

単語レベルを設定可能な レベルチェッカー

-単語レベルデータ共有の提案-

住田真一 北村達也

甲南大学知能情報学部

kabutoeditor@gmail.com

リーディング・チュウ太

The screenshot shows a web browser window with the URL `basil.is.konan-u.ac.jp/chuta/`. The page title is "リーディング チュウ太" (Reading Tutor) and it includes a sub-header "日本語読解学習支援システム". The main content area features a large text input box for pasting text, followed by buttons for dictionary lookups: "辞書: 日→日, 日→英, 日→独, 日→蘭, 日→葡, 日→西, 日→葡, 日→英(ふりがな)". There are also buttons for "レベル判定: 読義, 漢字" and a "検索" (Search) button. A logo for "Reading tutor" with a mouse character is visible. Below the main area, there are sections for "ニュース" (News) with two entries, "チュウ太について" (About Chuta), "工具箱" (Toolbox) with links to "辞書ツール, レベル判定ツール", "日本文学 珠玉の小品集" (Japanese Literature: Precious Short Stories), and "読解教材バンク" (Reading Material Bank).

チュウ太のレベルチェッカー

- 文章を単語のレベルごとに色分け
- 各レベルの単語数を表示
- 文章全体の難しさも表示



チュウ太のレベルチェッカーの処理例

Japanese word level ch... x

basil.is.konan-u.ac.jp/chuta/level/

よく見るページ Firefox を使いこなそう Web スライス ギャ... おすすめサイト

甲南大学
KONAN UNIVERSITY

チュウ太のレベルチェッカー
開発中のため予告なく仕様が変更されることがあります。
様々な情報をアナウンスするためのブログを開設しました(2013/4/15), <http://chutainfo.blogspot.jp/>

- 2014.11.11 [かよはてなでウェブ記事](#)
- 2014.10.28 [シンポジウムプログラム公開](#)
- 2014.10.27 [1万人アンケート集約の単語リスト公開開始](#)

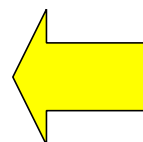
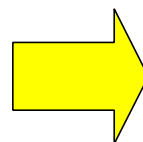
入力した文章を単語ごとにレベル分け・色分け・カウントします。
↓文章を入力して下さい。

香港は思っていた以上に暑い。

col kolcol colnbun keizakol tanso reset

freq faml mEF yEF freq

- col: 旧日本語能力試験出題基準に基づく分類
- kolcol: 介詞用語のレベルに基づく分類
- colnbun: colの構図に加え、一文の平均文字数および文章のレベルを表示
- keizakol: 経済用語のレベルに基づく分類
- tanso: 筑波大学との共同研究による新基準に基づく分類
- freq: 朝日新聞記事における頻度情報に基づく分類(10段階表示)
- faml: 単語頻度に基づく分類
- mEF: 朝日新聞記事データベースから作成した語に基づく分類



result

basil.is.konan-u.ac.jp/cgi-bin/switcher.cgi

よく見るページ Firefox を使いこなそう Web スライス ギャ... おすすめサイト

level5: Black, level5: SaddleBrown, level4: Purple, level3: Blue, level2: Green, level1: Magenta, others: Red

香港は思っていた以上に暑い。

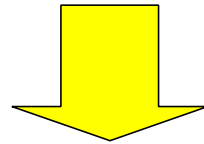
総数	語彙総数	level4	level3	level2	level1	その他	助詞助動詞接続詞	記号類
10	9	2	2	0	0	1		4
111.1%	100.0%	22.2%	22.2%	0.0%	0.0%	11.1%		44.4%
(10)	(9)	(2)	(2)	(0)	(0)	(1)		(4)
111.1%	100.0%	22.2%	22.2%	0.0%	0.0%	11.1%		44.4%

- Level4 2 (2)
 - 暑い 1
 - いる 1
- Level3 2 (2)
 - 思う 1
 - 以上 1
- その他 1 (1)
 - 香港 1
- 助詞助動詞接続詞 4 (4)
 - た 1
 - は 1
 - て 1
 - に 1
- 記号類 1 (1)
 - 。 1

