

多重の制約を利用した英語教育用例提示

佐野 洋[†]

東京外国語大学 外国語学部

〒183-8534 東京都府中市朝日町 3-11-1

e-mail: sano@tufs.ac.jp

概要

外国語学習は技能科目なので言語知識の学習だけでなく、学習した言語知識を意識させた用例を使った運用練習が不可欠である。いわゆる語感獲得に関係するから練習用例数は多いほどよい。文法学習書や辞書では、学習項目毎に、その用例が、紙媒体という記録制約もあって 1 例しか掲載されていないことが通例である。自然言語処理技術を使い、コーパスから英文を抽出することで多くの用例を得ることができる。また電子化テキストなどの複合メディア媒体とインタフェース技術を用いることで、効果的な運用練習を実現できる。そこで筆者は、学習項目用例を英語コーパス(British National Corpus, 1 億語)から抽出した。提示システムは、運用練習で効率的に学習できるよう語法の気づきを支援する仕組みを、Ajax フレームワークを用いて開発している。

一般に学習項目は、言語知識を学習しやすいように、言語現象を 1 視点に絞り込んでいる。学習効果を向上させるため、インタフェースは、学習項目に対応する 1 つの制約だけでなく、学習対象の中心語の他の語との共起制約など、言語が持つ多重の制約を多視点的に提示することができる。本稿は、多視点的に用例提示することを意図した BNC からの句動詞用例抽出の方法と、それら抽出された句動詞用例を提供するインタフェースについて説明する。

1. はじめに

1.1 研究目的

筆者は、現代的で学習効果の高い教材開発の方法を研究開発している[5,6]。その一つとして、コーパスや自然言語処理技術とウェブインタフェース技術を活用した教材作成に取り組んでいる。2005 年には、著者の研究室と(株)小学館コーパスネットワークの共同研究により英語文型教授のための教育用英語用例サイト(1.2 節を参照)を開発した。

本稿は、句動詞用例の提供について述べる。既に公開している英語文型教授のためのサイトが、目次方式のインタフェースを提供しているのに対し、句動詞用例の提供では、索引方式のインタフェースを提供することが特徴である。

外国語学習の観点から見ると、後者は多重の制約を意識することのできる用例提供の仕組みである。ここで多重の制約とは、複数の言語特徴が用例に明示されていることをいう。

初学者には、一つの用例に複数の学習項目が含まれることは習得負担となる。一方、中級あるいは上級者には、一度に複数の語法が学べるから効率的な学習教材となり、学習効率を向上させることができる。

句動詞用例について、用例提示の際に複数の言語特徴を明示的にインタフェースで示すことができるように、(1) 句動詞の文型分析のモデルを提案する、(2) そのモデルに従って用例検索の手続きを改良し、用例抽出を行っている、(3) 同時に Ajax フレームワークを使ったインタフェース実装を行っている。

なお、句動詞用例集は、2006 年 9 月に(株)

[†] Linguistic Constraints and Interface Logic about Manipulation of English Teaching Material
SANO, Hiroshi
Faculty of Foreign Studies, Tokyo University of Foreign Studies.

小学館コーパスネットワークから公開予定である。

1.2 英語文型教授のための用例サイト

2005年度、著者の研究室と(株)小学館コーパスネットワークの共同研究により英語文型教授のための教育用英語用例サイト²を開発した。本用例サイトには、BNC(British National Corpus)から、コンピュータを使って自動抽出した 1,320 項目の英語文型用例(およそ 80 万例文)がアップされている(2005年9月に公開済)。

