

小学校教科書の学習者に応じた電子的再構造化について

中尾桂子[†], 森下淳也^{††}

神戸学院大学人文学部[†], 神戸大学国際文化学部^{††}

nakao@ccs2000.cla.kobe-u.ac.jp

jm@kobe-u.ac.jp

教科書読解には教科書の文脈スキーマを認識する能力が不可欠である。そのため、言語的な弱者には、重要な教科書スキーマをクローズアップすることにより、読解能力を補助する必要がある。そこで、学習者のレベルや問題点にあわせられるように教科書を電子的に再構造化して提示するシステムを構想している。本稿では、文章中の思考活動を誘導する単語の流れや学習の流れを形成する文章構造と、それらと同価値に扱われている画像を同列に扱う場合における教科書内のコンテンツの再現性、再構成化を論じる。

Resequencing of a digital textbook according to the pupil's request

Keiko Nakao[†], Jun-ya Morishita^{††}

[†]Faculty of Humanities and Sciences, Kobegakuin University

^{††}Graduate School of Cultural Studies and Human Science, Kobe University

We discuss about the resequencing of a digitized contents in schoolbook. The schoolbook is constructed with small text discourse and many pictures. And some picture is constituted semantic content as much as text part in discourse. We make use of these equivalent digitized contents to realize a reconfigurability of schoolbook according to the foreign pupil's reading ability in Japan.

0. はじめに

一般に、文章に書かれた内容を理解するためには、文章の文脈をつかみ取ることが重要な課題となる。文脈を掴むには、読解力が不可欠であるが、さらに、読解には、文脈の意味的な構造形態であるフォーマルスキーマを認知する能力が重要である[1]。

ところが、日本語を母語としない非漢字圏の外国人など日本語の理解力がなんらかの理由で乏しいといった言語的弱者の場合、日本語自体の語彙的、文法的知識や文化的背景についての知識が少なく、日本語特有のフォーマルスキーマを知らないために、文脈を把握することが難しい[2]。特に、言語的弱者が日本の小学校に編入した外国人児童である場合、母語においても読解力

が身につけていないため、日本語の読解力やフォーマルスキーマの認知能力自体を開発していかなければならない[3]。

しかし、日本の教育制度の都合上、外国人児童には、言語的知識や社会的知識を一から積み上げていく程の時間的余裕はない[4]。そのため、乏しい知識を利用して小学校の教科書を読解するためには、日本語の教科書のフォーマルスキーマ認知能力を補助することが重要になる[5]。

そこで、我々は、小学校教科書の構造化により、フォーマルスキーマの視覚化を目指し、教科書内のコンテンツの可逆的パーソナライゼーションに取り組むことにした。そのために、可逆性と再構造化に適したデータ形式として XML を採用し、ニーズに基

づき評価した結果を別の形式に変換して提示させることや文脈という意味的な構造の明示について検討してきた。

テキストの内容における意味的な構造を視覚化するためには、通常データのやり取りに利用される汎用的な文書用 XML フォーマットを利用するだけでは不十分である。

それは、文脈構造が、書物の章節構造とは違った次元の構造であるため、文章の内容に応じて変化するものであること、また、構造化の指標が文中に散らばっているという性質を持つことが原因である。

したがって、XML 化したデータの意味的な文脈を構造化するためには、書物として構造化するタグに加えて、構造を形成する骨格となる指標にマークアップし、各々の指標の文脈構造化における機能に基づいて属性をつけることが不可欠となる。

しかし、意味的な構造化のためにデータを XML 化するには、テキストの構成要素全てに対するマーキングと、書物の構造のためのマーキング、さらに文脈構造指標のマーキングといった 3 種類のタグをつけなければならない。①指標の基礎調査、②データに特化した緻密なタグのデザイン、③相当数のデータに対するタグの付与といった 3 つの点で、時間的、人為的コストが高くなるという問題がある。

しかしながら、意味構造の指標にまでマークアップした XML データを利用することで教科書のもつ種々の構造を解明することができる。

先行研究で明らかにしてきたように、教科書特有のフォーマルスキーマと照合する読解力がなければわからない文脈構造や [6]、文章表現上の工夫 [7]、教科書をパー

ソナライズするためにこれらの利用の可能性を示すことができた [8]。

現在の教科書は、一連の先行研究で明らかにしたように、視覚的には文章と画像が同列に配置されており、画像が重要な情報を与えていることが多い [6]。また、文脈構造のパターンにいくつかの例外があり、さらに、書かれていない情報を、教科書中に提示された情報に基づいて読者が補って読むという「推測読み」が要求されることも多い。その上、推測読みの際には、文章中の情報だけではなく、図や写真といった画像で表現される内容が重要な意味を持ち、文脈の構造化にまで大きく関与することがある。そのため、視覚的な情報操作を行って文脈構造をクローズアップする場合、学習すべき内容を把握する上で意味がある図や写真も、テキストに組み込んで構造化することができるようにしておくべきである。

