

第7回例会プログラム

日 時 1996年4月20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 京都外国語大学 赤野 一郎

開会の辞 13:00 会長 齊藤 俊雄

総 会 13:10

研究発表 13:30-14:50

1. 「Word Basic によるプログラミング活用例

—Micro-OCP との比較」 追手門学院大学 稲木 昭子

司会 大手前女子短期大学 西村 道信

2. 「London-Lund Corpus における pauses と silence fillers

—その頻度の言語心理学的分析」 神戸市外国語大学大学院 森 庸子

司会 帝塚山短期大学 梅咲 敦子

シンポジウム 15:05-17:30 《Helsinki Corpus をめぐって》

「総論—現状と展望」 司会・講師 神戸大学 西村 秀夫

「Penn-Helsinki Corpus について」 講師 佐野国際情報短期大学 保坂 道雄

「コーパス検索ツール —いくつかの事例」 講師 日本大学 塚本 聡

「有効性の検証—中英語期における動名詞の発達について」

講師 日本橋女学館短期大学 伊藤 礼子

「検索とデータ処理の方法について」 講師 大阪大学大学院 永尾 智

閉会の辞

東海大学 朝尾幸次郎

懇親会 17:40-19:30

会場

大阪大学待兼山会館

第7回例会発表のレジュメ

○ 稲木昭子

「Word Basic によるプログラミング活用例—Micro- OCP との比較」

ルイス・キャロル(Lewis Carroll) の『鏡の国のアリス』(Through the Looking-Glass and What Alice Found There)を一読すると、強意副詞の1つである **very** が多用されていることに気づく。ところがこの **very** はイタリック体で出てくる。イタリック体は一般にその単語を強調するために使用されているが、言葉の魔術師であるキャロルの手にかかると、通常の使用だけではなく、意識的な利用がなされていると考えられる。

この点を明らかにするため、原文における、1) **very** の出現頻度数、2) **very** の含まれる文、3)イタリック体の **very** の頻度数、4)イタリック体の **very** の含まれる文、をコンピュータで抽出し、検討することが必要になる。1)および2)については、テキストファイルを対象とする従来の Micro-OCP (Oxford Concordance Program)で分析可能である。しかし3)および4)の文書ファイルに関係するものは分析できない。そこで Microsoft Word の Word Basic を活用して、1)から4)を抽出するプログラムを作成することにする。今回の発表では、作成したプログラムの内容について詳細を述べ、これを用いて抽出した結果を Micro-OCP のものと比較検討することにする。

○森 庸子

「London-Lund Corpus における pauses と silence fillers

—その頻度の言語心理学的分析—」

London-Lund Corpus では、100 の口語テキストが 12 のセクションに分類されている。Silent pause (SP)と、non-verbal filler(例: [a:], [a:m])や verbal filler (例 well, you know) との相互関係を探るため、これらの頻度 frequency distribution を活用して調査した結果、

以下のような事柄が明らかになった。

(1)長いポーズ(--または---と表記される)は、聞き手の **attention** と密接な関係にあると考察される。すなわち身振り言語や社会的圧力などにより聞き手の **attention** が充分注がれている場合 (例 **sermons, university lectures**)には増加し、**speech** 以外に **attention** を引きつける手段を持たない場合 (例: **telephone conversations, radio discussions**)や、**attention** を引き付けることが要求される場合 (例: **political debates**)には減少する ($p<0.01$ 0.05)。

(2) **unit pause** (-)と **brief pause** (.)は、ポーズの合計のそれぞれ 31%と 57%を占め、長いポーズを補うように生起する。その結果、各セクション間のポーズ合計の頻度差は小さくなり (S.D.=0.87)、平均 6.6 語につき 1 ポーズという頻度が得られた。(3) **Non-verbal fillers** ([@], [@:], [@m], [@:m])の合計の頻度は、セクション間で大きく異なり、即興の会話の内、電話や親しくない間柄での会話でより頻繁に用いられる($p<0.01$)。

以上のデータから、口語の形式を問わず、平均約 6 語 (約 7.8 音節) に 1 度の割合で、ポーズか **non-verbal filler** またはその結合体が生起するという結果を得た(S.D.=0.77)。

