

英語コーパス学会第23回大会

ワークショップ 《自然言語処理技術を活用したコーパスツール》 10:30 - 12:00 7号館3階733教室
講師 松本裕治(奈良先端科学技術大学)他
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 2004年4月24日(土)
会場 京都外国語大学 1号館7階 多目的ホール
(〒615-8558 京都市右京区西院笠目町6 : 075-322-6054 <http://www.kufs.ac.jp/>)
[JR 京都駅から市バスで約30分、阪急京都線西院駅から徒歩15分または市バスで5分]

受付開始 12:00

開 会 13:00

1. 会長挨拶
2. 会場校挨拶
3. 総会
4. その他

総合司会 赤野 一郎(京都外国語大学)
中村 純作(立命館大学)
堀川 徹志(京都外国語大学学長)

研究発表 13:35 - 15:20

- 司 会 中條 清美(日本大学)
1. N-gram モデルを利用した日本人英語学習者の発達指標特定を試み
品詞の共起関係による構造複雑性の分析の可能性
木村 恵(東京学芸大学連合大学院生)

- 司 会 畠山 利一(大阪国際大学)
2. Why do you 'make a decision' instead of 'decide'?
ロジスティック回帰分析に基づく英語軽動詞に関するケーススタディ
見目 卓之(元大東文化大学大学院生)

- 司 会 佐藤 恭子(プール学院大学)
3. Corpus of Spoken Professional American-English における文体とジェンダーにかかわる諸問題
接続詞の分析から
家人 葉子(京都大学)
家口美智子(摂南大学)
岡部 浩子(神戸市外国語大学大学院生)

休 憩 15:20 - 15:40

シンポジウム 15:40 - 18:00 《コーパスとコロケーション》

コロケーション研究概観 司 会・講 師 堀 正広(熊本学園大学)
コロケーションと文法・語法 講 師 東海林宏司(茨城リソテック短期大学)
コロケーションと文体 講 師 堀 正広(熊本学園大学)
コロケーションと英語教育 講 師 Joseph Tomei(熊本学園大学)

閉会の辞 石川 保茂(京都外国語大学)

《懇親会 18:15 - 20:00 於: インターナショナルホール(9号館7階) 会費 5,000円(一般) 4,000円(学生)》

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」としての参加も受け付けております(1,000円)。

英語コーパス学会 第 23 回大会資料

日時:2004 年 4 月 24 日(土)午後 1 時より(正午受付開始)

会場:京都外国語大学 1号館 7 階 多目的ホール(四条通に面した建物)

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

〒615-8558 京都市右京区西院笠目町 6

TEL: 075-322-6054

ワークショップ会場 7 号館 3 階 733 教室

第 23 回大会プログラム

ワークショップ 10:30-12:00 7号館 3階(733 教室)

《自然言語処理技術を活用したコーパスツール》講師 松本裕治(奈良先端科学技術大学) 他

会場 1号館 7階 多目的ホール
受付開始 12:00
開会 13:00

総合司会 赤野 一郎(京都外国語大学)

1. 会長挨拶 中村 純作(立命館大学)
2. 開催校代表挨拶 堀川 徹志(京都外国語大学学長)
3. 総会
4. その他

研究発表 13:35 ~ 15:20

司会 中條 清美(日本大学)

1. N-gram モデルを利用した日本人英語学習者の発達指標特定の試み
一品詞の共起関係による構造複雑性の分析の可能性—

木村 恵(東京学芸大学連合大学院生)

司会 畠山 利一(大阪国際大学)

2. Why do you ‘make a decision’ instead of ‘decide’?
—ロジスティック回帰分析に基づく英語軽動詞に関するケーススタディー—

見目 卓之(元大東文化大学大学院生)

司会 佐藤 恭子(プール学院大学)

3. Corpus of Spoken Professional American-English における文体とジェンダーにかかわる諸問題
—接続詞の分析から—

家入 葉子(京都大学)

家口美智子(摂南大学)

岡部 浩子(神戸市外国語大学大学院生)

休憩 15:20 ~ 15:40

シンポジウム 15:40 ~ 18:00

《コーパスとコロケーション》

コロケーション研究概観

司会・講師 堀 正広(熊本学園大学)

コロケーションと文法・語法

講師 東海林宏司(茨城リソト教学園ソソ短期大学)

