

# 英語コーパス学会第38回大会

## ■第1日目

ワークショップ1《BNCを読む—機械を使わずしてコーパスの構造を理解する—》

会場：全学教育推進機構 B 棟 1 階 108 教室

日時：9月29日（土）10:00–12:00（9:30 受付開始）

講師：高橋 薫（豊田工業高等専門学校）

定員：定員 60 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、ワークショップ1・2の別、氏名、所属、会員・非会員の別を記入の上、電子メールで jaecs.workshop@gmail.com まで。

日時 2012年9月29日（土）  
受付開始 12:20（全学教育推進機構 B 棟 1 階 107 教室前）  
開会式 13:20（全学教育推進機構 B 棟 1 階 118 教室）

1. 会長挨拶
2. 開催校挨拶
3. 総会
4. 学会賞審査報告
5. 事務局からの連絡

司会 田畑智司（大阪大学）  
堀 正広（熊本学園大学）  
木村茂雄（大阪大学大学院  
言語文化研究科長）

〈研究発表第1室（全学教育推進機構 B 棟 1 階 118 教室）〉

司会 和泉絵美（京都外国語大学/同志社大学）

研究発表 1 14:20–14:50

二言語コンコーダンサーWebParaNewsとAntPConcを利用したDDL授業の実践  
中條 清美（日本大学）  
西垣 知佳子（千葉大学）  
アントニ ローレンス（早稲田大学）  
内山 将夫（情報通信研究機構）

研究発表 2 14:55–15:25

Advancing AntConc: Design and Performance Improvements for Multi-Language, Multiplatform  
Corpus Analysis by Researchers, Teachers, and Language Learners Laurence Anthony（早稲田大学）

研究発表 3 15:30–16:00

The design, development and purpose of the International Corpus of Crosslinguistic Interlanguage (ICCI)  
Yukio Tono（東京外国語大学）

〈研究発表第2室（全学教育推進機構 B 棟 1 階 108 教室）〉

司会 浮網茂信（大阪大谷大学）

研究発表 1 14:20–14:50

派生における新語形成についての考察—否定接頭辞付加派生語の場合—  
岡田 晃（大東文化大学非常勤講師）

研究発表 2 14:55–15:25

性格の違いは使用される語にいかに関与するか—二人の福音派牧師の比較  
石井昌子（京都大学大学院生）

〈休憩 16:00–16:20〉

特別講演 16:20–17:50（全学教育推進機構 B 棟 1 階 118 教室）

Lexical priming and the properties of text

司会 堀 正広（熊本学園大学）  
講師 Michael Hoey（リヴァプール大学）

《懇親会 時間：18:00–20:00 場所：学生交流棟 1 階、「宙 (Sora)」；会費：5,000 円》

■第2日目

ワークショップ2《XML文書としてのBNCの利用法》

会場：全学教育推進機構B棟1階108教室

日時：9月30日(日)9:30-11:00(9:00受付開始)

講師：園田勝英(北海道大学)

定員：定員60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：29日の参加費納入者は不要)。

※予約申し込みは、ワークショップ1・2の別、氏名、所属、会員・非会員の別を記入の上、電子メールでjaecs.workshop@gmail.comまで。

日時 2012年9月30日(日)  
受付開始 10:40(全学教育推進機構B棟1階107教室前)

〈研究発表第1室(全学教育推進機構B棟1階118教室)〉

司会 大谷直輝(埼玉大学)

研究発表1 11:15-11:45

英語接頭辞over-の数量的分析

木山直毅(大阪大学大学院生)

研究発表2 11:50-12:20

副動詞を含む移動構文と使役移動構文の類似点と相違点

森下裕三(神戸大学大学院生)

〈研究発表第2室(全学教育推進機構B棟1階108教室)〉

司会 堀田秀吾(明治大学)

研究発表1 11:15-11:45

コーパスを活用したシノニム研究—英和辞典の記述改善に向けて—

島田祥吾(広島大学大学院生)

研究発表2 11:50-12:20

高汎用性コーパスシステムを用いた英語教育および学習支援に関する研究

藤野玄大(東北学院大学大学院生)

岡田毅(東北大学)

坂本泰伸(東北学院大学)

〈昼休憩 12:20-13:20〉

シンポジウム 13:20-15:20(全学教育推進機構B棟1階118教室)

《私のコーパス利用》

司会 赤野一郎(京都外国語大学)

講師 滝沢直宏(名古屋大学)

深谷輝彦(椋山女学園大学)

岡田毅(東北大学)

家入葉子(京都大学)

野ロジュディー(武庫川女子大学)

投野由紀夫(東京外国語大学)

井上永幸(広島大学)

田畑智司(大阪大学)

堀正広(熊本学園大学)

閉会式 15:25(全学教育推進機構B棟1階118教室)

閉会の辞

田畑智司(大阪大学)

英語コーパス学会 (Japan Association for English Corpus Studies)

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: @JAECSS2012 郵便振替口座:00930-3-195373

URL: <http://english.chs.nihon-u.ac.jp/jaecs/>

- ◆ 大会当日、入会受付も致しますので、お誘い合わせの上ご参加下さい(年会費 一般5,000円 学生3,000円)。
- ◆ 「当日会員」としてのご参加も受け付けております(ワークショップも含め、二日間を通して2,000円)。

