

## 外国人児童・生徒のための教科書読解支援 View システム

中尾桂子<sup>1</sup> 森下淳也<sup>2</sup>

<sup>1</sup>大手前大学社会文化学部

<sup>2</sup>神戸大学大学院総合人間科学研究科

EMAIL: {nakao@ccs2000.cla., jm@}kobe-u.ac.jp

教科書特有のストリームを日本語で認識する能力が乏しい外国人児童・生徒に、その要点や筋を読み取る支援を行うため、テキスト中の重要な部分をクローズアップさせて内容の筋や流れを示す View システムを提案する。本来論述文である小学校の教科書には、学習指示のために、文末にモダリティー形式を持つ文が利用されている。この学習指示文に見られる表現形態を利用して流れを視覚化することで、文章の構造化が行えるため、日本語の文章構造に対する知識が乏しくても、構造化の指標を学びながら文章のストリームを把握することが可能になる。文章の構造化には、XML を利用する。このマークアップ機能を応用することにより、元データを復元可能な形式で、データ構造自体の再構成を行うことができることを示す。

### 1. はじめに

この 10 年、日本の社会的、経済的変動に伴い、日本の小学・中学に編入する外国人児童・生徒が増加してきた。外国人児童・生徒の多くは日本語がわからないため、編入先の小学・中学で日本語指導を受ける。

しかし、一般に知られるように、日本語と外国語とは論旨や論の組み立て方や論拠の立て方が異なるということは、テキスト構造やディスコースストリーム、また、トピックにまつわる知識（トピックとトピックに関連する事物—場面スキーマ）、ディスコースの場面やその前後の文脈に関連した言語の背景的な知識が異なるということである。通常、ネイティブは、就学年齢に達するまでに、1つ1つのことばを覚える過程でことばと背景的な知識とを関連させながら習得していき、学校教育を通して、それらの知識を確固たるものとする。

ところが、外国人児童・生徒は母国語での学校教育が未完全で来日している。また、日本語特有の言語的知識が少ない。そのため、特に、論述や説明の多い学校の教科書というアカデミックな場面で使われる日本語は、外国人児童・生徒にとって難しい。日本語のテキスト構造に慣れていない外国人児童・生徒にとっては、文法や語彙が理解できても、抽象的な思考を述べる日本語の文章の流れがつかめなため、教科書で表わされる概念や文章内容の理解が困難になるのである。中でも、低学年で来日した外国人児童・生徒は、母国語でも

日本語でも時期的に言語能力を発達させるための教育が不十分であることから、言語能力に問題が生じやすい。

このような問題に対応するためには、文法や語彙指導だけでなく、読解や作文指導において、テキスト構造やディスコースストリームを掴むための訓練を行う必要がある。また、教科書学習に重点を置いてその内容を理解する支援を行うためには、教科書のテキスト構造やディスコースストリームが掴める手助けが必要になる。しかしながら、現在、そのような指導を行うための指導材料はほとんどない。そこで、本稿は、教科書特有のストリームを日本語で認識する能力が乏しい外国人児童・生徒に、その要点や筋を読み取る支援を行うため、文章の表現形態の解析結果を応用してテキスト中の重要な部分をクローズアップさせて内容の筋や流れを示す View システムを提案する。

### 2. 小学校教科書の構造とその視覚化

テキスト中の重要な部分をクローズアップさせて内容の筋や流れを示す指標は、テキスト構造レベル、文単位の意味的な構造レベルにある。視覚化は、このテキスト構造を示す指標に基づいて意味的な構造を示す指標を検索し、検索の結果をフロー形成部分としてテキストに射映することで行う。

テキスト構造レベルの指標としては、テンズとモダリティーといった文法形式があげられる。意味的な構造レベルの指標としては名詞や動詞といった単語があげられる。以下、それらが指標となる根拠を述べる。

