

読解支援のための小学校教科書の再構造化と動的拡張

—語彙的特性に基づいたテキストの簡略化表示とルビの補填—

中尾 桂子

神戸大学 大学教育推進機構

森下 淳也

神戸大学 国際文化学部

本研究は、学習者の能力に応じてデータを再編成させる教科書読解支援システムを実現し、教科書のマルチメディア化の一つのあり方を提案するものである。これまで、大きく分けて3つの観点、すなわち、(1)テキスト部分の強調表示による簡略化、(2)情報の外部からの補填、(3)内部情報の再編成による参考資料補充、を検討し、これらの融合に、xmlの特性を利用して取り組んできた[1][2]。これは、構造化したテキストデータを、内在する指標に基づき、別の構成に組み替えるというものである。今回、本稿では、2点、プロトタイプに改良を加えた結果について考察し、日本語弱者に教育的であり、且つ、優しい教科書を提示し、一斉指導用教材としての教科書のマルチメディア化実現に向けた一つの方向性を提案する。1点目の改良は、先に取り組んできた補填をさらに拡張し、外部データとして置いた情報を読み込むことで「ルビ打ち教科書」を実現するものである。2点目は、語の重要度を、強調表示や簡略化に反映し、テキストの簡略化を改良することである。強調表示は、要約結果のみの出力ではなく、読解時の着眼点を教示するために簡略化の過程を示した形式で行っているが、その表示を変更し、重要部分をさらに強調するものである。

Re-structurizing and dynamic enhancing of elementary school textbook for comprehension support

Keiko Nakao

Institute for Promotion of higher Education
Kobe University

Jun-ya Morishita

Faculty of Cross-cultural Studies
Kobe University

This research achieves the textbook comprehension support system that makes data reorganized according to learner's ability, and proposes one ideal way of making of the textbook a multimedia. In this text, it is considered that the two points add to the result of reconstructed text data with XML for improvement to the prototype system. And, we proposes one directionality for the achievement of making to the multimedia of the textbook.

1. はじめに

本研究は、小学校の教科書を読解するための支援として教科書のマルチメディア化を検討している。

年少者の日本語ノンネイティブ児童をその具体的な対象とし、小学校の教科書という学習用日本語の読解支援をシステム化に取り組んでいるが、現段階のシステムはプロトタイプであり、この段階での利用者として想定する対象は、ノンネイティブ児童を指導する教員の方であり、事実上は、教材開発支援ということになる。

日本語弱者としてのノンネイティブ児童の抱える主な問題を以下にあげる[3][4][5]。

- ① 語彙量の少なさによる認知・理解への弊害
- ② 文法知識・機能語の量的な問題により、日本語の文章理解に時間がかかる
- ③ 生活経験の内容と、既存知識が一致させられないことから物事の認識に時間がかかる

かるが、特に書き言葉における負担が大きい

- ④ 上記①②③の量的な問題により、日本語の文章構造の知識を体系的に把握しきれない
- ⑤ 日本の義務教育における一斉指導にあわせた学習方式に不慣れである（授業での教師の指示の意味を理解し、それに反応する方法、予習、復習の方法、日本語の文章から情報を収集する方法など、学習上のテクニック全般）

このような問題を抱える日本語弱者への教育では、多くの教員が独自に、教材開発や教科書学習の支援に取り組んでいる。

教材開発では、通常、人的、時間的制約により、初期から教科書を教材として利用するため、簡略化や翻訳に目を向けられることが多く、その場で意味がわかればよいという学習内容の理解に重点が置かれている。

しかしながら、ことばは、個人の経験量、学習量、言語的環境に影響される性質がある。また、学習者の能力や嗜好は、常に変化していく。同一教材が利用できる時期は短く、教材は、個人のニーズや学習時期によっても、常に流動的で、変化していく。