# 英語コーパス学会 第 38 回大会資料

日時：2012 年 9 月 29 日（土）－ 30 日（日）  
会場：大阪大学（豊中キャンパス）全学教育推進機構 B 棟  
(<http://www.osaka-u.ac.jp/>)  
〒560-0043 大阪府豊中市待兼山町 1-16

■第1日目

ワークショップ1《BNCを読む—機械を使わずしてコーパスの構造を理解する—》

会場：全学教育推進機構 B棟1階108教室

日時：9月29日(土) 10:00-12:00 (9:30受付開始)

講師：高橋 薫 (豊田工業高等専門学校)

定員：定員60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、ワークショップ1・2の別、氏名、所属、会員・非会員の別を記入の上、電子メールで jaecs.workshop@gmail.com まで。

日時 2012年9月29日(土)  
 受付開始 12:20 (全学教育推進機構 B棟1階107教室前)  
 開会式 13:20 (全学教育推進機構 B棟1階118教室)

- 司会 田畑智司 (大阪大学)  
 堀 正広 (熊本学園大学)  
 木村茂雄 (大阪大学大学院  
 言語文化研究科長)
1. 会長挨拶
  2. 開催校挨拶
  3. 総会
  4. 学会賞審査報告
  5. 事務局からの連絡

〈研究発表第1室 (全学教育推進機構 B棟1階118教室)〉

司会 和泉絵美 (京都外国語大学/同志社大学)

研究発表1 14:20-14:50

二言語コンコーダンサーWebParaNewsとAntPConcを利用したDDL授業の実践  
 中條 清美 (日本大学)  
 西垣 知佳子 (千葉大学)  
 アントニ ローレンス (早稲田大学)  
 内山 将夫 (情報通信研究機構)

研究発表2 14:55-15:25

Advancing AntConc: Design and Performance Improvements for Multi-Language, Multiplatform Corpus Analysis by Researchers, Teachers, and Language Learners  
 Laurence Anthony (早稲田大学)

研究発表3 15:30-16:00

The design, development and purpose of the International Corpus of Crosslinguistic Interlanguage (ICCI)  
 Yukio Tono (東京外国語大学)

〈研究発表第2室 (全学教育推進機構 B棟1階108教室)〉

司会 浮網茂信 (大阪大谷大学)

研究発表1 14:20-14:50

派生における新語形成についての考察—否定接頭辞付加派生語の場合—  
 岡田 晃 (大東文化大学非常勤講師)

研究発表2 14:55-15:25

性格の違いは使用される語にいかにも現れるか—二人の福音派牧師の比較  
 石井昌子 (京都大学大学院生)

〈休憩 16:00-16:20〉

特別講演 16:20-17:50 (全学教育推進機構 B棟1階118教室)

Lexical priming and the properties of text

司会 堀 正広 (熊本学園大学)  
 講師 Michael Hoey (リヴァプール大学)