## 2-1. 日本語の文章タイプを区別するテンスとモダリティーの特徴

テキストタイプの違いにより、テンス形式の現れ方に違いがあることについては、よく知られる事実として内外で研究されている[3]. 池上(1986)は同一時点の「語り文」における時制の交替について、出来事を前景化したり、背景化したりするのに利用される点が諸外国語でも観察される言語的特徴であることを紹介している。修辞学上で「歴史的現在」と言われるこの時制の交替については、多くの先行研究で以下のようなパターンが報告されているという。(a) 英語等の言語では通常は30%程度の「歴史的現在」が現れるが、直接話法と関連する状況では63%の利用率になる。(b) 目まぐるしい時制の交替はなく、切り替わったならば同じものが続く。(c) 「過去時制→歴史的現在→過去時制」の順に切り替わる。(d) 「歴史的時制」そのものよりも、切り替え自体に意味があるようである。(e) 従属節-主節間、同一節内で同じ主語を持つ動詞の間では起こらない。(f) 主語のある文においてのみ起こる。主語が省略されることが許される言語においても主語省略時には起こらない。

ただし、日本語の「語り文」における現在形の頻用には諸外国語には見られない現象が観察されることから、池上(1986)は日本語特有の言語的特徴と関連した複雑な要因があると指摘する。池上(1986)は、「語り文」における日本語の時制の区別は過去と現在と不定時が発話時間に融合する傾向にあるとしているが、その中で、日本語の「語り文」における時制の具体的な特徴として、以下のような点をあげている。(ア) 状態動詞は現在形になりやすい。(イ) 過去の出来事を現在言及する場合、過去形の文を現在形の助動詞で包むという二重構造を取る。(ウ) narration と comments の区別があいまい。(エ) 語り手と登場人物と一般不変現象との区別は言及しない。(オ) 主文の時制が優先するため従属節の時制と一致していなくてもよい。また、日本語では時制が目まぐるしくかわる点はこの5つの前提として諸外国語には見られない特徴だとしている。

以上から考えると、テンスの交替は、述べられる出来事の時点を語り手の発話時点に摺り替えたり、臨場感を醸すために利用される修辞学上の効果の一つであるが、テキスト構成における統語的な面との結びつきが強く、それは文の形態から判断できるものである。テ

ンスの役割として考えれば、出来事を前景化したり、背景化するといった内容における構造化の役割を担っていると考えられる。

一方、眼前の相手と直接会話する場合の発話ディスコースと、不特定多数の相手に対して行う発話や書物におけるディスコースとは、テンスの現れ方に違いがあること、また、モダリティーの現れ方にもディスコースタイプの違いが反映されることについても先行研究で指摘されている。仁田(1996)は、一方的な非対話状況で形成される「説き語り」のテキストについて、説き語られた内容の類型とモダリティーやテンスなどの表現機構の異なりから「語り物」(池上1986のいう「語り文」)と「論述」の2タイプに分け、テキスト内容の展開に大きく機能している認識的なモダリティーに着目して分析することで、テキストタイプを形成する機構の一部にモダリティー表現が関わることを示している[4]. それによると、「語り物」では、モダリティーを顕在化しない文と点在するモダリティー形式を取る文とがあり、両者は、小説の地の文と会話の文のような関係にあつて、テキスト展開におけるレベルの変化を生じさせ、語り手の目から見た登場人物の心内発話等を表すと言う。モダリティー形式を取る文は、主節の展開とは異なった機能を担い、テキストを「語り物」たらしめているという。そして、このモダリティー形式の出現位置を追うことで、「語り物」中で、テキストレベルの変化する領域が明示できることを示している。一方、「論述」では、モダリティー表現はこのタイプのテキスト形成には大きな機能を担っていないが、従属節のうち、条件節が「論述」というテキストタイプを形成するために重要な役割を担い、同じく従属節ではあるが、原因節は、テキスト形成には影響せず、主節を展開させる文として働いていると述べている。これら2つのテキストタイプでは、テキストタイプを形成する機構として、文の構造や文末の形態が、表現されている事態の想定性と事実性において大きな役割を担うと指摘している。

上の2つの指摘にもとづけば、モダリティー形式のあらわれる文とテンスの変化を追うことにより、テキストの流れがつかめることがわかる。そこで、テキスト(ディスコース)を具体的に視覚化するためのテキストの構造的なパターンや文章間の関係を示すものとして、テンス、モダリティー、条件節をテキスト構造を捉えるための指標とする。

