英語コーパス学会
第 **43** 回大会資料

日時：2017 年 9 月 30 日（土）–10 月 1 日（日）
会場：関西学院大学（西宮上ケ原キャンパス）
〒662-8501　兵庫県西宮市上ケ原一番町 1-155

英語コーパス学会 2017

英語コーパス学会　第 **43** 回大会　プログラム

■第 1 日目

ワークショップ 1【TCSE を用いた TED Talks の全文検索と英語教育への応用】
会　場：関西学院大学（西宮上ケ原キャンパス）第 5 別館 3 階　第 307 教室
日　時：9 月 30 日（土）10:00–11:30（第 5 別館 1 階正面にて 9:30 受付開始）
講　師：長谷部　陽一郎（同志社大学）
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、2 日間共通）。

日　時　2017 年 9 月 30 日（土）
受付開始　11:30（第 5 別館 1 階正面）
開会式　12:30（第 5 別館 1 階　第 2 教室）

司　会　石井　康毅（成城大学）
投野　由紀夫（東京外国語大学）
小菅　正伸（関西学院大学　副学長）
西村　秀夫（三重大学）

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 セッション（場所：第 5 別館 1 階　第 1 教室）　司会：佐竹　由帆（駿河台大学）

研究発表 1：13:30–14:00
Charles Dickens の *The Mystery of Edwin Drood* と Thomas Power James によるその続編の文体類似性評価　後藤　克己（中部大学大学院生）

研究発表 2：14:05–14:35
TF-IDF を用いた Alice Bradley Sheldon の計量文体分析　木村　美紀（明治大学大学院生）

研究発表 3：14:40–15:10
機械学習アプローチによる小説テクストの計量的分析：
―アーサー・コナン・ドイルの作品から―　黒田　絢香（大阪大学大学院生）

研究発表 4：15:15–15:45
Agatha Christie 作品の修辞的特徴に関する分析　土村　成美（大阪大学大学院生）

●研究発表第 2 セッション（場所：第 5 別館 1 階　第 2 教室）司会：藤原　康弘（名城大学）

研究発表 1：13:30–14:00
Investigating the Impact of Extensive Reading with Data-Driven Learning　Gregory HADLEY（新潟大学）
ハドリー　浩美（新潟大学）

研究発表 2：14:05–14:35
日本人英語学習者の関係代名詞の回避
―CEFR レベルを用いた検証―　高橋　有加（東京外国語大学大学院生）

研究発表 3：14:40–15:10
英語教材としてのアニメ分析　寺島　芙由（東京外国語大学大学院生）

研究発表 4：15:15–15:45
小中連携に向けた英語授業コーパスデータ構築と
インタラクション分析の試み　片桐　徳昭（北海道教育大学）
大橋　由紀子（ヤマザキ学園大学）

●研究発表第 3 セッション（場所：第 5 別館 3 階　第 307 教室）　司会：後藤　一章（摂南大学）

研究発表 1：13:30–14:00
複数コーパスの統一的処理を可能にした
高速コーパスデータベースシステム MyCo の開発
西村　祐一（元名古屋大学大学院生）

研究発表 2：14:05–14:35
構文構造を活用した学術論文における頻出コリゲーションの抽出
田中　省作（立命館大学）
徳見　道夫（九州大学名誉教授）
宮崎　佳典（静岡大学）
金丸　敏幸（京都大学）
田地野　彰（京都大学）

研究発表 3：14:40–15:10
Word2Vec による文学作品の時代比較
―コーパスを軸とした異分野融合研究の試み―
内田　諭（九州大学）
下條　恵子（九州大学）
渡邉　智明（九州大学）
斎藤　新悟（九州大学）
谷口　説男（九州大学）

研究発表 4：15:15–15:45
構文情報などを表す木構造の配列による情報処理
田中　省作（立命館大学）
宮崎　佳典（静岡大学）
田辺　利文（福岡大学）
田村　昌彦（立命館大学）

＜休憩　15:45–16:05＞

●研究発表第 4 セッション（場所：第 5 別館 1 階　第 1 教室）　司会：石川　有香（名古屋工業大学）
研究発表 1：16:05–16:35
強調語の調査による Popular Music の歌詞の文体研究　渡部　文乃（京都大学大学院生）

研究発表 2：16:40–17:10
ホテルのオフィシャルウェブサイトにおける概説文のストラテジー　近藤　雪絵（立命館大学）
―Move の構築と分析を中心に―

研究発表 3：17:15–17:45
一般教書演説から見る米国大統領の関心事の変遷　木山　直毅（北九州市立大学）
―トピックモデルと時代背景―

●研究発表第 5 セッション（場所：第 5 別館 1 階　第 2 教室）　司会：大谷　直輝（東京外国語大学）
研究発表 1：16:05–16:35
【賛助会員発表】コーパスの示す科学的データと学習性・商品性との　井上　永幸（広島大学）
両立―『ウィズダム英和辞典』の編集にあたって―　西垣　浩二（㈱三省堂辞書出版部）

研究発表 2：16:40–17:10
英語辞書レーベルとコーパス　田畑　圭介（神戸親和女子大学）

研究発表 3：17:15–17:45
怒りを表す類義語と概念メタファー　南澤　佑樹（大阪大学大学院生）

●研究発表第 6 セッション（場所：第 5 別館 3 階　第 307 教室）　司会：森下　裕三（環太平洋大学）
研究発表 1：16:05–16:35
日英対訳コーパス中の「～ことになる」構文と　大矢　政徳（目白大学）
その英訳文間の構造的不一致

研究発表 2：16:40–17:10
医学研究論文ジャンルにおけるコーパス作成ツール AntCorGen を活用　浅野　元子（大阪大学大学院生）
した教育の可能性―Construction of Corpora for Discipline-Specific
Learning in Medical Research Article Genres

研究発表 3：17:15–17:45
Applying Topic Models to Describe a Corpus’s Compositionality: How can the　Tomoji Tabata (University of Osaka)
external criteria be associated with meaningful sets of internal evidence?

＜懇親会：18:15–20:30（会場：関学会館、会費：5,000 円）＞

■第 2 日目

日　時　　2017 年 10 月 1 日（日）
受付開始　　9:30（第 5 別館 1 階正面）

ワークショップ 2【機械学習を用いたコーパス分析入門】
会　場：関西学院大学（西宮上ケ原キャンパス）第 5 別館 3 階　第 307 教室
日　時：10 月 1 日（日）10:00–11:30
講　師：小林　雄一郎（日本大学）
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、2 日間共通）。

＜休憩　11:30–12:30＞

●講演　　12:30–13:30（第 5 別館 1 階　第 2 教室）

A Frontier in Learner Corpus Studies: For Better Understanding of L2 Learners

司会：投野　由紀夫（東京外国語大学）
講師：Shin’ichiro Ishikawa (Kobe University)

＜休憩　13:30–13:50＞

●シンポジウム　　13:50–15:20（第 5 別館 1 階　第 2 教室）

話し言葉コーパスの構築と利用　　司会：野口 ジュディー（神戸学院大学名誉教授）

The ICNALE：中間言語対照分析の精緻化とアジアにおける学習者コーパス研究の発展を目指して　　講師：石川　慎一郎（神戸大学）

International corpus of Japanese as a second language：日本語学習者の言語研究と指導のために　　講師：迫田 久美子（広島大学・国立国語研究所）

JECPRESE：JSL と EFL ユーザーのために　　講師：野口 ジュディー（神戸学院大学名誉教授）

TED Corpus Search Engine：TED Talks を研究と教育に活用するためのプラットフォーム　　講師：長谷部　陽一郎（同志社大学）

閉会式　　15:30（第 5 別館 1 階　第 2 教室）

閉会の辞　　井上　永幸（広島大学）

■9 月 30 日（土）
【ワークショップ 1】

TCSE を用いた TED Talks の全文検索と英語教育への応用

長谷部　陽一郎（同志社大学）

TED Corpus Search Engine（TCSE）は TED が公開している約 2400 件の英語プレゼンテーションのトランスクリプトを解析してデータベースに格納し、英語テキストと翻訳テキストの全文検索を可能にした Web システムである。TCSE には英語教育や言語学研究のために TED Talks を役立てるための各種機能が実装されており、本ワークショップでは特に英語教育への応用を念頭においた解説を行う。予定している内容は下記のとおりである。

＜基本編＞
（1）　TED Talks と TCSE の概要
（2）　英語と日本語を用いた基本的な事例検索
（3）　教育素材となる Talk を探すために役立つ機能
＜応用編＞
（4）　高度な検索式を使った事例検索と結果の保存
（5）　「Pause and Check」機能を活用したリスニング／スピーキング学習
（6）　実践例：TCSE を用いた英語談話標識の指導
＜発展編＞
（7）　TCSE の内部構造と理論的背景
（8）　TCSE の実験的機能

まず「基本編」では TCSE で何ができるかについて大まかな理解を得ることを目指す。次に「応用編」では主に 2 つのことを行う。1 つは品詞や基本形などの語彙情報を用いてイディオムや構文の事例を効果的に検索する方法を知ること、もう 1 つは TCSE の「Pause and Check」機能をリスニングやスピーキングの学習や指導に利用する方法を知ることである。最後に「発展編」では TCSE の構造や背景について簡単に解説するとともに、いくつかの実験的な機能について言及する。

本ワークショップは基本的に講義形式で進めていくが、インターネット接続された機器を持参して実際に試していただくのも良いだろう。TCSE は PC（Windows, MacOS）の Web ブラウザ上での使用を基本としているが、多くの機能はスマートフォンやタブレットでも利用可能である。

■9 月 30 日（土）
【研究発表第 1 セッション】
【研究発表 1】

Charles Dickens の *The Mystery of Edwin Drood* と Thomas Power James によるその続編の文体類似性評価

後藤　克己（中部大学大学院生）

米国人 Thomas Power James (以下、T. P. James)は、Charles Dickens の遺作となった *The Mystery of Edwin Drood*（以下、原典）に続編を加えて完全版とし 1873 年に発表した。この続編は T. P. James がこの続編を"By the Spirit Pen of Charles Dickens, through a Medium."とアピールしたこともあって大きな論議を呼び、物語性、人物造型の一貫性、言語的側面等の観点から多くの批評がなされた。当時 W. H .B、George F. Gadd、John Cuming Walters 等がそれらの観点から否定的に批評したが、言語的側面への言及は、抽象的・感覚的また部分的なものにとどまっている。そこで原典と続編に、発表時期が 1864-5 年と原典（1870 年）に近い *Our Mutual Friend*（以下、*OMF*）を加えたコーパスを用いて、語彙使用の観点から文体類似性の数量的評価を試みた。

