

プラン  
地図を携えない発話生成の研究

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)  
〒 619-02 京都府相楽郡精華町光台 2 丁目 2 番地  
hajiri@mic.atr.co.jp http://www.mic.atr.co.jp/~hajiri/

概要: 音声対話における発話はじつに多種多様である。対話における自然な発話を生成するためには、柔軟な言語処理技術が必要である。本研究は複雑系や人工生命研究において議論されている "Simple Rules to Complex Behavior" というパラダイムを積極的にとり入れたダイナミックな言語処理を用いて、自然で状況に対して敏感な発話を生成することを目指している。我々は、発話は複雑な認知過程を経て創発する言語のパターンである、と考えている。これは従来の論理的・記号主義的な言語処理技術とは一線を画する考えであり、いわゆる Good-Old-AI と New-AI のような対比が出来ると言えよう。従来の発話生成を精密な地図 (プラン) を携えての慎重な行動計画と例えるなら、我々は精密な地図 (プラン) を携えない代わりに自由で柔軟な行動として発話を創り出すことを目指している。

キーワード: 発話生成, 自然性, 創発

## Study on spontaneous utterance generation without maps

Koichiro HAJIRI (ATR Media Integration & Communications Research  
Lab./Ritsumeikan Univ.), Michio OKADA (ATR Media Integration & Communications  
Research Lab.), Hitoshi OGAWA (Ritsumeikan Univ.)  
2-2 Hikaridai, Seikacho, Sorakugun, Kyoto, 619-02 JAPAN  
hajiri@mic.atr.co.jp http://www.mic.atr.co.jp/~hajiri/

Abstract: Utterances in speech communication is very rich in variety. To generate the spontaneous utterance, a flexible NLP technology is needed. In our research, we use a dynamic NLP technology, which is affected from the paradigm of "Simple Rules to Complex Behavior", that comes from Artificial Life and Complex Systems study, to generate situated utterances. We view that utterances are emergent patterns of cognitive process as a complex system. This viewpoint is quite different from ordinal logical and symbolic NLP technic. This can be viewed as the difference of Good-Old-AI and New-AI. Metaphorically speaking, the ordinal utterance generation is like a action planning with a detailed map, and we focuses on generating free and flexible utterances without detailed maps.

Keywords: utterance generation,spontaneousness,emergence

### 1 はじめに

M. マクルーハンはその著書「メディア論」[1]の中で、メディアを"ホット"と"クール"の2種類に分類している。彼のいうホットなメディアには映画やラジオ、講演(独話)などがあり、一方クールなメディアとしてテレビや電話、そして対話などがあげられている。彼の言葉で説明すると、ホットなメディアとは単一の感覚を高精細度で拡張する機能を持っているという。それに対してクールなメディアとは低精細度であるために、聞き手による補間が必要になる。つまり、

クールなメディアはホットなメディアに比べて受容者の参与性が高いのである。当然のことであるが、両者の違いはメディアとしての性質が大きく異なることに起因している。つまり、同じ音声を媒体としていても独話と対話はまるで違うのである。

しかしながら、これまでの多くの音声対話システムは短時間の独話が二人の人間によって交互に行なわれることを対話と見做していたのではないだろうか。つまり、音声処理研究においては音声、自然言語研究においては言語という媒体そのものに対する研究が先行しており、その影響で対話に対しても媒体主導の研究

アプローチがとられた。その結果として明瞭な発話や明確な構文こそが基準であり、他はそれから逸脱したものである、という「硬い」対話観が蔓延し、人間の本来持つ柔軟性や自然性という貴重な性質を捉えにくくしてしまったのではないだろうか。ホットなメディアのために培われた技術をクールなメディアに適用しようとしたために、ミスマッチが起こっているのである。マクルーハンのような視点からもう一度音声対話を見直せば、「書き言葉」から「話し言葉」へ、もしくは「独話」から「対話」へといった単純な拡張は効かず、それらの間には本質的な問題が横たわっていることを認識できるのではないだろうか。

ホットなメディアは収束、単一化、中央制御、ブロードキャストなどの特徴をもっており、クールなメディアは発散、多様化、偏在制御、ローカルインタラクションなどに特徴づけられる。講演が一人の意見を拡声器で多数の観衆に伝えるのに対して、対話は個々人がめいめいに目的や指向を持って会話に臨むものであり、その方向性や結末は一個人が管理できるものではない。

このように自然な音声対話の性質をよく考えてみると、いわゆる Traditional AI の目的指向性や手段目標解析とそれに基づくトップダウンな設計などが苦手とするものばかりが目立ち、むしろ新しいパラダイムである分散協調、局所性、創発計算などに基づく New AI の問題として捉えるべきである、という方向性が見えてくる。