《懇親会 時間：18:00-20:00 場所：学生交流棟1階、「宙 (Sora)」；会費：5,000円》

■ 第 2 日目

ワークショップ 2 《XML 文書としての BNC の利用法》

会 場：全学教育推進機構 B 棟 1 階 108 教室

日 時：9 月 30 日（日）9:30-11:00（9:00 受付開始）

講 師：園田勝英（北海道大学）

定 員：定員 60 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：29 日の参加費納入者は不要）。

※予約申し込みは、ワークショップ 1・2 の別、氏名、所属、会員・非会員の別を記入の上、電子メールで jaecs.workshop@gmail.com まで。

日 時 2012 年 9 月 30 日（日）  
受付開始 10:40（全学教育推進機構 B 棟 1 階 107 教室前）

〈研究発表第 1 室（全学教育推進機構 B 棟 1 階 118 教室）〉

司 会 大谷直輝（埼玉大学）

研究発表 1 11:15-11:45

英語接頭辞 over- の数量的分析

木山直毅（大阪大学大学院生）

研究発表 2 11:50-12:20

副動詞を含む移動構文と使役移動構文の類似点と相違点

森下裕三（神戸大学大学院生）

〈研究発表第 2 室（全学教育推進機構 B 棟 1 階 108 教室）〉

司 会 堀田秀吾（明治大学）

研究発表 1 11:15-11:45

コーパスを活用したシノニム研究—英和辞典の記述改善に向けて—

島田祥吾（広島大学大学院生）

研究発表 2 11:50-12:20

高汎用性コーパスシステムを用いた英語教育および学習支援に関する研究

藤野玄大（東北学院大学大学院生）

岡田 毅（東北大学）

坂本泰伸（東北学院大学）

〈昼休憩 12:20-13:20〉

シンポジウム 13:20-15:20（全学教育推進機構 B 棟 1 階 118 教室）

《私のコーパス利用》

司 会 赤野一郎（京都外国語大学）

講 師 滝沢直宏（名古屋大学）

深谷輝彦（椙山女学園大学）

岡田 毅（東北大学）

家入葉子（京都大学）

野ロジュディー（武庫川女子大学）

投野由紀夫（東京外国語大学）

井上永幸（広島大学）

田畑智司（大阪大学）

堀 正広（熊本学園大学）

閉 会 式 15:25（全学教育推進機構 B 棟 1 階 118 教室）

閉会の辞

田畑智司（大阪大学）

## ■ 1 日目

### 【ワークショップ 1】

#### BNC を読む —機械を使わずしてコーパスの構造を理解する—

講師 高橋 薫（豊田工業高等専門学校）

英語コーパスの研究がより複合的になる中、入門者が初期のコーパスにあまり触れずに先進的な研究を進める状況もあるように思われる。コーパス間の比較分析を行うにせよ、学習者コーパスを構築するにせよ、コーパスについての基礎的知識を持つことは、コーパス活用の発展的研究のためには必須であると考えられる。

本ワークショップでは、今では幾分古典的コーパスとも言える The British National Corpus (以下 BNC) のバージョンアップに伴う進化の過程をタグ、属性、コーパスサイズ、ファイル形式、ウェブアクセスという観点で考察する。これを効率的に概観するため、パソコンを用いた実習形式とはしない。

この BNC のワークショップでは、分析が手つかずの部分が少ないことを次第に明らかにして、今後の言語分析の注目点を提案する。とりわけ、社会言語学的な視野、言語スタイルの観点から研究を進めたい参加者には有益なものとしてほしい。また、これまで予想外に BNC の分析が滞っている分野について、その原因を明らかにして改善策を提起したい。そのため、BNC の活用条件が整っている参加者には、今後の研究の支援として解析用のソフトウェアを配布する予定である。

### 【研究発表第 1 室】

#### 【研究発表 1】

##### 二言語コンコーダンサー WebParaNews と AntPConc を利用した DDL 授業の実践

中條 清美（日本大学）  
西垣 知佳子（千葉大学）  
アントニ ローレンス（早稲田大学）  
内山 将夫（情報通信研究機構）

本研究は、新しく開発され、公開された二言語検索サイト WebParaNews とダウンロード可能な二言語検索ツール AntPConc を使って Data-driven Learning (DDL) を英語授業に取り入れた授業実践の結果の報告である。DDL を実際の外国語指導に取り入れるには、英語コーパステキストの難しさ、コーパス検索ツールの操作性の問題など解決すべき問題点が多い。そのため、これまでのところ、教育現場からの DDL 教材作成の試みや実践の報告は多くない。それに対して発表者らは、DDL における学習者の英文理解の負荷や不安を軽減するために、検索結果の英文に対応する日本語が示される日英パラレルコーパスを使って、DDL を用いた英語授業を過去 8 年間行って効果を検証してきた。その実践成果を踏まえ、DDL を教育現場に広げるために、2008 年にユーザーフレンドリーな二言語コーパス検索エンジン AntWebConc-Parallel の開発を開始し、検索スピードやコンコーダンス画面仕様の改善を行ってきた。その結果、2011 年には教室指導および家庭学習での DDL に活用できる実用レベルに達した。2012 年 8 月には日英新聞記事対応付けデータ（内山・井佐原, 2003）を AntWebConc-Parallel で検索できる WebParaNews のサイトを無償公開した（本公開は情報通信研究機構との知的財産利用契約に基づくものである）。また、ダウンロード可能な二言語検索ツール AntPConc の開発も進行しており、すでに試作版が公開されている。

発表では、コーパスを利用した DDL の普及に向けて、1) 日英新聞記事対応付けデータを Web 上で利用できる WebParaNews を利用して、大学初級・上級レベル英語学習者が DDL 学習を行った教育利用例と指導効果の報告、2) 教師が作成した DIY 二言語コーパスと AntPConc を利用して大学リメディアルレベル学習者が一般英語授業において、苦手な文法項目を中心に自立的に文法学習を行った DDL 指導実践例、また、3) 中・高英語教員や英語教員志望者が、AntPConc を教材作成に利用して作った DDL 教材の例を報告する。

## 【研究発表 2】

### Advancing AntConc: Design and Performance Improvements for Multi-Language, Multiplatform Corpus Analysis by Researchers, Teachers, and Language Learners

Laurence Anthony (Waseda University)

One of the key components of any corpus study is the software used to perform the analysis. In recent years, the *AntConc* software tool has grown in popularity among researchers, teachers, and language learners due to its rich set of features, freeware license, multiplatform support, and easy-to-use interface. In 2011 alone, the software was downloaded 120,000 times by users in over 80 countries. For researchers, the software performs speedily and accurately on a wide-range of small and large corpora. It also offers flexible handling of tags, metadata, and language encodings, and provides a wealth of functions and features. For teachers and learners, the software can perform basic corpus functions, such as concordancing and producing keyword lists, in a quick and easy way. Also, it can be used both inside the classroom and as part of student homework projects on Windows, Macintosh OS X, and Linux computers. Finally, to motivate learners to use corpora in their learning, it offers a modern and attractive-looking interface.