文法項目名(144 種類)と文の種類(14 種類)をメニューで指定すると、対応する用例がブラウザ画面に表示される(図 1)。

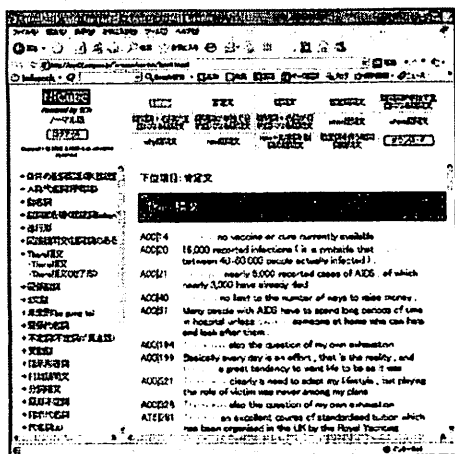


図 1 用例サイトの様子

1.3 句動詞用例

句動詞(Phrasal Verbs)は、動詞と前置詞もしくは副詞の連鎖で構成される動詞相当語句である³。例えば、eat away, eat in, eat into, eat off, eat out, eat up のような語句である。

句動詞は、簡単な動詞のあとに前置詞や副詞を付けて、さまざまな意味を表すことができるので、口語表現を中心に使用頻度が高いといわれている。

使用頻度が高いのだが、日本語には、前置詞や句動詞に対応する文法カテゴリがないの

² <http://www.corpora.jp/>

³ 群動詞(Group Verbs)とか2語動詞(Two-Word

で、一般に学習や習得が難しい項目であるともいわれている。学習のための説明や用例は、前置詞の項目にもあり、動詞の項目でも現れられ、しばしば副詞の説明の中にもある。

最近では、イメージ解釈やメタファー解釈によって前置詞/副詞を理解しようとする試み^[1]もある。

句動詞別の用例集は、使用頻度の高い表現を学習するための教材として重要であると考えられる。

筆者は、およそ 4300 エントリーにおよぶ句動詞見出しについて、文型記述を基にCQL⁴式への展開を行った^[6,11]。CQL 式に展開することで BNC から用例を抽出することができる。

1.4 用例集と用例提示

1.2節で示した用例サイトは、学習文型が目次としてウェブ画面に提示され、目的の用例を探す仕組みである(目次方式の検索インタフェース)。学習インタフェースとして考えると、目次の文字列そのものが学習内容を示すから、目次を辿る検索の操作が学習用例へのガイドになっている。

ところで、検索インタフェースには、目次方式と索引方式がある。目次方式の代表は YAHOO!™⁵で、人手で整理した分野別一覧から目的のページを探す仕組みになっている⁶。

一方、Google™は後者の代表である。2002年に Google が、キーワードとそれに関連する広告を表示する仕組みを提供して以来、ネット広告を通じた購買確率が上昇した⁶。検索連動型の広告は、検索結果に複数の概念を提示することに相当し、学習心理的には、連想

Verbs)とも呼ばれることがある。

⁴ CQL は、(株)小学館コーパスネットワークが提供するコーパス検索言語(Corpus Query Language)である。この言語様式を使って文型を記述し、コーパス検索エンジンを通じて BNC から用例抽出を行う。

⁵ 現在は、索引方式も併用されている。

<http://www.yahoo.co.jp/>

⁶ 2004年のネット広告市場は1800億円といわれ、そのうち検索語に連動した広告は350億円(約20%)という(2005年8月26日付け朝日新聞記事から引用)。

記憶を活性化させ、いわゆる気づき (consciousness raising) を促進していると考えられる。

目次方式の検索は、初心者にも使いやすいが細かい検索が難しい。図1で示す用例提示では、文型(144種類)が基本的な制約になっている。ただし、肯定文、否定文や疑問文などの文の種類を意識的に利用すると、同一文型中で、どの文の種類が多用されるのか、文の種類毎に使われる語彙の分布は何かなどを調べることができる。したがって、この2点を意識して注意深く用例を集めるという条件の下で、3つの言語的な制約を気づかせる練習用例ができあがる。

