

英語コーパス学会 第 41 回大会資料

日時：2015 年 10 月 3 日（土）－ 4 日（日）
会場：愛知大学名古屋キャンパス
(<http://www.aichi-u.ac.jp/profile/campus-nagoya.html>)
〒453-8777 名古屋市中村区平池町 4-60-6

■ 第1日目

ワークショップ【Webアプリケーションを使ったコーパス研究のための統計的手法】

会 場：愛知大学名古屋キャンパス講義棟7階708教室

日 時：10月3日（土）10:00-12:00（9:30受付開始）

講 師：水本 篤（関西大学）

定 員：定員50名（先着順・要予約）

参加費：会員無料。非会員2,000円（当日会員としての大会参加費二日間共通）。

※予約申し込みは、<http://goo.gl/forms/9t0OBSvXz1> にアクセスして、参加申し込み用 Web フォームに必要な事項をご記入下さい。（申込締切日：9月30日）。

| | | | |
|---|-----------------------------|-----|--------------------|
| 日 時 | 2015年10月3日（土） | | |
| 受付開始 | 12:00（愛知大学名古屋キャンパス講義棟9階ロビー） | | |
| 開会式 | 13:00（同 講義棟9階903教室） | 司 会 | 田畑智司（大阪大学） |
| 1. 会長挨拶 | | | 堀 正広（熊本学園大学） |
| 2. 開催校挨拶 | | | 砂山幸雄（愛知大学副学長） |
| 3. 総会 | | | |
| 4. 学会賞審査報告 | | | 新井洋一（中央大学） |
| 5. 事務局からの連絡 | | | |
| 〈研究発表第1セッション（講義棟9階903教室）〉 | | 司 会 | 石川有香（名古屋工業大学） |
| 研究発表 1 | 14:00-14:30 | | |
| 東海大学発話・作文学習者コーパスの構築と | | | 長沼君主（東海大学） |
| CEFR 準拠到達目標及び指標改善の試み | | | 宇佐美裕子（東海大学） |
| 研究発表 2 | 14:35-15:05 | | |
| CEFR 基準特性としての関係代名詞： | | | |
| 学習者コーパス・教材コーパスによる分析 | | | 高橋有加（東京外国語大学大学院 S） |
| 〈研究発表第2セッション（講義棟9階904教室）〉 | | 司 会 | 水野和穂（広島修道大学） |
| 研究発表 1 | 14:00-14:30 | | |
| 近代英語コーパスの代表性の検討—3 コーパスの比較検討 | | | 塚本 聡（日本大学） |
| 研究発表 2 | 14:35-15:05 | | |
| 否定接頭辞 <i>un-</i> の外来語付加派生語の通時的考察 | | | |
| — <i>in-</i> 派生語との比較において— | | | 岡田 晃（小山工業高等専門学校） |
| 研究発表 3 | 15:10-15:40 | | |
| <i>Hold—oh, hold! Stay, my Lords! Ere ye commit A deed,</i> —近代英語における | | | |
| 「待て」を表す語彙の役割分担、競合と入れ替え | | | 山崎 聡（千葉商科大学） |
| 〈休 憩 15:40-16:00〉 | | | |
| シンポジウム | 16:00-18:00（講義棟9階903教室） | | |
| 《コーパスからの認知言語学へのアプローチ》 | | 司 会 | 長 加奈子（北九州市立大学） |
| コーパスを用いた構文の分析 | | | 大谷直輝（東京外国語大学） |
| <i>Having said that:</i> コーパスを利用した使用基盤による分析 | | | 大橋 浩（九州大学） |
| 事態把握の観点から見た二重目的語構文・ <i>to</i> 与格構文 | | | 川瀬義清（西南学院大学） |
| 構文に現れる事態把握の違い：コーパスに基づく実証 | | | 長 加奈子（北九州市立大学） |
| 《懇親会 時間：18:15-20:00 場所：キャンパスレストラン 会費：4,000円》 | | | |

■第2日目

日時 2015年10月4日(日)
受付開始 9:10(講義棟9階ロビー)

〈研究発表第3セッション(講義棟9階903教室)〉

司会 阪上辰也(広島大学)

研究発表1 9:30-10:00

テキスト整形における文字修飾の効用:

Tex-Edit Plus を中心としたエディター機能の比較

安間一雄(獨協大学)

研究発表2 10:05-10:35

コーパス分析ツールの選択肢の一つとしての CasualConc

今尾康裕(大阪大学)

研究発表3 10:40-11:10

「ダブルツールDDL」実践授業の試み:

WebParaNews と LWP for ParaNews を組み合わせて

中條清美(日本大学)

水本 篤(関西大学)

西垣知佳子(千葉大学)

〈研究発表第4セッション(講義棟9階904教室)〉

司会 佐野 洋(東京外国語大学)

研究発表1 9:30-10:00

Time-away 構文の意味的特徴について

西原俊明(長崎大学)

研究発表2 10:05-10:35

日英語パラレルコーパス中の対訳文間における

格助詞「が」を伴う名詞句の依存関係の構造的不一致

大矢政徳(目白大学)

研究発表3 10:40-11:10

日英語における動詞語彙の量的構造

園田勝英(北海道大学)

〈休憩 11:10-11:20〉

講演 11:20-12:30(講義棟9階903教室)

《コーパスからの情報抽出と抽出データの意味づけ: 具体的表現からの抽象化》

司会 中尾佳行(広島大学)

講師 滝沢直宏(立命館大学)

〈昼休憩 12:30-13:30〉

〈研究発表第5セッション(講義棟9階903教室)〉

司会 三浦愛香(東京農業大学)

研究発表1 13:30-14:00

日本人英語学習者ナラティブコーパスによる

「話し言葉」「書き言葉」の研究

山口有実子(東海大学)

宇佐美裕子(東海大学)

研究発表2 14:05-14:35

日本人大学生の書く英語論証文の論理構造を探る

—4つの分析的枠組みを用いて

山下美朋(関西大学大学院S)

研究発表3 14:40-15:10

中学英語教師向け英日発話パラレルコーパス構築と

教師教育応用への展望

片桐徳昭(北海道教育大学 旭川校)

〈研究発表第6セッション（講義棟9階904教室）〉

司 会

高見敏子（北海道大学）

研究発表1 13:30-14:00

話し言葉コーパスを用いた談話辞 *absolutely* と
その類義表現の分析

山本五郎（広島大学）

研究発表2 14:05-14:35

男女両名義を使用する作家の作品判別
—Rowling と Sheldon

木村美紀（明治大学大学院 S）

久保田俊彦（明治大学）

研究発表3 14:40-15:10

マルチモーダルコーパスを用いた英語母語話者による
パブリックスピーキング・パフォーマンスの分析：
発話速度，ジェスチャー，語彙特性の観点から

冬野美晴（九州大学）

山下友子（九州大学）

閉会式

15:20（講義棟9階903教室）

閉会の辞

塚本倫久（愛知大学）

■10月3日(土)

【ワークショップ】

Web アプリケーションを使ったコーパス研究のための統計的手法

水本 篤(関西大学)

これまで英語コーパス学会でもワークショップが開催されるなど、近年、データ解析環境 R への関心が高まっている。R は(高価な)商用統計分析ソフトではブラックボックスになりがちな分析過程の理解を深めるために有用であり、また、コードが残るため結果の再現性という点からも優れている。しかし、R は GUI(グラフィカル・ユーザー・インターフェース)ではなく、自分でコードを書いて動かす必要があるため多くの人が敷居が高いと感じるのも事実である。そのため、初学者でも R を使ってどのようなことができるのかを体験できるように、また、R 利用への橋渡しとして、Web アプリケーション“langtest.jp”を作成した(<http://langtest.jp/>)。この Web アプリケーションでは、Excel のようなスプレッドシートのデータをサイト上でコピー&ペーストするだけで、基礎的な統計的分析を実行することが可能である。

本ワークショップでは、Web アプリケーション“langtest.jp”を用いて、コーパス研究でよく使われる統計的検定や、多変量解析、データ視覚化などの紹介を行う。また、R で langtest.jp と同じ分析をローカルな環境で行う方法についても触れる予定である。なお、本ワークショップの参加者は初学者を想定しており、参加にあたって統計的手法に関する事前知識は必要としない。

■10月3日(土)

【研究発表第1セッション】

【研究発表1】

東海大学発話・作文学習者コーパスの構築と CEFR 準拠到達目標及び指標改善の試み

長沼君主(東海大学)・宇佐美裕子(東海大学)

