

正規表現を使った英語コーパス検索システムの開発
と帰納的言語教育の可能性
中島浩二

Development of a Corpus-based Search System using *Regular
Expression*: An Inductive Approach
NAKASHIMA Kohji

Abstract

In English language classrooms in Japan we have experimented with a wide range of teaching methods over the decades, such as the Grammar-Translation Method, the Direct Method, the Oral Approach and the Communicative Approach. However, this has not lead to the discovery of a superior teaching method. These approaches, despite their variety, share the common theme of being deductive top-down approaches. It is relatively straightforward to teach English grammar and collocations to students because the answers are fixed and absolute. We all know that this methodology is unnatural and completely different from the first language acquisition process. Although I recognize that the traditional deductive and top-down approach is very efficient in teaching a second language (L2) in environments which are not conducive to natural second language acquisition, I think that we should make an attempt to fuse the less efficient but more natural inductive and bottom-up approach with the traditional approach in order to develop incentives for the students to learn the L2.

I have developed a CGI-based English search system named SKWICS with an easy-to-use interface and *regular expression* search function, which will help teachers to exploit both inductive and bottom-up English teaching approaches as a study aid. SKWICS may also serve to free students from the typical drudgery inherent in rote memorization.

1. 演繹的学習と帰納的学習

語学教育および語学学習の方法論について、これまで様々な理論が提唱され、また、様々な実践が試みられてきた。その方法論も時代によって流行り廃りがあり、一概にこの指導法・学習法で決まりというような決定打がある訳ではない。良くも悪くも日本の学校教育で長い間主導的な指導法であった Grammar-Translation Method (文法訳読法)、target language を用いて指導する Direct Method (直接教授法)、アメリカ構造主義と行動主義心理学を理論的根拠に C. C. Fries によって開発された Pattern Practice が特徴的な Oral Approach (OA)、さらに、英語を国際コミュニケーションの媒体として用い、国際社会で通用する人材を養成しようとして近年好んで用いられている Communicative Approach (CA) と、主要なものだけでいくつもの教授法が存在する。勿論より良い方法を求めてこういった様々な指導法が開発されてきたわけだが、新たに開発された方法論にも必ずと言って長所短所があり、万能でないがゆえにまた新たな方法論が開発され採用されていくのである。つまり、大まかに言えば「理論と実践」といったキーワードでとらえることが可能な「ゆれ」が語学教育、とりわけ英語教育においては常に存在してきたと言えるのではないか。時代とともに、流行とともに、振り子がどちらかに大きく振られるということはあっても、必ず揺り戻しがくるのであって、そのことがどちらも語学学習には必要であるということを図らずも証明しているのではないかと思う。一つの方法論が万能ではないわけだから、古い方法論だからといって無下に否定する必要はなく、また、一つの方法論を金科玉条のごとく盲信することもなく、それぞれの方法論の中の長所を活かしながら複数の指導法を組み合わせることが肝要ではないだろうか。

語学教育・語学学習の流れとか方向性といった視点でとらえると、演繹と帰納、あるいは、‘top-down’ と ‘bottom-up’ という上記とは異なったパラダイムで事象を切り取ることが可能である。これらは、学習者がどういう過程を通じて言語知識を習得し、自らの中で広義な意味での文法をどう構築していくかというベクトルを表している。つまり、演繹とか ‘top-down’ というのは、あらかじめ target language の文法が他者（通常は教師）によって提示され、それを学習者が習う形で言語知識を習得していくということで、日本の学校教育で行われている語学教育の最も一般的なプロセスである。「理論と実践」という「ゆれ」に対しては、上述したように、教える側の間で常に議論があり、試行錯誤があり、様々な方法論が提示され興味の対象であり続けてきたわけだが、言語知識の獲得・習得プロセスの方向性に関しては、語学教育はもっぱら演繹

的で top-down 的な方法論・指導法を採用してきた。つまり、はじめに文法あるいは文法理論ありきで、それを教師が何らかの指導法を用いて top-down 的に生徒・学生に教えるというやり方が一般的に行なわれてきたのだ。Oral Approach で多用される Pattern Practice や、Communicative Approach などの生徒・学生側のフィードバックを大切にしている指導法においてさえも基本的にそのベクトルは変わらない。それに対して、人が成長する過程で母語の運用能力を獲得していく場合は、周囲から文法あるいは文法理論を教えてもらって言語獲得しているわけではないことは明らかである。それは母語話者の多くがその言語の自在な運用能力を持ちながらも、その言語の文法や仕組みについてメタ言語的に説明することができない場合が大半であることから伺える。Noam Chomsky のいう Universal Grammar (UG, 普遍文法) が生得的に実在するか否かの議論は別として、母語を獲得していく過程は、Ferdinand de Saussure のいう Parole に日々接しながら自らの脳の中で言語知識を構築していくという作業であり、明らかに帰納的要素が大きく、bottom-up 的なベクトルに沿って行なわれていると言えるだろう。