それには、教科書のテキストと画像、則ち、コンテンツの全てを用いて、教科書をありのままの形で再現できるように構造化し、そのように電子化した教科書の中に、文脈構造における重要性に基づいて情報量に差をつけること、また、重要な価値が与えられた箇所の抜き出しやクローズアップの方法やその方法の妥当性について検討しなければならない。それは、学習者のレベルや問題点にあわせられるように教科書を電子的に再構造化して提示するシステムをより実用に足るシステムとして実現するためである。

本稿では、書物全体の構造化のために、XML を使い、そのタグのデザインを考え、実際の再構造化に耐えるものであるかを試す。実験的に行う構造化であるため、文章

が短く、画像が多用されている理科 6 年上 1 章「ものの燃え方と空気」を概観し、教科書の画像利用のパターンを探る。特に、文章ユニットと同様に扱われている画像の位置の重要性を調べたい。

以下、1 章では、教科書における画像の機能について述べ、2 章で、画像の機能に基づき、学習の流れを構成するためのキーとなる表現と関連する画像の XML 化の方法について述べる。3 章で、教科書内の画像の持つ特性と文章との関連性を表わすための方法を論じ、4 章で、実験的に行った構造化について述べる。

1. 理科の画像の機能

理科の教科書で提示されている抽象的な概念や現象は、文章と、絵、写真、グラフといった画像でも表現されている。そして、それらは、小さなユニットを形成して、同列に、章節の各階層に、書物としての流れに従って順序よく配置されている。ただし、画像の多くは、節内の視覚的な効果に配慮されることで、大きさや位置において、流れにそっていることが分かりにくい場合もある。

文章のユニットを文章ブロックと呼ぶことにしているが、この文章ブロックと画像の関連性という観点から、画像の機能を分類すると、画像は 3 つのタイプに分けられる。

- ① 文章の内容を補足するもの
- ② 文章に書いてある内容とは意味的に直接は関係ないもの
- ③ 文章の内容を補足するものではなく、文章に書くべき学習内容を説明するためや、文中で投げかけられた疑問

に相当する内容を表わすもので、文章同様、積極的に学習すべき内容に関する情報を与えているもの

画像の多くは、①に相当するもので、各々の機能別ブロック内の文や単語と関連することが多く、関連性のある文章ブロックの上下左右にその文章ブロックの内容を補足する目的で配置されている。

また、②の画像は絵が多く、それらは、通常、文章内に関連する箇所を持っていない。教科書の構成上、課題提示や囲み記事であることを示すマークのようなものである。また、学習活動における注意や別の学習活動の方法が書かれた文章とともに利用されていることもある。

③に相当する絵は、補足的なものではなく、学習すべき内容を表わす情報を与えたり、学習した知識をより詳しくまとめるために利用されることを想定して提示されたものである。これらは、短い文章で説明が十分できない場合や、活動を誘導する疑問表現への答えとなる状況を表わしていたり、学習内容の主目的である現象を紹介する場合に出現するもので、「百分は一見に不如」の実現であり、情報として価値が高いと考えられる。さらに、③の画像には、本文で説明されていることを別の表現でもう一度表わしたキャプションや、さらに、本文では十分述べられていない内容を補足するためのキャプションがつけられていることが多く、この画像の周辺には学習内容における重要な情報が集められている。

2. 理科の画像の XML 化

基本的には、画像も文章ブロック同様、上下左右の順に優位性がある書物としての流

れに応じて配置されているが、1章で見たように、画像は、文章ブロックと同列に章節介そうに従属し、表現内容に応じて重要性が異なっている。

そこで、①②③の機能別に出現させる位置を考え、以下のようなタグをつける。

①は、補足的であるため補足される単語や文に従属すると考え、関連先のブロック内に従属させる。図1にタグ付け例を示す。図は、スペースの都合で、詳細部分は省略している。

②は、文章内容における文脈の流れとは直接関係が薄い、注意事項等であることから、視覚的にはなんらかのブロック内に配置しているように見えても、関連させられている文章ブロックの近くに、その文章ブロックと同階層に置かれるようにする。図2に例示する。

