

解説



自然言語処理技術の最近の動向

自然言語生成への招待†

徳永 健 伸竹

1. 自然言語生成とは？

自然言語生成とは何であろうか？ と思ひ手元にあった計算機関係の辞書、ハンドブックのたぐいを調べてみた。調べたのは、A. Barr & E. A. Feigenbaum 編の *The Handbook of Artificial Intelligence*³⁾, S. C. Shapiro 編の *Encyclopedia of Artificial Intelligence*³⁴⁾, 人工知能学会編の人工知能ハンドブック⁴²⁾, 岩波書店の情報科学辞典⁴³⁾の4つである。

The Handbook of Artificial Intelligence (1981) には、“Understanding Natural Language” というタイトルの第4章の第5節に“Text Generation” というタイトルの1節がある。この節は次の1文で始まる：

“Computer text generation is the process of constructing text (phrases, sentences, paragraphs) in a natural language—in a sense, it is the opposite of natural language understanding by machine.”

そして、text generation の研究目標を、(1)文法や文法理論をテストするためにランダムに文を生成すること、(2)計算機の内部表現で表現された情報を自然言語に変換すること、の二つであるとしている。節のタイトルこそ“Text Generation”であるが、内容は1文の生成に関するものに限られている。また、ここであげている第1の研究目標は、今日では自然言語生成の研究とはみなされていない点も注意しておこう。

Encyclopedia of Artificial Intelligence (1987) には、“Natural Language Generation” という項目があり、13ページにわたる解説がある。その書

き出しは次の1文で始まる：

“Natural language generation is the process of deliberately constructing a natural-language text in order to meet specified communicative goals.”

この解説は自然言語生成の草分け的な研究をおこなった D. McDonald の手によるもので、1980年代におこなわれた研究も網羅している。

人工知能ハンドブックには「文生成」という項目がある。出だしは以下のとおり：

「自然言語処理の研究は、文の表層構造を解析してその深層構造を抽出する文理解と、文の『意味』を表わすある種の内部表現（深層構造）から表層構造を生成する文生成の2種類に大別できる。」

解説は13ページにわたり、1970年代の研究から最近の研究まで、主なものを解説している。項目名は「文生成」であるが、内容的には1文にとどまらずテキスト*の生成に関する話題も網羅している。

岩波書店の情報科学辞典には、「自然言語理解」という項目はあるが、「自然言語生成」あるいは「文(章)生成」という項目は収録されていない。この辞書には項目の KWIC が付いているので「*生成*」**で項目を探すと「形態素生成」、「統語生成」がそれぞれ「形態素処理」、「統語処理」という項目で説明されていることが分かる。しかし、生成に関する記述はいずれもごくわずかしかない。

さて、書評を書くのが本稿の目的ではないので、このへんでやめておこう。これらの辞書、ハ

† An Introduction to Natural Language Generation by TOKUNAGA Takenobu (Department of Computer Science, Faculty of Engineering, Tokyo Institute of Technology).

†† 東京工業大学工学部情報工学科

*「テキスト」は“text”の訳語である。“text”は「文章」と訳すことが多いが、「文章」は「書かれたもの」という印象を与えるので、本稿では、発話、文章を含めて「テキスト」という用語で表現する。

** *は任意の文字列を表わす。

ンドブックを調べると、自然言語生成の研究が自然言語理解の研究に比べて遅れていることが分かるだろう。その理由として、(1)自然言語の生成は内部表現を自然言語に変換するだけで、理解に比べれば容易であるという意識があったこと、(2)柔軟な自然言語の生成を必要とする現実的な応用分野がなかったこと⁴¹⁾、(3)研究者の絶対数が少なかったこと、(4)自然言語の生成は理解よりも困難であること²⁵⁾などが指摘されている。ここで(1)と(4)は矛盾するが、これには若干の説明が必要だろう。(2)とも関係するが、これまでの自然言語生成の応用分野では、あらかじめ用意しておいた文を出力したり、一部を空欄にしたテキスト*を用意し、空欄に適切な語を適宜埋めこむことによってテキストを生成する手法で十分であった。つまり、「自然言語生成は容易だった」のである。しかし、われわれが(筆者自身も今こうして)文章を書くときに大変な苦勞をしていることを考えれば、「よい」文章を書くことは決して容易ではない。もちろん、自然言語理解にも未解決な問題が山積しており、理解より生成が困難であるとはいえないが、質の高いテキストを生成しようとするれば、「自然言語生成は十分困難な問題である。」

自然言語生成が理解と比べて特徴的なのは、システムへの入力が見えにくいという点である。これが自然言語生成の研究を難しくしているひとつの原因である。つまり、研究者が共通に生成の出発点として合意できる対象がないのである。これは自然言語理解の研究者が、文字列を出発点として研究できるのとは大きな違いである。自然言語生成の研究者なら、「で、このシステムの入力は何かですか？」という質問を受けたことが1度ならずあるだろう。McDonaldの定義にもあるように、最近では、生成の出発点は話し手の意図あるいは伝達ゴールであるべきだという点で合意が得られつつあるが⁹⁾、話し手の意図とは何か？それはどのように記述されるのか？などに関しては現在もさかんに研究されている段階である⁵⁾。このように、出発点が見えにくい、その記述形式もはっきりしないために、研究成果を横並びで比較することが困難となっている²⁵⁾。自然言語生成の研究におけるこのような難しさは、自然言語理解