Although *AntConc* has many strong features, when compared to the most popular commercial tools, it has a number of weaknesses. To address these, various design and performance improvements have been introduced in the latest version of the software. In this presentation, I will explain these changes and compare the new version of *AntConc* with previous versions and other comparable freeware and commercial tools. As part of the discussion, I will explain the importance of including a flexible and explicit token definition that can incorporate characters from any of the world's languages. I will also explain how different tools handle whitespace issues and show that the implementation adopted in *AntConc* offers far greater transparency and flexibility. At the end of the presentation, I will demonstrate how the new version of *AntConc* can be used effectively by both teachers and learners in a Data-Driven Learning (DDL) classroom environment, with practical examples taken from the technical writing classroom.

## 【研究発表 3】

### The design, development and purpose of the International Corpus of Crosslinguistic Interlanguage (ICCI)

Yukio Tono (Tokyo University of Foreign Studies)

While learner corpus research is gaining grounds and the types of data and research topics have been diversified, there is still one serious problem. That is, research from the beginning stages of learning or acquisition is still very scarce. This is primarily because most learner corpus building projects to date focus on adult learners, especially university students, due to the ease of data collection from university students. Most large learner corpora were constructed by collecting learners' essays, either online or in an electronic format (e.g., Word documents), whereas younger learners usually do not access PCs or the Internet for writing. The same thing can be applied to commercial publishers' learner corpora, such as the Cambridge Learner Corpus (CLC). It consists of Cambridge exam scripts written mainly by those who wanted to study at universities in the U.K. Thus, the primary population for the corpus consists of young adults and older people who have already finished their secondary school education. Although they have exams for younger learners, their corpus breakdown shows that there is a lack of such data.

With this background in mind, the project called the "International Corpus of Crosslinguistic Interlanguage (ICCI)" was recently completed, as a part of the five-year Global COE project (2007-2011), granted to Tokyo University of Foreign Studies (TUFS) by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT). It consists of 9,000 English novice to lower-intermediate level learners across seven countries/ regions (Austria, China, Hong Kong, Israel, Japan, Poland, Spain, and Taiwan). In this presentation, the results of the project will be reported and several important issues will be discussed, including the challenge of collecting young learners' production data, the corpus design criteria and the reality of data collection, and the development of a user-friendly web corpus query tool to suit the needs of accessing the ICCI. As an example of possible applications, the on-going work on analysing learner data to find criterial features for CEFR levels will be presented and implications for future research and pedagogy will be discussed.

## 【研究発表第2室】

### 【研究発表1】

#### 派生における新語形成についての考察—否定接頭辞付加派生語の場合—

岡田 晃（大東文化大学非常勤講師）

英語の語彙増大に大きく関わっている語形成過程には、主に、複合と派生の2つがある。いわゆる自由形態素を2つ以上組み合わせる新語を形成する複合とは違い、派生は既存の語に接頭辞や接尾辞といった拘束形態素を付加し、新語を形成することをいう。この派生における新語というのは母国語話者の意識とは関係なく偶発的に出現することがある。それは、母国語話者の脳内にある既存の語を介して推測を行い、接辞付加の選択を無意識的に行っているということである。実際これまでの研究では、Jespersen (1917)や Marchand (1969)等に見られるように、in-は外来系の接頭辞であることから同じ外来系の語に付加され、un-は本来語と外来語に付加が可能である、といった史的観点からの考察が多かった。しかし実際は英語話者の脳内に史的な観点が即座に浮かび上がり接辞選択をしているとは到底考えにくく、英語内に定着した語を基に、そこから導き出される類推が接辞選択のカギとなっていると考えられる。Chapman and Skousen (2005)は形容詞 *abrogated* の否定接頭辞付加の可能性を既存の語を介し、接頭辞と接尾辞との関連、基体との音声的密着性を考慮に入れ、類推(analogy)で分析し、un-とdis-のどちらの接頭辞が今後使われていくかを述べている。

当発表では British National Corpus を使い、頻度数1の in- / un- 付加二重語(同じ基体に異なる2つの接辞 in-と un-が付加された派生語)をいくつか抽出し、その派生語が今後どちらの接頭辞を有したまま英語内で使われていくかを考察していく。Chapman and Skousen の分析方法と同様、既存の語から推測して、(1)接頭辞と接尾辞との関連性、(2)接頭辞と基体との音声的関連性を主に考慮し、この2つの視点に、(3)史的関連性、を加えた3つの点に注目して考察を進めていく。また、in- / un-付加派生語だけでなく、比較考察する上で dis-や non-付加派生語を考慮に入れている。考察を進めていく上で必要な語彙情報は主に *The Oxford English Dictionary 2nd edition* と *Middle English Dictionary* に依っている。

### 【研究発表2】

#### 性格の違いは使用される語にいかにも現れるか—二人の福音派牧師の比較

石井昌子（京都大学大学院生）

本発表では、ジョージ・エリオット (George Eliot, 1819-80) 作『ジャンネットの悔悟』(*Janet's Repentance*, 1858) に登場するエドガー・トライアンとシャーロット・ブロンテ (Charlotte Brontë, 1816-55) 作『ジェイン・エア』(*Jane Eyre*, 1847) に登場するセント・ジョン (St. John) という、ともに19世紀初頭、熱心に布教活動に取り組み若くして死んだ誠実な英国教会牧師の性格を、sympathy と arrogance という仮説を立てて読み解き、その性格が2人の使用する語にいかにも反映されているかを、quantitative に検証した上で qualitative に分析する。文体論の先行研究は、Burrows (1987)に始まり、Craig (1992), Tabata (1994, 1995, 1998, 2002), Hori (1993, 1999, 2002) などがあるが、同じ時代の同じ職業の人間の性格と言葉の関係に言及したものは見当たらない。