句動詞用例の提示を考える。用例検索インタフェースとして目次方式を想定すると、単純にアルファベット順に句動詞(語形)が配置されることになるし、語形自身を言語制約として考えると1制約でしかない。見出し語数が4300もあるので使いづらい。目的の特定句動詞用例以外に文型情報も提示可能だが、文型の種類も、(考え方に依存するが)最大で20程度であることから言語的な気づきも増えない。したがって、句動詞用例の提示では、言語が持つ多重の制約を多視点的に提示することがポイントになる。

以下、本稿では、多視点的に用例提示することを意図したBNCからの句動詞用例抽出の方法と、それら抽出された句動詞用例を提供するインタフェースについて説明する。

2. 文型分析モデル

2.1 動詞(原形)文型

一般に文型説明は動詞を中心に行われるが、動詞を原形と仮定し、動詞の補語部分を強調することによって学習内容を示す。例えば、[10,493頁]の文型説明は次のようである(枠内は引用。但し、例文記述部分は省略した)。

(I) 前置詞付き動詞：
動詞+前置詞+名詞句
この場合の前置詞は、名詞句を動詞に結びつける働きをする。
(II) 句動詞：

動詞+副詞

この場合の副詞は、動詞の意味を変える働きをする。

(III) 前置詞付き句動詞：

動詞+副詞+前置詞+名詞句

この場合、副詞は動詞の意味を変える働きをし、前置詞は、名詞句を動詞(+副詞)に結びつける働きをする。

また、いわゆる「基本5文型」を基礎にして拡張している文型表現も多いが、拡張部分は、補語の種類を細分化したもの(to-do, that-節, doing, 過去分詞, wh-節など)である。

言語制約としての視点は、「(主)動詞の右側に現れる要素配列に注目する」ことで、学習項目は、要素配列の種類になり、学習者は要素配列の様子に注目する。

このモデルでも新たな視点を導入することができる。一つは前置詞/副詞の分類で、もう一つは名詞句のクラスタリングである。

前者は、言語知見の導入である。句動詞データの電子化のために参照した文献[6]では、前置詞/副詞を、その語形ではなく、「テーマ」という概念で意味分類している。テーマ名を制約に加えることができる。

後者は、例えば、前置詞付き動詞(動詞+前置詞+名詞句)を取り上げると、名詞句部分に現れる語彙の頻度計算を用いて、名詞句に現れやすい語彙分布が分かる。クラスタリングの結果である語彙分布特徴⁷も言語制約である。

2.2 述語(定形・非定形)文型

筆者が提案する文型分析のモデルは、述語を動詞部と補語部からなると仮定し、補語部分の要素配列だけでなく、主語と動詞部の要素配列も強調する。説明と特徴を模式化した図2に示す。

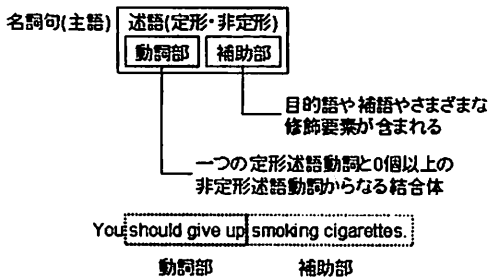
必要に応じて主語との一致を示す動詞の形を定形(finite form)という。動詞の形が、その主語によって定まるかどうかを基準とする分類である[6,7,8]。英語では、一般動詞は、

⁷ コーパス言語学の研究領域では活発に調査研究が行われている。

現在形と過去形だけが主語と結びつく形である。過去分詞形、-ing 形、不定詞は非定形(non-finite form)と呼ばれている。

一般動詞の非定形は、助動詞(be, have, do, will 等)や助動詞相当句(be going to など)と結合し、その結合形の用法は複雑である。結合形を構成したとき、本動詞は、結合形の最後に位置し、非定形をとり、実質的な意味を担う。主語と結びつく定形の助動詞は、時間関係や話し手の心的態度(モーダル表現)を担っている。"may have been Ving"のように、非定形の助動詞が連続することもある。

