

英語コーパス学会第25回大会

ワークショップ《Excelによるコロケーションの統計的指標の計算》 10:15~11:45(9:45受付開始)
会場：恒心館2階 情報語学演習室1(724教室)
講師：高見 敏子(北海道大学)
定員：先着50名 参加費：会員無料・非会員1,000円(申し込みは電子メール・郵便で事務局まで)

日時 2005年4月23日(土)
会場 立命館大学衣笠キャンパス 恒心館2階727教室
(〒603-8577 京都市北区等持院北町56-1 <http://www.ritsumei.ac.jp/>)
[アクセスについては上記ホームページをご覧ください。土曜日は快速バスがありませんのでご注意ください]

受付開始 12:00(恒心館2階)

開 会 13:00

1. 会長挨拶
2. 総会
3. 事務局からの連絡

総司会 赤野 一郎(京都外国語大学)
中村 純作(立命館大学)

研究発表1 13:30~14:00

British National Corpus を用いた body 形および one 形不定代名詞の比較研究

紙谷 一彦(大阪大学大学院生)

研究発表2 14:00~14:30

take a picture of はイディオムか?—コーパスとウェブを利用して—

廣瀬 絵美(名古屋大学大学院生)

休憩 14:30~14:40

司 会 新田 香織(近畿大学)

研究発表3 14:40~15:10

policeman と police officer の使われ方—PC 表現の視点から—

畠山 利一(大阪国際大学)

実践報告 15:10~15:40

コーパスを活用する学習活動の提案

中條 清美(日本大学)

内山 将夫(情報通信研究機構)

休憩 15:40~16:00

シンポジウム 16:00~18:15 《コーパスと英語史研究—Helsinki Corpus 以後》

英語史研究における構文解析コーパス

コーパスと初期近代英語研究

コーパスと後期近代英語研究

司 会 西村 秀夫(姫路獨協大学)

講 師 塚本 聡(日本大学)

講 師 福元 広二(鳥取大学)

講 師 水野 和穂(広島修道大学)

閉会の辞

朝尾幸次郎(立命館大学)

《懇親会 時間：18:30~20:00 場所：レストランカラム(末川記念会館地下) 会費：4,000円》

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

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 E-mail: i_akano@kufs.ac.jp 郵便振替口座 00940-5-250586

URL: <http://muse.doshisha.ac.jp/JAECS/index.html>

- ◆ 大会当日、入会受付もいたしますので、お問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」としての参加も受け付けております(1,000円)。

英語コーパス学会 第 25 回大会資料

日時:2005 年 4 月 23 日(土)午後 1 時より(正午受付開始)

会場:立命館大学衣笠キャンパス 恒心館 2 階

(<http://www.ritsumei.ac.jp/>)

〒603-8577 京都市北区等持院北町 56-1

第 25 回大会プログラム

ワークショップ	10:15～11:45 (9:45 受付開始)	恒心館 2 階 情報語学演習室 1 (724 教室)
《Excel によるコロケーションの統計的指標の計算》		講師 高見 敏子 (北海道大学)

会 場 恒心館 2 階 727 教室
受付開始 12:00
開 会 13:00

会長挨拶
総会
事務局からの連絡

司会 赤野 一郎 (京都外国語大学)
中村 純作 (立命館大学)

研究発表 1 13:30～14:00

British National Corpus を用いた body 形および one 形不定代名詞の比較研究

紙谷 一彦 (大阪大学大学院生)

研究発表 2 14:00～14:30

take a picture of はイディオムか? コーパスとウェブを利用してー

廣瀬 絵美 (名古屋大学大学院生)

休憩 14:30～14:40

司会 新田 香織 (近畿大学)

研究発表 3 14:40～15:10

policeman と police officer の使われ方ーPC 表現の視点からー

畠山 利一 (大阪国際大学)

実践報告 15:10～15:40

コーパスを活用する学習活動の提案

中條 清美 (日本大学)

