

# 言葉をデザインする

## ーデザインされた産業日本語ー

一般財団法人日本特許情報機構 特許情報研究所顧問 **横井 俊夫**

### PROFILE

1966年に電気試験所（現在：産業技術総合研究所）。1982年より第五世代コンピュータプロジェクトの推進に従事。1987年より電子化辞書プロジェクトの推進、運営に従事。1995年よりフィリピン国にてODAプロジェクトの推進、指導に従事。1997年より東京工科大学、2008年より現職。工学博士、東京工科大学名誉教授。

## 1 言語への新たなアプローチ

言語への新たなアプローチとして、合成的なアプローチ、すなわち、言語を工学的にデザインするというアプローチを提案する。そして、このアプローチに基づいて、構造化日本語、さらには、諸外国語に普遍化した構造化言語という言語系を提案する。構造化言語によって表現（記述）された構造化テキスト（構造化文書・構造化文章）によって、文書処理（作成、検索、翻訳、要約、情報抽出等々）が格段に効率化され高機能化されることを紹介する。

本稿では、日本語と和文テキストを中心に、すなわち、構造化日本語と構造化和文テキストを中心に説明する。構造化日本語は、表現対象となる情報が本来保有する構造により密着できるようにするためにグラフィカルな表現要素が組み込まれた日本語である。この構造化日本語に対し、通常の日本語を線状化日本語と呼ぶ。

Mind Map [1] や Conceptual Mapping 等々、様々なグラフィカルな表現形式が提案され、利用されている。構造化日本語は、これらの既存の図式表現形式と大きく異なる考え方に基いてデザインされた言語である。構造化日本語は、あくまでも、日本語である。素直に、日本語として、読み書くことができる。

構造化和文テキストを読み下すことによって、線状化和文テキストが得られる。構造化和文テキストは、読み方の方略にしたがって、多数の読み下し方が可能である。すなわち、ひとつの構造化和文テキストに多数の、時には、無数の線状化和文テキストが対応付けられる。この

事実が、構造化和文テキストが表現対象の情報の構造に如何に密着した表現形式であるかを示している。

構造化日本語、構造化和文テキストは、論理的な明晰さや厳密さが強く求められる情報の表現に適している。仕様書、契約書、特許文書、レファレンス文書、設計・運用マニュアル等々の産業技術文書、さらには、法令・司法関連文書などへ効果的に適用できる。また、従来、散文的に語るしかできなかったテクニカルライティングが、体系立った工学的技術として議論できるようになる。すなわち、構造化日本語は、デザインされた産業日本語である。

### 1.1 分析的アプローチから合成的アプローチへ

現在、言語に対する主流となるアプローチは、分析的 (analytic) なアプローチである。自然言語という自然現象、あるいは、自然言語によって創り出される自然現象を分析することによって、言語学は、言語共通の仕組みや個別言語の多様性の説明原理を追求する。説明原理は、あくまでも、現象を体系付けるための分析モデルである。そして、言語工学（言語処理）も、自然言語という自然現象を分析することによって、処理モデルや処理アルゴリズムを追求する。目指すところは、あくまでも、現象をあるがままに、出来るだけ幅広く、出来るだけ高速に処理できるようにすることである。

言語工学において、合理主義か経験主義かの論争がある [2]。言語モデルに基づくルールセットとして言語現象を捉えようという合理主義的方法論と大規模データに対する統計処理によって言語現象と捉えようという経験主義的方法論の論争である。この論争は、言語学をも

巻き込む言語研究に広範に係わる論争である。しかしながら、この論争は、分析的アプローチの枠内での論争である。実は、この論争の中に、合成的 (synthetic) アプローチの可能性と必要性を読み取ることができる。

まずは、合成的アプローチの可能性である。合成的アプローチにとって重要なことは、それぞれの分野の言語現象に対する大規模な現象データが集積され、大規模データに対する分析手法が進化したことである。この進化には、合理主義も経験主義も共に貢献したといえるであろう。分野ごとの言語現象を多角的に分析し、分野ごとの言語運用の実態をモデル化することが可能となった。このモデル化が、合成的アプローチにおける言語デザインのための枠組みを与えてくれる。さらに、言語学におけるもうひとつの進展がある。認知言語学等における進展である。認知言語学は、人間が持つ一般的な認知能力の枠組みのもとに言語能力を捉え、言語現象を一般の知能現象として分析する方法論を提供してくれる。この方法論が、言語デザインのためのもうひとつの枠組みを与えてくれる。なお、認知言語学の進展は、合理主義と経験主義の論争にはあまり登場しない。

つぎに、合成的アプローチの必要性である。言語現象は、典型事例の周りに周辺事例が非常に裾野広く分布するという境界を確定し難い現象である。これは、言語表現とそれが表現する内容とを対応付ける仕組がもつ高い柔軟性による。この柔軟性は、人間知能が情報の表現メディアに求める本質的な恣意性に基づく。このため、言語現象へ分析的にアプローチする限りは、周辺事例や例外的事例への際限のない対応に悩まされることになる。この対応作業は、多くの場合、アドホックな賽の河原の石積みのような作業となる。実際の言語運用では、周辺事例をある範囲内に限ったとしても、支障をきたすことはない。言語をデザインすることによって、不必要な周辺事例に対応しなくともよいようにできる。