本稿の述語のとらえ方



一般の述語(動詞)のとらえ方

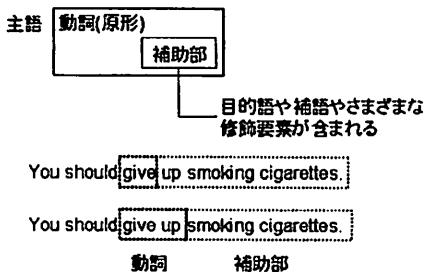


図 2. 述語文型とその比較

このモデルの採用によって、さらに新たな視点を導入することができる。動詞部の文型と名詞句(主語)のクラスタリングである。

2.3 動詞部文型

筆者が動詞部文型と名付けている文型の特徴は、主語と表現意図の共起関係を扱うことである。従来の辞書や文法書にはあまり記述がない。

例えば、"account for"は、"account"自身が定形で使われることが多く、しかも現在形の場合が多い(助動詞等と共に起さない)。それに

対して、"acquaint with"は、しばしば、辞書などで"be acquainted with"として語法説明がされるように、"acquaint"は非定形(過去分詞形)をとり、実質的な意味を担うけれども、主語と結びつくのは、be 動詞である。be 動詞は定形をとり、BNC から抽出した用例を見ると、現在形(is/am/are)は使われずに、過去形(was/were)が使われていることが分かる。

さらに特定動詞部文型に対する名詞句(主語)部分に現れる語彙の頻度計算を用いて、主語位置に現れやすい語彙分布が分かる。クラスタリングの結果である語彙分布は言語制約である。

この分析モデルを採用することで、抽出用例の細かな分類ができる。そして、句動詞ごとの語法観察の視野がより広くなると考えられる。

3. 句動詞用例抽出

3.1 句動詞補助部文型

句動詞一覧をデータ化するために参照した『英語句動詞文例辞典』[3]は、句動詞を、単に見出し語リストとして収集したのではなく、以下の考え方で分類し、且つシンボルグラマーと呼ばれる独自の方法で補助部に 37 の文型を認めている。

- 前置詞と副詞を区別せず、パーティクル(particle: 不変化詞)とする
- 補助部(動詞の目的語やいわゆる前置詞の目的語)の下位分類を設ける
- 補助部を義務的にとる文型と任意にとる文型に分ける
- 上記の方針に沿って文型分類し、文型が異なれば、インデックスが同じでも異なる句動詞とする

[3]では、動詞に後続する要素(いわゆる述語要素)を句動詞の特性によって 37 の類型を分類している。[3]によると、補助部を構成する要素は 7 種類である。なお、明示的に指定される 7 要素以外の要素という指定、各要素が義務的に必要な場合と、任意の場合があること、さらに語順の違いがあることから 37

の文型がある。なお、文型の CQL 式への展開と用例抽出については、[6]に詳しい。

3.2 前置詞/副詞とテーマ

[3]は前置詞/副詞を、その語形ではなく、「テーマ」という概念で意味分類している。33のテーマが設定されていて、さらにサブテーマがある。一つのテーマの中に複数の前置詞/副詞が所属し、それぞれサブテーマが与えられている(表 1)。これを元に個々の句動詞エントリーとテーマの対応データを作成した。

表 1. テーマの例(一部)

	テーマ
1	動き(狭い範囲:小さい/一時的な)
2	動き(広い範囲:大きい/連続的な/目的を持った)
3	接近・到着・訪問
4	出発・去ること・発送
5	位置・(非)存在
6	接触・発見・近隣
7	攻撃(言葉による/物理的な)
8	抑圧・打ち勝つこと
9	拒絶・回避
10	説得・奨励・強制
...	...

