

「漢文エディタ」の試み

国語科 篠崎 秀樹

インターネットという便利な情報メディアのおかげで個人でもほとんど元手いらずで情報を発信することが可能になったのは、'90年代の半ば頃からだったろうか。'97年に原念齋・東條琴台『先哲叢談』を1年間かけて古本から入力したのを皮切りに「日本古典文学テキスト」を中心として著作権の切れた電子データを無料公開するサイトを立ち上げ、その他参考文献類・近代文学テキストのほか電子図書館・電子テキストのマークアップ方法に関する半ば以上趣味的なデータの収集や公開を個人的に行ってきた。たまたま公開10年目の2006年から漢文のマークアップについて試行錯誤を始め、ExcelやWord等、MS-Officeソフトに標準で付属するVBAという簡易マクロ言語を使って「漢文エディタ」という漢文入力・変換のアプリケーションを作り始めた。Excelで入力・変換し、WordやInternet Explorer等で表示するというものであり、Office連携スクリプトによる漢文資料作成補助ツールである。

周知のごとく、普通に「漢文」と呼ぶテキストには漢字だけが並んだ「白文」、それに句読点・返り点・送り仮名（本来は「添え仮名」と呼ぶのがよい^[注1]と言われるが）等の訓読記号（訓点）を施した「訓読漢文」、そして返り点に従って読み下し、漢字平仮名交じりの文語文として表記した「書き下し文」等、一つの文に対して表記のしかたが幾通りか存在し、訓読には振り仮名処理など日本語固有のマークアップが加わる。「古典」学習にいわゆる「漢文」とは「訓読漢文」を指し、高校生は返り点の約束や構文（句形）・文字の用法を学習しながら、「書き下し」を通して伝統的な翻訳（訳読）になじむわけである。教員にとっては、学習プリント・小テスト作りから定期考査問題の作成に至るまで、「訓読漢文」の入力・表示にはずいぶん悩まされるならわしで、だいいち通常のワープロソフトでは返り点と送り仮名を左右に等しく按配することさえ難しい（一太郎の漢文入力機能は字間が空く）。TeX（テフ）には漢文表示用のマクロを作った人があり^[注2]、一応教科書並みのきれいな出力を得られるが、これも第三・第四水準等の漢字の表示には特別な工夫が要るらしく、EasyTeXなどでは表示されないようである。TAIMUS標準のMS-Wordで訓読漢文を表示しようとすると、送り仮名には「上付き文字」、返り点と再読文字の送り仮名には「下付き文字」といった苦肉の手法で表示するしかない。しかも、いちいち文字ごとに書式設定を行うのもWordのインテリジェントな文字列操作の影響で文字選択の段階から難渋する羽目になる。

「漢文エディタ」と銘打った拙作Excel・Wordその他の連携マクロは、こうした漢文の入力の手間、表示の手間を極小に抑えることと併せて、漢文データの「電子テキスト」としての蓄積を目指している。ちなみに、TeXの環境をインストールしておけば^[注3]前記sfkanbun.sty（注2参照）のデータ形式に変換してTeXエディタにデータを渡すことも可能である。生徒に対して有用であるというよりは、教員に対して日常実務の省力化に役立つという位置づけのマクロだ。昨年前任校でIT関係のディスプレイをせざるを得なくなった際、授業の合間に15分ほど時間をとり、レ点（雁点）や一二点の簡単な表示変換の様子をPowerPointに埋め込んだ動画で流して、返読の概念を視覚的に把握する一助として生徒に披露した。そんなこともしてみたが、生徒は特に興味を引かれた様子もなかった。国語とコンピュータとの奇妙な付き合い方に興味を持つような少し風変わりな人の出現でも羨つしかなさそうだ。

それはさておき、ここで「漢文エディタ」^[注4]の使い方を紹介したい。尤も、完成品ではない。仕事の合間にたらたらと10年近くいじっているものの、ver. 0.65くらいの代物だ。それなら、どうしておこがましくもここで紹介するのか。簡単に言えば、この辺りで中仕切りをしておく必要があるからだ。これは、ユニークな作品である。「書き下し」の自動処理について、これの他にこれまで類似の「試み」を殆ど聞かない^[補足]。世に「漢文エディタ」と称する専用ソフトや一太郎のマクロは概して「入力・表示の省力化」のツールにとどまっていると言ってもそれほどの外的外れではない。データベースでなくとも、