性格に関する仮説は、これまでに批評家から指摘されてきた各々の性格描写に基づく。

*OED2 on CD-ROM Version 4.0*によれば、sympathyとは the fact or capacity of entering into or sharing the feelings of another or others であり、arroganceとは undue assumption of dignity, authority, or knowledge; aggressive conceit である。

作家の文体の違いを考慮して、『ジャンネットの悔悟』では、主人公ジャンネットとその夫デンプスター、『ジェイン・エア』では、主人公ジェインとその夫となるロチェスターの言葉も比較対象に入れて調査した。

方法は、まず Project Gutenberg を利用して当該人物の言葉(引用符号に入っていれば心の中のつぶやきも含む)を抽出し、Oxford Version (各1985年版と2008年版)と対照して修正した。次に AntConc 3.3のキーワード配列機能を利用して彼らの代名詞の使い方の違いに着目し、さらに N-grams 機能と concordance 機能を利用して、代名詞の使用頻度およびその周辺の語の違いを調べた。語数の比較は、1000語あたりのそれに直して行った。

その結果例えば、一人称複数代名詞を、トライアン(以後 T と略す)は15.1語(話し相手を含む意味では14.4語)、セント・ジョン(以後 S と略す)は4.7語(同3.3語)しか使わない。二人称代名詞の数は、逆に T が34.4で S の51.3より少ない。*we*の支配する語数(*we*の右側で新たな節に至る前に存在する語の数)の平均は、T が10.0、S は4.6である。主語 *you* の支配する語数(同)は、T が3.6、S 4.4で

あった。さらに主語 *you* に続く動詞 (助動詞) は、T は *know, can* など相手を肯定する語が多く、S は *must, may, cannot* など相手に命令したり、相手を否定する語が多い。I に続く動詞 (助動詞) を頻度の高い方から挙げると、T が *had, think, have, am, shall*、S は *have, am, know, shall, can* と続く。

その他の点を含め、同じ時代と同じく誠実に勤めを果たした牧師二人の性格について立てた *sympathy* と *arrogance* という仮説を、コーパスを使って彼らの言葉を分析することにより検証することができた。

## 【特別講演】

### Lexical priming and the properties of text

Michael Hoey (University of Liverpool, UK)

Lexical priming is a psycholinguistic theory of language that seeks to explain the corpus-linguistic categories of collocation, colligation and semantic association, as well as other matters such as grammatical category and pragmatic association, by drawing upon psycholinguistic research into the way that words are connected in the mind that is well established in the field of psychology but has been largely ignored by corpus linguists. Lexical priming theory can be used to generate corpus-linguistic hypotheses and one of the most important hypotheses that it generates is that traditional descriptions of the properties of text can all be explained by the notion of lexical priming. After a relatively brief description of the features of lexical priming theory and its psycholinguistic background, this paper will report recent corpus-linguistic research into the application of lexical priming theory into three of the key properties of text. The first of the properties looked at will be the lexical nature of cohesion; I shall seek to show that textual cohesion and corpus-linguistic collocation are two aspects of the same phenomena and contribute considerably to our ability to acquire a language. The second of the textual properties I shall look at will be the fact that the organisation of texts can be partly described in terms of the semantic relations, signalled by writers and perceived by readers, that hold between parts of the text. I shall present preliminary evidence to support the claim that ordinary non-signalling words may be associated in our minds (and in the texts we read) with particular types of semantic relationship between parts of the text. The final property of text that I shall look at will be the way we begin and paragraph texts. I shall argue on the basis of research funded by the AHRC and undertaken with Matt O'Donnell, Michaela Mahlberg and Mike Scott, that the way we paragraph and the way we begin texts is much more lexically-driven than is normally assumed by English Language teachers.

## ■第2日目

### 【ワークショップ2】

#### XML 文書としての BNC の利用法

講師 園田勝英 (北海道大学)

EPUB, Open XML, など、コンピュータ上で接する「文書」は、知らない間にXML形式だらけになってしまった。私達に直接関係するコーパスや辞書データも、近年入手できるものは、ほとんどがXML形式である。このような流れの中で、BNCも2007年の版で従来のSGMLからXMLに変わった。BNCがSGMLであったころは、SGMLの処理系そのものが一般的ではなく、またBNC自体が従来の平テキストを扱う方法の延長で処理することができるように配慮されていた。しかし、XML版になると、さすがに従来の平テキストを扱う手法では、太刀打ちできない。幸い、XMLの処理系は至る所にあって簡単に手に入るようになっている。本ワークショップでは、XML文書としてのBNCをどのように利用したらよいかを、具体的な例を交えて紹介する。また、この点について私自身手探りの状態であるので、フロアからのフィードバックを期待したい。