まず、原典、続編および *OMF* に生起する語彙頻度に着目した。作中の発話部は登場人物によって大きく異なるため除外し、作者の本来の文体を最もよく反映していると考えられる地の文のみを用いた。各作品を章単位でサブコーパス化した原典 20[1]、続編 23 および *OMF*67 のサブコーパスについて、AntConc を用いて使用語彙のレマでの生起頻度を抽出／構成して語彙頻度表を作成し対応分析を行った。（結果：図 1）

つぎに、語彙クラスター(n-gram)には作者の好みが反映されると考えられることから、Mahlberg (2013)で Dickens 作品に特徴的とされている 5 語クラスター、ならびに類義語 as if / as though を選び、原典、続編および *OMF* での生起頻度を比較した。なお、ここではいずれの作品とも発話部を含むフルテキストを使用した。（結果：図 2）

[1] Dickens の原典は 23 章で構成されているが、T. P. James による完全版の原典部分は、一部の章が合体して 20 章構成となっている。

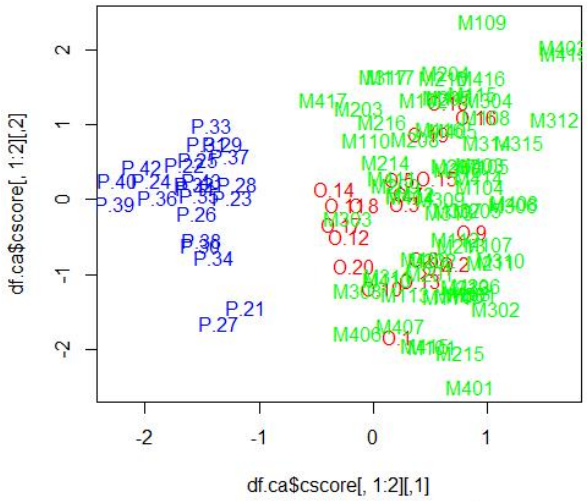


図 1　対応分析による 1—2 軸の列スコア [2]

5 語クラスター グループ	頻度計（10万語換算）		
	原典	続編	*OMF*
Body Part	6.3	0.8	8.9
As If	8.5	0.8	6.4
計	14.8	1.6	15.3

2 語クラスター	頻度計（10万語換算）		
	原典	続編	*OMF*
as if	112	15	120
as though	29	92	11
計	141	107	131

図 2　クラスターの頻度比較

その結果、原典（O.xx）と *OMF*（Mxxx）のサブコーパスは、まとまった一つのクラスターを形成するものの、続編（P.xx）のそれは明らかに別クラスターを形成しており、また Body Part と As If グループの 5 語クラスターの続編での生起頻度は原典／*OMF* より圧倒的に少なく、また、as if / as though の生起傾向は、原典／*OMF* と続編とで逆転していることが判明した。これらの結果は語彙使用面で、続編と Dickens 作品との文体類似性が低いことを示しており、続編を"By the Spirit Pen of Charles Dickens, through a Medium."とする T. P. James のアピールへの懐疑的見方を支持するものでもある。

【研究発表 2】

TF-IDF を用いた Alice Bradley Sheldon の計量文体分析

木村　美紀（明治大学大学院生）

本研究では、コーパス言語学の一分野である計量文体論の手法を用いて Alice Bradley Sheldon 作品群に関する文体調査を行う。Alice Bradley Sheldon (1915-1987: 米国) はデビューから約 10 年間正体不明・性別不明の作家 James Tiptree, Jr. として著作活動を行ってきた作家である。この作家の文体は文芸批評において、Ernest Hemingway と比較されることが多い。特に Silverberg (1975) では “Hemingway was a deeper and trickier writer than he pretended to be; so too with Tiptree, who conceals behind an aw-shucks artlessness an astonishing skill for shaping scenes and misdirecting readers into unexpected abysses of experience. And there is, too, that prevailing masculinity about both of them.” というようにいささか主観的・印象論的に Alice Sheldon の男性名義である James Tiptree, Jr. 作品群と Ernest Hemingway 作品の文体比較が行われている。

計量文体論の手法を用いた Alice Bradley Sheldon 作品群の文体比較では、Alice Bradley Sheldon 全 72 作品を収録したコーパス (延べ 865,802 語) と Ernest Hemingway69 作品を収録したコーパス (延べ 271,475 語) を用いて、Burrows (1987) に基づき高頻度語彙上位 50 語を指標としてランダムフォレストを行った。その結果、分類正確率は 92.20%であり基準値 51.06%を有意に上回っているということが判明した。ここから、文芸批評で主張されているようなこの 2 著者における文体の類似性はとらえることができなかったといえる。

しかしながら、この文体比較は、サンプルサイズが小さいという問題点や、ジャンルの相違などの様々な交絡因子が存在するという問題点がある。このため、発表ではこれらの問題点を克服するために Alice Bradley Sheldon と同時代・同ジャンルで活躍した作家のコーパスを構築して文体比較を進める。具体的には、Alice Sheldon と親交のあった Ursula K. Le Guin の 45 テキストを含むコーパス (延べ 589,481 語) 、Octavia E. Butler の 93 テキストを含むコーパス (延べ 867,396 語)、Arthur C. Clarke の 104 テキストを含むコーパス (延べ 467,983 語) 、Theodore Sturgeon の 222 テキストを含むコーパス (延べ 1,777,561 語) を構築した。サンプルサイズの違いの影響を減じるため、Theodore Sturgeon 全 222 作品からランダムに 70 作品抽出したコーパス (延べ 475,704 語) を分析に使用した。

分析には、Baayen (2008) で分類感度の高さが指摘されているサポートベクターマシン(SVM) と、金・村上 (2007) において分類感度の高さが指摘されているランダムフォレストという統計手法を採用した。指標としては、定義や実行の容易さとその実績から単語頻度、単語の bigram、trigram、小町 (2016) や小林 (2016) などの情報処理の分野で使用されている TF-IDF の unigram、bigram、trigram という 6 種類を用いて分類正確率の変化を検証する。6 種類のコーパスでのランダムフォレストを用いたテキストの分類正確率はそれぞれ、86.31%、77.70%、66.45%、85.43%、77.26%、62.03%だった。また、SVM を用いたテキストの分類正確率はそれぞれ、99.56%、97.35%、90.29%、99.56%、97.57%、81.23%だった。本発表では、TF-IDF という情報処理の分野で広く使用されている重みづけの概念を言語研

[2] いずれかの作品に 70%以上偏在する語、人称代名詞および人名を除き、計 50Tokens 以上の上位 547 語による。

究の場に導入し、これらの指標を用いた分類正確率の比較提示をおこなう。また、分類正確率の提示だけではなく、作品の類似度の視覚化を行い、コーパスを用いた計量的研究が文芸批評で行われている研究へどのように貢献できるのかを提示する。

【研究発表 3】

機械学習アプローチによる小説テクストの計量的分析―アーサー・コナン・ドイルの作品から―

黒田　絢香（大阪大学大学院生）

Arthur Conan Doyle はシャーロック・ホームズシリーズの著者として広く知られる作家であるが、自身の本分と考え注力していた歴史小説はこれまであまり批評や研究の対象となっていない。また、既存の研究はいわゆる 'close reading' のアプローチが主体で、対立概念である 'distant reading' (Moretti, 2000)、つまり客観的な量的データに基づいて分析するアプローチは少ない。そこで本研究は、Doyle の推理小説と歴史小説を対象として、語彙頻度や生起パターンなど量的な観点から考察を行うことで、文学研究に新たな視点を提案することを目的とする。リサーチクエスチョンは以下の二つである。

1) Doyle の推理小説群と歴史小説群は統計的手法に基づき機械的に分類することは可能か、またどれほどの精度で行うことができるか。
2) それぞれのジャンルを特徴づける語や表現は何か。

分析対象は、Doyle による推理小説 7 作品と歴史小説 9 作品の計 16 作品である。分類には、アンサンブル学習のアルゴリズムである Random Forests (Breiman, 2001)を用いた。分類ののち、Tabata (2015)で提案されている Random Forests による分類に寄与した語に注目する特徴語抽出法を参考に、各ジャンルの特徴語をリスト化した。また、単語単体だけではなく、関連性のある単語をグループ化した『トピック』にもジャンル間の差異が現れるのではないかと考え、潜在的ディリクレ配分法(LDA) (Blei et al., 2003)に基づくトピックモデリングを行った。MALLET という java ベースのツールキットを用いて、トピックごとの単語の生起確率、テクストごとのトピックの生起確率を算出し、得られた結果をネットワークグラフに可視化した。

Random Forests による分類の精度は 96.58%で、十分に高い値で分類することが可能であった。また、推理小説側の中で部分的に歴史小説と誤分類されたファイルは、推理や謎解きではなく過去の出来事を描写する箇所で、この 'retrospective narrative'が歴史小説と似た語彙パターンを持っているのではないかと考察した。分類に寄与した特徴語には、歴史小説側は'cried'や'fight'、また'head', 'faces', 'arms', 'eyes'のような体の部位を表す語、一方で推理小説側は'case'や'found'など捜査に関する語や、'house', 'chair', 'room'など家や家具に関する語が抽出された。

トピックモデリングの結果も Random Forests の結果とある程度一致し、推理小説側に頻出する『犯罪捜査トピック』や『家・家具トピック』などを発見した。設定するトピック数を変化させて実験を行い、最も効果的にジャンル間の差異を発見できるトピック数設定について考察した。

二つの機械学習に基づく手法から得られた結果を総合し、量的な観点からそれぞれのジャンルを特徴づける語やトピックを発見した。

【研究発表 4】

Agatha Christie 作品の修辞的特徴に関する分析

土村　成美（大阪大学大学院生）

本研究では、Agatha Christie 作品の文体的特徴について、彼女と同時代に活躍したミステリー作家 Dorothy Sayers との比較を通して分析を行う。文学作品の文体に関する計量的な研究では主に語の出現頻度が分析の指標として用いられており、Christie 作品に関する計量的な分析も語彙を指標とした分析が多い(Lancashire & Hirst, 2009; 稲木, 2009; 稲木, 2013)。テキストに修辞的アノテーションを施すことは処理が複雑になるため、修辞的項目を指標とした分析は語彙や POS タグ・統語構造を指標とした分析ほどには多く行われていない。そのため、本研究ではテキストに修辞的アノテーションを行い、修辞的項目を指標として計量的分析を行う。修辞的項目を変数とした 2 作家の作品の分類実験を行い、Christie に特徴的な修辞的表現の抽出・検討を行う。