③は、主に、実験の手順が記された文章ブロックの後ろに出現する実験結果を説明する文章ブロックの周辺に現れることから、直接関連する単語

や文がなくても、実験ブロックの次の文章ブロック内に従属させる。また、この③タイプの画像の多くは、

```
<block bkind="experiment">
<title>
  <chasen parspch="名詞-一般">・・・
  <part pkind="inst_question">
<ol>
<li>
  <paragraph>
    <sent ence skind="explain_body">
      <chasen parspch="名詞-一般">・・・
    </sent ence>
    <sent ence skind="explain_body">
      <chasen parspch="名詞-一般">・・・
    </sent ence>
  </paragraph>
  <image href="1103.jpeg"/>
<ul>
<li>
  <paragraph>
    <sent ence skind="inst_question">
      <chasen parspch="名詞-一般">・・・
    </sent ence>
  </paragraph>
</li>
</ul>
```

図2:実験の場合

```
<section>
<title>
  <chasen parspch="名詞-一般">・・・
</title>
<image href="1101.jpeg"/>
<image href="1102.jpeg"/>
<block bkind="story">
  <paragraph>
    <sent ence skind="explain_body">・・・
    <sent ence skind="explain_body">・・・
  </paragraph>
  <paragraph>
    <sent ence skind="inst_question">・・・
  </paragraph>
</block>
</subsection>
```

図1:章頭の場合

```
<block bkind="story">
  <paragraph>
    <sent ence skind="explain_body">
      <chasen parspch="名詞-一般">・・・
    </sent ence>
  </paragraph>
  <image href="1104.jpeg"/>
  <block bkind="image">
    <image href="1105.jpeg"/>
    <block bkind="headline">
      <paragraph>
        <sent ence skind="explain_body">
          <chasen parspch="名詞-一般">・・・
        </sent ence>
      </paragraph>
    </block>
  </block>
</block>
```

図3:イメージが下位にブロックを持つ場合

```

<?xml version="1.0" encoding="UTF-8" ?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0">
<!-- 【アドバイス】とよい -->
<!-- 注目部分「とよい」を持つ文 -->
<xsl:template match=
  "sentence[
    /chasen[text()='と' and
    @parspch='助詞-接続助詞' and
    following-sibling::chasen[1]/text()='よい' and
    following-sibling::chasen[1]/@parspch='形容詞-非自立' and
    following-sibling::chasen[2]/text()='。' and
    following-sibling::chasen[2]/@parspch='記号-句点']
  ]"
>
<xsl:apply-templates mode="em" />
</xsl:template>
</xsl:stylesheet>

```

図 4:XSL のサンプル

結果や学習内容における重要な意味を解説するキャプション的な文章ブロックを従属していることが多いため、キャプション情報を従属させている画像は、ブロックを形成するとして、画像ブロックを設定する。また、キャプションに表わされている内容は、実験結果や学習活動のまとめと関連することが多い。そのため、この画像ブロックも、実験手順ブロックの後ろの文章ブロックに従属するものとする。図 3 に、③のタイプの画像につけたタグの例を示す。

その他、③に関連する場所であるが、どの文章ブロックとも関連していない画像がある。これについても、読者が推測読みで理解しなければ内容を補佐するものとしてあげられていると考え、③のタイプの例として図 3 で示すように、「まとめ」機能を持つ文章ブロックに従属させることにする。

3. 文脈構造の強調表示と画像

文章の内容が持つ意味的な文脈構造を表示するということは、以下のような処理を行うことにより、XSLT での検索結果を HTML 形式でブラウザに強調して表示することである。

- (1) 文章の内容における文脈構造の指標にマークアップして、属性を付与する。
- (2) マークした文脈構造の指標を XSL で条件指定して検索する。条件は、名詞、動詞といった単語単位のもの、句単位の学習指示表現である。
- (3) 検索の結果、得られる文脈情報の指標を強調して HTML 形式でブラウザに表示させる。

例えば、図 4 は、注意を促す学習指示を読者に与える表現を検索条件として指定したスタイルシートである。これを検索条件として指定した場合、XSLT は学習指示においてアドバイスを行う表現を検索する。そして、その結果を、別のスタイルシートで指定された表示形式に基づいて HTML 形式