## 【研究発表第1室】

### 【研究発表1】

英語接頭辞 *over-* の数量的分析

木山直毅 (大阪大学大学院生)

構文文法(Croft 2001, Langacker 2005)において、「形態素+語幹」という「語」のレベルも構文と認める動きが今日では主流となっている。ところが、構文文法の立場から語のレベルの研究を押し進めたものは少ない(cf. Iwata 2004, Asao 2007)。そこで本研究は英語の接頭辞 *over-* を構文文法の立場から、量的に分析する。

接頭辞 *over-* は一般的に「過剰」と「時空間的」意味の2つが想定されている(由本2005, Iwata 2004)。そして歴史的には不変化詞や前置詞 *over* との関連が強く、それ故に例えば “*overflow the bank*” と “*flow over the bank*” は「言い換え」が可能とされる(由本 2005)。ところが、“*flow over X*” と “*overflow X*” を COCA で検索すると、目的語のタイプの重複が極端に少なく ( $n = 4$ )、目的語に *bank* をとるのは *overflow* のみであった。次に、“*overgrow*” と “*grow over*” を観察すると、生起する構文に差がある。“*grow over*” は現在形、現在分詞、過去分詞など、様々な使用があるのに対し、“*overgrow*” は過去分詞の用法に使用にほぼ限定される(BNC: 96%)。このような使用実態を考慮すると、接頭辞と前置詞を同等に扱って良いものかは疑問である。さらなる先行研究の問題点は、使用実態を重視していない点にある。故に、構文の下位分類をする際に、使用実態には沿っていない分類をしている(cf. Iwata 2004)。そこで本発表は以下の2つの構文を量的な視点から分析し、さらに (1) をより深く考察する。

(1). 接頭辞 *over-*: *over-V + Obj*

Either gives access to overflying Auckland city and harbour, perhaps visiting the airfield on Great Barrier Island. (BNC: CAU-1215)

(2). 前置詞 *over*: *V + over + Obj*

Legs fly over tambourines held steady by one or the other partner. (ibid: A12-1176)

本研究ではBNCから(1)と(2)に該当するデータを入手し、両構文に共通する動詞(*overlook vs look over* etc.)を分析対象とした。分析の手続きとして①Collostructional Analysis (Stefanowitsch and Gries 2003)によってデータの絞り込み②得られたデータのタグ付け③タグから構文の好む意味を抽出④②で付与したタグを変数として(1)の構文のみRandom Forestsによって分類、⑤その結果をクラスター分析で分類した。すると、理論言語学においてなされてきた分類とは異なった分類が提示された。

そこで本発表では、理論言語学において言われてきたことを、量的な側面からサポートしつつ、これまで追求されてこなかった接頭辞 *over-* の持つ多義性を考察する。

【研究発表 2】

副動詞を含む移動構文と使役移動構文の類似点と相違点

森下裕三 (神戸大学大学院生)

本発表は、言語類型論で副動詞 (Converb) と呼ばれる「非定形で副次的な意味を持つ」動詞 (*-ing* 形の動詞) を含む2種類の構文の類似点と相違点を明らかにする。具体例は以下の通りである。

(1) a. Joe **came running** through the plastic curtain. (BNC-HGF)

b. A mother **ran screaming** from the building...

(2) a. Then he **sent** the car **leaping** out of the garage. (BNC-B0B)

b. ...he **brought** it **slamming** down onto the desk top. (BNC-BMW)

(2)は(1)の使役形である。そのため、(1)を副動詞移動構文、(2)を副動詞使役移動構文と呼ぶ。これらの構文の特徴として、*-ing* 形の動詞が、主語または目的語が移動する際の様態 (e.g. *running, flying*) または付随的動作 (e.g. *whistling, screaming*) を表すという点が挙げられる。これら両構文の類似点と相違点について頻度を基にした議論を行う。

まず、類似点についてだが、*-ing* 形の動詞として高頻度で生起する動詞の多くが両構文に共通するという点が挙げられる。具体的には、副動詞移動構文では *running* が 129 例、*rushing* が 71 例、*crashing* と *flying* が 50 例、*tumbling* が 40 例、*walking* が 28 例、*hurtling* と *screaming* が 25 例生起することが確認できた。一方、副動詞使役移動構文では *crashing* が 50 例、*flying* が 23 例、*scurrying* が 22 例、*tumbling* が 21 例、*sprawling* が 16 例、*rushing* が 13 例、*running* が 10 例生起することが確認できた。これらの動詞のうち、*running, crashing, flying, tumbling, rushing* は両構文に高頻度で生起する。また、*walking* と *sprawling* を除いた全ての動詞には、比較的激しい動作を伴うものだという共通点がある。

次に、両構文の相違点についてであるが、定形の動詞と *-ing* 形の動詞との間に見られる関係が異なる。具体的には、副動詞移動構文では、直示的移動動詞 (e.g. GO と COME) が定形の動詞として生起する時には移動様態動詞 (e.g. *running*) が *-ing* 形の動詞として生起しやすい。しかし、移動様態動詞 (e.g. RUN) が定形の動詞として生起する時には活動動詞 (e.g. *screaming*) が *-ing* 形の動詞として生起しやすくなる。一方、副動詞使役移動構文では、定形の動詞の直示性ではなく、移動の開始時にのみ使役が行われる動詞 (e.g. SEND) であるか、移動中に継続して使役が行われる動詞 (e.g. TAKE や BRING) であるかによって、*-ing* 形の動詞として生起する動詞のタイプが異なる。