の研究の目標が漠然としていることの裏返しの結果でもある。すなわち、「何をもって計算機で自然言語を理解できたというのだろうか？」という古くて新しい疑問が生成研究の難しさの根源であるといえよう。

生成の研究を困難にしているもうひとつの原因として、出力の評価が難しいという問題がある。自然言語理解における重要な課題のひとつであるあいまい性の解消は、多くの候補からなんらかのヒューリスティクスによって、より正しい解を選択する過程である。本特集の長尾・丸山氏の解説にもあるように、あいまい性を解消するためのヒューリスティクスがこれまで数多く提案され、その有効性も確かめられている。このようなヒューリスティクスをうまく発見できるのは、正しい解が何か分かるからである。自然言語生成も種々の項目についての選択過程と考えることができるが、正しい解が何であるのか、すなわち、「よい」テキストとは何なのか明確でないために、有効なヒューリスティクスを発見することが困難となっている。現状では生成システムの出力の評価基準についての合意は得られていない。この難しさは機械翻訳の出力を評価することを考えれば容易に想像できるだろう。この点については5.でもう一度ふれる。

自然言語理解の研究が進むにつれ、より深い自然言語の解析が可能になってきたために、さまざまな応用分野でより柔軟な自然言語の生成が必要となってきた。そして、自然言語を柔軟に生成するためには、多くの興味深い問題があることが次第に明らかになってきた。本稿では、これまで読者にあまりなじみのなかった自然言語生成に関する最近の研究動向をできるだけ網羅的に紹介し、今後の研究課題について述べる。

2. 生成のモデル

自然言語理解では、(1)形態素解析、(2)統語解析、(3)意味解析、(4)談話(文脈)解析の4つの処理を想定するのが一般的である(図-1(a)参照)。これに対して、自然言語生成では、図-1(b)のように(1)what-to-say*の決定と(2)how-to-say**の決定の二つに処理を分けることが多い。

*何をいうか
**どういうか

*これをテンプレートと呼ぶ。

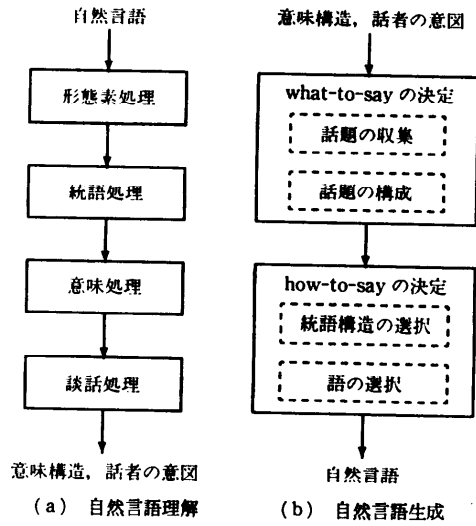


図-1 処理のモデル

what-to-say の決定では、発話時の状況、聞き手のモデルなどを参考に、システムがもつ情報の中から出力すべき情報内容を選択し、その構成に関する種々の決定をおこなう。一方、how-to-say の決定では、主に言語的知識を参照しながら統語構造、語の選択、省略、照応などに関する決定をおこない、出力すべき情報内容を1次元の文字列として生成する。

1980年代に提案された代表的な生成システムの多くがこの what-to-say/how-to-say モデルを踏襲し、決定項目をあらかじめ決められた順序で決定する戦略をとっている^{19), 27)}。しかし、個々の決定項目は互いに依存しており、最適な決定の順序をあらかじめ決めることは困難であるという指摘もある^{2), 15)}。たとえば、一般に語の選択は how-to-say の決定に属するが、語の選択に依存して、表現できる情報内容や情報の順序、すなわち、what-to-say が制約を受けることも十分考えられる。また、what-to-say あるいは how-to-say のみについて考えても、決定項目は互いに依存している。この問題を解決するために Appelt²⁾や Hovy¹⁵⁾ は決定項目の間の依存性を扱うためのモデルを提案している。

3. 文の集合がテキストではない

1980年前半までの生成に関する研究は1文を生成することが主な目的であった。複数の文からなるテキストの生成は1980年代における生成研

究の中心的なテーマであった。特に、文の集合がテキストとなるための要件として多くの研究者が注目したのが結束性 (coherence) と結束構造 (cohesion) という概念である*。結束性と結束構造の定義は研究者によって微妙に異なるが、本稿では、結束性を意味的・概念的なつながりを指すもの、結束構造を表層的なつながりを指すものとする。つまり、結束性は what-to-say に、結束構造は how-to-say に関係する。

テキストを生成するためには、2.で述べたように、まず、適切な内容を選択する必要がある。この選択には、たとえば、以下のような要素が関わってくる⁴⁾：

- 話し手の意図を達成するのに適切か
- 聞き手にとって了解可能で、有益な情報を含むか

• テキストが伝達される場面に適切か
さらに、選択した命題内容からひとつのテキストを構成する場合、情報を時間的順序で提示したり、原因を提示した後で結果を述べるなど、聞き手が意味的なつながりを容易に連想できるような順序にテキストを構成する必要がある。たとえば、次の例文(a)では三つの文の間に意味的な関係が読みとれるが、例文(b)では文と文の間に意味的な関係が読みとれず、全体としてひとつのテキストを構成しているとは言いがたい。

