

英語コーパス学会第21回大会

日時 2003年4月26日(土)
会場 関西外国語大学短期大学部中宮キャンパス (〒573-1001 大阪府枚方市中宮東之町 16-1
Tel: 072 805 2801(代), 京阪枚方市駅「北4番」のりばより、京阪バス「関西外大」下車、この間約8分、詳細は <http://www.kansaiuidai.ac.jp/www/j0000001.htm> 参照)

ワークショップ 10:30 - 12:00
《コンコーダンスラインから何を読み取るか》 講師 京都外国語大学 赤野 一郎
徳島大学 井上 永幸
定員 先着50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 摂南大学 今井 光規
2. 開催校挨拶 関西外国語大学・関西外国語大学短期大学部学長 谷本 真人
3. 総会
4. その他

研究発表 第1セッション 13:30 - 14:30

- 司会 東京成徳短期大学 大和田 栄 大手前大学 西村 道信
1. 既存語彙と借用語の使い分け BNCにみる競合する語彙の分布
文化女子大学室蘭短期大学 木村まきみ
 2. 英語とフランス語の<コピュラ+不定詞>構文 パラレルコーパスを利用した分析
大阪女子大学 内田 充美
中部大学 柳 朋宏

休 憩 14:30 - 14:40

研究発表 第2セッション 14:40 - 15:40

- 司会 熊本学園大学 堀 正広 名古屋大学 山下 淳子
3. 学習英和辞典における -ly 副詞の効果的記述についての提案
慶応義塾大学非常勤講師 吉村 由佳
 4. 社会言語学的視点による BNC の解析
豊田工業高等専門学校 高橋 薫
豊田工業高等専門学校学生 白井 翔悟

休 憩 15:40 - 16:00

特別講演 16:00 - 17:20

- 私のコーパス言語学研究への道
- 司会 摂南大学 今井 光規
講師 大阪大学名誉教授 齊藤 俊雄

閉会の辞

関西外国語大学短期大学部 西村 公正
《懇親会 17:50 - 20:00 会費5,000円》
司会 関西外国語大学 田中 廣明

英語コーパス学会 (Japan Association for English Corpus Studies)
会長 今井光規 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: 088-656-7129 E-mail: jun@ias.tokushima-u.ac.jp 郵便振替口座 00940-5-250586
URL <http://muse.doshisha.ac.jp/JAECs/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5,000円 学生4,000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 21 回大会レジュメ

ワークショップ《コンコーダンスラインから何を讀み取るか》

(講師 赤野 一郎・井上 永幸)

コーパスによる語の分析は KWIC コンコーダンスに始まり, KWIC コンコーダンスに終わると言っても過言ではない. この表示形式は, ある語が隣接する語とどのように結びついているか, その結びつきの傾向を示してくれる. 従って語と語の典型的結合, すなわち語のコロケーションを明らかにするのに最適である.

本ワークショップの前半(赤野講師担当)では, KWIC コンコーダンスの特徴を解説したのち, データの基本的読み方を実習する. 次に語のコロケーションを手がかりに, どこまで分析対象を拡大できるかを, 具体的データを提示しながら参加者ととも考えてみたい.

ワークショップ後半(井上講師担当)では, corpus-driven な手法を取り入れることで, これまでの辞書・教材から抜け落ちていた部分の検証が可能になることを示す. さらにその応用面として, corpus-driven な手法による効率のかつ実践的な語彙指導が期待できることを示すとともに, KWIC コンコーダンスを使った教材作成の方法を紹介する.

大会開催校の事情で当日コンピュータを使用することができないとのことなので, ハンドアウトとパワーポイントのスライドを併用した形でワークショップを行う. 使用するコーパスは WordBanksOnline, コンコーダンスは Txtana を予定している.

研究発表 第 1 セッション

既存語彙と借用語の使い分け BNC にみる競合する語彙の分布

(木村 まきみ)

本発表では既存語彙と新語彙(主に借用語)の競合の実態を BNC を用いて明らかにする. 英語には本来語やラテン語・ギリシャ語由来の語などの古い語彙が存在する一方で借用・造語などによって日々新語彙が追加されている. 語彙借用は主に新しく導入された概念が既存の語彙では端的に言い表せない場合に起こるが, 既存の語彙と同じ意味を持つ語が借用

されたり, 借用後に意味が変化して既存の語彙と同じ意味を持つようになることもある. そのような場合には既存語彙と借用語の間で競合が起こり, 二語が共存, あるいは一方が他方を駆逐する道を辿る.

本研究の目的はそのような競合する語彙がどのような共存関係を築いていて, それぞれの用法にどのような類似点・相違点があるのかを解明することである. BNC で競合する語彙を検索した結果, 以下のような共存関係があることが明らかになった.

- ・既存の語彙と新語彙が同義語として共存し, どちらも同じように使用される.
- ・それぞれが意味を特化して区別して使用されるようになる.
- ・新語彙は使用範囲が限られ, 同義の既存語彙はより広範囲に用いられる.
- ・新語彙が既存語彙に取って代わり使用範囲を拡大する.