文書の入力作業という少からぬ労力がかかるからには、「電子データ」の蓄積と利用という目標まで開発の射程範囲を広げる必要はないか。しきりとそう思うのだが、このような大風呂敷を広げるためには、少々草臥れた。ネットの諸規格の変貌は速いのだ。(この機会に補足しておく、「日本古典文学テキスト」アーカイブの発足は'97年であると前に書いた。これはかの「青空文庫」と同じ年のはずである。多数の参加による共同作業がどれほどの成果を挙げたかは、このサイトが今日の電子ブックのコンテンツとして揺るぎない地位を築いていることでも分かる。これに反して、一人相撲でのごたごたやってもこの程度というのが私のサイトの現状だ。しかし、電子テキストという《コンテンツの充実》に向けて、この20年ほどの間に画期的な進歩を示す事例はいかに乏しかったか。国語関連ではNDLのオンライン検索や「近代デジタルライブラリ」、国文学研究資料館、早稲田大学の「古典籍総合データベース」等、PDF画像を中心とした公開が進んでいるが、テキストベースでは民間の自助努力があるばかりとあってよい。できるかぎり信頼度を確保したテキストの提供を目指せば量はこなせない。それにも誤入力は必ず伴うのだ。)

長い前置きは以上で擱き、私が漢文テキストの普及に向けて工夫した点について具体的な紹介に入りたい。このアプリケーションの最大の特長は、一定のマークアップを施して入力したテキストを、上記のさまざまなスタイルのテキスト形式に自在に変換するところにある。主な機能は、次の通りである。

- ◎ 白文・訓読漢文・書き下し文への変換（書き下し文の変換精度は現在も調整中である。）
- ◎ TeX(LaTeX), HTML への変換とファイルへの書き出し
- ◎ Word ファイル、テキストファイルへの書き出し
- 文字種の変換、検索・置換、注記の抽出、データの結合・並べ替え（ソート）、印刷等
- 新旧漢字変換、ユニコード16進文字実体参照へのエンコード等

このうち、検索・置換・抽出・結合・ソート等の機能は、データ蓄積や印刷と同じくExcel本来の機能を利用している。ひらがな・カタカナ変換や新旧漢字変換、文字実体参照(Character Entity Reference)等への符号化(encode)と復号(decode)の機能は、テキスト処理のためにぼつぼつ貯め込んできた自作マクロを組み込んだものだ。特に文字実体参照は便利なもので、このおかげで通常ならエディタやHTML等の制限で表示できない「環境依存文字」等に対して、これを数値に変換することで「文字化け」することなくテキストベースで交換したり表示したりすることを可能にする。デザインされたフォントという文字実体に対応させた「コード番号」と似た考え方、エイリアス(alias)である。例えば、「槩」を槩に、「玉珎」を「玉玦」というように当該文字だけエンコードするとShift-JISの範囲でデータ交換が可能になる。そうでないと、これらの「環境依存文字」は高機能ワープロソフトやユニコード対応エディタ等でないと表示すらできず「?」のような表示に置き換わってしまうか、空白になってしまう。国語の教員は従来「今昔文字鏡」「indexfont」(16万字)等の高価なソフトに付属するフォントを使って表示するしかなかった(ただし、MojCharMapという昔のツールは今日でも入手可能)。今日ではATOKその他のIMEでも「手書き文字入力」や「文字コード表」等のツールがあり、フリーのフォントではIPA明朝やとくにIPAmj明朝(6万字)ができたおかげで、よほど表示が楽になった。とはいえ、文字の検索・データ交換に際しては依然として従来の障壁が立ち塞がっている。参考までに、私は自作「漢文エディタ」と合わせてフリーソフトの「Mana2」というエディタを使って古典作品をHTML化している。これはユニコード表示も可能なエディタだが、たいていはShift-JISモードで使い、マクロを登録しておき、異体字の取り扱い等を除いて「ほとんどあらゆる」文字を取り扱う。マクロの一例を挙げる。(選択状態にあるテキストを変数に格納し、これに様々な加工を施してもとのテキストに代入する。加工にはVBScript等を使うが、これはWindows標準で、言語の仕様も複雑ではないので文系人間でも取り扱える。)

