

WWW上における英単語ハイパー辞書の作成及び公開について

大出真† 下島崇† 蔡東生† 池辺八洲彦‡

† 筑波大学 ‡ 会津大学

我々は、「語学教育の改革は辞書の改革なくして有り得ない」という立場から、一般の英単語学習者に役立つ英単語ハイパー辞書データベースの研究を推進している。最近の研究では、辞書に導入すべき情報の一つとして、印欧語根を用いた英単語分類法について着目してきた。印欧語根とは比較言語学によって構成された仮想の言語である。本稿では、印欧語根を利用した英単語学習方法とその調査状況について説明する。そして、調査分類した印欧語根情報を基にした英単語ハイパー辞書データベースをWWW上で検索できるよう作成したシステムについて報告する。

Database System of Hyper English Dictionary on World Wide Web

Makoto Ohde†, Takashi Shimojima†, DongSheng Cai†, Yasuhiko Ikebe†

† University of Tsukuba
1-1 Tennodai, Tsukuba-shi, Ibaraki 305 Japan

‡ University of Aizu
Ikki-cho, Aizuwakamatsu-shi, Fukushima 965 Japan

We are taking up the position that there's no reformation of linguistic education without the reformation of dictionaries. We propel the research of database of Hyper English Dictionary for learning English words. Now we notice the concept of Indo-European roots for classifying English words. That is a virtual language and was organized by comparative linguistics. In this paper we explain the new way of learning English words which were classified by the concept of Indo-European roots. Also report the trial of producing database system of Hyper English Dictionary on World Wide Web.

1. はじめに

語学学習における辞書の重要性については改めて本稿にて強調する必要はないであろう。長い間辞書は、紙という記録媒体に縛り付けられてきた。線形構造しか扱えない紙による情報の記録では、

- 扱える情報の量や種類が制限される
- さまざまな構造を持った情報が扱えない

といった欠点があった。

しかし近年の計算機分野の急激な発展により、高速な計算機と CDROM などに代表される大容量の記録媒体を利用して辞書をデータベース化することで、それらの欠点を克服した新しい辞書を開発することが可能になっている。

我々の研究室では、「語学教育の改革は辞書の改革なくしては有り得ない」という立場から、一般の語学学習者に役立つ、計算機の特性を生かした新たな英単語辞書の形を模索し研究を進めてきた。その研究範囲は、辞書が持つべき情報の検証分析から、表示インタフェースの作成、公開方法の検討まで多岐にわたっている。

本稿では、これまで我々が注目してきた印欧語根について簡単に説明し、辞書の要素として印欧語根を導入することで可能となる英単語学習方法について簡単に紹介する。また今回作成した英単語ハイパー辞書データベースを、インターネットを利用してWWW上で検索できるシステムについても説明する。

2. 印欧語根情報の有用性

2. 1. 印欧祖語と印欧語根について

印欧祖語 (Proto-Indo-European) は、それ自体現実の言語として存在したものではなく、19 世紀からヨーロッパにおいて急激に発展した比較言語学によって、印欧諸語 (Indo-Europeans) に見られる規則的な音韻変化をた

どるなど、理論的な再構成によって得られた仮定の言語である。そして印欧祖語の個々の言葉の語幹を幾らかのパターンにまとめたのが印欧語根 (Indo-European Roots) である。現存する言語がどのように印欧祖語から派生したかを図1に示す。ただし現実には、このような単純な派生の他に、各段階において借入語などが存在し、遥かに複雑な関係となる。

AHD 第三版[1]において言語学者 Calvert Watkins は、印欧語根を 596 種に分類しており、我々はこの印欧語根の分類をもとにして、英単語がどの印欧語根に由来するかを調査している。