現在の言語現象には、特筆すべきもうひとつの側面がある。この側面は、分析的アプローチでの論争ではあまり注目されていない。現在の言語現象は、もはや、言語だけの現象、プレーンなテキストの集積という現象ではない。Web ページテキストは、HTML によってテキ

スト構造やテキスト間のリンク情報が注記された構造化されたテキストである。この構造化情報を活用するテキスト処理によって、高度な処理機能が実現されている。

Google の検索エンジンは、全文検索というプレーンテキストに対するテキスト処理だけで成り立っているわけではない。HTML による注記情報を活用することによって、ソフトウェアロボットである「クローラ」は、Web スペースを駆け巡り、検索のための情報を収集し、検索エンジンは、重要度を計算し、検索されたページにランク付けを行う。

高度なテキスト理解やテキスト処理のためには、文脈情報の利用が必要である。これは、言語学においても言語工学においても常識である。しかし、一般的に文脈の議論を始めると、現在の言語学や言語工学では手の付けようの無い話しとなる。ただし、文脈情報を Web ページ間のリンク情報、ユーザーの利用履歴情報であるとすれば、Web スペース上では、膨大な文脈情報を機械的に捕捉することが可能となっている。言語処理は、言語のみの狭いテキスト処理に拘泥するのではなく、より多様な構造化情報を利用する、より多様な構造化情報が得られるようにする、これらへの努力も加味すべきであろう。

本稿は、構造化への努力に向けた提案である。ただし、その努力も、HTML の普及の経緯を十分に踏まえたものでなければならない。構造化言語、構造化テキストも、一般の利用者にとっての敷居が至って低いこと、利便性を直裁的に実感できること、この 2 点が重要である。

## 2 言語をデザインするとは

言語への合成的アプローチとは、設定した目的を効果的に実現するための言語を工学的にデザインし、デザインされた言語によって、高度なテキスト処理機能（作成、検索、翻訳、要約等々）や知識処理等への効果的な連携機能を実現することである。

まず、デザインするとはである。デザインするとは、対象のアフォーダンス (affordance) を高め、人にとっ



での利用価値を高めることである。アフォーダンスとは、米国の生態心理学者の J.J. ギブソンによって提唱された概念である [3]。ギブソンは、(人間を含めた) 動物が (外界) 環境から受け取る「意味」であるとアフォーダンスを定義付けた。動物は、環境とのインタラクションを通じて、環境を理解していく。このインタラクションを成り立たせるための動物への環境の働きかけの関係性がアフォーダンスである。アフォーダンスは、単なる外界からの刺激ではない。

アフォーダンスは、米国の認知科学者である D.A. ノーマンによって、いささか、拡大解釈され、ヒューマンインタフェースのデザイン原理を定義付けるものとして位置付けられた [4]。ノーマンは、人間が知覚するアフォーダンスを高めることをデザイン原理と定義づけた。

つぎに、言語をデザインするとはである。言語をデザインするとは、人間に知覚される言語のアフォーダンスを高め、高度な利用価値をもたらす、効果的で効率的なインタラクションを実現するような言語を設計することである。人間と言語とのインタラクションには、2つの側面がある。ひとつは、言語そのものとのインタラクションであり、もうひとつは、言語を媒介とした表現対象とのインタラクションである。

まず、言語そのものとのインタラクションにおけるアフォーダンスである。ユーザーの母語、ユーザーのその分野における日常的な母語の用法、これらから外れば外れるほど、そのユーザーにとってのアフォーダンスは低下する。したがって、日本人ユーザーにとっては、日本語であること、そして、産業日本語としては、その産業技術分野の日常化された用法に則ること、これらがアフォーダンスを高める。

このアフォーダンスの観点に立てば、エスペラントをはじめとする国際共通語 (Common Language) が何故普及しないのかが明らかとなる。共通化のために、誰にとってもアフォーダンスの低い言語が無理やり設計されるからである。その点、Globish [5] のように、すでに世界で広く使われている言語 (英語) をベースにする国際共通語は、健全である。

次に、表現対象とのインタラクションにおけるア

フォーダンスである。表現対象となる情報を出来るだけ素直になぞることが出来れば、アフォーダンスは高まる。ただし、この「なぞる」というところに課題がある。情報は、言語から独立して存在するのではなく、言語によって表現されて、初めて、情報となる。したがって、アフォーダンスを高めるには、人間による情報の理解過程・伝達過程に沿うことが肝要となる。

このアフォーダンスの観点に立てば、制限英語や制限日本語などの制限言語 (Controlled Language) における「制限する」ということと「デザインする」ということが本質的に異なることが明らかとなる。機械処理の都合のために制限したりすることや簡易化したりすることは、多くの場合、人間のためのアフォーダンスを高めることと真逆の行為であるからである。

### 3 構造化日本語ーデザインされた産業日本語

デザインされた産業日本語として、構造化日本語を提案する。構造化日本語は、あくまでも日本語である。日本語として書き、日本語として読むことが出来る。この、日本語であるということが重要である。