Japanese word level checker © Reading Tutor

研究の目的

- チュウ太のレベルチェッカーでは単語レベル判定時にページが切り替わってしまう。文章の編集目的で利用するには不向き。
- レベルの判定基準が変更できない。



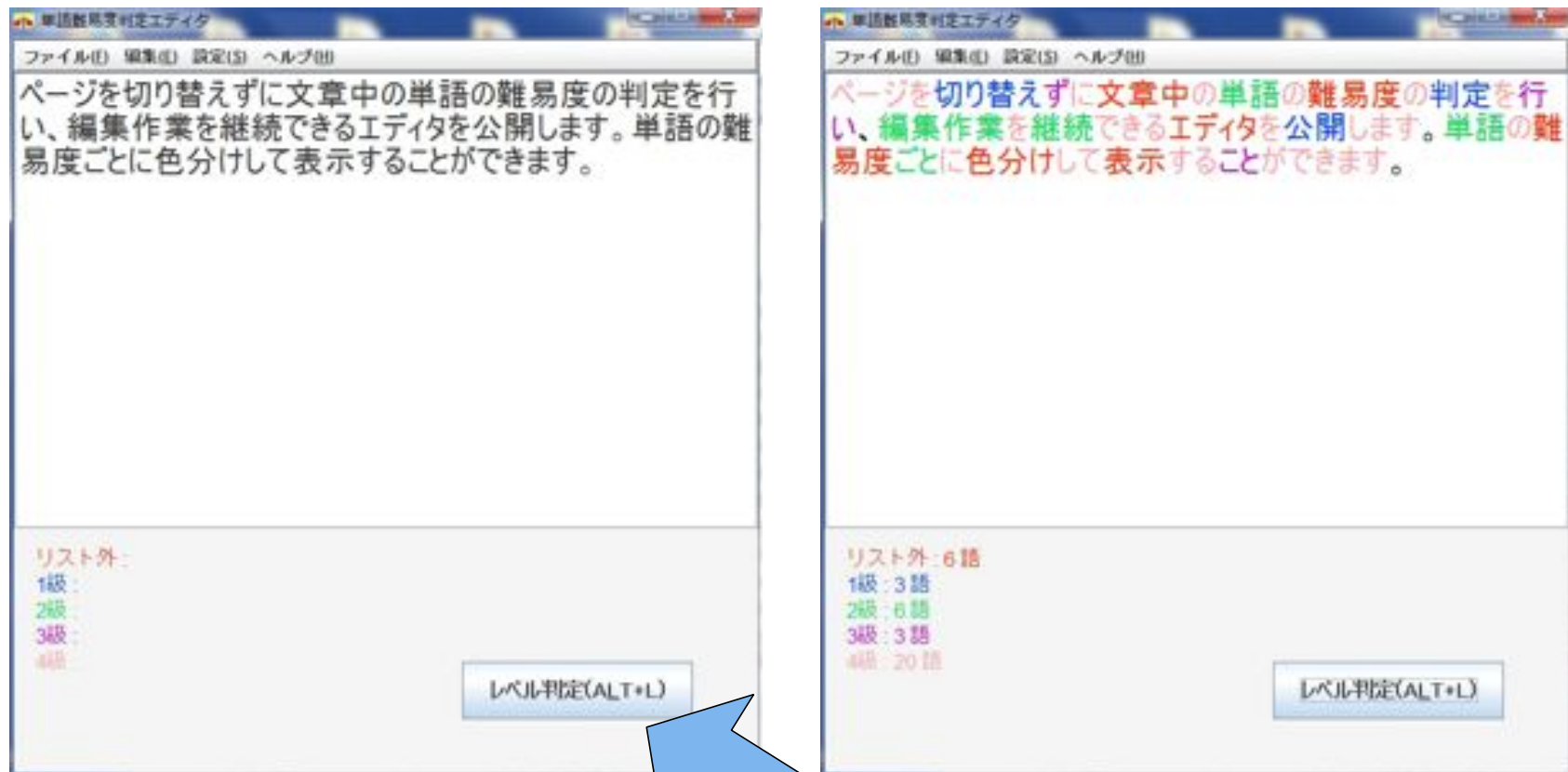
単語リストを変更することでレベルの判定基準を変更でき、文章を編集しやすいエディタを開発した。