(a) 雨がやんだ。風がやんだ。日がさしてきた。

(b) 一生懸命勉強した。風がやんだ。ポチが死んだ。

つまり、例文(a)のほうが結束性が高いといえる。

結束性が高くても、話題間の意味的なつながりを正確に相手に伝えるためには、さまざまな言語表現による努力が必要である。このためのひとつの手段として結束構造を考えることができる。結束構造を考える上では、論旨の展開や同一対象の指示などを聞き手に分かりやすく示すためにどのような言語表現を用いるかが問題となる。たとえば、論旨展開を明確にするために「したがって」や「ところが」などの接続詞を用いたり、既出の

* "cohesion", "coherence" の訳語はゆれがあるが、本稿では文献4)にしたがい、"coherence" を「結束性」、"cohesion" を「結束構造」とする。なお、文献13)では、"coherence" を「結束性」、"cohesion" を「結束作用」と訳している。

登場人物を人称代名詞などで指示したりするのがこれに当たる。Halliday と Hassan は、英語において結束性を実現するために用いられる結束構造として、参照 (reference)、代用 (substitution)、省略 (ellipsis)、語彙結束 (lexical cohesion)、論理的接続 (logical connectives) の 5 種類をあげている¹¹⁾。

3.1 結束性

what-to-say を決める過程は、話し手の意図からテキストの談話構造 (discourse structure) を構成する過程としてとらえることができる。ここでは、談話構造を「各話題が他の話題とどのような意味的關係にあり、どのような順序で話題が並んでいるかを表わす構造」と定義しておく。

談話構造の構成には、どの話題を聞き手に提示するかを決める「話題の収集」とどのような順序で話題を提示するかを決める「話題の構成」の二つの処理が必要である。以下では、談話構造を構成するためのこれら二つの処理を指してプランニングと呼ぶ。

結束性をテーマとする研究には、談話構造のモデルを用意し、そのモデルにもとづいてプランニングをおこなうものが多い。自然言語生成の分野でこれまで用いられてきた談話構造のモデルは、記述の方法という観点からスクリプト的談話構造モデルと抽象的談話構造モデルとに大きく分けることができる¹⁵⁾。以下、それぞれの談話構造モデルについて簡単に説明する。

3.1.1 スクリプト的談話構造モデル

スクリプト的談話構造モデルは、テキストの談話構造をある程度具体的なシナリオとしてあらかじめ定義したものである*。われわれが実際に書いたテキストを分析してみると、テキストの目的や分野に応じてある程度固定化した談話構造を用いていることが分かる。このような分析にもとづいて、実際のテキストの談話構造から一般化できる部分を抽出したものがスクリプト的談話構造モデルである。たとえば、岩波の情報科学辞典¹⁹⁾では各項目の説明を、(1)同義語、反義語、(2)内包的定義、(3)外延的定義、(4)構成的定義、(5)性質・属性・機能・目的的定義、(6)発生的定義、(7)文脈的定義の順で記述することを原則

とし、(4)以降の記述については省略も可能であるとしている⁴⁴⁾。これは、項目の定義を分かりやすく提示するための一種のスキーマであると考えることができよう。スクリプト的談話構造モデルは、エキスパートシステムの説明文生成に使われたテンプレートを抽象化したものだと考えることができる。

スクリプト的談話構造モデルを用いてテキストの生成をおこなうには、まず、目的に合ったスキーマを選ぶことから始める。各スキーマには、そのスキーマの目的あるいはそのスキーマを使用したときの効果が記述されており、これを手がかりとしてスキーマを選択する。また、システムがテキストとして表現したい内容をどのくらいそのスキーマによって表現できるかもスキーマ選択の手がかりとなる。スキーマを選択したら、そのスキーマに沿って必要な情報を収集し、スキーマにあてはめる。一般にスキーマは談話構造を決定的に定義するのではなく、スキーマをたどる途中でいくつかの枝分れがあるのが普通である。実際にどのパスに沿って談話構造を作るかは、焦点移動に関する制約*などを使って決定することが多い。

スクリプト的談話構造モデルを用いた代表的なシステムとして、McKeown の TEXT²⁷⁾、Kukich の Ana²²⁾などがある。

スクリプト的談話構造モデルを用いると、われわれが日常用いるテキストの構成を容易にシステムに与えることができる。しかし、システムの適用分野やテキストの目的・話題に応じてそれぞれ別々のスクリプトを用意する必要があるため、談話構造モデルの記述の負担が大きい、システムが適用分野に依存しやすいなどの欠点もある。

3.1.2 抽象的談話構造モデル

テキストを構成する節・文・段落などの各ブロックは「理由と結論」、「説明と具体例」などの意味的な関係でつながっている。これらの意味的な関係の種類は無数にあるわけではない。抽象的談話構造モデルでは、ブロック間の意味的な関係を表現するプリミティブの集合を用意し、各ブロックの間をこれらのプリミティブで関係付けることによって談話構造を記述する。抽象的談話構造モデルを用いたプランニングでは、このプリミティブ

*このように個々のシナリオを記述したものをスキーマあるいはスクリプトと呼ぶ。

*談話の各時点で注目されている対象を焦点というが、この焦点が自然に移動するための制約のこと⁴⁴⁾。

を用いてテキストの談話構造を再帰的に構成する。どのような関係をブロック間の関係のプリミティブとして認めるかは、研究者によって意見が分かれるところだが、この点については Hovy が詳しく調査している¹⁷⁾。

