（スベナか傘か）

★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）

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る。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（下記参照）、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。

Subject(NorNP) + Case - particle(PP) → Object(NorNP) + Case - particle(PP) → Verb(VorVP) + Voice - particle" RERUorRARERU"

↑PPとして今回の実験で使用さ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)たたく(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。

3.2 分析

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産出において、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を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

3.3 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たいと考える。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。

3.4 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。

4 おわりに

筆者の所属する研究グループで行なわれている、発話産出の計算モデル構築の試みと、モデルへの反映を目指しての心理言語実験の概要を紹介した。まだ結果が完全にまとまっていないので本稿へ記載しなかったが、心理言語実験においては複数の言語運用的／言語能力的制約の多重関与による、複雑な言い直し形成が行なわれていることがわかってきている。こうした制約の多重関与は、機能モジュールの相互作用という我々の計算モデルで採用している形式へと還元可能なものであると考えている。今後はこのようなエンピリカルな分析結果の計算モデルへの反映を進めたいと考えている。

参考文献

- [1] T. Dijkstra and K. de Smedt. *Computational Psycholinguistics*. Taylor & Francis, 1996.
- [2] T. Gunji. *Japanese Phrase Structure Grammar*. Reidel, Dordrecht, 1986.
- [3] W. J. M. Levelt. *Speaking: From Intention to Articulation*. MIT Press, 1989.
- [4] P. Maes. Situated agents can have goals. In *Robotics and Autonomous Systems*, volume 6, pages 49-70, North-Holland, 1990.
- [5] C. Pollard and A. Sag. *Head-Driven Phrase Structure Grammar*. The Univ. of Chicago Press, 1994.
- [6] C. van Wijk and G. Kempen. A dual system for producing self-repairs in spontaneous speech. *Cognitive Psychology*, 19:403-440, 1987.
- [7] 益岡 隆志. モダリティの文法. くろしお出版, 1991.