使用データは、Christie の 221 作品(5,071,282 語)、Sayers の 55 作品(1,375,645 語)を用いる。両作家で作品数・総語数共に大きく異なるため、それぞれの作家から 50 作品を無作為抽出し、分析に用いた。

修辞的アノテーションには DocuScope を使用した。DocuScope は Kaufer & Butler (1996)におけるレトリック理論と、Kaufer & Butler (2000)における言語表象理論を基礎として構築されたテキスト分析・視覚化のためのツールであり、言語表現を 101 の Language Action Types (LATs)に分類し、タグ付けを行う。

DocuScope を用いて得られた作品ごとの LATs の頻度情報を指標として、機械学習の一種 Random Forests (Breiman, 2001)を使用した分類を行った。分類の正解率は 92%~94%であった。Random Forests の分類において寄与率の高い項目について検討を行う。

寄与率の高い項目の中で、Sayers 作品と比較して Christie 作品において最も過剰使用されているのは、一人称代名

詞と動詞や前置詞の組み合わせ(*I think, I feel, for me* など)で一人称としての意識を表す Self-Disclosure であった。推理小説故に会話文でこの LAT が多く出現することは明らかであるが、Sayers と比較すると Christie 作品では地の文でも一定数使用されていることが確認された。Christie 作品は作中の登場人物が語り手を務める作品も多く、両作家の地の文での語り手の違いが反映されている可能性があると考えられる。また Christie は熱中や傾倒を表す Intensity(*very, indeed, I do, great* など)の使用も多いことが明らかになった。特に Intensity に関しては同一作品における同じ語の繰り返しが顕著に見られ、Christie の語彙多様性が晩年になるに従い低下したこと(Lancashire & Hirst, 2009; Le et al., 2011)を反映していると考えられる。以上のような修辞的特徴から、Christie 作品の文体的特徴の分析を試みる。

■9 月 30 日（土）
【研究発表第 2 セッション】
【研究発表 1】

Investigating the Impact of Extensive Reading with Data-Driven Learning

Gregory HADLEY（新潟大学）・ハドリー　浩美（新潟大学）

This presentation discusses an ongoing research projected investigating the use of data-driven learning (DDL) as a means of stimulating greater lexicogrammatical knowledge and reading speed among lower proficiency learners in an extensive reading program. From April 2015 to July 2017, students from six extensive reading classes were chosen for this study. For 16 weekly 90-minute sessions, an experimental group (21 students) used DDL materials created from a corpus developed from the Oxford Bookworms Graded Readers, which contained 186 books from all seven levels with a total of 1,715,160 tokens (17,670 word types). The control group (28 students) had no DDL input. All students were required to read a minimum of 200,000 words during the course. Students not reaching the 200,000 word threshold were removed from this study. Quantitative data from a C-test (Klein-Braley & Raatz, 1984) constructed from an upper-level Bookworms reader. A speed reading test by Quinn, Nation, & Millett (2007) was also selected. A pre-test post-test design was used, and dependent as well as independent samples *t* tests were used. Pre-test analysis found that the experimental group was in statistically distinct from the control group in terms of having lower levels of second language proficiency. Post-test scores found that, within both groups, the learners improved significantly, with high impact factors. However, the experimental group improved more to the point that the entered the same statistical bands as the control group. Post-test findings also indicate that students using the DDL materials were reading more books and reading faster than the control group. The study concludes that an informed use of DDL can work with lower proficiency learners, and that the methodology can be used to improve receptive learning and lexicogrammatical proficiency better than extensive reading alone.

【研究発表 2】

日本人英語学習者の関係代名詞の回避―CEFR レベルを用いた検証―

高橋　有加（東京外国語大学大学院生）

本研究では、同一実験参加者による①通常の英作文と、②関係詞を使うように指示のある英作文の、2 つのタスク内での関係詞の使用頻度とエラー率を比較することにより、知識はあっても使用が回避される傾向が CEFR レベル別にどの程度あるのかを明らかにすることを目的とする。

関係詞は日本人英語学習者にとって難しいとされるものの 1 つであり、使用が義務的でない文法項目である。Shachter（1974）は、学習者にとって難しいものは使用が回避される傾向があるとし、その例として関係詞を取り上げている。また、近年日本の英語教育にも大きな影響を及ぼしているヨーロッパ言語共通参照枠（Common European Framework of Reference：CEFR）の 6 つのレベルを同定する言語特性である基準特性（criterial features: Hawkins & Filipović、2012）を学習者コーパスから抽出する研究が行われており、Hawkins（2009）は基準特性として有効な文法項目の 1 つとして関係詞の使用を挙げている。しかし、自発的な産出が求められるパフォーマンステストなどでは、知識があっても実際に産出されないと評価することができないため、自然に産出される関係詞と、意識的な関係詞の使用を求められるタスク内における頻度及びエラーについて比較した。研究設問は以下のように設定した。

RQ1. 関係詞の回避または不使用の現象が実際にどのくらいあるのか？
RQ2. 意図的に関係詞を使用するようにした場合、エラー率は増加するのか？

対象とする関係詞は、that、which、who、whose、whom の 5 つで、省略形は含まない。また、SLA 理論の習得難易度階層を示す Noun Phrase Accessibility Hierarchy（NPAH: Keenan & Comrie、1977）、SO Hierarchy Hypothesis（SOHH: Hamilton、1994）の分類ごとにも集計した。実験参加者は、英検級を保持する日本人英語学習者 93 名である。文部科学省（2015）の対照表によると、5-3 級: A1、準 2 級: A2、2 級:B1、準 1 級: B2、1 級: C1 であるため、それぞれのレベルから 20 名程度の実験参加者を集めた。

データ収集方法として、全ての実験参加者に 2 種類の英作文タスク（20 分間辞書なし）を異なる日時に受験して

もらった。1つ目は通常の自由英作文（描写タスク）で、2つ目は同一のタスクに「できるだけ関係詞を使うように」という指示を加えたタスクである。データ処理として、全ての英作文を書き起こし、(a)関係詞の表層形、(b)NPAH、SOHH のタイプ、(c)エラー情報について、1文ごとに人手でアノテーションを施した。

調査の結果、関係詞を使うように指示のあるタスクでは、全レベルで関係詞の頻度が大幅に増加した。このことから、関係詞の知識があっても間違いを恐れたり、関係詞の使用が不必要であると判断した場合には使用されない関係詞が多くあることが分かった。エラー率に関しては、A2 レベルにおいて1回目より2回目の英作文内に特にエラーが多く見られたことから、間違いを恐れるために関係詞の使用を避ける傾向がある可能性が示唆された。

【研究発表 3】

英語教材としてのアニメ分析

寺島　芙由（東京外国語大学大学院生）

アニメ人気の高まりを受け、アニメを英語教材として提案する研究（佐々木, 2005）や、英語版アニメを活用した授業（吉田, 2012; 佐藤, 2014）がみられる一方、実際にアニメが効果的な学習方法かどうか検討を行っている研究は稀である。そこで本研究では、アニメの英語学習への活用の可能性を、コーパスを使って語彙の観点から調査した。

対象は英語に吹き替えられた日本のアニメ4作品で、比較対象としてアメリカで制作された映画1本も分析した。以下、3つの研究設問に沿って、調査の手法と結果を示す。

1. アニメの視聴は学習に適しているか。

アニメ全てのテキストと、BNC spoken corpus の上位 100 語および 200 語、バイグラムおよびトライグラムの上位 100 件を比較したところ、その多くがアニメに高頻度で現れたことから、アニメは効果的なインプットである可能性が示された。ただし、AntWordProfiler を用いて分析を行ったところ、General Service List（GSL）と Academic Word List（AWL）に含まれる語彙がアニメのテキストの約 90%を占めていたため、最低でもこれらの語彙知識がない初級・中級学習者にはアニメによる英語学習は適切ではない可能性が示された。

2.　アニメごとにどのような特徴があるか。

分析結果を以下の表に示す（括弧内は割合（%））。

	時間	延べ語数	TTR	GSL 1000, 2000	AWL	その他
アニメ 1	110 分	14291	0.163	12478 (87.31)	294 (2.06)	1519 (10.63)
アニメ 2	110 分	11459	0.127	10390 (90.67)	107 (0.93)	962 (8.4)
アニメ 3	110 分	13062	0.145	11642 (89.13)	207 (1.58)	1213 (9.29)
アニメ 4	110 分	11419	0.131	10268 (89.92)	126 (1.1)	1025 (8.98)
映画	124 分	7748	0.197	6747 (87.08)	149 (1.92)	852 (11)

簡単な語や同じ語の繰り返しが多いアニメ2は、難解な語を多く使い、延べ語数や同じ語の繰り返しが少ない映画と比べ、学習者にとってより効果的である可能性などが示された。

3. アニメによる自主的な学習のためにどのような援助を行うのが効果的か。

AntConc の keyword list を使ってアニメ2を分析すると、アニメに特有な単語として、固有名詞やアニメにおける用語などが発見された。加えて、アニメの内容に基づいてとりわけ多くみられるコロケーション（例: wish for）が存在することがわかった。これらの単語や表現を事前に提示することで、学習者の負担を減らし、学習を促すような援助方法が考えられる。

発表では、本研究の課題や、今後必要とされている語彙以外の面での研究も考察する。

【研究発表 4】

小中連携に向けた英語授業コーパスデータ構築とインタラクション分析の試み

片桐　徳昭（北海道教育大学）・大橋　由紀子（ヤマザキ学園大学）

1.　背景と目的

平成 32 年(2020 年)度から新学習指導要領の完全実施となり、小学校で英語が教科化される。文科省が謳う小中連携の一つに、小中 9 年間の学びの中の「学習指導の継続性」という視点がある。そこで本研究では、小中学校での英語授業のコーパスの構築において、インタラクションタグ付与が小中の学習指導の継続性を調べる上で有用なアノテーションとなるかの調査を試みた結果の報告をする。以下の研究課題に基づき, 英語の授業内でのインタラクション分析を行い、小中の接続点である、小学校 6 年生の終了時と中学 1 年生の開始時の英語授業の「継続性」がど