したがって、自立を念頭に置いた、ノンネイティブ児童に対する教材においては、以下の点をおさえた教材の提供が望ましい [6] [7]。

- A) 教材は、段階別に、様々な形で準備し、学習者の能力や嗜好、学習方法の変化に対応させて、題材をアレンジすることができる
- B) 日本の社会生活を基盤としたものについては、解説や関連語彙を補填し、内容理解を補足する情報を、コンテキストに合わせて提供する
- C) 生の教材と変更後の教材の状態の違いを明示し、文章構造のどの部分に着目した簡略化が行われているかを明示する
- D) 学習用語彙の使用法や関連語彙を確認することができるように、教材内での語の使い方一覧を例示することができる
- E) 談話や論述の流れに合わせた学習のテクニック自体を指導することにつながっている

以上、支援対象の問題と自立のための教育的学習支援における課題を踏まえ、多様な学習者に応じられる読解支援システムの構築を目指す。

動的な教材という考え方と、教科書テキストを用いた読解システム、マルチメディア教科書の一つの形を提案するものである。

2. 教科書テキストの特徴

2.1. 教科書の情報提示法

2.1.1. 教科書のレイアウト

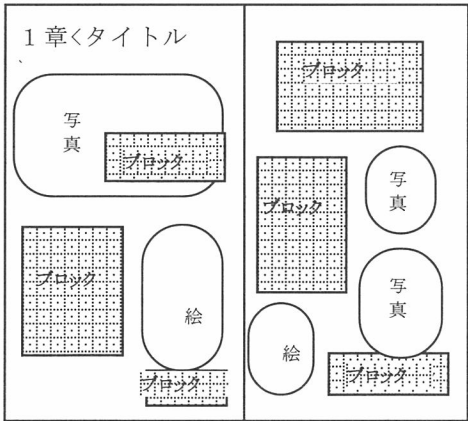


図1:教科書の見開き図

小学校の教科書には、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。

図1に示すようなブロックにまとめられ、テキストが見開き1ページに配置されているが、ブロックは、小学校教科書の章節構造に基づいて配置されている。

2.1.2. 教科書の標識

教科書は、図1のようなレイアウトであり、また、学習の流れや教育的示唆といった複数の文脈が輻輳した文章構造を持つ。

この学習の流れや教育的示唆、副流といった複数の文章の流れは、それぞれの文脈ごとに、色、枠、マーク、文末表現、文体、画像に統一したものを使用し、視覚的に区別されている。

2.2. 教科書の学習の流れと文章構造

教科書の文章は小さなまとまりごとに二次元に内地されているが、その文体やページ内の位置により、機能や内容に特徴がある。

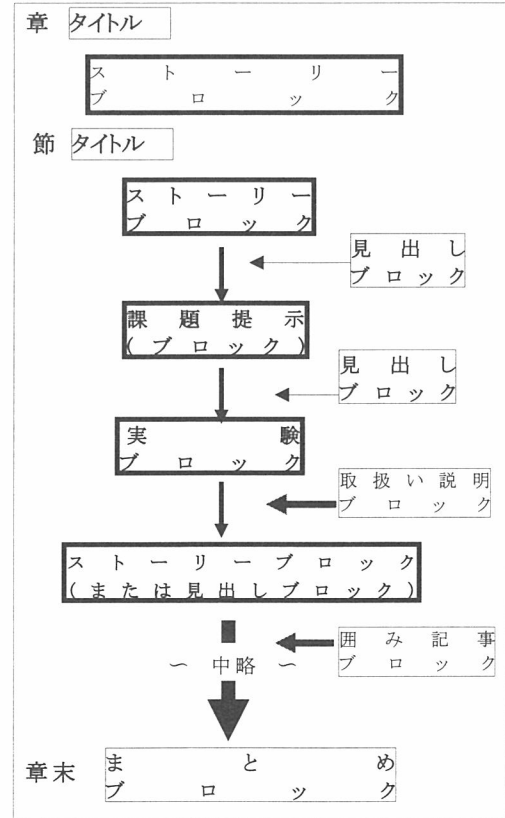


図2:章頭から章末までの流れと節内の主流ブロックの出現モデル

啓林館新訂理科の場合、ブロックは、全部で7種類あり、各々、(1)ストーリーブロック、(2)課題提示ブロック、(3)実験ブロック、(4)見出しブロック、(5)囲み記事ブロック、(6)取り扱い説明ブロック、(7)まとめブロック、となっている。