## 2-2. 読解に必要なスキル

一般的な読解指導では、一字一句逐次理解する「精読」の他に、文章の筋やテーマを掴むことを目的として「多読」や「速読」が指導される。特に、大要を把握する「掴み読み」や一定の情報だけを探して拾い読みする「検索読み」などが情報収集能力の向上につながるとされている[5]。この「速読」や「流し読み」をするには、構造やディスコースの流れを知っている必要がある。外国人児童・生徒は日本語のテキストの流れに関する知識が不足しているために流れがつかめないのであるから、内容やテーマに関係する単語のフローを見ながら、概念のイメージを掴む訓練を行うことが内容理解のための有益な訓練や支援となる。

これらのことから、内容に直接関係のあるフローを掴むことが概念を理解するために必要であるとして、内容と直接に関連する単語の配列を、フローを掴むための指標とする。

詳しく読み込むには述部のフローを名詞と関連させながら見ていくことが必要であるが、概念イメージを掴むには名詞に着目すればよい。そこで、名詞と、名詞の関係性を詳しく把握するために動詞の2種類の品詞に焦点を当てて、その配列をクローズアップする。

## 2-3. 小学校教科書の文の文末形態

テンス、モダリティーは文末に現れるもので、その意味や機能により、形態的に区別できる。小学校教科書において、どのようなものを実際にディスコース顕示の指標とするかを決定するために、中尾・森下(2001a)は小学校教科書内の文末の形態を調べた[6]。

その結果、教科書のディスコースは、算数や理科といった教科の別なく、いずれの教科においても話者の意図を表すモダリティー表現が利用されていた。話者の意図を表す表現とは、この場合、終助詞や助動詞、接尾辞などを組み合わせて利用した表現の形態的な特徴を持ち、意味として、勧誘、問いかけ、誘導、疑問、義務、禁止等、一般に相手のある発話場面で話者の意図を示すために利用される表現形式である。これらモダリティー表現の教科書での使われ方を調べてみると、ディスコースの本流の始めや終わり、または、その本流とは離れた場所で利用されることが多いということが観察された。

## 2-4. 小学校教科書の構造と文章の流れ

そこで、次に、小学校教科書の本流や支流がこれらの表現とどのように関わっているかを

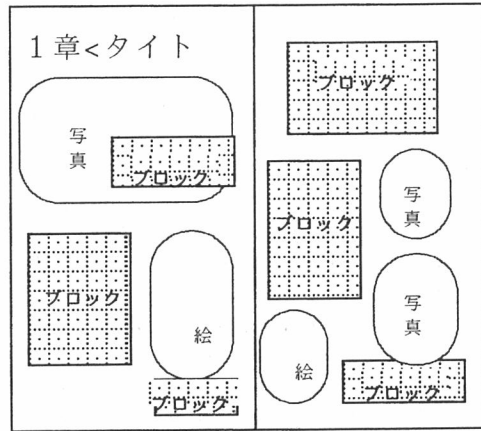


図1:教科書の見開き図

観察し、この文末形態とテキストの構造との関係を調べた。

小学校の教科書はいずれの教科にも単元といわれる章がある。教科により文章の長短に違いはあるが、各章節以下の階層では、意味的に関連のある文が小さな文章のまとまりに分けられ、二次元に配置されている。長いものが「本文」や発展的な「補足説明」にあたるが、この文章の一まとまりをここではブロックと呼ぶ。その状態を図1に示す。

このような小学校教科書の章節構造に基づいてモダリティー形式の現れる位置を観察すると、まず、モダリティー表現を持つ文は、教科書で展開される学習活動というディスコースにおいて、手順や考察、注意等、読者である児童・生徒に指示を行うための文に現れることが多い。モダリティー表現は、教科に関わらず、ブロックの最後尾の文末に見られる。ブロック末にモダリティー表現を持たないブロックで

A:指示・呼び掛け	・・・ましよう。てみよう。よう
B:命令	・・・なさい。
C:指示・疑問	・・・だろう。だろうか。
D:指示・注意	・・・てはいけない。 ・・・なければならない。
E:説明	・・・普通体
F:解説	・・・丁寧体
G:誘導	・・・終助詞
H:アドバイス	・・・～とよい。

