

英語コーパス学会第12回大会

日時 1998年10月10日(土)

会場 椋山女学園大学 (名古屋市千種区星ヶ丘元町17-3)

(JR名古屋駅で地下鉄東山線藤ヶ丘行に乗換え20分、星ヶ丘下車、六番出口より徒歩5分。CONF\_12.files¥www.sugiyama-u.ac.jp 参照)

ワークショップ 10:30-12:00

《テキスト処理: 検索と簡単な計数処理の実際》  
梗概

講師 浜口 崇

先着 30名 (予定) 参加費 会員無料・非会員 1,000円

(申し込みは電子メール・郵便で事務局まで)

受付開始 12:30

開 会 13:00

1. 会長挨拶  
齊藤 俊雄

姫路獨協大学

2. その他

研究発表 13:15-15:15

第1室  
学 水光 雅則

司会 九州工業大学 許斐 慧二

京都大

1. 「現代英語における ed-adjective 形成の条件」  
伸子 梗概

東京大学大学院 國森

2. 「英語における SVCO パタンについて」

明治大学

久保田 俊彦 梗概

3. 「go ?ing 構文の ?ing のところに来る動詞について」 関西外国語大学 岡  
田 啓 梗概

4. 「否定強意副詞（句）の統語的特性」 中央大  
学 新井 洋一 梗概

第 2 室 司会 大阪大学 田畑 智司  
名古屋大学 滝沢 直宏

1. 「語源情報を用いたテキスト分析」 早稲田大学大学院 長  
田 哲男 梗概

左良木機械翻

訳研究室 左良木 昌 梗概

2. 「多変量解析法による因子の言語的解釈について」  
— LOB Corpus における第 3 因子の解釈 — 豊田工業高等専門学校 高橋  
薫 梗概

3. 「日英翻訳で見る曖昧な日本語の『など（等々）』」 関西外国語大学短期大  
学部 西村 公正 梗概

〈休憩 15:15-15:30〉

シンポジウム 15:30-17:30

《コーパスと英語教育》 司会 東海大学 朝尾  
幸次郎

「コーパスの英語教育への利用の歴史」 講師 東海大学 朝尾幸

次郎 梗概

「コーパスによる学習支援システム」

講師 名古屋大学

杉浦 正利 梗概

「学習者コーパスの構築」

講師 中部大学

尾関 修治 梗概

「学習者コーパスによる事例研究」

講師 立命館大学 野

澤 和典 梗概

「第二言語習得と学習者コーパス」

講師 名古屋外国語大

学 原田 邦彦 梗概

閉会の辞  
彦

梶山女学園大学 深谷 輝

《懇親会 17:40-19:30 梶山女学園大学 生活社会棟 1 階会議室 会費  
4,000 円》

英語コーパス学会 (Japan Association for English Corpus Studies)

会長 斎藤俊雄

事務局 770-8502 徳島市南常三島町 1-1 徳島大学総合科学部 中村純作研究室

TEL 0886-56-7129 郵便振替口座 00940-5-250586 (英語コーパス学会)

E-mail: (E-mail address deleted)

URL [../index.html](#)

◆大会当日、入会受付もいたしますので、お誘い合わせの上ご参加下さい (年会費 一般  
4,000 円 学生 3,000 円)。また「当日会員」としての参加も受け付けております(1,000 円)。

## 英語コーパス学会第12回大会レジュメ

### ◆ ワークショップ《テキスト検索：検索と簡単な係数処理の実際》 講師 浜口 崇

今回のワークショップでは、MS Windows 環境(Windows3.1+Win32s を含む)で検索ソフト等のテキスト検索・各種処理ツールと市販アプリケーションなどとの連携を含めたテキスト処理の実習を行います。

現在テキストファイルとして CD-ROM や WWW などでは大量のファイルが入手可能ですが、それらのファイルから様々なデータを引き出すために、複数のツール・アプリケーションを組み合わせて利用する技術を身につけていただく事を目標とします。

前半はテキスト処理の基本といえる単語数計数・単語出現頻度(全体と個別ファイル単位の比較)と、正規表現の利用を含めたテキスト検索および検索結果に対する絞り込み・形式変換などの基本操作について、拙作ソフト Corpus Wizard, Boreal, Dr.Count, Parallel Boreal を用いて確認していきます。

後半は、エディタ・ワープロソフトへのデータ転送とタグジャンプ、表計算ソフトへのデータ転送後の簡単な計数処理、簡易検索マクロなどアプリケーションとの連携を含めた応用のための技術について、いくつか紹介します。