代表的な抽象的談話構造モデルに Mann らが提案した RST (Rhetorical Structure Theory)²⁴⁾がある。RST は、人間が書いた文章を分析することによって、隣接するブロック間の意味的關係を〈evidence〉、〈sequence〉、〈elaboration〉など約 20 種類に分類し、形式化したものである。

意味的關係で結ばれた二つのブロックはそれぞれ nucleus, satellite と呼ばれ、たとえば、〈evidence〉スキーマであれば、satellite ブロックが根拠となって nucleus ブロックを支持するという関係を表わす。この 2 項関係を RST スキーマという形式で表現する。nucleus と satellite はさらにそれぞれ別の RST スキーマで表現することができ、テキストの談話構造は木構造として表現される。たとえば、次の段落の談話構造を分析すると図-2 のような木になる*。

[1RST スキーマはテキストの目的や話題に依存しない基本的な意味的關係に対応しているため、]
[2RST スキーマを構造的に組み合わせることによって]
[3多様な談話構造が表わせる]。したがって、
[4プランニングに RST を用いると、]
[5適用分野に対して比較的柔軟なシステムを作ることができる。]
[6RST を用いたプランニングに関しては Hovy¹⁶⁾や Moore³⁰⁾の研究などがある。]

抽象的談話構造モデルを用いてテキストの生成をおこなうには、RST スキーマにいくつかの情報を付加する必要がある。たとえば、Hovy は RST スキーマに「nucleus に関する制約」、「satellite に

関する制約」、「意図する効果」などの情報を付け加え、これらを手がかりに与えられたゴールからトップダウンに談話構造を展開する手法を提案している¹⁶⁾。

3.2 結束構造

結束構造には大きく分けて参照表現と接続表現がある¹¹⁾。参照表現は代名詞や指示詞など特定の対象を指示するための言語表現であり、接続表現は接続詞などブロックとブロックとの意味的關係を示す言語表現である。これらの表現は言語に大きく依存するため、特定の言語に限定した研究が多い。たとえば、英語では既出の対象に対して代名詞化が頻繁に起こるが、日本語では代名詞化の頻度は比較的低い。これまでに開発された自然言語生成システムのほとんどがこれらの表現手法の一部を備えているが、それを中心に扱った研究は少ない。また、言語学の研究成果が生成システムに直接適用できるほど十分に形式化されていないために、これらの表現を生成するための規則にはアドホックなものが多い。

3.2.1 参照表現

参照 (reference) には大きく分けて談話内照応 (endophora) と談話外照応 (exophora) がある¹¹⁾。談話内照応はテキスト中にすでに登場した対象を指し、談話外照応は、談話がおこなわれている場面や談話の参加者があらかじめもっている知識など談話の環境にある対象を指す表現である。談話内照応には、すでに出てきた対象を指す前方照応 (anaphora) と、今後登場する予定の対象を指す後方照応 (cataphora) があるが、生成の分野では後方照応はほとんど扱われていない。前方照応を実現するには代名詞、指示代名詞、語の繰り返し、上位語、省略などの表現を使うことができる。

Grosz らは、談話構造が参照表現に影響することを指摘し、その関係の形式化を試みている¹⁰⁾。Dale は、Grosz らの理論を拡張し、談話構造を用いた参照表現の生成を実現している⁶⁾。

談話外照応を扱ったものには、スクリプトの談話構造モデルと聞き手のモデルを用いることによって物理的対象を適切に指示する名詞句の生成をおこなう Paris らの研究³¹⁾がある。Appelt の KAMP も談話参加者の個別知識、共有知識を用いて適切な談話外照応の生成を実現している²⁾。

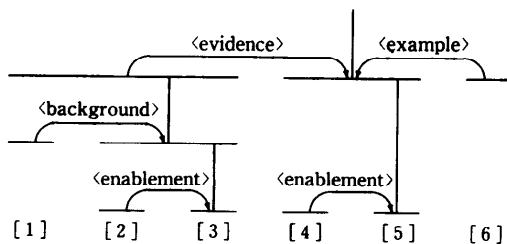


図-2 談話構造の例

* 各節の先頭に付加した番号はブロック番号を表わす。この段落も本文の一部であることに注意。

3.2.2 接続表現

テキストの中では、節と節、文と文、あるいは段落と段落など、ブロック間の意味的関係を聞き手に明確に提示するためにさまざまな接続表現が用いられる。接続詞などを用いてブロックとブロックの意味的関係を明示することは聞き手の理解の助けになるが、逆に接続表現を多用するとテキストが冗長になるおそれがある³⁹⁾。たとえば、次の例では、(b)に比べ(a)は冗長な感じを受ける。また、(c)は接続関係を明示的に表現していないにもかかわらず、容易に接続関係を推測することができる。

(a) 空が暗くなった。そして、風が吹いてきた。そして、雨が降ってきた。

(b) 空が暗くなって、風が吹いてきた。そして、雨が降ってきた。

(c) 雨がやんだ。日がさしてきた。

この例のように、接続詞などを用いて意味のつながりを表層で明示することと、冗長性を減らすこととは、トレードオフの関係になる場合がある。また、接続表現の選択は、複数の話題をいくつの文に分けて述べるかというような問題とも直接的な関係にあり、文章のスタイルなども考慮しなければならない。

