

英語コーパス研究会第9回例会

日 時 1997年4月19日(土)

会 場 同志社大学田辺校地(京都府綴喜郡田辺町[4月より京田辺市])

→京都(大阪・奈良)方面から田辺校地への交通路

恵道館2階 マルチメディア講義室(KD-202)→田辺校地地図

(1) JR学研都市線「同志社前」下車 徒歩10分

(2) 近鉄京都線「興戸」下車 徒歩15分、あるいは

「新田辺」下車、バス10分、タクシー6分 →バス時刻表

ワークショップ 10:30-12:10 恵道館1階 情報処理教室(KD-103)

《WordSmithを使ったコーパス検索の技法》(梗概)

講 師 赤野一郎(京都外国語大学)・井上永幸(島根大学)

先着 40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・FAX・郵便で事務局まで)

受付開始 12:30

開会の辞 13:00 姫路獨協大学 齊藤俊雄

総 会 13:10-13:30

研究発表 13:30-15:30

1. 『結果』を表す現在分詞構文について

—マッキントッシュ環境での電子テキスト処理—(梗概)

大阪大学大学院生 内田充美

司会 三重大学 綾野誠記

2. 抽象名詞の修飾と不定冠詞の関連性について  
－COBUILD Direct のデータから－ (梗概)

聖母女学院短期大学 小寺正洋

司会 龍谷大学 正保富三

3. 動詞の分布に基づく LOB Corpus の構造 (梗概)

徳島大学 中村純作

司会 大阪大学 田畑智司

<休憩 15:30－15:45>

シンポジウム 15:45－17:50

《コーパスを基礎としたコロケーション研究》(梗概)

司会 大東文化大学 山崎俊次

「ディクンズにおける副詞のコロケーション」(梗概)

講師 熊本学園大学 堀 正広

「小辞と動詞の位置について－Penn-Helsinki Corpus を検索して」(梗概)

講師 豊田工業高等専門学校 神谷昌明

「英語変異の基になるコロケーション

－LOB Corpus と Wellington Corpus の比較分析－」(梗概)

講師 大東文化大学 山崎俊次

閉会の辞

同志社大学 西納春雄

《懇親会 18:00-19:30 紫苑館 教職員ラウンジ 会費 4,000 円》

英語コーパス研究会 (Japan Association for English Corpus Studies)

会長 齊藤俊雄

事務局 657 神戸市灘区鶴甲 1-2-1 神戸大学国際文化学部  
西村秀夫研究室

TEL/FAX 078-803-0737 E-mail: (E-mail address deleted)

郵便振替口座 00940-5-250856

◆例会当日、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）。

◆「当日会員」としての参加も受け付けております（1,000 円）。

## 英語コーパス研究会第9回レジュメ

### 目次

#### ◆ワークショップ

《WordSmith を使ったコーパス検索の技法》(講師 赤野一郎・井上永幸)

#### ◆研究発表

●「『結果』を表す現在分詞構文について—マッキントッシュ環境での電子テキスト処理—」  
(内田充美)

●「抽象名詞の修飾と不定冠詞の関連性について—COBUILD Direct のデータから—」(小寺正洋)

●「動詞の分布に基づく LOB Corpus の構造」(中村純作)

#### ◆シンポジウム

《コーパスを基礎としたコロケーション研究》(司会 山崎俊次)

●「ディケンズにおける副詞のコロケーション」(堀 正広)

●「小辞と動詞の位置について—Penn-Helsinki Corpus を検索して」(神谷昌明)

●「英語変異の基になるコロケーション—LOB Corpus と Wellington Corpus の比較分析」  
(山崎俊次)

◆ワークショップ《WordSmith を使ったコーパス検索の技法》(講師 赤野一郎・井上永幸)

このワークショップでは、いくつかの言語分析プログラムを統合した、Windows(3.1 以上)対応の WordSmith(OUP)というソフトの実習をおこないます。このソフトは Concord, WordList, KeyWords, Dual Text Aligner, Splitter, Text Converter の6つのモジュールからなっていますが、特定の語の文脈上の振舞いを教えてくれる Concord, テキストの全単語

の一覧表を作成する **WordList**, テキスト中のキーワードを探し出す **KeyWords** の操作方法を学びます。時間があればコンコーダンスラインの読み方のコツについても説明します。使用するコーパスは **MicroConcord Collection** を予定しています。出席者には **WordSmith** のデモ版を配布します。

◆研究発表

●『「結果」を表す現在分詞構文について—マッキントッシュ環境での電子テキスト処理—』  
(内田充美)

この発表は、現在分詞構文のうち、分詞節が主節に対して後置され、その分詞の意味上の主語が先行主節の内容であるものを研究対象とする。たとえば、次の例文 (1)、(2)において

(1) With the exception of the Salang Highway, roads into the city are cut, resulting in shortages of bread, diesel fuel, sugar, kerosene and other basics; . . . (TIME 1989)

(2) But in those with glaucoma, the drainage canals are somehow blocked, leading to an increase in pressure. (TIME 1990)