本稿ではこのように新しい AI のパラダイムである雑系や人工生命研究で提唱されている "Simple Rules to Complex Behavior" [8] という考えに基づいたダイナミクスによる自然言語処理、特に発話生成モデルへの適用という試みについて論じる。

## 2 自然な発話という複雑な問題と非決定的なアプローチ

自然な発話とは何なのだろうか。再びマクルーハンの言葉を借りれば、それはクールな言葉の列である。

従来からの発話生成モデルの多くは、発話を問題解決過程として捉えていた [3]。発話目標が明確に存在し、それを言葉の列として組み立てるための文法規則を積極的に用いていた。規則を問題解決におけるオペレータもしくは解決手続きとして適用し、徐々に発話という目標を実現していく。このような視点には、目標の達成度が高ければ高いほどよいという考えが暗黙のうちに込められている。目標を精密かつ正確に達成することを目指し、その過程には最適性が求められる。

これはまさにホットな視点であり、結果として従来の様々な発話生成モデルで生成される発話は一貫性や収束性、結束性などを追求することを良とする了解

が明に暗に存在していた。いや、実際にはそれ以外に工学的に健全なアプローチが思い浮かばなかったのかも知れない。しかしながら、真に自由な発話は、最適性というホットな定規で計ることの及ばない、混沌としたものである。例えば自由な対話を分析すれば、たちまちのうちに冗長語と呼ばれる分類不能で機能や目的が一見して不明の表現や、言い直しや言い間違いなどの現象が大量に見つかる。これらの "無駄" な言葉遣いは従来、あくまで短期記憶など認知能力の有限性に根ざした失敗にすぎず、最適な発話からの逸脱と捉えられていた。その結果、対話相手の人間のためにわざわざ逸脱処理を盛り込んで疑似的な自然性を実現する、といったことが行なわれ、計算機に発話生成メカニズムを実装する際には "無駄" は基本的に必要ないとして最適性の追求だけがなされていた。

しかし本当に発話生成モデルの開発は最適性の観点から行なわれるべきなのだろうか。多様で柔軟な発話を生成する高度な能力が、冗長な言葉や言い間違いをする低度な能力と同居しているのだろうか。むしろ両者はある能力の表と裏の顔であり、じつは一つのシステムの振舞いを観察者が主観によって分類しているにすぎないのではないだろうか。ここにも他分野で議論されているデカルト的二元論の綻びを見い出すことが出来る、と我々は考える [2]。

類似した問題として進化論における多様性と適応能力の関係がある。進化とは環境への適応という複雑な問題を解くために自然が生み出したひとつの戦略である。ある環境に最適に適応するだけの戦略は、環境の変化に対して脆弱である。むしろ環境に不適応な種を多数用意しておけば、全体として多様な環境に適応出来る可能性が上がる。しかし、不適応を単純に無駄と見做すなら、多様な環境への適応能力を否定することになる。

もちろん、進化論と言語の認知過程が別の問題であることは明白である。単純なアナログが通用するとは思えない。また、進化論自体も壮大な仮説に過ぎず、その信憑性も定かではない。しかし、無駄と多様性、柔軟性が密接な関係にあることを認識するには十分である。

進化の過程こそ、クールそのものである。進化は発散、多様化、偏在制御、ローカルインタラクションに満ちている。そして、クールな技術とは進化の過程と同じく、最適性というホットな定規で計ることが出来ない。なぜなら、クールな技術とは非決定性を有し、完全性と引き替えにその挙動の柔軟性を手にしているからである。我々に求められているのは、最適性を見直し非決定性を受け入れることである。非決定的ながら合目的に動くシステムこそが、無駄を有しつつも柔軟に動作することができるのである。

本章ではこのように非決定的ながら合目的に動く

システムの根幹にある、創発という考え方について述べ、創発計算による発話生成の性質について考える。

### 3 発話生成における創発のダイナミクス

発話とは認知過程という複雑系から創発する言葉のパターンである、という考え方は魅力的であり、前章までの議論のように、真に自然で自由な発話は線形なメカニズムで形成されるのではなく、複雑系の非線形なメカニズムが生み出していると考えられる。ただし、従来の発話生成モデルに非線形性を組み込めばいいという、単純な話ではない。発話生成という複雑系は多様な諸要素によって構成されるべきである。本章では多様な発話生成における諸要素と発話生成行為の相互作用について述べる。

#### 3.1 思考と発話生成行為の一体化

従来からの発話生成モデルでは、大きく分けて What-to-Say と How-to-Say という二つの要素を想定していた。さらに最近では漸次的な動作を考慮した発話生成モデルの研究がいくつかなされており、そこでは When-to-Say という新たな要素を考慮することが提案されている。