ブロック間の意味的関係には接続表現を用いて明示する必要のないものもある。Scottらは、RSTで表現された談話構造からどの話題を取り出してひとつの文を構成するかについて議論し、埋め込み文や等位接続などを生成するヒューリスティックを提案している³⁹⁾。

日本語における接続表現の生成には高橋らの研究³⁹⁾がある。高橋らは、小学校低学年の国語の教科書の分析から得られた接続表現生成のアルゴリズムを汎用文章生成システム上に実現している。

伊藤らは、接続詞の可換性にもとづいて接続詞を分類し、読み手が文間の接続関係を理解する上で接続詞がどのように役立っているかを実験的に調べている³⁸⁾。

4. 生成のための文法

文法は、自然言語理解の研究では統語処理をおこなう上で重要な役割を担っている。直観的には、自然言語生成においても自然言語理解で使用されている文法を使えばよさそうである。実際、

解析器と生成器で文法を共有しようという研究がおこなわれている。しかし、これらの研究における生成器への入力には論理形式かあるいはそれに近いもので、生成器にはすべての情報が細部まで決定されて渡されるのが普通である。いいかえれば what-to-say は完全に決定されていることが前提となっている。したがって、このような研究の成果を自然言語生成にそのまま利用しようとする、必然的に2.で説明したような what-to-say と how-to-say の決定を切り分けたアーキテクチャを採用することになる。現在の生成研究の関心が what-to-say の決定と how-to-say の決定の間でどのようにして緊密に情報を交換するかという点であることを考えると、これらの研究の成果を採り入れるためにはなんらかの工夫が必要である。この点に関しては本特集の橋田氏の解説が参考になるだろう。

初期の生成システムでは文法をプログラム中に直接埋め込むものが多かったが、最近では、言語理論の成果を採り入れ、文法を宣言的に記述するものが多くなってきた。自然言語理解では、文法の枠組として句構造文法の枠組を使用することが多いが、自然言語生成では、FUG²⁰⁾やシステミック文法¹²⁾が多く使われている。その主な理由として、次の2点が考えられる。

- 句構造文法の対象が基本的に単文に限られているのに対し、3.で述べたように自然言語生成の主な興味は1文を越えたテキストにあること

- FUG やシステミック文法の枠組では生成で重要となる選択操作が扱いやすいこと

FUG は、統語的な情報とともに、話題や焦点などの文を越えた情報を属性として容易に文法中に取り込める特徴をもち、属性の単一化によって適切な表現を選択することができる。ただし、FUG の単一化は一般に計算コストが高いので実現には注意しなければならない。McKeown の TEXT²⁷⁾、Appelt の TELEGRAM¹³⁾などが FUG を使っている。

システミック文法は文法的素性* を枝にもつ一種の弁別ネットワークの集合によって文法を表現する。おのおのの節点をシステムという。生成は、

*たとえば、能動態か受動態か、平叙文か疑問文か命令文かなどを表わす素性。

● 選択器という特別な手続きが環境と対話しながらネットワークをたどり、適切な文法的素性を選択する

● 選択した素性の集合から文を生成するという二つの段階を経ておこなう。システミック文法に関しては文献 36) の第 6 章に詳しい。

システミック文法を使ったシステムの例としては、Mann の Penman²³⁾、Patten の SLANG³²⁾ などがある。

そのほかに、生成システムで用いられる文法の形式として phrasal lexicon という考え方がある*。Phrasal lexicon は、文法規則を語彙の中へ埋め込んだ形で記述したもので、初期の生成システムで使われていたテンプレートをより抽象化したものと考えることができる。Phrasal lexicon は、人間も定型句を組み合わせて言語を使っている側面があるという観察にもとづいている。

Kukich は、人間がよく使う言い回しを phrasal lexicon としてシステムに与えることと流暢な自然言語の生成との関連性を指摘し、phrasal lexicon の有用性を主張している²²⁾。また、最近では、Fillmore が慣用句などを扱うために“Construction Grammar”という文法理論を提案している⁹⁾。Fillmore も認めているように、Construction Grammar と phrasal lexicon の間には共通する考えかたが多い**。Phrasal lexicon には体系的な理論がなく、これまで、研究者ごとにアドホックな扱い方をしていたが、筆者は Construction Grammar が phrasal lexicon の理論的な支えになるのではないかと期待している。

Kukich の Ana²²⁾、Jacobs の PHRED¹⁹⁾、Hovy の PAULINE¹⁵⁾などが phrasal lexicon を用いている。

紙面の都合で、文法の扱いに関して具体例をあげて説明できないが、興味がある読者は文献 45) を参照して欲しい。

5. 自然言語生成はどこへ向かうか？

1980年代は自然言語生成にとって実り多い10年であった。研究の関心が、1文の生成から複数の文の生成に移ってきた点、生成の出発点も出力すべきテキストの意味構造から遡って話し手の意図へと移ってきた点が大きな進歩だろう。この理由

として、自然言語理解の研究の関心が形態素や統語といった浅い解析から意味や談話といったより深い解析に移ってきたことが考えられる。自然言語生成は深い理解の結果を出発点として処理を始めなければならないからである。また、これと関係して文章の要約システムや対話システムなど、自然言語生成の応用範囲が広がってきたことも理由のひとつだろう。最後に、残された研究課題と最近の傾向について述べて本稿の結びとする。