東海大学では2010年度より、CEFR を参照した到達目標レベル設定に基づいた統一カリキュラムによる英語教育を進めており、Can-Do リストの形式での到達目標記述文を作成し、初回、中間、最終と自己評価チェックリストを実施している。また、中間及び期末評価においても、統一ルーブリックによるライティング及びスピーキング評価を行っている。しかしながら、統一ルーブリック自体は CEFR を参照して作られたものではあるが、独自に観点や基準を工夫し、スケールもレベル表記ではなく、50点満点の点数化をしており、目標が到達されたかは必ずしも明らかではない。

そこで CEFR に準拠したより妥当性の高い到達目標及び基準作成のため、学内の学部等研究教育補助金を受けて、2013年度にはライティング学習者コーパス、2014年度にはスピーキング(インタビュー)学習者コーパスのパイロット的な構築を試みた。学習者コーパスの構築にあたっては、中間・期末評価における学習者の作文・発話データを収集した上で、授業担当者による統一ルーブリックを用いた評価スコア(全体及び観点別スコア)の情報を付加した。また、外部評価者2名による CEFR の評価観点(assessment criteria grid)に基づくレベル評価も行い、レベル情報を付与した。

本研究では、統一カリキュラムの到達目標のレベル設定が妥当なものであるのか、また、評価ルーブリックに基づくクラス内の観点別評価が実際にはどの程度 CEFR のレベルと対応しているのかを検証するため、CEFR レベルと内部評価スコアとの対応関係を分析した。さらには、学習者コーパスにおける言語データと評価データを結びつけることで、CEFR レベルごとにどのような基準特性が見られるか分析した。分析の結果、CEFR レベルと統一ルーブリックスコアは関連を示し、それぞれのレベルを弁別するスコア間に直線的関係が見られた。また、平均語数や文長から目標となるテキスト指標も得られた。平均文長は English Profile で示唆されている基準との一致が概ね見られた。CEFR-J の語彙リストとの関連においても、レベルが上がるほど、延べ語数、異なり語数の両面で産出語彙のレベルもあがっていくなどの関係が見られた。

今年度より科学研究費補助金を受けて、ライティング及びスピーキング学習者コーパス及び評価間の関連性の検討や学年間での到達目標の検証を試みている。今後、学習者コーパスと評価データを結びつけることで、Evidence Centered Design に基づいた、実際のパフォーマンスの証拠(evidence)を反映した到達目標及び指標の改善を進めて行きたい。また、言語ポートフォリオの構築による実際に参照可能な証拠に基づいた Can-Do リストの開発にもつながって行きたい。

【研究発表 2】

CEFR 基準特性としての関係代名詞：学習者コーパス・教材コーパスによる分析

高橋有加（東京外国語大学大学院 S）

本研究は、日本でも影響力を拡大しつつあるヨーロッパ言語共通参照枠（CEFR）の英語に関する参照レベル記述（reference level description）の一環として、A1～B2 レベルを同定する際の基準となる「基準特性（criterial feature）」として関係代名詞をとりあげ、日本人英語学習者の書き言葉・話し言葉における関係詞のタイプごとの使用上の特徴（正用法・誤用法のタイプや分布）を明らかにするとともに、CEFR 準拠のコースブックでの分布の特徴と比較することを目的とする。関係詞を選んだのは、Hawkins (2009) で基準特性として有効と考えられる言語項目に関する「20 の仮説」の中で関係詞の使用が挙げられており、日本人英語学習者にとっても困難な項目であることで基準特性としてさまざまな学習困難点や習得の側面を観察できると考えたからである。具体的な研究方法としては以下の手順を踏んだ：

・対象とする関係詞：

対象とする関係詞は *who*, *which*, *that*, *whose*, *whom* の 5 つとした。今回は省略の関係節は対象外とした。関係詞の分類に関しては既存の言語類型論に基づいた Keenan & Comrie (1977) の Noun Phrase Accessibility Hierarchy (NPAH) および関係節の種類と埋め込み位置をもとに難易度階層を提案した Hamilton (1994) の SO Hierarchy Hypothesis (SOHH) を用いた。

・コーパス：

使用したコーパスは、日本人中高生の英作文コーパスである「JEFLL コーパス」(約 70 万語, 約 10000 人) (投野, 2007), 日本人英語学習者の話し言葉コーパスである「NICT JLE コーパス」(和泉他, 2004), そして科学研究費基盤研究 (A) (2012-15 年度, 代表: 投野由紀夫) で整備された「ELT Coursebook コーパス」の 3 つを用いた。

・分析方法：

- Sketch Engine にコーパス・データをアップロードし、各種形態素解析を行い、各コーパスから関係詞を含んだ文を Sketch Engine の CQL をもとに抽出した。
- 結果をエクセルにエクスポートし、NPAH, SOHH の分類およびエラーとそのタイプを各例文に人手でタグ付与を行った。
- CEFR レベルごとの NPAH, SOHH の頻度を集計し、CEFR レベルとの独立性の検定を行った。またエラータイプのレベル別分布状況も観察した。

・主要な結果は以下ようになった：

- 学習者コーパスにおける関係代名詞の使用頻度は CEFR レベルが上がるにつれ多くなっており、正用法か否かに関わらず関係詞の使用自体が基準特性として有効である可能性が示唆された。これは書き言葉・話し言葉両方に関して当てはまった。
- 学習者コーパスにおける関係代名詞のタイプ別出現率は NPAH にほぼ従っていたが、間接目的語 (IO) など若干順序が逆転するものが見られた。
- 学習者コーパスにおけるエラータイプに関しては (a) 関係詞のタイプの誤り, (b) 関係節内の代名詞残留, (c) 前置詞の脱落, (d) 日本語構造に起因すると思われる構造全体の誤り, などが見られた。ただし、エラー頻度が基準特性として有効かどうかは統計的には検証できなかった。
- 教材コーパスにおける分析は日本の教科書に比べると海外の CEFR 準拠のテキストは主格の *who* の用法が有意に頻度が高く、他の用法よりも重点的に扱われていることがわかった。

■ 10月3日 (土)

【研究発表第 2 セッション】

【研究発表 1】

近代英語コーパスの代表性の検討—3 コーパスの比較検討

塚本 聡 (日本大学)

A Representative Corpus of Historical English Registers (ARCHER) コーパスは、Helsinki Corpus に次いで早い時期から、初期近代英語をカバーするコーパスとして知られてきた (齊藤他 1998)。長らく公開がされていなかったが、現在では Lancaster 大学ホームページで ARCHER 3.2 が公開されている。その後、史的コーパスとして *Penn Corpora of Historical English* の一連のコーパス (PPCME2, PPCEME, PPCMBE), *The Corpus of Late Modern English Texts* (CLMET3.0), *The Corpus of Historical American English* (COHA) などが編纂され、公開されてい

る。これらのコーパスに共通する特徴は、品詞標識が付されていること、また ARCHER および COHA ではレンマが付され、これらの情報を利用することにより、広範な調査が可能となる。

早くから作成された ARCHER コーパスを利用した研究事例は多数存在する。Biber, Conrad and Reppen (1998, 2009), Denison (1998), Rissanen (2004), Akimoto (2014) のように現在でも広く研究に利用されている。

しかし、ARCHER コーパスは、当初170万語規模で計画され (Biber et al. 1994)、現在では約330万語まで拡大されて収録されているが、1600年から1999年までをカバーし、12のジャンルに分けられ、かつ英・米の地域にまたがる地域変種を扱うなど、後発の PPCMBE (170万語)、CLMET (340万語)、COHA (4億語) と比較して収録語数やジャンル構成が異なる。収録テキスト数は1710と多数であるが、収録語数が100語程度のテキストも多数存在する。

この3つのコーパスでは、言語項目の生起数が異なる。例えば、もっとも頻度の高い語である *the* の3コーパスにおける1800年ころの1000語あたりの生起数は、68.6, 65.8, 59.5 (それぞれ ARCHER, PPCMBE, CLMET の順。以下同)、Modal Verb と品詞分類された語の生起数は、15.0, 16.1, 15.8と異なっている。現在分詞は、ARCHER では1800年から1900年に向けて、線型的に使用頻度が高まるが (1000語あたり12.2から14.7)、他のコーパスでは同種の変化は見られない (PPCMBE: 15.1から15.6, CLMET: 12.1から14.6。ただし1850年にはすでに14.5)。このように、同時期のコーパスでありながら、コーパス間で語の生起数には相違がみられる。