ワークショップで使用するオンラインソフトは PACK <http://www.vector.co.jp/> 等から入手可能です。

### ◆ 研究発表 第1室

#### ● 現代英語における ed-adjective 形成の条件 國森 伸子

現代英語において、名詞が接尾辞 **-ed** をとって修飾的に使われる場合、「その名詞の表すもの(部分)を被修飾名詞の表すもの(全体)が所有する」という意味の形容詞になるとされる。(例: 'horned frog' = ツノのあるカエル) OEDによると、この接尾辞 **-ed** は、あらゆる名詞について「所有」「特徴づけ」の意味の形容詞となると記述されているが、現実にはそのように機能しない例(\*wifed man, \*carred man, etc.)も存在する。では、どのような名詞が ed-adjective として許容され、どのような名詞がされないのだろうか、その有契的な理由はあるのだろうか? 本稿では、この疑問を解消することをめざし、「多数のデータの分析を通して、ed-adjective 形成の条件を導出すること」を目的とする。構成は以

下のとおりとする。

(1) 先行研究の結果と問題点の確認

(2) 既を示されている「ed-adjective 形成の条件」の妥当性の検討

・ OED を中心に辞書、コーパスから ed-adjective 用例を抽出し、分析する。

(3) 先行研究ではほとんど言及されていない、許容されない例の分析

・ 計算機プログラムによって創作した ed-adjective 新規表現に対してネイティブスピーカーによる許容度テストを試み、許容されない例の要因を考察する。

(4) より適切な「ed-adjective 形成の条件」の導出、残る問題点の確認

先行研究で言及されてきた「所有関係（部分と全体）」という条件が重要であることは多数のデータの分析から統計的に確認できた。さらに、その関係の種類についてデータを詳しく分析することにより、これまで深入りされなかった「許容されない例」の理由にある程度、体系的な説明が与えられ、より適切な ed-adjective 形成の条件が求められるものとする。

#### ●英語における SVCO パタンについて 久保田俊彦

いわゆる SVOC という語順パタンを持つ動詞のいくつかは、同時に SVCO という語順を持っている。一例を挙げるならば、動詞 shoot はその補部に次の二種類のパタンを持っている。

- a. The burglar shot the customer dead.
- b. The burglar shot dead the customer.

研究書、辞書等において、このパタンに従来十分な検討が加えられてきたとは言い難い。Cobuild English Dictionary は第二版（1995）でこのパタンの用例を収録しているが、これを独自のパタンとしては認識していない。大多数の辞書ではそもそも用例が収録されていない。本発表ではこの SVCO というパタンの特徴・成立範囲を統語、意味の両面から探ると同時に、社会言語学的な考察も加える。

SVCO パタンのように比較的稀な言語現象の検証には大型のデータベースにあたること

が必要であり、100万語規模のデータベースでは不十分である。今回は一般に利用可能なコーパスのうちで最大のものの一つ、COBUILD Direct を主に使用するが、ここで現在提供されているデータベースでもなお規模的に不足することも指摘することになる。

● 「go ~ing」構文の「~ing」のところに来る動詞について 岡田 啓

この発表の目的は、コーパスを用いて「go ~ing」の「~」のところに来る動詞を調査し、それを分類するとともに、できるだけ多くの具体例を列挙することである。

コーパスとしては、現代の小説・随筆などを収録した関西外大コーパス（136MB）、特に児童書を中心に集めた私的コーパス（20MB）英国の The Daily Telegraph 紙（1992年分、185MB）、米国の The Washington Times 紙（1996年分、203MB）、Time Magazine（1989年・1994年、66.5MB）などを用いた。語数にして1億語を越えるコーパスからの集積であるから、ここで挙げた類型以外の「go + ing」が存在する可能性は極めて小さいといえよう。この発表では、コーパス利用の最大の利点である、網羅的（といってもやはり漏れるものもあるが）用例収集を試みた。その結果を分類し、なぜこの構文が特定の使用のしかたをされるのか、なぜ限られた動詞のみと結びつくのかについて考察と行いたい。なお、本発表では~ing が分詞であるか動名詞であるかなどの論議は行わない。学習者にとっては、動詞 go と組み合わせになる ~ing はどのようなものがあるか、またそれらはどのような意味基準で用いられているかを理解できれば十分であると思うからである。

調査の結果、go の後ろに来る動詞は、およそ次の4群に分類される。

1. 社会的慣行（Social Institution）として確立され

た行為で、原則として目的語を取らないもの

- a. go dancing/shopping （社会生活上の行動）
- b. go swimming/skiing （スポーツ）
- c. go hunting/fishing/blackberrying （狩猟・採取）
- d. go birdwatching/sightseeing
- e. go hiking/picnicking（道行きを含む移動）

2. 話者が行為にたいして「不適切」「不本意」「是認できない」という気持ちを表し、多く否定の脈絡で用いられるもの

Now, don't go talking behind my back.