副動詞移動構文と副動詞使役移動構文には *-ing* 形として生起する動詞に共通点が見られるものの、定形の動詞と *-ing* 形の動詞の間に見られる関係は異なっている。これらの類似点と相違点は、動詞のス

ロットが2つあるという特殊性、および、移動のタイプの違いによる物理的な制約の違いによるものであると言える。

## 【研究発表第2室】

### 【研究発表1】

コーパスを活用したシノニム研究—英和辞典の記述改善に向けて—

島田祥吾（広島大学大学院生）

本発表では、コーパスを活用したシノニム研究の一環として接尾辞-lyを持つ副詞 *drastically* と *radically* に焦点を当て、両語の類似性と相違性を明らかにし、それらを英和辞典の記述にどのように反映させるかを提案する。本発表ではコーパスを活用したシノニム研究の1つの方向性を示すために、まずは意味領域が狭く、かつ兼ねてから EFL/ESL 辞典において記述が薄いと感じていた項目である接尾辞-lyを持つ副詞を取り上げる。同レベルの辞書で、副詞についてはまだまだ記述が不十分な点が多いが、その一例として、特に両語に意味分布の違いのわかる訳語が与えられていなかったのをこれらを選んだ。

語法研究の一環として扱われることが多かったシノニムに関する記述は、現在では各種 EFL/ESL 辞典においてはコラムの形で扱われているものが多く、重要視されていることが分かる。コーパス言語学が発展して以来、語法研究においてコーパスを活用したものが多くなったが、シノニム研究としてコーパスを利用しているものは多くない (e.g. 井上 (2001, 2010), 仁科 (2008), 鈴木 (2010))。シノニム研究の歴史を概観するには Gove (1984) がすぐに役に立つが、現在シノニムについて詳しく論じている代表的なものとして小西 (1976) を取り上げ、シノニムという場合どのようなものをシノニムと呼ぶか言及し、さらに自分の立場を明らかにする。

両語の比較研究を行う際に WordBanks Online を利用する。分析の際には *t*-score と MI-score を基に適宜 KWIC で実際を確認しながら分析してゆく。まず *t*-score から *drastically* と *radically* はどのような語と結びつきが強いのかを確認する。上位 50 位を見ると両語に共通して現れる動詞が確認できる。次に MI-score を見ると *t*-score で確認された語だけでなく *t*-score ではわからなかった結びつきが強いと思われる語があり、それらを観察すると大きく3つに分類することができる。すなわち、1. 変化を否定的に捉える動詞、2. 変化を中立的に捉える動詞、3. 変化を肯定的に捉える動詞である。この分類法に従うと、*drastically* と *radically* では両語に共通して現れる動詞はあるが、前者は1の動詞と共起する傾向が、後者は2の動詞と共起する傾向が見られる。これらのデータから両語の類似性や相違性を英和辞典の記述にどのように反映させるとよいかを考察する。

EFL/ESL 辞典で当該副詞を見ても両語の違いに関する情報は得られず、大型辞典でさえ一般サイズの辞書と類似した記述が多くみられた。辞書の記述の際には頻度に関わる問題や文字制限など様々な制約が考えられるが、英語学習者にとってより良い英和辞典となるように工夫が求められる。

### 【研究発表2】

高汎用性コーパスシステムを用いた英語教育および学習支援に関する研究

藤野玄大（東北学院大学大学院生）

岡田 毅（東北大学）

坂本泰伸（東北学院大学）

本発表では、高汎用性を実現したコーパス解析システムを活用しての英語教育および学習支援のあり方とその可能性について述べる。英語教育・学習を促進するために、これまでに開発・運用されてきている各種コーパスの多くは、外部からは見えにくい独自のフォーマットと処理モジュールによって構成されており、ユーザー間でのデータ共有や、ユーザーが独自の属性を付与し次の段階の分析に供したいような場合には、多大な労力と相当レベルのプログラミング能力を必要とすることがある。また、分析・研究の成果を受け取る側の研究者や学習者は、受け身的に処理結果を享受するだけの立場に置かれるという問題もある。これら諸問題を解決すべく、教育者・学習者の要求に応え、ユーザー間での情報共有を保証する汎用性の高いコーパス解析システムの開発を進めている。

英語教育・学習支援用のコーパスに求められる要素として、レマやコロケーション、ユーザー指定の語彙集という単語の集合体の他に、文書中で注目されるべき単語(群)や品詞そのものやその出現位置も重要である。本システムでは、Allocation of Remarkable Element (ARE)と呼ぶ概念を用い、文書中・センテンス中の特定箇所を管理している。これにより、教員が学習者に注目させたい部分をハイライトして表示することが可能となる。また、語彙集やレマリスト、同・反意語といった個々のユーザーで異なる定