内山 将夫 (情報通信研究機構)

休憩 15:40～16:00

シンポジウム 16:00～18:15

《コーパスと英語史研究ーHelsinki Corpus 以後》

英語史研究における構文解析コーパス

司会 西村 秀夫 (姫路獨協大学)

コーパスと初期近代英語研究

講師 塚本 聡 (日本大学)

コーパスと後期近代英語研究

講師 福元 広二 (鳥取大学)

講師 水野 和穂 (広島修道大学)

閉会の辞

朝尾幸次郎 (立命館大学)

懇親会 18:30～20:00 《レストランカルム (末川記念会館地下) 会費 : 4,000 円》

発表要旨

【ワークショップ】

Excelによるコロケーションの統計的指標の計算

高見敏子(北海道大学)

“You know a word by the company it keeps.”という言葉に象徴されるように、コロケーション研究はコーパス言語学の諸分野の中でも特に多くの研究者が関心を寄せる分野です。そして近年は他の分野と同様にコロケーション研究においても、出現度数や相対頻度といった基本的計量データの利用から一歩進んで、MI や t-score といった統計的指標が広く使われるようになってきています。

ところで、これらの統計的指標に関してその計算式は文献などで知ることができても、実際にどうすればその値を自分で計算できるのか、という具体的な手順についての情報を見つけることはなかなか難しいのではないのでしょうか。また、すでに MI や t-score を使った研究をしている場合でも、実のところはコーパス解析ソフトがはじき出してくれた数字を使っているだけで、どのような計算によるものなのかということまでは知らないままになっているケースもあるのではないかと思います。

そこで今回のワークショップでは、コロケーション研究で最も広く用いられている MI と t-score の2つの計量的指標について、その計算式の成り立ちを知り、多くの人が使い慣れている表計算ソフト(Microsoft Excel)を使って、「度数情報さえあれば自分で計算できるようになる」ことを目標とした実習を行います。この方法を習得すれば、例えば関心のあるコロケーションだけを並べて MI と t-score の値を比較する、後から見つかったデータの不備を修正した値で再計算を行う、あるいは架空の値を様々に代入して結果の変動を観察するなど、コーパス解析ソフトでは実現が面倒だったり難しかったりするようなことも比較的簡単に行えるようになります。

ワークショップでは、主として Geoff Barnbrook, *Language and Computer* (1996, Edinburgh UP) の第5章, Uri Zernik (ed.), *Lexical Acquisition: Using On-line Resources to Build a Lexicon* (1991, Lawrence Erlbaum) の第6章, Jeremy Clear の “Re: Corpora: T-score in collocational analysis” (1999年12月12日付 Corpora メーリングリスト投稿記事: <http://torvald.aksis.uib.no/corpora/1999-4/0146.html>)などのデータを取り上げる予定です。(『英語コーパス言語学(改訂新版)』(2005, 研究社)に関連した内容を書きましたので、機会があればご一読ください。)

【研究発表】

British National Corpus を用いた body 形および one 形不定代名詞の比較研究

紙谷一彦(大阪大学大学院生)

anybody と anyone, somebody と someone, everybody と everyone および nobody と no one (no-one を含む)の各不定代名詞は, any-, some- などの限定詞形態素および -body, -one の名詞形態素からなる(Quirk et al. (1985))が, 限定詞形態素が同じで名詞形態素が異なる対(例: anybody/anyone)の意味あるいは機能が同じであるか異なるかについては, 先行研究で見解が分

かかれている。Jespersen (1914)や Quirk et al. (1985)は、body 形と one 形の実質的な相違はないとしているが、Bolinger(1976)は、one 形は“closeness to the speaker”や“narrowness of the range”について有標であると主張している。Svartvik & Lindquist (1997)は、コーパスを使用して Bolinger (1976)の主張を裏付けているが、コーパスの規模が小さい、共起語の調査では明瞭な結果が得られていない、という理由で、さらに調査を進める必要があると思われる。