本発表では、ARCHER, PPCMBE, CLMET の3コーパスが重なり合う1700年から1900年の約200年間に着目し、これらのコーパスの語彙、品詞、レンマの生起数を相互比較することにより差異が存在することを確認する。さらに、ジャンル構成や収録語数の違いなどから差異となる要因を推定し、それぞれのコーパスが有する特徴を明らかにすることを目的とする。

【研究発表 2】

否定接頭辞 *un-* の外来語付加派生語の通時的考察—*in-* 派生語との比較において—

岡田 晃 (小山工業高等専門学校)

本発表は、自身の否定接頭辞付加派生語の通時的研究の一環としていくつかの事例を取り上げ、今後の更なる研究に貢献していくものである。

現代英語には多数の否定接頭辞が存在している。中でも *un-* は古英語期から英語母国話者によって使用されてきた最も生産性の高い接頭辞の1つである。中英語期以降の外来要素流入にも関わらず、廃語となることなく使用され続け、今日においても新しい造語を作り出すその生産性の高さは否定できない。*Un-* の生産性の高さが窺える1つの特徴として *uncommunicative* や *unadvisable* などのように外来由来の語、いわゆる「借用語」にも付加が可能であるということが挙げられる。一方、*un-* と言語学的特徴が似ていると言われているラテン語系接頭辞の *in-* は、通例、*incommunicative* や *inadvisable* などのように同じラテン語系の単語に付加され、新たな語を作り出してきた。これら両接辞は形容詞を主な付加品詞としているため、外来由来の形容詞に対しては *in-* と *un-* の両方の付加が理論上可能であるということになる。その結果として *in-/un-communicative*, *in-/un-advisable* などのような「二重語」が派生されるのは珍しい現象ではない。これまでもこのような二重語の存在は指摘されてきたが (Jespersen, 1917; Zimmer, 1964; Funk, 1971; Horn, 2001)、接尾辞との関連性からの考察を通時的に行っているものは少ない。また、接頭辞の生産性について言及している Katamba (2004) では、形容詞派生の代表的な外来由来接尾辞 *-able/-ible/-uble*、及び *-ive* と *un-* の関連性が指摘されているが、他の外来由来形容詞派生接尾辞との関連性については言及されていない。

本発表では外来由来の形容詞に古英語由来の *un-* が付加された派生語の現代英語における使用状況を British National Corpus (BNC) で頻度検索を行い、さらに中英語期以降の使用例を *The Oxford English Dictionary 2nd edition (OED)* を使用して通時的に調査を行う。調査方法は以下の通りである。

- ① *un-* [[+latinate]-ic] を BNC で使用頻度の高い 10 語を検索
- ② ① の 10 語の *un-* を取り、*in-* を付加した派生語の使用頻度を検索
- ③ ② の頻度が確認されたらそれらの使用域 (テキストタイプ、年代、性別、地域など) を調査
- ④ *OED* を使用して、①及び②の語の過去における使用例を調査

以上のような考察を行うことにより、否定接頭辞と接尾辞との関連性、本来語と外来語との融合的語形成過程など、今後の更なる研究発展の可能性を示唆できると思われる。

なお、本発表における「二重語」という用語は、形態論的に「異なる2つの接頭辞が共通する基体に付加されて生産された派生語」を意味している。また、便宜上、*in-(il-, im-, ir-)* だけでなく他の否定接頭辞 (*non-*, *dis-* など) も考慮に入れる。

【研究発表 3】

Hold—oh, hold! Stay, my Lords! Ere ye commit A deed,—近代英語における
「待て」を表す語彙の役割分担, 競合と入れ替え

山崎 聡 (千葉商科大学)

表題の *Hold* と *Stay* は共に「待て」という、命令文で相手の行為を阻止しようとする効果をもつが、両者にはどのような違いがあるのだろうか。また、*hold* と *stay* はその後、いつ頃、どのようにして *stop* や *wait* に取って代わられていったのだろうか。本発表では、命令文で主に指示的な発話行為を表す、日本語の「待て」に相当する近代英語期の語彙・表現の役割分担と通時的な語彙の入れ替えの実態を明らかにすることを目的とする。日本語の「(ちょっと) 待て」がそうであるように、近代英語期に(比較的)よくみられる *stay*, *hold*, *soft*, *stop*, *wait* には、いずれも次の用法がみられる。

- [1] 去って行く人を引き留めて (*OED Online*, s.v. *stop* 34 a 参照)。
- [2] 相手の行為を阻止したり、発言を遮ろうとして (*OED Online*, s.v. *stay* 2c; *hold* 27; *stop* 37d 参照)。
- [3] 相手に異議を唱えたり、抗議して ([2] (以下、ブラケットに入れた数字は特定の用法を表す) の場合も異議を唱える意が含まれることが多いだろうが、[3]は[2]の効果に直接つながらないもの)。
- [4] 新たな状況に気づいたり、ある考えを思いついたり、考え直す時に用いて (*LDOCE Online*, s.v. *wait* 3c, *hang on* b; *OALD Online*, s.v. *wait* Idioms ; 日本語の「待てよ…」参照)。
- [5] *Let's see...* のように、次の言葉を探したり、考え中に挿入されて (*OED Online*, s.v. *stay* 2c 参照)
- [6] 以上の[1]~[5]の効果をもたず、単に「そのまま少し待って、時間を下さい」といった意味を表して

発表では、The Penn-Helsinki Parsed Corpus of Early Modern English, A Corpus of English Dialogues 1560–1760, The Corpus of Late Modern English Texts, Version 3.0 及び初期近代英語期の 20 数編の戯曲作品をデータに、1570 年–1899 年を、70 年ごとに輪切りにして (最後の期間は 50 年)、まず、*stay*, *hold*, *soft*, *stop*, *wait* 等の用例は上記[1]から[6]のいずれの用法で用いられたかを調査する (form-to-function mapping, Jacobs and Jucker 1995 参照)。その結果、1570–1639 年では *stay* では[1], *hold* では[2], *soft* では[4]がそれぞれの用例の約半数を占め、それぞれ好みがみられた。*Soft* はこの後データから姿を消すが、18 世紀後半頃まで、関連する *tarry*, *stand*, *cease* 等の(sense) synonyms (Murphy 2010) を含めて、[1]は *stay* が、[2]は *hold* がそれぞれ主に担うという役割分担がみられた (function-to-form mapping)。

一方では 18 世紀に入ると *stop* がデータに現れ、おそらく、*hold* の[1]から[6]の用法の分布が (*stay* に比べて) *stop* のそれと似ていることから、*hold* は *stop* との競合により直接的にさらされ急速に頻度を落とし、19 世紀の終わりには[2]でもほぼ *stop* に取って代わられる。[1]は主に *stay* が担っていたが、[1]も、[2]を主に担った *hold* の場合よりもやや遅いペースで *stop* との入れ替えが進んでいく。現代英語 (BNCweb の fiction) では最も頻繁な *wait* (*a minute* 他) は *stop* より 70 年ほど遅れて出現するが、当初は[6]の用法が多く、*hold* と *stay* の主要な活躍の場を奪ったのはまずは *stop* であることが分かる。

■10月3日(土)

【シンポジウム】

コーパスからの認知言語学へのアプローチ

司会 長 加奈子 (北九州市立大学)

1980年代以降、新しいパラダイムとして発展してきた認知言語学の言語観を支えるものの一つとして「動的用法基盤モデル」というものがある (Langacker, 2000)。それまでの言語理論においては、文法は言語使用から切り離されたルールの集合体であるとされてきたが、動的用法基盤モデルでは、文法は実際の言語使用の場面からボトムアップ的に習得される抽象化された概念構造 (スキーマ) のネットワークであると考えられている。その為、実際の言語使用の場面における具体的な発話の集積であるコーパスと認知言語学の理論との間に高い親和性があることは容易に想像できる。認知言語学の分野においても、コーパスによる言語理論の実証の重要性が認識されてはいるものの、まだ十分な研究がなされているとは言い難い。

そこで本シンポジウムでは、コーパス言語学が認知言語学に対してどのように貢献できるかというテーマのもとで、個別の言語現象に対する理論的分析とそれを裏付けるコーパス分析を行う。

コーパスを用いた構文の分析

講師 大谷直輝 (東京外国語大学)

認知言語学では、語・句・節などの言語ユニットは、形式と意味の対からなり、部分からは予測されない全体の意味を持つ構文であると考えられる。また、構文は言語使用の中から創発し、ゆらぎ定着するという用法基盤モデルの考え方を採用する。本発表では、構文文法と用法基盤モデルの考え方に基づいた言語研究の実証的な方法論を紹介する。