このように一般的な教授法は、演繹的で top-down 的であるのに対し、母語の言語獲得は帰納的で bottom-up 的であり、その違いゆえに、日本の中学・高校・大学における現状の語学教育が不自然であり、ことばを身につける方法としては適切でないとしばしば批判を受けるのである。しかしここで考慮しなければいけないのは、中学・高校・大学における英語教育は第 2 言語 (L2) の習得であり、母語の場合は第 1 言語 (L1) の獲得であるということだ。母語の場合は 7 歳から 12 歳頃までと言われるいわゆる言語形成期／臨界期(これらの概念自体の信憑性や期間についての議論はここではしない) までに運用能力の大半を獲得するのに対し、L2 は言語形成期をほぼ過ぎてから学習によって習得される外国語であることに注意して議論しなくてはならない。母語獲得の場合は、家族や友人などと日々生活していく中で言語音や語の指し示す外延との結びつきが文字通り四六時中繰り返し実感でき、しかも、相手と双方向的コミュニケーションがとれる状況下における学習過程であるのに対して、L2 はせいぜい週にわずか数時間、主として教科書等の文字、CD 等による音声、ビデオ等の映像で人工的で断片的に触れる中での受動的な学習にすぎないのである。母語獲得と同じような帰納的体験学習的手法をこのような貧弱な言語体験の中で用いれば、言語獲得に必要なインプットを得るのはきわめて困難である。言語形成期までとそれ以後という違いがどの程度影響するかについてはもう少し科学的な検証が必要だが、少なくとも target language に触れてインプット・アウトプットす

る量には歴然とした差があるということを認めた上でL2の指導法・教授法について検討していく必要があると考える。

演繹的で top-down 的な教授法の特徴としては、その効率の良さが考えられる。様々な過去の研究成果に基づいて既に構築されている文法や言語の仕組みについて教えるわけだから、極論を言えば、最初に問題の正解を提示した後それを暗記するよう指示するようなものであり、教えられる側はなぜそうなるのかについてじっくり時間をかけて深く考える必要もなく、ただただ教えられることを丸暗記に近い形で覚えればよいということになる。このような指導法・学習法が、「英語は暗記科目」というような単純な図式化を学習者に芽生えさせてしまう決定的な要因になっているとも言えるだろう。そこには母語獲得においては避けて通れない試行錯誤的な要素は微塵もない。母語の獲得過程とは全く正反対とっていいこのようなやり方は、言語学習過程の方向性としては不自然極まりないものと言わざるを得ないだろう。それにも関わらず圧倒的にこのような top-down 的なやり方が行われ続けているかということ、それはとりもなおさずその「効率性」であり、限られた時間内で詰め込まざるを得ない非母語学習の持つ宿命的側面に理由があると言えよう。しかし、このような指導法だけを行ってはいは「英語（外国語）は暗記科目」という観念が学習者に染み付いてしまうばかりか、言葉そのものが暗記科目的なものとして誤解される危険性まではらんでしまうことになりかねない。演繹的で top-down 的な外国語指導法が効率的であることは認めつつ、ところどころで立ち止まり、非効率的で時間はかかるが、母語獲得と同様な方向性を持つという意味で自然な帰納的で bottom-up 的な語学学習の仕組みを取り入れていくべきだろう。本来はこういう過程を通して言語は習得するものなのだが、物理的諸事情によって top-down 的な手法をとらざるを得ないということを学習者に理解してもらうことは、無味乾燥な暗記科目として捉えられがちな英語・外国語学習に意味付けを行い、学習の動機付けを促すという意味で大切なことではないか。

この論文では、非効率的で時間はかかるが、自然で本来的な習得の方向性に沿っている帰納的で bottom-up 的な学習を補助するためのシステムを構築し、そのシステムをどのように英語教育に活かすことが可能であるか、またその可能性と問題点について論じていく。