こうした一般的な考え方とは別に、マクニールなどの発話生成に対する独特の捉え方が存在する [5]。ひとことと言うならば、思考と発話生成行為の一体化である。何を発話するべきか (What-to-Say) と、どのように発話するべきか (How-to-Say) の関係は実に微妙であり、両者の間には密な相互作用が働いている。発話生成行為は実時間の現象であるため、当然両者の相互作用には時間的要素 (When-to-Say) も密接に関係してくる。これらの諸要素を異なった記述で扱い、しかる後にそれらの間の相互作用を考えるという、いわゆる分割統治の考え方は Traditional AI では至極自然な考え方であるが、それによって実現されるシステムの出来を芳しく思わない人々が多々いるのも確かである。では、どのようなアーキテクチャを用いればいいのか。どのような枠組ならば思考と発話生成行為を一体化できるのか。次章では発話環境と発話生成行為の相互作用という観点から、両者の統合について述べる。

#### 3.2 発話環境と発話生成の相互作用

極論かもしれないが、発話生成行為と思考は渾然一体となったものである、と我々は考えている。発話することは考えることであり、考えは発話を駆動する。何かから切り出せばよいか分からない時に、とりあえず何か言葉を発すれば、関を切ったように思いが言葉として溢れ出す、という経験は誰もが持っているはずである。そのような発話に勢いや独特の説得力を感じる

こともしばしばある。これこそ自然な発話の好例である。このような場合、文法的な規則を強く意識することはまれであり、むしろ頭に浮かんだ言葉を次々と発話するだけである。それでも、意味をなさない言葉の並びがやたらと出てくるわけではなく、やはりそれらには何らかの秩序が見受けられる。この秩序こそ、自然な発話でも最低限守られなければならないものである。しかしながら、秩序と自然さの共存をどのように進めればいいのかという問題がある。そこで我々は発話生成行為の創発的な複雑系のメカニズムを考えることにした。創発とは混沌からの秩序の発生であり、複雑系のなかに秘められた、局所的な相互作用から秩序を生み出す力がそれを実現しているのである。

こうした複雑系としての発話生成モデルを考えるとき、純粋に言語的なリソース、つまり文法や辞書などについて考えるだけでは不十分である。そこで我々は、発話環境を扱うことにした。発話環境とは認知的制約のみならず、さらに広い意味で発話生成行為に関係する様々な要素のことである。環境という言葉状況を置き換えることも可能であるかもしれない。例えば、先に発話された言葉が次に続く発話を”引き出す”といった現象においては、先行する言葉の役割が重要である。また、対話の履歴も環境となる。先に対話相手の喋った印象的な言葉が次発話に与える影響は、表層的な修飾語の選択などのレベルに留まらず、発話目標にも大きな影響を与え得る。さらに、発話目標は発話の大まかな内容を保持し、発話生成行為に大きな影響を与えるという、発話環境の主要素である。ただし従来からの発話生成モデルでは発話目標は宣制君主のように絶対的な存在であったが、実際の発話においてもそうではない。むしろ発話目標が他の要素との相互作用によって形を変えようといった柔軟性が、自然な発話のためには重要である。そして発話者や対話相手の内的状態や指向、感情などが発話の形成に密接に関わることは様々な研究から指摘されており、これらも重要な環境である。

発話が生成される過程を認知的に観察するならば、短期記憶などの認知的性質もしくは制約を意識することが重要である。文法が無制限の文章を生成できる能力を有していても、実際の発話ではそのようなことはあり得ない。それゆえ、こうした実世界の認知的制約を考えることは、自然な発話の生成において重要である。しかし、こうした様々な発話環境を構成する要素のそれぞれについて、これまでの研究は独自の記述とメカニズムを用意している。なかには洗練されて汎用性の高いものもあるが、全てを統合できるような一般的な記述が現時点では不可能であることは、周知の事実である。我々の立場では、こうした様々な要素の様々な記述は、全て発話生成という複雑系の局所的なリソースとなる。例えば対話の履歴と次発話を関係づ

ける記述があれば、両者の局所的な相互作用が計算できる。しかし、それが全てではない。他の要素との局所的相互作用も同様に計算され、それらの総和としてひとつの大局的な秩序が形成された時にはじめて、自然な発話となるのである。

### 3.3 記述形式とダイナミクス

一般的な言語モデルにおいては、大局的な秩序の記述が随所に見られる。統語解析における書き換え規則などはその典型である。しかし、より複雑な統語構造を柔軟に解析するため、またその際の計算オーダの爆発を防ぐという観点から、HPSG[6, 7], JPSG[4]など、局所的な文法的規則の記述が提案されている。これは発話生成行為の基本的な単位同士の局所的な関係を求めるための重要なリソースとなり得る。その最底辺に存在する記述は語彙の素性情報である。素性情報を言語のアトムと考えることで、複雑系の最底辺の計算対象とするのである。もちろん、アトムの下位レベルにクォークやストリングがあるように、素性の下位記述も十分に想定できるであろうが、発話環境と発話生成行為の間で十分に柔軟な関係の計算が出来るかどうかの問題であり、その点で素性情報は完璧とまではいわないが、かなりいい粒度である。