Don't go getting married one week and throwing it up the next..

3. go という「移動」に伴う、音声・様態などを付加するもの  
動きを表す副詞辞の about, around, off などを用いることが多い。
  - a. 音声 : A motorcycle went chugging by the house.  
go puffing past
  - b. 様態 : He went striding across the lawn.
  - c. その他 : They went swarming up the ladder.
  
4. 「～ing で与えられている状態に入る」という意味を持つもの  
go missing/hiding  
go fasting/dieting

● 否定強意副詞（句）の統語的特性 新井 洋一

英語の副詞（句）の中でも、whatever, at all, in the least 等は否定文の中で用いられ、否定を強める働きをしている強意副詞（句）であることはよく知られている。また否定環境 (Negative Contexts) の中でのみ生起する意味において、否定極性項目 (Negative Polarity Items) とみなすことができる。

また、次の各文において、

- a. I wouldn't hurt her for the world.
- b. I wouldn't touch her with a ten-foot pole.
- c. I wouldn't go back to that job for the tea in China.

for the world, with a ten-foot pole, for the tea in China はどれも、同じ機能を持つ副詞句である。この発表では、英語におけるこれらの副詞（句）について、コーパスを用いて考察しながら、それぞれの副詞（句）の意味的特徴、および統語的特性について考察することにした。特に、any や ever 等の他の否定極性項目と比較しながら、生起する環境面の違いを明確にすることによって、これらの副詞（句）の否定極性認可 (Negative Polarity Licensing) の条件について考察する予定である。そして、否定極性項目の認可条件を一元的に規定する提案は不十分なものであり、それぞれの制約をより精緻に規定する必要がある。

ることを論じたい。

#### ◆研究発表 第2室

##### ●語源情報を用いたテキスト分析 長田 哲男, 左良木 昌

言語表現の語源属性は文法・意味属性とともにその表現形式のあり方を決定する重要な要素である。従って、英語に特有であるような表現構造の分析にあたっては、文法や意味の側面のみならず、その表現を構成する個々の単語の語源を考察することが大きな課題となる。

例えば、P. M. Roget は *Thesaurus* 初版の *Introduction* において、*Resistance-Attack-Submission*、*Beginning-Middle-End* などの「相関語 (correlative word)」を挙げているが、ここで各語について前組はロマンス語系、後組はゲルマン語系という語源情報を得ることで、なぜこれら三語が一組となって使用されるのかが明らかとなる。

このような、語源情報をもとにした英語表現の分析を目的として開発されたのが「英語語源辞書 (Dictionary of English Etymology for natural language processing)」(以下 DEE) とタグ付与プログラムである。DEE はおよそ 25000 の英単語について、その語源と借用過程を示したデータベースであり、このデータベースをもとにタグ付与プログラムが任意の英文テキスト内の個々の単語に語源情報を付与する。

本発表では、これまでの文法・意味分析では明らかにされることのなかった英語の表現構造を取り上げ、DEE とタグ付与プログラムによる分析の事例を提示する。

##### ● 多変量解析法による因子の言語的解釈について

—LOB Corpus における第3因子の解釈— 高橋 薫

*Text Typology* の研究にコーパスが欠かせなくなった昨今、特に注目されるのは、*linguistic feature* の問題である。LOB Corpus や BNC にあるような文法範疇標示が付されたコーパスにおいては、それらの出現頻度の様相により、テキストジャンル間の類似性を測ることができる。幾分、文法に偏重した *feature* の設定である。

なお、多変量統計解析法を用いることによって、言語学的に解釈可能な複数の尺度上で、



ジャンル間の類型化を行うことができる。

筆者はすでに、LOB Corpus において、第一因子として、literary/colloquial の対比による、統語的複雑さの尺度を、第二因子は意味的な尺度として、specific/general という概念で区別される尺度の存在を見いだした。