表現対象となる情報は、三次元空間に時間軸を加味した四次元空間を構成すること、物理的な具象物に係わる情報はそうである。そして、抽象物に係わる情報は、具象物へのメタファーを用いて構成される。一方、言語は一次元である。その誕生の過程からも、また、幼児の言語習得の過程からも、言語の主軸となるのは、音声言語である。発声器官や聴覚器官の線状性 (一次元性) が、音声言語に線状構造を課することになる。文字言語も、音声言語に準じて、基本は線状構造となる。

言語による情報の表現や言語表現からの情報の読み取りには、四次元の情報と一次元の言語表現との間のマッピング操作が求められることになる。もちろん、言語には、マッピングのための色々な仕組が、形態的にも構文的にも用意されているが、十分に手当てされているというわけではない。手当ての不十分さや手当ての不適切な使用が、言語表現に曖昧さや非明晰さを生む要因となる。

構造化日本語では、表現対象とのインタラクションにおけるアフォーダンスを高めるために、(自然)日本語の線状性という制約を緩める。そして、日本語としての基本的な仕組みを保持することによって、言語そのもののインタラクションにおけるアフォーダンスを維持し、あるいは、より高める。このようにデザインされた日本語が構造化日本語である。構造化日本語は、情報を論理的に明晰に表現するのに適していることから、デザインされた産業日本語となる。

### 3.1 構造化をデザインする

構造化日本語をデザインするためには、まず、構造化そのものをデザインしなければならない。情報は四次元空間を構成するとした。しかし、構造化が、この四次元に直接対応するわけではない。構造化が対応すべきなのは、四次元情報が言語表現されたときに持つ構造である。情報の構造は、言語表現されることによって定まる。言語によって、情報をどのように理解していくのか、情報をどのように伝えるのか、すなわち、言語による情報の理解過程・伝達過程が創り出す構造がデザインすべき構造化である。

なお、言語の媒介を副次的なものとして情報の構造をどう表現するのかについては、情報デザイン (Information design) や情報可視化 (Information visualization) などの分野で議論されている [6]。「言葉をデザインする」は、「情報をデザインする」に相補的に連携することになる。

そこで、言語による情報の理解過程・伝達過程が創り出す構造化とはである。一例を挙げよう。事柄としての情報を< A 社が B 装置を開発した > であるとしよう。この情報の言語表現として、以下のようなものを対応付けることができる。

- ① 「A 社が B 装置を開発した。」
- ② 「A 社は、B 装置を開発した。」
- ③ 「B 装置は、A 社が開発した。」
- ④ 「A 社が開発したのは、B 装置である。」
- ⑤ 「B 装置を開発したのは、A 社である。」

この5つの文は、事柄としての情報としては、同じものを表現している。事柄としての情報の構造としては、以下の事象構造 [7] がすべての文に対応付けられる。

[事象 時間性 = 過去 :

事象基 → < 開発する >

動作主 → < A 社 >

動作対象 → < B 装置 >

それでは、5つの文の違いは、何を反映しているのだろうか。これらの違いが反映するのは、人が事象構造をどのように理解しようとしているのか、どのように伝えようとしているのか、そこにおける違いである。人間による情報の理解や伝達の過程をまとめて情報の処理過程と呼ぶとすると、反映するのは、この情報の処理過程の違いである。そこで、この情報の処理過程の構造を処理構造と呼び、その構造化に取り組むことになる。

情報の処理構造は、言語学における談話構造に関する議論の中で、適切とはいえない用語ではあるが、情報構造と呼ばれているものにおおよそ対応する。情報構造では、文の構造を2つの成分からなる構造に見立てる [8]。旧情報成分と新情報成分、あるいは、既知情報成分と未知情報成分の2つである。

言語学においては、この情報構造をもう一歩進めて、文連鎖にまたがる理解や伝達を対象にした議論の展開がある [9]。この議論では、3つの成分からなる構造が見立てられる。そう見做すことができる。質問対象成分と質問成分と回答成分、あるいは、予測対象成分と予測成分と結果成分の3つである。構造化日本語では、この3つの成分からなる構造を前提にして、構造化をデザインする。基本的な構造化として、< 対象 > < 質問 > < 回答 > という三つ組み構造を採用する。

5つの文の情報の処理構造は、その違いを反映し以下のようになる。対象成分は、それぞれの文の前に位置する文の回答成分である。そして、回答成分は、後に位置する文の対象成分となる。

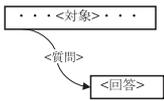
① [処理 : 対象 → < 【共有知識】 >

質問 → < (何ごとが) >

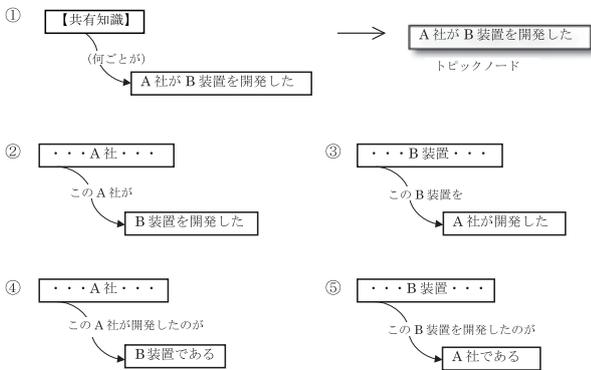
回答 → < A 社が B 装置を開発した >]

