

日本語読解学習支援システムにおける辞書ツールの役割

The Role of the Dictionary tools in a Japanese Language Reading Tutorial System

東京国際大学教授・ウィーン大学客員教授 (2000.10-2002.3)

川村 よし子 kawamura@tiu.ac.jp

はじめに

日本語学習者のための読解学習支援システム「リーディング・チュウ太」はリンク集・学習支援ツール・読解教材バンクから成り立っている。このシステムを、1999年10月からインターネット上で公開してきたが、2001年3月現在、日本語版・英語版あわせてすでに5万件以上のアクセスがある。月曜と金曜には利用者がやや増え、土日の利用者はやや少ない等、多少の増減はあるものの、ほぼ毎日100件以上のアクセス数である。利用者は日本語学習者ばかりでなく、日本語教育に携わる教師、あるいは日本語研究者等も多く、辞書引きツール・レベル判定ツール・リンク集・教材バンク等をそれぞれ目的に応じた方法で利用している。ところが、日本語学習者が一番利用するのは、やはり辞書引きツールである。そのため、筆者らは読解学習支援システムの一般公開後も、日本語学習者や日本語教育関係者等からのフィードバックをもとに、日本語学習者にとってより使いやすい辞書ツールを目指して改良を続けてきた。今回、これら一連の改良をもとに、従来の辞書引きツールを大幅に改訂した新しい辞書ツールを完成し、公開することにした。本論では学習者の視点から見た従来の辞書引きツールの問題点とその原因、さらに新しい辞書ツールの改良点とその評価について報告する。

1. 「リーディングチュウ太」における辞書引きツールの位置づけ

読解学習支援システム「リーディング・チュウ太」は、リンク集・学習支援ツール・読解教材バンクによって構成され、日本語学習者のための読解学習支援環境をインターネット上に提供しているシステムである。その各々は次のような特徴を持っている。(このシステム全体の詳しい仕組みや使い方に関しては川村 2000 参照)

リンク集は日本語学習者にとって役立つホームページへとリンクが張られ、インターネット上の情報検索や教材探しが簡単に行える仕組みになっている。ネット上には無尽蔵とも言える情報が発信されているのだが、学習者が自らが必要とする情報にたどり着くのは容易ではない。日本語学習者にとって利用価値が高く、わかりやすく、情報量が多いことを基準にして、厳選したホームページにリンクが張ってある。学習者はここで見つけた情報を用いて読解学習を進める。その際に役に立つのが学習支援ツールである。

学習支援ツールには次の4つのツールが整備されている。

辞書引きツール

- ・ 日英辞書ツール テキスト中の単語の読みと意味情報を英語で表示する
- ・ 日日辞書ツール テキスト中の単語の読みと意味情報を日本語で表示する

レベル判定ツール

- ・ 語彙チェッカー 日本語能力試験に準拠してテキスト中の単語の難易度を判定する
- ・ 漢字チェッカー 日本語能力試験に準拠してテキスト中の漢字の難易度を判定する

学習者はこれらのツールを用いて、インターネット上の情報や電子メールなど電子化された文章（以下「テキスト」と呼ぶ。）を自由に教材化することが出来る。辞書引きツールで処理された文章では、テキストに含まれる単語の各々が辞書情報とリンクされている。学習者がわからない単語に出会ったら、その単語をクリックするだけで、右側に辞書情報（単語の読みと意味）が提示されるという仕組みである。さらに一連の学習が終了した時点で学習履歴（クリックした単語）を提示する機能も備えている。（北村・川村ほか1999）いっぽう、レベル判定ツールで処理された文章では、テキストに含まれる単語や漢字の各々のレベルが日本語能力試験の1級から4級までの基準をもとにして表示される。級ごとの語彙の含有率も示されるため、各々の文章にどのレベルの語彙がどの程度含まれているかが即座にわかる。また語彙のレベルをもとに文章そのものの難易度も判定される。（川村1998・1999）