コーパスには大量の言語データが収集されているという特徴と、収取されている文が単文の集合ではなく、前後の文脈と関連しているという特徴があるため、実際の使用の場で用いられる構文の特徴を定量的に捉えるうえで適している。コーパスを用いた構文研究では、構文は形式と意味の対であるため、語句の表層分布や頻度等の形式的な情報だけでなく、構文の意味的・談話的な特徴を捉える必要がある。本研究では、構文の意味的・談話的な要素をコーディングする方法を紹介しながら、形式と意味の両面に注目した実証的な構文分析を行う。

Having said that: コーパスを利用した使用基盤による分析

講師 大橋 浩 (九州大学)

言語への認知的アプローチは、言語体系は言語使用者による個々の言語事象に基盤を持ち、ボトムアップ的に創発するという、使用基盤 (usage-based) の立場をとる (Kemmer and Barlow (2000), Langacker (2000), Bybee (2010)など)。言語要素がユニットとして構造化されるためには、同じパターンが繰り返し使用されなければならないが、また、パターンとして定着するためには高い頻度で繰り返される必要がある ((Kemmer and Barlow (2000), Bybee and Hopper (2001), Bybee (2007)など)。このように、頻度を重視する使用基盤による言語分析には大規模コーパスからのデータが極めて重要な意味を持つ。本発表では大橋 (2013) などに基づき、分詞構文由来の *having said that* 節をとりあげ、COCA や COHA からの例を詳細に分析し、その統語的、意味的、談話的特徴を明らかにすることにより、コーパスを利用した使用基盤的分析の有効性を示したい。具体的には *having said that* について以下の点を中心に論じる。i) 構文的特徴、ii) 文中の位置、iii) 共起する接続詞、iv) *that* 使用の認知的動機、v) 主節との関係、vi) *that being said*, *that said*, *having said this*, *this being said*, *this said* などの関連構文との分布、使用域、通時的変化に関する比較、v) 談話機能と新たな用法。

事態把握の観点から見た二重目的語構文・to 与格構文

講師 川瀬義清（西南学院大学）

認知言語学では、構文の違いは概念化者の事態把握の違いを反映していると考えられる。このような見方に立つと、客観的には同じ出来事を表していると考えられる二重目的語構文と to 与格構文の違いは、事態把握の違いを反映したものであるととらえることができる。二重目的語構文も to 与格構文もいわゆる直接目的語と間接目的語の二つの名詞句が現れるが、二重目的語構文では動詞の直後に間接目的語が、to 与格構文では動詞の直後に直接目的語が現れる。しかしながら、この語順の違いがどこから生じているのかについてはこれまでほとんど言及されることはなかった。二重目的語構文も to 与格構文も、二者関係から三者関係への拡張であるが、その拡張の方向、すなわち事態把握の違いが、この語順の違いとなって現れていることを、様々な言語データに基づき明らかにする。ここでの分析で得られた事態把握の違いが実際の言語使用例にどのように現れているかについては、次の長講師の発表でコーパスデータに基づき見ていく。

構文に現れる事態把握の違い：コーパスに基づく実証

講師 長 加奈子（北九州市立大学）

認知言語学は言語使用者が実際の言語使用の実例を基盤に言語を習得するという動的用法基盤モデルを取っている。このことは、認知言語学による言語現象に対する理論的説明は、言語使用の実例を集積した大規模コーパスによって、その説明に対する裏付けを示すことが可能であることを意味する。すなわち、大規模コーパスを用いて言語現象に関する理論的説明を実証することで、認知言語学の知見をより確かなものにするのが可能となる。そこで本発表では、言語現象に対する理論的説明を実証するためのツールとしてのコーパスの活用という観点から、コーパスデータを定量的に分析することにより理論的説明の検証を行う。

本発表では、特に、認知言語学の理論的な枠組みから二重目的語構文と与格構文の記述を行う川瀬講師の発表を受け、目的語の定性という観点からコーパスを用いて川瀬講師の理論的説明を実証する。

■10月4日(日)

【研究発表第3セッション】

【研究発表1】

テキスト整形における文字修飾の効用：Tex-Edit Plus を中心にしたエディター機能の比較

安間一雄（獨協大学）

本報告は、文字修飾情報を識別要素として利用することで、テキストエディターを用いたコーパス処理が柔軟性を上げ複雑な手順を簡略化できることを示すものである。言語コーパスの整形にテキストエディターは不可欠のツールである。ウィンドウズアプリの「秀丸」や pPad、マッキントッシュアプリの mi や JeditX などのよく使われるツールは、正規表現はサポートしているのでパターン化された文字列の編集や単語末尾の認識といった初歩的操作は容易であるが、少し複雑な操作（例えば特定の語を含む論理行のみを抽出する）になると手の込んだ手続きを取るか、あるいは諦めて awk などのプログラミング言語に頼ることになる。しかしながら、検索および置換に文字修飾情報（特に色）が利用できれば、エディターでの操作だけで高度な処理が可能になり、コーパス研究の初心者にもツール使用の敷居を低くすることに繋がる。また、文字修飾情報はテキスト本体を書き換えないので安全であり、結果を複数のアプリで同一状態として共有することが可能である。さらに、文字修飾情報を利用する／しないを切り替えることができれば、特定の色の文字列のみを取り出したり、異なる色に跨る文字列を色に依存せずに抽出することも可能である。現在市場に出回っているフリーウェアおよびシェアウェアでこれが可能なのは、唯一マッキントッシュ用のエディター Tex-Edit Plus である。このアプリはワープロソフトであるので WYSIWYG による印刷用編集のみならず、フォント種やイタリックなどの文字修飾ができるのが当然であるが、POSIX 文字クラスを含む正規表現もサポートされているので、テキスト中の色を付けたタグ部分とそれ以外の部分とで異なる処理（例えば地の文に含まれるカンマやピリオドなどパンクチュエーション文字のみに改行コードを付ける）を行うことも可能である。本報告では、文字列マーカーの付加／削除のみを用いる他のアプリでどこまでの操作が可能になるのかを並行して示す予定である。現行のテキストエディターは殆どがプログラミング言語もしくは HTML のソースコード作成用であるため、コードの枠構造を示す目的以外には色に関心を持たない。また、ワープロソフトはユーザーが文書を作成するためのものなので、文字列のパターンの「発見」は意図されていない。自然言語テキストの編集処理にこそ新たな次元の識別要素の活用が高い効果を発揮する。

【研究発表2】

コーパス分析ツールの選択肢の一つとしての CasualConc

今尾康裕（大阪大学）

コンピュータの処理速度が上がり、記憶装置やメモリが大容量になるにつれ、個人所有のコンピュータでテキストコーパスを検索できるようになってから久しい。1990年代には、MicroConcord や Conc, MonoConc などが公開され、MS-DOS や Mac OS などで、容易にコーパス分析ができるようになった。その後、Windows 向けには様々な有料の高機能コンコーダンサーが開発されたが、2000年代半ばを過ぎに Anthony 氏による AntConc が Mac に移植されるまで、Mac で動くまともなコンコーダンサーは存在しなかった。ただ、AntConc も、当初は X11 に依存していたため様々な制限があり、使いやすいアプリケーションとは言えなかった。

そのような状況の中、2006年後半に Mac でネイティブに動く CasualConc の開発を始めて無料配布し、Mac が再び普及し始めた海外で使用されるようになってきた。現在まで、機能追加などのアップデートを繰り返してきたが、開発に利用しているスクリプト言語も古くなり、新しい Mac OS への対応も難しくなってきたため、別の言語でほぼ一から書き直す次期バージョンの開発を開始した。

開発中の新バージョンは、現時点では、搭載予定のすべての機能が備わっているわけではないが、これまでのファイル及びコーパス・データベースを作成した上での検索に加えて、Web 上で見つけたテキストなどをコピー&ペーストして簡単に処理を実行するモードを備えている。また、マルチスレッド処理や結果テーブル上での絞り込みなどの OS X 標準の機能への対応も進めている。

本発表では、KWIC 検索、単語・ n -gram リストやコロケーションリスト作成などの基本的な機能を紹介しつつ、ファイルごとの単語リスト作成や TF-IDF の計算、指定文字列の頻度集計やコロケーション集計など、CasualConc ならではの機能を使ったコーパス研究での利用方法の提示を行うとともに、新たに追加した、頻度集計されたデータを R に送ってグラフの描画を行い、その結果を表示させる機能を紹介する。

今後は、マニュアルの整備とともに、文字列抽出の機能を充実させ、さらに高度な分析を行うための前処理を行えるようなツールへと進化させたいと考えており、そのための議論を行いたい。