3.3 句動詞動詞部文型

複数の文法書[2,8,9,10]に従って動詞部文型の調査を行い、述語の定形・非定形パターン展開を行った。そして、CQL 式で検索式を記述した。

定形動詞が現在形で且つ肯定の場合、30の定形パターンがある。ただし、助動詞相当語句の認め方などでパターン数(30)は増減する。定形の種類が現在形、過去形の2種類あり、述部には、否定辞を含めることができるので、約4倍のパターン数になる。

例えば、“[am, is, are]+[able/going]+to+have+been+Ving”は6語から構成される動詞部文型である。ちなみに CQL 式に展開すると次のようになる。

```
{P="N.*|PN.|CRD"}{W="am|is|are"}
P="VBB|VBZ"}{W="able|going"}
L="able|go"}{W="to"} P="TO0"}{W="have"}
P="VHI"}{L="@Var1"} P="V.N"}
```

なお、文型の CQL 式への展開については、[2]に詳しい。

3.4 用例検索工程

句動詞ごとの補助部文型(シンボルグラマーのタイプ)によって CQL 式数に違いがあるが、肯定文だけで、おおよそ 100 行程度の CQL 式に展開される。この CQL 式は、CQL スクリプトとしてコーパス検索エンジンで実行され、英語コーパス(BNC)から用例が抽出される。

例えば、“act out”は、act+out+名詞句と act+名詞・代名詞+out の2つの補助部文型を持つ。このように見出し語レベルで、2つの補助部文型がある場合があるので、全展開すると 5,772 の CQL スクリプトが生成される。さらに動詞部文型が 97 あるので、総計で 559,884(5,772×97)の抽出スクリプトがある。

動詞部文型 97 のうち、各句動詞はおそらく 2~3 程度の抽出がある見込みで、一般的な辞書記述の点からみても語法の精緻化につながるデータが得られることが期待される。

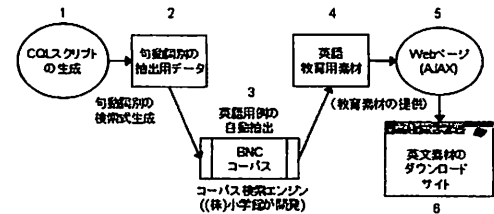


図 3. 作業工程の概念図

図 3 には、作業工程の概念図を示す。図中の 1,2 の工程を終え、3 の抽出作業を行っている。4,5 のインターフェースについては、現在、設計を実施し、一部プロトタイプを作成を行っている。

4. 多重制約とインターフェース

4.1 電子化データと言語上の制約

[6]をもとに、句動詞の見出し語と対応するシンボルグラマーのタイプだけでなく、不変化詞と対応するテーマ、サブテーマ名も電子化した。動詞部文型も整理し電子化したので、最終的に、抽出結果としての各用例には、表 2 に示す言語上の制約情報が対応することになる。

表 2. 電子化データと言語上の制約

句動詞	数	言語上の制約
見出し語数	4332	語形、訳
補助部文型	37	文型、語彙分布(目的語、補語)
動詞部文型	97	文型、語彙分布(主語)
前置詞/副詞	46	語形、テーマ、サブテーマ

句動詞用例の提示に目次方式を想定すると、アルファベット順に句動詞(語形)が配置され、語形と文型で制限された用例が表示されるだろう(図 4は図 3にあるコーパス検索エンジンを使い BNC から用例検索を行った結果である)。

Many are highly adept at adding up their calorie intake .
 little wooden abacuses to add up your bill .
 vel interest , but when we add up those pledges and cost
 a figure he arrived at by adding up the ages of people in t
 How do we add up the value of a school ?
 iht visits was obtained by adding up the four quarterly tota
 Adding up everyone 's several sou
 Of course , if we were to add up the tora in evry marke
 the may take quite a while adding up the wine list . "
 relative figure obtained by adding up the cost of treating a
 g from A to B is found by adding up the waves for all the p
 iten to the song it has to add up you know . it ca n't ju
 not be solely a matter of adding up figures on a balance sh
 But add up the ledger : psychic te
 • Add up your logos As you make your