かぶとエディタ

- Microsoft Windowsのメモ帳のような文章編集ソフト
- 文章中の単語をレベルごとに分けて色分け
- Windows、MacOSXで動作



処理の例



かぶとエディタの機能

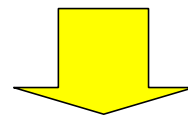
- 文章を形態素解析し単語ごとに分割。そして利用者が指定した単語リストに基づいて単語の色分けを行う。
- 単語リストを変更することで判定基準を変えられる。
- カット、コピー、ペースト、印刷などエディタの基本的な機能をもっている。

形態素解析

- 文法のルールや辞書を参考にして、文章を意味がわかる最小の単位に分ける作業

(例文)

今日は雨が降っています



今日 | は | 雨 | が | 降っ | て | い | ます

名詞

助詞

名詞

助詞

動詞

助詞

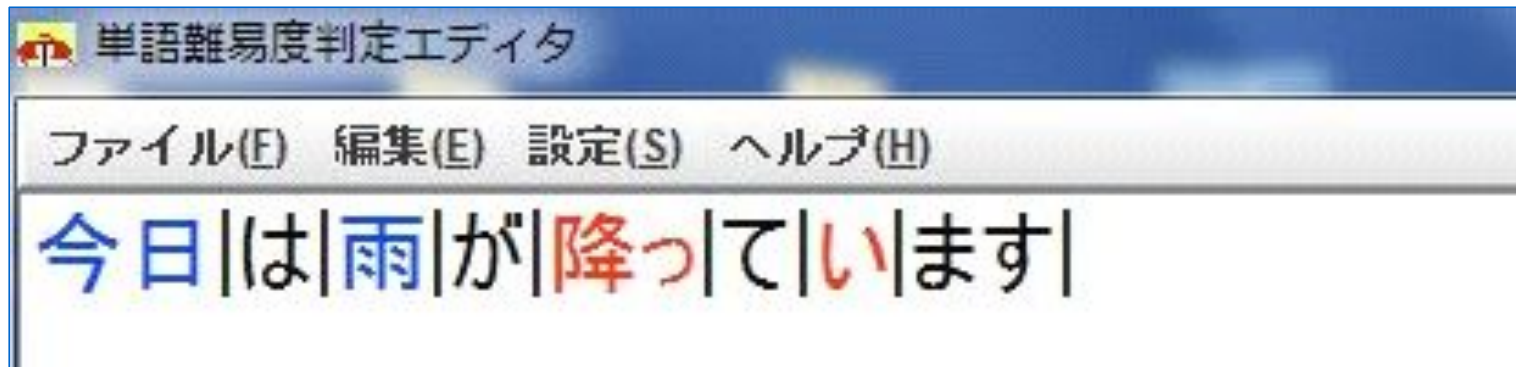
動詞

助動詞

形態素解析ツール

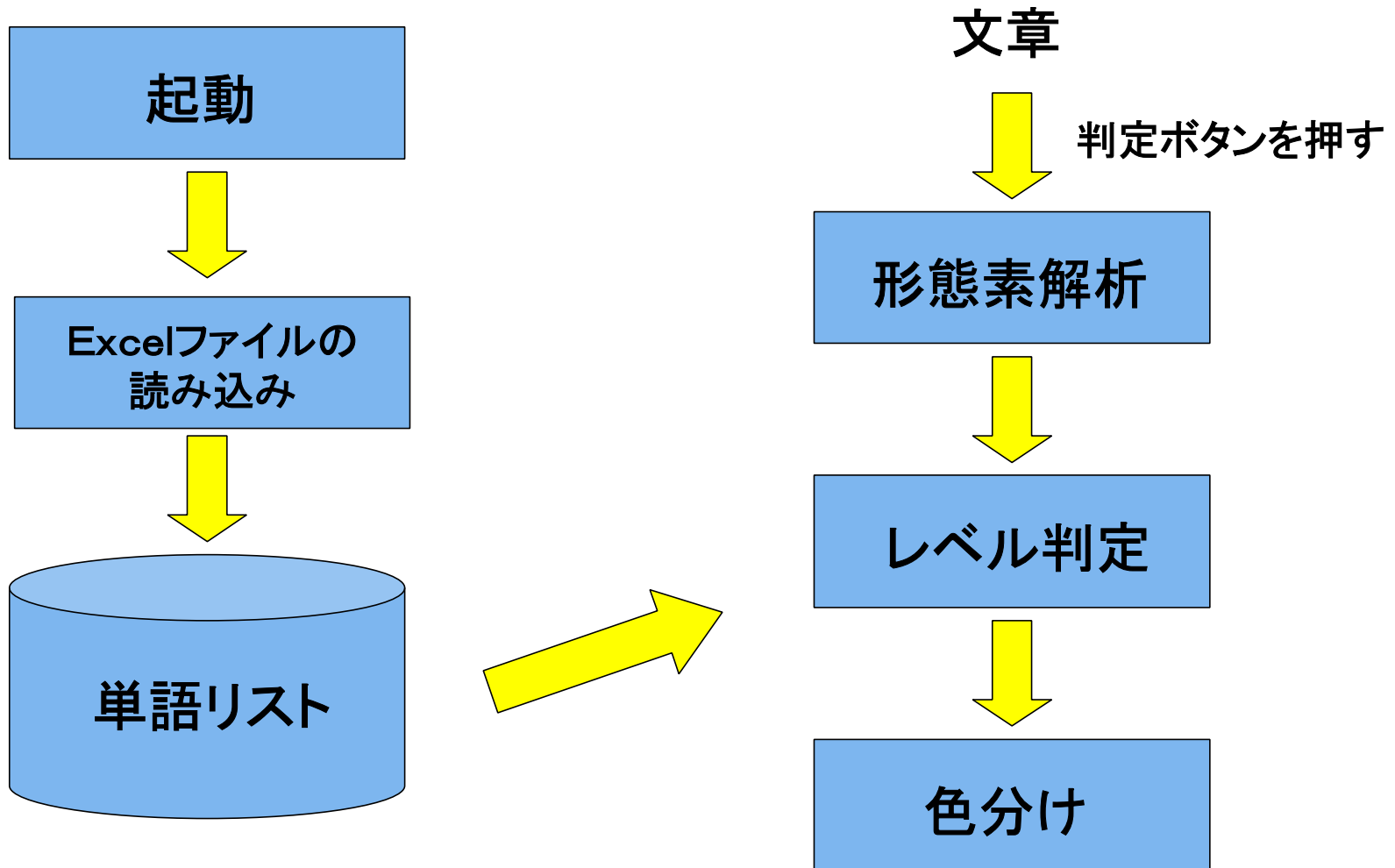
- 形態素解析ツールはYahoo! APIとMeCabから選ぶことができる。
- Yahoo! API
 - Yahoo! Japanが提供しているテキスト解析ツール
- MeCab
 - 日本の研究機関によって開発されたオープンソースの形態素解析ツール