【研究発表 3】

「ダブルツール DDL」実践授業の試み： WebParaNews と LWP for ParaNews を組み合わせて

中條清美（日本大学）・水本 篤（関西大学）・西垣知佳子（千葉大学）

本発表では、大学の一般英語クラスにおいて、2 種類のタイプの検索ツールを組み合わせて実施した Data-driven Learning (DDL) 指導実践の背景、手順、教育効果、学習者の反応について報告する。

発表者らは 2004 年より毎年改訂を加えながら DDL 授業を実践しているが、近年検索ツールの実用性が飛躍的に高まっている。具体的には、無料で使えるユーザー・フレンドリーな Web 検索ツール WebParaNews (以下 WPN) が 2012 年に、LWP for ParaNews (以下 LWP) が 2013 年に公開された。本発表では、効果的に DDL を授業に取り入れるひとつの方策として、異なる検索機能を持つ WPN と LWP という 2 種類の検索ツールを組み合わせて授業に取り入れ、1 種類の検索ツールを使った DDL 授業の効果と比較した。

1 つ目の検索ツールである WPN は、AntConc などのように、検索語を画面中央に据えた Key Word In Context (KWIC) を用いて、文の一部に色を付け、注目すべき箇所を目立たせて検索語の用法を理解しやすく示す。KWIC は習熟度の低い学習者の学習負担を軽減すると言われるが、検索される共起語やパターン数が多くなると、検索語の文法的振る舞いの全貌を把握することは困難となる。この欠点を補う 2 つ目の検索ツールが LWP である。LWP は Sketch Engine のようなレキシカル・プロファイリング型検索システムで、検索語の文法的振る舞いを分析し、コロケーション/コリグレーション情報を単語ごとに提供し、語彙・文法の規則を集約して、わかりやすく結果を表示する。

本実践では、最初に WPN の KWIC 表示から得られる情報を利用して学習者の仮説形成を促した。次に LWP によるプロファイリング・サマリーを使って学習者が自発的に形成した仮説の検証をサポートした。本指導実践ではこのアプローチを「ダブルツール DDL」と呼ぶ。この DDL アプローチは、1 種類の検索ツールを使って行う「シングルツール DDL」よりも深い情報処理を伴うことから、学習項目の理解度をより高められるであろうと考えた。

実践は、2013 年から 2015 年まで行われ、参加者は TOEIC 360-395 点レベルの大学生 129 名であった。処置群では半期 10 回、名詞句構造の習得を目指してダブルツール DDL を行い、対照群ではシングルツール DDL を行った。指導効果の検証は事前・事後テストとして 3 種類の名詞句テストを実施し、さらに学習者からフィードバックを収集した。評価テストおよび学習者からのフィードバックからダブルツール DDL の有効性が明らかになった。

■ 10 月 4 日 (日)

【研究発表第 4 セッション】

【研究発表 1】

Time-away 構文の意味的特徴について

西原俊明（長崎大学）

英語には、(1)に示す V[Time NP][away] の連鎖をもつ文が可能であり、時間表現と away の語順を入れ替えることが可能である。

(1) a. We danced the night away. b. We danced away the night.

Jackendoff (1997)(2002)によれば、Time-away 構文は、(2)の統語構造と意味構造をもち、「～して無頓着に無駄に時間を過ごす」ということを表す構文であると分析されている。また、Jackendoff (2007)では、Time NP と away がイデイオムをなしているとみなされている。

(2) a. [VP V NP [Prt away]] b. 'waste [Time NP] heedlessly V-ing'

この発表では、COCA などのコーパスを用いて、それぞれ V [Time NP] [away] と V [away] [Time NP] の連鎖に生起する動詞の種類を明らかにし、そこに存在する意味的特徴を明らかにする。具体的には、1)V

[Time NP] [away]の連鎖の場合、party, waltz, sing, danceなど、ある行為を楽しんだ結果、時間が経っていたことを表す動詞が多く生起すること、また、文脈からもそのように解釈できること、2)V [Time NP] [away]の連鎖の場合、無駄な時間を過ごしてしまったという解釈は、構文そのもの意味によるものでなく、動詞が本来もつ意味によるものであること(idle, gamble, snooze, sleepなど)、3)V away [Time NP]の連鎖の場合に「無駄な時間を過ごした」という意味が多くみられること(flap, flirt, snore, sipなど)、また、この種の動詞の生起からの類推により、V [away][Time NP]の連鎖の場合に「無駄な時間を過ごした」という意味がでやすいこと、また、自発的な意図をもって行う行為の場合、この語順がとれないこと、4) time NP以外のNP、距離を表すNPやmind, frustrationなどのNPが同じ形式に生起することから、結果構文との関係について中村(2003)によるX' 意味論における意味合成を用いた分析を試みたい。

【研究発表 2】

日英語パラレルコーパス中の対訳文間における 格助詞「が」を伴う名詞句の依存関係の構造的不一致

大矢政徳（目白大学）

本研究の目的は、日英語パラレルコーパス中の日本語文と英語対訳文との依存文法木の構造的不一致の実態を明示的に理解するために、日本語の格助詞「が」を伴う名詞句と動詞との依存関係に着目し、それが英語対訳文ではどのような依存関係として表現されるかをコーパスデータに基づいて提示することである。Oya (2015)では、Wikipedia 日英京都関連文書対訳コーパス(2011)中で「～となっている」という表現を含む日本語文とその英語対訳文のごく一部を抽出し、それらを構文解析器にかけて得られた依存文法木の不一致を、Mel'čuk and Wanner (2006)の Syntactic Mismatches の概念を援用して分類した。この概念は、翻訳元の言語の単文中の単語の依存関係と、これに対応する翻訳先の言語の単文中の単語の依存関係のずれを6種類に分類したものである。この6種類は以下の通りである。(1) Head switching: 翻訳元の単文では単語 A は単語 B に依存していたが、これに対応する翻訳先の単文では単語 B に対応する単語 B' が単語 A に対応する単語 A' に依存している場合、(2) Conversion: 翻訳元の単文では単語 A は単語 B に依存関係タイプ C として依存しているが、これに対応する翻訳先の単文では依存関係タイプ C' で依存している場合、(3) Part-of-speech change: 依存元あるいは依存先の単語の品詞が交替している場合、(4) Lexical fission/fusion: 翻訳元の単文では一単語だったものが翻訳先の単文では複数単語である場合、またはその逆の場合、(5) Function-word introduction: 翻訳元では不在だった機能語が挿入される場合、(6) Branch transposition: 翻訳元の単文では単語 A は単語 B に依存していたが、翻訳先の単文では単語 A に対応する単語 A' が単語 B に対応する単語 B' 以外の単語に依存している場合)。Oya (2015)では、「～となっている」表現を含む文の中から代表的なもののみを取り上げ、それらの日英対訳文中にある Syntactic mismatches の数を提示するにとどまっていた。本研究では、Oya (2015)で利用されたものと同一の日英語対訳コーパス中のより多くの翻訳ペアの中で、一般的に日本語の主格を担うとされる日本語の格助詞「が」を伴う名詞句を含む日本語文とこれに対応する英語文の中からランダムに抽出した100文を対象を絞り、日本語文中の当該名詞句と動詞との依存関係が英語文の依存木ではどのように表現され、各 Syntactic Mismatches がどの程度出現するかを数値的に把握することを試みた。その結果得られた日本語の格助詞「が」を伴う名詞句122個の中で、対応する英語文で上述の Syntactic mismatches が見られないものが66個、いずれかの Syntactic mismatches を見せるものが56個あり、この中で最も頻繁に観察されたのは(2) Conversion の32個であった。この結果は、格助詞「が」を伴う名詞句はどのように英語に訳されることが多いか、という点に関する数値的・客観的な知見を与えるものであり、英語教育や日英機械翻訳などへの応用が期待される。

【研究発表 3】

日英語における動詞語彙の量的構造

園田勝英（北海道大学）

現代英語の動詞についての調査研究は、過去半世紀以上にわたり活発に行われ、統語的、形態的、意味的な特性の解明が大きく進んでいる。その中でコーパスが大きな役割を果たしていることはよく知られていることである。しかしその一方で、現代英語の動詞全体を対象として、コーパス等を用いて計量的に観察調査することは、今のところ単純な頻度表作成などに限られていて、余り進んでいるとは言えない。(表題の「動詞語彙」は動詞全体を意味する術語として用いている。)

BNC-XML の Written part に出現する一般動詞の数(基本形に直した異語数、以下同様)は約 35,000 であるのに対し、LDOCE 等の上級用英語学習辞書に収録されている一般動詞の数は 6,000 弱である。