のように観察されるかについて調査した。
研究課題 1. インタラクション情報は小中英語授業の「継続性」分析に有用なアノテーションか。
研究課題 2. 連続する小中英語授業の「継続性」はインタラクションに観察されるのか。
2.　データ収集と分析方法
同じ国立大学に附属する小学校 6 年生の最後の英語の授業 4 回と連続する年度の中学校 1 年生の最初授業 6 回の授業データを収録した。計 10 回の授業データについて教師・生徒の発話(日英両語)について、話者タグ、言語タグに加えてインタラクションタグを付与した。インタラクションタグとして Walsh (2006)の言う SETT (Self-Evaluation of Teacher Talk) から 4 つの interaction mode (managerial, materials, classroom context, skills and systems)と Ellis (1984)の述べる social という考えを属性値として組み込んで分類した。インタラクションタグは主に発話ターンごとに付与したが、質的変化があると判断された場合には、同 1 ターンを細分割してインタラタクションタグを付与した。
3.　結果
インタラタクションタグ数は小学校 100 回 (*M*=25.0)・中学校 136 回(*M*=22.7)となり、小学校・中学校ともに授業体制を構築する managerial mode が最頻出(小学 51 回[*M*=12.8]、中学 74 回[*M*=12.3])となり、教師主導の授業展開の傾向が見られた。また、skills and systems mode による言語材料の習熟活動(小学 26 回[*M*=6.5]、中学 33 回[*M*=5.5]), classroom context mode による教師生徒間のインタラクション(小学 17 回[*M*=4.3]、中学 18 回[*M*=3.0])も同等の傾向を示した。しかし、授業目標の核となるコミュニケーション活動を示す materials mode において、小学 4 回[*M*=1.0]、中学 13 回[*M*=2.2]となり中学校ではわずかながら発展性が示唆される結果が観察された。
発表当日は、(1)コーパス構築の手順, (2)データ整理の方法, (3)分析結果の詳細ついて述べ、(4)インタラクションタグ付与に関して Walsh (2006, pp. 82-91) が言っている mode switching, mode side sequences, mode divergence といった deviant cases の問題についても触れ、データの利用の拡張性について考察を加える。

■9 月 30 日（土）
【研究発表第 3 セッション】
【研究発表 1】

複数コーパスの統一的処理を可能にした高速コーパスデータベースシステム MyCo の開発

西村　祐一（元名古屋大学大学院生）

記録方式の異なる様々なコーパスが併存している今日、複数のコーパスを統一化された環境で高速に処理できるデータベースシステムを開発することの意義は大きい。そこで、リレーショナルデータベース管理システム MySQL を利用して、英語コーパスのデータベースシステム（MyCo と呼称）を開発することとした。例えば、BNC と COCA は設計が大きく異なるが、MyCo を用いれば、これらを同一の利用環境で、選択的にあるいは統合して処理することが可能である。随時、様々なコーパスを追加できる設計にしたことによって、入手した複数のコーパスを比較利用できる点も研究上、有益である。また、データ処理にかかる時間を大幅に短縮できたことも、研究遂行上、極めて重要な点である。本発表では、コーパス利用の観点から MyCo の特長を述べる。

1. データベースに記録するデータ
最小単位は WLP（Word, Lemma, POS-tag）である。研究目的に応じて W, L, P を組合せた検索式を指定してデータを抽出する。
2. データ抽出処理
コーパス利用の中心作業は、検索する語（またはレマ）とその周辺に共起する要素の文字列を抽出することである。MyCo では抽出範囲を検索語の前後各 10 要素に固定し、検索語の WLP をノードとする 21 個の WLP から成る文字列 (21gram) をテキストファイルに出力する。さらに、このテキストファイルを利用者が直接利用することも可能である。
3. 資料の加工
MyCo は、抽出した 21gram をもとにピクチャー、kwic リスト、コロケーションリストを作成できる。ピクチャーを例に説明すると、表示範囲、集計対象（Word, Lemma, POS）、値（頻度、MI-score、t-score）、特定の位置に現れる語またはレマを含むテキストの一覧表示、などである。
4. 検索処理時間
BNC によるデータベースを例に、because（頻度 100,509）、something（頻度 50,062）、maybe（頻度 10,025）について 21gram 抽出時間の実測値を例示すると、それぞれ 2.7 秒、1.3 秒、0. 3 秒である。1 億語規模のコーパスから頻度 5 万件程度の語の 21gram をほぼ 1 秒で得ることができ、高速レスポンスを実現している。これは、研究上、全くストレスを感じずに済む速度である。また、単一の語またはレマの検索に加えて、イディオム kick the bucket のような複数要素を組み合わせた検索も可能である。
5. 実行環境
MyCo を、Linux (CentOS 6.9) 、Perl v.5.10、MySQL v.5.1 で開発し、利用している。同等環境の PC であれば MyCo を容易に搭載できる。

【研究発表 2】

構文構造を活用した学術論文における頻出コリゲーションの抽出

田中　省作（立命館大学）・徳見　道夫（九州大学）・宮崎　佳典（静岡大学）・
金丸　敏幸（京都大学）・田地野　彰（京都大学）

現在、分野別の学術論文コーパスをもとにした、分野に依存しない頻出表現の整備を試みている。コーパスからの頻出表現の抽出には、論文コーパス内の各英文で、分野依存性の高い語を適当な痕跡や品詞に置き換え、n-gram で計数すれば良いように考えられる。しかし、n-gram は、“take 〜 into account”のように自由項を挟んだ不連続な頻出表現の捕捉が難しく、また適切な n の設定も問題となる。そこで、本研究では、痕跡に相当する情報を含む英文から、構文構造を考慮した頻出表現の抽出法を提案する。構文構造の考慮と n-gram の重み付けによって、痕跡が木構造における節表示となったコリゲーションの抽出も可能となる。

提案手法は次の通りである。まず、痕跡の取り扱いである。構文構造を参照し、内容語を含まない、痕跡をまとめて導出する最上位の節表示に置換する。たとえば、“The algorithmic procedure takes supremum norm into account”という情報系論文の英文で、“algorithmic”, “supremum”, “norm”を分野依存性の高い語とし、痕跡に相当したとする。SNLPG(n.d.)で与えられる構文構造を参照し、痕跡を節表示に置換すると、“<NP> procedure takes <NP> into account”となる。ただし、<α>は α という節表示、NP は名詞句を表す。コーパス内の英文を全てこのように置き換える。

次に、このように節表示を含んだ英文の n-gram の重み付け計数である。節表示を含む n-gram の重みはそれが元の英文で内含している語の n-gram 数とする。例えば、n=4 のとき“<NP> procedure takes <NP>”の重みは 3、“takes <NP> into account”の重みは 2 と計数される。高次の節表示を多く含む n-gram の方が重く計数される傾向となる。

京大論文コーパスに対して、分野依存性の高い語を痕跡に見立て、上記方法で n を適当に動かし、累積的に計数した。その結果、“between <NP> and <NP>”や“according to <NP>”という具合に節表示を含む、より分かりやすい形で、様々な長さの頻出表現を抽出できることを確認した。

【研究発表 3】

Word2Vec による文学作品の時代比較―コーパスを軸とした異分野融合研究の試み―

内田　諭（九州大学）・下條　恵子（九州大学）・渡邉　智明（九州大学）・
斎藤　新悟（九州大学）・谷口　説男（九州大学）

近年、コンピュータの発達と言語データの蓄積により、自然言語処理の技術は大幅に向上してきている。ニューラルネットワークによる翻訳精度の向上(Wu et al., 2016)、単語行列をベクトル化して単語の意味を数値で表す Word Embedding の手法(Levy and Goldberg, 2014)の発展など、その進歩は目覚ましく、大きな注目を集めている。一方でこれらの手法を言語学や文学、文体論などのいわゆる人文社会系の研究への応用については、十分に議論されているとは言えず、未開拓の部分が多く残されている。

本研究の目的は、Word Embedding の代表的な手法の一つである Word2Vec (Mikolov et al., 2013)を用いて単語の用法を独自に作成した文学コーパスで検証し、文学研究への応用の可能性を探ることである。文学作品について計量的あるいは統計的なアプローチを用いる研究は増加傾向にあるが、Word Embedding の手法を用いた研究は限られており、その応用方法については現段階では未知数であるといえる。この手法の最大の特徴は、共起関係に代表される syntagmatic な関係にある単語ではなく、類似のベクトルを持つ単語を探索することで paradigmatic な関係にある単語を検証することができるという点である。特定の単語と類義的に使用されている単語を調査することで、その単語が使用される文脈や含意などを明らかにすることが可能となる。

本研究では 1900 年代前半の文学作品からなる「前半コーパス」（約 180 万語）と 1900 年代後半からなる「後半コーパス」（約 171 万語）を作成した。対象となる作品については、アメリカ文学作品の中から、金融業界を取り扱った作品及び現実を写実的に描いたとされるリアリズム小説を中心に選定した。20 世紀のアメリカは 1945 年の第二次大戦終戦を機に国際政治の舞台で超大国の地位を獲得しただけでなく、経済面でも規制緩和を繰り返して活性化を図るなど、社会的に大きな変化を経験している。そしてこのような変化は文学作家たちの意識形成に影響を及ぼしており、作品内容だけでなく語彙レベルでその変化が生じていると考えられる。

これらの 2 つのコーパスを入力としてそれぞれの Word2Vec のモデルを構築した。その後、単語の頻度表から高頻度の名詞を選定してコサイン類似度から同義的に使用されている単語のリストを生成した。その結果、類義語のリストは年代によって興味深い違いを示すことが明らかになった。例えば、money は「前半コーパス」では pay, work, dollars などと類似度が高い。一方、「後半コーパス」では interest, market, bonds などが類似度の高い語としてリストされた。これは「前半コーパス」の時代には労働に対する具体的対価として金銭が語られているのに対し、「後半コーパス」では金融商品など利益を生む無形の財やサービスとして語られているということを示唆する。これらの実験結果に対して、本研究では文学・政治学等の観点からの解釈を試みる。言語データの裏側にある事実や社会変化などを多角的に読み解き、コーパスを基軸とした異分野融合の可能性についても議論を行う。