txt = Editor.GetSelText ‘vbsではtxtなど適当な名前前で、タイプの定義なしに変数を使用できる。

txt = “《” & txt & “》” ‘この変数に選択範囲のテキストを格納し加工。ここは前後に文字を加えている。

txt = Editor.SetSelText(txt) ‘入出力の命令文はエディタによって規則が異なるが、Mana2ではこのようにし

て変換したテキストを代入する。この3行の枠組みを基本形として、加工過程を順次複雑化していく。

Script の本体は、下線を施した3行だけだ。これを、例えば「kakko.vbs」として保存する。エディタに登録して実行すれば「吾が輩は猫である。」という文が「吾が輩は《猫》である。」となる。この伝で、入力行と出力行の間(2行目)を、vbs の操作を少しずつ覚えながら自由に加工していけばよい。

話がとにかく脇道に逸れていくので、「漢文エディタ」に戻る。慣れるまでちょっと手間なのは、最初の入力時だけである。(ファイルはLZH形式で圧縮してあるので、Lhaca等のアーカイバ(archiver)を使って解凍する。名前は好きなものに改めて使えばよい。)xlsm ファイルを開くと(残念ながら TAIMUS ではマクロはほとんど働かないようだ)、1・2行目にボタン等のついた Excel 画面が出るので、「コンテンツの有効化」を押してマクロを有効にし、次に左上の「漢文入力フォーム」をクリックすると下のフォームが表示される。

起動



図1 起動時の画面とセキュリティの警告



図2 「漢文エディタ」入力画面

ここから入力操作を始めることになる。書誌事項はフォーム下方に入力できるが、これらは空白でもよいので、まずは真ん中の大きなテキストボックスに文字を打ち込んでいく。(本当は縦書きにしたいところだが、Excelにはまだ縦書きフォームというものがなく、リッチテキストフォームも標準では付属しないので仕方がない。)途中で入力や操作を間違えたら、DEL キーや BS キー等で修正する。決して Esc キーを使ってはならない。Esc を2回押すと、テキストボックスの性質として、それまで何百字入力していても、そのカードで入力したすべての内容が消えてしまって復活できない。そこだけ注意すれば、たいいていのは大丈夫である。一通り入力を終わったら、「登録」ボタンを押して内容を登録する。カードは次のデータを入力する状態になる。いったん終了して、表計算画面にデータが(入力時とはスタイルが少々変わって)転記されているのが確認できる。(この表計算画面でもデータの統合・配置換えなどの作業ができる。)

※ Excel・Word のバージョンや設定によって起動できない場合がある。Web にヘルプを出し、対処法を示してある。

データ入力 具体的な入力手順は、次のようになる。

【例1】 右のタイトル文の場合

- ① 「守る」と漢字ひらがな交じり文で入力。(自動でカタカナに変換。)
- ② 上の「レ」のボタンを押す。「^」という独自記号が入力される。
- ③ 「株を」と入力して、Enter を 2 回押す。1 行スペースができる。

【例2】 続けて本文 1 行目を作成

- ④ 上の「ルビ」ボタンを押す。「` `」というやや見苦しい独自記号が入力される。カーソル位置に「宋人」を入力し、キーボードの「→」ボタンで 1 文字右に移動し、「そうひと」と入力、さらに「→」ボタンで記号の外に出て「に」と入力する。(漢字以外は全てひらがなでよい。)
- ⑤ 続けて「有り」を入力、フォームの上の「二」点ボタンを押す。「2」のような記号が入力される。そのまま(半角スペースを削除してはならない)「耕す」と入力。「レ」点ボタンを押し、「ルビ」ボタン、「田」→(カーソルを移動)「でん」→(カーソルを移動)「を者」と続けて入力し、「一」点ボタンを押す。「守る^株を」「宋人`そうひと`に有り 2(耕す^田`でん`を者)1。」のように打つことになる。慣れるまでは、多少違和感があるが、要は印刷の順次に従ってぼつぼつと漢字ひらがな交じり文で入力していく。この方式は、慣れると意外な速度で打てるようになる。(図4)