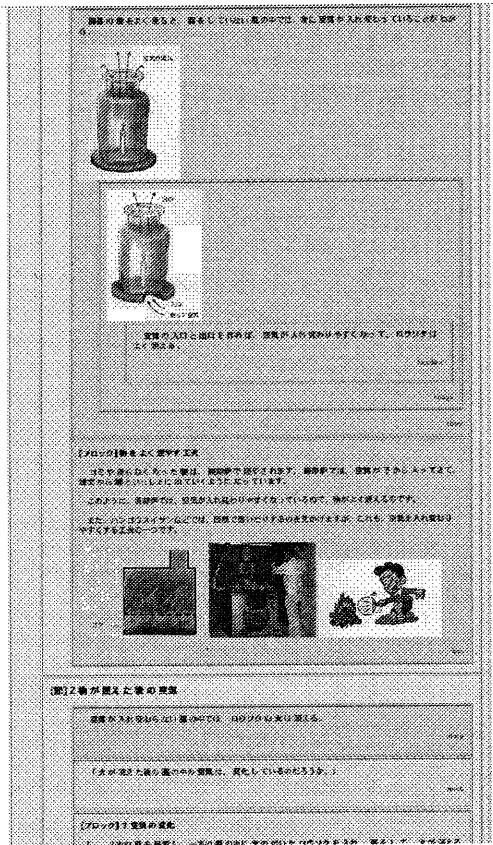


図6: Viewerで表示した画像付電子化テキスト

に変換し、ブラウザーに表示させる。

以上のような手順で強調表示する教科書の文章構造に、各々の文章に関連性の高い画像を加えて、ブラウザーに表わしたものを図5に例示する。

検索の結果、画像の出現位置を見ると、表1に見られるように、上よりは下、右よりは左への画像の配置が多いことがわかる。ただし、節内に属さない章頭のイメージは、より上で、より左に置かれている。また、実験結果を示す画像は、結果を解説する文章と実験手順を示す文章ブロックの間に置

	上	下	左	右
該当画像数	10枚	17枚	9枚	3枚

表1: 文章ブロックからみた画像の位置

かれており、直接関係する文章ブロックよりも上に置かれることになる。このように文章ブロックの上に置かれる画像に、考察のための情報となる物が多く見られる。さらに、学習活動をまとめる記述がある箇所では、文章の下に画像が置かれるが、結論に関する重要なものはより右に置かれるため、数は少ないが、右に位置する画像には重要な物が多いことが観察された。

図6に、学習の流れである教科書の文脈構造とそれに関する画像の出現位置を一般化したモデルを示す。色の濃い丸や四角は

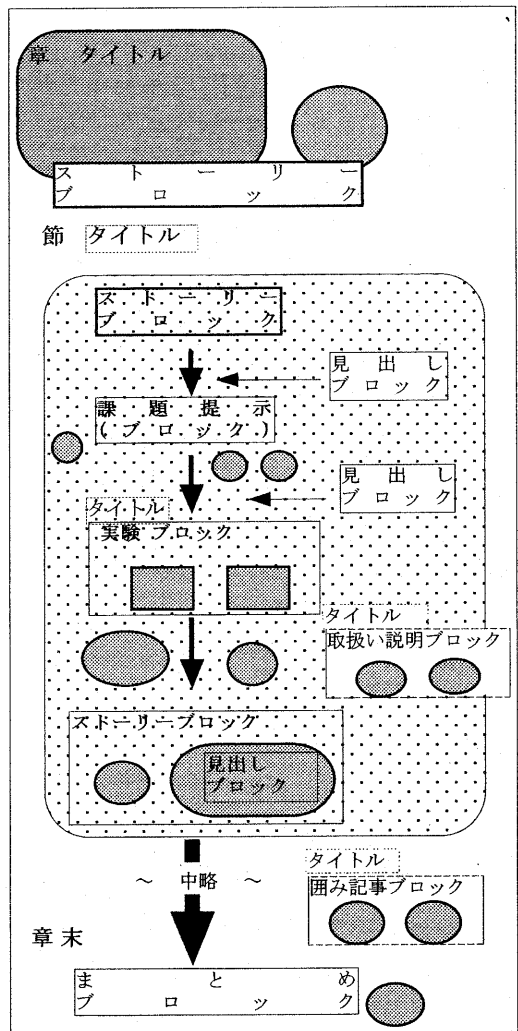


図6: 章頭から章末までの節内の主流ブロックの流れ

画像を表わしている。

4. 考察

小学校の教科書には、通常の本物とは異なった性質がある。新たな概念を学ぶ学習活動で利用されるものであるということから、語彙は、専門用語や抽象概念を表わす物が多く、読解には、フォーマルスキーマ認知能力が不可欠である。そのような教科書は、これまでの調査の結果、章節構造や学習の流れといった意味的な文章構造の他に、単語単位での既知・未知における知識の提示法への配慮があることが分かっていた[6][7]。本稿では、さらに、それらテキスト部分の情報に加えて、画像という視覚的な情報がいかに関わるかについて教科書の性質を調べた。そして、「推測読み」が必要な箇所、課題を提示する表現が記載されたテキストと導き出される結論を表わす画像を結び付けて内容を理解する能力が要求される部分があることがわかった。