教材バンクには初級・中級学習者用の教材が集められている。初級・中級学習者にとってはネットサーフィン自体も容易ではない。また、そのレベルの教材そのものもネット上では見つけることがむずかしい。そこで、やさしい教材、読みやすい教材、日本語学習に役に立つ教材を集めた教材バンクを用意した。収められた教材は、すべて辞書引きツールを利用して教材化したものであり、辞書機能、学習履歴表示機能を備えている。（川村・北村 2001）

読解学習支援システム「リーディング・チュウ太」は、これらを有機的に統合したもので、日本語学習者が自らのレベルとニーズに合わせて自由に学べる学習環境を提供している。このシステムはインターネット上で公開され（<http://language.tiu.ac.jp>）無償で提供されているので、日本語学習者はインターネットで日本語が読める環境にさえあれば、世界のどこからでもアクセスして利用することが可能である。

本論で扱う辞書引きツールは、上述のように、この読解学習支援システム「リーディング・チュウ太」の学習支援ツールの一つとして提供されているものである。学習者は、辞書引きツールを利用することによって、単語の読みや意味を即座に調べることが出来る。読解作業が辞書引きによって中断されることがないため、文章の意味を読み解く作業に集中することが出来る。このように、辞書ツールはこの読解支援システムの要ともいえる部分である。

2. 入力画面の改善

学習支援ツールのうち学習者が最も多く利用するのは辞書引きツールである。辞書引きツールは、読みたいテキストを入力すると辞書引き作業を自動的に行う道具である。ところが、その入力画面は学習支援ツールとして語彙チェッカー・漢字チェッカー等と一体化した構成になっていた。学習者は読みたいテキストをまず学習支援ツールのテキストボックスに切り張り（コピー＆ペースト）した上で、改めてテキストボックス下のボタンによって自らが必要な道具を選択するという仕組みである。（図1参照）

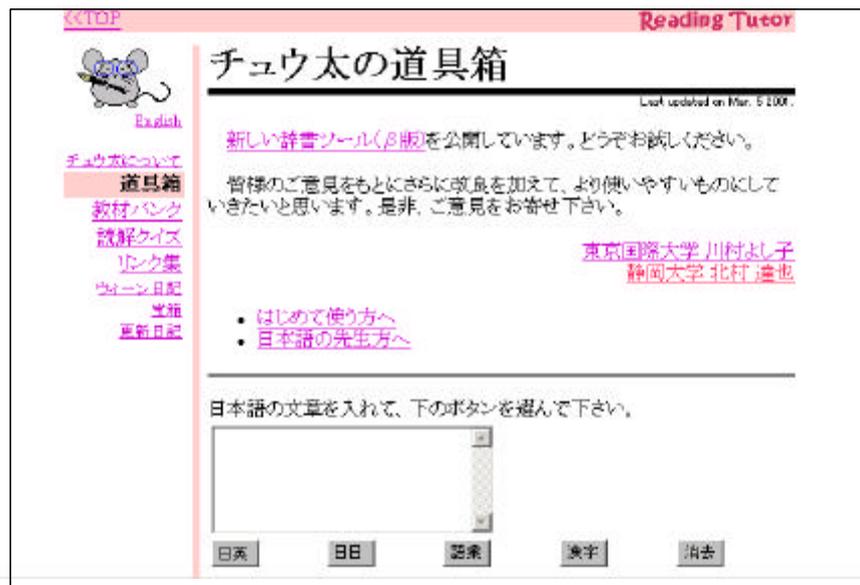


図1 学習支援ツールの入力画面

上図の一番下にある「日英」「日日」「語彙」「漢字」のボタンは、それぞれ日英辞書ツール、日日辞書ツール、語彙チェッカー、漢字チェッカーを意味している。そこで、例えばこの図の一番左の「日英」のボタンを押せば「日英辞書ツール」が選択され、テキストに含まれる単語の辞書引きが自動的に行われ、日英辞書の情報が表示される。

