

英語コーパス学会 Newsletter No. 54

Aug. 25, 2006

■会長: 中村 純作
■事務局: 〒615-8558 京都市右京区西院笠目町 6 京都外国語大学 赤野一郎研究室
■TEL: 075-322-6103 ■郵便振替口座: 00940-5-250586 (英語コーパス学会)
■URL: <http://muse.doshisha.ac.jp/JAECS/index.html> ■E-mail: i_akano@kufs.ac.jp

JAECS
Japan Association for English Corpus Studies

第 28 回大会のご案内

英語コーパス学会第 28 回大会は、10 月 7 日 (土) 北海道大学で開催されます[〒060-0817 札幌市北区北 17 条西 8 丁目 札幌駅から地下鉄南北線「麻生方面行き」に乗車、2 つめの「北 18 条」駅下車、徒歩西へ約 500m <http://www.hokudai.ac.jp/index.html>]。会場校の園田勝英先生、高見敏子先生のご尽力に感謝いたします。

詳しくは、同封の「大会資料」をご覧くださいのですが、今大会では恒例の午前中のワークショップのほか、研究発表 4 件と特別講演を準備いたしました。

研究発表につきましては、運営委員会の査読を経て、7 月 16 日 (日) に中央大学で開かれた大会準備委員会での最終審査の結果、北本徹平氏 (大阪大学大学院生) の「live a happy life と live happily の交替可能性」、村形舞氏 (東京大学大学院生) の「法助動詞を伴う文における until 節の節順選択」、木村恵氏 (獨協大学)、田中省作氏 (立命館大学)、八島等氏 (東京都立城東高等学校)、依田みずき氏 (元東京学芸大学大学院生) の「中高教科書コーパス分析と習得困難度要因に基づいた自動語彙レベル判別の試み」、岡田毅氏 (東北大学) による「コーパス分析と中学校英語教科書 動詞活用形の観点から」の 4 件が選ばれました。

特別講演では、『今日から使える発話データベース CHILDES 入門』(ひつじ書房)の編者、宮田 Susanne 氏 (愛知淑徳大学) に「発話データベース CHILDES の概要とその成果」の演題でお話しいただきます。

恒例となっております午前中のワークショップでは、特別講演とリンクさせ、宮田 Susanne 氏を講師に「発話データベース CHILDES 入門」と題して、CHILDES の解析プログラム

CLAN の実習を行っていただきます。参加御希望の方は、電子メール(件名「ワークショップ申込」)で、所属と会員・非会員の別を明記の上、事務局宛にお申し込みください。英語コーパス学会の会員であれば参加費は無料です。非会員の場合は当日会費 1,000 円で、ワークショップと大会の両方に参加していただけます。

John Sinclair 先生の講演会報告

Sinclair 先生は 8 月 18 日 (金) から西南学院大学を会場に開催された AsiaTEFL での Plenary Session と Demonstration のため、Thomson ELT の招きにより来日されました。この機会を利用して、*Newsletter* や JAECS-ML を通じてお知らせしておりました講演会を福岡だけでなく関西地区と東京で開催しました。

関西地区の講演会は 8 月 19 日 (土) 午後 3 時より立命館大学末川記念会館 1 階ホールにおいて行われ 82 名の参加者を得ました。講演は「Exploring a Corpus」と題した主に新しい 2 つのソフトウェアの紹介でしたが、それに先立って、それら 2 つのソフトウェアの必要性に関する理論的な背景について詳しい説明が行われました。まず、従来の Grammar と Lexis を峻別する枠組みそのものに疑問を投げかけるとともに Phraseology の重要性を意識する枠組みへの転換が起きていること、従来は非常に少ないデータから如何に結論を導き出すかが問題であったが、コーパスが巨大化するにつれて大量のデータを如何に効率的に利用するかが問題になっていること、数量的分析が必要になっていることが指摘されました。さらに従来の slot-and-filler 型の分析から successivity を問題にする必要性、paradigmatic model から syntagmatic model への転換の必要性などが指摘され、これ