図3 サンプル文



図4 入力文と使用したボタン

- ⑥ 入力が終了したら、「登録」ボタンを押して入力内容を確定する。

※ なお、文字実体参照はあくまでデータの記録・交換用であり、各種変換はこれに対応していない。

次に、「2 データ操作」のタブを選択する。ここで現れる画面から、ほとんどの変換操作を行う。



図5 「データ操作」タブをクリック。タブが切り替わる。



図6 データ操作画面

変換1 「訓読漢文」のWordへの書き出し

- ① 図6の番号順に操作。まず下向きの「▽」ボタンを押すと、最後(最新)のデータが表示される。
- ② 「訓読漢文」ボタンを押す。下段に記号スタイルの若干変換した文が表示される。(図7)
- ③ 続けて「MS-Word」ボタンを押す。Office 連携マクロが動作し、Word が起動して Word 流の「訓読漢文」が表示される。(図8)

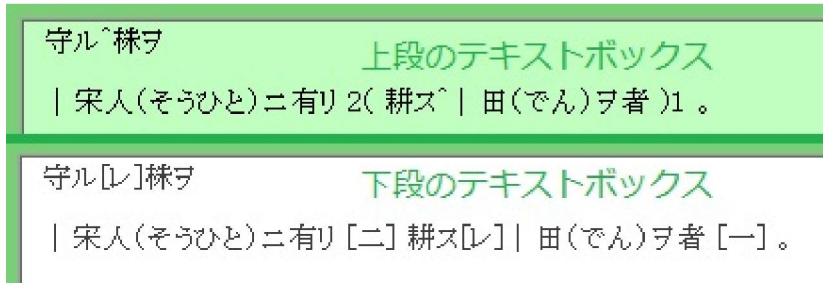


図7 「訓読漢文」ボタンを押した状態。(下段に表示される。)

図8(右) ワードが起動し、それなりに体裁を作った文ができる。
(ルビ処理や割注処理は、マクロ化できない。Wordの制限である。)



変換2 「白文」への変換

- ④ 「白文」ボタンを押す。(図9)結果の書き出しもちろんできる。

変換3 「書き下し文」への変換

- ⑤ 「書き下し文」ボタンを押す。(図10)

※ 万一フリーズした場合は、Ctrl+Shift+Delでタスクを消去。

- ⑥ その状態で「MS-Word」ボタンを押せば、「書き下し文」が表示される。(図11)
- ⑦ また、下の「<ruby>」ボタンを押すと、HTMLのルビ・タグが付加される。(次項)変換結果を下の「Clip」ボタンでクリップボードに送ることができ、Wordやエディタに貼り付けることができる。

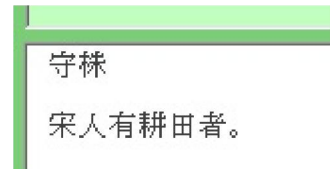


図9 白文の表示

変換4 HTMLへの変換

- ⑧ データ操作画面にある「HTML」の文字をクリックするとボタンの種類が入れ替わる。(図12)
- ⑨ この状態で「TAG(IE)」を押す。(難字を表示するため)ワードパッドが起動し、HTMLファイルのタグ付けを施したテキストが生成される。ワードパッドのままではワードと同じくリッチテキストになってしまうので、(難字は文字実体参照にするなど処理した上で)「○○.htm」等の名を付け、ファイル形式では「テキストドキュメント」を選んで保存する。

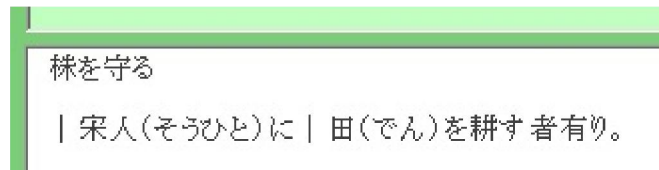


図10 書き下し文への変換

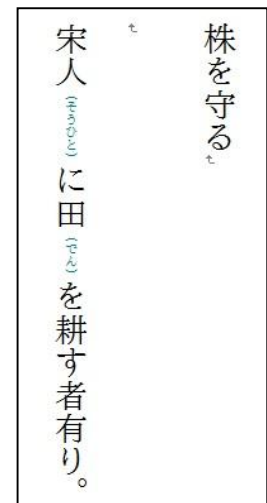


図11 Wordに転記



図12 HTML関連のボタン類

これを Internet Explorer で開くと《縦書き》表示の HTML ドキュメントとなる。
 (縦書きは、今なお IE でしかできない。) (図 13)