さて、JPSGなどの文法的な素性情報は語彙、すなわち発話生成行為の基本単位の文法的な関係を計算するのに十分であるが、環境要素との関係を計算するにはどのような素性情報が必要かを考えてみると、それは、環境をどのように記述するかに依存する。状況意味論などの記述は命題論理や述語論理を拡張したものである。従来の発話生成モデルでは発話目標などは独自のオペレータを論理的記述で表現したものが多く、意味ネットワークなどの幾何的な知識表現も論理的な表現に変換することで計算機に実装されている。最近では意味素性と文法素性の統合も行なわれているが、それは素性集合とその計算のレベルにおいてであり、プリミティブとしての素性には依然として論理的表現が用いられている。このような背景から、我々の発話生成モデルにおいても環境要素の記述は論理的表現を採用している。しかし、これは全ての環境要素を一括した論理表現で記述することを意味するわけではない。局所的な一貫性が保たれればよい。

こうしたプリミティブな記述の枠組を決定することにより、とりあえず局所的な計算が可能となる。具体的には単一化や関数などのプリミティブに対するオペレーションを明確に定義し局所計算を繰り返すことで、部分の総和として大局的な挙動が求められる。しかし、これだけでは発話を創発するモデルとしては十分ではない。創発とは下位の記述から上位の挙動を得ることであり、記述のレベルと観測できる顕著な現象との間にある程度のギャップが存在することを意味す

る。これに対して発話環境と発話生成行為のインタラクション、発話生成行為同士のインタラクションが局所的な計算として定義されているだけでは、結果として得られる発話はたかだか局所計算の単純な演繹にすぎず、多様な発話を生み出す能力とはなりにくい。また、大局的な一貫性のない偏在制御では、恣意的に与えられた要素間の関係計算の能力や記述形式、およびその内容が、全体を一方向的に制御してしまう可能性がある。これでは、生成される発話の多様性が狭められてしまう。“Simple Rules to Simple Behavior”もしくは“Complex Rules to Complex Behavior”であり、システムの記述の複雑さと挙動の複雑さの相関関係が1次直線を描くだけである。

ここで複雑系に再び注目すると、そこにはダイナミクスの非線形性が多様性を生み出すという特徴が見い出せる。ダイナミクスとして代表的なものはシステム内の各要素と環境情報との関係の導出、および活性/抑制などの関係を連続量で表現している。これはコネクショニズムから複雑系研究へと脈々と受け継がれているアナログ指向に根ざした技術であり、複雑な情報の関係を明瞭なパラメタに集約することで、統合的で非明示的な挙動を実現することが出来る。発話生成の局所的な計算のダイナミクスとしては、単一化や関数の機能に連続性を導入すること、発話環境の諸要素および発話生成行為の間に連続量によるメタ制御のメカニズムを導入することで解決する。具体的には環境要素は活性源として、また発話生成行為の基本単位間は活性の授受を関係に基づいて計算することで実現できる。各環境要素はそれぞれ独自の活性送出の「強さ」のパラメタを有する。発話生成行為の個々の基本単位は他の基本単位との関係や環境要素との関係から、活性の授受を能動的に行なう。こうした活性値という連続量の授受を通して得られるそれぞれの行為の活性の強さから、それぞれの状況、すなわち環境の状態や発話生成行為の基本単位の状態にふさわしい行為の選択が可能となる。これは明示的な競合解決規則やメタ規則によって制御されるものではない。連続量は汎用性の極めて高い情報表現であるため、活性授受においてどのように局所的な動作を設定しようとも、それは連続量の値に還元されるにすぎない。

このような連続量の一元性と関係計算の非線形性から、局所的な相互作用の総和が複雑な挙動を創発することが可能となるのである。

## 4 Spontaneous Speech の生成モデル

本章では今回実装した発話生成モデルの概要を説明する。

## 4.1 モデルの諸要素

今回実装したモデルでは、大きく分けて3つの要素が存在する。すなわち、発話目標記述 (Goal Description Units=GDU), 発話生成行為の基本単位の集合 (Behavior Objects for Utterance=BOU), そして先に発話された言葉の記述 (Spoken Words Descriptions=SWD) である。それぞれの要素は独自に指向性を持つようにデザインされ、それぞれが原目標 (Proto-goal) を持っている。発話生成行為に貢献するという原目標の達成に向けて、他の要素と協調や競合をするのである。GDU と SWD は発話生成行為を取り巻く”環境”であり、BOU を活性化させる役割を担っている。これが環境の基本的な指向性である。これに対して BOU は個々が発話されることを原目標として保持しており、自分が発話されるために必要な条件を認識できる。条件を認識した結果としてどのような振舞いをすれば自分が発話されるかを自律的に判断し、その場その場での振舞いを決定する。この自律的な判断とそれに基づく振舞いの決定は個々の”戦略”という概念に基づいて実装されている。