5.1 マルチメディア化

最近の研究の傾向として、音声や映像などのようにテキスト以外の媒体とテキストをどのように組み合わせ人間に提示するかという研究が増えてきたようである。隔年で開かれている自然言語生成に関する国際ワークショップの今年のテーマは“Multimodal generation”である⁷⁾。しかし、筆者の個人的な考えでは、他の媒体との融合を考える以前に、テキストだけをとりあげても生成の研究にはまだ解決すべき問題が数多く残されていると思う。

5.2 心理学モデル

自然言語生成の研究も幅が広がり、より人間に近い生成のモデルを目指す認知科学的なアプローチをとる研究も生まれてきた。Kempen は、人間は発話をおこなう際に、その発話の内容をすべて決定してからしゃべりはじめるのではないという心理言語学の知見に注目し、人間の発話過程を計算機上に実現するために漸次的生成モデルを提案している²¹⁾。このモデルでは、次々と想起される概念のストリームから漸次的に統語構造を構成し、表層化できるままとまりができ次第発話がおこなわれる。

確かに発話に関しては、人間は「思いつくまま次々にしゃべっている」ようである。しかし、文章を書く場合には若干事情が異なる。もちろん、最初の草稿は「思いつくままに書きなぐる」ということはあっても、それを最終稿とする人はほとんどいないだろう。よい文章を書くためには多くの複雑な要求や制約を考慮する必要があるが、すべての要求や制約を一度に考えるのは人間にとって困難である。これはおそらく情報処理装置としての人間の作業記憶の制限によるものだろう*。

*たとえば文献 19) を参照。
**私信による。

*この解説を書くのに何度書き直しをしたことか。もし、これに異論ある人がいれば筆者に連絡されたい:-)

そこで普通は草稿を何度か書き直したり部分的な修正をする。つまり、推敲をする。推敲によって、複雑な要求や制約を少しずつ考慮することができるからである。実際、人間が文章を書く作業の大部分は推敲に当てられているという心理学的観察³⁷⁾も報告されている。このような知見にもとづき、生成過程に推敲機能を導入するモデルがいくつか提案されている^{18), 23), 29)}。もし、推敲が人間の作業記憶の制限だけのために必要なのなら、計算機で文章を生成する場合には、作業記憶を十分大きくとれば推敲は不要かもしれない。しかし、現在の計算機上に人間の作業記憶より大きい記憶領域を確保できるのかどうかは明らかではない。この点では、生成システムに推敲過程を導入することは、心理学的見地からだけでなく工学的にも意味があるだろう。

5.3 よいテキストとは何か？

生成されたテキストをどのように評価するかについてはこれまであまり論じられてこなかった。計算機を使おうが、人間が書こうが、テキストを評価するのは一般に困難である*。しかしながら、生成システムを横並びで比較するためには、テキストの評価基準が必要である。

前述した推敲機能を生成システムに導入する際にもテキストの評価は重要となる。これまで推敲機能をもつシステムの実現が困難であった主な原因は、生成したテキストの評価が困難であることに起因している。

テキストの評価に関しては、機械翻訳システムの訳文の評価に関する研究、あるいは実験心理学の手法が参考になると思われる。また、評価自体をテーマとした研究もいくつか報告されている⁴⁰⁾。このほかに、テキストを解析し推敲に役立つ情報を書き手に提供することを目的とした文章推敲支援システムの研究において、文章の評価法が模索されている⁴⁶⁾。

5.4 「入力は何ですか？」

1.で述べたように、最近では話し手の意図が自然言語生成の出発点であるべきだという大方の合意はあるが、具体的に意図がどのようなものであるか、またそれをどのように表現したらよいのかについては明らかでない。

一方、生成システムへの入力を仮定することすら疑問視する考え方もある。McNeill は、心理言語学の立場から発話の自発的な生成能力を重視し、思考と言語(記号)はそもそも不可分であるとしている。発話の入力として意図を生成システムに与えるアプローチ、すなわち思考過程と切り離れた形で発話の入力を仮定するのでは真の発話とはいえないというのが McNeill の主張である²⁸⁾。McNeill は、より人間に近い発話のモデルを作るためにはコネクショニストアプローチのほうがまだ見込みがあるとしている。

この問題は、工学の立場から入力の解釈あるいは概念化の問題として取り上げられている。これまでの生成システムでは、生成すべき内容をなんらかの意味表現形式で与えることが多い。この場合、情報内容は入力の時点ですでに特定の概念を用いて表現されており、その概念に対応する語を用いて言語化していることになる²⁶⁾。これに対し、Hovy は伝達すべき情報をどのような概念でとらえるかがまず重要であると指摘している⁴⁴⁾。たとえば、野球の試合の流れを記述するのに、どちらのチームのファンかによって記述内容、表現は当然変化するだろう*。Hovy の PAULINE は談話状況の語用論的特徴**に依存して同一の事象を異なる概念化を経由して言語化することができる。Hovy はこれを話題の解釈と呼んでいる。

しかしながら、入力が概念の組み合わせで表現されている以上、すでになんらかの解釈を含んでいることになる。概念で表現されていない入力を扱った例は少ない。たとえば、Kukich の Ana²²⁾ は、株式の数値データを入力として株価の変動状況を文章化する。たとえば、ある数値の変動を上昇とみるか、無変化とみるかは解釈に依存する。