図 4. "add up" の用例表示例

見出し語数が凡そ 4300 もあるので使いづらいし、言語制約は見出し語の語形だけである。仮に"add up"の日本語訳を表示しても語法上の気づきに結びつかない。1.4節で述べたように、句動詞用例の提示では、言語が持つ多重の制約を多視点的に提示することがポイントになる。

不変化詞"up"の意味分類(「テーマ」)や"add up"の意味分類(「サブテーマ」)、"add up"の動詞部文型ごとの用例表示など、表 2で示す言語上の制約情報を活用した索引方式のインタフェースが必要である。

検索条件としての言語上の制約を、用例学習に活かすには、利用者が、現時点でどの視点で検索をしているかが意識でき、且つ、用例を絞り込むための他の視点をナビゲートできるようなインタフェースが理想である。

4.2 Ajax によるインタフェース実装

現在、Ajax でのインタフェース実装を目指し、ソフトウェア開発を行っている。

Ajax⁸フレームワークは、サーバ(本稿の場合、用例サイト)との通信結果に応じて DHTML で動的にページの一部を書き換えるインタフェースを実現するというアプローチを取る。句動詞用例提供では、できるだけ画面遷移を伴わない動的な Web アプリケーションとして実装を計画している。

テーマ/サブテーマについては、テーマをイメージ(アイコン)化してテーママップを作成する。33 項目あるテーマもイメージ化することで、一画面でナビゲート可能になる。また、同画面には、索引入力の入力ボックスを設け、GoogleTMサジェスト風に、(1) 句動詞見出し語の入力支援、(2) 前置詞/副詞(不変化詞)の入力支援を行う。選択肢の入力支援及び選択肢の候補の自動リストアップを行いながら、利用者が求める用例にアクセスできるようにする。

語彙分布の利用者への提示方式と、動詞部文型(選択幅として文型数(97)を絞り込む視覚化の方法)については、インタフェースの実現方法をもっか検討中である。

4.3 インターフェイス

図 5 に実際の用例検索のためのインターフェイスの様子を示す。なお、図中で「テーマアイコン」とある部分は、イメージ化されたテーマである。現在、デザイナーに作画を依頼中である。

⁸ Ajax(Asynchronous JavaScript and XML)は、XMLHttpRequest (HTTP 通信を行うための JavaScript 組み込みクラス)を使った非同期通信を利用するユーザーインタフェース構築技術の総称である。なお、開発のベースには Ruby on Rails を用いている。

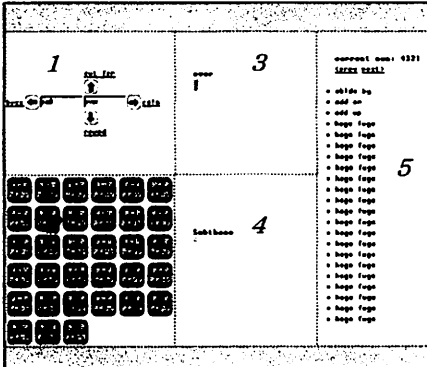


図 5. 用例インタフェース

インタフェースは5つの部分から構成されている。

1. 動詞・前置詞検索フィールド

Google™サジェスト風に動詞・前置詞の選択候補が自動補完で提示される(図 6)。見出し語語形の入力を支援する。

2. テーママップフィールド

テーマをアイコンによって視覚的に表現する。アイコンがクリックされるとポップアップのサブウィンドウを表示し、サブテーマを表示する。

3. 前置詞の制限による動詞選択フィールド

1のフィールドで選択された前置詞/副詞について、共起する動詞が提示される。スライダのつまみを上下することで選択できる。前置詞/副詞を固定視点にした動詞語群を意識することができる。