2. 帰納的学習を手助けするためのコーパス検索システムの開発

母語獲得の過程と第2言語としての外国語習得の過程は、前章で述べてきたように、水と油とっていいくらいの方向性の違いがある。今後小学校の授業

で本格的に英語学習が取り入られると事情は少し変わってくるかもしれないが、一般的には12歳頃から学習し始め、母語に比べて接したり使用したりする機会が極めて少ないという時間的・物理的制約の中で第2言語を教授するには、「一方通行的」と非難されようとも、短時間で効率的に教授できる演繹的でtop-down的な従来型の手法をとることが基本となるべきだろう。しかし、そこに母語の言語獲得と同じような自然なプロセスを少しでも組み入れたいと考える。前述したように、言語体験を通じてそこから言語の持つシステムを自らの中で構築していくというプロセスには多大な時間と豊富な言語体験が必要となる。その過程を体験学習的にそのまま第2言語学習の中に持ち込むのは限られた学習時間内で可能なはずもない。生徒・学生にとって外国語は学習すべきことのほんの一部に過ぎず、それ以外にも様々なことを学習しなくてはならないわけだから。

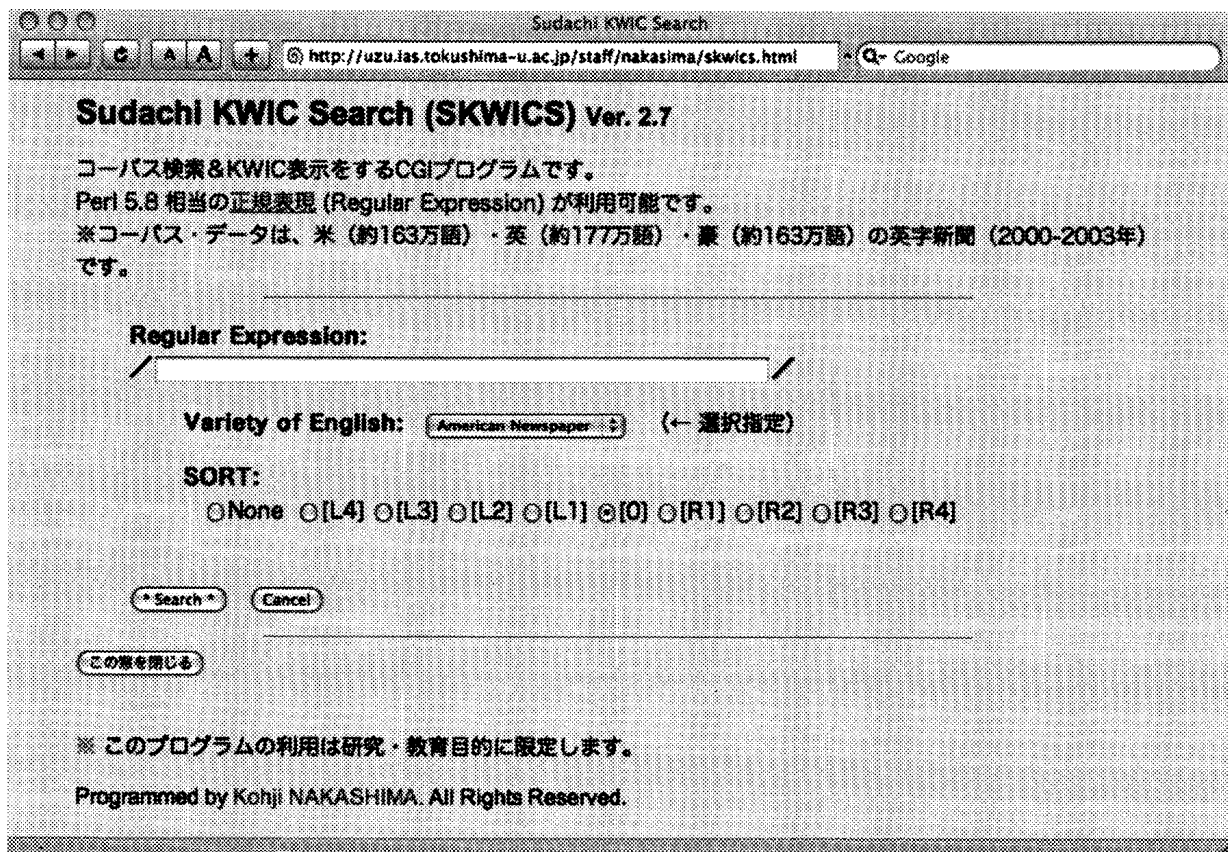
それではごく限られた時間内に、Parole 的言語データから帰納的に言語システムや文法の仕組みを抽象し自ら構築する疑似体験をさせるのにはどのような手法が考えられるだろうか。一つの可能性として、IT技術を用いて、誰でもどこからでも簡単に利用可能な言語データ（コーパス）検索システムの活用が有効ではないかというのが筆者の提案である。母語獲得の際は長い期間にわたって豊富に音声と文字によって言語に触れることになるが、時間のきわめて限られた第2言語学習ではどんな言語データと出会えるかも予想がつかないまま悠長に構えているわけにはいかない。大量の言語データから瞬時に該当する言語事例を提示することが可能なコーパス検索システムは、きわめて短い時間に凝縮して帰納的言語習得環境を提供できる可能性があるのではないか。その可能性を少しでも広げるために筆者が開発したのが、今回提示する‘Sudachi KWIC Search (SKWICS)’である。