図2:指標とする文末形態

あっても、そのブロックの外側のどこか近くに、モダリティー形式を持つ文だけで構成されるブロックがある。そして、文末にモダリティー表現を持つ文の前後は、手順の説明や考察、実験などが説明されたり、それらの結果が表わされたりしている。以上のことから、文末にモダリティー形式を持ついくつかの文は、ブロックどうしを関係付け、学習活動という内容の流れを形成する要となっていると考えられる。

そこで、中尾・森下(2000)は、特にディスコース形成に大きな役割を担う図2に示すような文末のモダリティー表現をブロック単位の流れを構造化するための指標とし、文末にモダリティー表現のあるブロックをマークすることでその文末表現の意味機能から談話の特性を規定し、理科の教科書において学習の流れを明らかにした[1]。

## 2-5. 学習の流れと個々の文の関係

小学校教科書の学習の流れに基づいた構造のパターンと、文末のモダリティー形式との関係を中尾・森下(2000)で明らかにした。しかし、これは、学習の流れを教科書という書物の中で大きく捉えるために必要なことであるが、流れはブロック単位で見出すものであった。

読解における「検索読み」「拾い読み」といった速読の指標は、単語である。大きなディスコースの構成を把握しながらも、ディスコースの内容の流れを読み取るには、単語の現れ方と内容の流れとの関係を明らかにする必要がある。具体的な内容に関わるディスコースを明らかにしてさらに具体的な情報を得るためには、文中の語レベルでの文脈情報を把握する必要がある。

そこで、中尾・森下(2001b)は、ディスコースの下位構成要素である単語単位の並びにおける形状パターンを解析し、章節構造により階層化された教科書の文章において、文章を構成する文と文の関係を、単語(名詞、動詞)を指標に、文頭からの出現順序を観察した。そして、単語の出現パターンと前後の文配列との比較した結果、(1)名詞が概念理解への視点の移動を示す道標となっていること、(2)名詞、動詞の出現分布が学習概念の性質を説明する文脈であるか、学習概念の意味を説明する文脈であるかと関係があること、(3)章節構造とブロックの関係が単語からも明らかになること、(4)文脈を主になって形成するブロックとそうでないものがあり、文脈判

断の指標となるブロックとそうでないものとを区別することが具体的に可能であること、を明らかにした。

以上、2-4, 2-5 に述べた中尾・森下(2000, 2001ab)で確認された結果を利用し、教科書の構造と文末形態と名詞ディスコースの流れの視覚化には、図2に示すような文末形態、「～と、～ば、～たら、～なら」の条件節、それらの出現位置をテキスト構造視覚化の指標とする。

## 3. View システム

### 3-1. 小学校教科書の XML 化

構造との関係に基づいて、教科書の構成要素である談話や文の特性を明らかにするために汎用のデータ形式、XML(eXtensible Markup Language)を利用する[6]。XMLはデータを階層化し、データの持つ上下関係や順序関係を木構造で表現するデータ形式である。これはデータをマークアップすることで構造化するが、その際、元のデータを壊すことなく階層構造を表現することができるものである。また、その構造の解析や変換を行なうソフトウェアが充実しているため、階層に依存する探索や順序関係を抽出するのに適している。つまり、多様な目的に応じてタグを数種類用意し、それらのタグをテキストに付け替えたり重ねたりして階層を再構成することが可能である。また、検索の結果をさらにデータの下位構造として位置づけることが可能である。

小学校教科書のXML化にあたっては、教科書自体を視覚的にイメージすることができるように、ページという単位のタグや文章の横に配置されて重要な情報を与える役割を持つ絵や図等に対しても、それらの位置を示すタグを用意する必要がある。ただし、小学校の教科書は、著作権が多重にかかっているため、実際の教育活動の場でのみ利用が許されるものである。本稿では、テキストのみを紹介するにとどめるが、現場で使用を試みる際には、Web上に教科書自体を提示して行いたい。