ところが、初めて使う学習者を観察していると、あらかじめ黒板や OHP 等を用いてクラス全体に利用方法を説明しておいても、実際に使う段になるとどのボタンを押したらいいか迷う学習者が必ず存在する。迷ってしまう最大の原因は、一番右の「消去」も含めると合計5つのボタンが並んでいることにあるようだ。特にレベル判定ツールのための「語彙」「漢字」というボタンの意味が直感的にわかりにくいらしい。「語彙」で言葉の意味、漢字「漢字」で言葉の読みが調べられると思ったという学習者もいた。インターネット等で直接アクセスして来る学習者も多いことを考えると、初めて使う学習者にとってもわかりやすい形であることが重要である。

そこで今回、利用目的の異なっている辞書引きツールとレベル判定ツールとを分けることにした。新しい入力画面は次の通りである。

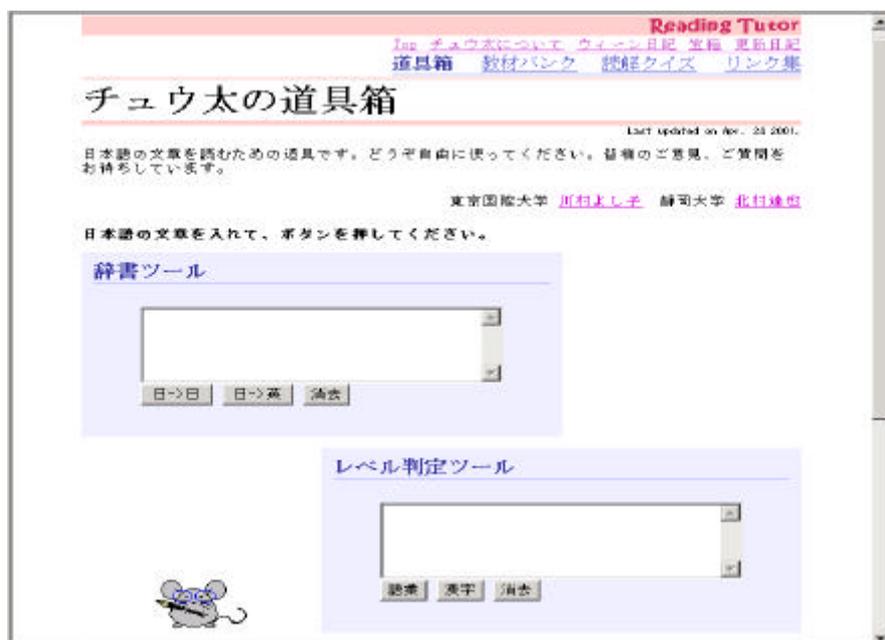


図2 新しい支援ツールの入力画面

辞書ツールとレベル判定ツールを分離することによって、今後他の言語に対応した辞書ツールを加えることも容易になる。また語彙チェッカー、漢字チェッカーがレベル判定ツールであることが明示されているため、初めて利用する人にもわかりやすい。

3. 辞書情報の改善

3.1 学習支援の視点からみた辞書表示方法の問題点

図3が従来の辞書ツールの結果表示画面である。画面左の本文の単語をクリックすると右の辞書情報が自動的にスクロールする仕組みになっている。



図3 従来の辞書ツールの結果表示画面

辞書引きツールでは、まず奈良先端科学技術大学院大学の開発した形態素解析システム「茶筌」によって単語の自動切り出しを行う。次に、茶筌の解析結果を元に辞書情報を検索しこれを表示する。辞書としては日本電子化辞書研究所（EDR）の日英対訳辞書の辞書情報を利用している。EDRの辞書そのものは多言語対応の翻訳システムを作り上げることを目的に開発されたものである。

こうした事情もあって、日本語学習者が用いる学習支援ツールという役割から見たとき、従来の辞書ツールはいくつかの問題を抱えていた。主な問題点としては次の3点があった。