図2に章節構造下のブロックの出現の様子を表すが、図2は、啓林館新訂理科[8]においては、全ての学年の教科書で見られる章節構造下の基本的なパターンをモデル化したものである。

章のはじまりには「ストーリーブロック」が現れ、その章の導入となる状況が説明されている。節内でも、通常、まず、節毎の導入部分となる「ストーリーブロック」があり、次に、考察ポイントでもある作業課題が提示され、さらに、実験手順が説明される。

2.3. 教科書の不足情報

中尾・森下(2003)は、物理的には見えない文脈、すなわち、学習の流れである文章内容に内在する論理的な構造と、それらが構成する書物としての意味的、機能的な文脈をより明示的に視覚化した[9]。

その結果、教科書の記述、補足資料の画像や説明には、文脈の中で不完全な部分があることを明らかにした。

教科書では、学習の流れの中で、問題提起に対応する答えが省略されていたり、画像で示されていることがある。

また、自明である内容や、別の単元で既に学習した内容に関する情報等は、文字で記載せず、ヒントや誘導的な発言で内容を示唆するにとどめることもある。

このような情報は、教科書を読解する上で問題となる。外にこれらに関する知識を求める必要がある。

2.4. 外部情報による内容補填の必要性

通常、外国語学習における e-learning 型読解支援では、テキストの外に別の参考資料を求める場合、辞書や百科事典の添付機能を追加し、語訳を表示させるのが主流である。

しかし、年少の日本語弱者の場合は、母語自体の意味の把握が不完全であることから、翻訳や辞書が母語で与えられていたとしても、意味や概念が十分に理解できるとは限らない。

また、単語の意味が理解できても、文単位で表される内容や文脈がつかめるかどうかは別の問題となる。

なぜなら、文章の内容は、内容を構成する個々の語が、それぞれどのように結びついているかといった関連性を把握していなければ、全体像を思い浮かべるのは困難だからである。

把握、理解というのは、ある種の、概念の分類作業である。概念の分類、位置づけとは、すなわち、概念の説明時に使用される語彙とその語彙が織り成す文脈の違いを区別、判断することなのである。

したがって、概念理解には、個々の語彙の組合せ、語彙の共起が生み出す概念世界をイメージできなければならない。

これがいわゆるイメージスキーマ、概念の位置づけというものである。

文章を読んで理解するには、個々の単語によって構成されるある概念のイメージスキーマを得なければならないが、言語習得の過程でこのイメージスキーマを形成するためには、経験量と反復練習が必要になる[10]。

ところが、e-learning には、実物に触れる体験が重要となるような学習には不向きという弱点がある。

そこで、画像や音声といった情報を利用し、ある語彙のイメージや概念の体験がなくとも、また、語彙を多く知らなくとも、理解できるように、視覚情報と聴覚情報をテキストに補填し、年少者の語彙力、経験の乏しさをカバーすべきである。

3. 教科書の構造化と指標

以上、2節で見たように、教科書のテキストには、教科書特有のデータ構造がある。

日本語弱者の支援として、教科書の構造上の特徴を強調して表示するために、元のテキストの構造自体を電子化し、さらに、内容に埋め込まれた文脈情報を視覚化するために、再構造化する。

元のテキストの目に見える物理的な構造は、書物内の章節を階層化することによって実現する。すなわち、教科書の場合は、①平成8年度教科書、②教科、③学年、④上下巻、⑤単元、⑥章節、6つの階層である。