- ② [処理：対象→<・・・A社・・・>  
質問→<A社が(したのが)>  
回答→<B装置を開発した>]・・・
- ③ [処理：対象→<・・・B装置・・・>  
質問→<B装置を(したのが)>  
回答→<A社が開発した>]・・・
- ④ [処理：対象→<・・・A社が開発した・・・>  
質問→<A社が開発したのが>  
回答→<B装置である>]・・・
- ⑤ [処理：対象→<・・・B装置を開発した・・・>  
質問→<B装置を開発したのが>  
回答→<A社である>]・・・

処理構造の構造化をグラフィカルな表現形式で表わすと、以下となる。



5つの文に対するグラフィカルな表現は、以下となる。



## 4 構造化日本語の表現機能

前章でデザインされた基本的な構造化を土台として、より複雑な言語表現による情報の処理構造の構造化を紹介する。本稿の目的は、直観的に理解していただくことである。すべて、簡単な例を用いて説明する。

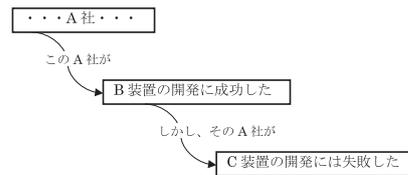
### 4.1 複文や連文に対応する構造化表現

論理的な関係が接続詞相当語句で表現される複文や連

文の構造化である。接続詞相当語句は、事象に対して、その論理的な関係を質問し、事象として回答得るという構造化として捉える。例えば、

「・・・A社・・・。A社は、B装置の開発に成功した。しかし、C装置の開発には失敗した。・・・」

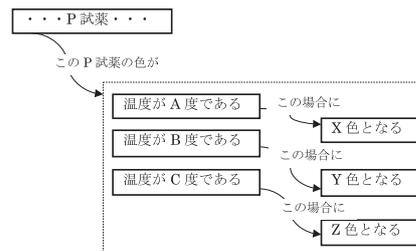
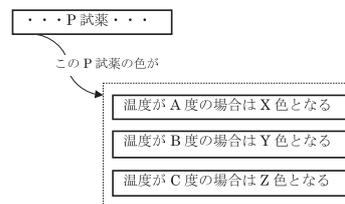
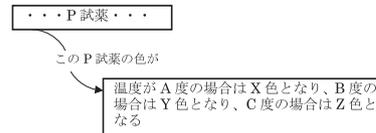
この言語表現は、以下のように構造化表現される。構造化日本語では、事象は、<質問>と<回答>の組として表現される。



もう少し複雑な条件分岐を表わす表現の構造化の例である。

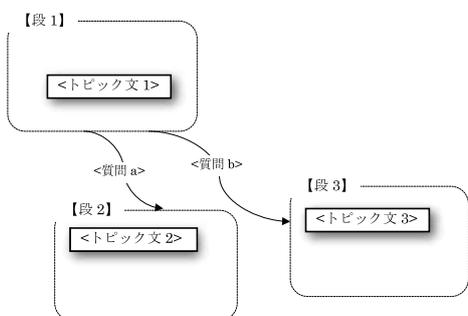
「P試薬の色は、温度がA度の場合はX色となり、B度の場合はY色となり、C度の場合はZ色となる。」

この例は、以下のような一連の構造化に対応付けられる。どの構造化を採用するかは、それぞれの目的による。鎖線ノードは、グループノードである。



### 4.2 文章構造に対応する構造化表現

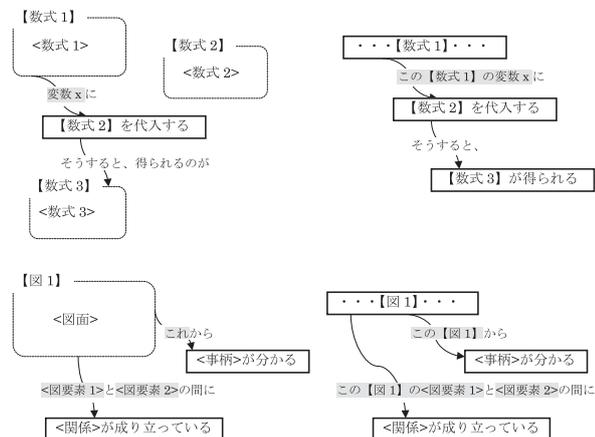
文章は、複数の段（パラグラフ）からなり、段は、複数の文（センテンス）からなる。段内の文は、話題文（トピックセンテンス）と支持文（サポーティングセンテンス）に分けられる。段のひとつが話題段（トピックパラグラフ）となる。このような文章構造も構造化される。例えば、3つの段（【段1】【段2】【段3】）からなる文章が、以下のように構造化される。角丸の鎖線ノードは、ブロックノードである。ブロックノード内の影付きノードは、トピックノードである。段をどう線状化するのか、段内の文をどう線状化するのか、そこには、各国や各文化圏で好まれる文章構成スタイル [10] が反映されることになる。



