

8. 漢字の読み方辞典と単語親密度チェッカー

Kanji Pronunciation Dictionary and Word Familiarity Checker

あらまし 新しい詳細読み策定のために開発した策定支援システムの機能は他の研究者・教育関係者にも便利だと思われたので、その一部を一般公開用に作り直した。漢字の読み方辞典では、JIS 第1水準、及びJIS 第2水準に含まれる漢字 6355 字の読み情報を検索できる。読み方は、教育段階で習う読み、常用漢字表に記載されている読み、それ以外の読みなど、難易度に分けて表示される。その漢字が属している文字集合、教育漢字であればその配当学年、シフト JIS コードも表示する。単語親密度チェッカーでは、複数の単語の親密度を1度に求めることができる。

1. 漢字の読み方辞典

1. 1 機能

漢字の読み方辞典には JIS 第 1 水準、及び JIS 第 2 水準の漢字合計 6355 字の読み情報が収められている。その機能は以下の通りである。

- 1) 音読み，訓読みを区別した表示
- 2) 対象漢字が属する文字集合の表による読みと，それ以外の読みを区別した表示
- 3) 対象漢字が属する文字集合（教育漢字，常用漢字，JIS 第 1/2 水準）の表示
- 4) 教育漢字の配当学年の表示
- 5) シフト JIS コードの表示

1. 2 典拠

漢字の読みの典拠としたのは，以下に示す資料である。

- 1) 教育漢字
 - 『学習指導要領準拠—漢字指導の手引き 第四版』（教育出版，2001）
- 2) 常用漢字
 - 常用漢字表（文化庁，1981）（表内の読み）
 - 『漢字必携 一級』（日本漢字能力検定協会，2003）（表外の読み）
- 3) JIS 第 1 水準漢字・JIS 第 2 水準漢字
 - 『JIS X0208:1997 7 ビット及び 8 ビットの 2 バイト文情報交換用符号化漢字集合』（日本規格協会，1997）（表内の読み）
 - 『漢字必携 一級』（表外の読み）

1. 3 動作環境

漢字の読み方辞典は Excel のマクロであり，以下の環境で動作を確認している。

- OS : Microsoft Windows XP Professional
- Excel : Microsoft Excel 2003 または Microsoft Excel XP

ここで挙げた環境以外でも動作すると思われるが，動作確認は行っていない。

1. 4 使い方

漢字読み辞典は五つのワークシートから成っている。読みを調べる作業では [検索] ワークシートを使用する（図 1）。ワークシートの上部にある入力フォームに，読みを調べたい漢字 1 字を入力してから，Enter キーを押すか，あるいはフォームの右隣にある [検索] ボタンをクリックする。すると，読み情報，その漢字が属している文字集合，