そして、これら6つのノード下のそれぞれに、テキストを内包する「ブロック」や、図表がある。

なお、テキストは、さらに、形態素ごとに分割され、下位情報を属性として保有している。

図3に、構造とXMLタグの関係を示す[*]。図3は単元以下のツリーをモデル化したものである。

以上の物理的構造とは異なる文脈の場合、その構造化は、この書物としての構造化に加え、ブロックの位置、文章に埋め込まれた表現により定義している。

ブロック単位の文脈を構造化するための指標としては、文末のモダリティ表現を利用している。

文末形態の違いとは、丁寧体、普通体といった文体や、助動詞相当語句で主観的に筆者の立

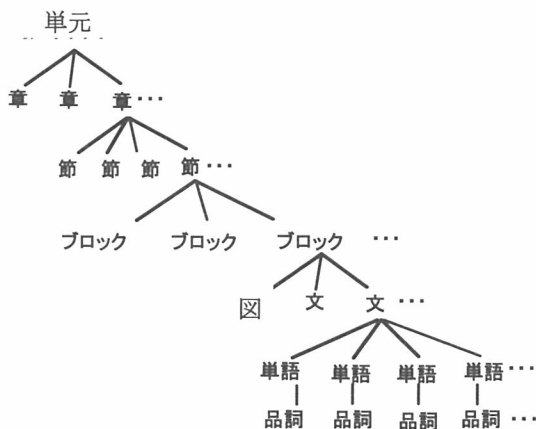


図3: 教科書の構造化モデル

場や読者への働きかけを表明するという文の調子を変えるものである。

これらは、ブロックの階層上の位置により、文末の形態に違いがみられることから、教科や、教科書で展開される学習活動と関係が深い。文末形態の違いと、ブロックの位置は、教科書の文章、文脈といったディスコースの形成、変換に大きな役割を担うものである[12]。

4. 教科書の動的拡張

4.1. 補填箇所

教科書に不足しているテキストや学習者側にはないイメージの補填として、教科書テキストへ画像情報、文の使用例、辞書による漢字の読み方、言葉の意味、訳語を追加している。

これらは、一般に公開されているインターネット上の情報や電子辞書の情報を教科書テキストに取り込むものである。

学習者のニーズに応じて利用できるよう、各語をキーとして検索できるよう、単語に検索結果表示用のリンクを埋め込んでいる。

4.2. ルビの追加

4.2.1. 未習と既習区別

小学校の教科書は、該当学年の学習漢字の既習、未習により、特殊な漢字かな混じり表記が行われている。

しかし、外国人児童等、Non-native の日本語話者は、必ずしも段階を追って学年別に漢字を学習しているとは限らず、漢字の読み書き能力に差がある。

通常、漢字かな混じり文の読解のために、教科書にはルビを打つ。

ルビは、担当教員が打つ場合もあるが、主に、同学年の児童生徒など、日本人ボランティアに頼るところも多い。そのため、ルビは手書きで

つけられることが多くなるが、手書きのルビは、けて読みやすいものだけではなく、消しゴムで覚えたものを消しても跡が残る。

また、教育的指導を考えた場合、学習者の能力に合わせて、テキストの漢字にルビを打つ必要がある。

既に覚えた漢字にはルビを振らず、まだ読めない漢字につけるなどの配慮が必要になる。

標準は、全て表示としておき、学習者の覚えた漢字の読み方については、非表示にし、読みたくなれば、いつでも手軽に見られるように、表示、非表示の切り替えが簡単に行えるよう、利用者の漢字の能力に応じて、ルビ表示を切り替えられることが望ましい。

4.2.2. 漢字ルビの補填

ルビは様々な方法、形式でつけることができる。ルビをつけることを考えれば、ブラウザの表示設定上の問題として扱う類のものである。

例えば、Web2.0 を利用すれば、漢字の上にルビをつけて表示できる。他のブラウザでも各漢字の後ろに () で括り、読み仮名が表示できるなど、何らかの形で可能である。