【研究発表 4】

構文情報などを表す木構造の配列による情報処理

田中　省作（立命館大学）・宮崎　佳典（静岡大学）・田辺　利文（福岡大学）・田村　昌彦（立命館大学）

近年、実用的な構文解析ツールも増え、構文構造などの語や品詞の連鎖よりも高次の言語情報処理が可能となり、コーパス研究においてもその活用が期待される。本発表は、このような構文情報付き言語データをコーパス研究で活用するための技術的提案である。

コンコーダンサ等の既存分析ツールを超える処理をする際、文字列・語列程度の取り扱いであれば自身でプログラミングするコーパス研究者も少なくない。一方、構文情報は複雑で「木構造」とよばれるデータ構造で表されることが多く、その取り扱いは格段に難しくなる。木構造は「構造体」「ポインタ」「再帰」など、語列処理では無縁のプログラミング構成概念が求められる(情報処理推進機構, 2016)。情報系学生でも慣れないうちは混乱することもあり、コーパス研究者にとってはなおさらである。そこで、本発表ではプログラミングの初期に学習し、語列処理等でも使う「配列」、具体的には 2 つの配列 c、d で木構造を表し、処理する方法を提案する。

配列の一つ c は句表示や語といった節・葉の言語情報、もう一つの配列 d は根（頂点の節）からのパス長を格納する。格納順序は、木構造を根から深さ優先最左探索した順序である。配列では木構造が平坦化してしまったようにみえるかもしれないが、次の基準で配列を先頭から走査し、節 i を解釈すると、木構造が表現されていることがわかる。

1. i=0 ⇒ 根
2. d[i]=d[i-1]+1 ⇒ i は、i-1 の最右子節
3. 1、2 以外 ⇒ i は、j<i で d[j]=d[i]-1 となる最も大きな j の最右子節

なお、d[i]は、配列 d の i 番目の要素を表し、最右子節は i-1 までに導出される子節のなかで最も右側の子節という意味である。

このように木構造を配列に直すことによって、木構造間の照合を比較的単純な配列間の照合に帰着でき、語列処理とさほど変わらない複雑さで実装できる。本発表では、構文情報が付された英文に対して、構文的な特徴を考慮した、本方式による検索事例も紹介する。

■9 月 30 日（土）

【研究発表第 4 セッション】

【研究発表 1】

強調語の調査による Popular Music の歌詞の文体研究

渡部　文乃（京都大学大学院生）

1. 背景・目的

本研究の目的は、Popular Music の歌詞の文体を明らかにすることである。Popular Music とは、不特定多数の聴衆に向けられた、利益を目的とする音楽 (cf. Tagg 1982: 41-42) のことで、その起源は 19 世紀末にも遡る (cf. Frith 1986: 79) ことができる。しかし、Popular Music が本格的に学術分野として扱われ始めたのは最近のことであり (Tagg 1982: 37)、歌詞の研究はほとんどない。文体は、社会言語学や歴史言語学などの様々な言語研究の議論において考慮に入れるべき重要な事項であるため、本研究はそのような研究を行う前段階として、Popular Music の歌詞の文体について調査を行う。

2. 方法

本研究は、強調語の分布に注目する。強調語を研究対象とした理由は、この項目の調査によって、① 話し手と聞き手との関わり (cf. Hyland 1998)、② 文体のフォーマリティー (cf. Yaguchi et al. 2009)、③ 主観的性 (cf. Hyland 1998) などの文体的特徴が明らかになると期待されるからである。本研究では、申請者が構築した American Popular Music Corpus of English (PMCE-US) という英語歌詞コーパスを使用し、20 個のジャンルから構成される均衡コーパス Manually Annotated Sub-Corpus (MASC) と比較して調査するため、Popular Music の歌詞の④ジャンル間における位置づけについても言及する。本研究は形容詞を修飾する副詞の強調語をすべて抽出し、頻度をジャンルごとに比較した。

3. 結果・考察

PMCE-US と MASC の比較から明らかになったのは、歌詞が独特の文体をもつということである。強調語全体の頻度に関して、PMCE-US には 200 個以上の強調語が出現したが、この頻度は MASC の中で最も強調語の出現頻度が高いジャンルである twitter と比べても著しかった。さらに強調語の種類に関しても、幅広いジャンル (e.g. journal, court, e-mail) で見られる very や、口語のジャンル (e.g. face-to-face) や新しいジャンル (e.g. blog) に見られる really はほとんど現れず、出現数の 8 割を占めていたのは so であった。

このような結果から、歌詞は MASC のどのジャンル以上に、話し手が積極的に聞き手を説得する（＝積極的に聞き手と関わろうとする）、主観的でインフォーマルな文体であることが明らかになった。また、強調語 so が歌詞の強調語の全体頻度の 8 割を占めていたのは、音楽のリズムが短い音節の語を好む傾向があることや、最も収益が見

込まれる若者世代の言葉に近づけるための作詞者の意図的な言葉の選択と考えられ、歌詞はそのように他のジャンルでは見られない特徴をもつジャンルであることが分かった。

【研究発表 2】

ホテルのオフィシャルウェブサイトにおける概説文のストラテジー—Move の構築と分析を中心に—

近藤　雪絵（立命館大学）

本研究はロンドンのホテルのオフィシャルウェブサイトに掲載された概説文を Swales（1990, 2004）が提唱したジャンル分析の手法を用いて分析し、読み手にアピールするストラテジーを探求することを目的としたものである。ロンドンの 3-5 つ星のホテルのウェブサイトに掲載された概説文（3 つ星ホテル 11、4 つ星ホテル 66、5 つ星ホテル 47 の計 124）を集積し、書き手の意図と特徴的な言語表現を元に分類したところ、3 つの Move と 3 つの Step が構築された。Move と Step の概要を表 1.に、ホテルのグレード（星）別 Move 採択率を表 2.に示した。

表 1. Move・Step とその機能

Move		機能
Move 1	Defining self	ホテル自身を定義する
Move 2	Establishing features	ホテルの特徴を確立する
	Step 1: Describing the history/architecture	歴史／建築を述べる
	Step 2: Describing the location	所在地を述べる
	Step 3: Describing the facilities	設備を述べる
Move 3	Establishing connections	ホテルと読み手との関係を築く

表 2. ホテルのグレード別 Move 採択率

	3-star	4-star	5-star
Move 1	81.8%	83.3%	83.0%
Move 2	100.0%	92.4%	87.2%
Move 3	81.8%	75.8%	53.2%

Move 1、2 は採択率が全てのグレードにおいて 8 割を超えており、自身を定義付けてから特徴を確立することがホテルの概説文の典型パターンであることがわかった。ホテルのグレードが下がると Move 2 の採択率は高まり、3-star では全ての概説文に Move 2 が採択された。一方で、Move 3 は 3-star ホテルでは Move 1 と同じ 8 割強の採択率であるが、5-star ホテルでは 5 割強にとどまった。Move 3 では二人称代名詞を用いて読み手に呼びかけたり、予約を促したりする表現が見られ、中グレードホテルではこのような表現を使い読み手と関係を築くストラテジーが使われていた。概説文は一見すると自由に創作された文章のように思われるが、Move 分析を行うことで典型的なパターンが存在し、中グレードホテルと高グレードで読み手にアピールするストラテジーが異なることがわかった。

【研究発表 3】

一般教書演説から見る米国大統領の関心事の変遷 —トピックモデルと時代背景—

木山　直毅（北九州市立大学）

米国大統領（以下、大統領）は毎年 1 月に一般教書演説（State of Union Address）を行い、その時の関心事を連邦議会に対し演説を行う。本研究ではこの演説原稿をコーパスとし、トピックモデルと呼ばれる手法を用いて大統領の主要政治課題がどのような社会背景に影響を受けてきたのかを明らかにする。

本研究では、様々なウェブサイトで公開されている一般教書演説を 1 つにまとめ、表 1 のようなコーパスを作成した。このデータに対し、ストップワード（田畑（2017）が利用したものを改訂）処理を施し、潜在的ディリクレ配分法（Latent Dirichlet Allocation）（Blei 他、2003）によってトピックを解析することで、35 個のトピックから 3 個の主要トピックを得ることができた。

表 1: 一般教書コーパス情報

総ファイル数	トークン頻度	タイプ頻度
228	1,751,570	32,204

まず、米国の初代大統領が就任した 1790 年から 1900 年までの 110 年間と、その後、断続的に 1916 年までの間、*congress* や *government*、*country*、*duty*、*duties*、*law(s)*といった語彙が多く現れる。これらのうち、*duty* のコロケーシ

ョンを調査すると、大統領の仕事を強調する *my duty* という表現は 134 例のうち 116 例がこの期間に現れている。また *duties of the federal/general government* といった、政府の役割を強調する表現は 27 件中 26 件がこの期間に現れる。このことから本トピックは「国の役割」と言える。

次に、1914 年、1915 年、1917 年、そして 1932 年から 1980 年までの間、*peace* や *world*、*freedom*、*national*、*security*、*defense* といった表現が頻出するようになる。これらの語彙のコロケーションを見ると、*peace and freedom in the world* や *national security*、*national defense* といった表現が目立つ。そのため、この時代の主要トピックは「軍事」であったと言える。

最後に、1960 年代からトピックの重要性が上がり始め、1982 年以降、最も重要であるとされるトピックは、*jobs*、*children*、*families*、*health*、*care*、*working* といった語彙が目立つ。これらをコンコーダンスラインで確認すると、*working families* や *health insurance*、*create new/good/more jobs* といった表現が目立つ。このことから、近現代の大統領の関心事は「社会福祉」に関してであると言える。

以上を総括すると、大統領の関心事は[[国の役割 => 軍事関連 => 社会福祉]]というトピックの変化を辿ってきたことになる。本研究では、一般教書演説のトピックは建国時の内政、世界大戦や冷戦、そして冷戦後の米国内景気の悪化という社会背景に影響されていることを、コロケーションなどの観点から論じる。