以下ではそれぞれの要素の具体的な内容を示しながら、より詳しい説明を行なう。

### 4.1.1 GDU

以下に GDU の書式を示す。

[*Prop* : *Nprop* : *ActiveLevel*]

ここで、*Prop* は発話の核となる命題情報、*Nprop* は命題以外の情報を意味する。*ActiveLevel* 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る。命題以外の情報とは、それら発話の核となる情報以外の助詞、終助詞、モダリティなどである。

GDU には意味的な集合とその集合の前後関係が設定できる。これは、必ずしも発話の順番を規定するわけではない。なぜなら、環境や BOU 同士の相互作用によって適応的な発話文の形成がなされる時、適応の度合と GDU の集合間関係のどちらが優先されるかは、モデルのダイナミクスによって動的に変化するからである。ただし、GDU の集合間の前後関係が組織化にどの程度影響するかは、モデルのパラメータとして設定することが可能である。そのため、決定的ではないが、組織化におけるダイナミクスに大局的な影響を与えることは可能である。この GDU の集合間の意味的な前後関係が、発話目標における談話構造である。

### 4.1.2 BOU

発話生成における組織化において我々が想定した小さな発話生成の基本単位 BOU は、いわゆる心的辞書

の項目に対応する。具体的には語彙が BOU に相当する。それぞれの BOU には HPSG[6] や JPSG[4] に基づく局所的な文法的情報や意味情報、音韻情報などが記述されている。また、BOU が環境や状態に応じて自律性を発揮できるように戦略選定知識と戦略知識も記述されている。戦略選定知識は個々の BOU のおかれている状況、すなわち環境との関係や他の BOU との関係、さらに時間的制約や BOU 自身の内的状態を参照し、ふさわしい戦略を選定するのに用いられる。

現在のモデルでは発話に対する制限時間を制約として設けており、この時間制約が十分に緩い場合は共同体的戦略、時間的制約が厳しい場合には利己的/利他的戦略のいずれかが選択される。利己的戦略は環境との関係の数が一定以上の場合で、一定以下の場合には利他的戦略がとられる。この戦略選定知識の設定は、時間的余裕がある場合には落ち着いて話すことで文法的に正しい文章を話そうとし、時間的余裕のない時には大切なことを簡潔に話そうとする、という特性を表現している。大切なこととは、より多くの環境要素から支持されていることである、との見地から利己的/利他的戦略の選定方法を上記のように設定している。

戦略知識は数種類が設定でき、それぞれが異なった BOU の振舞いをもたらす。本モデルでは3種類の戦略、すなわち共同体的戦略、利己的戦略、利他的戦略のいずれかが選択される。共同体的戦略は文法的な関係を守るためにデフォルトの方向に活性エネルギーを送る。利己的戦略は周辺関係を持つ諸要素から活性エネルギーを奪い、いち早く自分が活性化される。利他的戦略はその反対で、他の諸要素を活性化するために、自分の保持する活性エネルギーを関係をもつ周辺の BOU に分配する。

図1に BOU の記述例を示す。なお、図1の Local Syntactic Knowledge において *X* と *Y* は変数、NIL は匿名変数である。また、Strategy Selecting Knowledge において *ConnectionThreshold* と *TimeThreshold* は各 BOU が保持する閾値、*TimeLimit* は制限時間の残りを、*EnvConnections* はその BOU と環境との関係の数を示す変数である。

### 4.1.3 SWD

SWD には先に発話された言葉が記述されている。先に発話されたとはこの場合、1回の発話において漸次的に形成される発話文の単位ステップ毎の記録である。例えば言いたいことが明瞭に意識できないような時でも、ひとこと言葉を発すれば次々と続く言葉が溢れ出す、といった経験は誰もが持っていると思われる。このとき、ひとこと先に発話された言葉は環境として発話主体に働きかけ、次に発話すべき言葉を引き出す作用をもっている、と考えることができる。このように SWD は環境として短期的に作用する発話され

```

‘Local Syntactic Knowledge’
WORD: 獲-r
HEAD: [POS: v, FORM: root, GR: NIL]
SUBCAT: [HEAD: [POS: p, FORM: ga, GR: sbj],
          SUBCAT: NIL,
          SEM: _X]
[HEAD: [POS: p, FORM: wo, GR: obj]
          SUBCAT: NIL,
          SEM: _Y]
SEM: Get, sbj:_X, obj:_Y
‘Strategy Selecting Knowledge’
if(TimeThreshold < TimeLimit){
  Community Oriented Strategy
}else{
  if(ConnectionThreshold < EnvConnections){
    Selfish Strategy
  }else{
    Altruistic Strategy
  }
}
‘Strategy Knowledge’
Community Oriented Strategy:
  send activation to daughter BOU
Selfish Strategy:
  snatch activation from connective BOU
Altruistic Strategy:
  send activation to connective BOU

