

英単語選択システムの開発と応用

八鳥 吉明* 大野 佳代子**

(2014年11月27日受理)

1. 背景

英語を習得するためにはさまざまな学習活動が必要となるが、中でも語彙学習は必須である。筆者は、高等専門学校（以下「高専」と略称する）における英語の授業を通じた語彙習得のための効果的方法を求め、学生が意識的に語彙力を強化することができる学習環境の研究を行ってきた。

もともと外国語習得のためには語彙の学習に「計画的に時間をかけることが絶対に必要」（千野 64）であるが、現在進めている研究には、この基本的認識に加えて二つの契機が存在している。

一つは高専の学生を取り巻く環境の変化である。大学へ編入したり、さらには大学院に進学したりする学生が増えるにつれ、受け入れ側の大学や大学院は高専生により高い英語力を求めるようになっており、入学試験、あるいは近年ではTOEICやTOEFLのスコアによって、英語力を検査や評価の対象とすることが多い。また、就職後も、経済のグローバル化に伴って業務で英語が必要になる機会が増えるにつれ、英語で書かれた文書の読み取り、海外の取引先との英語による口頭あるいはメールでのやり取り、さらには海外出張や海外勤務での英語によるコミュニケーションなど、技術者に期待される英語力も高まっている。

もう一つの契機は、高専4年次以降の語彙学習環境である。学生が進学や就職に向けた学習準備を意識するこの時期、授業での使用テキストは検定教科書から授業毎に異なる教科書へと多様化する。英語学習の実践的側面が強化される反面、語彙習得という観点からは、学習は体系的というよりもむしろ拡散的になる傾向がある。

こうした状況への対策として、学生が授業を通して意識的・計画的に単語を学習することができる環境の整備が必要になると思われた。そこでまずは語彙学習の目標を設定することにし、覚えるべき単語数におおまかな制限を加えた。当面の目標は高校までの学習範囲に相当する約2,500語から3,000語とし、選別の基準は使用頻度数の高さ（千野 55-56；白井『英語』163-64）とした。

実際の選別にあたっては、この目標や基準とほぼ同じ観点に沿って単語がデータ化されている既刊の数種類の英単語本を試行的に参照することにした。

このようにして授業で使用するテキストの中から重要単語を選択し、授業で解説を行った。さらに授業の後には、選択した単語を重要単語リスト（図-1 参照）としてまとめ、紙媒体並びにWeb上で情報を提供するシステムを構築した（八鳥・大野）。単語は文脈の中で覚えることが効果的であるとされているが（白井『外国語』172-73；白井『英語』80-81, 161-62）、今回の方法は授業教材と語彙学習を密接に関係付けることで、単語を文脈の中で理解することを促すものとなっている。

No.8	[3]	
[1]	9 meditation	瞑想
1 necessities	10 physical	肉体[身体]の → No.10 [3]10
2 duty (to do)	11 observe	を観察する
3 provide A with B	12 suffer from A	Aを患う
4 citizen	13 gain	を得る
5 available (to A)	14 patient	患者
6 vital	15 extend	を伸ばす、拡大する
7 responsibility (to do)	16 recommend	…すべきたと勧める (that-節)
8 welfare	17 practice	練習する
[2]	[1]	
9 deny	1 fold	を折りたむ
10 physical	2 theory	学説、理論
11 environment		
12 environment		

『SUPREME スプリーム英語構文 109 活用ワーク英文解釈演習』に基づいて作成

図-1 学生に提供している重要単語リスト

目標と基準に沿った単語の選択、単語と文脈の接続、授業での指導、Web環境の整備——これらの試みによって、語彙学習をより効率的なものにし、学生が意識的に語彙学習を継続する環境をある程度整備することができたように思われる。

しかし、この方法では授業準備に時間を要することが問題となっていた。重要単語を選び出すためには、複数の英単語本を参照しながら表を作成する作業が必要になるが、それにはどうしても時間がかかるからである。そこで今回、この問題に対処するため、英単語本における単語に関する情報を含む表を自動で出力できる「英単語選択システム」を開発し、それを運用することで作業

時間の削減を実現することが可能となったので、以下にその報告をする次第である。

2. 従来手法と問題点

従来、重要単語を選択するためには、表-1 に示す6冊の既刊の英単語本を試行的に参照し、その内容を比較・検討しながら表の作成を行ってきた。その際、同時にTOEICや科学・工業英語などで使用される語彙も考慮に入れ、6冊の内容を補うようにした。