変換 5 TeX への変換

- ⑩ データ操作画面にある「LaTeX」の文字をクリックするとボタンの種類が入れ替わる。(図 12) ※ TeX への変換には、予め TeX の環境を作っておく必要がある。
- ⑪ この状態で「LaTeX-漢文」ボタンを押す。入力文が sfkanbun.sty マクロの指定形式に変換される。そのまま (TeX 本体と EasyTeX エディタが正しくインストールされていた場合)「EasyTeX」ボタンを押せば EasyTeX が立ち上がる。(図 15)
- ⑫ EasyTeX 画面で「コンパイル」する。(この時、sfkanbun.sty が「漢文エディタ」の「保存先」テキストボックスで指定した箇所に入っていないとコンパイル・エラーが起こる。)ファイル保存画面が現れるので、「oo.dvi」という形で DVI ファイル (表示用ファイル) を保存する。

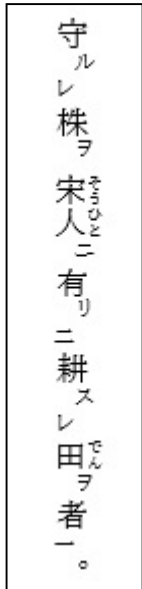


図 13



図 14 LaTeX のマクロパッケージ sfkanbun.sty 用書式に変換

- ⑬ 続けてエディタ画面から「DVI」メニューを実行すると dvi ファイルが実行され、sfkanbun.sty で整形された TeX 形式の表示が出来上がる。(図 16)

TeX についての知識と関心を私はあまり持ち合わせていない。一応 Word よりはやや見栄えのする表示ができるというに過ぎない。

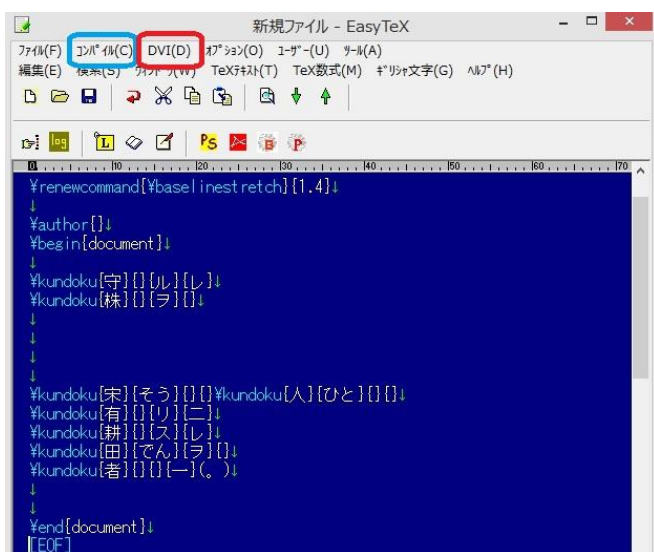


図 15 EasyTeX に出力して DVI ファイルにコンパイル

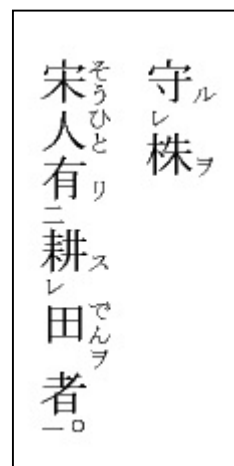


図 16 DVI への出力結果

以上、駆け足で入力から各種表示形式への変換までの作業手順を示した。その他、Notepad 等にテキストファイルとして書き出すことや、データの結合・並べ替え・検索・抽出等も可能である。HTML は横書きが今なお基本であるが、これでは訓読漢文の視認性が極端に劣ってしまう。Internet Explorer は Ver.10 から HTML5 に対応したため、これまでの多彩な表示を維持するためのタグ付けの仕様がすっかり変わってしまった。それでも、《縦書き》表示はこのブラウザに依存するしかない。なお、現在は思いつきに止まるが、「青空文庫」形式に変換することで、このタグ付けに対応した AIR 草紙（シェアウェアだが優秀なツールである。）や T-Time などのビューア（viewer）を活用でき、「電子読書」が可能になる。その他補助ツールを通しての EPUB 形式への変換など。これらの意味は今日大きいかもしれない。