### 3-2. アノテーションタグの導入

XML化された教科書に基づいて、構文情報や語彙情報を活用するためには、元々の教科書データに対して解析を行なう必要がある。そして、個別の解析結果と教科書の構造内における出現順序に基づいて評価するために、効果的な解析結果の表現方法を考えなければならぬ。そのため、教科書の構造を表すタグとは別に、解析結果を格納するタグ、アノテーションタグを導入する。

アノテーションタグは、構造記述のタグとは異なり、その随所に現れるものとして、定義される。即ち、解析の対象となるものに随時付けられるようにする。このタグの特性は、HTMLにおける<A>タグ(ハイパーテキストリンク)や<EM>などの装飾タグと同様である。

また、教科書データとアノテーションされた部位に対する解析情報とを区別するために、解析情報はテキストノードではなく、属性としてタグに付与することにする。

これによって、アノテーションタグは、教科書の構造記述タグとは独立に付与でき、且つ、アノテーションタグを取り除く、或いは不可視のものとする事で、元の教科書を復元することが出来る。

内容と直接関係のある単語をクローズアップするには、単語の機能的な性質を規定した品詞情報を利用しなければならない。そこで、XML化したテキストデータの日本語の部分に対して形態素解析を行い、同時に、形態素解析の結果得られる品詞情報を文の構成要素である単語の下位に関連づけて格納する。解析は chasen2.2.8[7]利用する。

元データに対して形態素解析を行い、その

1語ずつにXMLを付与することは、一文中における単語単位まで構造化することである。また、形態素解析の結果得られる品詞情報を単語の下位に組み込むことは、元データの各々の単語に対して属性を補うということになる。このような方法を採用することで、文中の出現位置を正確に保持したまま単語単位まで構造化された下位に品詞情報を保有する文データが得られる。これは、XMLに基づいたデータ形式を再構成し、拡張するという事である。しかし、XMLを利用して元データに属性を埋め込んでも、元データは復元が可能である。属性も取り外し自由である。

### 3-3. 読解支援 View システムの構成

小学校教科書アーカイブには、章、節、ブロック、段落、文、単語といったカテゴリーで書物を階層化するタグ、ならびに、文の構成要素を単語単位で区切り、各要素下に品詞情報を組み込んだ文中の単語レベルでの構造化タグといった意味的に2種類の性質に分けられるXMLタグをつけている。また、文末にモダリティ表現が見られる文を最後尾に持つブロックには、各文末表現に応じた意味機能を示す属性を付与している。XMLタグつき例を図3

```
<?xml version="1.0" encoding="Shift_JIS" ?>
<textbook>
:
<section><title>物の燃え方と空気</title>
  <block bkind="story">
    <part pkind="explain_body">
      <paragraph>
        <sentence skind="explain_body">は .. けが出ることもある。 </sentence>
        <sentence skind="explain_body">し .. く燃えるようになる。 </sentence>
      </paragraph>
    </part>
  </block>
  <subsection><title>火と空気</title>
  <block bkind="story">
    <part kind="explain_body">
      <paragraph>
        <sentence skind="explain_body">はんごうすい .. ている。 </sentence>
      </paragraph>
    </part>
  </block>
  <block bkind="inst_question">
    <part kind="inst_question">
      <paragraph>
        <sentence skind="explain_body">「隙間を作ると .. どうしてなのだろう。 </sentence>
      </paragraph>
    </part>
  </block>
  <block bkind="inst_question">
    <part kind="inst_question">
      <paragraph>
        <sentence skind="explain_body">ロウソク .. 付被せる。 </sentence>
        <sentence skind="explain_body">一方の瓶には、蓋をする。 </sentence>
        <sentence skind="inst_question">ロウソクの .. うなるか。 </sentence>
      </paragraph>
    </part>
  </block>
:
```

図3:XML文書

に示す。

XML を利用した検索の条件は基本的に自由である。しかし、学習内容の筋や話の流れを示すには、いくつかの条件に該当するものを複数定義して検索する必要がある。そこで、ディスコースの流れを効率良く提示するために、通常の検索条件指定方法に加え、学習の流れを示すために必要な条件を、あらかじめ項目毎にユニットにまとめた「検索条件指定メニュー」を用意する。