この差は、語彙の多様性や流動性を反映するものと考えられるが、本発表は語彙の多様性・流動性をテーマとしていないので、辞書に収録されている約 6,000 の一般動詞を一応「現代英語の動詞全体」とする。これらのうち、約 1,000 語の動詞は *en-, be, de-, ...; -ify, -ize, -en* などの接辞を付加することによって生み出されている。さらに、約 1,000 の動詞が名詞からの転換によって作られる名詞転換動詞(denominal verbs)である。残りの約 4,000 の動詞が「純粋な」動詞ということになる。また、これらとは別にイディオム化した句動詞が約 2,500 ある。

現代日本語についても、『現代日本語書き言葉均衡コーパス』(国立国語研究所)や各種の電子化辞書を用いて同様の観察調査を行うことが可能である。

このように、主に形態的派生に基づく動詞の下位類の語数計算をすることは、日本語においても英語においても、いまだになされていない。本発表では、一般に入手できる日英語の一億語規模の書き言葉コーパスや各種の電子化辞書を用いて、日英語における動詞語彙の下位類リストを作成し、さらにそれら下位類ごとの使用頻度等を調査することによって、日英語における動詞語彙の量的構造を明らかにしたい。

■10月4日(日)

【講演】

コーパスからの情報抽出と抽出データの意味づけ：具体的表現からの抽象化

講師 滝沢直宏(立命館大学)

英語学・英語教育学関係の研究者が一般に用いているコーパスは BNC や COCA などであるが、これらのコーパスの annotation 情報は、品詞情報とレンマのみである。このことは、統語構造をはじめとした抽象的情報に依拠した検索が不可能であることを意味する。したがって、少し気のきいた情報を得ようとすると、語、品詞情報、レンマのみに基づいて抽象的情報をあぶり出す方法を個別事例に即して考える必要が生じる。

本講演では、様々な種類の個別事例においてコーパスからの情報抽出をどのように行うか、抽出されたデータをどのように意味づけるのかを、講演者自らの試行錯誤の経験、また失敗事例も交えて考えたい。更に、コーパスに基づいて何が主張でき何が主張できないかなどについても問題提起を行いたい。

■10月4日(日)

【研究発表第5セッション】

【研究発表1】

日本人英語学習者ナラティブコーパスによる「話し言葉」「書き言葉」の研究

山口有実子(東海大学)・宇佐美裕子(東海大学)

コーパス言語学の分野において、学習者コーパスを英語教育に応用した研究は比較的多くなされているが、第二言語習得分野に応用した研究はまだ少ない(投野・金子・杉浦・和泉, 2013)。また、一般公開された学習者コーパスは書き言葉が中心のため、話し言葉は少なく、特に、同一教材を用いた同一学習者による話し言葉と書き言葉を収集した学習者コーパスは稀である。一方、第二言語習得研究では、小規模データに基づく語彙や文法の発達に関する研究が多く、大規模学習者コーパスに基づいた検証は十分になされていない。近年多くの第二言語習得研究の枠組みとして用いられている「処理可能性理論」(Processability Theory, Pienemann, 1998; Pienemann, Di Biase, & Kawaguchi, 2005) で予測される普遍的発達段階についても、大規模学習者コーパスによって検証した先行研究はなく、各発達段階において出現が予測されている文法項目が話し言葉と書き言葉で同じように出現するのかについても Tan and Zhang (2014) による小規模データの研究(中国人英語学習者 5 人)のみである。

そこで本研究では、日本人英語学習者の大規模ナラティブコーパスを構築し、英語学習者の話し言葉と書き言葉における文法、語彙の使用の違いを解明することを目的とする。まず、日本人大学生による文字のない絵本を題材とした口頭ナラティブと同じ題材を用いた筆記ナラティブを収集し、口頭、筆記を行う順番が文法項目と語彙の産出数や正確さに影響する可能性を踏まえて、口頭から始める学習者と筆記から始める学習者の数なるべく同数になるよう考慮した。口頭ナラティブは録音して文字化し、筆記ナラティブは紙面に書かせたものをコンピュータに入力した後、CLAWS による品詞タグ付け、処理可能性理論に基づく各文法項目の正用・誤用のコード付けを行い、アンケート調査で得た英語学習歴についての情報を属性として加えたコーパスを構築した。本発表では、WordSmith による語彙リストや

n-gram, 文の複雑さ, のべ語数と異なり語数, 文法項目の産出と使用の正確さ等の分析に基づいて, 学習者の話し言葉と書き言葉にどのような違いがあるかを明らかにすると共に, 日本でのみ英語を学んできた学習者と海外で英語を学んだ経験のある学習者における文法と語彙の使用を比較し, 海外経験の有無が話し言葉と書き言葉にどのような影響を及ぼしているかを検証する。また, 処理可能性理論で予測される普遍的発達段階について, 学習者の話し言葉, 書き言葉にて検証する。

【研究発表 2】

日本人大学生の書く英語論証文の論理構造を探る —4つの分析的枠組みを用いて

山下美朋 (関西大学大学院 S)

日本人の談話における論理構造や修辭的特徴を分析した研究は, 対照修辭学の一つとして Kaplan (1966) 以降 1970 年代を中心に広く行われてきた。特に日本人学習者の英作文の特徴としては, 西洋人が演繹的 (topic-first) であるのに対し, 帰納的 (topic-delayed) であること, 読み手による書き手が意図する内容の推測 (Hinds, 1983) 及び論理展開に一貫性がないことや論拠の欠落 (Oi, 2005) などが挙げられ, これらの特徴の母語転移の影響を分析した研究も多い (Kubota, 1998; Hirose, 2005)。一方で英作文コーパスによる誤用分析により, 談話分析の観点から日本人学習者の書いた英作文の特徴を明らかにする試みも行われているが, 語彙や文法に焦点が置かれた研究が多く, その数は少ない。

本発表では, まずテキストの論理構造や修辭的特徴を分析するために発表者が考案した 4 つの分析的枠組み—1) Structural-Functional Analysis, 2) Keywords-Chain Analysis, 3) Metadiscourse Markers (MDM) Mapping, 4) Rhetorical Structure Theory (RST) (Yamashita, 2015) —について説明する。この枠組みは, 1) 英文には適切な構成要素がある (染谷, 1994; Oshima et al., 2006), 2) テキストには一貫性, 結束性がある (Halliday & Hasan, 1976), 3) 文と文, パラグラフ間には適切な論理関係がある (Mann & Thompson, 1988), といった英文の特徴を根拠とし, 妥当な論理分析が可能となるよう考案した。(なお, このうち 4) については Mann & Thompson (1988) の RST モデルを改変したものである。) 続いて, これらの枠組みを用いて習熟度の異なる日本人大学生が英語で書いた論証文 (60 編) と英語母語話者の論証文 (20 編) を分析した結果明らかになった各々の論理構造や展開, 修辭的特徴について計量的かつ質的に詳述する。なお, この分析的枠組みでは, テキストの論理構造が特定のタグの連鎖により視覚的に表わされ, 研究者のみならず学生が自らのテキストの論理構造を分析できるようになっている。本研究は, 将来的な大量データ分析の予備的研究として分析はほぼ手作業で行った。

分析の結果, 日本人学生の英作文が英語母語話者のそれよりもより「型」にはまった構造を持ち, 特定の論理連結詞を使用するなど, 論理展開に特徴的なパターンがあることが分かった。特に, 日本人に特徴的な「論理の逸脱」が顕著に見られ, 様々な逸脱, つまり本論に関係のない話題や意味不明な内容の挿入, 及び説明不足な記述, などが見られた。またこれらの特徴は習熟度に関わらず発見された。本研究は, 現在編纂中である「関西大学バイリンガルエッセイコーパス」(山西・水本・染谷 2013) のデータの一部を使用したものであり, 本発表ではこのプロジェクトの概要についても簡単に触れる予定である。

【研究発表 3】

中学英語教師向け英日発話パラレルコーパス構築と教師教育応用への展望

片桐徳昭 (北海道教育大学 旭川校)

本発表では, 国立教員養成系大学附属中学校での英語非母語話者教師 5 名による授業中の発話に焦点を当てた日英パラレルコーパスの構築の試みと, 日本語発話を英語に翻訳したコーパス開発とその分析例の報告を行う。ここで定義する「パラレルコーパス」とは元の発話が日本語の場合にのみ英訳したものを併記し, 元の発話が英語の場合, 日本語訳は併記しない形式をとる。