3. コーパス検索システム SKWICS の開発とその概要

「どこからでも簡単に利用できる」ということが大前提のシステムなので、自ずと開発の方向性は決まってくる。「どこからでも」を実現するためには、インターネットブラウザによるアクセスで利用可能になるシステムが現段階では最も有力だろう。様々な通信機器やOS (Operating System) が提供されている現在、de facto standard となっている汎用的アクセス通信プロトコルは、HTTP (HyperText Transfer Protocol) であり、今後もしばらくは主流であり続けるだろう。インターフェイスの記述言語としては、現状では HTML, DHTML, XML, JavaScript, CSS など様々なものが用いられているが、やはりどのブラウザを

使っていようと動作上問題が少なく汎用性の高いHTML、それもバージョン的にはやや以前のものになるが、'HTML 3.2' (1997年1月14日付W3C勧告)、あるいは現在主流の'HTML 4.01' (1999年12月24日付W3C勧告、'HTML 4.0'の修正版)をベースにインターフェイスを構築するのが最適な解となるだろう。教育現場で用いる場合は流行や最先端の技術を用いることがかえって徒になる場合が少なくない。学校にもよるが、およそ5年スパンで端末機器の更新をする場合が多いと思われる。つまり、一般に5年間程度は同じシステムで運用され続けるわけだし、予算の都合上10年程度前の古いシステムを使い続けざるを得ないケースも出てくる。また、学校ごとに導入する機器やOSも様々である。これらをできるだけ吸収して動作するシステムを作るためには古くて安定している'HTML 3.2'が無難とも言えるが、このバージョンは非西欧言語に対応していないという将来性面での不安がある。今回作成したSKWICSでは、英語を対象とした検索システムであるため直接的な問題は生じないが、将来英語以外の言語も検索するシステムにアップグレードしていく可能性を考慮するなら、国際化に対応しているバージョンである'HTML 4.01'を採用しておいた方が良いでしょう。ただ、古くて推奨されないタグ類(タグなど)やCSS (Cascading Style Sheets)のサポートについては、ブラウザごとにばらつきがあるため、Strictな4.01では互換性に問題が出る可能性が高い。そういったことをすべて考慮して、今回作成したSKWICSでは後方互換性の高い'HTML 4.01 Transitional'を採用することにする。

「簡単に」検索できるようにするためには、何の説明書を読むこともなく直感的に操作できる単純で分かりやすいインターフェイスが必要であり、検索オプション類は機能を極力制限しない範囲で最小限にしておくことが肝要である。検索式を入力する欄は文字の打ち込みをせざるを得ないが、それ以外のオプションはセレクトメニューやラジオボタンを使いユーザーの負担をできるだけ減らした。また、オプション部分は何も意識しなくてもデフォルトで最小限の検索ができる形にデザインした。



Sudachi KWIC Search (SKWICS) Ver. 2.7

コーパス検索&KWIC表示をするCGIプログラムです。
Perl 5.8 相当の正規表現 (Regular Expression) が利用可能です。
※コーパス・データは、米 (約163万語) ・英 (約177万語) ・中 (約163万語) の英字新聞 (2000-2003年) です。

Regular Expression:
/ /

Variety of English: (← 選択指定)

SORT:
 None [L4] [L3] [L2] [L1] [0] [R1] [R2] [R3] [R4]

[この際を閉じる](#)

※ このプログラムの利用は研究・教育目的に限定します。
Programmed by Kohji NAKASHIMA. All Rights Reserved.

図1 Sudachi KWIC Search (SKWICS) の表紙画面

簡単に検索できるようにしたがゆえに高度な検索ができないと言えないよう設計することも重要である。何の説明書も読まずに簡単に検索するという使い方もできるし、高度で複雑な検索をしたいユーザーの要求を満たすことも可能であるというシステムが望ましい。それを実現するために正規表現を使った検索が可能に設計した。正規表現の優れているところは、単なるアルファベットの文字は通常の文字としてそのまま認識し、メタキャラクタ (特殊文字) を用いればさらに高度な文字列マッチングが可能であるという点にある。基本的な動作メカニズムとしては、表紙ページの HTML 内の<FORM>タグ内に入力された値は、Unix 系 OS である Fedora 上で動作する CGI (Common Gateway Interface) を通じて Perl (ver. 5.8) で書いたプログラムに送られる。受け取った値をもとに様々な処理をした後、検索結果を再び CGI 経由で HTML として返すことになる。内部的には Perl 5.8 の正規表現処理を利用しているため、SKWICS で検索に使える正規表現も基本的にそれに準じたものとなる。