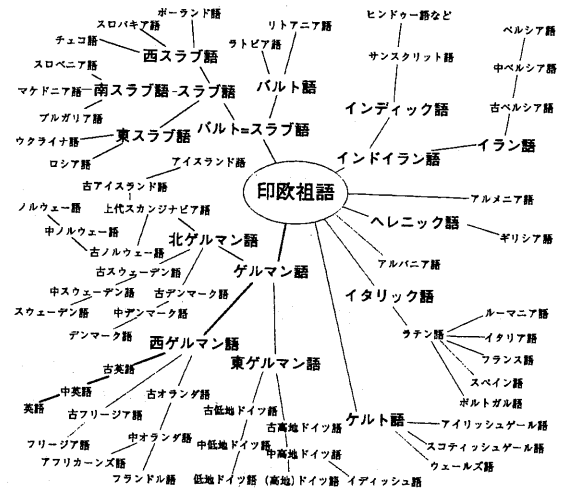


図 1 印欧祖語から派生した諸言語

2. 2. 印欧語根を用いた学習法

我々は、辞書データベースの要素として印欧語根を導入することにより、

- 同一の語根に由来する単語 (同根語) をまとめて学ぶ。
- 同義語の微妙な差異をそれぞれが由来する印欧語根から理解する。

といった新たな学習方法を容易におこなえることができると考えている。ここではそれぞれの

学習方法とその有効性について、簡単に具体例を挙げながら説明を行う。

2. 2. 1 同根語をまとめて学ぶ

印欧語根[teua-]は「ふくれる」という概念をあらわす。同根語としては、butter (バター)、thousand(千)、thumb(親指)、tomb (墓)、tumor (腫瘍) などがある。

thousand は hundred がふくらんだものと考えられる。ここでは詳しく説明しないが、他の各単語も「ふくれる」といった概念を語義に含んでいることが認められる。

このように同一の語根に由来する単語をまとめて学ぶことにより、単語間の関係やその基本的な語義概念を容易にイメージすることができ、個別に単語を学ぶより、広く深い理解が得られると考えている。

2. 2. 2 類義語を学習する

choose と select は、いずれも日本語では「選ぶ」の意味に当たるが、英語では微妙なニュアンスの差が生じる。この差異を印欧語根から推察してみる。

choose は印欧語根[geus-]に由来している。一方 select は印欧語根[s(w)e-]と[leg-]の二つに由来している。これらの印欧語根について、その簡単な概念と、同じ印欧語根に由来する英単語をいくつか挙げてみる。

印欧語根	概念	同根語
[geus-]	To taste, choose	Disgust
[s(w)e-]	Further appearing in various forms referring to social group as an entity	Self, secret, separate
[leg-]	to collect	-logy, collect

表 1 語根の概念と同根語

このような調査結果から、choose はより広い意味での「選ぶ」という概念を持ち、それに対して select は比較的排他的なニュアンスを

持っていることがうかがえる。勿論印欧語根情報だけで類義語の差異がわかるものばかりではなく、例文などの補助的な情報も必要となる場合も多いが、上に挙げた例のように、印欧語根の違いからみた類義語の使い分けの説明は、類義語の微妙な差異を理解する上で役に立つ一つの要素であるといえよう。

2. 3. 印欧語根の調査状況

印欧語根を利用した学習方法とその有効性については先に述べた通りであるが、では実際英単語について、どのくらいの割合で印欧語根が判明するのであろうか。

我々は、AHD[1]を利用し、英単語に対して由来する印欧語根を調べ、その結果をデータベース化する作業を行ってきている。調査対象は、受験生などによく使われている英単語学習参考書などにおいて、見出し語扱いで収録されている英単語であり、現在総語数 9054 個となっている。単語集別の判明率は以下のとおりである。

単語集の出典	単語数	判明率 (%)
資料[1]の印欧語根概説より important Derivatives	3959	100
VIS-ED No.1 (Visual Education assn. 刊)	1000	74.0
VIS-ED No.2 (Visual Education assn. 刊)	998	72.0
How to prepare for the GRE (Barron's Educational Series, Inc 刊)	3737	73.1
How to prepare for the TOEFL (Barron's Educational Series, Inc 刊)	634	69.1
英単語頻出案内 (桐原書店 刊)	1526	85.1
英単語連想記憶術 第1集～第3集 (青春出版社 刊)	1706	77.4
英検準1級参考書 (旺文社 刊)	557	76.3
全体	9054	80.4

表 2 印欧語根判明率

全体の判明率は約 80%であり、この調査結果により印欧語根による分類に則った学習法が実用に耐えるものであるといえると考えてい

る。また印欧語根による分類を補完する意味で、接頭辞 (prefix)、語幹 (stem) といった要素についても調査しデータベース化を行っている。

3. www 上での公開

3. 1. 目的

一連の研究での最終的な目標は、幅広い範囲の人々に利用される媒体で、完成した辞書データベースを公開することである。そのための方法として現在もっとも適していると思われるのは、インターネットを利用した WWW 上での公開である。