形態素表示モード



かぶとエディタにはこんな機能もあります！

かぶとエディタの仕組み



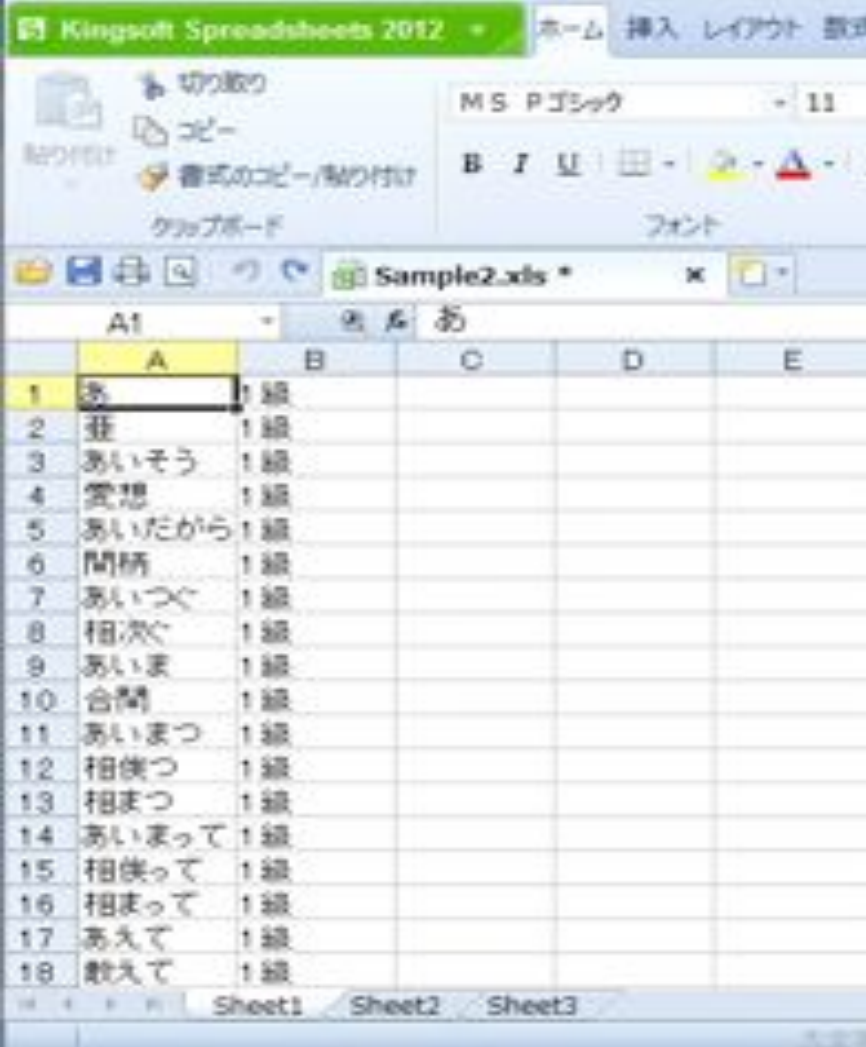
「げんき」シリーズの単語リスト

- The Japan Times社の教科書「げんき」シリーズに対応した単語リストを公開しています。
- 7つの単語リスト
 - 「会話・文法編1～9課」
 - 「会話・文法編10～17課」
 - 「会話・文法編18～23課」
 - 「読み書き編1～9課」
 - 「読み書き編10～17課」
 - 「読み書き編18～23課」
 - 「げんきあいさつG課」



単語リストの作り方

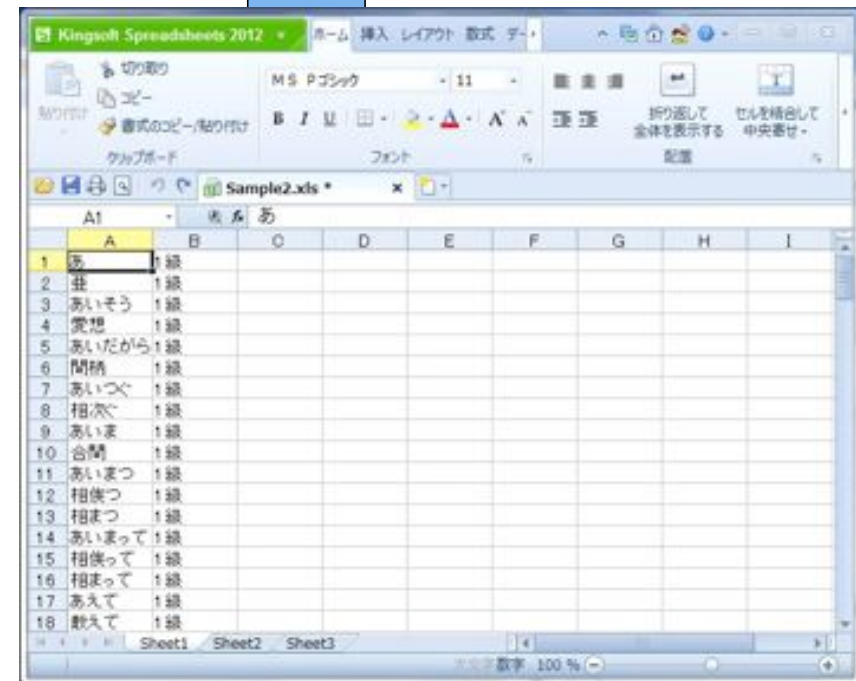
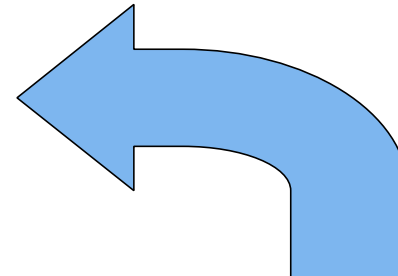
- Excelで簡単に作ることができる。
- 単語とレベルを並べて書く。