本発表では、BNC をもとに統計学的分析手法も用いて、主に body 形と one 形の頻度の差異および共起語の差異について考察する。調査の結果、body 形と one 形は話し言葉と書き言葉で頻度がかかなり異なること、限定詞形態素が同じ場合、共起語は全般に類似しているが、頻度の大きく相違する共起語も存在することが確認できた。

take a picture of はイディオムか？—コーパスとウェブを利用して—

廣瀬絵美(名古屋大学大学院生)

本発表の目的は、take a picture of という連鎖がイディオムかどうかという問いにコーパスとウェブを利用して答えることである。生成文法理論の中には、take a picture of をイディオムだと考えれば、移動操作や解釈規則が関わる諸現象をうまく説明できるという考え方がある。しかし、この提案はそれらの諸現象に理論的な説明を与えやすくするために提出されたにすぎない。本当にこの連鎖がイディオムかどうかは、一般的にイディオムと言われる連鎖と比較検討し述べられるべきであるし、実際にイディオムだとしたら、イディオムの中にも様々な種類があるため、イディオムの中の位置付けについても調査する必要がある。

本発表では、コーパスとウェブを利用して、たとえば take advantage of のような、一般にイディオムと言われる動詞句と take a picture of を始めとするある類の動詞句を、2種類の統語的特性に関して比較する。2種類の統語的特性とは、具体的には、目的語の位置に現れる名詞句が無冠詞で現れたり、特定の動詞や前置詞と共起したりするといった、動詞句内の形式的な特性と、受動操作による名詞句内部からの抽出の可能性である。

この2つの観点において、一般にイディオムと言われる動詞句と take a picture of が同じように振舞うのか、あるいは異なるのかを、コーパスとウェブを利用して得られる証拠をもとに示す。また、コーパスで検出されない場合でもウェブを検索すると検出されることがしばしばあるが、ウェブ上の言語資料がどの程度信頼のおけるデータとなりうるかについても考察する。

PC 表現の使われ方

—policeman と police officer の使用を中心に—

畠山利一(大阪国際大学)

man という語を「(男女を問わず)人,人間」の意味で使うのはセクシズムに通じるので使用を避け、公正な(politically correct)表現(PC 表現)を用いるのがよいとされている。このような言語上の性差別廃止の動きは1960年代中頃からおこり、1970年代にはLakoff (1973) 'Language and Woman's Place'が現れるなど、活発に議論された。1980年代にはいると、PC 表現のためのマニュアルや語彙集が多く出版され、今では広く社会に行き渡っているような印象を受ける。

本発表では、主に policeman とその PC 表現の police officer を取り上げ、アメリカで発行されている主要な新聞・雑誌の中での使用実態を調査する。コーパスとして LexisNexis データベースを用いる。policeman の使用が減っていき、police officer の使用が増えていくという一般的な傾向はうかがえるが、その変化は一様ではない。次のようなタイプがある。(1)1970 年代は policeman が police officer よりも使用例が多いが、年を追うごとに police officer の使用が増え、1980 年代はじめに逆転をして、police officer が policeman よりも多く使われるようになり、現在は policeman の使用例は少数に過ぎないもの(例: *The Washington Post*)。 (2)年を追うごとに police officer の使用が増えるが 2000 年代になってはじめて police officer が policeman よりも多く使われるようになったもの(例: *Business Week*)。 (3)徐々に police officer の使用が増えていき、1990 年頃に policeman と police officer の使用がおよそ同数になり、その後もそのままのもの。(例: *Newsweek*) (4)1970 年代から 2004 年にいたるまで、終始 policeman が police officer よりも圧倒的に多く使われているもの(例: *The Economist (US edition)*)。

このようにメディアによって policeman と police officer についての扱いが異なる。policewoman など関連する語の使用状況もみながら、なぜこのような違いが生じるのか、その理由を考察したい。

【実践報告】

コーパスを活用する学習活動の提案

中條清美(日本大学生産工学部)