■9 月 30 日（土）

【研究発表第 5 セッション】

【研究発表 1】

コーパスの示す科学的データと学習性・商品性との両立―『ウィズダム英和辞典』の編集にあたって ―

井上　永幸（広島大学）・西垣　浩二（㈱三省堂辞書出版部）

『ウィズダム英和辞典』は、本格的にコーパスを活用して編纂された初の英和辞典として、初版以降 2 回の改訂を行い、現在第 3 版が刊行されている。コーパス分析の成果をいかに盛り込んでいるかというような点については、過去に本学会やその他の機会に何度か発表をしているが、本発表ではコーパスから得たデータ、あるいはそこから漏れるものを、いかにして掘り下げ、学習者にとって有益な記述を作り上げてゆくかという点について、実例をもとに論じることとしたい。

例えば「頻度」や「コロケーション」などについて、コーパスの分析により得られる知見は、コーパスを用いた辞書編纂にあたってはもっとも参考とすべきところであるが、当然「学習者が効率的に学習を進めるための教材」を編纂するという立場に立つのであれば、頻度の高い表現であれば何でもすべてを採用するという考えには立てない。学習性、日常生活への密着度（これはコーパスには反映されにくい場合もある）、記述分量と学習効果のバランスなど、さまざまな点について熟慮を加えたうえで、紙面に反映しているのである。これは、語法・類義解説、用例採用基準、語義分析など、辞書のすべての記述において言えることである。コーパスの産出する科学的データはそれ自体が目的となるのではなく、それをもとに学習者にとって必要な情報を取り出し、いかに学習効率が上がるように情報を配置してゆくかという点が、学習者にとって使いやすく、結果的に商業的にも成功する教材を編纂するにあたって肝要な点である。

本発表においては、コーパスから得たデータをいかに昇華させ、学習効率の高い教材を編纂してゆくかという点について論じることとしたい。また、開示できる範囲内で、コーパス構築・活用の裏話等にも言及する。井上からは、『ウィズダム英和辞典』における記述について、実例を挙げながらどのようなデータを反映しているのか、コーパスデータに「加えて」、どのような点を編者の「経験と勘」によって補っているのかについて論じる。西垣からは、成功する「商品」としての辞書の編集にあたって、記述内容とターゲットユーザの求める情報との整合性、記述分量のバランスなどについて補足をする。

【研究発表 2】

英語辞書レーベルとコーパス

田畑　圭介（神戸親和女子大学）

現在の英語学習辞典において学習者に語彙情報を効果的に伝達する手法の一つにレーベルの提示がある。レーベルは該当語の用法だけでなく他の語との差別化を図る機能も持ち合わせている。本発表では最初にレーベルにヘッジが付与されうる集合タイプのものと付与されない段階的タイプのものに二分されることを例証する。そして[disapproving]と[offensive]のレーベル、[informal]と[spoken]のレーベルへの細分化の必要性を COCA とテレビドラマコーパスのデータをもとに論証する。Longman Dictionary of Contemporary English (LDOCE)では[offensive]は採用しておらず、Oxford Advanced Learner's Dictionary では[spoken]は採用していないが、condescending、conventional、brigade、bowdlerize、foolhardy といった語の使用状況から、[disapproving]、[非難して]のレーベルの必要性、foreigner、fruit などの使用状況から[offensive]、[けなして]のレーベルの必要性が帰結される。you could have fooled me は COCA で FIC10 例、SPOK3 例、MAG1 例、NEWS1 例の計 15 例検出されるが、各用例はいずれも会話文で用いられている。

これは会話で用いられる表現であることを示すと共にCOCAのジャンル分析に注意が必要であることも暗示している。fortunately と luckily を COCA の SPOKEN で検索すると、fortunately 1107、luckily 532 となる。日常的な表現である luckily の使用が予想外に少ないが、これらは COCA の SPOKEN がくだけた会話体でなく報道番組の発話データが中心となっていることに起因している。これらの事実は fortunately を[ややかたく]として luckily と差別化し、LDOCE のように spoken≠informal と捉える必要性を示すものとなる。

【研究発表 3】

怒りを表す類義語と概念メタファー

南澤　佑樹（大阪大学大学院生）

本発表の目的は､コロケーションを抽出する統計手法を用いてメタファー表現を収集し､怒りを表す類義語 *anger*、*rage* に見られる概念メタファーの違いを示すことである。

感情のメタファーに関しては、概念メタファー理論の枠組みに基づき盛んに議論が行われてきた。怒りの感情に対しても複数のメタファーが提案され、怒りを容器内の熱い液体とみなす ANGER IS A HOT FLUID IN A CONTAINER (“*boiling with anger*”) はその中でも最も中心的なメタファーとされている (Kövecses, 1990, 2000)。しかしながら、多くの先行研究では *anger* と *rage* が区別されておらず、類義語間の違いにはこれまであまり関心が向けられてこなかった。この理由には、従来のアプローチでは結果を量的に分析することが困難であったという点が挙げられる。

近年では、概念メタファー分析にもコーパスが用いられるようになっているが、現段階ではメタファー表現をコーパスより網羅的に収集することは難しい。これを踏まえ Stefanowitsch (2006) は、根源領域に属する語と目標領域に属する語（本発表では感情語）が共起する表現であるメタファーパタンを分析する手法を提案している。Turkkila (2014) は、この手法を用いて怒りの類義語に見られる概念メタファーを分析しており、*anger, rage, fury* に見られるメタファーが概ね同じであると主張している。しかし、コーパスを用いた手法では、メタファー表現の出現頻度によって概念メタファーの重要度を決定することが多いことから、*boiling with anger* や *his anger welled up* のような感情の様態を具体的に表す表現よりも *in anger* のような意味内容の薄いメタファー表現を重要とみなす傾向がある。しかしながら、そのようなメタファー表現は他の感情等にも用いられる一般性が高い表現であるため、それらによって類義語間の違いを見いだすことは難しい。

これらを踏まえ本発表では、British National Corpus よりそれぞれ *anger, rage* を検索語としてメタファー表現を収集し、その相違点を指摘する。本研究では、抽象的で一般性の高いメタファー表現ではなく怒りの様態を具体的に表すメタファー表現を収集するため、意味的に結びつきの強いコロケーションを測る MI スコアを用いる。その結果、*anger* は *vent, seethe, well, simmering* といった語と結びつきが強く、従来の主張通り ANGER IS A HOT FLUID IN A CONTAINER が最も中心的な概念メタファーであると言えるのに対し、*rage* では、*anger* と比較して *howl, bristle, murderous* といった語がリストの上位にきていることから ANGER IS A HOT FLUID IN A CONTAINER だけでなく ANGER IS A DANGEROUS ANIMAL (“*bristling with rage*”) とも結びつきが強いと言えることを主張する。また *anger* と結びつくメタファー表現は怒りの様々な側面を表すのに対し、*rage* と結びつきの強いメタファー表現は怒りの特定の側面を表す傾向にあることも主張する。

■9 月 30 日（土）

【研究発表第 6 セッション】

【研究発表 1】

日英対訳コーパス中の「～ことになる」構文とその英訳文間の構造的不一致

大矢　政徳（目白大学）

機械翻訳の分野における Bitext word alignment(BWA)手法では、翻訳元の文中の単語と翻訳先の文中の単語との対応関係（単語アライメント)を統計的に算出する。しかしながら、日本語の表現には、英語では複数の表現で翻訳される場合や、日本語には存在しない要素を補完しなければならない場合が高頻度で存在し、この領域は BWA 手法ではカバーしきれない。例えば､「健は明日帰国することになっている」という日本語文は､英語では典型的には"Ken is supposed to come back to Japan."と訳され、「～ことになる」は英語では"be supposed to"に対応しているが、実際の「～ことになっている」構文は、必ずしも"be supposed to"と訳されているとは限らない。また、日本語では主語や目的語が省略される場合が多く、それらを英語のように主語や目的語が省略される頻度が低い言語へと翻訳する場合に問題となる。

本研究では、このように BWA 手法ではカバーしきれていない領域を補完することを目的とし、翻訳元文と翻訳先文との対応関係を、単語間のアライメントだけではなく、単語間依存関係木の統語的不一致パターン(syntactic divergence patterns)として提示することを提案する。コーパスデータとして『Wikipedia 日英京都関連文書対訳コーパス』の日本語文と英語対訳文をそれぞれ構文解析して得られた単語間依存関係情報を用い、日本語構文「～ことに

なる」を含む日本語文 1197 文から無作為に 100 文を選び、これらの文中の「～ことになる」が英語対訳文のどの単語・フレーズに対応しているかを統語的不一致パターンとして人手で抽出し、各パターンの発生確率を算出する。特に、依存文法の枠組みでは、統語依存木の根の位置にある述語（主節の動詞がこれに該当する場合が多い）がどのような要素を述語項として要求するか（例：形容詞を含む名詞句か、それとも関係節を含む名詞句か）がその統語依存木全体の構造を決定する点を鑑み、日本語構文「～ことになる」が統語依存木の根の位置にある場合に、これに対応する英語対訳文では統語依存木の根の位置にないという統語的不一致パターンの発生確率が高いことを示す。さらに、統語的不一致パターンを人手によらず自動で抽出することを目的として、当該 100 文から得られた「～ことになる」統語的不一致パターンを、上述コーパス中の「～ことになる」を含む日本語文 1197 文からこれら 100 文を除いた 1097 文からさらに無作為に選んだ 100 文に対して、正規表現でマッチングすることによって自動的に抽出する手法を試み、その可能性と改善点について論じる。

【研究発表 2】

医学研究論文ジャンルにおけるコーパス作成ツール AntCorGen を活用した教育の可能性
—Construction of Corpora for Discipline-Specific Learning in Medical Research Article Genres

浅野元子（大阪大学大学院生）

本発表の目的は、最近公表されたコーパス作成ツール AntCorGen (Anthony, 2017a) を用いて構築した比較的大規模な医学研究論文コーパスにおけるテクストの言語的特徴を量的質的に検討し、教育応用の可能性を探索することである。

英語が医学論文での国際共通語となり、小規模言語を母語とする大学院生や研究者 (Giannoni, 2008) は専門家集団に受容されるために分野特有の修辞パターンの習得が必要といわれる (Flowerdew, 2013)。研究論文を用いてミニコーパスを構築するデータ駆動型学習 (Data-Driven Learning) が提案されて久しいが (Anthony, 2017b; Lee & Swales, 2006; Noguchi, 2004; 朝尾・投野, 2005)、多様な専門分野の範囲をどのように限定して論文を使用すべきかについての報告は少ない。本稿では、医学研究論文の言語的特徴を検討した研究 (Nwogu, 1997) を参考に、同一誌での種々の医学研究論文に異なる言語的特徴があるかどうかについて AntCorGen に実装された PLOS ONE 誌の学術論文をコーパス化する機能を用いて検討した。