さて、このような考え方を進めるとどうなるだろう。「入力は何ですか？」という問いに対する究極の答は「文字列です」ということになるのではないだろうか。自然言語を生成するためには必ずなんらかの解釈が必要である。では、何を解釈するかといえばやはり環境からの刺激であろう。もちろん、人間の発話は必ずしも環境からの刺激に対する反応ではなく、自由意思にもとづいた自発的な発話の場合もあるだろう。しかし、自由意思

* 筆者も今書いている文章が「よい」ものかどうかは自信がない。もししたらもっとよい文章が書けるかもしれない。何が最終稿になるかは時間の制約で決まるものである。

*たとえば、報知新聞の巨人の激いのみよ。

** 談話の雰囲気、話し手、聞き手の特徴、話し手と聞き手の関係など。

とは何か? という問いにまともに答えるのは困難であるし、自分の言いたいことをただしゃべるだけのシステムがあまり役に立つとは思えないのでここでは考えないことにしよう。さらに、環境からの刺激にもいろいろあるが、なぜ文字列なのか? ひとつには計算機に入力しやすいという利点がある。しかし、何よりも自然言語理解の研究成果が利用できるという利点大きい。そう、自然言語理解の研究の目的はまさに与えられた言語表現(文字列)の解釈をすることだったではないか!

すなわち、自然言語生成の入力は何かという問いをまじめに考えると自然言語理解に辿り着くわけである。かくして、自然言語理解と生成の研究がめぐりあうことになる。1.で述べたように、生成の研究が理解の研究に比べて遅れていたこともあって、理解の研究成果を生成の研究が利用することはあってもその逆は少なかった。「こういう応答を生成するにはどのような解析をおこなえばよいか」という観点に立った自然言語理解の研究がもっとおこなわれてもよいのではないだろうか。両者は車の両輪であり、互いの研究成果をフィードバックしながら研究を進めてゆくべきだろう。このことと理解と生成の両方の要素を含んでいる対話の研究が最近さかんになっていることは無関係だろうか?*

参 考 文 献

- 1) Appelt, D. E.: TELEGRAM: A Grammar Formalism for Language Planning, In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 595-599 (1983).
- 2) Appelt, D. E.: *Planning English Sentences*, Cambridge University Press (1985).
- 3) Barr, A. and Feigenbaum, E. A., editors: *The Handbook of Artificial Intelligence*, William Kaufmann, Inc. (1981).
- 4) Beaugrande, R. and Dressler, W. U.: *Introduction to Text Linguistics*, Longman Group Limited (1981). 邦訳: 池上嘉彦, 三宮郁子, 川村喜久男, 伊藤たかね: テクスト言語学入門, 紀伊国屋書店 (1984).
- 5) Cohen, P. R., Morgan, J. and Pollack, M. E., editors: *Intentions in Communication*, The MIT Press (1990).
- 6) Dale, R.: Generating Recipes: An Overview of Epicure, In Dale, R., Mellish, C. and Zock, M.,

editors: *Current Research in Natural Language Generation*, chapter 9, pp. 229-256, Academic Press (1990).

- 7) Dale, R., Hovy, E., Rösner, D. and Stock, O., editors: *Aspects of Automated Natural Language Generation*, Vol. 587 of *Lecture Notes in Artificial Intelligence*, Springer-Verlag (1992).
- 8) Dale, R., Mellish, C. and Zock, M.: Introduction, In Dale, R., Mellish, C. and Zock, M., editors, *Current Research in Natural Language Generation*, chapter 1, pp. 1-15, Academic Press (1990).
- 9) Fillmore, C. J.: The Mechanisms of "Construction Grammar", In *the Proceedings of the 14th Berkeley Linguistic Society Conference* (1988).
- 10) Grosz, B. J. and Sidner, C. L.: Attention, Intention, and the Structure of Discourse, *Computational Linguistics*, Vol. 12, No. 3, pp. 175-204 (1986).
- 11) Halliday, M. A. K. and Hassan, R.: *Cohesion in English*, Longman (1976).
- 12) Halliday, M. A. K.: *An Introduction to Functional Grammar*, Edward Arnold (1985).
- 13) Halliday, M. A. K. and Hassan, R.: *Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective*, Deakin University Press (1985). 邦訳: 笥 壽雄: 機能文法のすすめ, 大修館書店 (1991).
- 14) Hovy, E. H.: Interpretation in Generation, In *Proceedings of the National Conference on Artificial Intelligence*, pp. 545-549 (1987).
- 15) Hovy, E. H.: *Generating Natural Language under Pragmatic Constraints*, Lawrence Erlbaum Associates (1988).
- 16) Hovy, E. H.: Planning Coherent Multisentential Text, In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 163-169 (1988).
- 17) Hovy, E. H.: Persimonious and Profligate Approaches to the Question of Discourse Structure Relations, In *Proceedings of the Fifth International Workshop on Natural Language Generation*, pp. 128-136 (1990).
- 18) Inui, K., Tokunaga, T. and Tanaka, H.: Text Revision: A Model and Its Implementation, In Dale, R., Hovy, E., Rösner, D. and Stock, O., editors, *Aspects of Automated Natural Language Generation*, pp. 215-230, Springer-Verlag (1992). *Lecture Notes in Artificial Intelligence* Vol. 587.
- 19) Jacobs, P. S.: PHRED: A Generator for Natural Language Interfaces, *Computational Linguistics*, Vol. 11, No. 4, pp. 219-242 (1985).
- 20) Kay, M.: Functional Unification Grammar: A Formalism for Machine Translation, In *Proceedings of the International Conference on Computational Linguistics*, pp. 75-78 (1984).
- 21) Kempen, G. and Hoenkamp, E.: An Incremental