らの必要性に応えるソフトウェアとして Concgram@ と PhraseBox™ が紹介されました。

Concgram@は n-gram と Concordancer の機能を合わせ持っており、2 語以上を対象にその組み合わせを、前後関係の違いだけでなく、途中で何語か挿入された場合も含めて全て抽出するプログラムで、“work” と “hard” の検索結果が示されました。PhraseBox™ は Sinclair 先生がスコットランド政府の語学教育のために実用化しようとしている非常に簡便なソフトウェアです。2 語以上のキーワードにも対応した CQL Query が可能なコンコーダンサーで、同時に vocabulary words と grammatical words によるコロケーションを見ることができるものでした。BNC の 70%、Bank of English の 50% に Scottish English を加えた 3 億語のコーパスを利用した “gamut” のコンコーダンスとコロケーション、“a,” “of,” “grass” から “a bit of grass,” “a strip of grass,” “a blade of grass” 等が抽出される様子や、“she” の重要な共起語が “when” であり、さらに “when she was” に対しては “approached” が典型で、何か良くないことが起こっている文脈で使用されることなどが次々と紹介されました。1 時間 30 分の予定の講演が 15 分ほど長くなり、その後の 30 分にわたる質疑、応答も熱がこもったものでした。

その後、会場を移動、5 時 30 分から 35 名が参加しレセプションに移りました。主催者の挨拶の後、JAECS から第 10 号の名誉会員証を差し上げ、Sinclair 先生のお言葉、トムソンの代表者からの乾杯と続きました。Sinclair 先生との懇談、彼の著書にサインを頂いたり、記念写真を撮ったり、参加者同士での情報交換などで、瞬く間に 2 時間のレセプションが終了しました。

翌日は 10 時過ぎの新幹線で東京に移動。途中、原宿に立ち寄りお子さんたちへのお土産の購入と昼食。2 時過ぎに会場の日大文理学部 100 周年記念館に到着。3 時からの講演会では、基本的には前日の中身を踏襲しながら、最初にコーパス言語学の歴史を少しばかり紹介されました。こちらでも 15 分超過し、たっぷり 30 分の質疑にお答え頂きました。講演会参加者は 67 名でした。5 時 30 分からは JAECS では、初めての試みとしてケイタリングサービスを利用

したレセプション。参加者は 22 名とこじんまりとしたパーティでしたが、とても良い雰囲気です。終えることができました。

Sinclair 先生は翌日、成田発で帰国。スコットランドでご家族との休暇中にもかかわらず、73 歳のご高齢をおして、遠路はるばる来日、5 日間のみ日本滞在中、2 日間は JAECS に割っていただきました。そのお陰で我々の長年の懸案であった Sinclair 先生の講演会とそれを契機に名誉会員になっていただくことが、実現しました。このことを可能にいただいたトムソン ELT、特に、6 ヶ月にわたり準備いただいた小嶋里佳さんにはお世話になりました。日大では塚本聡先生、秋山孝信先生に会場、パーティに関して特にお世話になりました。そのほか、関係大学の事務の方々、院生諸君にもお世話になりました。この紙上を借りて篤くお礼を申し上げます。最後になりましたが、酷暑のおり、遠方より駆けつけていただいた会員諸氏に心より感謝いたします。

中村 純作 (立命館大学)

学会賞応募規定

第 6 回の学会賞を募集致します。応募規定は次の通りです。

【対象】英語コーパス学会の目的に照らし、英語のコーパス言語学に関する優れた研究業績をあげた学会員(個人またはグループ)とする。ただし、奨励賞は英語のコーパス言語学に関する優れた研究業績(論文の場合、学会誌『英語コーパス研究』に掲載されたものに限る)をあげた 35 歳以下または大学院修了後の研究歴 5 年以下の学会員個人に限る。