内山将夫(情報通信研究機構)

半数の大学が英語授業にコンピュータを導入している現在、コーパスと検索ツールがあれば、学習者自身に文法規則や語彙の意味・用法等を帰納的に学習させる「データ駆動型の発見学習」が可能になる。しかし実際には、日本では期待されるほど英語コーパスは授業で活用されていない。コーパスは学習材として「自然な英語」の魅力を持つ反面、読み取る英文の難易度が高く、日本人学習者の英語力レベルによっては語法の検証や発見に至らないことがある。本報告では、英語初級者を対象に、英語テキストデータに対応する日本語対訳テキストデータを併せ持つ「日英パラレルコーパス」を活用して英文理解の負荷を軽減し、CALL 形態の一般英語授業においてコーパス指導を試みた。

2004 年後期、TOEIC スコアの向上を目指す週 1 回の CALL 授業 3 クラスにおいて、90 分のうち 20~25 分間、「読売新聞/Daily Yomiuri」コーパスと検索ツールである ParaConc を使用した学習活動を 4~5 回実施した。学習者はパートナーと相談しながら教師が作成した「語彙指導用コースウェア」に沿って自律的にタスクをこなした。実践の結果、新聞英語という難易度の高い学習材にもかかわらず、日英パラレルコーパスと周到に用意されたタスクによって、初級学習者を対象としたデータ駆動型の発見学習が可能になり、英語に苦手意識を持つ初級学習者がコーパスを有効な学習手段であると好意的に受け入れたことが確認された。また発見学習を通して英語と日本語は一つでないということばの多義性に学習者自らが気づいたことも重要である。なお、検索ツールの操作の簡易化を図るため、日英パラレルコーパスを Web で検索できるサイトを情報通信機構より開設した。

【シンポジウム】

コーパスと英語史研究—Helsinki Corpus 以後

司会 西村秀夫(姫路獨協大学)

1991年に公開された Helsinki Corpus (750-1710年, 約160万語)は英語史研究のあり方に大きな影響を与え、今日では、コーパス言語学的手法を用いた研究が英語史研究の大きな柱となっている。

Helsinki Corpus の編者たちは公開後5年以内に拡大改訂版を出すと言ったが、今日に至るまでそれが公開されることはなかった。しかしながら、近年、史的コーパスに関して以下のような注目すべき動向が見られるようになっている。

- (1) Helsinki Corpus のファイルを元に、品詞標識・構文解析を施したコーパスの編纂、および語数の拡大 (YCOE, PPCME, PPCEME など)
- (2) 時代や地域を限定したコーパスや、特定のジャンルを対象を絞り込んだコーパスの編纂 (CEEC など)
- (3) Helsinki Corpus がカバーしていない後期近代英語期を対象としたコーパスの編纂 (CONCE など)
- (4) 現代英語の通時的な変化への関心 (Brown, LOB と Frown, FLOB の比較)
- (5) コーパスを利用した研究領域の拡大 (文法化, 史的語用論など)

本シンポジウムでは「Helsinki Corpus 以後」のコーパスを利用した英語史研究の方向性・可能性について、3人の講師による研究事例の報告を基に考えていく。時間の制約があるので、上記のうち、(1)(3)(5)を中心に扱うことになるが、(2)(4)についても可能な限りの情報提供を行う予定である。

英語史研究における構文解析コーパス

講師 塚本 聡(日本大学)

Helsinki Corpus 完成後、史的コーパスの分野ではいくつかの拡張が行われたが、古英語、中英語におけるその例として、The Penn-Helsinki Parsed Corpus of Middle English (PPCME)や The York-Toronto-Helsinki Parsed Corpus of Old English Prose (YCOE)などの Treebank の枠組みによる構文解析コーパスを挙げることができる。

構文解析コーパスの特徴は、語の品詞表示および構文の解析にある。従来のコーパスでは同綴異義語の区分が困難であったが、品詞区分がなされることで、より精密な検索が可能となる。また、構文に関する情報は、構文の変遷を研究する際に有益である。すなわち、質的な変化がもたらされたと言える。