現在データとして存在するのは、各単語とそれが由来する印欧語根、語幹、接頭辞についての情報だけである。これだけでは辞書としての情報は不十分である。

しかし公開方法、ユーザインタフェースの研究もなるし、これまで調査分類をしてきた印欧語根をはじめとする情報によって生じる学習効果について実証するためにも、英単語ハイパー辞書データベースとして公開することは十分に意義のあることだと考え、今回試験的に公開する事とした。

実装は以下の環境で行った。

計算機	HP Vectra XU
OS	Windows NT Workstation 4.0
DBMS	Access97
コンパイラ	Visual Basic 5.0
WWW サーバー	Peer Web Services 2.0
ブラウザ	Internet Explorer 4.0

表 2 実装環境

CGI (Common Gateway Interface) プログラムを Visual Basic を用いて作成することにより、比較的容易に WWW ブラウザを用いてデータベースを動的に検索するシステムを構築することが可能になる。

3. 2. データベースの作成

これまでの印欧語根などの調査結果は、Microsoft Access を用いて入力管理してきた。今回はこれを利用して簡単な英単語辞書データベースを作成した。

これまで作成してきたデータベーステーブルは以下の 4 つである。

- 単語マスター [見出し識別子、綴り、コメント、語根 1、語根 2、語根 3、語根 4、接頭辞、語幹]
- 語根 [語根識別子、表記、説明]
- 接頭辞 [接頭辞識別子、表記、語根、説明]
- 語幹 [語幹識別子、語根 1、語根 2、説明]

今回は、印欧語根の学習方法の一つとして挙げた類義語学習の効果をより持たせるため、登録された単語の中で類義語を持つ単語同士に「類義語」という関連を持たせ、「類義語」データベーステーブルを新たに作成した。

- 類義語 [類義語識別子、共通意味、候補 1、候補 2、候補 3、候補 4、候補 5]

また、辞書データベースとしてある程度の実用性を持たせるため、今回便宜的に「単語マスター」テーブルに「品詞」と「意味」の属性を追加した。これにより、印欧語根を用いた学習についての効果を一応計る事ができるようになった。

3. 3. 機能

ユーザは WWW ブラウザ上で以下の検索を行うことができる。(図 2)

- 単語を入力し、単語の由来する印欧語根などその単語に関する情報を引くことができる。
- 印欧語根を入力し、その印欧語根に関する説明と、同根語のリストを得ることができる。