### 4.3 数式、図、表による情報に対応する構造化表現

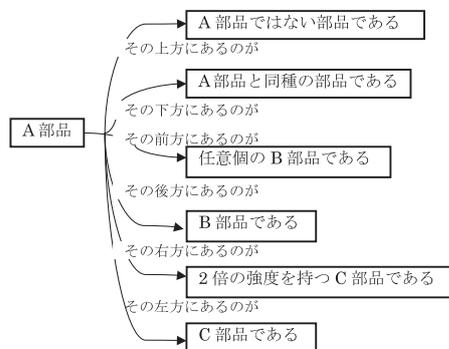
文章中には、言語以外の表現メディアによる情報表現が用いられる。数式や図面等である。それぞれの数式や図面は、それぞれがひとまとまりの情報を表現する。段（パラグラフ）と同様に、ブロックノードを用いて構造化が表現される。

例えば、数式と図面の例である。

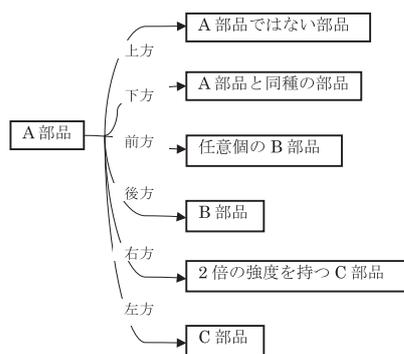


### 4.4 物理的構造体情報に対応する構造化表現

物理的な構造体は、構造体を認識するための一群の<対象><質問><回答>として構造化される。例えば、以下である。



この構造化表現は、簡略化の便法を用いて、<質問>や<回答>を以下のように語句化することができる。



### 4.5 論理的構造体情報に対応する構造化表現

#### (1) 物理的構造体の構造化を抽象化・一般化

質問対象となる<物理的事物>、質問となる<物理的関係>、回答となる<物理的事物>、これを<論理的事項>と<論理的関係>と<論理的事項>へと抽象化する。そして、知識表現における最も基本の枠組みである<対象><特性><特性値>へと一般化する。すなわち、



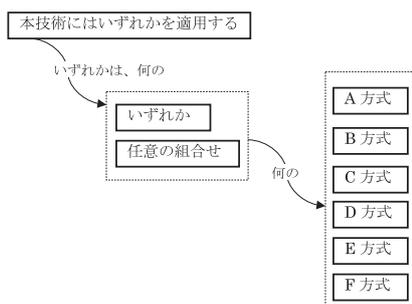
#### (2) 集合とその要素選択に対応する構造化表現

法令文等では、複雑な入れ子構造となる並立表現が多用され、それに対応するための便法的な表現形式が慣用化されている。例えば、ORの入れ子に対しては、一番

外側には「又は」を用い、その内側には「若しくは」を用いる。さらその内側がある場合は、再び「若しくは」を用いる。法令担当者は、入れ子となる2つの「若しくは」に対し、内側を「小若し」、外側を「大若し」と呼んで区別する（田島信威：法令用語ハンドブック改訂版、ぎょうせい（2005年））。これらは、入れ子構造を線状化するための苦肉の策である。

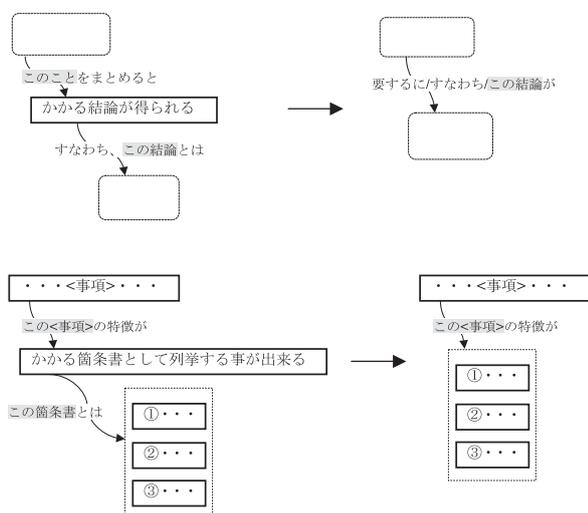
並立表現を集合とその要素選択（限量表現）の表現に一般化し、集合自身を要素とする集合へと入れ子構造を一般化することができる。構造化日本語では、簡明な構造化によって、明解な表現が可能となる。例えば、以下である。

「本技術には、A方式、B方式、C方式、D方式、E方式若しくはF方式又はそれらの任意の組合せを適用する。」



#### 4.6 構造化表現のマクロ化・縮退化

一連の<対象><質問><回答>の連鎖が定型的に頻出する場合がある。そのような場合には、その連鎖をマクロ化・縮退させた簡潔な表現を導入することが出来る。



## 5 構造化文章のテキスト処理

構造化日本語による構造化和文テキストに対するテキスト処理、すなわち、作成、検索、翻訳、要約について概観する。構造化の情報を活用することによって、通常線状化テキスト処理に比べ、高機能で高精度なテキスト処理が可能となる。