本発表では上記の各タイプの共存関係を事例を挙げて概説し, BNC から検出した用例を様々な側面から分析した結果を述べる. 借用語の語法やコロケーションは競合する既存語彙や類似した語彙の語法・コロケーションを踏襲する傾向があること (e.g., blitz on ~, tsunami of ~), その一方で, 英語辞書や類語辞典では同義だと定義されている語の間にも使用範囲や頻度, 指示対象, 意味やニュアンスなどに相違が見られるものがあることなどを示す. 今回は日本語からの借用語と, 英語に既存の語彙と同じ意味を持つ語が多く借用されることが分かっているドイツ語からの借用語を対象とした. この 2 言語から 19~20 世紀に借用された語彙とそれらと競合する既存語彙 (e.g., angst vs. anxiety, ersatz vs. substitution, doppelganger vs. wraith, tsunami vs. tidal wave, kami-kaze vs. reckless, tycoon vs. magnate/lord/king) を対比し, 特徴を論じる. また, 既存語彙を用いて翻訳借用された語句と直接借用された語彙の両方が使用されている場合 (e.g., Gestalt vs. configuration, shiatsu vs. acupuncture, kanban vs. just-in-time) についても同一の事物を指す 2 つの表現があることから, 一種の新

旧語彙の競合と見なし、その使用実態についても言及する。

英語とフランス語の<コピュラ+不定詞>構文 パラレルコーパスを利用した分析

(内田 充美・柳 朋宏)

本発表では、英語とフランス語における<コピュラ+不定詞>構文の類似点・相違点について、パラレルコーパスの特徴を活かして収集した資料に基づき、両言語の持つより一般的な特徴の観点から論じる。調査に使用したパラレルコーパスは、Official Journal of the European Community に掲載された、書面による質議応答のデータからなる JOC コーパスと、カナダ議会の議事録データを収めた Hansard コーパスである。

当該の構文は、英語では、(1) *The budget for the programme, which is to run until 1996, is US\$ 2,9 billion.*(JOC Corpus) のように *be* 動詞 (*is*) + 不定詞 (*to run*) という形式が用いられる。これと同じように、フランス語においても、(2) *... une hausse des prix des médicaments est à prévoir* (lit.) *... an increase of the prices of drugs is to expect* (Hansard Corpus) のように、動詞 *être* (*est*, *be* 動詞に相当) + 不定詞 (*à prévoir*) という形式が用いられる。以下、英語の<コピュラ+不定詞>構文を *be to* 構文、フランス語の<コピュラ+不定詞>構文を *être à* 構文と呼ぶ。これら2つの構文は、上述のように形式的な共通性に加えて、いずれの構文も、「義務・必然」などのモダリティを表現するという点において、意味的にも共通する特徴を持っている。

これら2つの構文は、形式的にも、意味的にも共通する特徴を持っているが、異なる側面もある。たとえば、*be to* 構文では、*be* 動詞のあとにくる不定詞が(1)のように形式的に能動形である場合、基本的に能動的な意味を持つものとして解釈されるのに対し、(2)のような *être à* 構文では、形式的には受身形ではないにもかかわらず、受身の解釈が可能である。つまり、*be to* 構文では、モダリティ的な意味は構文そのものが、また、受身の意味は構文以外の要素が担っているのに対し、*être à* 構文では、モダリティ的な意味だけでなく受身の意味関係も、構文そのものが担っていると考えられることができる。

また、両構文の使用傾向においても差異が観察される。パラレルコーパスを調査した結果、*be to* 構文の使用頻度に比べて、*être à* 構文の頻度は著しく低いことがわか

った。

本発表では、英語とフランス語における<コピュラ+不定詞>構文に関わる現象を、「義務」モダリティの表現方法、態の選択、主語の選択など、両言語の一般的な特徴と関連づけながら検証していく。

研究発表 第2セッション

学習英和辞典における -ly 副詞の効果的記述についての提案

(吉村 由佳)

-ly 副詞は対応する形容詞の派生形として扱われ、英和・英英を問わず、辞書では従来あまり十分な記述が見られなかった。形容詞の最後に ~ly と書き足すだけで、英和辞典では訳語すら出ていないものも多い。これでは英語学習者は当該の副詞の意味を理解する情報を十分に得られないばかりか、近年その必要性が求められているコミュニケーションを重視した発信型の情報はまったく得ることができない。actually, certainly などの語用論上の重要な働きを指摘されている一部の -ly 副詞には十分な記述スペースが当てられるようになったが、こうした副詞以外にも、コーパス検索の結果を精査すれば、明らかな語法上の特徴が見取れる -ly 副詞が数多くある。