さらに、今回取り上げるコーパスは収録語数にも特徴がある。元となった Helsinki Corpus より拡大され、規模の点で大きく異なっている。一例を挙げると、Helsinki Corpus と PPCME 共通に共通するテキストから収録された語数は、前者が約32万語であるのに対し、後者は約100万語にまで拡大されており、量の面でも変化している。

本発表では、この質と量の変化という観点から、1)品詞表示や構文情報の有無によって、コ

ーパスで可能な調査にどのような変化が生じるか、2) サンプルの量が異なることでどのような問題が生じるか、について、補文構造(節・不定詞句などの区分)を例に検討し、これらの構文解析コーパスの存在意義を考えたい。

コーパスと初期近代英語研究

講師 福元広二(鳥取大学)

初期近代英語研究でよく利用されるコーパスは、Helsinki Corpus の他に、OED2 CD-ROM, Corpus of Early English Correspondence (CEEC), Chadwyck-Healey Texts, Gutenberg Texts などであるが、今年になって、Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME)も公開され、コーパスを用いた初期近代英語研究もますます活発になると思われる。本発表では、まず最近の初期近代英語研究で利用されるコーパスについての動向を概観し、それぞれの研究目的に応じて各種コーパスを使い分ける必要があることを述べる。

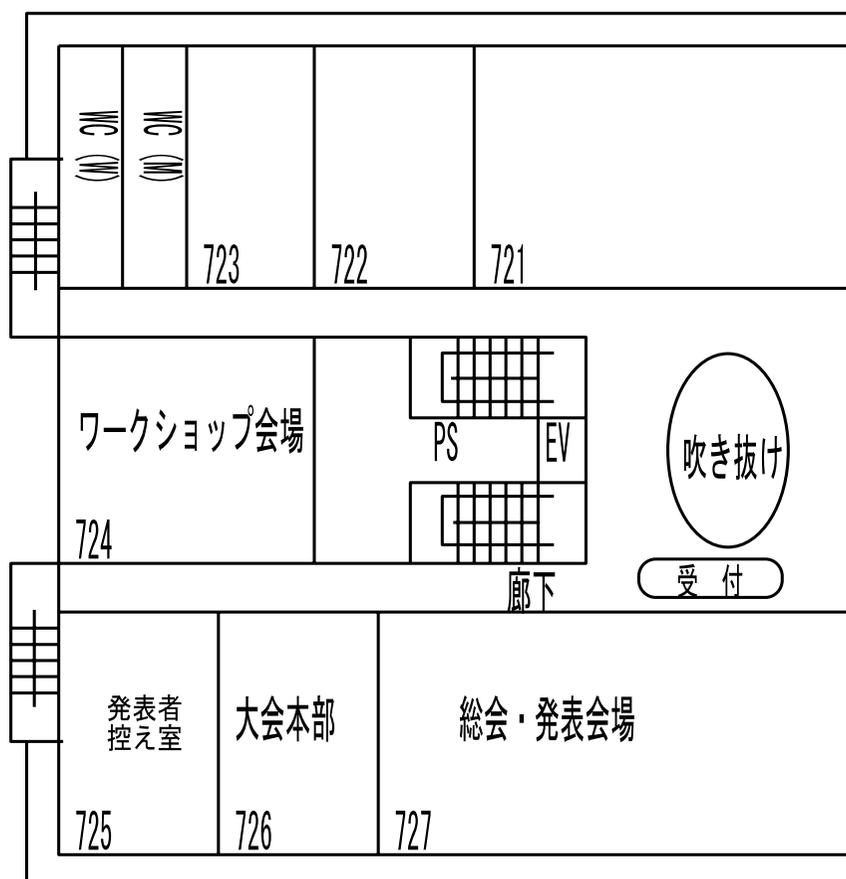
次に、近年盛んに行われるようになった文法化研究においてもコーパスを用いた研究の有効性が証明されているが、コーパスを利用した文法化の実践例として、現代英語において間投詞や強調詞のように用いられる *I say* や *I tell you* のような 1 人称単数人称代名詞と発話動詞とのコロケーションを取り上げる。これらのイディオムが英語史、特に初期近代英語期を通してどのように文法化、主観化されてきたのかを Shakespeare Corpus, Helsinki Corpus, Modern English Drama Corpus の 3 つのコーパスを利用して分析してみたい。そしてそれぞれのコーパスからどのようなことが言えるのかを検証したい。