表-1 6冊の英単語本の情報一覧

No.	1	2	3	4	5	6
書籍名	ドラゴン・イングリッシュ必修英単語 1000	英単語ターゲット 1900 5訂版	DUO3.0	システム英単語改定新版	速読英単語(1)必修編改訂第5版	COCET 2600
出版社	講談社	旺文社	アイシーピー	駿台文庫	Z会	成美堂
区分	○:見出し語 △:解説	○:見出し語 △:派生・関連	○:見出し語 △:派生・関連	○:見出し語 △:上記以外	○:通し番号 △:上記以外	○:見出し番号 △:派生・関連
Index	16ページ	26ページ	49ページ	27ページ	24ページ	20ページ
単語数	2,377	3,473	8,231	3,712	3,250	3,135
内訳○	1,298	1,900	2,940	2,407	1,923	2,600
△	1,079	1,573	2,217	1,305	1,327	535
その他	0	0	3,074	0	0	0

単語リストの提供そのものは2011年度から開始していたが、複数の英単語本のデータをもとにした重要単語リストの提供を開始したのは2013年度(表-2 参照)である。その結果、2011年から2013年までの間に、合計1,822単語を重要単語として選択してきた。

表-2 重要単語リストの提供履歴

年度		2010	2011	2012	2013	2014
項目						
担当学年						
本科 1, 2年	紙媒体	1年	2年			→ 予定
本科 4年	ホームページ					
専攻科 1年	紙媒体					
専攻科 1年	ホームページ					
重要単語リストの提供						
本科 2年 4年	紙媒体		△	△	△	△
本科 2年 4年	ホームページ				△	△
専攻科 1年	紙媒体				△	△
専攻科 1年	ホームページ				△	△
ホームページ開発 Web提供						
					△	△

※) 筆者判断+『ドラゴン・イングリッシュ』参照

ここで問題となったのは、重要単語を選択するためにかかる作業時間である。テキストの中から、まず重要単語の候補を選び、それらを6冊の英単語本で確認しながら結果を表にまとめ、作成したこの表の情報をもとに最終的に重要単語を確定していく。この調査には、どうしても相当程度の時間が必要であった。

3. 英単語選択システムの開発

上記の問題を解決するために、作業効率を上げることを目的とし、6冊の英単語本における単語の掲載状況を自動で作表できる「英単語選択システム」を開発した。

3.1 アプリケーションの選定

システムの開発にあたり、プログラム開発言語とアプリケーションの選定を行った。本システムは、英単語本のインデックス(索引)の情報をデータベース化して格納した情報の中から、調査対象の英単語と一致するデータを出力するものである。データベース開発ツールとしては、OracleシリーズやMySQLなどが有名である。今回の英単語選択システムの実行は、ユーザ自身が行うことから、日常業務で使用しているアプリケーションで、プログラムの組み込みが可能な、Microsoft Office 2010のExcelとし、プログラミング言語はExcelに標準装備されているVBA(Visual Basic for Applications)とした。ExcelのVBAでシステム開発を行うためには、VBAそのものの知識と、プログラミングのスキル、そしてシステムを作り上げるノウハウが必要となる(立山 14)。

3.2 開発日程

表-3 に、本システムの開発日程を示す。2014年度の教材から適用を開始することを目標とし、2013年11月から仕様検討を始め、2013年12月にPhase 1として、1冊の英単語本を対象とした英単語選択システムを作成した。その後レビューを行い、問題点を検証して制御プログラムに改良を加えた。更に、残り4冊の英単語本をデータベース化して、2014年3月22日にPhase 2を完成させた。

表-3 開発日程

No.	工程	開始日	終了日	工数
1	仕様検討・作成	2013/11/30(土)	2013/11/30(土)	0.5人日
2	Phase1: 開発	2013/12/01(日)	2013/12/08(日)	3人日
3	Phase1: レビュー	2013/12/08(日)	2013/12/08(日)	1hr.
4	Phase2: 開発	2013/12/14(土)	2014/03/16(日)	5.5人日
5	Phase2: レビュー	2014/03/22(土)	2014/03/22(土)	1hr.
合計				9人日