医学・健康科学分野のうち Cardiology (心臓病学)、Gastroenterology and hepatology (胃腸病学と肝臓学)、Pulmonology (呼吸器学)、Oncology (腫瘍学) の領域での論文各 5000 報を得た後に 100 報ずつを無作為抽出した。個々の論文をタイトルならびに抄録におけるムーブ (Swales, 1990; Salager-Meyer, 1990 & 1992) とヒント表現 (Tojo, Hayashi, & Noguchi, 2014) を手がかりに医学研究の種類すなわち動物での研究や症例コホート研究など (国立国際医療センター, 2009) に分類した。論文テクストの語彙を計量し (Imao, 2015)、本文の頻度上位語を変数としてウォード法、ユークリッド距離を用いてクラスター分析を行った (田畑, 2004)。

本文は総語数が 1,727,472 語、異なり語数を総語数の平方根で除して算出した Guiraud Index が 36.8 で、内容が詰まった文書であることが示唆された。語彙によるクラスター分析では医学研究の種類による類似性が示唆された。

本ツールは学術論文コーパス構築において利便性が高く有用であると考えられた。教育現場では、本ツールを用いて構築したコーパスにおける論文からタイトルや抄録を頼りに対象研究と同一種類の研究に関する論文を選択して各自のコーパスを作成すると、目標とする専門家集団が慣れ親しむ修辞パターンをより実践的に学習することが可能性となることが示唆された。

【研究発表 3】

Applying Topic Models to Describe a Corpus's Compositionality:
How can the external criteria be associated with meaningful sets of internal evidence?

Tomoji Tabata (University of Osaka)

Topic modelling is a machine learning method for uncovering hidden semantic structures in a corpus of texts. Based on a probabilistic inference algorithm, latent Dirichlet allocation, the technique makes it possible to identify sets of frequently co-occurring words, or topics, that characterize a text as well as classify texts into meaningful groups defined by inferred sets of strongly associated topics.

One of the major advantages topic modelling has over traditional key-word detection techniques, such as the Chi-squared test, log-likelihood ratio test, Mann-Whitney's *U* (or exact rank) test, or somewhat modestly but robustly applied Welch's *t*-test, employed in many stylometric/corpus linguistic studies is that topic models do not simply provide typical dichotomous or polarized sets of key-words for a target corpus versus a reference corpus, but enable us to spotlight key-words of multiple sets of texts, thereby making it possible to classify texts into reasonable subsets clustered in terms of word co-occurrence patterning. Outputs obtained from a topic modelling run range from a word-topic association table, which tells us what and how many topics a particular word belongs to and how much weight the word has in each topic; a topic composition table, which illustrates what types of words knit up a particular topic and to what extent individual words contribute to composing a given topic; to a text composition table, which accounts for topic density in each of analyzed texts, or to what extent a text is occupied by words

belonging to each of the topics associated with the text. Of further interest in the context of corpus linguistics is that results of a topic modelling can be visualized in the form of a topic box-plot, network diagram of topics and words as well as that of topics and texts, and a summarizing heatmap of topics and texts under investigation.

The present study applies topic modelling to the FLOB corpus with a view to analyzing latent semantic structures/patterns underlying in the corpus and mapping its subcorpora (or, registers) in the network of words, topics, and texts. What is of special interest is that by means of this approach it is now possible to shed new light on thematic/topical structures composed by a large number of infrequent words, which would otherwise escape the net of key-word statistics due to infrequency of occurrence or lack of a proper classification of lexical items.

This paper summarizes results of multiple topic modelling runs on the FLOB corpus. The paper reports that the text categories A—J (informative prose registers) are clearly distinguished from the texts that belong to the imaginative prose writings, or fiction (categories K—R) according to topical structures underlying in the corpus. The fifteen text registers in the FLOB corpus are classified into the two distinct clusters: informative versus imaginative proses. To turn our attention to topic distributions across registers, we can notice that the generated topics are divided into two sets: topics contributing more to the informative registers and those talked about more in the fictional registers.

Emerging results from this research are expected to open up a new avenue of inquiry into key semantic patterns in a large collection of texts, thereby suggesting a possibility of building a bridge between findings from machine learning text mining and traditional stylistics, distant reading and close reading, with an empirical interplay of insights that will benefit modern text analysis. Of further note is that in interpreting results of a topic modeling, we are likely to confront a sea of multitudinous contentious interpretations. Topic modeling is not just a cutting-edge machine-learning technique, it involves a highly humanistic interpretation and insight: the true value of machine-learning can only be judged by the depth of human insight, ironically but interestingly.

■10 月 1 日（日）

【ワークショップ 2】

機械学習を用いたコーパス分析入門

小林　雄一郎（日本大学）

本ワークショップでは、近年コーパス言語学の分野でも盛んに利用されるようになってきた機械学習 (machine learning) の技術を紹介します。機械学習は、人間が持つ学習能力をコンピュータに持たせることを目指す人工知能の研究分野です。また、コンピュータにデータを解析させることで、データの背後に潜むパターンを発見（学習）させる技術のことを指します。そして、多くの場合、データから発見されたパターンは、新たなデータの予測に活用されます。

機械学習の技術を用いることで、手作業では扱えないような大量のテキストデータを効率的に分析できるようになります。そして、パターンを発見するための十分な量のデータを用意すれば、人間が予測するよりも高い精度で予測を行うことが可能になります。さらに、予測に寄与したパターンを吟味することで、分析対象のテキストを特徴づける言語項目を特定することができます。

コーパス言語学における機械学習の活用事例としては、テキストの著者推定やジャンル推定、英作文の自動採点、語彙や文法の使用に関する通時的分析などがあります。本ワークショップでは、このような事例を紹介しつつ、機械学習の基本を講義形式で詳しく説明します（ハンズオンの実習形式ではありません）。

ワークショップの流れとしては、(1) 機械学習とは何か、(2) データの準備方法、(3) 具体的な仕組みと手順、(4) 分析結果の検証方法、(5) コーパス言語学における活用事例、を予定しています（諸般の事情で若干変更する場合があります）。なお、本ワークショップは初学者を対象としており、統計学などの事前知識を参加者に求めません。また、機械学習の手法を説明するにあたっては、可能な限り、分かりやすい言葉やイメージを使うことを心がけ、四則演算（足し算・引き算・掛け算・割り算）以外を使った数式は出しません。

【講演】

A Frontier in Learner Corpus Studies: For Better Understanding of L2 Learners

Shin’ichiro Ishikawa (Kobe University)

Various learner corpora have been developed to date and they have greatly contributed to improvement of L2 teaching. However, a more carefully designed corpus would be needed for a reliable contrastive interlanguage analysis. Thus, recent learner corpora have come to pay much more attention to controlling variety in the collected data.

The International Corpus Network of Asian Learners of English (ICNALE) is one of the largest learner corpora ever compiled. It includes more than 10,000 speeches and essays produced by L2 English learners in ten countries and regions in Asia as well as English native speakers. Its unique feature is that the topics are carefully controlled. All the participants are required to speak or write about two kinds of common topics: (A) It is important for college students to have a part-time job and (B) Smoking should be completely banned at all the restaurants in the country. Such a topic control is expected to lead to a greater reliability in

varied types of contrastive analyses (Ishikawa, 2013).

The ICNALE currently consists of four modules: Spoken Monologue (1,100 participants, 4,400 samples, 500,000 tokens), Spoken Dialogue (under construction), Written Essays (2,800 participants, 5,600 samples, 1,300,000 tokens), and Edited Essays (290 participants, 580 samples, 140,000 tokens).

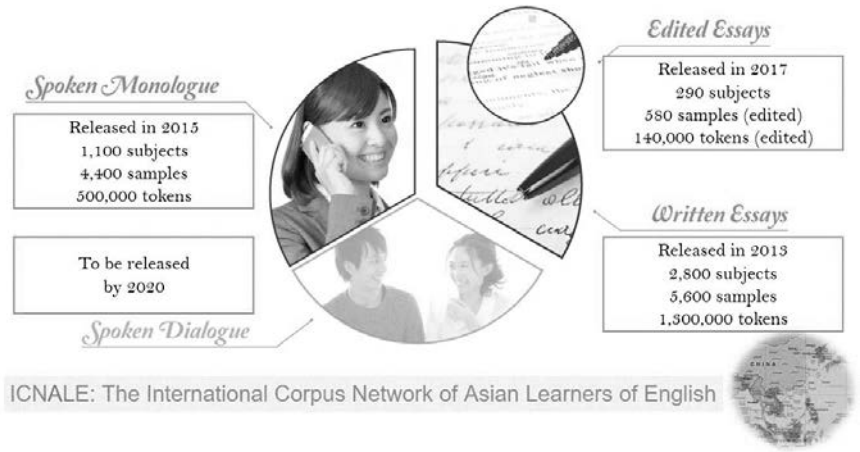


Fig. 1 The structure of the ICNALE

The ICNALE development team believes that comparing something comparable is a key to further development of learner corpus studies.

