

第5回例会プログラム

日 時 1996年4月8日(土) 13:00-

会 場 大阪大学大学院言語文化研究科棟大会議室

総合司会 神戸大学 西村 秀夫

開会の辞 13:00 大阪大学 齊藤 俊雄

総 会 13:10

研究発表 13:30-14:50

1. 「Helsinki Corpus における"DO"の数量的考察」

佐野女子短期大学 保坂 道雄

司会 日本大学 塚本 聡

2. 「口語英語研究のためのマルチメディア型

データベースの開発」 専修大学 佐藤 弘明

司会 同志社大学 西納 春雄

シンポジウム 15:05-17:30

《Bank of English を使ったコーパス言語学研究》

司会 島根大学 井上 永幸

「Bank of English における“woman”の世界」

講師 徳島大学 中村 純作

「コーパスに見る談話辞—I mean を例に—」

講師 京都外国語大学 藤本 和子

「間接疑問文を導く if と whether—Bank of English のデータから—」

講師 京都外国語大学 巳波 義典

「話し言葉における many について—Bank of English を使った分析—」

講師 島根大学 井上 永

幸

閉会の辞 17:30

大阪大学 今井 光規

懇親会

17:40-19:00

会場 大阪大学言語文化部大会議室

第5回例会発表レジュメ

◇研究発表

○保坂 道雄「Helsinki Corpus における“DO”の数量的考察」

助動詞 do についての研究が近年再び、脚光を浴びているようである。1953年の Ellegard の研究により、一時はその起源と発達に関する問題は決着が付いたかのようであったが、Denison (1985), Stein (1991) 等の研究により、再びその問題に火が付いた。日本でも昨年、中尾祐治、天野政千代等により、『助動詞 Do』(英潮社、1994) が出版され、その起源・発達・機能について議論された。

今回の発表では、英語の史的研究に有効と思われる Helsinki Corpus を用いて、do の起源と発達の様子について、数量的に考察してみたい。Rissanen (1991) において、1500年以降についてはかなり詳細に述べられているので、ここでは、それ以前の do の様子について中心に扱いたい。特にどのような構文上で使われる率が高かったか、その数量的変化はどうだったか等に注目し、その起源に関して何らかの新しい視点が得られればと考える。また、do の発達を通じて、機能範疇の成り立ちについても言及できたらと思っている。

○佐藤 弘明「口語英語研究のためのマルチメディア型データベースの開発」

筆者が開発した「英語字幕データベース」は、検索された語句が使われる場面をレーザーディスク映画で再生し、その語句がどのような場面で使われ、どのように発音されるかを確認できるマルチメディア型のデータベースである。

筆者は、辞書で調べてもわからない口語英語の表現があると、これまで英語のネイティブスピーカーに尋ねていた。しかし、彼らの言葉による説明では、語句の使われる状況を十分に理解できないことが多かった。現在は、「英語字幕データベース」を使って、調べたい語句が使われる場面を見ることができるので、話し手・聞き手の年齢、性別、顔つき、服装、身振りなど、細かな情報までを直接確認できる。さらに、古い映画と新しい映画で使われる英語を比較することができるので、口語英語の変遷も調べることができるようになった。長年の英語圏での生活経験がなければ獲得できないと考えられている、このような情報を、「英語字幕データベース」では、映画から調べることができる。

本発表では、「英語字幕データベース」のシステムを紹介し、口語英語の研究と英語教育での活用について述べる。

◇シンポジウム

《Bank of English を使ったコーパス言語学研究》（司会 井上 永幸）

Brown Corpus, LOB Corpus, Helsinki Corpus をはじめとする主要コーパスの CD-ROM による配布、昨年の British National Corpus の完成、さらに Bank of English の 2 億語達成と、コーパス言語学の研究環境は次第に整いつつある。British National Corpus のヨーロッパ圏以外での利用や、Bank of English のインターネットを通じてのアクセスなど、今後のさらなる発展が期待される場所である。