コロケーションと文体

講師 堀 正広(熊本学園大学)

コロケーションと英語教育

講師 Joseph Tomei(熊本学園大学)

閉会の辞

石川 保茂(京都外国語大学)

懇親会 18:15 ~ 20:00 (9号館 7階 インターナショナルホール)

発表要旨

【ワークショップ】

自然言語処理技術を活用したコーパスツール

松本 裕治 (奈良先端科学技術大学)
橋本喜代太 (大阪女子大学)
投野由紀夫 (明海大学)
浅原 正幸 (奈良先端科学技術大学)
森田 敏生 (総和技研)

現在、3年計画で、先端的な自然言語処理技術を援用したコーパス研究に必要な諸ツールの開発を進めている。ツール類は Windows 上で動作するものであり、KWIC ツールを中核として、コーパスそのものの整備に必要な機能として形態素解析技術を利用した自動タグ付け、それに必要な辞書整備などの機能も盛り込む予定である。また、近年の研究動向に合わせて、MI スコア等の統計数値算出もサポートする。まだ開発途上であるが、このワークショップでは、本ツール類の背景となる基礎技術を解説しつつ、特に自分で編纂しているコーパスデータをより有効に活用するためのツールとしての使い方を紹介していきたい。

【研究発表】

N-gram モデルを利用した日本人英語学習者の発達指標特定を試み： —品詞の共起関係による構造複雑性の分析の可能性—

木村 恵 (東京学芸大学連合大学院生)

本研究は、異なる習得段階にいる学習者の発話を分析・比較することによって、第2言語習得のプロセスを解明するための手がかりである「発達指標」を特定することを目的とした一連の研究の一つである。日本人英語学習者のインタビュー・テストによる話し言葉を収集し編纂された SST (Standard Speaking Test) コーパスのトライアル・バージョン 100 ファイルを用い、習得段階の異なる学習者グループ間にどのような「違い」が見られるかを探る。第2言語習得の発達指標には、1) 流暢さ、2) 正確さ、3) 文法的複雑性、4) 語彙使用などが関わると考えられるが (cf. Wolfe-Quintero et al. 1998)、その中でも特に「文構造の複雑性」を扱う。

本研究では N-gram モデルによる品詞の共起関係、特に連続する3つの品詞が出現した回数 (trigram の頻度) が、異なる習得段階にいる学習者グループを判別する手段として有効であることを検証する。コーパス・データへの品詞タグ付与には CLAWS (タグセットは C7) を用い、N-gram 統計を取るためのプログラムである ngram を利用してグループごとの各品詞連鎖の頻度を算出した。得られた頻度情報を基に、1) 習得が上がるにつれ出現する trigram、2) 習得段階が上がるにつれ減少する trigram、さらに 3) 中位層の学習者にのみ特徴的に用いられる trigram のそれぞれを見つけ出し、詳細な品詞タグ情報や使用されている語彙などを関連させ考察を加える。

Why do you 'make a decision' instead of 'decide'?

—ロジスティック回帰分析に基づく英語軽動詞に関するケーススタディー—

見目 卓之 (元大東文化大学大学院生)

英語には軽動詞と呼ばれる動詞が存在する。これは、動詞派生の名詞をその目的語とする動詞のことであり、代表例として *do, give, have, make, take* などが挙げられる。軽動詞句 (e.g. *have a drink, take a breath*) において動詞は、統語的には必要であるが、句全体が表す意味には貢献していない。つまり意味的に「軽い」のである。また軽動詞は、Poutsma (1926) や Jespersen (1942) による文法書でも扱われているが、これまで「なぜ意味的に等価の 1 語動詞の代わりに軽動詞句が使われるのか」という点に関する分析はなされなかった。

本研究は、軽動詞句 'make a decision' に焦点を当て、それと意味的に等価の 1 語動詞 'decide' との間に判別要因があると仮定し、英語コーパスに基づいて、その要因を計量的観点から明らかにすることを目的とする。先行研究に基づいて、'make a decision' と 'decide' の判別要因となり得る 13 の因子群 (Verb-finiteness, aspect, tense, mood, modal, polarity; Subject-type, semantic weight, animacy; Complement-type, length; Register; Modifier) を設定し、ロジスティック回帰分析を行なった。その結果として生成された確率モデルから、「否定語の有無」、「補文の種類」、「修飾語の有無」が両者の判別要因となることが明らかになった。本発表では、各判別要因に対して詳細な分析を行ない、「make a decision' と 'decide' との使い分けに影響を与える母語話者の直感を検証していく。