4. サブテーマ選択フィールド

サブテーマを選択する。スライダのつまみを上下する操作は、近接した意味を意識させる。

5. 項目表示フィールド

1~4のフィールド内における操作で制限された用例項目を提示する。多視点の制限によってインターラクティブに用例項目数が絞り込まれる。

絞り込み過程は、複数の言語視点で制限を行う過程である。すなわち表示される項目が減少することを観察する行為は、複数の言語制約がどのように課されているかを知る行為となる。

なお、用例全体は、絞り込まれた項目をクリックすることで、ポップアップの別ウィンドウ上に表示する。

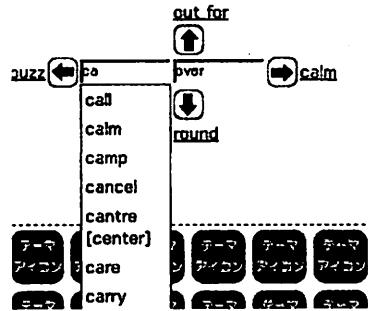


図 6. Google サジェスト風の選択候補表示

学習者は、1~4の4つの領域で制約を加えることで、多重の制約を意識し、多視点な検索を行うことができる。

5. おわりに

本稿は、句動詞用例の提供について述べた。既に公開している英語文型教授のためのサイトが、目次方式のインタフェースを提供しているのに対し、句動詞用例の提供では、索引方式のインタフェースを提供することが特徴で、多重の言語上の制約を意識することのできる用例提供の仕組みである。句動詞動詞部文型部分のインタフェース実現を早急に進めたい。

本インタフェースは、学習項目に対応する1つの制約だけでなく、学習対象の中心語の他の語との共起制約など、言語が持つ多重の制約を多視点的に提示することができる中級あるいは上級者には、一度に複数の語法が学べるから効率的な学習教材となり、学習効率を向上させることができると考えている。評価実験を通じて効果を検証したい。

本稿では、句動詞用例について、用例提示の際に複数の言語特徴を明示的にインタフェースで示すことができるように、句動詞の文型分析のモデルを提案し、そのモデルに従った用例検索の手続きのインタフェースについて説明した。

謝辞

査読者からは有益なコメントが与えられました。感謝します。

参考文献

- [1] アンドレア・タイラー, ビビアンエバンス, 『英語前置詞の意味論』, 研究社, 2005.
- [2] AJトムソン/AV マーティネット, 江川泰一郎訳注, 『実例英文法 第3版』, オックスフォード大学出版局, 1983.
- [3] クリストファバーナード, 『英語句動詞文例辞典』, 研究社, 2002.
- [4] 佐藤尊範, ドナルド C. ウッド, 片平昌幸, 中村 彰, 「正規表現全文検索を利用した英文 CORPUS 利用教育と学習用具」, Computer & Education, Vol.20, コンピュータ利用教育協議会(CIEC), 2006.
- [5] 佐野洋: 「ESP 適合の教材コンテンツを実現する語学教育支援システム」, 『最新外国語 CALL の研究と実践』, コンピュータ利用教育協議会(CIEC)・外国語教育研究部会(34~44,10頁),2003年3月.
- [6] 佐野洋, 「BNC を利用した教育用英語句動詞用例集」, コンピュータと教育研究会, 2006-CE-84, 1pp-6pp, 2006.
- [7] ジェフリー・リーチ, 田中春美, 樋口時弘訳, 『ネイティブ英語運用辞典』, MACMILLAN LANGUAGEHOUSE, 1999.
- [8] 鈴木英一, 『統語論』, 開拓社, 1995.
- [9] 安井稔, 『英文法総覧』, 開拓社, 2003.
- [10] 綿貫陽他, 『ロイヤル英文法』, 旺文社, 2004.
- [11] Nakamura, T., Tateno, J. and Tono, Y. (2004) Introducing the Shogakukan Corpus Query System and the Shogakukan Language Toolbox. Williams, G. and Vessier, S. (eds) EURALEX 2004 Proceedings . The Eleventh EURALEX International Congress, July 6-10, 2004, Lorient, France, pp. 147-152.