正規表現 (Regular Expression) とは、いくつかの文字列集合のパターンを表現するための方法である。これを利用してマッチングを行えば一つの表現で複数の文字列を効率よく検索・置換することが可能になるため、文字列処理には

欠かせないものとなっている。もう少し具体化するために、SKWICS で英文検索をする際に有用であると思われるいくつかのメタキャラクタを挙げてみる。

表 1. 英文検索に役立つ正規表現のメタキャラクタ (一部)

.	改行文字以外の任意の 1 文字にマッチ
<i>a</i> *	文字 <i>a</i> が 0 文字以上続いているときにマッチ
<i>a</i> +	文字 <i>a</i> が 1 文字以上続いているときにマッチ
<i>a</i> ?	文字 <i>a</i> が 0 文字か 1 文字あるときにマッチ
(<i>a</i>)	正規表現 <i>a</i> をグループ化。¥n で参照可。
¥n	() でグループ化した正規表現にマッチした内容がそのまま入る
[<i>c1c2</i>]	[] で囲んだ文字群のいずれか 1 文字にマッチ
[^ <i>c1c2</i>]	[] で囲んだ文字群以外のいずれか 1 文字にマッチ
<i>a</i> { <i>n</i> }	<i>a</i> を <i>n</i> 回繰り返す文字列にマッチ
<i>a</i> { <i>n</i> , }	<i>a</i> を <i>n</i> 回以上繰り返す文字列にマッチ
<i>a</i> { <i>n</i> , <i>m</i> }	<i>a</i> を <i>n</i> 回から <i>m</i> 回繰り返す文字列にマッチ
<i>a</i> <i>b</i>	<i>a</i> または <i>b</i> にマッチ
¥b	語の境界 (boundary) にマッチ
¥s	空白文字 (スペース文字、タブ、改行) の 1 文字にマッチ
¥S	空白文字 (スペース文字、タブ、改行) 以外の 1 文字にマッチ
¥w	[a-zA-Z0-9_] と同じ
¥W	[^a-zA-Z0-9_] と同じ
¥d	[0-9] と同じ。数字 1 文字にマッチ
¥D	[^0-9] と同じ。数字以外の 1 文字にマッチ

表 1 でイタリック体になっている箇所は文字そのものを表しているのではなく、任意の文字あるいは文字列を表している。これらがどうして効率良い英文検索に有用であるかこの表を見ただけでは即座に理解できないかもしれないが、次章であげる具体的な英文検索の応用例とともに表 1 を見ればその意味と有用性が理解できるだろう。

4. SKWICS を使ったコーパス検索の語学学習への応用

帰納的語学学習の最大の弱点はその効率の悪さである。母語の言語獲得には極めて豊富なインプット・アウトプットの機会があり、ターゲット言語に接する時間は寝ているとき以外はほぼすべてといってもいいほど長い。それゆえに

少々効率の良くない帰納的学習プロセスであってもかまわないわけだし、成長に合わせて徐々に学習する方法としてはむしろこの方が適しているとも言える。たまたま出くわした単語や表現を徐々に学習していくというスタイルは集中的に脳に負荷を与えないため、万人にとって負担が少ないと考えられる。しかし、それをそのまま第2言語学習に応用するとどうなるだろうか。自然に任せてしまうと、教師側からみて理解してほしい単語・表現や文法的構造に出会えるのがいつになるのかさえ予想できず、いつになったら教えたい事項の大半に出会い尽くせるのかも分からない。臨界期以前以降という違いを度外視して単純に考えれば「母語の獲得に要した期間と同じくらいかかる」ということになるだろうが、母語に比べて第2言語習得の場合、インプット・アウトプットが致命的に貧弱であるということと接する物理的時間が極めて短いということを想起する必要がある。言語の帰納的学習は人にとって自然なプロセスではあるが、第2言語学習の環境でそのやり方を踏襲するのは無理がある。とは言え、演繹的で top-down 的にあらかじめ用意された答えを教えるような不自然な言語学習法では、教えられる側からすると押し付けられるばかりの単なる暗記科目のように感じられてしまい、その言語がどうしても生きていく上で必要不可欠だと感じている学習者以外にとっては、さほど興味のわかない苦痛な修行になってしまう。