Corpus of Spoken Professional American-English における文体とジェンダーにかかわる諸問題

—接続詞の分析から—

家入 葉子 (京都大学)

家口美智子 (摂南大学)

岡部 浩子 (神戸市外国語大学大学院生)

Corpus of Spoken Professional American-English を使ってプロフェッショナル・スピーカーの英語における文体とジェンダーの問題を調査するプロジェクトの一環として、今回の発表では接続詞を扱う。分析の対象とするのは、当該のコーパスに比較的頻繁に出現する *and, but, or, as, if, because, when* の 7 つの接続詞である。分析にあたっては、これまで同様、発話者の性別がわかっている部分を取り出したものを利用する。ただし、今回のテーマである接続詞については細かい頻度を扱うので、さらに、このコーパスから 15,000 語を目途に小さなコーパスをできるだけ多く切り出したもの (小分けしたもの) を分析の対象とした。分析の結果、文体やジェンダーとの関連で、次のような点が明らかになった。

コーパスには、(1) *White House press conferences*, (2) *Faculty meetings*, (3) *Mathematics committee meetings*, (4) *Committee meetings on Reading* の異なる場面のファイルが含まれているが、この 4 者において、明確な文体上の違いがある。ジェンダーによる言語の違いは、場面の違いによる差ほど大きくはないが、特に (3) (4) において、注意すべき差異が見られる。全体として、双方向的なコミュニケーションの形態を取る会議形式の (3) (4) において接続詞の使用が多く、特に *or, if, because, when* がこれらのファイルを特徴付けている。一方、*White House* においては、*as* が多いのが特徴的である。*Faculty meetings* は、場面としては会議形式であるものの、文体的には *White House* に近く、全体としては報告型の英語の特徴を示している。また、男女の差が比較的大きく見られる (3) (4) においても、男性の方が報告型の英語に傾いている。一方、女性は接続詞を多用し、特に発話のつながりの部分にこの傾向が強く見られる。

【シンポジウム】

コーパスとコロケーション

司会 堀 正広 (熊本学園大学)

近年、コロケーション研究はこれまで以上に注目され、様々な関連領域の中でその重要性が指摘され、さらに理論的にも深められている。たとえば、Hunston & Francis (1999)は language patterning の考えに基づいて collocational patterns を文法と意味の面から論じている。Hunston (2001) では英語教育における collocation の重要性を説き、Lewis (ed.) (2001) は collocation を重視する lexical approach と実際の現場での collocation の教授法を提示している。また、Singleton (2000) は辞書編纂における lexical partnerships としての collocation の記述の重要性を強調している。そのような考えに沿って『ウィズダム英和辞典』(三省堂、2003 年)は、他の学習辞書に先駆けて使用頻度の高い collocation の表示を積極的に行った。文学作品の文体研究に関しては、Hori (2004)は Charles Dickens の言語文体を collocation の面から分析し、新たな文体分析の方法を提示している。さらに、Stubbs (1996)はコーパスによって異なる keywords の意味の違いを通して collocation 研究を文化の問題にまで発展させている。理論の面においても深められ、collocation を collocate、semantic prosodies、colligation の三方面から区別して分析することが提唱され始めている。

このような最近の collocation 研究の動向を踏まえて、各講師が関心を持っている collocation 研究の一端を発表していただき、フローと共に collocation 研究の重要性を認識し、研究のさらなる可能性を探ってきたい。

コロケーション研究概観

講師 堀 正広 (熊本学園大学)

“The habitual juxtaposition or association of a particular word with other particular word”の意味での collocation の研究は OED の collocation の初例が示すように 1950 年代以前に既に行われている。たとえば、Harold Palmer (1938)では英語を外国語として学ぶ学習者を念頭に基本語の collocation の例が挙げられている。しかし、collocation が音声、語彙、文法の研究と同じように記述言語学においては重要な研究領域であることを指摘し、collocation 研究を唱道したのはロンドン学派の J. R. Firth であった。