しかし、本稿では、学習者の能力に応じてデータ表示を可変にすることを第一に考えるため、システム上で処理する。

システム上でルビうちを実現するには、リソースをどこに置くかという点で2つの場合が考えられる。

一つ目の方法は、テキストデータの内部属性として漢字の「読み」情報をデータの下位に組み込むことである。

二つ目は、リソースをテキストデータの外部に置き、そこから情報を読み込むことである。

一つ目の内部属性に組み込む方法は、形態素解析[13]の結果と学校の学習漢字表に基づいたリソースを、現在のデータの下位に属性として組み込み、その組み込んだデータ属性を漢字とともに表示させるという方法である。

二つ目の外部参照とする場合は、サーバに読み込ませたリソースを参照し、合致するデータには、リソース内からデータに結果を反映させるというものである。

今回は、リソースとなる「読み」情報のメンテナンスが行いやすいことを考慮し、「読み」情報をデータの外部におく。学習者側の指定行為を参照の際の条件にし、それに依りて外部のリソースを参照するという仕組みである。参照と利用の流れを図4に示す。

ここで言う学習者側の指定行為とは、すなわち、ブラウザで HTML 形式で表示しているテキストに対して、漢字表記をポイントすることで、外部情報を参照させるということである。

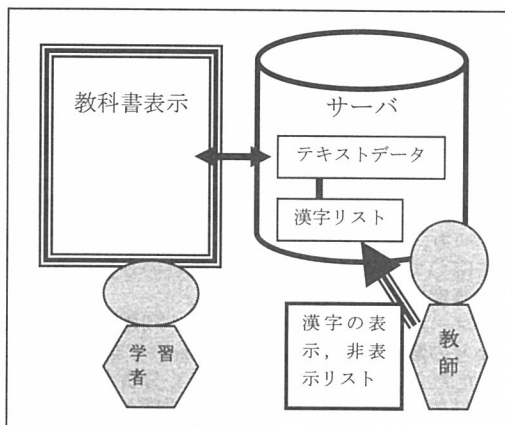


図4: 「読み」データの外部参照と利用の流れ

JavaScript の機能を利用し、サーバ参照後変更結果をダイナミックに表示する。

ルビは、これまでのプロトタイプシステムでは、個別に学習者が必要とした単語に対して表示させることにしていたが、今回、標準で全漢字表記に補填することに変更した。

その理由は、現在、利用しているテキストデータの性質と複数の指定を埋め込む必要が生じたためである。

現在、本システムで利用しているテキストデータは、構造化の際に形態素解析を行い、単語単位に分割した後、単語単位までの階層を構造化している。その際、各単語の下位に、属性として品詞情報を組み込んでいたが、形態素解析の精度を向上させるため、全てのテキストに対

して、可能な限りの漢字変換を行っている。また、未知語として処理されたものもそのままである。

また、漢字変換の際、教科書特有の漢字かな混じり表記、及び、教科書特有の分かち書きは排除している。

分かち書き排除の理由は、データを部分的に変更するのではなく、一律同条件で形態素解析を行うようにする方がよいという判断による。

先に述べたように、Non-native の学習者が段階を追った漢字学習を行っているとは限らない。

また、分かち書き表記は低学年のみであり、出版社や教科により、分かち書きの方法が若干異なっていることによる。

以上のような変更を加えた現状のテキストデータにより正確に直し、また、教育的示唆からルビ表記の表示を切り替えることを行えるようにする。

新たに、「読み」情報を対応させ、利用するといった準備と利用の流れを、以下の1、2の準備と、3の利用にまとめる。

1. 現在利用しているテキストデータの漢字の読み方を、形態素解析システムの解析結果を利用して準備し、データベースに格納する。
2. データベースのテーブルは、「漢字の読み方」、「ID」、「小学校1年から6年までの学習漢字表との関連情報」「表示、非表示指定」からなる。
3. 学習者は、ホームページに用意された学年指定、その他の指定を複数かけられ