- a. 単語の意味情報が数多く列挙されているために学習者にとって意味の特定が難しい。特に、基礎語にあたるもののなかには、意味情報の多いものが多い。そのため、例えば「行く」では28個、「目」では26個もの訳語が列挙されてしまう。
- b. 意味情報の配列順が概念番号順になっているため、中心的な意味やよく使われる意味が上位にないことも多く、学習者にとってわかりにくい。例えば、「言う」では、その10個の訳語が 1.put in words 2. inexpressible 3. produce a sound 4. make a sound 5. (pig) squeal 6.(animal) whine 7.enunciate 8.speak 9.say 10.express という順で並んでいて、最も代表的な speak, say といった訳語が下の方に来てしまっている。

- c. 単語の読みが複数ある場合辞書ツールの読みの表示が正しいとは限らないことがある。

読みの情報は茶筌の解析によっている。茶筌は一定の条件下で最も可能性の高い読みを選び出している。そのためその読み情報はほぼ正しいのだが、100%ではない。日本語学習者を支援する道具として開発するのであれば、間違った情報を与えることは極力避けなくてはならない。EDR 自体が25万語の見出し語を持つ非常に大きな辞書である。意味情報の記述も非常に詳しい。ところが、詳しくすぎる情報は学習者に負担を与える。辞書ツールでは、文の単語への分割、読みの表示、辞書引き作業、その表示等をコンピュータがすべて自動的に行っているため、避けられない問題もあるのだが、日本語学習者のためには可能な限りの改良が必要であった。

3.2 辞書表示方法の改善と評価

上記の問題のうち一番大きな問題は列挙されている意味情報をどうすれば少しでもわかりやすくできるかであった。この場合、「辞書ツールはあくまでも辞書であるべきである」という理念から意味情報そのものを絞り込むことはしなかった。

各々の対応方法は次の通りである。

- a. 単なる意味の列挙を避けるために EDR の概念説明を利用して、訳語情報を意味概念ごとにまとめ、概念説明を加えることにした。
- b. EDR の「単語辞書」にある意味概念ごとの頻度情報を活用し、頻度順に並べ替えを行った。この意味概念ごとの頻度情報は EDR のコーパス内での頻度情報をもとにしたものだが、調査の結果十分に活用できると判断した。(北村・川村 2000)
- c. 茶筌で読みが複数表示されている単語に関しては全ての読みを表示することにして、さらに各々の読みごとに意味を表示することにした。読みの提示順もコード順ではなく、頻度順に並

べている。

実際の改良による変化を比較してみることにしよう。図4が従来の辞書情報の表示によるもの、図5が今回新たに改良した表示方法によるものである。両者共に「山」に関する辞書情報の表示内容である。

山 (ヤマ)
1. mine
2. speculation
3. long shot
4. guess <examination>
5. speculation <stock>
6. heap
7. pile
8. high point
9. peak
10. offense
11. crime
12. top
13. mountain-shaped festival float
14. stack of
15. hill
16. mountain

図4 従来の辞書の表示

山【 やま 】****
1 [mountain] a very high hill, mountain mountain / hill
2 a very large amount of something mountain / pile / heap
3 a mine for the extraction of ore from the ground mine
4 the higher part of a thing crown / top / peak
5 the climax of an event apex / top / climax / peak / high point
6 speculative action venture / gamble / long shot / speculation
7 little success venture / speculation / guess
8 a case of crime crime / offense
9 a festival float mounted with a decorative halberd, called “yamaboko” mountain-shaped festival float

図5 新しい辞書の表

新しい表示方法では、実際に表示されている辞書情報は増えたのだが、項目数としては16項目が9項目に減少している。また概念に適切に当てはまる英訳がある場合には1の**[mountain]**のように色