### (1) 作成

構造化テキストライティングは、以下の3つのライティングステップが交錯しあいながら展開するプロセスとなる。

#### ① 俯瞰と接近

テキスト全体の全容を把握するために俯瞰し、テキスト各部分の詳細を詰めるために接近する。俯瞰と接近がダイナミックに繰り返される。

#### ② 書き加え

不全感を感じるノードに対して質問アークと回答ノードを書き加える。感じ取られた不全感を質問として表現し、不全感を満たすように回答を書き足す。

#### ③ 書き換え

質問アークを付け替えることによって、部分構造化テキスト全体の移動操作がなされる。その際に求められる<質問>や<回答>の調整内容も明解に把握される。

いずれ、このライティングプロセスを支援するグラフィカルなエディターが用意されることになる。ワードプロセッサではなく、構造プロセッサである

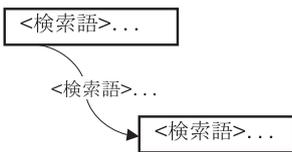
この構造プロセッサには、構造化テキストと線状化テキスト（通常のテキスト）との間の相互変換の機能も用意される。広く流通する文書では、現状のままの形式が継続することになる。したがって、構造化テキストを線状化する機能が必要となる。一方、膨大に蓄積されたテキストを再利用するために、線状化テキストを構造化する機能が必要となる。

構造化テキストの線状化の操作は、二次元グラフを

たどりながら構造化日本語テキストを読み下す操作となる。ひとつの構造化テキストに多数の、場合によっては、非常に多数の読み方、すなわち、線状化テキストが対応付けられる。そこで、基本的な読み方として、いくつかの推奨方略が用意されることになる。

### (2) 検索

構造化日本語テキストは、情報内容に即した構造を明示的にもつテキストである。この構造を利用することによって、通常のプレーンなテキストに対するより、より高度な検索機能が実現できる。基本となる構造化である<対象><質問><回答>という構造を利用した検索を設定することができる。検索エンジンには、以下のような構造化された検索要求文が入力されることになる。



### (3) 翻訳

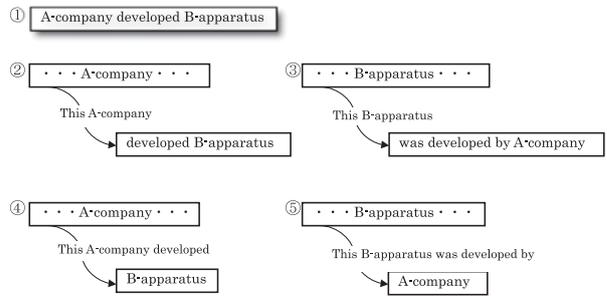
構造化日本語が表現する構造は、情報が本来持つべき構造である。したがって、基本的には、この構造は、言語に依存しない、あるいは、そう仮定してもそれほどの不都合は生じない。ただし、<質問>と<回答>という構造は、言語に依存せずに保存されるとしても、<質問>部分と<回答>部分の言語表現の方式は、言語に依存することになる。日本語のような主題（話題）優勢言語と英語のような主語優勢言語 [10] では、異なってくる。すなわち、

- 主題優勢言語：< 質問 >=< 主題成分 >
- < 回答 >=< 題述成分 >
- 主語優勢言語：< 質問 >=< 主語成分 >
- < 回答 >=< 述語成分 >

この<質問>部分と<回答>部分の表現方式を変換する仕組みを整備すれば、ほぼ現状の機械翻訳システムを用いて、構造化言語間に高精度な機械翻訳を実現することができる。なお、<質問>と<回答>の基本的な表現方

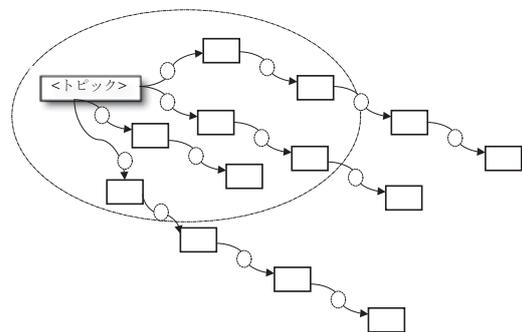
式のパターンは、20～30程度に整理されると思われる。

例えばとして、3.1節で用いた5つの例を英訳すると以下となる。



### (4) 要約

構造を利用することによって、高機能な情報抽出や要約が実現される。例えば、トピックノードからの（重み付けされた）距離によって要約部分を切り出すことも可能となる。



## 6 構造化言語系

日本語だけではなく、すべての言語に構造化を導入することが出来る。構造化英語、構造化中国語等々である。そして、構造化国際共通語を設定することができる。ここでの国際共通語は、コンピュータのための国際共通語（e 国際共通語と呼ばれる）である。

エスペラントを始めとして、人工的な言語による人が理解する国際共通語はことごとく普及に失敗した。国際共通語は、人が理解するものではなく、コンピュータに理解させるべきものである。人は、それぞれの母語を用



い、母語と共通語間の翻訳・通訳はコンピュータが自動的に行う。構造化国際共通語とすることによって、現状の機械翻訳技術・自動通訳技術でも、高精度な自動翻訳・自動通訳を実現することができるようになる。

コンピュータが理解する構造化された国際共通語を構造化 e 国際共通語と名付ける。構造化 e 国際共通語は、プログラミング言語などのいわゆるコンピュータ言語ではない。コンピュータのための自然言語、コンピュータの母語である。人間のための自然言語、すなわち、人間の母語には、国や地域の特性を反映しながら人間活動や社会活動の微細にいたるまでの表現能力や伝達能力が求められる。一方、コンピュータの母語、構造化 e 国際共通語には、まずは、そのような微細な表現能力や伝達能力は必要無く、国際的に共有化すべき情報を表現できればよい。