*たとえば、人工知能学会には最近「言語・音声理解と対話処理研究会」が新設された。

- tal Procedural Grammar for Sentence Generation, *Cognitive Science*, Vol. 11, No. 2, pp. 201-258 (1987).
- 22) Kukich, K.: Fluency in Natural Language Reports, In McDonald, D. D. and Bolc, Leonard, editors, *Natural Language Generation Systems*, chapter 8, pp. 280-311, Springer-Verlag (1988).
- 23) Mann, W. C.: An Overview of the Penman Text Generation System, In *Proceedings of the National Conference on Artificial Intelligence*, pp. 261-265 (1983).
- 24) Mann, W. C. and Thompson, S. A.: Rhetorical Structure Theory: Description and Construction of Text Structures, In Kempen, G., editor, *Natural Language Generation*, chapter 7, pp. 85-96, Martinus Nijhoff (1987).
- 25) McDonald, D. D.: Natural Language Generation, In Shapiro, S., editor, *Encyclopedia of Artificial Intelligence*, pp. 642-654, Wiley-Interscience (1987).
- 26) McDonald, D. D.: On the Place of Words in the Generation Process, In Paris, C. L., Swartout, W. R. and Mann, W. C., editors, *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, chapter 9, pp. 229-247, Kluwer Academic Publishers (1991).
- 27) McKeown, K. R.: *Text Generation*, Cambridge University Press (1985).
- 28) McNeill, D.: *Psycholinguistics: A New Approach*, Harper & Row (1987). 邦訳: 鹿取廣人他訳, 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).
- 29) Meteer, M. M. (Vaughan) and McDonald, D. D.: A Model of Revision in Natural Language Generation, In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 90-96 (1986).
- 30) Moore, J. D. and Paris, C. L.: Planning Text for Advisory Dialogues, In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pp. 203-211 (1989).
- 31) Paris, C. L.: Tailoring Object Descriptions to a User's Level of Expertise, *Computational Linguistics*, Vol. 14, No. 3, pp. 64-78 (1988).
- 32) Patten, T.: *Systemic Text Generation as Problem Solving*, Cambridge University Press (1988).
- 33) Scott, D. R. and Souza, C. S.: Getting the Message Across in RST-Based Text Generation, In Dale, R., Mellish, C. and Zock, M., editors, *Current Research in Natural Language Generation*, chapter 3, pp. 47-74, Academic Press (1990).
- 34) Shapiro, S. C. editor: *Encyclopedia of Artificial Intelligence*, John Wiley & Sons, Inc. (1987).
- 35) Sidner, C. L.: Focusing in the Comprehension of Definite Anaphora, In Brady, M. and Berwick, R. C., editors, *Computational Models of Discourse*, pp. 267-330, MIT Press (1983).
- 36) Winograd, T.: *Language as a Cognitive Process*, Vol. 1: Syntax, Addison-Wesley (1983).
- 37) Yazdani, M.: Reviewing as a Component of the Text Generation Process, In Kempen, G., editor, *Natural Language Generation*, chapter 13, pp. 183-190, Martinus Nijhoff (1987).
- 38) 伊藤俊一, 阿部純一: 接続詞の機能と必要性, 心理学研究, Vol. 62, No. 5, pp. 316-323 (1991).
- 39) 高橋 晃, 桃内佳雄, 宮本衛市: 文章生成における接続詞の生成方略について, 情報処理学会自然言語処理研究会, Vol. NL 62-3 (1987).
- 40) 高橋善文, 牛島和夫: 計算機マニュアルのわかりやすさの定量的評価方法, 情報処理学会論文誌, Vol. 32, No. 4, pp. 460-469 (1991).
- 41) 上原邦昭: 文生成, 人工知能学会(編), 人工知能ハンドブック, pp. 274-286, オーム社 (1990).
- 42) 人工知能学会(編): 人工知能ハンドブック, オーム社 (1990).
- 43) 長尾 真他(編): 情報科学辞典, 岩波書店 (1991).
- 44) 長尾 真: 辞典形式での専門分野の知識の体系的構成法, 人工知能学会学会誌, Vol. 7, No. 2, pp. 320-328 (1992).
- 45) 徳永健伸, 乾健太郎: 1980年代の自然言語生成一三一, 人工知能学会学会誌, Vol. 6, No. 5, pp. 651-662 (1991).
- 46) 林 良彦, 高木伸一郎: 日本文推敲支援機能の実現方式, 人工知能学会全国大会, pp. 399-402 (1989).

(平成4年3月23日受付)



徳永 健伸 (正会員)

1961年生. 1983年東京工業大学工学部情報工学科卒業. 1985年同大学院理工学研究科修士課程修了. 同年(株)三菱総合研究所入社. 1986年東京工業大学大学院博士課程入学. 現在同大学工学部情報工学科講師, 博士(工学). 自然言語処理, 計算言語学に関する研究に従事. 認知科学会, 人工知能学会, 計量国語学会, Association for Computational Linguistics各会員.