3. 3 システム概要

図-2 にシステム構成図を示す。

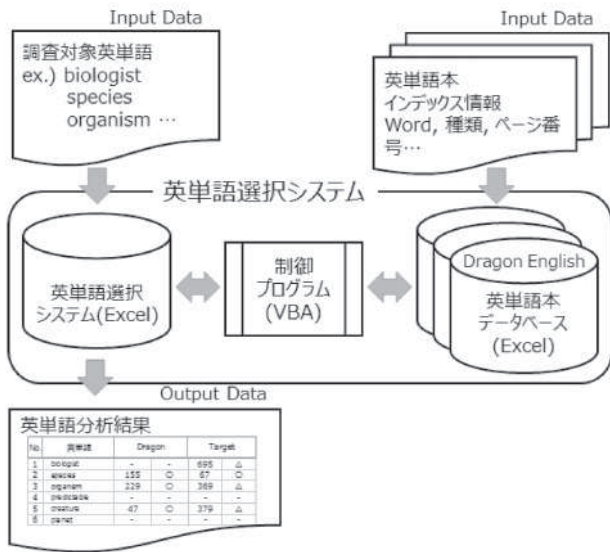


図-2 英単語選択システム構成図

本システムの動作環境としては、アプリケーションはMicrosoft社のOffice Excel 2010以上、オペレーティングシステム(OS)はWindows 7以上である。入力データは、調査対象の英単語である。

制御プログラムの実行は、Excelのファイル上で行う。Excelのシートに調査対象となる英単語を入力し、制御プログラムを実行すると、英単語本データベースの中から一致する情報を検索し、出力データである一覧表を作成する。

英単語本データベースは、市販英単語本の改訂版の発行に伴う差替えや、調査対象とする単語本の追加・削除などが簡単に行えるように、英単語本1冊につき、Excelデータベースを1ファイル作成する。

機能性に優れたデータベースを作成するためには、「利用者にとって使いやすい」という点を常に念頭に置き、利用者の心理的な側面についても十分に配慮した手順を考える必要がある(石桁、他 148)。そのため、英単語本データベースには、単語とその区分だけではなく、単語が掲載されているページ番号や通し番号の情報も登録しておき、出力する項目に追加する。

さらに、出力された結果を見やすくするために、Excel自体が持つ条件付書式である指定データ強調表示機能や、指定文字をカウントする関数であるCOUNTIFを活用する。条件付書式を活用することにより、データの変化を視覚的に表現できるので、出力結果である表が読みやすくなる(小舘、他 163)。この条件付書式とカウント関数は、事前にExcelシートに設定しておくことができ、VBAプログラムには組み込む必要が無いため、制御プログラムの処理時間に影響することはない。

3. 4 処理アルゴリズム

システムをExcelで開発するため、VBAプログラムの実行処理に時間がかかることが懸念される。プログラムの処理速度は、データ構造と処理アルゴリズムの仕様によって決まる。

本システムの処理の大部分は、データベースに格納された文字列の検索である。検索処理は、プログラミングにおいて、頻繁に現れる操作の1つであり、これまでに多くの異なるアルゴリズムが開発されてきた(ヴィルト 53)。ただし、検索に複雑な処理を施せば、その分、処理時間は長くなる。そのため、調査対象の単語と、データベースに格納された単語の比較は、対象文字が「含まれる」かどうかのみを判断基準とすることとした。これにより、プログラムの処理スピードの劣化を防ぐことができる。しかし、この方法では、完全一致ではない単語を誤って出力する可能性がある——例えば、調査対象単語の“portion”に対する“proportion”の出力。このため、調査対象単語と完全一致したデータベース内の単語の区分と番号を一覧表に出力し、完全一致ではない場合は、その旨の表示を出力する。

また、ブランク文字やタブ文字など、目に見えない文字も、検索単語と不一致の結果をもたらす。その対策としては、検索対象の単語の前後の空白を除く関数であるTrimを使用する。

検索単語を含む候補が複数存在する場合は、それぞれの単語本に対してあらかじめ作成されているシートに、その詳細を出力する。その中で完全一致した単語には、一目で判断ができるように、先頭に星印を付ける。完全に一致するものがなく、同じ文字列を含む単語のみが存在する場合は、結果出力一覧表に、その個数を出力する。

一致する単語が存在しなかった場合は、空欄ではなく、ハイフンを表示する。この記号を出力することで、処理が正しく行われたことを示す。

最後に、ユーザの操作や運用による、実行時エラー発生時に対処する機能を追加する。これは、制作者が意図していない操作をユーザが行う可能性があることを念頭に、あらゆる操作パターンを想定してエラー対策を組み込むことが必要である(立山 377)という認識に基づいている。本システムでは、調査対象の単語の入力がない場合や、プログラムの処理範囲を超える個数のデータが入力された時などは、エラーメッセージを表示してプログラムを終了させる処理機能を採用した。