コーパスと後期近代英語研究

講師 水野和穂(広島修道大学)

後期近代英語は、古英語、中英語、初期近代英語と比較すると大きな言語構造上の変化に乏しいため、一昔前までは積極的に研究が行われてきたとは言えなかった。しかし、*The Cambridge History of the English Language vol. IV* (1998)の出版以後状況は一転し、とりわけ社会歴史言語学的アプローチによる研究成果が数多く報告されている。後期近代英語のコーパス編纂に関しても、まだ決して十分とは言えないが、Helsinki Corpus と Brown, LOB, BNC 等現代英語コーパスとの時期を埋めるものがいくつか編纂、公開されてきている。

本発表では、まず近年の後期近代英語研究動向について概観する。次に、主な後期近代英語を射程としたコーパスを紹介する。具体的には、Century of Prose Corpus, A Corpus of Late Modern English Prose, A Corpus of late Eighteenth-century Prose, ARCHER Corpus 等である。(ARCHER Corpus については未だ公開に到っていないが、現時点での最新情報を報告できればと思っている。)そして、後期近代英語期において変化が生じた統語現象(指示代名詞と所有代名詞の共起、関係詞構造等)について、前述のコーパスを利用した調査結果を述べたい。現代英語が確立する直前の後期近代英語研究の側面をお伝えできれば幸いである。



恒心館 2 階の教室配置図

《大会参加者へのご案内》

- 車でのご来場はできません。
- ワークショップの受付は「恒心館」2階 724 教室前で午前9時45分から行います。
- 大会の受付は「恒心館」2階で正午から行います。
- 昼食につきましては、「存心館」地下生協食堂、「末川記念会館」地下レストランカラムが営業しています。「中川会館」地下生協購買部ではサンドイッチ、コンビニ弁当類が購入できます。
- 校内は分煙措置がとられています。指定場所での喫煙にご協力ください。
- 会員でない方も、「当日会員」として参加していただけます(1,000円)。

ワークショップ・総会・発表会場

懇親会場(地下1階)