義が要求される単語の集合体も、ユーザー自身が自由にシステムに組み込むことを可能にさせている。特定語彙の位置情報や単語の集合体に対し、RDBMS の特性を活かし利用目的に応じた複数の属性情報を関連付けることで、ユーザーの要求に合わせた教育支援や学習支援が実現できると考えている。

本発表では、システム概要と、それを活用した教育および学習支援の具体例の一部を示す。例えば“*Ask me no questions...*”のような、日本語的センスでは産出しにくいような英語独特の表現などを ARE 機能によって強調することで、学習者に対し英日の発想のコントラストなどを実感させることができる。また、ユーザー定義の自由な属性を ARE で指定する語群に対し付与させ、その情報を教員がリアルタイムで閲覧したり、学習者間で共有することにより、柔軟で効率的な相互学習が可能となる。さらに、レマに対し JACET 8000 に代表される単語難易度レベル属性を付与し、その情報を利用することで習熟度や学習目標に合致した学習支援が実現できるということを示す。

## 【シンポジウム】

### 私のコーパス利用

司会 赤野一郎（京都外国語大学）

1960 年代初頭、世界で初めての電子コーパスである Brown Corpus が編纂されて以降、多種多様なコーパスが利用可能になっている。そしてコーパス（とコンピュータ）のおかげで、今まで見えなかったことが見えるようになり (Kettermann 1995)、新しい言語の見方が可能になっている (Moon 2007)。また言語研究のみならず、さまざまな分野でコーパスが利用されている。とりわけ外国語教育でコーパスを用いた新たな試みや教材開発が行われており、辞書編纂もコーパスを活用することで、飛躍的に質が向上している。

では具体的にコーパスはどのように利用されているのか、また利用すべきなのか。コーパスの使い方に決まった方法や手順があるわけではなく、各研究者が研究テーマに応じた使い方をしている。このような現状を踏まえ、本シンポジウムでは、コーパスの利用経験が豊富な方々に以下の 4 項目を中心に、それぞれのコーパス利用について報告していただく。それを受けて参加者からもコーパス体験を語っていただきたい。

- (1) なぜコーパスを使うのか
- (2) どのようなコーパスをどのように使っているのか
- (3) コーパスの威力と限界
- (4) コーパスを使う際に注意していること

講師にお願いしたのは以下の方々である。

滝沢直宏（名古屋大学）、深谷輝彦（椋山女学園大学）、岡田毅（東北大学）、家入葉子（京都大学）、野口ジュディー（武庫川女子大学）、投野由紀夫（東京外国語大学）、井上永幸（広島大学）、田畑智司（大阪大学）、堀正広（熊本学園大学）（発表順）

さまざまな利用方法を共有することで、会員の中から、新たな研究成果が生み出されることを願っている。

## 《大会参加者へのご案内》

- 大会・ワークショップの受付：会場の全学教育推進機構 B 棟 1 階 108 教室前の前で、1 日目は午前 9 時 30 分から、2 日目は午前 9 時から受付を行います。
- 昼食について：土曜日は次の店舗が営業しています：大阪大学博物館（阪大坂下の駐輪場横）ミュージアムカフェ「坂」（10:30-17:30）、学生交流棟 1F 食堂「宙（Sora）」（8:30-20:00）、同 1F コミュニティストア（8:00-23:00）、総合図書館下食堂（11:00-15:00）、カフェテリア「ら ふおれ」（11:00-13:30）。キャンパス内の店舗はいずれも**日曜日は営業していません**。構外では、中央環状線を挟んで正門の真向かいにファミリーレストラン「Friendly」があります。また、キャンパス東側を走る府道沿い（国際交流会館の北東側）にはファミリーマートがありますが、**日曜日は昼食を持参されることをお勧めいたします**。なお、阪急石橋駅付近にはコンビニエンスストア数軒がありますが、モノレール柴原駅近辺には店がありませんので、あらかじめご承知下さい。
- 構内での喫煙について：構内では、「指定喫煙場所」（言語文化研究科 A 棟玄関脇、同 B 棟北出入口脇、および全学教育推進機構大講義室北側、いずれも灰皿設置）においてのみ喫煙が許可されています。「指定喫煙場所」での喫煙にご協力いただきますようお願いいたします。
- 当日会員について：会員でない方も、「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）。

### ◆会場案内図◆



公共交通機関アクセスマップ《会場は豊中キャンパスです》



○ 大阪伊丹空港からのアクセス：

大阪モノレール 大阪伊丹空港駅から2駅目(6分)、柴原駅下車 徒歩約15分

○ JR 新大阪駅からのアクセス

(1) 新大阪より地下鉄御堂筋線（江坂駅以北は北大阪急行）で千里中央まで13分、千里中央にて大阪モノレール大阪空港行きに乗り換え、2駅目（5分）、柴原駅下車 徒歩約15分（※初めてお越しの方にはこのルートの方が下記（2）よりも簡単です。）

または、

(2) JR 新大阪からJR線で大阪駅にて下車、阪急梅田駅にて阪急宝塚線に乗り換え、急行で約15分、石橋駅下車、東へ徒歩 全学教育推進機構まで約15分

2012年9月1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://english.chs.nihon-u.ac.jp/jaecs/>