ゴウドウ セイタカケイラ、チョウテンガ イナンニ アツマルヨウニ キタ シキツ マテミ
合同な 正多角形を、頂点が1点に集まるように隙間なく敷き詰めてみよ

ウ。タノセイタカケイダト ドウカナ。ドウシテ ウマク シキツメラレナイノヨカンガ
う。他の正多角形だとどうかな。どうしてうまく敷き詰められないのか考え

テミヨウ。シタンズノヨウニ、セイサンカケイラ テカイロイロナカタヲ ツクエ
てみよう。下の図のように、正三角形を6つ使って、色々な形をつくり

ナサイ。セイロッカケイ ノカクノワラ、タノニ ノケンガエカデ ヌマシタ
なさい。正六角形の6つの角の和を、次の2通りの考え方で求めました。

図5: 漢字のルビ表記と CaboCha を利用した色分け強調表示

るようにした切り替え指定ボタンを利用し、表示、非表示の切り替えを行う。

これにより、利用者の問合せに応じて、テキストに付与したルビを表示、非表示と切り替える。

指定された条件は、ベースとなる XML テキストに対して XSLT[14]で検索した結果、再変換されて提示されるテキストを、HTML 形式で表示している。

4.3. 強調表示による簡略化

本稿のプロトタイプシステムで言うテキストの簡略化とは、テキスト内部の内容語（content word）の強調表示を意味する。

すなわち、名詞、動詞といった概念語に対する読解時のスキニングに習い、その軌跡を、ブラウザー表示の際に、HTML で色をつけてクローズアップするというものである。

これは、概念語の配列全体を一律同格に扱っているということである。

そこで、構文、意味といった以下の 2 つの観点から見た文内の語彙の重要度を、ブラウザー表示に反映させる。

4.3.1. 文内の構造上の主要素の強調

文は、核となる述語を中心として、構造を持ち、意味は、主要な文の構成要素に基づき、判断される。

つまり、修飾、被修飾語といったレベル差のように、構造を踏まえた重要度の違いというもの、明確に表示に反映させることで、内容語の中の中心のものを探すという読解ポイントのスキニングがより簡単になることを利用するものである。

読解の際の着眼点を示唆する場合、一律同列に扱うのではなく、構造上の重要度が繁栄された表示にすることが好ましいため、述語を基準とした格関係を反映する表示が望ましい。

そこで、Cabocha[15]を利用し、必須格を他の内容語とは区別して、色を変えて表示する。図 5 はその例である。

4.3.2. 文章内の意味上の主要素の強調

一方、構造とは異なり、文章の内容やテーマに応じて、よく利用される語という観点からの重要度評価がある。

コーパス言語学の分析でよく利用される頻度などの数量的な情報に基づくものである。

ただし、通常の頻度に基づく分析だけでは、教科の学習における語彙の重要度が特定できない。そのため、いくつかの統計的手法を複数利用する必要がある。

そこで、基本的な頻度やレンジといったコーパス言語学における基本分析手法に加え、残差 IRF、エントロピー等を用い、教科別の特徴語を洗い出した。

表 1 に国語科の比較の場合をあげる。表 1 から、頻度の高い語は国語科でさも利用しそうな語であることがわかるが、エントロピー降順で見ると、教科として、全般に出現するものとは異なっていることがわかる。また、残差 IDF で特異なものを見ると、各単元のキーワードとして重要なものがある。残差 IDF を昇順に見ると、局所的に出現する頻度の高いものがわかり、各単元毎の概念説明のために必要な語を特定することができる。ことがわかる。

ただし、このような順位で特徴を見る場合、どこまでを特徴語とするかは不明な点も多いため、さらに、主成分分析による教科別の全体との比較などを複数重ねることにより、基本語と特徴語を区別する必要があるだろう。

表 1：頻度とエントロピーと残差 IDF (RIDF) の語彙比較（国語科）

番号	全学年で出現頻度の高い語	教育基本語彙 DB と不一致な語	エントロピー降順語	一般的なもの (RIDF 昇順)	特異なもの (RIDF 降順)
1	言葉		縦	名	じいさん
2	人		名	縦	アナ
3	次		他	楽しみ	兵
4	自分		お母さん	休み	藁沓
5	漢字	漢字	楽しみ	歯	うち
6	声		外	部屋	ガン
7	文	文	横	窓	おじさん
8	音		遊び	一つ一つ	白馬
9	目	目	歌	庭	トラ
10	お母さん		うち	色	ガマ