初期の教師教育においてはコーパス使用の利点がわかっており, 教師が使用すべき言語的気づき (linguistic awareness) を助長することにコーパス応用領域の一つがあるとされる (O'Keeffe & Farr, 2003, p. 393)。録音した発話の書き起こしで作成する発話コーパスは時間と費用がかかり, 発話コーパスとして大規模コーパスと呼ぶには 100 万語が目処とされるが, 教師教育を考えた場合は, 10 万語以下でもその発話データは教育現場の内容に即した豊かな内容が含まれる (O'Keeffe & Farr, 2003, p. 409)。

発表では, 主にこれまで収集した約 7,000 語の教室英語発話と, 日本語の部分英語に翻訳した約 7,500 語のデータをもとにした研究報告をする。*n*-gram を用いた語彙分析では, 英語の発話と翻訳とでは高頻度語(上位語)で英語と翻訳に共通するものが多く存在した。日本語の発話と英語発話とのコンテ

クスト分析では、日本語と英語の発話の出力に幾つかのパタンがあることがわかった。また、日本語から翻訳した英語に対するネイティブのコメントから「等価」の問題(若松・石井・中條, 2015, p. 39 など)やデータの蓄積に関するフォーマットの課題も明らかとなった。

「データ駆動型学習」といったようなコーパス資料を基にした学習者支援の研究は盛んであるが(中條, 2013; 内堀・中條, 2010 など多数), 教師の教室英語発話支援については研究例が少ない。よって、発話パラレルコーパスの蓄積とそれを応用した、英語非母語話者教員への英語発話支援は、中学校現場や、将来的には小学校教員養成課程ならびに現職の小学校教員支援において潜在需要があると考えられる。

■10月4日(日)

【研究発表第6セッション】

【研究発表1】

話し言葉コーパスを用いた談話辞 *absolutely* とその類義表現の分析

山本五郎(広島大学)

本発表では、談話辞 *absolutely* と類義の相槌表現の語法を話し言葉コーパスを用いて分析する。強意副詞としての *absolutely* については多くの語法書(Quirk et al. 1985, Biber et al. 1999)で取り上げられており辞書でも詳細な記述がなされているが、被修飾語を伴わず談話辞として単独で用いられる *absolutely* とその類義語については十分な研究がなされてこなかった。

Tao(2007)では、Michigan Corpus of Academic Spoken English をはじめとする複数の話し言葉コーパスに基づき、話し言葉における *absolutely* の35%が相槌表現として用いられることを確認した上で、その意味と用法について分析を行っている。しかし研究の焦点は、被修飾語を伴う強意副詞としての *absolutely* と談話辞 *absolutely* の関連に置かれており、相槌表現としての *absolutely* の談話上の機能や用法については十分な分析がなされているとは言えない。また、類義語については、語法書や辞書に記載されている被修飾語を伴う強意副詞の類義語 (*totally, completely, perfectly, entirely, utterly* など) は、相槌表現として用いられる *absolutely* の類義語ではないことを述べるにとどまっており、類義の談話辞についての用法や *absolutely* との談話機能の違い等についての分析はなされていない。

この談話辞 *absolutely* について、内田(2009)では談話上の機能に注目し、その用法を4つに分類している。内田では、(1) 相手の言ったことばに対して強く同意を表す、(2) 感想や意見を求められて強い賛同の意を表す、(3) 勧誘に対して喜んで承諾することを伝える、(4) *Not* を伴うと、意見を求められて強く否定することや要請を強く拒否する応答になる、という用法を提唱し、それぞれについて文例を挙げて説明している。また、類義の談話辞として *definitely* と *certainly* を取り上げ、*definitely* については「*absolutely* とほぼ同様に用いられる」、*certainly* については「*absolutely* の方が強意的」という説明をそれぞれ与えている。

内田では、*certainly* についても項目を設けており、文例を挙げながら(1) 依頼・勧誘に対する快諾、(2) 強い同意、という用法を認めているが、これらの用法は *absolutely* の用法とほぼ同じであり、*absolutely* の方が強意的であること以外に用法や談話機能の違いについての記述はない。談話辞としての *definitely* については独立した項目としては取り上げられておらず、用例や例文の提示はされていない。

上記の先行研究に見られるように、談話辞 *absolutely* に関しては、特に類似表現との比較について十分研究対象になっているとは言えない。本発表では、*absolutely* にあわせて内田で類似表現として提示されている *certainly* と *definitely* を取り上げ、WordBanksOnline の話し言葉コーパスや映画英語コーパスの用例を提示し、オーラルコミュニケーションにおけるこれらの談話辞の意味と用法について考察する。

【研究発表2】

男女両名義を使用する作家の作品判別—Rowling と Sheldon

木村美紀(明治大学大学院S)・久保田俊彦(明治大学)

一般に計量的な著者判別は著者の意識的コントロールの及ぶにくいと考えられる要素を変数とし、作品群を著者毎に分類、帰属させようとする試みである。複数の名義を使って作品を発表する作家は変数の有効性検証に格好の材料を提供してくれる。複数名義の使用とは作品群を複数構成するために著者が意識的に何らかの書き分けを行っているということであり、著者判別のための変数は、著者間の相違と著者内の書き分けの両者に異なる対応をすることが求められるからである。

日本語作品による複数名義と著者判別の問題を扱った金・樺島・村上(1993)では、長谷川海太郎の3名義による作品群が対照群とは異なる「長谷川群」として区別されるものの、長谷川群内で名義毎に区

別「されない」事実を示すことによって、読点前文字という変数の有効性を示した。

本発表ではこの問題の延長上に、男性名義を使った女性作家作品群判別を考える。歴史的には、Bronte 姉妹、George Eliot に代表されるように男性名義を使用した女性作家は少なくない。本発表では近年の例となる Joanne Rowling (J.K. Rowling 名義と Robert Galbraith 名義) と Alice Bradley Sheldon (James Tiptree, Jr 名義と Raccoona Sheldon 名義) の二人の作品を取り上げる (Rowling 10 作品, 1,581,756 語, Sheldon 72 作品, 865,802 語)。Rowling, Sheldon とも男女両名義の作品があるだけでなく、Sheldon はアイデンティティー「発覚」の後も両名義の作品が発表されているため、この発覚が作品に与えた影響、文芸批評との異同についても比較可能となる。

比較は 2 著者、4 名義の直接比較を行うと共に、Galbraith, Tiptree と似たジャンル、あるいは批評上比較された作家の作品群 (Galbraith に対する Val McDermind, Tiptree に対する Ernest Hemingway) を加えた判別を試みる。分析には金・樺島・村上 (1993) と同様に教師無し分類として主成分分析およびクラスター分析 (階層および k -means) を用いると共に、教師あり学習による判別 (判別分析および SVM) を併用する。変数には著者判別研究において評価の高い高頻度語彙使用率を使用する。名義によって対象読者、ジャンルを変えている Rowling の作品群を類似ジャンルの作品と比較することで、一見著者のコントロールの範囲内にあるように思われるこの変数の著者判別における有効性を確認する。Rowling, Sheldon の直接比較では、教師なしの分析において両著者 (2 群) の区別に成功するが、名義別の分析では Tiptree/Sheldon の判別結果が Rowling/Galbraith の判別結果を上回る。

【研究発表 3】

マルチモーダルコーパスを用いた英語母語話者による
パブリックスピーキング・パフォーマンスの分析：
発話速度、ジェスチャー、語彙特性の観点から

冬野美晴 (九州大学)・山下友子 (九州大学)

コーパス研究において、従来のテキストベースのモノモーダルコーパスに加え、音声・動画データ等を含むマルチモーダルコーパス研究が盛んに行われている (McCarthy, 1998; Adolphs & Carter, 2013; Fuyuno et al., 2015)。マルチモーダルコーパス研究においては、コミュニケーションの言語要素のみならず発話ポーズ、リズム、ジェスチャーなどの非言語要素も分析可能となる (cf. Rohrbach et al., 2012)。本研究の目的は、マルチモーダルコーパスの枠組みを踏まえ、英語母語話者によるパブリックスピーキングに着目し、テキスト・音声データ・動画データから成るマルチモーダルコーパスを分析することで、英語パブリックスピーキング教育に役立つ指標を抽出することである。具体的な分析対象は次の 3 点である：1) 発話速度 2) アイコンタクト動作の頻度と時間長 3) リーダビリティを中心とした語彙特性。