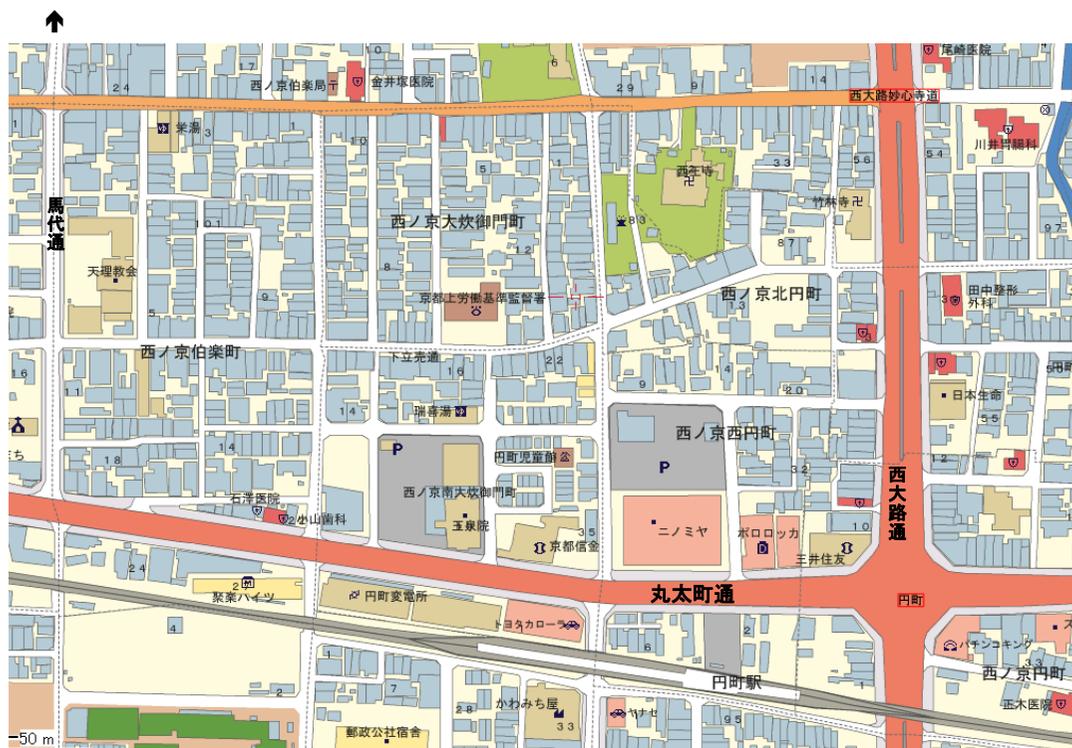
- | | | | |
|----------|-------------|------------------------|----------------------|
| ① 図書館 | ⑪ 啓明館 | ⑳ 末川記念会館 | ㉓ 由来の碑 |
| ② 存心館 | ⑫ 洋洋館 | ㉑ 第1体育館 | ㉔ 敬学館 |
| ③ 興学館 | ⑬ 恒心館 | ㉒ 第2体育館 | ㉕ 西側広場食堂
(仮称、建設中) |
| ④ 志学館 | ⑭ 諒友館 | ㉓ 学生会館 | |
| ⑤ 以学館 | ⑮ 修学館 | ㉔ アート・リサーチセンター | |
| ⑥ 研心館 | ⑯ ランゲージセンター | ㉕ アカデミア立命21 | |
| ⑦ 学而館 | ⑰ 尽心館 | ㉖ (株)クレオテック | |
| ⑧ 有心館 | ⑱ 明学館 | ㉗ いこいの広場 | |
| ⑨ 西園寺記念館 | ㉒ 尚学館 | ㉘ 創思館(ヒューマン・リサーチ・センター) | |
| ⑩ 清心館 | ㉓ 中川会館 | ㉙ 中央広場 | |

立命館大学キャンパスマップ

- 会場の「恒心館」はキャンパスの一番西の奥に位置します。
- 正門からは③〇 「中央広場」を抜けて右方向へ徒歩8分、東門からはほぼまっすぐ徒歩5分の位置にあります。

 JR 京都駅に到着される方は、山陰本線(嵯峨野線普通列車は32番ホーム)に乗り換え、円町(京都駅から3つ目;前頁地図参照)で下車(9-10分)し、そこからタクシーを利用するのが早くて便利です。円町駅前の「丸太町通」で西行きの、もしくは「西大路通」で北行きのタクシーに乗り、小松原北の立命館大学東門付近(小松原郵便局前)で降りる(東門まで徒歩3分)か、立命館大学正門と指示してください(前者でワンメーターか1回、後者で2,3回変わる程度の距離)。JR 京都駅からタクシーに乗ると2,500円前後。前頁の15番と204,205番のバスも利用できます。「西ノ京円町」のバス停は「西大路通」にあります。

至小松原北町



JR 円町駅周辺

2005年3月7日 発行
 編集・発行 英語コーパス学会
 代表者 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 赤野一郎研究室内
 TEL: 075-322-6103 FAX: 075-322-6246
 E-mail: i_akano@kufs.ac.jp
 URL: <http://muse.doshisha.ac.jp/JAECS/index.html>