現在指定できるユニットは、先行研究でディスコースの流れを明示することが可能だと判断される構造上と内容との2種類の指標に基づくものである。一つはブロック単位での学習の流れを示す学習支持表現で、もう一つは単語単位での意味的な概念イメージの流れを示す文中の名詞と動詞の配列である。ブロック単位で視覚化するのは、主流ブロック、主流ブロックと副次的ブロック、副次的ブロックのみ、といった3種のユニットが考えられる。また、文単位では、学習指示を表わす表現、その学習指示文と対応する動作、学習指示文と活動後の結果、名詞の流れ、動詞の流れ、名詞と学習テーマとの関連性、動詞と学習活動との関連性があげられる。さらに、これらを

組み合わせたユニットや、ブロック単位と関連づけたユニットも考えられる。しかし、これらは、実際の使用の結果再構成されるものであるため、本稿では、ブロックレベルでは主流ブロックと副次的ブロックとを合わせた全体的な流れを見るのみにとどめる。

条件を指定する単位は、章、本、学年(上下2冊)、教科といった単位から選択するようになる。これにより、単語の配列を追ったり、同じテーマの学習内容において、学年別の単語の使用状況を追ったり、各教科における学習指示文の状況を比較するといった目的に応じた結果に焦点を当てやすくする。また、教科単位で比較する場合は、各単位のタグを横断して結果を返すため、行や列に関係なくマトリックスの作成が容易になる。

以上のように、ユニット単位に分けて検索の準備をし、それを選択するためのメニューを利用して検索条件を指定することにより、希望する教科書の見せ方が選択される。これにより、XML タグが自動的にテキストに付与され、その自動付与されたタグに基づき、教科書が表示されるため、システムの使い方に慣れていなくてもいくつかの文脈が提示される様子を見ることが可能である。