【シンポジウム】

話し言葉コーパスの構築と利用

司会：野口　ジュディー（神戸学院大学名誉教授）

コーパス言語学の多くの研究は書き言葉を扱っていますが、人間の自然言語は元々話し言葉からスタートしています。しかし、話し言葉は扱いにくいものです。コーパス作成には、発話者から許可を得ないとレコーディングすることさえ難しいでしょう。また、生の言葉を捉える環境がレコーディングに適していません。このようなハードルを越えての、時間のかかる書き起こしが必要になります。最終的にはコーパスを利用しやすいインタフェースを用意しなければなりません。そういった作業の苦労話を研究者たちは少なくともいくつかは抱えています。そういった研究者たちの努力によって、異なる 4 つコーパスが利用できるようになりました。英語を lingua franca として使用する、様々な母語を持つ学習者（大学生や大学院生を含む）からプロフェショナルまでの話し言葉がそれにあたります。この 4 つの異なるコーパスをどのように研究や教育に利用できるかを紹介いたします。各講師が構築したコーパス〔学習者の書き言葉・話し言葉（英語）〕ICNALE、学習者話し言葉日本語（テーマ別）I-JAS、話し言葉日英（理系プレゼン）JECPRESE、TED コーパス〕に関して、構築とその利用方法について話していただきます。

The ICNALE：中間言語対照分析の精緻化とアジアにおける学習者コーパス研究の発展を目指して

石川慎一郎（神戸大学）

The ICNALE（The International Corpus Network of Asian Learners of English）は、アジア圏 10 か国・地域において、英語学習者の L2 産出データを収集するプロジェクトで、すでに、Written Essays、Spoken Monologue、Edited Essays の 3 つのモジュール（計約 190 万語）が公開され、現在は、Edited Essays モジュールの拡充と、初のマルチモーダル版となる Spoken Dialogue モジュールの開発が進められています。The ICNALE の特徴は、プロンプトやテキストの長さなどが一定の範囲で統制されていることで、これにより、信頼性の高い国際比較研究が可能となります。The ICNALE は、ダウンロード版のほか、オンライン版があり、専用の検索用インタフェースが開発されています。

International corpus of Japanese as a second language：日本語学習者の言語研究と指導のために

迫田　久美子（広島大学・国立国語研究所）

International corpus of Japanese as a second language (I-JAS, http://lsaj.ninjal.ac.jp/)は、12 の言語を母語とする日本語学習者の発話と作文のコーパスです。JFL 学習者と JSL 学習者、さらに国内の学習者は教室環境と自然環境学習者のデータが収められ、さらに同じタスクを実施した日本語母語話者のデータも含まれています。完成は 2020 年春を予定しており、現在は 450 名のデータが公開されています。I-JAS の特徴としては、英語、中国語、韓国語、西語、独語、仏語、露語、タイ語、トルコ語、インドネシア語、ベトナム語、ハンガリー語の母語話者の複数のタスクのデータを所収しており、検索システムを備えていることが挙げられます。全員が同じテストを受けており、成績や背景事情も公開しています。

JECPRESE: JSL と EFL ユーザーのために

野口　ジュディー（神戸学院大学名誉教授）

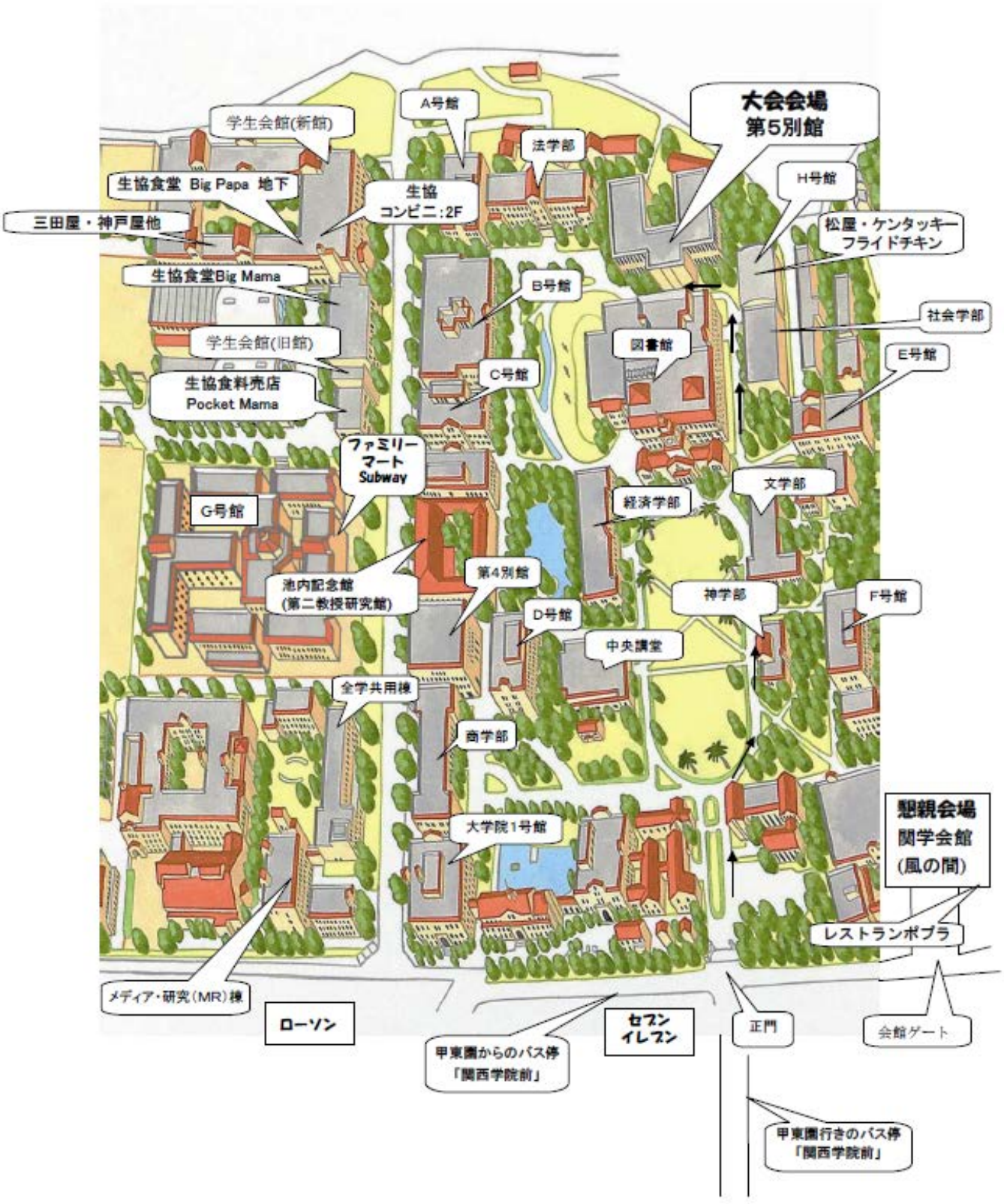
JECPRESE, the Japanese-English Corpus of Presentations in Science and Engineering (http://www. jecprese.sci.waseda.ac.jp/), は留学生のための専門日本語教育（JSL, Japanese as a second language）を支援する研究発表コーパスでスタートしました。日本の大学院生の日本語プレゼンテーションに加えて、アメリカの大学生や国際学会の英語プレゼンテーションも収められていて、EFL（English as a Foreign Language）の学生にも利用できるコーパスになりました。理工系のプレゼンテーションの特徴をわかりやすくするために、各発表を ESP（English for Specific Purposes）の手法であるジャンル分析に基づいてセクションやステップで検索できるようにしました。単語や表現の検索もできます。

TED Corpus Search Engine: TED Talks を研究と教育に活用するためのプラットフォーム

長谷部陽一郎（同志社大学）

TED Corpus Search Engine（http://yohasebe.com/tcse）は TED Talks の英語トランスクリプトを検索するためのウェブ・システムです。定期的な更新によってデータは増え続けていますが、現時点では約 2400 のプレゼンテーションから抽出されたテキスト（延べ語数は約 600 万語、異なり語数は約 8 万語）が収められています。語の品詞や基本形を指定可能な検索機能を実装しており、特定の構文や談話標識の実例を取得して研究や教育に役立てることができます。その他の特徴としては、得られた発話セグメントの動画をピンポイントで再生する機能や、日本語を含む 28 の言語による対訳データを表示／検索する機能が挙げられます。本発表では本システムを言語研究に活用する方法を中心に論じていきたいと思います。

≪会場案内図≫


大会会場
第5別館
A号館
学生会館(新館)
法学部
生協食堂 Big Papa 地下
生協
コンビニ:2F
三田屋・神戸屋他
H号館
松屋・ケンタッキー
フライドチキン
生協食堂Big Mama
B号館
社会学部
学生会館(旧館)
図書館
E号館
C号館
生協食料売店
Pocket Mama
ファミリー
マート
Subway
文学部
G号館
経済学部
池内記念館
(第二教授研究館)
第4別館
神学部
F号館
D号館
中央講堂
全学共用棟
商学部
懇親会場
関学会館
(風の間)
大学院1号館
レストランポプラ
メディア・研究(MR)棟
ローソン
セブン
イレブン
正門
会館ゲート
甲東園からのバス停
「関西学院前」
甲東園行きのバス停
「関西学院前」

《大会参加者へのご案内》

- 会場での無線 LAN 接続サービスの提供はございません。
- 会場には駐車場の用意はございませんので、公共の交通機関をご利用ください。第 1 日（9 月 30 日）は大学の行事があるため、周辺の駐車場も満車が予想されます。
- 大会（ワークショップを含む）への事前参加予約は不要です。ただし、懇親会（下記）への参加には予約が必要です。
- 第 1 日目、第 2 日目のワークショップの受付：会場の関西学院大学（西宮上ケ原キャンパス）第 5 別館 1 階正面で、9 時 30 分から行います。
- 大会受付：第 1 日（9 月 30 日）は第 5 別館 1 階正面で 11 時 30 分から行います。第 2 日（10 月 1 日）はワークショップ受付と併せて 9 時 30 分から行います。
- 構内での喫煙は指定の喫煙所にてお願いいたします。
  （喫煙場所：http://www.kwansei.ac.jp/students/attached/0000101114.pdf）
- 昼食について：第 1 日（9 月 30 日）は、生協食堂（BIG PAPA と臨時営業の BIG MAMA）、隣の H 号館内の松屋、学生会館新館内の三田屋、神戸屋他が利用できます。第 2 日（10 月 1 日）は生協食堂（BIG MAMA）とその手前にある飲食料売店（POCKET MAMA）のみ営業しています。
  （各店舗の場所・営業時間等の詳細：http://www.kwansei.ac.jp/students/students_003898.html）
- 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ので、お誘い合わせの上ご参加下さい（当日会費 2,000 円、2 日間共通）。懇親会（下記）へもぜひご参加下さい。大会当日に入会受付もいたします（年会費：一般 6,000 円、学生 3,000 円）。
- 大会第 1 日の学術プログラム終了後の懇親会は、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。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 43 回大会懇親会
  - 日時：9 月 30 日（土）18:15–20:30
  - 場所：関学会館
  - 会費：5,000 円

  ※懇親会参加ご希望の方は、参加申込 Web フォーム（https://goo.gl/forms/X4UQJUmuQQKHCJPm2）から 9 月 22 日（金）までにお申し込み下さい。

Memo

Memo