社会においてグローバル人材育成の重要性が認識される中、英語による発信能力の向上が求められている。スピーキングスキルの一つとしてパブリックスピーキング・スキルが挙げられるが、日本人向けの英語パブリックスピーキングの教育は十分に研究されているとは必ずしも言えない (田村他, 2014; 仁科, 2012)。特に問題となるのが、パブリックスピーキング教材や学習者の評価方法において定量的・客観的指標に基づくものが少ない点である (山下・冬野, 2015; Fuyuno et al., 2015)。そこで本研究では、英語母語話者によるパブリックスピーキングデータを定量的に分析することで具体的指標を抽出する。

分析に用いたデータは英語母語話者によるアメリカの大学の卒業式祝辞スピーチである。アメリカの大学において、卒業式祝辞スピーチは政治家や起業家等を招聘し行われることが多く、訓練されたパブリックスピーカーによるデータが得られる。また、祝辞スピーチはデータの発話時間長や話者環境、話者と聴衆の配置等の要素も比較的均一であり、定量的分析に適していると言える。本研究では 10 名のスピーカーによる祝辞スピーチについて、音声データ・動画データ・テキストデータから成るマルチモーダルコーパスを用いて分析を行った。データ分析の結果、主に以下のことが明らかになった。1) 発話速度：日常会話よりも遅い速度であることが示唆された。2) アイコンタクト動作：各話者で頻度と時間長の傾向が異なるがスピーチパートごとの特徴に共通点が見られた。3) リーダビリティ：ガニングフォグ指数を基に分析すると、平易な表現を意識して構成されていることが分かった。

《大会参加者へのご案内》

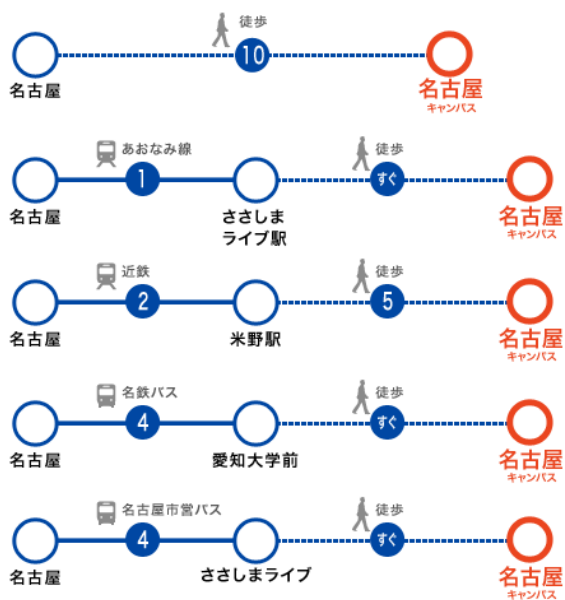
- ・ ワークショップの受付：会場の愛知大学名古屋キャンパス講義棟 7 階 708 教室前で、午前 9 時 30 分から受付を行います。
- ・ 大会受付：第 1 日（10 月 3 日）は愛知大学名古屋キャンパス講義棟 9 階ロビーで正午から行います。第 2 日（10 月 4 日）は午前 9 時 10 分から受付いたします。
- ・ 構内での喫煙は原則できません。構内禁煙にご協力いただきますようお願いいたします。
- ・ 第 1 日（10 月 3 日）の昼食は、学生食堂は閉店していますが、学内の「すき家」は営業しています。なお、一階にコンビニの「ミニストップ」が入っています。大学を出て約 20 メートルのところに、「サイゼリヤ」の他、ハンバーグ、イタリアン、オムライス、カレー等を提供するレストラン街があります。また、キャンパスから約 50 メートルのところに JICA のレストランがあり、11:30-14:00 までランチメニュー 700 円があります。第 2 日（10 月 4 日）は、学内の「すき家」は閉店していますが、その他は JICA も含めて土曜日と同様です。
- ・ 当日会員について：会員ではない方も、「当日会員」としてご参加いただけます（会費 2,000 円、二日間共通）。懇親会へもぜひご参加下さい。
- ・ 大会第 1 日の学術プログラム終了後の懇親会では、インフォーマルな雰囲気の中、参加者同士さまざまな意見交換、情報収集ができる場です。大会ご出席の方々には、ぜひ奮ってご参加いただけましたら幸いです。名古屋ならではの料理をはじめ、ソフトドリンクやビールに加え、地元の美味しい日本酒、ワインや焼酎も用意しております。どうぞご期待ください。なお、会場準備の都合上、参加ご希望の方には事前の予約をお願いしております。ご協力のほどよろしくお願い申し上げます。

- ・ 英語コーパス学会第 41 回大会・懇親会
- ・ 日時：10 月 3 日（土）18:15-20:00
- ・ 場所：キャンパスレストラン
- ・ 会費：4,000 円

※懇親会参加ご希望の方は、9 月 21 日（月）までに <http://goo.gl/forms/qdMdYXzADa> にアクセスして参加申込用 Web フォームに必要事項をご記入下さい。（申込締切日：9 月 21 日）。

◆愛知大学名古屋キャンパスへのアクセス◆

● 丸内の数字はおおよその所要時間です。

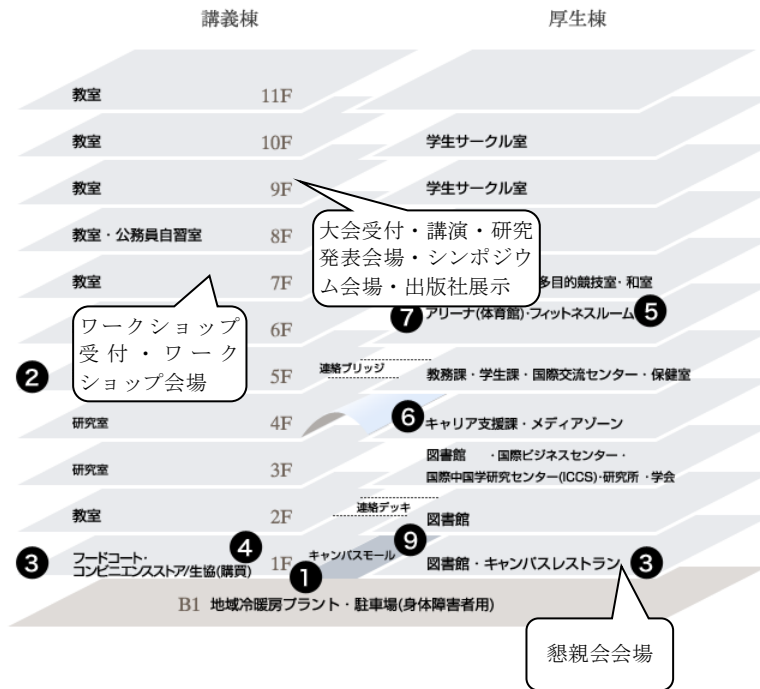


※名古屋駅から愛知大学名古屋キャンパスへは「あおなみ線（1 駅 1 分）」を利用するのが便利です。料金は片道 200 円です。

あおなみ線 名古屋駅発 金城ふ頭駅行時刻表

| 時刻 | 土日・祝日 | | | | |
|----|-------|----|----|----|----|
| 5 | 30 | | | | |
| 6 | 0 | 20 | 40 | | |
| 7 | 0 | 17 | 29 | 40 | 50 |
| 8 | 3 | 16 | 28 | 40 | 50 |
| 9 | 0 | 15 | 30 | 45 | |
| 10 | 0 | 15 | 30 | 45 | |
| 11 | 0 | 15 | 30 | 45 | |
| 12 | 0 | 15 | 30 | 45 | |
| 13 | 0 | 15 | 30 | 45 | |
| 14 | 0 | 15 | 30 | 45 | |
| 15 | 0 | 15 | 30 | 45 | |
| 16 | 0 | 15 | 30 | 45 | |
| 17 | 0 | 15 | 30 | 43 | |
| 18 | 0 | 12 | 29 | 45 | |
| 19 | 0 | 15 | 30 | 45 | |
| 20 | 0 | 15 | 30 | 45 | |
| 21 | 0 | 15 | 30 | 45 | |
| 22 | 0 | 18 | 40 | 57 | |
| 23 | 17 | 36 | | | |

◆会場案内図◆



詳細は下記URLより愛知大学名古屋キャンパスのウェブサイトをご覧ください。

<http://www.aichi-u.ac.jp/profile/campus-nagoya.html>

2015年8月29日発行

編集・発行 英語コーパス学会

会長 堀 正広

事務局 〒560-0043 大阪府豊中市待兼山町 1-8

大阪大学大学院言語文化研究科

田畑 智司研究室気付

電話：06-6850-5866

e-mail: jaecs.hq@gmail.com twitter: @JAECs2012

URL: <http://jaecs.com/>