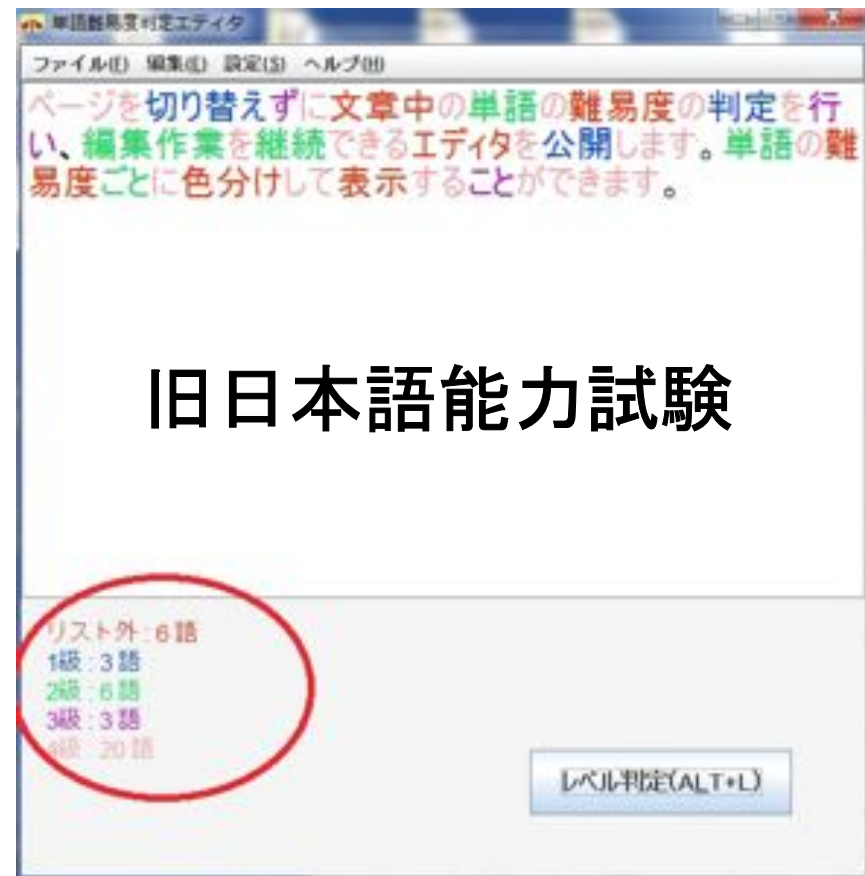
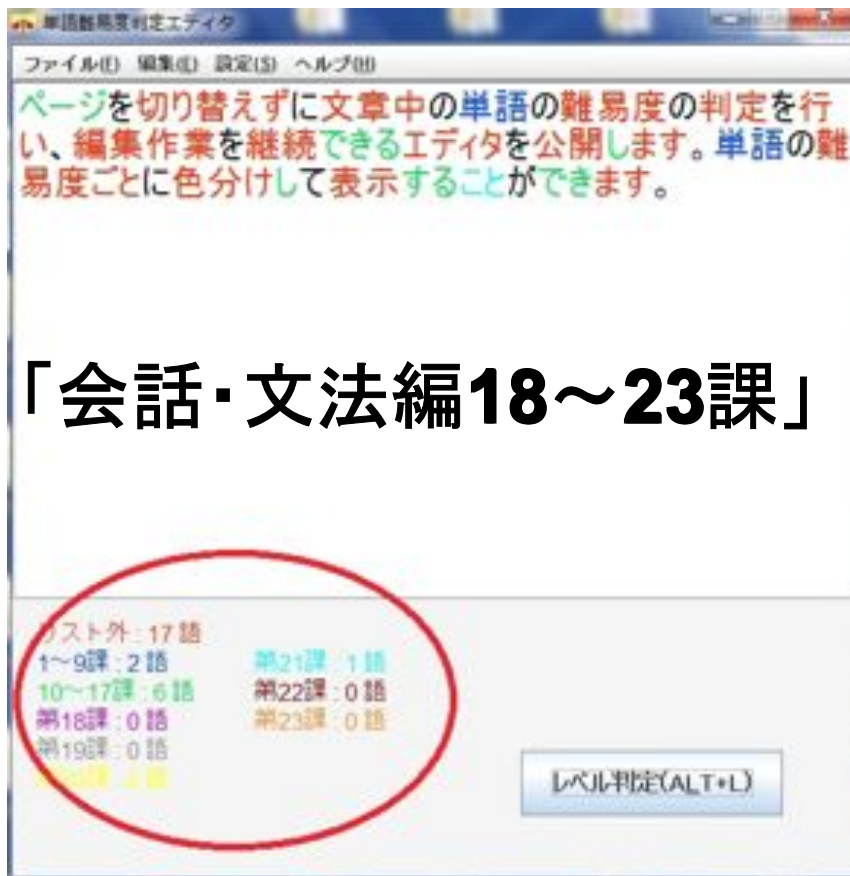
The screenshot shows a spreadsheet application window titled 'Kingsoft Spreadsheets 2012'. The active window is 'Sample2.xls'. The spreadsheet has columns A through E and rows 1 through 18. The data is as follows:

	A	B	C	D	E
1	あ	1級			
2	亜	1級			
3	あいそう	1級			
4	妄想	1級			
5	あいだから	1級			
6	間柄	1級			
7	あいつ	1級			
8	相次ぐ	1級			
9	あいま	1級			
10	合間	1級			
11	あいまつ	1級			
12	相俟つ	1級			
13	相まつ	1級			
14	あいまって	1級			
15	相俟って	1級			
16	相まって	1級			
17	あえて	1級			
18	敢えて	1級			

単語リストの切り替え

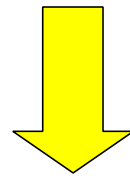


単語リストを入れ替えて判定



どんなことに便利か

- 教科書に出てきた単語をリストにすれば、すでに学習した単語かまだ学習していない単語かどうか分かる。
- まだ学習していないと判定された単語を書き直せば文章をやさしくすることができる。



教材や課題・テスト作成に便利！
やさしい日本語を作成するのに便利！

単語リストの共有

- 単語リストは自作できる。
- 作成した単語リストをお互いに共有しませんか？



まとめ(1/2)

- かぶとエディタは文章を単語のレベルごとに色分けすることができる。
- 「げんき」シリーズの単語リストを使うことができる。
- 利用者が単語リストを作ることもできる。
- 授業の準備や、やさしい日本語文章作成に役立つ。

まとめ(2/2)

- 作成した単語リストをお互いに共有し、日本語教育分野を盛り上げていきましょう！

げんきオンラインのデータの使用許可を下さったThe Japan Times社に深謝いたします。本研究の一部は平成26年度科学研究費(24320096)及び私立大学等経常費補助金特別補助「大学間連携等による共同研究」の支援を得て行われた。

かぶとエディタを配布します

- かぶとエディタのCDを持ってきています。
- ほしい人は私に言ってください。
- 「げんき」シリーズの単語リストも入っています。
- USBメモリからかぶとエディタをコピーすることもできます。



無料！