(5)さらに 36 の **verbal fillers** の頻度を調べた結果、即興の親しい間柄での会話のほうが、親しくない者同志の会話より多くの **verbal fillers** を含むことが分かり、これら **lexical items** の **silence fillers** および **intimacy signals** としての機能が認められた。

(6)これらの 50 万語に及ぶ口語データより得られた結果は、言語心理学の理論と合致する。

◆シンポジウム《Helsinki Corpus をめぐって》

(司会 西村秀夫)

本シンポジウムは、平成7年度文部省科学研究費重点領域研究「人文科学とコンピュータ」の公募研究「コンピュータコーパスを利用した英語発達史研究」(研究代表者 渡辺秀樹)の研究成果の発表を目的として企画されたものである。この研究課題には、史的統語法研究・ Helsinki Corpus の有効性の検証・検索プログラムの開発という3つの主要なテーマがあり、齊藤俊雄会長による呼びかけ (NewsletterNo.5) に応じて集まった、英語史研究者10名で発足した HC 研究グループのメンバーが参加している。

今回のシンポジウムは、司会者による総論(Helsinki Corpus 編纂の経緯と今後の展望)

とペンシルバニア大学で公開されたタグ（文法標識）付き **Helsinki Corpus** の紹介（保坂氏）の後、上述の 3 つのテーマについて、検索ソフトウェア全般（塚本氏）、中英語の動名詞を例とした有効性の検証（伊藤氏）、検索およびその後のデータ処理まで含めた言語研究の 1 例（永尾氏）の順に進める予定である。史的英語コーパスを対象とするシンポジウムではあるが、現代英語のコーパスにも関連する点が多いと思われるので、現代英語の研究者にも積極的に議論に参加していただければ幸いである。

○保坂道雄「Penn-Helsinki Corpus について」

The Penn-Helsinki Parsed Corpus of Middle English が、1994 年 12 月にペンシルバニア大学により公開された。このコーパスは、**Helsinki Corpus** の **Middle English** のパートに文法標識を付加したものである。散文に限られているという点はあるが、史的言語資料に文法標識が付けられたのは画期的なことである。私のパートでは、この **Penn-Helsinki Corpus** の文法標識がどのように付与されているかを紹介し、その利便性と問題点を指摘していきたい。特に、生成文法に基づく文法標識もあるので、その点に関して注目したい。

○塚本 聡「コーパス検索ツール—いくつかの事例」

この発表は、コンピュータコーパス使用時に PC 上で利用可能なツールについて、どのようなものがあり、どのような長所・短所があるのかについて調査・検討するものである。また、**Helsinki Corpus** にはテキストに関する情報タグが付加されているが、既存のツールがこれに対してどの程度対応できるのか、あるいはそれらの検索時の簡便さについて考える。**awk**, **grep** 等の汎用ツール、また **Micro-OCP** にふれながら、できるだけ手軽に **Helsinki Corpus** を利用できる自作プログラムを紹介する。

○「有効性の検証—中英語期における動名詞の発達について」（伊藤礼子）

動名詞は名詞の派生形に由来しているが、動詞的性格を持つ動名詞は中英語期に発生し、近代英語期に発達・確立した。その過程や分析については、**Visser (1966)**, **Mustanoja (1960)**, **Tajima (1985)** 等が述べている。齊藤(1993)は初期近代英語期について **Helsinki Corpus** を検索し、その有効性を実証した。**Koma (1987)**, **Ito (1994)** は、限られたテキストではありながらも、動詞的動名詞の発生を前置詞との関連でとらえている。

本研究は、動名詞の発生について **Helsinki Corpus** の中英語期のテキストを検索し、その有効性を検証しようとするものである。

○永尾 智「検索とデータ処理の方法について」

この発表では、**Helsinki Corpus** を使った初期近代英語期の **will** と **shall** に関する調査を例にして、解析から類別に至る一見煩雑に見える作業過程が、市販ソフトウェアで軽易に進むことを具体的に示す。**Helsinki Corpus** を利用した試験的研究では、分析結果の統計資料によってコーパス自体の有用性は示されるものの、その作業過程は示されない。集計に手間取りそうな多岐にわたる参照部も、データベースのクロス集計で、大量データであっても即座に集計できる。本コーパスが手軽に利用できるものであることを示したい。