平成8年度版の啓林館新訂理科6年上1章、2章では、特に、章頭の導入のための文章や、内容において重要な「まとめ」のための解説が行われている文章の所で、その推測読みが要求されることが多い。そして、学習概念やその学習の状況をイメージし、結論を推測して導き出しやすくするために、「まとめ」のための文章による記述が行われ、そこに画像やキャプションが多用されている。また、学習の流れという文脈構造や、タイトルや色分け等による章節構造の明示にも配慮されている。

本稿では、以上の教科書の性質に基づいて、画像も文章と同様の価値をもつとして、同列に章節構造ツリー上に記載してみた。

今回は、章節構造ツリーにシークエンスに出現する文章ブロックと画像との関係を示すXMLタグを付与し、画像も含めたコンテンツをツリー化した。そして、これにより、文章ブロックだけでは散在して配列されていたために相互の関係が不明であった文章ブロックの位置が、明確になった。

しかし、画像は解析により、その重要性を判断することはできないため、手作業で重要性をマークアップすることが必要なものである。今回は、重要性の判断は行わずに、解析によって得られたテキスト部分の重要性を利用し、重要テキストの周辺に出現する画像は重要であるとして位置付けている。その是非は検討しなければならない。

また、今回の調査により、画像が読解された内容のイメージ化や行間をうめるべく利用されていることから画像がフォーマルスキーマ認知に大きく影響するコンテンツであることが再確認できた。しかし、教科書のパーソナライズにおいてどのような情報の取捨を行うべきかが今後の大きな課題となってくる。今後は、これにより、学習すべき内容が記述されている文章ブロックの周辺に現れる画像に対して、さらに工夫を加えてパーソナライゼーションを考えていきたい。

また、今回は、平成8年度版理科の教科書を見た。現行の平成12年度版の教科書では、さらに、画像の多用や視覚的な効果への配慮が進んでいると考えられる。したがって、今後、現行の5教科で、画像の持つ情報の重要な役割がどうなっているかを比較する必要がある。5教科で異なったテキスト構造があるのか、また、日本語の文章構造と教科書との関係を調べ、教科書特有

のテキスト構造や日本語の文章の中での位置付けを明らかにしていきたい。

さらに、読解の補助を目指す場合、具体的に教育効率が上がるようなインターフェースを考え、より実際の教材利用で応用できるようにしなければならないだろう。

以上、本稿では、電子化されたテキストの内容における意味的な構造化を行い、電子化テキストにおける再現性の際、情報価値の高い画像を組込む方法について考察した。画像は、画像が関連するテキスト部分との関連性を考慮して、章節階層に組み込んだ。そして、評価の結果をブラウザーに表示し、テキストと画像の両方が教科書の構造を反映した状態で表示されることを確認した。

参考文献

- [1]Tannen,D.1977.What's in a frame? Surface evidence for underlying expectations. In R. O. Freedle (Ed.), *New direction in discourse processing*. Norwood, Nj:Ablex.
- [2]Kintch, W., & van Dijk, T.A. 1978. Toward a model of text comprehension & production. *Psychological Review* 85, 363-394
- [3]Hudson, T. 1982. The effects of induced schemata on the "short circuit" in L2 reading: Nondecoding factors in L2 reading performance. *Language Learning*, 32,1-31.
- [4]総務庁行政監察局, 1997 "教育の国際化を目指して", 大蔵省印刷局.
- [5]早川勝広, 1992.2 "言語形成と物語構造", 日本語学, 明治書院.
- [6]中尾桂子, 森下淳也, 2000, "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", 神戸大学国際文化学会『国際文化学』第4号, pp, 133~151.
- [7]中尾桂子, 森下淳也, 2001, "文章ブロックの構造化における形態素タグと XM 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細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. 2001, No. 18, pp, 173~180.

[8]中尾桂子, 森下淳也, 2002, "外国人児童・生徒のための教科書読解支援 View システム", "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.2002, No.19.