また、プログラムが自動で判断できない部分は、次のように手動で事前処理を行う必要がある。例えば、英単語本のインデックス(索引)に記載されている“analytic(al)”のような省略記述の指示内容を、人間は“analytic”または“analytical”であると判断できる。しかし、

コンピュータは、この“analytic(al)”という文字列すべてを1つの単語として判断してしまう。そこで、英単語本データベースを作成する際には、その記述を省略せず、区切り文字として空白を入れて“analytic (analytical)”と入力し直すことにした。これにより、人間と同じ判断がプログラム上で実現可能となる。

以上の処理アルゴリズムを組み込んで完成した制御プログラムのソースコードは、コメント行と空行を含め、220ステップであった。プログラムとしては、制御部分を極小規模で実現できたと言える。図-3 に制御プログラムの開発画面を示す。

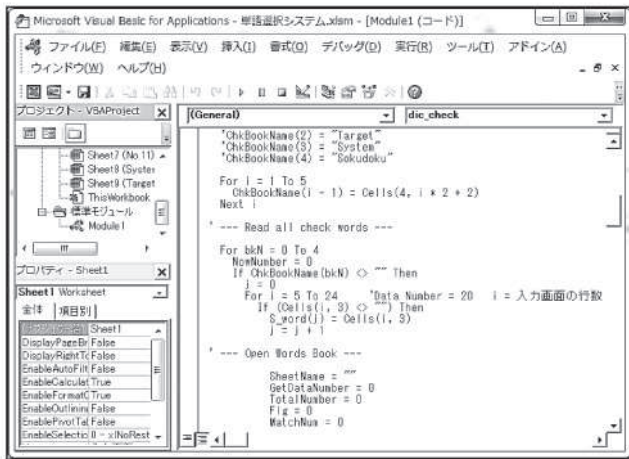


図-3 制御プログラム開発画面

3. 5 英単語本データベース

英単語本データベースの作成方法を図-4 に示す。

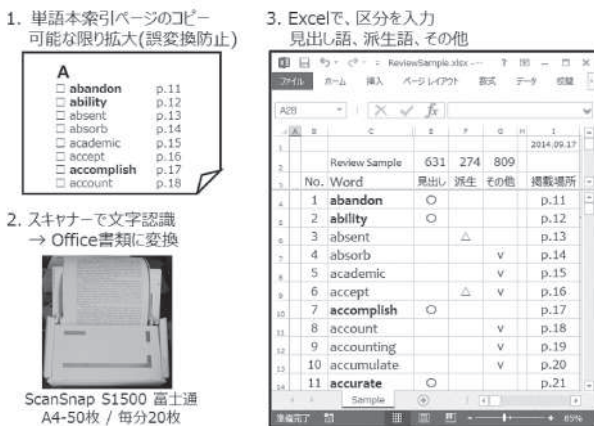


図-4 英単語本データベース作成方法

データベースの作成にあたっては、人の手動操作によるデータ入力方式は、速度が遅く、誤りも多いため、できる限り自動化することが望ましい（石柝、他 12）とされる。そのため、まず初めに、英単語本のインデックス（索引）ページのコピーを取る。この際、可能な限り拡大してコピーをする。拡大することで、次の工程で

ある、スキャナーに文字を自動認識させる処理をより円滑に行うことができる。使用スキャナーは、オートフィーダー付きで、文字認識アプリケーション対応のものを活用した。拡大コピーをしないで文字認識をさせた場合、文字の誤認識率は20%であった。しかし、拡大コピーをした後に文字認識をさせた場合では、誤認識される確率はほぼなくなった。

自動認識された単語はExcelフォーマットに変換し、その後、区分情報である見出し語には丸印、派生語・関連語などには三角印を入力してデータベースを完成させる。

3. 6 実行方法と出力ファイル

本システムの実行例を図-5 に示す。Excelの「入力画面」シートのセル内に、調査対象の単語を入力する。マクロボタン「単語本サーチ」を押すと、プログラムが実行され、英単語本6冊分の検索を行い、掲載されているページ番号や区分情報などが実行シート上に一覧表で出力され、最後にプログラムの処理が正常終了したことを示すメッセージボックスが表示される。