人々は、構造化 e 国際共通語の介在を意識することなく、それぞれの構造化された母語で全世界とコミュニケーションできる。情報サービスや情報システムの人間と係る部分には、構造化 e 国際共通語が用いられることになる。すべてのサービスやシステムは、一気にグローバル化を達成することができる。ローカライゼーションの手間隙は必要なくなる。

なお、構造化はされていないが、e 国際共通語としては、国連大学の関連機関が推進する UNL (Universal Networking Language) [11] がある。UNL は、CWL (Common Web Language) と呼ばれることもある。構造化 e 国際共通語は、UNL の多くを参考に進められることになる。

## 7 特許文章ライティングでの利用

構造化日本語の効用を実感できる例として、特許文章ライティングにおける請求項文ライティングでの利用を取り上げる。請求項文は、慣例的に、複雑な入れ子構造となる名詞句の形式に表現される。この名詞句形式は、英文クレームの記載形式を引き継いだ結果と思われる。英文では、修飾句（動名詞句）が被修飾語の後に置かれ

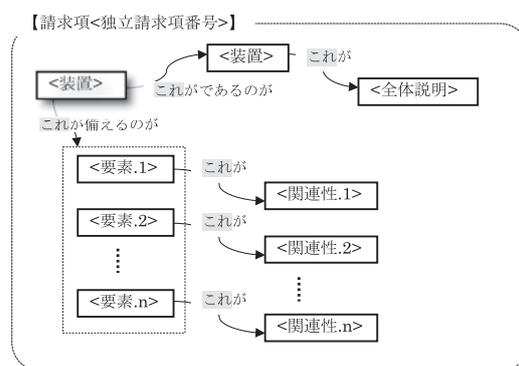
るのに対して、和文では、修飾節（連体節）が被修飾語の前に置かれる。この言語の構文上の違いによって、英文クレームが明晰さを失わない名詞句となるのに対して、和文請求項は二文制約（容認性を損なわないように中央埋め込みを2文以内とするという規則）に反する非明晰な名詞句となる。

このような請求項文をそのまま用いると、知財専門家が作成する過程、知財専門家と発明者とが議論する過程、知財翻訳者が翻訳する過程、いずれの過程も効率の悪い、誤りを生じる危険性を含んだコミュニケーション過程となる。そこで、これらの過程に、構造化日本語で請求項の記載内容を書くというステップを組み込む。そのことによって、請求項文の作成や変更の作業、さらに、翻訳等の作業が格段に明晰で効率の良いものとなる。

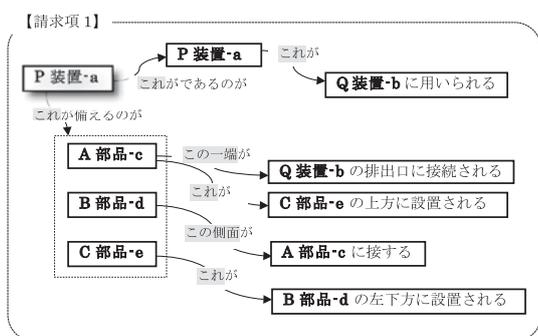
構造化和文テキストである請求項を読むことによって線状化された請求項文が得られる。読み方にいくつかの方略を設ける。それぞれの方略にしたがうことによって、名詞句形式の請求項文、複文形式の請求項文、連文形式の請求項文等々が得られることになる。名詞句形式は、特許請求の範囲の請求項文として、複文形式は、請求項の翻訳原稿文として、連文形式は、請求項文を明細書や要約書内で利用する際の文として、それぞれに適切な利用に対応することができるようになる。

以下では、基本的な仕組を分かり易く説明するために、簡略化された請求項パターンや意図的な作例を用いる。

まず、請求項の記載内容の構成パターンを（メタな）構造化日本語で書かれた 10 ~ 20 個の基本パターンに整理する。以下に示すのは、物に関する発明で、発明技術のみを取り上げ、構成要素を列挙する独立請求項の基本構成パターンである。



以下は、この構成パターンに当てはまる作例である。ただし、この作例には、装置構成としての意味的配慮はなされていない。太字で示される要素は、それぞれ同一事物を示す。添え字記号によって、同一であるもの同士が明示されている。本例のように自明な場合は、添え字記号を省略してもよい。



このように、構造化日本語で書くことによって、請求項は、格段に明晰となる。この構造化テキスト上で請求項の作成作業が行われる。技術要素をより少数に絞り込む、技術要素や関連性をより上位概念化する、関連性に矛盾が無いことを検証する、等々の検討が加えられる。

記載内容が確定すると、次は、この構造化テキストを読むことによって、通常の線状化テキストを得ることができる。読み方には、いくつかの方略が設けられる。例えば、「C 部品-e」に対する<質問>と<回答>に対して、以下のような読み方をすることができる。

- ㊤ <回答>を題述成分として読む  
「C 部品は、B 部品の左下方に設置される」
- ㊦ <回答>を連体節として読む  
「B 部品の左下方に設置される C 部品」
- ㊧ <回答>を形容詞句として読む  
「B 部品の左下方に設置の C 部品」