また、テーマ自体との関連を自動抽出させるべく、意味タグを利用する方法を検討する。

意味タグとは、ある語がどのようなカテゴリーに属する語かを示すタグである。

この意味タグを利用することにより、同義的、または、共起関係として、ある概念を表す際に結びつきが強いと考えられる語彙を特定することができる[16]。つまり、テーマとの関連を考慮して、文章の特徴を表す共起語を調べ、テーマに応じて重要度を計ることや、語の意味上の違いや、スキーマとして概念を支える語群という考え方から、意味的なまとまりで重要性を提示させるという方法である。

これは、語の使用頻度調査だけでは明らかにしない語、すなわち、頻度が低いが、共起度の高い関連語彙をより確実に認定することができるため、教科別、テーマ別の重要語彙を特定することに有益であることから、この方法でも重要度を評価する。

5. 教科書のマルチメディア化

本稿で提案しているシステムは、学習者の能力に応じ、教科書を誰でも読めるようにするための読解支援を目的としている。また、教育上の示唆を行い、自立的な読解能力を身につけることに配慮した、e-learning システムを念頭に

おいている。以上に基づき、教科書をマルチメディア化するという作業を行っているのである。

今回は、ルビと重要部分の表示法に改良を加えたものを紹介する。ルビは、教育的示唆という観点から表示の切り替えが行えるようにしており、また、より重要な情報をスキミングする観点を養うための表示切替に、単語の重要度評価を応用するという拡張を行った。改良の結果を表示したものが図6である。

e-learning という観点では、マルチメディア化ということが目的のとなりがちな点であるが、教育という観点で、個々に応じた支援を考えた場合は、厳密に言えば、教科書のマルチメディア化ではなく、教科書の多様化であり、そのため手段として、マルチメディア化があるということである。

個々の学習者によって、読解能力、つまりは、把握できる内容が異なっているという問題に対応することが、教育的示唆に配慮した読解支援システムで最も重要な観点である。

また、教科書には、編纂者の教育的な意図があり、また、それが学習の流れを形成している。さらに、教科書は、必ず、指導者と共に利用するものであることから、必ず、教える人の視点