授業や定期考査・補助プリント等に活用する場合には MS-Word への書き出し機能が役に立つはずである。インターネット（ウェブ）での公開には HTML への変換を、データの蓄積と交換には MS-Excel ファイルそのもの、または「漢文エディタ」独自形式のテキスト書き出しが使えるだろう。ただし、「環境依存文字」や Shift-JIS で対応できない文字を含むデータの蓄積・交換にはユニコード 16 進文字実体参照への変換を行い、数値・記号とアルファベットの組み合わせにより代替する。そうしてできたテキストファイルをブラウザで表示したり、ユニコード表示対応のワード・プロセッサかエディタ等でデコードして表示したりする。私は前述のエディタに自作マクロ^{〔注5〕}を組み込み、変換している。表示フォントも、「MS 明朝」等だと「文字化け」する場合があるので、「New Gulim」等の韓国フォント（表示に対応する文字種が多い。）か「MinLiU」「Simsun」「Simhei」「FangSong」等の中国フォントを当該文字に指定することでブラウザ等での表示が可能になる。（これらはほとんどが Windows 標準のフォントである。Word や一太郎を使う場合には、このように一変換する必要はなく、表示の問題は生じないだろう。また Word から HTML に書き出す場合にも「HTML 形式」で保存すると、自動で JIS10 進文字実体参照に変換されて表示されるようだ。なお、「鷗外」「啄木」等の旧字体表示に使える「癸羊明朝」という便利なフォントもある。）

〔例〕 李徴と袁愔 → 李徴と袁傪

以上の各種変換は「書き下し文」への変換以外は、ほとんど誤変換が無いはずである。なぜなら、機械的にタグを付加・削除するだけの機能だからだ。しかし、「書き下し文」への変換は、なかなか容易に問屋が卸さない。特に再読文字の変換には、何年もてこずっている有様だ。また置き字の処理なども中国語と日本語の言語の仕組みが大きく異なる点である。例えばいわゆる「於于乎」等の前置詞を単に削除する処理だけでは、「於戲（ああ）」「于定国」「炳乎」等の語句はどうするかという問題が出て来る。接続詞「而」なども同様だ。文章解析・構文解析等の高度な技術でもなければ満足いく結果は望めないに違いない。機械的なパターンだけでは生きた言語は容易に掌握しきれない。にもかかわらず、これが「漢文訓読」の最終段階であることも確かである。所詮完璧は期しがたいので、むしろコンピュータのツールは人間が修正する手間をなるべく抑制する補助手段であればよいと割り切って使うことが、コンピュータに振り回されないための健康法ともなるだろう。それでも、最近はわりあい練れてきて誤変換やフリーズ現象も比較的少なくなってきたので、この機会に紹介してみようと思ったわけである。雑然としたコードだが正規表現（Microsoft VBScript Regular Expressions 5.5）を使っているので、200 字や 400 字程度の文章なら瞬時に変換する。その代わりに点の付け間違いをすると、このパワフルなスクリプトによってまず間違いなく Excel は「暴走」して画面が真っ白になってしまう。（万一そうってしまった場合には、Ctrl+Shift+Del キーでタスクマネージャを起動し、メモリに残存している Excel を強制終了させる。その場合を想定して、データをバックアップする仕様にはしてあるが保存していないデータは救出できない場合がある。）最近では、代表的なミスの可能性については、「書き下し文」にする前にチェックするように作ってある。

この《「書き下し文」変換機能》も、《データの蓄積と再利用》と合わせて「漢文エディタ」の重要な特長である。先蹤は皆無とっていたので、変換の仕組みを一から作った。変換精度もまずまずだが、