シフト JIS コード等が表示される。

他のワークシートには、教育漢字、常用漢字、JIS 第 1 水準、JIS 第 2 水準の各表の読み方一覧を記している。

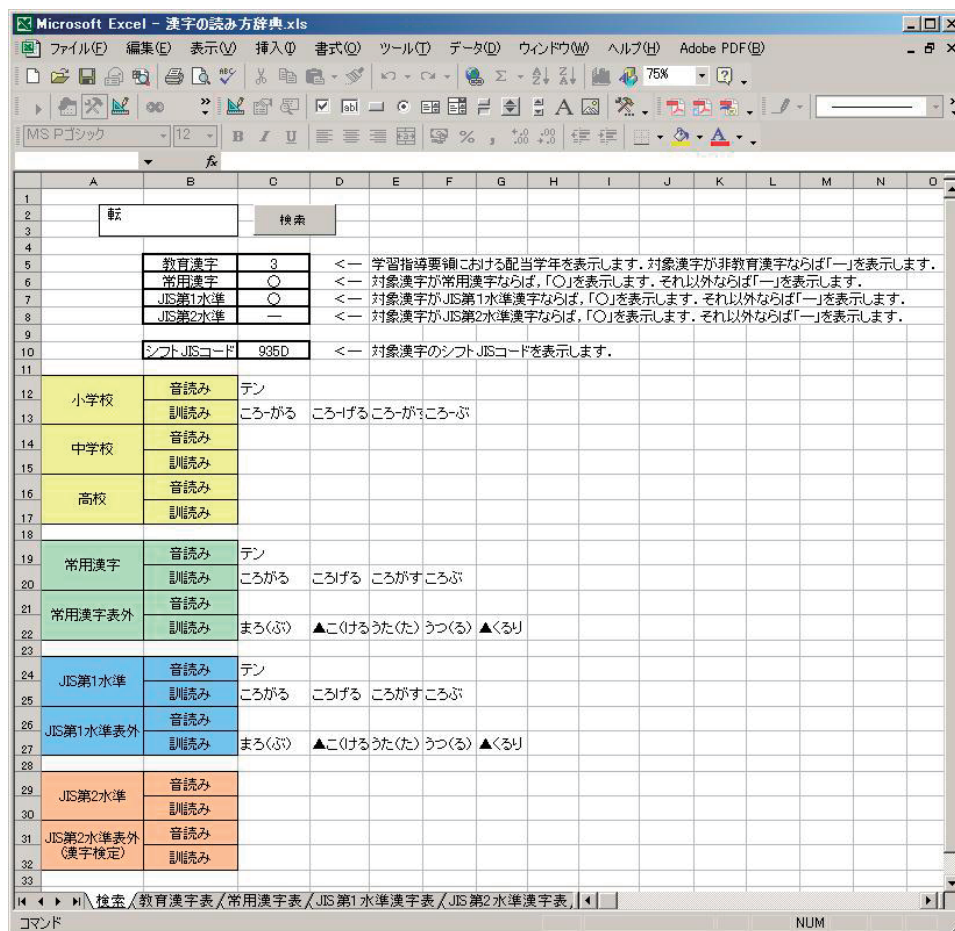


図 1 漢字の読み方辞典の実行画面 (Microsoft Excel XP)

2. 単語親密度チェッカー

2. 1 概要

単語親密度チェッカーは、『NTT データベースシリーズ 日本語の語彙特性 第 1 期 CD-ROM 版』(三省堂, 2003。以下、『日本語の語彙特性』と表す)に収録されている文字音声単語親密度, 文字単語親密度, 音声単語親密度のデータを表計算ソフト Microsoft Excel 上から調べるためのマクロプログラムである。『日本語の語彙特性』にも親密度検索プログラムが添付されているが, 1 度に一つの単語の親密度しか求められない。これに対して, 開発したマクロプログラムでは, 複数の単語の親密度を 1 度に調べられる点が特長である。実験で用いる単語群の親密度を調べる際などに便利である。

2. 2 動作環境

単語親密度チェッカーは Excel のマクロであり、以下の環境で動作を確認している。

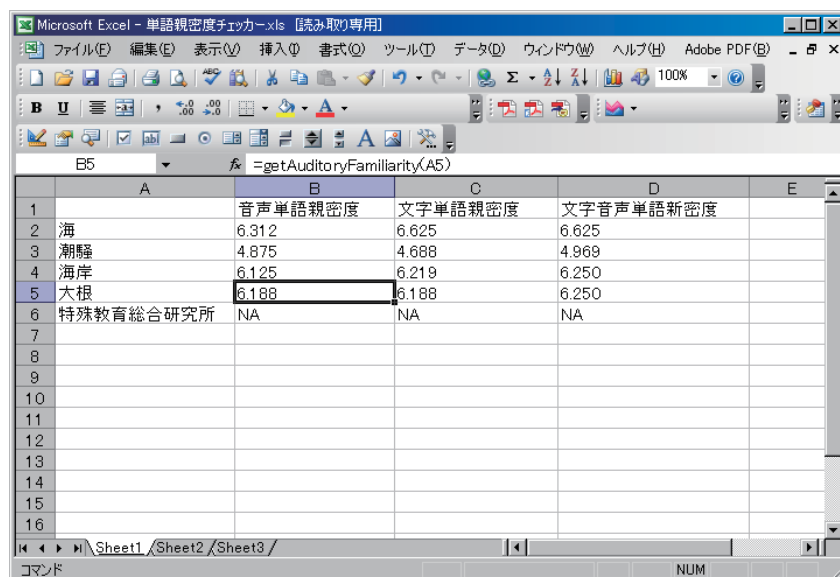
- OS : Microsoft Windows XP Professional
- Excel : Microsoft Excel 2003 または Microsoft Excel XP

2. 3 使い方

単語親密度チェッカーの実行画面を図 2 に示す。左端の A 列には単語親密度を調べる語を入力する。B, C, D 列にはそれぞれ、音声単語親密度、文字単語親密度、文字音声単語親密度を調べた結果が表示されている。日本語の語彙特性の中に収録されていない語を調べようとした場合は、図中 6 行目にあるように、「NA」が表示される。

単語親密度を調べるために利用できる関数は以下の通りである。いずれの関数も、引数にチェック対象の文字列かセルを指定すると、それぞれの親密度を得ることができる。

- getAuditoryFamiliarity() : 音声単語親密度を返す。
- getVisualFamiliarity() : 文字単語親密度を返す。
- getWordFamiliarity() : 文字音声単語親密度を返す。



	A	B	C	D	E
1		音声単語親密度	文字単語親密度	文字音声単語新密度	
2	海	6.312	6.625	6.625	
3	潮騒	4.875	4.688	4.969	
4	海岸	6.125	6.219	6.250	
5	大根	6.188	6.188	6.250	
6	特殊教育総合研究所	NA	NA	NA	
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					

図 2 単語親密度チェッカーの実行画面 (Microsoft Excel 2003)

謝 辞

膨大な量の漢字の読み情報の入力、当研究所で非常勤勤務している村上真理氏に御骨折り頂いた。