【応募方法】自薦、他薦を問わない。

【提出書類】1) 所定の推薦理由書(学会ホームページより入手)。
2) 論文の場合は現物またはコピー。単行本の場合は事務局で用意するので送付は不要。

【提出先】事務局

【応募期限】2007 年 3 月 31 日

【発表】2007 年度秋季大会

『英語コーパス研究』第14号について

前号のニューズレターにおいて『英語コーパス研究』第14号(2007)の原稿を募集しましたところ、論文4件、研究ノート1件の申し込みを頂きました。『英語コーパス研究』では、論文のほか、研究ノート、実践報告、書評、海外レポートなども募集しています。まだ執筆申し込みをされていない会員の方も、原稿締め切りの9月末日までに原稿を送付いただければ、審査の対象となりますので、ぜひとも奮ってご投稿ください。

なお、投稿規定、スタイルシートについては、今回より一部修正されています。詳細につきましては、http://muse.doshisha.ac.jp/JAECS/Guidelines/ECS_SGuide-j.html をご覧ください。

『英語コーパス研究』編集委員会委員長
塚本 聡(日本大学)

JAECS 東支部活動報告

JAECS 東支部では、第4回研究談話会を以下のとおり開催いたします。多数の皆様のご参加をお待ち申し上げます。

日時：2006年9月17日(日)

13時30分～17時30分

場所：大東文化大学大東文化会館 K-0302

<http://www.daito.ac.jp/exten/access.html#01>

東武東上線東武練馬駅前 徒歩1分

池袋駅から東武練馬駅まで、約15分

参加費：会員無料・非会員300円

(事前申し込み不要)

発表者・題目：

小林雄一郎(法政大学大学院)

「タグ頻度解析による学習者レベルの識別
The NICT JLE Corpus を例に」

大羽 良(早稲田大学非常勤講師)

「Web形式による日英対訳コーパスの検索
システムの開発とその活用例」

JAECS 東支部支部長

新井 洋一(中央大学)

新入会員紹介(8月25日現在、Sは学生)

黒崎 紫乃 ロンドン大学 Institute in Paris 博士
課程 S

秀野 作次郎

長井 みゆき 名古屋大学大学院 S

村形 舞 東京大学大学院総合文化研究科言語
情報科学専攻博士課程 S

森 茂 大分大学医学部

森本 由子 筑波大学大学院 S

所属の変更

井上 亜依 長崎外国語大学

岩崎 克己 広島大学外国語教育研究センター

大森 誠 東広島市立向陽中学校

木村 恵 獨協大学

小室 誠一 バベル翻訳大学院

田中 美和子 京都ノートルダム女子大学

谷村 緑 京都外国語大学

廣瀬 絵美 フリーランス

堀池 保昭 武庫川女子大学付属高等学校

森田 光宏 山形大学人文学部

FORUM

新刊紹介

中村純作(立命館大学)

jnakamur@li.ritsusmei.ac.jp

お送り頂いた新刊で、*New Letterer* 編集上の都合により53号でご紹介できなかった2点を今回は紹介させていただきます。

まず、荒木一雄、天野政千代両先生監修『英語学入門講座』の第10巻として、本学会の会員、大門正幸、柳朋宏先生(中部大学)による『英語コーパスの初歩: An Introduction to English Corpora』が英潮社より本年3月に刊行されました。著者による「まえがき」にもあるように、コーパスという言葉が定着してきた現在、研究者や英語教員だけでなく一般の学習者(主に言語・文化あるいは英語・英文学を専攻する学部

生か?)にも分かりやすい入門書の必要性から本書は書かれたものである。その特色として、コンピュータに詳しくない学習者にも分かりやすいように画像を多用、できるだけ金銭的負担を必要としない方法を優先し、Mac OS と Windows に対応して書かれていることの3点が挙げられる。