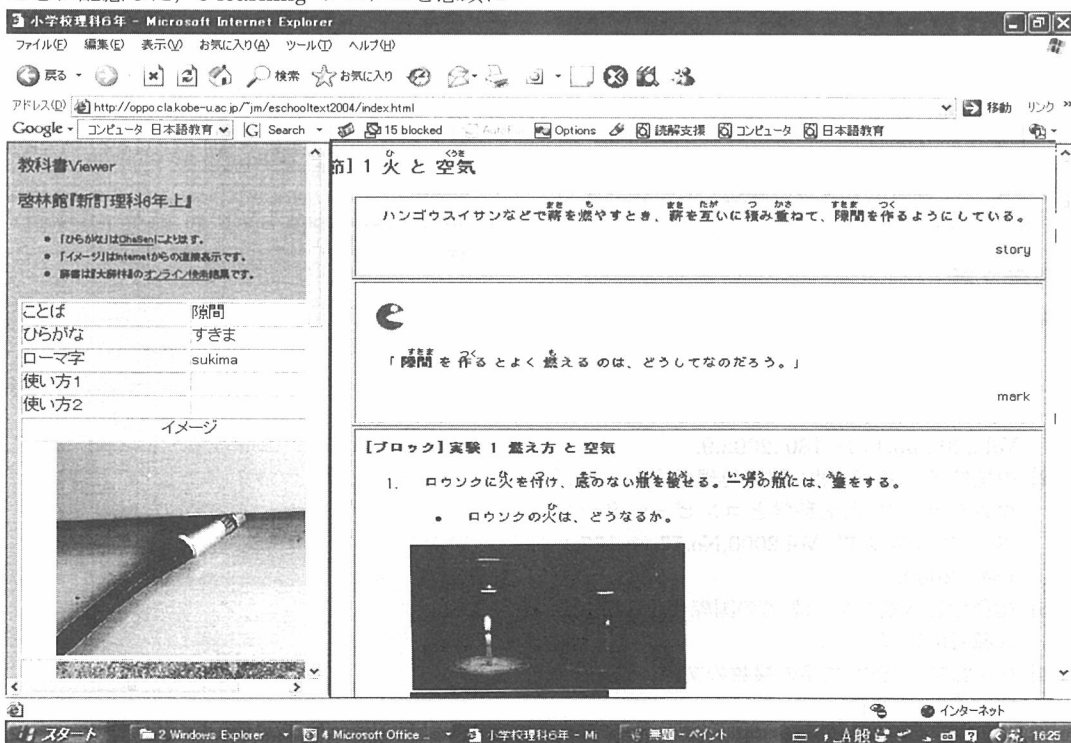


図6：教科書読解支援システムの表示例

を介するものでもある。教科書の編者の意図とは別に、教育的効果を考慮したコースデザインにより、教育内容は再編纂される。学習者は、教科書自体の学習の流れの意図、指導者の意図、学習者の能力に応じた要求に左右される。

教科書のマルチメディア化は、これらの編纂者の意図や指導者の配慮を変更するのではなく、その流れを、学習者に明示し、かつ、学習者のニーズに応じるための手段であるべきなのである。

多様化する教科書とは、教科書を中心に置いた3者の意図や意識が反映されて、いろいろな見方をされるものであるため、その見方に応じて可変なものである必要がある。この間点から逸脱しないシステムが、情報工学的にマルチメディア化され、かつ、教育的にマルチメディア化された多様な教科書となる。教科書が多様な学習者に応じて、必要な情報を与えられることが、支援というものだからである。

6. まとめ

本稿では、学習者の能力に応じて教科書を多様化した読解支援システムを提案している。現在のシステムでは、イメージや補填を中心としている。成人留学生に対する使用感調査でも指摘されているが、イメージだけでは、まだわかりにくい問題も多いため、仮想的な状況を設定し、現実をシミュレーションとしたもので訓練するような教材の提示や、利用履歴をフィードバックする方法も検討していかなければならない。より現実的なイメージを疑似体験することができるような多様化も今後の課題として検討すべきだろう。

参考文献

- [1] 中尾桂子・森下淳也: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, pp.173~180, 2002.9.
- [2] 中尾桂子・森下淳也: 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.
- [3] 総務庁行政監察局, :教育の国際化を目指して, 大蔵省印刷局, 1997.
- [4] 石井恵里子:教室談話の複数の文脈, 日本語学, 明治書院, 1997.
- [5] 早川勝広:言語形成と物語構造, 日本語学, 明治書院, 1992.2.
- [6] Hudson, T.: The effects of induced schemata on the “short circuit” in L2 reading: Nondecoding factors in L2 reading performance. *Language Learning*, 32,1-31, 1982.
- [7] Tannen,D.: What’s in a frame? Surface evidence for underlying expectations. In R. O. Freedle (Ed.), *New direction in discourse processing*. Norwood, Nj:Ablex., 1977.
- [8] 啓林館:新訂理科, 1996
- [9] 中尾桂子・森下淳也 “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”, 第 58 回人文科学とコンピュータ研究発表会., 2003. 5. 30,
- [10] Kintch, W., & van Dijk, T.A. :Toward a model of text comprehension & production. *Psychological Review* 85, 363-394, 1978.
- [11] XML: <http://www.w3c.org/XML/>
- [12] 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. 2001, No. 18, pp, 173~180, 2001.
- [13] ChaSen : <http://chasen.naist.jp/hiki/ChaSen/>
- [14] XSLT : <http://www.alphaworks.ibm.com/tech/LotusXSL/>
- [15] CaboCha: <http://chasen.org/~taku/software/cabocha/#performance>
- [16] <http://www.comp.lancs.ac.uk/ucrel/wmatrix/>