このような「苦痛」を少しでも緩和し、帰納的でありながらも効率性を犠牲にしないために、SKWICS による英文検索の効用について考えてみたい。SKWICS を使えば、極めて短時間に教えたい単語・表現・共起に的を絞って学習者に英文を提示することが可能である。表2にSKWICSで利用可能な正規表現による英文検索の具体例を掲げる。

表2. 正規表現を使った英文の検索例

- ・ tall という単語を検索
 /tall/ で検索してみる
 → tall という文字列を含む語が検索されてしまう
 (正規表現の検索は部分一致検索のため)
 ヒント: 語の境界 (¥b) を指定する
 /¥btall¥b/
- ・ st で始まり tion で終わる 11 文字の単語
 /¥bst.....tion¥b/
- ・ woman と women という単語のどちらかを含む文

/¥bwom[ae]n¥b/

- happy を語幹とする単語の検索

ヒント：語尾変化だけでなく、prefix が付く場合も考慮

/¥b(un)?happ(y|ier|iest|ily|iness)¥b/

- for ... reason(s) という熟語の検索

→ 間に3つまでの単語が入る可能性を考慮

ヒント：単語は、おおよそ ¥S+ で表現可

/¥bfor (¥S+){0,3}reasons?¥b/

- one by one などのような、by の両側に同じ単語が繰り返される表現

/¥b(¥S+) by ¥1¥b/

例えば、表2の最後に例示した正規表現を使って実際に SKWICS を使った検索を行うと瞬時に次のような結果が画面に現れる。(図2)