を変えて示した。意味の表示順に関しては、従来は mine, speculation, long shot と並び、最も中心的な意味であるはずの mountain が最後に現れるため、「山」自体の意味が捉えにくいという問題があった。これに対して、改良版ではきちんと mountain が一番上に来ているし、2. a very large amount of something 3.a mine for the extraction of ore from the ground 4.the higher part of a thing と続く。日本語教育の視点からみると 3.の意味はもっと後でもよいという感も残るが、一応それなりに納得のいく配列になっている。

図6が新しい辞書ツールの結果表示画面である。この図の「道」のように読み方の可能性が複数あるものに関しては全ての読みが表示され、さらにそれぞれの読みごとに意味が表示されている。しかも学習者にはできるだけ可能性の高い読み方を優先的に教える必要があるため、読みの提示順は50音順ではなく、頻度順に並べられている。例えば図6の「道」の読みは「みち」「どう」の順で並んでいる。

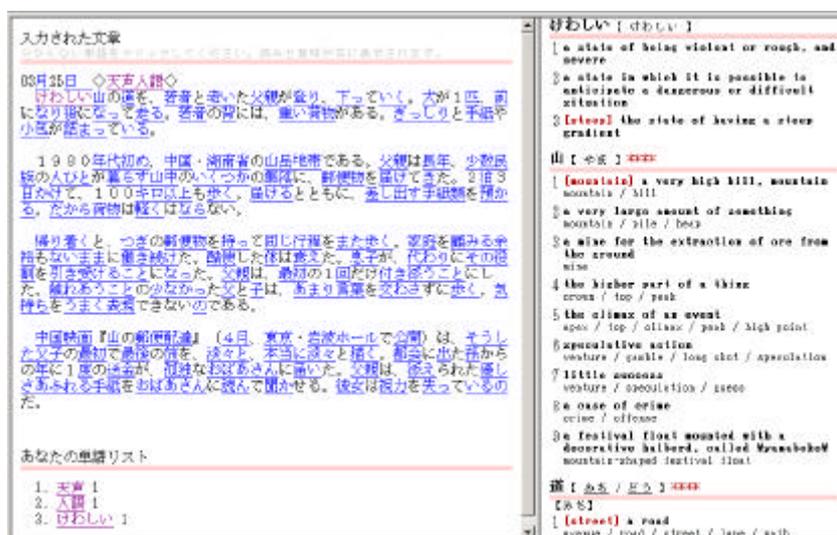


図6 新しい辞書ツールの出力画面

以上のような辞書の表示内容の改善に加えて、新しい辞書ツールではさらにいくつかの改良を行った。主な改良点は次の通りである。

1. 読みの表示方法の改良： 読みの表示はカタカナからひらがなに変えた。実際に利用している学習者へのインタビューの結果、ひらがなの方が読みやすいという意見が圧倒的に多かったためである。カタカナ語に関してもカタカナが苦手な学習者の便宜を考え、読みはすべてひらがなで表記した。
2. 基本単語の表示： 学習者に優先的に学ぶべき単語の目安となるように、各単語の難易度を、見出し語の右肩に星印によって示した。これは日本語能力試験に準拠したもので、1級から4級までの各レベルを星の数で表した。1級が星1つで4級が星4つである。星の数が多いほど基本単語ということになる。

4. 日日辞書ツールの改良

上記の辞書ツール（日英辞書ツール）と同様の改良を日日辞書ツールについても行った。



図6 新しい日日辞書ツール

日英辞書ツールと同様に、意味情報の表示順は概念の頻度順になり、読みはひらがなで表記し、読みの可能性が複数ある場合にはそれぞれの読みごとに意味を表示している。また、単語の難易度レベルも表示している。

この日日辞書ツール独自の改良としては、概念ごとに該当する英語を表示したことである。これまでの日日辞書ツールでは日本語による概念説明のみだったため、見出し語自体よりも説明の語句の方が難しいという問題があったが、英訳の提示により、多少はわかりやすくなったと思われる。