本発表は、近年出版された複数の学習英和辞典の -ly 副詞の記述内容を検討した上で問題点を指摘し、コーパスを利用した分析結果を盛り込むことで、どのように効果的な記述が可能となるかを探るものである。具体的手順は2つに分けられる。第一の手順は、客観的に母国語話者が頻繁に使用する -ly 副詞を選び、その語についての学習英和辞典の記述内容を調査することである。そのため、コーパスに基づいた語彙リストが公開・出版されている BNC を使用し、高頻度の -ly 副詞(50語)を選び出した。次に、これらの -ly 副詞が学習英和辞典ではどの程度のスペースを割いて記述されているかを調べた。使用した辞典は2001年から2003年にかけて出版された、9万語から10万語程度の学習英和辞典5冊である。この5冊の辞書のいずれにおいても記述スペース(具体的には行数)が少なく、内容の再検討が必要だと思われる数語の -ly 副詞を選択し (currently, effectively, increasingly, mainly など)、その内容を語義・コロケーション・例文・語法などの面から比較・考察した。第二の手順は、コーパスからのデータを使用して、

-ly 副詞の辞書記述を改善する方法を模索することである。そのために品詞タグの検索も可能な Wordbanks-Online を使用し、先に取り上げた -ly 副詞のコンコーダンスラインの分析を行なった。本発表の中心は、このコンコーダンスラインから読み取った情報をもとに、今までの辞書記述の内容との一致・相違点、記述の不備などを考察することにある。そして、次の5点に着目した辞書記述の改善が必要であることを、前述の -ly 副詞を例に取り上げて論じる。

- (1) 学習英和辞典に必要な語義の選択
- (2) 文(修飾)副詞の効果的記述方法
- (3) コロケーションを重視した例文選択
- (4) 語法を明示する手段としての選択制限の採用
- (5) 類義語との比較情報

さらには、考察した結果をどのようにして英語学習者に役立つ形で辞書記述に盛り込むか、その具体的な例示も行なう。

社会言語学的視点による BNC の解析

(高橋 薫・白井 翔悟)

本研究の目的は社会言語学の視点からイギリス英語の言語特性を分析することにある。書き言葉と話し言葉のドメインからなる British National Corpus (BNC) の各語に付加された文法範疇標示(タグ)の頻度を多変量統計解析法により分析し、その解析結果に言語学的な解釈を与えるという手法をとる。本発表では、これまで得られた解析結果の解釈を高橋が行う。また一連のデータ加工法について白井が詳説する。

社会言語学者 Coulmas (1997)* は、community を space と time という尺度により考察すること、さらに言語における社会集団的相違を計量的に分析する必要性を提唱している。

約1億語からなるイギリス英語の書き言葉と会話文からなる BNC は、複数のジャンルの書き言葉や、さまざまな条件下での会話文の中に固有の言語学的特質を持つことが予想される。数量化 III 類を用いることにより、すでに書き言葉のジャンル間の類似性の度合いとタグ頻度との関係が判明している。また、判別分析の手法を用いることで、タグの頻度あるいは特定の語句の頻度データの分析が、ジャンル間の判別をある程度可能にす

ることも判明した。

ここでは、BNC の各テキストに付加された記述情報により、書き言葉のサブコーポラを再構築して、数量化 III 類によりサブコーポラ間の類似性を測り、さらにタグ頻度との関連より言語特徴を見出し、言語学的な解釈を与える。

また、話し言葉については、話者の年齢、社会階層、性別、地域等の区分により、仮想的な speech community を構築して、上記と同様の解析を行う。これにより、Coulmas が捉える community の概念にいかなる提言を与えることができるか考察する。

特に今回は、BNC の話し言葉の指標である社会階層について解析を行った結果、社会階層のレベルと文章難易度の指標に対応関係が見られた。また、各階層を特徴づける語彙を抽出し、その単語頻度について判別分析にかけたところ、その単語群が9割程度の精度で社会階層のレベルを特定できることが判明した。

なお、発表の後半では、これらの解析を行うにあたっての解析手法の手順について述べる。特に、テキストデータの初期加工について、C 言語とエクセル使用の観点から概説する。

*Coulmas, Florian. (1997) *The Handbook of Sociolinguistics*. Oxford: Blackwell.

特別講演

私のコーパス言語学研究への道

(齊藤 俊雄)

1993年に英語コーパス研究会(現英語コーパス学会の前身)が発足した。この組織は日本におけるコーパス言語学研究の牽引車の役割を果たしてきたが、その創設に動いた一人である私がコーパス言語学の研究に携わるようになり、日本におけるその普及・発展に努めるにいたった軌跡をたどる。

私は文科系がコンピューターの利用できない時代に研究生活を始め、主として英語の史的变化の研究に携わってきたので、必然的に実証的な研究方法を取った。したがって、本の形をした文献資料を(今で言う)コーパスとして利用して、用例を手作業で抽出して統語的変遷を実証してきた。50歳代になってコンピューターにめぐり合い、コーパス言語学的アプローチを取るようになったわけである。したがって、この講演では史的コーパス言語学を中心に回顧することになる。