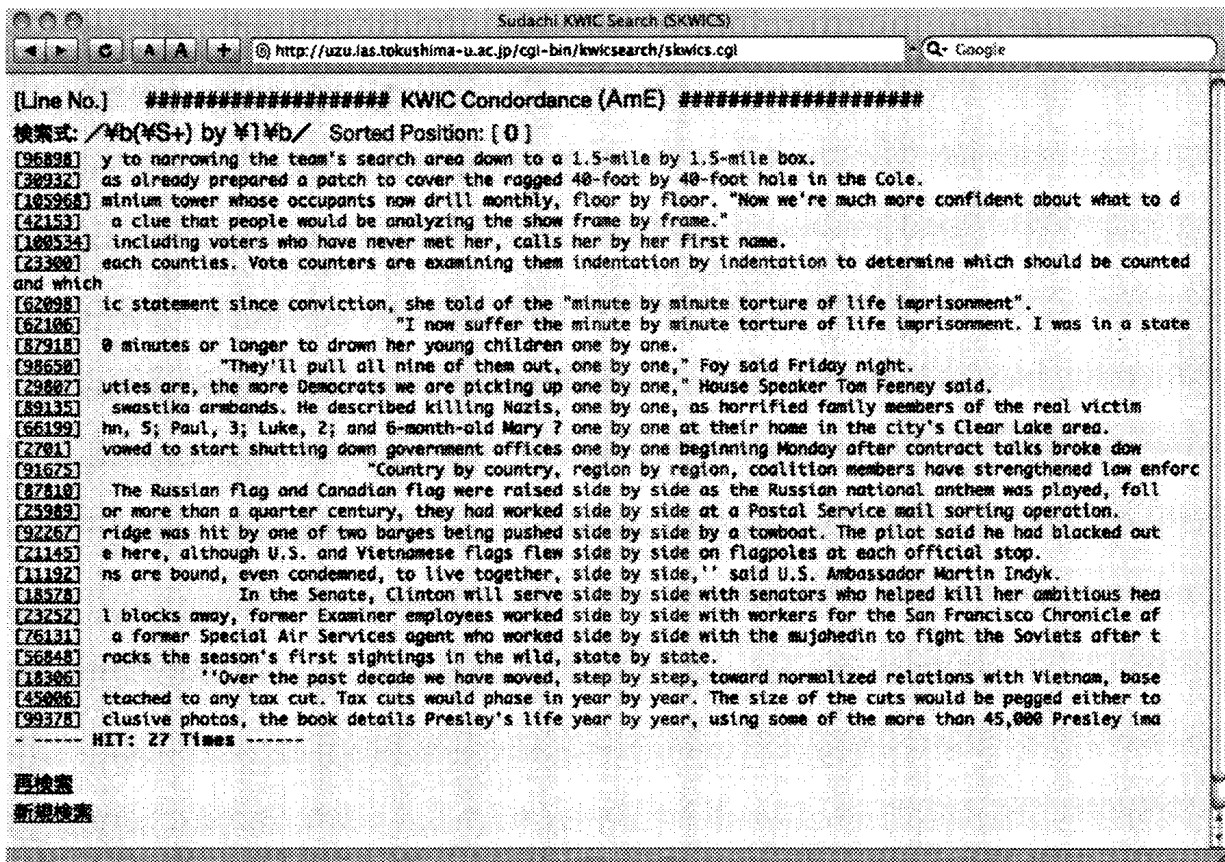


図2 SKWICS の検索結果提示画面

これは、検索対象コーパスをアメリカ英語の英字新聞記事（約163万語）と

して検索した結果画面の例である。検索対象文字列を Key としてその両側に文脈を表示するこのような表示形式は一般的に ‘KWIC (KeyWord In Context)’ と呼ばれている。

母語獲得の場合はこのような前置詞 ‘by’ の両側に同じ単語が現れる熟語的表現を長い言語体験の中で徐々に覚えていくわけだが、SKWICS を使った検索では瞬時に「経験」することが可能になる。換言すれば、時間を凝縮して言語体験が可能になるとも言えるだろう。‘by’ の両側にどういう単語が来るかあらかじめわからない状態で検索するというのは、アルファベット順に項目が並ぶ紙の辞書やアルファベットによる単純な検索機能しか持たない一般の電子辞書では到底不可能な検索例だが、このような柔軟で高度な検索が行えるのは正規表現を利用してシステムを構築しているからに外ならない。正規表現を使った検索を行えば、未知のしかも予想もしなかった言語表現に遭遇し新たな発見をすることさえ可能になるのだ。

さらにもう少し文脈を詳しく見たいときには、各コンコードンスラインの左にある [] に囲まれた数値 (Keyword を含む当該の英文の断片がコーパスファイル中の第何論理行目にあるか示している) をクリックすると、検索にヒットした箇所を含む段落およびその前後 1 段落分の英文が表示される仕組みになっている。教育・研究用途に限定しているとはいえ、著作権問題に配慮した最大限度の表示段落数であると考えられる。(図 3)

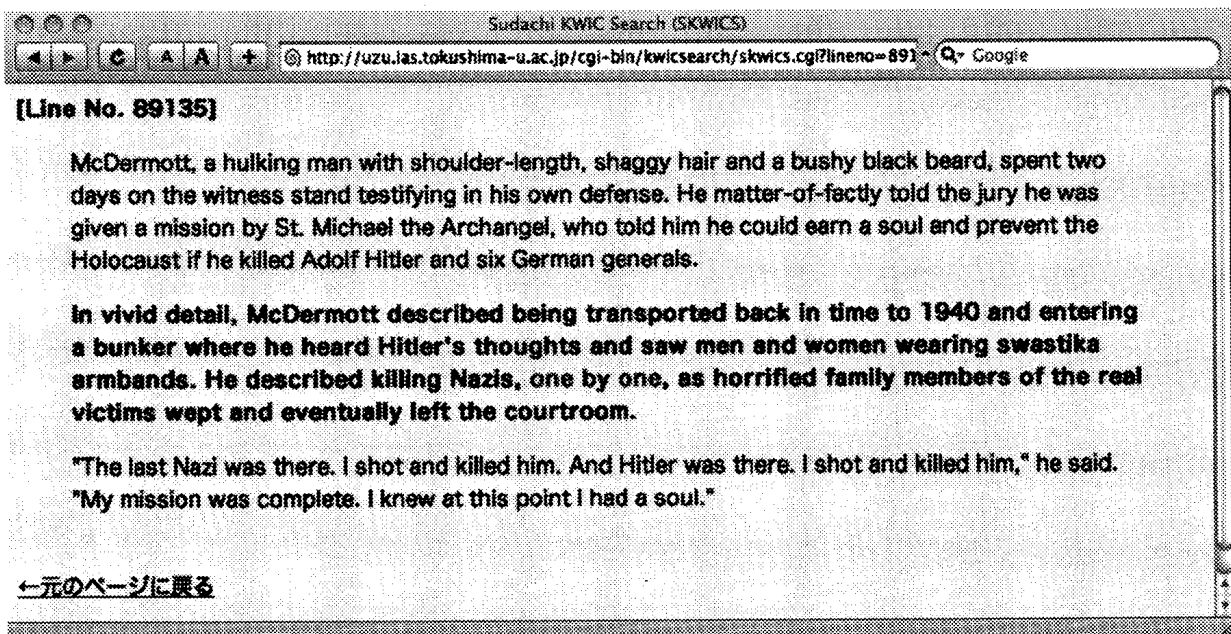


図 3 Concordance line を含む文脈を拡大した表示

検索対象にできる英文は、アメリカ英語の新聞記事だけでなくイギリス英語やオーストラリア英語の新聞記事に切り替えることも表紙画面（図 1）の“Variety of English”のセレクトメニューをマウス操作することで簡単にできる。英字新聞コーパス（テキストファイル）の構成は以下の通りで、いずれも 2000 年 9 月～2003 年 5 月の記事から筆者がランダムに集めたものである。