それぞれ、先行する主節全体が分詞節の動詞 **result in, lead to** の意味上の主語になっており、分詞節は主節の表す内容がもたらす結果を示すものである。学术论文やジャーナリズムなどの文章では比較的なじみのある用法だが、絶対的な頻度があまり高くないこともあって、分詞構文についての先行研究においては、あまり関心を寄せられておらず、例外的・周辺のものとして考察の対象外とされることもある。

しかし、まとまった量の電子テキストを検索することによって得た類例を観察してみると、このタイプの分詞構文はパラグラフの最後に現れやすい、先行する主節に能動態の動作表現が現れにくい、など一定の傾向をうかがうことができる。この研究発表では、一見特殊に見えるこの用法が、一般的な慣用表現と同様に、動詞の意味特性と、広義の文脈に支えられて、独自の機能を有していることを実例に基づいて検証する。

各種テキストの CD-ROM 化、インターネットの普及などに伴って、実際に使用されている英語を記録した、電子テキストが比較的容易に入手できるようになっている。しかし、その膨大なデータを活かすには、適切な検索手段が不可欠である。一般の電子テキストを言語研究に利用しようとする場合、LOB コーパスのような既成のコーパスを利用する場合とは異なり、品詞 **tag** が付加されていない点、テキストごとに編集スタイルが異なる点など克服

すべき問題がいくつかある。この発表では、電子テキストを検索して目的の用例を収集しようとする際に直面する問題点や、収集した用例を効率的に管理するための方法などについても触れてゆく。

●「抽象名詞の修飾と不定冠詞の関連性について—COBUILD Direct のデータから—」(小寺正洋)

本発表は、5千万語のコーパスをもつ COBUILD Direct のデータを利用したものであり、発表の内容に入る前に、COBUILD Direct についての簡単な紹介をする。COBUILD Direct は「コウビルド英英辞典」でお馴染みのコーパス、Bank of English (1996年6月現在3億2千万語)のうち5千万語(当初2千万語)をインターネット上で telnet を通じて利用できるようにした言語研究用コーパスであり、その利用方法やコーパスの内容、特徴、完成度、さらに利用する上で注意する点などを簡単に紹介する。その後、発表の本題である抽象名詞と不定冠詞の関連性を COBUILD Direct のデータを基に検討する。

抽象名詞は基本的に不可算であるが、'difficulty'や'hope'のように可算用法と不可算用法の両方を持つものと、'education'や'knowledge'のように一般的に不可算用法しか持たないとされるが、不定冠詞を伴うことのできるものがある。本発表では後者のタイプの抽象名詞がいかなる場合に不定冠詞を伴うかについて従来行われてきた説明を一つずつ COBUILD Direct のデータによって検証していく。

抽象名詞が不定冠詞を伴う条件としては、従来から統語論的観点、意味論的観点の両方から説明がなされてきたが、以下の5つに整理する事が出来る。

1. 「人間の持つ感情や精神活動を指示する名詞」

(e.g. dislike, distrust, education, knowledge, sensitivity, understanding)

2. 「何らかの形で意味が限定される場合」

(一般的には形容詞などの修飾語との関連性から説明される)

(修飾表現が長いほど不定冠詞を伴う可能性が高いといった説明もある)

3. 「種類を表す場合」

(一般的には形容詞などの修飾語との関連性から説明される)

4. 「始まりと終わりを持つ事象を指示対象にする場合」(e.g. rest, silence, sleep)

5. 「具体的でインフォーマルなコンテキストの場合」

第1の条件については検証の対象とする名詞の限定が難しいので、本発表では、Quirk et al.(1985)と Swan(1995)が例として挙げている上記の6つの抽象名詞に調査の対象を限定し、それぞれの名詞について、残りの4条件によって不定冠詞の用法がいかに関係を受けるかを検証する。修飾語に関しては、直前に形容詞を伴う場合、直後に'of'を伴った修飾表現を伴う場合、関係詞節による修飾表現を伴う場合などについて不定冠詞との関連から検討し、抽象名詞と不定冠詞の関係は修飾表現との関連性からだけでは説明できないことを示す。また、「種類」とか「始まりと終わりを持つ事象を指示対象とする」といった条件も不十分であり、「変種」とか「時空間的特定性」といった要件を意味限定の枠として考慮する必要があることを提案する。

#### ● 「動詞の分布に基づく LOB Corpus の構造」(中村純作)

語彙や文法範疇の頻度分布とテキストとの関係は多変量解析の対象として扱われており、その主な手法には Biber の因子分析、Burrows の主成分分析、発表者等の用いている林の数量化 3 類等が挙げられる。カテゴリー間、あるいはサンプル間の相関係数を計算した上で、更に複雑な処理を必要とする因子分析や主成分分析では、あまり多くのサンプルやカテゴリーを分析することは、メモリー、計算に要する時間等の関係で容易ではない。本発表は、古典的ではあるが Tagged LOB Corpus の Token 数にして約 10 万語 におよぶ動詞の 15 のカテゴリーにわたる頻度分布を基に林の数量化 3 類によりその構造を見ようとする Pilot Study である。