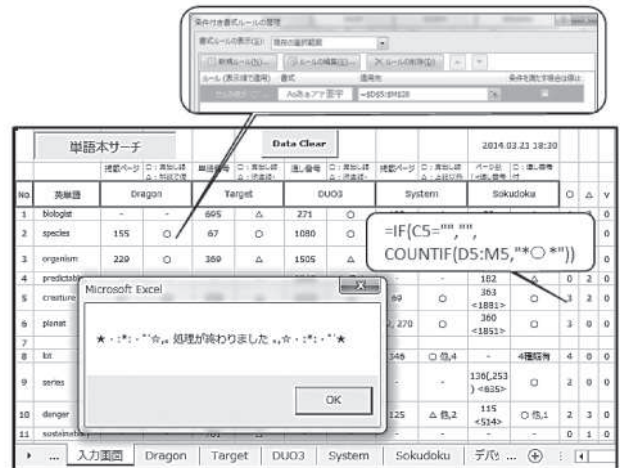


図-5 システム実行例

Excelの結果出力部分のセルには、見出し語を示す丸印が存在するセルを赤文字に変更する条件付書式機能と、列内の丸印を自動でカウントするCOUNTIF関数が組み込まれているため、ユーザは、出力された表から結果を視覚的に、かつ素早く判断することが可能である。

4. 適用結果

開発した本システムを適用し、自動で出力された表の情報を参考にして重要単語を選択した結果、2014年4月から本科4年と専攻科1年の各回の授業範囲に対して、授業1回分に2時間程度かかっていた作業時間を、約30分に短縮することができた。

2014年9月時点で、該当年度の半年間に本システムを利用して検索した単語数は、2学年合わせて1,072単語であり、その中から重要単語として選択した数は、399単語であった。

5. 今後の課題

今後は、筆者が2011年度から学生に提供している重要単語リストをもとに、独自のデータベースの構築を行っていく予定である。さらに、データベースに基づく英単語学習アプリケーションの開発に着手することを目指している。こうした試みを実践することを通して、学生にとってより効果的な英語学習が可能となる、新たな学習環境の整備を進めていきたい。

参考文献

- CHART INSTITUTE. 『SUPREME スプリーム英語構文 109——活用ワーク 英文解釈演習』. 改訂版. 東京：数研出版. 2011.
- 石桁正士、磯本征雄、江澤義典. 『データベース入門』. 東京：パワー社. 1984.
- ヴィルト, N. 『アルゴリズムとデータ構造』. 浦昭二、國府方久史訳. 東京：近代科学社. 1990.
- 風見寛. 『速読英単語（1）必修編』. 改訂第5版. 東京：Z会. 2009.
- 亀山太一、青山晶子、穴井孝義、武田淳、中井大造、森和憲、森岡隆. 『COCET2600——理工系学生のための必修英単語2600』. 東京：成美堂. 2012.
- 小館由典. 『できるExcel 2010 Windows 7/Vista/XP対応』. 東京：インプレスジャパン. 2010.
- 白井恭弘. 『外国語学習の科学——第二言語習得論とは何か』. 岩波新書. 1150. 東京：岩波書店. 2008.
- . 『英語はもっと科学的に学習しよう』. 東京：KADOKAWA. 2013.
- 鈴木陽一. 『DUO 3.0』. 東京：アイシーピー. 2000.
- 竹岡広信. 『ドラゴン・イングリッシュ——必修英単語1000』. 東京：講談社. 2008.
- 立山秀利. 『続 Excel VBAのプログラミングのツボとコツがゼッタイにわかる本』. 東京：秀和システム. 2009.
- 千野栄一. 『外国語上達法』. 岩波新書. 329. 東京：岩波書店. 1986.
- 刀裯雅彦、霜康司. 『システム英単語 Ver.2』. 東京：駿台文庫. 2005.
- 八鳥吉明、大野佳代子. 「英単語学習用 Web 教材開発——効果的な英語学習法を求めて」. 『群馬高専レビュー』. 32 (2013) : 61-65.
- 宮川幸久. 『英単語ターゲット1900』. 5訂版. 東京：旺文社. 2011.

Development and Application of a Computerized System for Selecting English Words

Yoshiaki HACHITORI and Kayoko OHNO

This paper reports an attempt to introduce a computerized system which makes it possible to select important English words in an efficient way. Vocabulary building is indisputably important in developing English skills, so we have adopted a web-based English vocabulary learning system that we use in conjunction with regular classes. To maintain this system, it is necessary to select target vocabulary words from textbooks while also referring to several wordbooks. Although this process serves an important purpose, it takes a certain amount of time each time, which inevitably constitutes a problem. As a means of addressing this issue, we have developed a computerized system for selecting important English words based on the data of multiple wordbooks. This new system has made it possible to reduce the time it takes to choose target words.