今回は、第三の因子が言語学的な説明のつく因子として捉えることが可能かどうかを考察する。これに関しては、LOB Corpus は第三因子として、その寄与率が微少であったため、結果の考察には戸惑うところがあったが、類似した linguistic feature を持つ BNC での多変量統計解析法による分析では（英国留学中に行われたもの）、第一因子、第二因子が LOB Corpus のそれぞれの因子と極めて似かよっていて、第三因子に関しても、同様の傾向が現れた。さらに BNC では、この因子がより高い寄与率で現れたことより、LOB Corpus における第三因子は寄与率が低いながらも、充分考察に値するものだと考えられる。

本報告では、この二つのコーパスの文法範疇標示（タグ）の相違点に注目した後に、BNC との関連で LOB Corpus における第三因子の解釈をするとともに、それによる、テキストジャンルの類型化を行う。

#### ●日英翻訳で見る曖昧な日本語の「など（等々）」 西村 公正

日本語の「など（等々）」は、曖昧な表現語彙とも受け取られるが、優れて豊かな表現には欠かせない副助詞である。英語文法でいう modal adverb という場合の "modal" という概念を借りれば、日本語の「など」は "法"副助詞である。しかも、その表すところは1つに限られず、「緩衝」、「軽蔑」、「謙譲」、「否定の強意」にわたる (cf. 『新潮現代国語辞典』)。

そこで、英語⇄日本語の翻訳をコーパスとして使って、日本語の「など」が英語ではどう表現されるか、反対に英語のどのような表現が「など」で表されるのかを調査・観察する。

まず、Project Gutenberg CD-ROM 所載の The Moon and Sixpence と、『CD-ROM 新潮文庫の 100 冊』所載の中野好夫訳を資料として比較する。さらにこの中野好夫訳が使用した「など（等々）」が、厨川圭子訳、阿部知二訳、瀧口直太郎訳においてはどうなっているかを比較し、英語小説の邦文翻訳における文体論を試みる。

つぎに、『CD-ROM 新潮文庫の 100 冊』所載の『雪国』（川端康成）の「など」とその該当英語訳を検討する。

このような手法により、日本語の「など(等々)」の意味と用法の姿を浮き彫りにするのが本発表の目的である。

◆シンポジウム《コーパスと英語教育》 司会 朝尾 幸次郎

「コーパスの英語教育への利用の歴史」 朝尾 幸次郎

コーパスを外国語教育に利用しようという考えは比較的新しいものである。しかし、その発展は著しい。この発表ではコーパスが外国語教育とどのようにかわるかを、その歴史をたどりながら考察する。

われわれがふつう、コーパスとよぶもの、すなわち、母語話者が書いたり、話したりした言語データをもとにしたコーパスは、言語研究への利用をめざして始まった。しかし、それと同時にコーパスは外国語教育への利用にも利用されてきた。学習辞書、テキストの編集はその典型的な例である。また、コンコードンスラインを直接、教材として利用し、これまでの演繹的な外国語学習から帰納的な学び方への模索が始まっている。

外国語教育への利用に関して、もうひとつの大きな柱は、学習者が書いたり話したりしたことばを系統的に収集した学習者コーパス (**learner corpus**) である。学習者の習得途上の発話を系統的に記録していくことにより、第二言語(外国語)習得の過程を実証的に研究することができる。

しかし、学習者コーパスの構築には一般のコーパスとは異なったむずかしさがある。それは、一般のコーパスが言語習得の終了した段階の発話を記録したものであるのに対し、学習者コーパスは言語習得の過程、すなわち、中間言語の現れを記録したものである。このため、学習者コーパスはエラーを忠実に記録し、学習者の発達段階における情報をくわしく盛り込むことが必要である。

学習者コーパスの構築、利用の歴史は新しく、このため、だれが先陣を切るか、各地で激しい競争が始まったところで、外国語教育研究において、もっともホットな分野のひとつになりつつある。

● 「コーパスによる学習支援システム」 杉浦 正利

本発表では、英語学習者のための、コーパスによる英語学習支援システムを紹介する。この学習支援環境システムは、辞書部とコーパス部とから成り立ち、WWW 上で利用できるようになっている。英語学習者は、WWW ブラウザーを使い、単語を辞書で引き、その例文をコーパスから得ることができる。

語句の検索は、英語でも日本語でもできるようになっているため、英和辞典としても、和英辞典としても利用できる。各単語の例文は、事前に静的に用意してあるのではなく、英文コーパスの中からその単語を含む文を、WWW サーバー上で CGI を通して Perl のプログラムで動的に検索するようになっている。