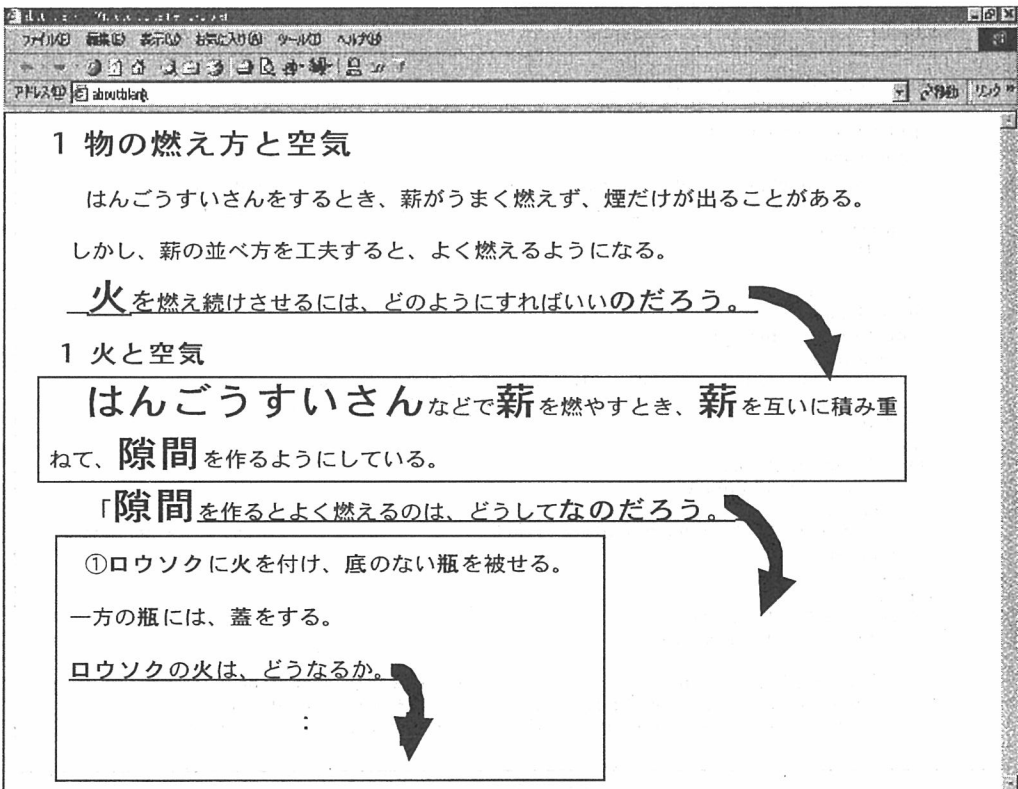


図4:情報を追うためにクローズアップさせた文章の流れ

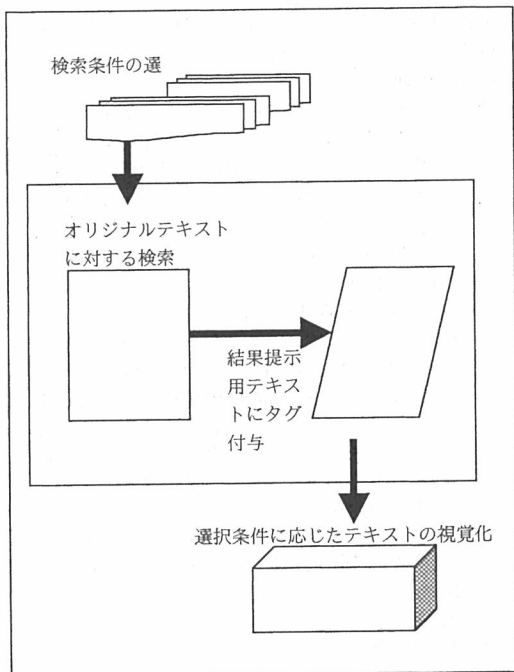


図 5: 検索から結果までの流

表示に際しては、大きさや色を検索条件に基づいて変更する。アンノテーションに対する検索の結果は、構造ではなく、表示の強調を返す。言い換えれば、XML で取り外しが自由な属性を補うことにより、データ利用者のニーズに応じた意味が表示に反映されるのである。検索結果の一例を図 4 に示す。

検索の結果を表示する際には、性質の異なるタグを数種類併用することになる。処理における時間的コストを低くするために、タグ付与のための専用のテキストを用意する。図

3 に示すような XML テキストに対して検索をかけるが、結果は、ユニットメニュー選択により自動生成されたタグを、毎回、別に用意した結果表示用テキストに自動的に付与して表わすのである。また、検索結果はタグ単位で保存する。検索の流れを図 5 に示す。

#### 4. 検索結果のグループ化

中尾・森下(2001b)で明らかになったように、単語の配列には、パターンがあり、そのパターンを形成する単語は一つの事柄を表すために利用される語彙としてグループ化できる。今回、日本語の統語的知識や文脈把握能力が乏しくとも、その流れを目で追えるような View システムを提案したが、これらの流れの中に見られる単語は、そのディスコースで表

される概念を表現するために用いられる語彙のグループである。言い換えれば、この語彙グループはその概念を学ぶために必要なものである。つまり、これらの語彙はある概念を形成するスキーマを示すものであり、教科学習においては、学習用語に相当するものである。これまでの小学校教科書の調査では、学習用語は頻度の多さや、学年や出版社を違えても利用が多いかどうかという延べ出現数の多さで定義しようとする流れが多かった。しかし、その頻度や出現分布の広さでは、ある概念を形成する語彙の重要性を十分に表し得るものではない。本稿で試みたように、学習指示文と関連づけた名詞の配列は、ある概念を説明する主要ストリームにおける語彙をグループ化したものである。その頻度や他学年での出現を参照せずとも、学習用語として捉えることができるという利点があり、学習用語の規定に重要な意味を持つと考えられる。このグループ化についての詳細は別稿に譲る。

これまでの言語学的なテキスト研究により、テキストタイプの区別は一般的に認められているが具体的なテキストタイプを各々のテキストに応じた形で詳細をパターン化することはほとんどなされていなかった。しかし、今回、小学校の教科書の文に対してその構造を元に流れを顕示する手法に基づけば、テキスト構造を具体的にパターン化する可能性が明らかになる。これは、日本語教育への応用だけでなく、学力が低いとされる日本人児童・生徒にも応用できる可能性を示唆するものである。