```

図 1: BOU の記述例

た言葉の影響を反映することのできる存在である。例えば倒置文などは SWD などの存在を仮定することで言語使用の視点から説明できる現象である可能性が高い。人が自然に発話を行なっている時、意識せずに倒置文が生成される背後には、認知過程において通常文から倒置文への変形が行なわれていると言うよりは、SWD のような存在が“結果として”倒置文を引き出した、と考える方がより自然であると思われる。このように、SWD という環境を想定することだけでも、多様な発話における現象をモデル化することが可能である。

## 4.2 動作過程

これまで説明してきたモデルの諸要素が、複雑なインタラクションをすることで自然な発話生成が行なわれる。本節では諸要素のインタラクションのメカニズムを説明することでモデル全体の動作過程を示す。具体的には各要素間の関係の計算と、その関係に沿った活性流の計算が行なわれることで環境と発話生成行為の間の相互作用や、発話生成行為の複雑な組織化が成される。

### 4.2.1 関係導出計算

各要素間では「関係」という概念に基づくコネクティビティを求めることが出来る。この関係の導出は各要素の記述内容の相互参照によって実現される。具体的にはユニフィケーションや代数的・論理的記号演算を用いることで、相互の情報の関連性を求め、その関連性の度合に応じて重みを持った「関係」が確立さ

れる。関係導出の際に相互参照される記述の組み合わせは以下のようにになっている。

- GDU と BOU : GDU の Prop もしくは Nprop と BOU の SEM 素性
- BOU 同士 : BOU における Head 素性と Sub-Cat 素性
- BOU と SWD : SWD における SubCat 素性と BOU の Head 素性

### 4.2.2 活性流によるダイナミクス

要素内の記述の相互参照により各種要素の間での関係の導出が終了すると、活性流の計算が行なわれ、状況に応じて発話される言葉が選択される。人間が発話生成行為を取り巻く環境に導かれるように自然に発話している時、環境が発話生成行為を組織化する駆動源となっていると捉えることができる。このような視点から、本モデルにおいては環境要素、今回の場合は GDU と SWD が駆動源、すなわち活性流の源となる。それぞれの GDU および SWD は局所的に確立された関係に沿って活性エネルギーを送出するのである。

活性エネルギーを受けとった BOU はそれを一旦保持し、戦略選定知識を用いて戦略を選択する。その結果として選定された戦略は BOU 同士の関係における活性流の方向と量を定める。たとえば共同体的戦略を選択した BOU は、自分から見て補語にあたる関係の BOU、すなわち自分の SubCat 素性とユニファイする Head 素性を持つ BOU に活性エネルギーを送出する。このようにして BOU 同士での活性エネルギーの授受が行なわれ、最も高い活性エネルギーを保持する BOU が発話される言葉として選択される。選択された BOU は発話され、BOU の集合から削除され SWD の集合に新たに登録される。それに伴い、発話された BOU を支持していた GDU はその目標を達成したため、GDU から削除される。このように BOU は GDU や SWD などの環境から影響を受けて組織化され、組織化された BOU は環境を変更するのである。

また、活性流は各種のパラメタによって大局的な制御される。各種関係はそれぞれの種類に応じて共通のパラメタを重みとして保持しており、その関係に沿って流れる活性エネルギーの量を増幅/減衰させることができる。これら複数のパラメタを調整することにより、GDU の影響や SWD の影響、さらに BOU 同士のインタラクションにおける各戦略の影響の大きさを調整することが可能である。これにより、組織化の合目的性や環境依存性などを大局的な操作で間接的に制御することができ、より多くの柔軟な発話生成行為を扱うことが可能となる。