本書は11章から構成され、コーパスの定義から始まり、コーパスの種類と変遷、ブラウザ上での検索、インターネットの検索エンジンの使用、エディターによるテキストファイル検索、MS-Wordによるワイルドカードを使った少し高度な検索、タグ付きコーパスの検索、コンコーダンスの作成、WordbanksOnlineの利用法とそのファイルの転送、正規表現を使った検索などを取り扱っている。具体的に取り扱われるトピックも“anathematize”の語義とその用例、“fill in/out”の使用頻度の英語変種による差異、A. Bierceの*The Devil's Dictionary*のダウンロードと検索、corpusの複数形、likeの品詞による分布、noneの数による一致など興味深いものがある。ただ、画像が多いこととMacとWindowsに対応させるためにスペースがとられることもあり、面白い言語現象を提示しつつ学生をコーパスに引き込んでいくための例がもっと欲しいところである。延々と続くSusanne Corpusの品詞標識リストなどは実際の検索例から品詞標識の有効性を示すことに使った方が良さそうである。Mac上でのコマンドモードでのコンコーダンスの作成はちょっと初心者には取っ付きにくいと思われるし、Windows版に対応したCygwinの使用法については説明が見られない。正規表現の説明にはWordbanksOnlineが使われているが、実際にアクセスできる読者は限られている。この辺りは、第2版以降改訂されることを切に希望したい。とまれ、コーパスを講義するものにとって、大いに参考になる一書であることに間違いはない。是非、ご一読下さい。(B5版306頁/定価3,400円(外税)/ISBN 4-268-00411-4)

次にご紹介するのはMichael Stubbs (2002) *Words and Phrases: Corpus Studies of Lexical Semantics* (Oxford: Blackwell)の翻訳で、本年5月研究社より出版された南出康世、石川慎一郎監訳『コーパス語彙意味論：語から句へ』である。著者のStubbs教授は現在ドイツのトリ-

ア大学の英語学教授。限られた時間とスペースの中で、著者の長年にわたる意味論研究とそれに基づいた講義の集大成である本書の内容を十分理解し、論評するだけの能力を筆者は持ち合わせていないが、従来のGrammarとLexisを厳密にレベル分けし、意味は個々のLexisにあるとする伝統的枠組みから出発し、Phraseologyの重要性を徹底的に追求したものが本書だと言える。多義語がより長い文脈の中で完全に非曖昧化し、予測される結合形の中で使用されていること、さらにこのことにより、評価の意味が与えられたり、テキストの構成にも影響が及ぶこと、文化的にも重要なキーワードとなったりするかが、句、テキスト、文化のレベルで検証されている。それも全てコーパスを用いた事例研究によることは本書のタイトルが示すとおりである。従来難解だと思われていた意味論の世界が、豊富な具体例で説明され、展開されていく様子にコーパスの威力を実感することができる。但し、言語学、英語学の概論を終えた院生、研究者が対象。

このようにコーパスを駆使する手法は、英国流の典型的な経験主義の流れを汲むもので、Malinowskiに始まり、Firthにより確立されるロンドン学派を現在に引き継ぐものと考えられる。StubbsとともにCollocationとPhraseologyを重視するJohn Sinclair、*Pattern Grammar*のSusan HanstonやGill Francis、*Corpus Linguistics at Work*でCorpus-drivenな手法を提案しているElena Tognini-Bonelliなどの一連の流れとその方向性を理解するのにも本書は重要である。

訳者の「あとがき」によると本書は発足して4年になる関西英語辞書学研究会(Kansai English Lexicography Circle, KELC)の2番目の輪読用テキストで、発表者を決め、ハンドアウトやパワーポイントを使う研究発表の形式で、約1年をかけて読んだ成果として生まれたものだという。難解な大著に挑戦し、出版にこぎつけた監訳者の南出先生(大阪女子大学名誉教授)、石川先生(神戸大学)と各章を担当された10名の先生方(スペースの都合でお名前は割愛)に敬意を表するとともに、出版を引き受けた『英語青年』編集長の津田正氏にも敬意を表したい。(B5版376頁/定価3,800円(外税)/ISBN 4-327-40143-9)