本シンポジウムは、コーパスを使った言語研究の可能性の一端を示そうとするものである。今回は Bank of English のデータを使って、それぞれの発表者のお得意の分野を引き受けていただいております、図らずもコーパス言語学研究の多様性を示す結果となった。

○中村 純作「Bank of English における“woman”の世界」

多変量解析による英文テキストの類型分析には、因子分析やクラスター分析を用いた Biber の手法や、主成分分析を用いた Burrows, Tabata の手法、あるいは林の数量化 3 類

を用いた Nakamura, 高橋の手法が見られる。これらは、いずれも、ある特定の 카테고리（例えば文法指標、代名詞等）のコーパス内での頻度分布に基づいて、直接テキスト間の相互関係を見ようとしたものであるが、同じ方法がある特定の 카테고리に関連した別の 카테고리（例えば、あるキーワードとその共起語）のコーパス内での分布に基づいた間接的な類型分析にも応用が可能である。

本発表では、Bank of English の内、1992 年時点でのサブコーパス 4 種類(Book Corpus, Times Corpus, Spoken Corpus および BBC Corpus)での “woman” の共起語の続き分布に数量化 3 類をあてはめ、サブコーパス間の関係、共起語間の関係、サブコーパスと共起語との関連を考える。その際、数量化 3 類の結果として与えられる数量をもとに、Bank of English における “woman” の世界を視覚的に再現し、単語 “woman” が Bank of English のサブコーパスにおいてどのように扱われているかを明らかにしたい。

○藤本 和子「コーパスに見る談話辞—I mean を例に一」

Bank of English には、書き言葉だけでなく話し言葉のデータも多く含まれている。話し言葉に関しては、現時点ではまだ London-Lund Corpus のように詳細な音声表記は付されていないものの、テキスト量が豊富であるため、多くの用例を抽出することができる。本発表では、Bank of English の話し言葉のデータを用いた分析の 1 例を、談話辞 I mean を中心に紹介する。

談話辞 I mean は、その振る舞いや意味が非常に多様であるため、用法については、多くの学習英和辞典で、挿入的に用いられ先行する語句の訂正等に用いられるといった程度の説明がなされており、どのような言語環境で用いられ、どのような談話的機能を果たすものかについては、あまり具体的に解説されてはいない。本発表では、Bank of English から抽出された多数の用例を基に、談話辞 I mean が、典型的には、どのような言語環境で用いられ、どのような機能を果たすものかを示したい。

○巳波 義典「間接疑問文を導く if と whether—Bank of English のデータから—」

ともに間接疑問文であると解釈される if 節と whether 節の間にも、統語的なふるまいに相違があることはよく知られている。例えば if 節は前置詞の補部や文の主語の位置には現れないなど、whether 節よりも課せられる制限が強い。この両方の節が生起できる共通の環境は、その節の時制が定形でかつ動詞の補部になっている場合に限られる。しかしながら、この条件を満たしていれば常に if と whether の交替が可能というわけではなく、そこには何らかの他の制約が働いているように思われる。

今発表はこの **if** と **whether** の交替に関してそこにどのような要因が係わっているのかを考察することを目標とし、その手段として、**Bank of English** から検索、抽出された様々なデータを検討してゆくこととする。

○井上 永幸「話し言葉における **many** について—**Bank of English** を使った分析—」

発表者は以前から **many, much, a lot of, few, little** などの数量を表わす語句について、それらが生起する文体や文の種類を始めとするさまざまな制約に強い興味を覚えるとともに、従来からの説明に疑問を感じるものがしばしばあった。

本発表では特に話し言葉における **many** に焦点を当て、**Palmer (1938)** などに始まる「**many** は一般に、(a) 疑問文・否定文・条件文で、(b) 主語または主語の一部として、(c) **a good [great] many, too many, so many, as many, how many, many a** などの特定の連語で用いられ、それ以外の場合は **a lot (of), lots of, a (large) number of** などが代わりに用いられる」という説明の再検討を試みてみたい。分析対象とした **Bank of English** の **Spoken Corpus** はラジオの聴視者参加番組、講義、会議、個人の電話の通話、くだけた会話からなる 400 万語の英国英語コーパスである。