## 5 動作例

以下に、ある単文を発話した場合のシステムの動作例を示す。また、併せて大まかな各要素の状態の変化を図2に示す。

(1) 「野茂、が、ね、レッズ戦、に、ね、登板し、たんだ、よ。」

GDU1:Prop:Nomo,Nprop:NOM,ActiveLevel:10

GDU2:Prop:vs\_Reds,Nprop:LOC,ActiveLevel:150

GDU3:Prop:take\_mound,Nprop:PAST,ActiveLevel:300

時間的制約: 800

(2) 「登板し、たんだ、野茂、が、レッズ戦、に。」

GDU1:Prop:Nomo,Nprop:NOM,ActiveLevel:200

GDU2:Prop:vs\_Reds,Nprop:LOC,ActiveLevel:300

GDU3:Prop:take\_mound,Nprop:PAST,ActiveLevel:10

時間的制約: 400

(3) 「野茂、レッズ戦、に、登板。」

GDU1:Prop:Nomo,Nprop:NOM,ActiveLevel:300

GDU2:Prop:vs\_Reds,Nprop:LOC,ActiveLevel:150

GDU3:Prop:take\_mound,Nprop:PAST,ActiveLevel:10

時間的制約: 320

発話文において、' ' で区切られた各要素がひとつのBOUに対応する。それぞれの動作過程を説明する。

まず例文(1)では、時間的制約が緩いため、全てのBOUは共同体的戦略をとっている。そのため、文法的に補語の関係にあるBOUすなわちあるBOUのSubCat素性にユニファイするHead素性をもつBOUに活性エネルギーが送られる。例文(1)の場合では「よ→たんだ」「たんだ→登板し」「ね→に」「に→レッズ戦」「ね→が」「が→野茂」「登板→が」「登板→に」が主辞→補語の関係である。この結果として、この場合BOU「野茂」がまず最初に発話され、SWDとなる。このとき同様の立場にあるBOU「レッズ戦」よりもBOU「野茂」が先に発話された理由は、関係に沿って授受される活性エネルギーを増幅/減衰させるパラメタの作用が原因である。SWDとなった「野茂」は次に続く言葉として「が」に活性エネルギーを送る。このようにして文法的な関係を守る方向でBOUの組織が行なわれた結果として、例文(1)のような発話文が生成されている。

例文(2)では(1)に比べて時間的制約が厳しいため、BOU「登板し」が利己的戦略をとっている。その結果としてまず「登板したんだ」という発話文が組織化された後で、BOU「野茂」とBOU「レッズ戦」の属する二つのBOUの集合の間で独立に組織が行なわれる。このときGDUにおけるActiveLevelの違いが作用した結果、「野茂が」「レッズ戦に」という順で発話が行なわれている。また、この例では終助詞の「よ」や「ね」が発話されていないが、これは終助詞が利他的戦略をとったために活性エネルギーの保

有量が過程全体を通して常に低く、ついに発話の制限時間内に活性化されることがなかったからである。

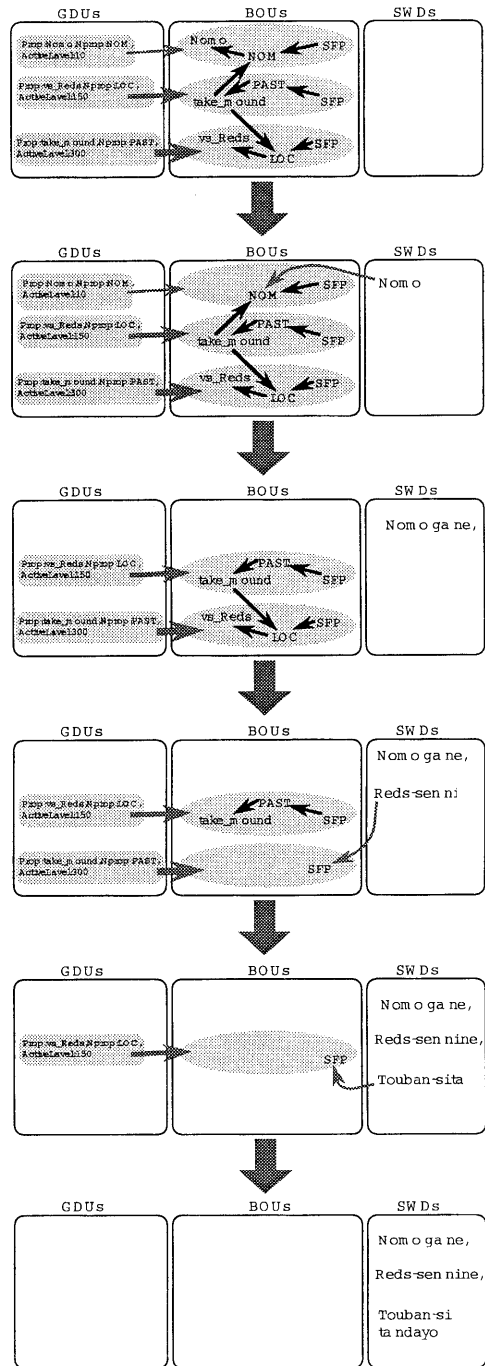


図2: 発話例(1)のモデルの動き

さらに時間的制約が厳しい例が(3)である。この場合利己的戦略をとったBOUは「野茂」「レッズ戦」「登板し」の3つで、他は全て利他的戦略をとっている。このため、必要最小限とも言える発話が組織化されている。