すべての<回答>を連体節として読み下すと、以下のような名詞句形式の請求項文が得られる。

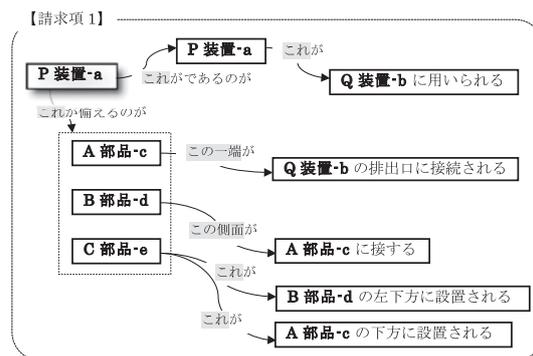
【請求項 1】  
Q 装置に用いられる P 装置であって、  
その一端が上記 Q 装置の排出口に接続され、  
かつ、下記 C 部品の上方に設置される A 部品と  
その側面が上記 A 部品に接する B 部品と、

上記 B 部品の左下方に設置される C 部品と  
を備える P 装置。

構造化テキストにおいては、太線で示されていた同一事物は、線状化テキストでは、先行詞と照応詞による照応表現に読み替えられる。すべての<回答>を連体節として読むという方略に従うと上記の下線部で示されるように後方照応表現が現れる。このような後方照応は、請求項文には馴染まない。そこで、この部分を題述成分として読み、最後に加えると、以下のような名詞句形式が得られる。

【請求項 1】  
Q 装置に用いられる P 装置であって、  
その一端が上記 Q 装置の排出口に接続される A 部品と、  
その側面が上記 A 部品に接する B 部品と、  
上記 B 部品の左下方に設置される C 部品と、  
上記 A 部品が上記 C 部品の上方に設置されることと  
を備える P 装置。

あるいは、記載内容の等価性が保たれるなら、以下のように「A 部品-c」に対する関連性を「C 部品-e」に対する関連性に書き換える。



すべての<回答>を連体節としてこの構造化テキストを読み下すと、以下のような名詞句形式の請求項文が得られる。

【請求項 1】  
Q 装置に用いられる P 装置であって、  
その一端が上記 Q 装置の排出口に接続される A 部品



と、

その側面が上記 A 部品に接する B 部品と、  
上記 B 部品の左下方に設置され、かつ、上記 A 部品  
の下方に設置される C 部品とを備える P 装置。

すべての<回答>を連用節の題述成分として読み下すと、  
以下のような複文形式の請求項文が得られる。この  
請求項文は、英文クレームの記載形式に直接的に対応す  
るため、翻訳原稿として用いることができる。

**【請求項 1】**

P 装置が、Q 装置に用いられ、  
P 装置は、A 部品と B 部品と C 部品とを備え、  
A 部品は、その一端が上記 Q 装置の排出口に接続され、  
B 部品は、その側面が上記 A 部品に接し、  
C 部品は、上記 B 部品の左下方に設置され、かつ、  
上記 A 部品の下方に設置される。

すべての<回答>を文の題述成分として読み下すと、  
以下のような連文形式の請求項文が得られる。この請求  
項文は、請求項文が明細書や要約書に転載されるとき  
の記載形式として用いることができる。短文化されてい  
るため、機械翻訳の入力原稿用として用いることもできる。

**【請求項 1】**

P 装置が、Q 装置に用いられる。  
P 装置は、A 部品と B 部品と C 部品とを備える。  
A 部品は、その一端が上記 Q 装置の排出口に接続さ  
れる。  
B 部品は、その側面が上記 A 部品に接する。  
C 部品は、上記 B 部品の左下方に設置される。かつ、  
上記 A 部品の下方に設置される。

**参考文献**

- [1] トニー・ブザン、バリー・ブザン著、神田昌典訳：ザ・  
マインドマップ、ダイヤモンド社（2005 年 11 月）
- [2] 特集「ポスト経験主義の言語処理」、人工知能学会  
誌（2012 年 5 月）
- [3] 佐々木正人：アフォーダンスー新しい認知の理論、  
岩波書店（1994 年）
- [4] ドナルド・ノーマン著、野島久雄訳：誰のための  
デザイン？ー認知科学者のデザイン原理、新曜社  
（1990 年）
- [5] ジャン＝ポール・ネリエール、ディピッド・ホン著、  
一般財団法人グローバル人材開発訳：世界のグロー  
ビッシュ、東洋経済新報社（2011 年 3 月）
- [6] マニユエル・リマ著、奥いずみ訳、久保田晃弘監修：  
ビジュアル・コンプレキシティー情報パターンの  
マッピング、ピー・エヌ・エヌ新社（2012 年 2 月）
- [7] 町田 健：言語構造基礎論 - 文の意味と構造、勁草  
書房（2011 年 10 月）
- [8] 堀川智也：日本語の「主題」、ひつじ書房（2012 年）
- [9] 石黒 圭：日本語の文章理解過程における予測の型  
と機能、ひつじ書房（2008 年）
- [10] 吉村公宏：英語世界の表現スタイルー「捉え方」  
の視点から、青灯社（2011 年 5 月）
- [11] Universal Networking Language ホームペー  
ジ、<http://www.undl.org/>