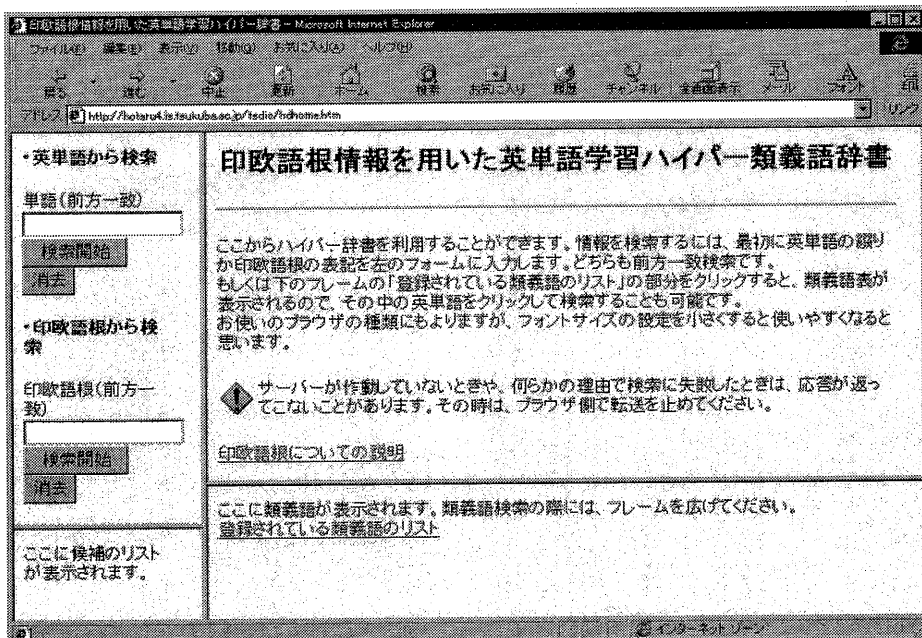


図2 スタートページ

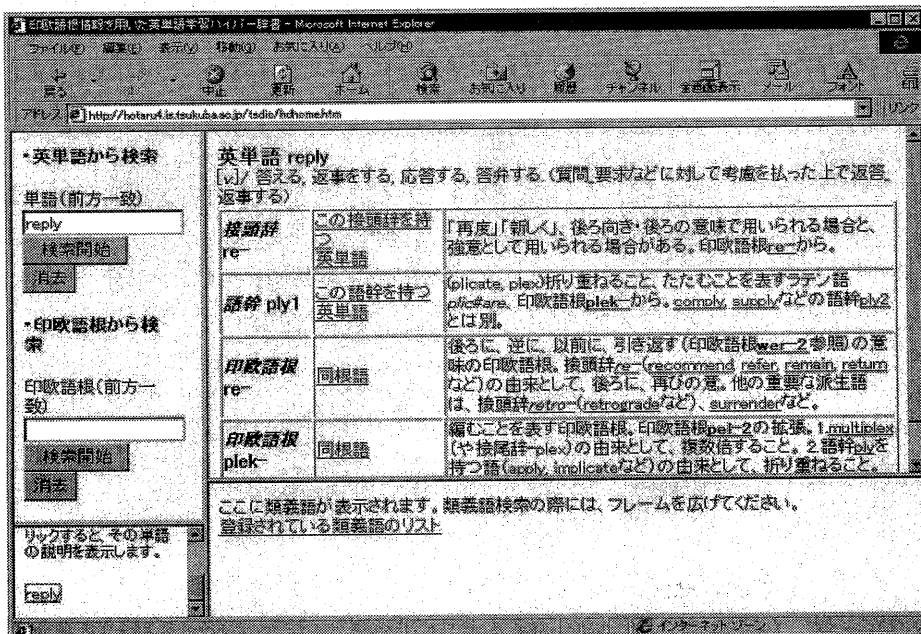


図3 英単語を検索した例

- 類義語リストに登録されている英単語を表示させ、そこから英単語をひく事ができる。

表示された結果は CGI プログラムによって作成されたリンクを含んだ HTML 文書である。

(図3) 従ってその文書においてリンクされている情報をクリックすることにより、さらにそれぞれの検索を行うことができる。

今回作成したハイパー英単語辞書は、
<http://hotaru4.is.tsukuba.ac.jp/tsdic/hdhome.htm>
にて試験的に公開している。

3. 4. 評価と問題点

WWWブラウザとCGIプログラムを組みあわせる方法により、比較的容易にハイパーテキストの特徴であるリンクを生かしたハイパー辞書データベースを検索するシステムが構築できることがわかった。

英単語辞書として最低必要な要素は、綴り、発音、語義、の3つであるといわれている。また実際の辞書には用途に応じて、派生語、成句、用例、語源などの要素が含まれている。残念ながら現在の辞書データベースでは、多くの要素が欠けており、実際の英単語辞書として使うことは難しい。それらのデータをそろえる事は今後の課題の一つとなる。

語義、用例に関しては、本研究とは別に歴史的原理に基づいて編纂された OED (Oxford English Dictionary) を利用したデータ構造の分析がおこなわれており、その成果をふまえた上で統合した概念スキーマ設計を新たに行う予定である。

4. まとめと今後の展開

本稿では、印欧語根情報を辞書の要素に加えることで生じる新たな学習法の可能性を示した。また将来の本格的な辞書データベースシステム

の公開に向けて、今回収集した印欧語根を中心とした辞書データベースをWWW上で検索できるシステムの構築を行った。

今後の展開としては、これまで調査結果をふまえ、実用的なレベルの本格的なデータベースの概念スキーマ設計を行う予定である。

またより高度なユーザインタフェースを実現するために、Java を用いた辞書データベースシステムの実装についても検討中である。

参考文献

- [1] The American Heritage Dictionary of The English Language, Third Edition : Houghton Mifflin Company, 1992.
- [2] 英語類語辞典 : 小稲 義男 他 編, 研究社, 1987.
- [3] 「印欧語根を用いた英単語ハイパー類義語辞書に関する研究」 : 下島 崇 著, 平成 9 年度 筑波大学情報学類学士論文, 1998.
- [4] 「OED を用いた前置詞 WITH のケーススタディ」 : 大出 真 著, 平成 8 年度 筑波大学情報学類学士論文, 1997.
- [5] 「印欧語根情報を用いた英単語ハイパー辞書に関する研究」 : 吉村 信吾 著, 平成 8 年度 筑波大学大学院理工学研究科修士論文, 1997.
- [6] 英語の辞書と語源 : 今里 智晃, 土家 典生 著, 大修館書店, 1984.