岡田毅(東北大学)
t-okada@lark.intcul.tohoku.ac.jp

6月23日、24日の両日にわたり、University of Nottingham の CRAL (Centre for Research in Applied Linguistics) がホスト校となって The 3rd IVACS (Inter-Varietal Applied Corpus Studies) International Conference が開催された。本学会からは兵庫県立大学の瀬良晴子先生と私が参加し研究発表を行った。本稿は、keynote speaker の一人として、現在はイタリアの The Tuscan Word Centre に席を置かれている John Sinclair 先生を迎えての大会の報告である。

約 90 の研究発表がパラレルセッション形式で行われ、それぞれのテーマは、EFL 教育へのコーパスの応用、phraseology の諸相、ESP におけるコーパスの役割、パラレルコーパスの可能性などであった。

開催校である Nottingham 大学の研究者の発表には CANCODE (Cambridge and Nottingham Corpus of Discourse in English) を駆使したものが多く、例えば Irina Dahlmann 氏の、pause を手がかりとしての spoken English における multi-word unit の抽出に関する発表などは、母語話者の言語直観の重要さと、コーパス分析から得られる数値的な結果との関連性を扱ったもので興味深かった。

コモンルームでは、Birmingham に拠点を置く Speechinaction 社の Richard Cauldwell 氏が大会の初日から Streaming Speech online courses のデモを行い、参加者から少なからぬ関心を集めていた。Speechinaction 社は 2004 年に British Council から Innovations in English Language Teaching に与えられる ELTON 賞を与えられている。中・上級学習者向けのこの発音トレーニング教材は、これまで British/Irish の spontaneous speech を素材とした双方向的システムであったが、最近では American/Canadian の素材もカバーするようになり、初級学習者向けのコースも準備中である。日本のメーカーが開発している発音矯正ソフトなどにも高性能なものが増えつつある(例えば株日本ビクター社の「ソフトウェアレコーダー」)が、問題になるのは、学習者がモ

デルとする音声素材である。この点、Streaming Speech で提供されるそれは、Nottingham 大学等とのコラボレーションの成果もあり、実際の場で用いられた発話であるために、より自然な発音や応答表現が習得しやすいのではないかと感じられた。使用料金等に関しては、以下の URL を参照されるかメールで Cauldwell 氏にお問い合わせ願いたい。

<http://www.speechinaction.com/>
richard@speechinaction.com

大会をしめくくる Keynote Speech は、“From Text to Tree: LUG, LUM and PUB” のタイトルで University of Helsinki の Anna Mauranen 先生と Sinclair 先生が担当された。LUG (Linear Unit Grammar) を用いて、linear なテキストから意味へと繋がるのに必要な最低限の hierarchical な情報を入力していくという考え方の中にあって、単純な chunk の連続体としてテキストを処理するのではなく、chunk を例えば message type や organizing type 等に分類し、これらを元に PUB (Provisional Unit Boundary) を動的に認識し、それを LUM (Linear Units of Meaning) としてテキストを処理していくという概念が具体的な発話データに言及しながら提示された。LUG は linear なテキストと hierarchical な文法との間の橋渡しの役割を担うものであって、従来の言語研究に比べて、一層 spoken テキストに重点を置いたこの研究を進めることによって spoken と written テキストの類似性が明らかになってくると主張された。

「皆さんが Power Point ばかり使われるから、私もようやくこれを使えるようになりました。どうですか？上手くってます？」と茶目っ気を交えて(その実、手馴れた操作をなさりながら)、Mauranen 先生との講演を終えられた Sinclair 先生であったが、実はこの後に、予定にないプレゼンテーションをして下さった。一応、学会の閉会を待って、しかも部屋を移動して、Sinclair 先生はアメリカ、カナダからご帰国直後のお疲れも全く見せずに、新しいコーパス・サービスを紹介して下さいました。確かに開発企業は Scotland にあるが、未公開・開発中であるために Sinclair 先生は「このサーバーは Scotland (西北端の) Skye 島の霧の中に隠れて