ただし、あくまでもこれは便宜的な方法に過ぎず、日日辞書ツールとしては各々の意味説明が出来るだけやさしい日本語（基本語）でなされていることが望ましい。現状ではあくまでも EDR の対訳辞書を用いてどこまで改良できるかを試みているに過ぎない。今後日本語教育関連の機関等で日本語学習者のための日日辞典が編集されることを強く願う次第である。

5. その他の改良点

今回の辞書情報提示方法の改良に伴って、辞書引きツールの学習履歴表示方法についても学習者がより使いやすい形への改良を行った。

学習履歴管理機能はこれまでの辞書引きツールにも備わっていた機能であり、学習者が辞書を参照した（単語をクリックした）学習履歴を保存し、一連の学習が終わった時点で学習者が本文下にある「List」ボタンを押すと、単語リストの形で学習履歴が表示されるというものである。この機能を活用することで語彙の学習効果が上がることはすでに立証されている。（北村・川村ほか 1999）

ところが自律学習を行っている学習者の中にはこの学習履歴表示機能そのものの存在に気づかない学習者も多い。そこで今回、図6のように画面左フレームの本文表示画面の下に学習履歴が自

動的に「あなたの単語リスト」という形で表示されるようにした。リストの右の数字は同一単語をクリックした回数である。このリストの単語もすべて右の辞書情報とリンクされているので、各単語をクリックすると辞書情報を再度参照できる。

6. 改良に対する利用者の評価

改良した新しい辞書ツールは2000年9月に試作版を公開以来、すでに4000件近い利用がある。2001年に入ってから1月627件、2月521件、3月690件と利用されている。リーディング・チュウ太のユーザーからの新しい辞書ツールに対する評価は高く、以前の辞書ツールのほうがいいという意見は寄せられていない。また、ウィーン大学の学生達に実際に両方の辞書ツールを利用した上で比較評価してもらったところ、新しい辞書ツールの方が見やすいし、わかりやすいという評価であった。また、現在わかっている問題点としては、1.英訳情報が異なっている場合に同じ概念説明が繰り返されて表示されてしまう。2.茶釜の辞書に登録されていないカタカナ語が自動的にサ変動詞として認定されてしまう場合があるという2点である。これらの問題に関してはさらに詳しい調査をした上で早急に対応したい。

謝辞：本研究には東京国際大学の海外研修助成を得ている。また、辞書の提示方法の改善に関しては国立国語研究所研究員の植木正裕氏から貴重な助言を得た。ここに記して感謝の意を表したい。

関連論文

川村よし子(1998)「読解のためのレベル判定システムの構築 - 語彙チェッカーの開発と活用」『日本語教育方法研究会誌』Vol.5, No.2, pp. 10-11.

川村よし子(1999)「漢字の難易度判定システム『漢字チェッカー』を用いたテキストの分析」『東京国際大学論叢』第59号, pp.73-87.

川村よし子(2000)「インターネット時代に対応した読解教育」『新世紀之日語教学研究国際会議論文集』台湾：東呉大学, pp.347-365.

川村よし子・北村達也(2001)「インターネットを活用した読解教材バンクの構築」『世界の日本語教育』第6号, pp.241-255.

北村達也・川村よし子ほか(1999)「学習履歴管理機能を持つ日本語読解支援システムの開発とその評価」『日本教育工学雑誌』23-3,pp.127-133.

北村達也・川村よし子(2000)「読解学習支援システムにおける辞書提示方法の改善」『日本語教育国際シンポジウム「21世紀型総合的日本語教育における語学・文学・文化及びメディアのあり方」』韓国.

日本語能力試験企画小委員会編(1993)『日本語能力試験出題基準』国際交流基金・日本国際教育協会.

日本電子化辞書研究所(1996)EDR電子化辞書仕様説明書,日本電子化辞書研究所.

松本裕治・北内啓・山下達雄・平野善隆・松田寛・浅原正幸(1999)「日本語形態素解析システム『茶釜』version 2.0使用説明書第二版」NAIST-IS-TR99012.