## 6 おわりに

自然な発話を柔軟に扱うために様々な見地から自然性や柔軟性、実世界性などについて考察し、自然な発話は複雑な認知システムからの創発であるという視点を得た。この視点から発話生成システムを考え、創発計算に基づく発話の組織化のメカニズムを提案した。

しかしながら、今回実装したシステムの生成する発話は、まだ十分な多様性と自然性をもっているとはいえない。その原因としては実装したシステムにはまだ明瞭な複雑形システムとしての特性を十分に備えていないことがあげられる。たとえば様々な要素や要素間の関係における各種のパラメタはその種類や性質が多数あり、現在はこれらを使用者及び開発者が試行錯誤的に操作することで多様なシステムの挙動を得ているのが実情である。このような問題に対しては、システムにおける学習を考慮することで大半の繁雑さが解消されると考えている。たとえば強化学習の手法を用いてパラメタの学習を行なわせることは有効であると思われる。強化学習は弱い情報源にも対応でき、教師なし学習などでも有効である。特に近年はマルチエージェントシステムにおける強化学習についての研究[9]も盛んであり、これらの研究動向を踏まえつつ効率的な学習メカニズムをモデルに盛り込むことを考えている。

また、各要素の記述内容自体もまだまだ十分に洗練されていない。HPSGやJPSGなどの局所的な文法記述をBOUの主要な文法知識として用いているが、果たしてこれらの記述のもたらす能力の全てが自然な発話において必要であるかどうかは検討の余地を残している。たとえばHPSGにおけるSLASH素性はUnbounded Relation(例えば関係代名詞などによる関係節の接続など)のために必要であるが、日常的な発話においてこのように文法的に複雑なものはあまり見受けられない。また、GDUの保持する情報量もまだ冗長である。たとえばNpropなどの情報を発話の際に意識することは稀で、むしろかなり無意識的な操作で整合性を保ちながら発話されていることが予想される。こうしたGDUに置ける過記述の問題の解決には、例えば想起順序など異なる次元の情報をうまく用いることなどが考えられる。

さらに、関係導出計算における相互参照だけで説明が出来ないような類の発話現象をどのように扱うか、という問題がある。例えば複数の主部の候補がBOUの集合に存在する場合、単純に次々と計算を行なうだ

けでは助詞などの整合性がとれなくなる。このような場合にはBOU同士でネゴシエーションを行ない、整合性をとる、と言った解決策が考えられる。このときネゴシエーションに必要な知識を各BOUに記述しては、結局記述の複雑性の増加を招くことになる。このような場合、分散制約充足問題で研究されている効率的なアルゴリズム[10]などを積極的に用いることで、出来るだけ簡潔な記述を実現するだけでなく、ダイナミクスの計算、具体的には活性流の計算によって実現されるダイナミクスを積極的に活用することで、計算過程から組織化を支援することなどを考えている。

このように今後これら複雑形システムに備えるべき要素をモデルに盛り込むことにより、自然で多様な発話を創発するシステムを実現できると考えている。

## 参考文献

- [1] M.McLuhan.「メディア論」,みすず書房,(1987)
- [2] 岡田美智男.「口ごもるコンピュータ」,情処学会編 情報フロンティアシリーズ,共立出版,(1995)
- [3] D.Appelt. *Planning English Sentences*, Cambridge University Press,(1985)
- [4] T.Gunji. *Japanese Phrase Structure Grammar*, Reidel, Dordrecht,(1986)
- [5] D.McNeill (鹿取廣人氏訳)。「心理言語学:ことばと心への新しいアプローチ」,サイエンス社,(1990)
- [6] C.Pollard (郡司隆男訳)。「制約に基づく統語論と意味論 - HPSG 入門 -」,産業図書,(1994)
- [7] C.Pollard,I.Sag. *Head-Driven Phrase Structure Grammar*,The Univ. of Chicago Press,(1994)
- [8] C.Langton(ed.). *Artificial Life I*,Addison Wesley,(1989)
- [9] 山村雅幸,宮崎和光,小林重信. エージェントの学習,人工知能学会誌, Vol.10, No.5, pp.683-689,(1995)
- [10] 横尾真. 弱コミットメント戦略を用いた制約充足問題の解法, 情報処理学会論文誌, Vol.35, No.8, pp.1540-1548,(1994)
- [11] L.Steels. Self-organising vocabularies, Proc. of ALIFE-V, Oral Presentation Proc., pp.136-141,(1996)
- [12] 小野哲雄, 東条敏, 佐藤理史. マルチエージェント・モデルによる文法の獲得, 日本認知科学会第13回大会論文集, pp.52-53,(1996)