分析に当たっては、まず、Tagged LOB Corpus の各ジャンル毎に文法指標 VB で始まる語 (VB, VB", VBD, VBG, VBN, VBZ) を全て抽出した上でその頻度を計算、15 のジャンルにわたる頻度表を作成し、更に、この頻度表を Lemma 毎に整理したものに、林の数量化 3 類を当てはめた。結果として与えられた数量の内 2 軸までの数量を使い、ジャンル及び動詞の位置関係を 2 次元の空間に図示することにより視覚的に各々の特徴を捉えた。ジャンルでは情報散文と創作散文が見事に分離されるほか、新聞英語、科学論文、Governmentese 等のジャンルも特徴的であり、当然動詞の位置関係はこれらのジャンルの特徴を反映したものである。

LOB Corpus の動詞の「宇宙」を垣間見せるように、約 3,000 におよぶ動詞の Lemma が次々コンピュータ画面に表示される様子は、使用したコンピュータとプログラムの関係で発表時には提示不可能なので、希望者には休憩時間などを利用してお見せしたい。

◆シンポジウム《コーパスを基礎としたコロケーション研究》(司会 山崎俊次)

コンピュータ技術の発展に伴い、研究の対象になるコーパスも Brown コーパスや LOB コーパスから文学作品のコーパス、さらに自作のコーパスと多種多様になってきた。今回のシンポジウムではディケンズの文学作品をコーパスとして副詞のコロケーションの研究、Penn-Helsinki Corpus を検索して小辞と動詞の位置についての研究、また LOB コーパスとそれに構成が類似しているニュージーランド英語の Wellington Corpus を比較分析して英語変異の基礎としてのコロケーション研究を発表する。コロケーションの定義も多様であり、コロケーション研究の方法もまた様々である。本シンポジウムによってコロケーション研究の一端がうかがえれば幸いである。

●「ディケンズにおける副詞のコロケーション」(堀 正広)

J. R. Firth は "Modes of Meaning" (1951) の中で、collocation の共時的、通時的研究の重要性を主張した。これまでの研究では、Sydney Greenbaum の Verb-Intensifier Collocations in English (1970), Ulf Baklund の The Collocation of Adverbs of Degree in English (1973), A Dictionary of English Collocations (1994), COBUILD の English Collocations on CD-ROM (1995) などがあるが、いずれも現代英語が対象で、文学の言語の collocation の研究は、Firth 自身が、さかんにその重要性を唱えたにも関わらず、ほとんどなされていない。

本発表では、Dickens の英語を collocation の点から見ていき、これまで指摘されなかった Dickens の英語の特徴を明らかにしていきたい。ここでは、主に副詞の collocation を扱い、18 世紀の代表的な小説の言語と比較しながら、Dickens の豊かで特徴的な collocation や collocational pattern および collocability の問題を論じる。

Dickens の作品は Levit の Dickens on Disk (Hall Design, 1995) や Like the Dickens (Multimedia PC, 1994) を、18 世紀の小説は Eighteenth-Century Fiction (Chadwyck-Healey, 1996) を、通時的な collocation の調査には、コーパスとして The OED2 on CD-ROM を用いた。

●「小辞と動詞の位置について—Penn-Helsinki Corpus を検索して」(神谷昌明)

初期中英語(eME)において、小辞(adverbial particle)は動詞の前後に位置することが可能であった。即ち、P(...)V型とV(...)P型である。Hiltunen(1983)によれば、古英語(OE)では、主節、従節、共にP(...)V型が一般的であった。MEに入り、V(...)P型が優勢になり、後期中英語(IME)では、P(...)V型はほとんど見られなくなる。タグ付きコーパス、Penn-Helsinki Parsed Corpus of Middle English を検索することにより、ME 散文において、P(...)V型が、どのように衰退していったのか検証する。また、MEにおける小辞の頻度や、できれば小辞と結びつく動詞の種類についても触れてみたい。

注) Penn-Helsinki Corpus は、Helsinki Corpus の diachronic part の ME 部分にタグ(文法標識)を付加したものであり、総語数は約 50 万語。Penn は散文のみに限られるが史的コーパスにタグを付けた世界最初のコーパスである。また、タグ付けは生成文法理論に基づいている。

●「英語変異の基になるコロケーション—LOB Corpus と Wellington Corpus の比較分析」  
(山崎俊次)

コロケーションの定義は様々であるが、本発表ではコーパスの構成が類似している LOB Corpus とニュージーランド英語の Wellington Corpus から形容詞+名詞のコロケーションを抜き出して、それぞれの文化的、歴史的、地域的な特徴の検索を試みる。そしてその形容詞+名詞のコロケーションが英語変異の基礎となっているか分析するものである。二つのコーパスはかつては同じ大英帝国の英語文化圏でありながら、特にコロケーションには興味ある特徴が見られる。また本発表では、それぞれの品詞の頻度数と形容詞+名詞のコロケーションのカテゴリ間における現れかたの違いについて、ICE Tagset を使って出した数値を使って二つのコーパスにおけるコロケーションの比較分析を試みたい。