“usa0009-0305.txt” from “USA Today” consists of:

51277 Types, 1625634 Tokens

“gua0009-0305.txt” from “Guardian” consists of:

49363 Types, 1774437 Tokens

“smh0009-0305.txt” from “Sydney Morning Herald” consists of:

46019 Types, 1636087 Tokens

また、Keyword の前後 4 単語の位置で ASCII コード順にソートすることも“SORT”項目のラジオボタンをクリックすることで可能になり、共起する語彙を見つける際に便利なよう設計したが、デフォルトで Keyword 直前の単語を ASCII コード順にソートして検索結果を表示するようにしているため、このオプションをさほど意識しなくても検索結果はある程度分かりやすくグループ化して表示される。

前述したように、このシステムはインターネットに接続されている環境で一般的なブラウザさえあればどこからでも利用することが可能なよう表示部は“HTML 4.01 Transitional”で作成している。学校等の一般的な CALL 教室や家庭で自学自習する際にも面倒なインストール作業や複雑な操作もなく簡単に利用できる。

5. まとめと今後の課題

演繹的で top-down 的な語学教授法は、極めて限定的で貧弱な言語接触の機会しか望めない第 2 言語学習の現場では、言語習得のプロセスとしては不自然ではあるがやむを得ない側面がある。しかし、「言葉は絶えず変化している」ことを頭では理解していながら、教師自身が数十年前に学習した古くなった知識をもとに教えてしまっていることも多いのではないか。また、教師が教える際に参考にする文法書や辞書に書かれている内容は、それ以前に書かれた文法書や辞書に基づいて書かれている場合が多く、それらもまた同様である。すなわち、その記述内容が今現在実際に使われている英語を反映しているか極めて危

ういと言わざるを得ないし、文法書や解説書では分かりやすくするために事象を単純化し多くの例外を捨象してしまいがちだ。実際、高校時代に「demand-that-節中では、アメリカ英語では原形動詞、イギリス英語では should+原形動詞となる」と教えられた人も多いと思うが、実際の所どうなっているか SKWICS を使ってイギリス英語の新聞記事を対象に /`demand(s|ed|ing)? that`/ という正規表現を使って検索してみたところ、should+原形動詞がわずか 4 例だったのに対し、原形動詞は 29 例も見つかった。もはやとても上記の「定説」が成り立つ現状ではないということが、SKWICS を使えば簡単に「体験」できるのである。

SKWICS の現バージョン(ver. 2.7)ではコーパスデータが新聞英語に限定されているため中・高・大学生が学習に利用する英語例文としては必ずしも適切でない。将来的にはもう少し一般的な英語学習に適した英文をコーパス化する必要があるだろう。そういう課題はかかえながらも、英語学習補助システムとして SKWICS を活用することにより、無味乾燥な暗記科目として捉えられがちな英語教科を発見的で創造的な教科に変えるきっかけにできるのではないだろうか。

NOTE

'Sudachi KWIC Search (SKWICS)' は、中島のホームページ：
<http://uzu.ias.tokushima-u.ac.jp/staff/nakasima/> よりアクセス可能。

参考文献

- 大名 力 (2001) 「コーパス利用のための Perl」 in 『英語研究と英語教育のためのコーパス利用』名古屋大学言語文化部・国際言語文化研究科
- 大名 力 (2005) 「正規表現によるテキスト検索」 in 『大規模コーパスと英語研究 - テキスト処理ツールを使ってコーパスの中身を見る』名古屋大学国際開発研究科
- 岡秀夫・赤池秀代・酒井志延著 (2005) 『「英語授業力」教科マニュアル』大修館書店
- 鷹家秀史・須賀廣 (1998) 『実践コーパス言語学』桐原ユニ
- 齊藤俊雄・中村純作・赤野一郎 (2005) 『英語コーパス言語学 -基礎と実践- 改訂新版』研究社
- ミューテック (2006) 『Perl テクニックブック』C&R 研究所
- 武藤健志・トップスタジオ (2004) 『独習 Perl 第2版』SHOEISHA
- 結城 浩 (1998) 『Perl で作る CGI 入門 -応用編-』SOFTBANK
- Jeffrey E. F. Friedl (1999) 『詳説 正規表現』O' REILLY