出来上がった「日本文」は、置き字や終尾詞等の助字類をいたずらに削除して誤読の原因とならぬよう、相当数をそのまま残している。置き字は彼我の言語仕様の違いを表し、終尾詞を訓読する可否の判断は個人差が大きい。その他、古典原文のカタカナ類は「読み仮名」「送り仮名」であるよりは「添え仮名」であるのが実態であろうから、これを「送り仮名」としてみると、出来上がったものが日本文としては極端に仮名文字が少い、あるいは送り仮名がおかしいということになる。どうしても多少怪しげな日本文とならざるを得ないが、とにかく読めることの便益は大きい。(副詞では「最後の一字を送るのが原則」というのも添え仮名ゆえだろう。他にも「者」が「ハ」であり「モノ」の場合「者ノ」と書くとか、「中」は「ウチ」と読み、「中カ」を「ナカ」と読ませるなど、名詞でも同様のことがある。活用語においても、誤読を防ぐために最小限添えた「添え仮名」なのである。戦後定められた「送り仮名の基準」通りに活用語尾を送ってあるわけではない。教科書本文だけを対象にするのではなく、多少なりとも汎用性を考えるなら想定どおりの「書き下し文」が出力できなくてもしかたがない。古典の原文に従う必要がない場合には、自分で考えた基準通りに読み仮名・送り仮名を入力すればよい。)

「書き下し文」変換、その他の変換の実例は、現に古典テキストの入力文を作成中なので、前記「古い訓読文」の用例なのだが、そちらで出来栄を参照してほしい^[注6]。

「漢文エディタ」はマークアップ方法が独特である。ありふれた記号類では各種変換に際してエラーにつながるので、なるべく「通常は使わない記号」を採用したからである。ただし、一二点では ()、上中下点では { }、甲乙丙点では [] を使った。それぞれの返り点の特徴が数学記号とも照応して、ちょっと見にも符号化・変換のイメージが把握しやすくなっていないだろうか。レ点は ^ を使い、最小スペースで片付けている。もとよりマークアップ一つでも「好み」というものは分かれるので、この方式が普及するとはさすがに想定していない。それでも、曲がりなりに「書き下し文」変換を含め、統合的な「古典漢文のマークアップ」の試作品として自ら位置づけ、ここに紹介するのである。

なお、TAIMUS ではセキュリティ・チェックによるせいか、Excel マクロ自体がそのままでは働かない。動いた場合でも 2003 以前の xls 形式に戻して保存し VB エディタを起動してから実行するなど、実用どころでなく起動にすら四苦八苦する有様である。自宅か他の環境で使っていただくしかなさそう。一方、文字コードの規格は日進月歩である。Windows8 では Unicode IVS という異体字取り扱いの符号化規格が採用された。「漢文エディタ」のいくつかの機能はどこでもたまたま古びてしまうか分からない。Excel に依存する寄生木状態なのだが、漢文の電子化にはこれでも頼もしいツールである。(2014.2)

[注記] (敬称略)

- (1) 「漢文の場合は『送仮名』ではなく、漢文を読むために必要な仮名を添えたものであって、語幹に対して語尾を送る送り仮名とは違うものである。例えば『^ミ自^ラ』は『みずから』、『^オ自^ラ』は『おのずから』を示すために添えた仮名である。そのほか最後の一拍を添えると便利な『故(も)と・唯(た)だ・猶(な)ほ』などのト・ダ・ホも送り仮名ではないのだから、添え仮名というのが非常によい称呼であるが、(以下略) (原田種成『漢文のすゝめ―諸橋「大漢和」編纂秘話―』(新潮選書 新潮社 1992.9.15) pp.112-113)
- (2) 藤田眞作「漢文の訓点文の組版」にマクロ・パッケージ (sfkanbun.sty) が公開されている。
(<http://homepage3.nifty.com/xymtex/fujitas/kanbun/kanbunex.html>)
- (3) 奥村宏「TeX Wiki - EasyTeX」(<http://oku.edu.mie-u.ac.jp/~okumura/texwiki/?EasyTeX>)
- (4) <http://www2s.biglobe.ne.jp/~Taiju/index.html>
- (5) unicode_entity_reference.vbs (エンコード用), unicode_entity_reference_d.vbs (デコード用) 等。
(http://www2s.biglobe.ne.jp/~Taiju/leaf/tjsoft49_unicode_entity_reference.htm) なお、MS-Word には Word VBA (oo.bas) を組み込む。Word 画面で Alt+F11 キーを押すと VB エディタ (Microsoft Visual Basic エディタ) が立ち上がるので、「ファイル」-「インポート」で bas ファイルを指定して読み込むと、Word から「ツール」-「マクロ」で呼び出して使えるようになる。
- (6) 「世説新語」(http://www2s.biglobe.ne.jp/~Taiju/5c_sesetsu_shingo_01.htm)、「聖徳太子伝暦」(http://www2s.biglobe.ne.jp/~Taiju/10c_shoutokutaishi_denryaku_1.htm) 等を参照。それぞれまだ入力途中のものだが、各種変換の出来上がりのイメージを掴むことができる。