#### 5. 目的に応じて成長するデータに対する XML の利用について

データが、情報として価値を持ったとたん、そのデータは、利用者のニーズに応じて詳細に分類され、また、それらに関連する情報により、複雑に構造化されていく。このように、個々の価値感に応じたデータは、各々のニーズに応じて、より個人にとって価値の高い情報となるべく、個々の利用者のイメージに応じてアドホックに成長していく。つまり、データを利用するということは、利用者の連想に応じ、元のデータに対して新たな関連データを追加したり、より詳細な基準に基づいて細分化されることである。また、大なり小なりデータが構造的に再構成されるということである。

従来、このように構造的に変化したデータは、元のデータから変更された時点で、新たな構造データとして保有されてきた。そして、デ

ータの構造自体が変化することが可能なデータ構造はあまり考えられてこなかった。しかし、元のデータをそのままにして、構造のみをその元データが保有する一つのデータとして保有することは、実際にデータを有効に利用していく過程では回避できないことであり、また、XML を利用すれば不可能なことではない。

XML は、近年、データを、テキストでありながら構造を表現できるため有益だとしてインターネットにおいて普及してきた。しかし、その本来のもう一つの機能であるマークアップに着目すれば、元のデータを保持したまま、その利用方法に応じて成長する構造を元データに従属するデータとして保有することができる。個人にとって意味を見出す部分は異なるが、各々の目的に応じて、データにマークすることができるし、その各々のマーキングの過程自体を、元データに従属するアトリビュートとして追加することができる。このようなアノテーション機能をフルに活用すれば、アドホックなデータの構造化が容易に実現するのである。

今回、本稿では、XML で付与されたテキスト構造に基づき、そのテキストというデータを吟味した。そして、意味的な文の構造である品詞を利用して、テキストの中の必要な部分を取り出し、その検索結果得られた情報を、元のデータに再度射影することにより、重要な意味を持つ部分の表示を強調するという使用方法を試み、このような XML の利用法の有効性を確かめることができた。今後、さらに、この XML のマークアップという機能を応用し、元データの構造を保持したまま、独自の利用に際して構造を再編成し、データを更に利用しやすいように自在に組換えて、様々な構造のデータ利用を試みていきたい。

## 6. まとめ

以上、文のテキストタイプを構成する形態的な特徴を指標として文脈を提示する半自動化読解支援システムを作成した。しかし、この読解支援 View システムを利用して表示される教科書の有効性は、実際の学習活動に応用して検証しなければならない。また、今回自動生成のために利用した文脈情報は 2 種類である。今後他の文脈情報の可能性を検討する必要がある。以上は今後の課題である。

本システムで表示するのはテキスト部分のみであるが、教科書は、写真や絵などの情報価値も高いため、実際、教育現場において有効性

を検証するにあたっては、これらを含めた教科書全体を表示して行う。

### [参考文献]

- [1] 中尾桂子・森下淳也，“日本語教育のための表現意図出現パターン調査における文書データベースと XML の活用”，情報処理学会シンポジウムシリーズ，人文科学とコンピュータシンポジウム論文集 Vol. 2000, No. 17, pp. 177～184, 2000. 12. 16
- [2] 中尾桂子・森下淳也，“文章ブロックの構造化における形態素タグと XML の活用—年少者日本語教育への応用にむけた小学校教科書の最小文章単位のパターン化—”，情報処理学会シンポジウムシリーズ，人文科学とコンピュータシンポジウム論文集 Vol. 200, No. 18, pp. 173～180, 2001b. 12. 15
- [3] 池上嘉彦，“日本語の語りのテキストにおける時制の転換について”，語り—文化のナラトジー，記号学会研究，1986, 日本記号学会編
- [4] 仁田義雄，“語り物の中のモダリティ”，阪大日本語研究 8 pp. 15-27, 1996
- [5] 日本語教育学会編，“読解指導”，日本語教育事典 pp. 604-605, 1982, 大修館
- [6] 中尾桂子・森下淳也，“年少者中級日本語教育における読解指導用シラバス検討のための小学校教科書の文章構成の解析方法”，国際文化学第 4 号, pp. 133～151, 2001a. 5, 神戸大学国際文化学会
- [7] <http://www.w3c.org/XML/>
- [8] <http://chasen.aist-nara.ac.jp/index.html>