コーパスから検索された例文の一覧を見て、学習者は、自分の検索した単語が、実際にどのように使われているかを自分で帰納的に学習を行う。例文コーパスは、実際に WWW 上で公開されている英語のページを集めてきたものと、英語学習者専用英語母語話者により作成された日英対訳例文集とが、用意されている。

用例の観察を通して語法や文法規則を学ぶということは、認知心理学的には、手続き的知識の帰納的学習と位置づけられる。こうした学習による知識は、長期記憶に蓄えられ、言語の生成に利用されると考えられている。

インターネットの爆発的な広まりにより、日常的に英語をまったく使わない日本においても、英語を実際に使用する機会がこれまでとは比べものにならないほど増えてきている。本学習支援環境システムは、英語学習者が英語の語法を学習する手助けになるというだけでなく、実際にインターネットを使いコミュニケーションを行なう際に本システムを同時に利用することにより、英語による発信能力を高めことができると考えられる。

## ● 「学習者コーパスの構築」 尾関 修治

本発表では発表者らが現在進めている学習者コーパス構築プロジェクトに関して、(1)データ収集の方法、(2)データの標準化とフォーマットの方法、(3)データの公開方法の3点について解説し、それぞれの問題点を整理する。

データ収集の方法については、現在、メーリングリストなどを利用した学習活動の中で発生したすでに電子的に蓄積されたデータを整形する方法と、新たに広く学習者の作文を収集して入力する方法の2つで進めている。

データの標準化については、電子メールのヘッダなどの不要部分の削除およびプライバシーに関わる部分の削除と、必要なプロフィール情報を残すことの両立をはかっている。フォーマットについては CHILDES を採用している。これは、言語習得過程の発話コーパスの記述に用いられる CHILDES が、多様なエラーを含みしかもそのエラーのコーディングに重要な意義がある学習者コーパスの性質を考慮してのことである。

データの公開は CD-ROM による配布と WWW サーバーによる提供の2つをすすめている。特に後者では検索用ソフトウェアの提供も同時に行う。

● 「学習者コーパスによる事例研究」 野澤 和典

応用言語学の一分野としてのコーパス言語学の研究は、Brown Corpus, LOB Corpus, The British National Corpus, CobuildDirect Corpus などの大規模なコーパスを利用して、国内、国外ともに盛んに行われてきた。これとともに、最近では第二言語（外国語）としての英語学習者の作文やスピーチの大規模データである学習者コーパスの構築と利用が世界各地で試みられ、新しい波として脚光を浴びてきている。

本発表では、世界各地に散在する大規模・小規模の学習者コーパスが外国語（英語）教育・学習にどのように利用されてきたか、いくつかの事例から、それらの理論的背景と実践的手法を紹介するとともに、今後の学習者コーパスを利用した英語教育・学習にどのような新展開が待ち受けているのかを考察する。特に、1990 年に出版された Chris Tribble と Glyn Jones の言語教育における教育資源としてのコンコーダンス利用、さらにその 1997 年の改訂版に示されたコーポラの言語教育への利用に関する考え方、また日本で先端的な研究をしている杉浦(1993, 1998)、朝尾(1998)、投野(1998)などの有益な実践についても触れる。

● 「第二言語習得と学習者コーパス」 原田 邦彦

コーパス (corpus) とコープス (corpse) は響きが似ている。コーパスは実際の言語を記録し集大成したものだが、そのまま保存し分析するだけで再び教育の場で生かさなければ死んだデータであり、その点で英語教員にとってはコープスと変わらない。

言語学でコーパスという言葉が使われ始めて久しいが、第二言語としての英語教育では重要な役割を果たしていない。その理由として、コーパスを作成し利用するにはコンピュータやソフトのツールに関する高度な技術や知識が必要でそれが壁になることが多かったこ

と、また、コーパスが作られても英語を母語とする者のものであったり、理論的な分析に使われるだけで第二言語としての英語教育に応用されなかったことなどが挙げられる。

たとえば、コーパスとそのツールを使うとすぐに平均的な文の長さ (MLU) がわかる。そして上級者の MLU は初級者の MLU よりも長いということがわかる。この分析結果は貴重であるが、それは直接教育効果に結びつかない。初級者にただ長い文を書けと指導しても、内容が伴わないだらだらとした長い文ができるだけだからだ。では、コーパスから得られるデータで語学教育に役立つものはどういったものがあるだろうか。

本発表では科学技術の進歩によりコーパスの作成及び利用が比較的容易になったことを提示し、コーパスから得られるデータにはどのようなものがあり、それが第二言語としての英語教育にどのように役立つかを事例を使って紹介する。