[補足]

(1) テキスト形式・MS-Word 形式・HTML 形式その他への自動変換と合わせて「書き下し文」への変換を試みたこの作品では、変換方法を一から工夫してきた。この機会に類似の試みを Web で探してみると、「漢文作成」マクロ(①)、「白文」「訓読漢文」「書き下し文」作成スクリプト(②)を試みているサイトがあった。①は Word での「表示」に特化しているが、フォントの大きさや文字間隔まで調整できる。②はブラウザでの表示を前提にしているようである。データの蓄積に関しては「青空文庫」のタグ付けを前提にしているが、入力用のツールはまだ無いようだ。(また、「漢字データベースプロジェクト」の更新それ自体が 2010 年で止まっている。)なお、このプロジェクトでは「花園フォント」(9 万字)を配布している。収録文字数からいえば、IPA フォントを超えるようである。

①大島裕明「漢文便利だなー」(<http://www.dl.kuis.kyoto-u.ac.jp/~ohshima/wiki/index.php?FrontPage>)

②漢字データベースプロジェクト(上地宏一(大東文化大学)作成か。)
「漢文訓読 JavaScript」
(<http://kanji-database.sourceforge.net/software/kanbun.html>)

なお、①の「フォントの大きさ指定」に着想を得て、Word に出力するフォントの種類・大きさ・色を指定するフォームをこの機会に作った。Excel から Word に値を渡す方法も調べ、より「連携」マクロらしくなった。(図 18)

(2) 独自のマークアップ付加は、誰しも抵抗感があるかと思う。また、難解な文字は文字実体参照を使う以外に、単語登録しておき「訂正保存」ボタンを使って再登録する方法もなぜか使える。便利な「漢文エディタ」だが、「漢文を作成するのに、どうして一々 Excel を起動しなければならないか」と面倒に思える時もあった。その時作成したのが Word VBA による「簡易漢文エディタ」(http://www2s.biglobe.ne.jp/~Taiju/leaf/tjsoft93_easy_kanbun_editor.htm)である。これは直感的なタグ付けにしたので、入力に際して感じる抵抗感はあまり無いはずである。(注記の制約等はある。) Word の入力フォームなので、結果は Word に直接書き出す。それを後で「漢文エディタ」の別のタブに組み込み、独自マークアップに自動変換することで細部は不十分なままに一応「統合」した。(図 19) もう一歩進んで、初めから抵抗感の少いやり方で進めそうなものだが、まだそういう工夫はしていない。現在は、古典漢文テキストの他、高校生向けの『漢文提要』の類の各種例文を入力し、「書き下し文」への変換精度を調整している。



図 17 再読文字は直感的に入力できる。



図 18 Word への出力文字の体裁を指定。

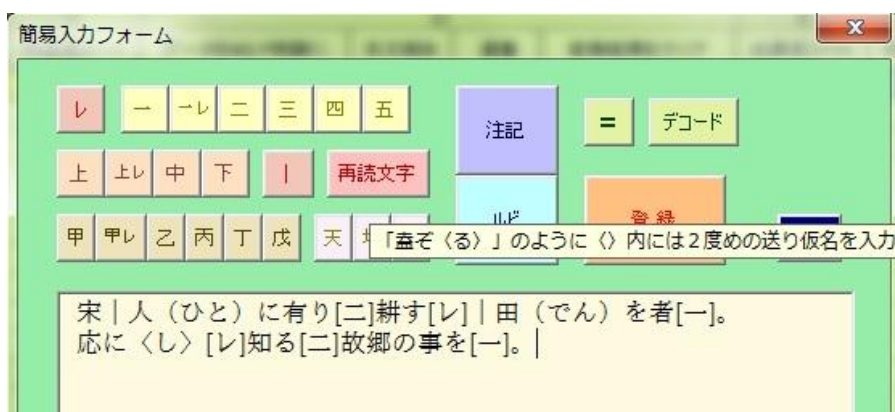


図 19 マークアップを見やすくした簡易入力フォーム。
(Word ベースの「簡易漢文エディタ」を組み込んだもの。)