いる」と、目を輝かせて部屋を移動した聴衆を文字通り煙(霧?)に巻いておられた。

「9月頃には公開できるだろう」とおっしゃるシステム PhraseBox™のデモ版をオンラインで見せて頂いた。開発中であるために、全貌を詳しく教えて下さらなかったが、concordance や collocation 集計・分析機能が相当に強化されていること、また対象とするコーパスの種類と量が圧倒的であろうことが、検索設定ページから垣間見ることができた。

本 Newsletter が発行される時期と Sinclair 先生が本学会で講演なさり、おそらく PhraseBox™ の紹介もなさるであろう時期とが微妙な関係になるが、少なくともこのシステムの本格稼働は秋以降となるために、ここで紹介させて頂いた。デモ版については以下のサイトにアクセスされたい。

<http://www.phrasebox.com/phrasebox/>

10年以上前に Birmingham でお目にかかった頃と比べて、少しお体が「丸く」なられたような感じを受けたが、まだまだエネルギーで新しい分野に意欲的に取り組まれている Sinclair 先生のお姿には頭の下がる思いであった。

発表前日に大学のバブで Japan vs. Brazil の WC フットボール試合を England サポーターたちと観戦することができた。彼らは(下心があるとは言え)一様に Japan 鼻根であった。「なぜ Japan はホームでもアウェイでもブルーしか着ないんだ?」というパーテンダーの質問には、即興の思いつきで答えるしかなかった。IVACS 2008 はアイルランドで開催予定と聞かされた。

Workshop “Corpus Approaches to the Language of Literature”について

堀 正広(熊本学園大学)
hori@kumagaku.ac.jp

2006年7月25日(火)にフィンランド東部の閑静な大学町 Joensuu で行なわれた標記大会について報告する。これは、Poetics and Linguistics Association (通称、国際文体論学会)の第26回

年次大会の前日に、Oxford Archive の Martine Wynne によって企画された。この workshop の目的は、タイトルが示すように、コーパス言語学の成果や方法論の文学作品の言語文体研究への援用である。

まず、企画者の Martin Wynne が文体研究とコーパス言語学の共通性と相違点を述べ、文体研究にはデータ収集は不可欠であるという点から、コーパス利用の利点と今後の可能性が強調された。

次に、ランカスター大学の Jonathan Culpeper が、現在共同制作中の Shakespeare's Dictionary の構想を述べた。たとえば、horrid はどのような語や特定の意味のタイプの語や品詞やレジスターと共起するのか。また、どの劇に頻出し、どの人物がよく使うのか、などをコーパスを使って収集した辞書である。内容語だけでなく“and”などの機能語や“ah”のような感嘆詞も収録される予定である。発表の後、実践的なワークショップが行われた。完成までには数年はかかるとのことだが、コーパス言語学の知見があって初めてなしえる業績である。

次に、リバプール大学の Michaela Mahlberg が Charles Dickens における word cluster について、Dickens は同時代の19世紀の作家達に比べて3語から5語の cluster がいかに多いかを豊富な例を挙げて指摘した。たとえば、as if he had been, in the course of the, a quarter of an hourなどは最も頻度の高い5語の cluster である。発表後、wordsmith を使って、このような cluster の見つけ方が実践的に教授された。

Mahlberg は昨年 John Benjamin から *English General Nouns: A corpus theoretical approach* を刊行している。拙著 *Investigating Dickens' Style: a collocational analysis* に刺激を受け、Dickens を次の研究対象に選んだとのことだった。私の本を3度読んだと言われたときは大変恐縮した。

詳細は <http://www.pala.ac.uk/signs/corpus-style/joensuu.htm> を参照していただきたい。ここには今回発表に使われた power point と練習問題が公開されている。日本人の参加者はみなコーパス学会の会員で、瀬良晴子、田畑智司、奥総一郎各氏と私の4人であった。