その collocation 研究の歴史を 1950 年代の J.R. Firth から概観する。1960 年代は M.A.K. Halliday, John Sinclair, Angus McIntosh によって collocation 研究の基本的な理論構築が行われた。1970 年代、Sinclair はコンピュータ利用を前提とした方法論を模索し、Sinclair に批判的な Sidney Greenbaum は native informant test を collocation 研究に導入した。1980 年代には、特に辞書作成の面で COBUILD project の目覚ましい成果がみられた。1990 年以降は Bill Louw, Michael Stubbs, Susan Hunston に見られるコロケーション研究の広がりや理論面での進展 (e.g. semantic prosody, colligation など)が見られる。辞書作成や英語教育においても collocation を教えることの重要性が強調され始めた (M. Lewis: 2001)。文学作品の文体研究においても作家や作品における collocational style 研究の重要性が指摘され始めた (Hori: 2004)。

このように、今後もさらにコロケーション研究は文法・語法の問題としてだけでなく、英語教育、文化論、英語史、辞書学、文体論などの領域でと大いに研究されて行くであろう。

コロケーションと文法・語法

講師 東海林 宏司 (茨城キリスト教学園シオン短期大学)

本発表では、いわゆる *non-literary language* の中でも、Web 上に公開されているメディアの英語 (BBC と CNN) に焦点を絞り、それをコーパスとして用いた場合に得られるデータを素材にして、コロケーションと文法・語法の問題について考察する。コーパスから得られるデータが有益なのは、数量的 (統計的) な分析、すなわち客観的な文法・語法研究ができることにあるのは言うまでもない。しかしながら、そこから一歩進めて質的 (意味的) な分析を加えることが、ある語の言わば「性格」、あるいは「親戚関係」を明らかにする手がかりとなる。

いわゆる *metaphor* が決して詩などの *literary language* の専売特許ではなく、日常言語の中に蔓延していることを指摘したのは Lakoff & Johnson (1980) であった。コーパスから得られるコロケーションのデータを分析すると、その指摘が裏付けられる。例えば本来物理的な位置関係 (高さ、深さなど) を表す語が、*metaphor* を伴って用いられる例は、特に強調副詞のコロケーションに多く見受けられることが、「形容詞＋名詞」のコロケーションと「副詞＋形容詞」のコロケーションを比較してみると浮かび上がってくる。また、*awfully sorry*, *dreadfully sorry*, *terribly sorry* と「恐縮」、*deeply sorry* と「深謝」などを比較してみれば、日常言語における *metaphor* は、複数言語において同じパターンが用いられる場合もあることがわかる。その一方で、*terribly important*, *terribly happy* などの例を見ると、*terribly* 本来の意味が薄れつつある度合いは高まっているのが見て取れる。外国語として英語のコロケーションを見る際に注意したい点である。

コロケーションと文体

講師 堀 正広 (熊本学園大学)

Firth は “*Modes of Meaning*” (1957) の中で記述言語学における意味の問題として *collocation* 研究の重要性を強調したが、同時に文体研究の重要な要素としての *collocation* 研究をも主張した。実際に Firth が論文の中で扱った用例はほとんどが文学作品で、Edward Lear の *limericks* (五行俗謡) である Gorboduc、Blake の *King Edward the Third*、Swinburne’s の詩などから取られている。このように Firth は文体研究の重要な柱として *collocation* 研究を唱道したが、その後の *collocation* 研究は Sinclair に代表されるようにもっぱら *non-literary language* の *habitual collocation* 研究に向けられ、文学作品の言語・文体研究における *collocation* はほとんど行われなかった。

本発表では、文学作品の言語・文体研究における *collocation* の重要性を強調し、コーパスを用いた *collocation* 研究の実践例を示し、*non-literary language* の *collocation* とは異なる文学作品の *collocation* を指摘する。さらに、*non-literary language* においては軽視されるが、文学の言語においては重要な低頻度の *collocation* の重要性と文学言語の創造性を指摘したい。

扱うコーパスは 18 世紀、19 世紀の主要な英国の小説である。作家の *collocational style* の例として 19 世紀の英国の小説家 Charles Dickens の *collocation* の特徴を 18、19 世紀の他の小説家の *collocational style* と比較しながら明らかにしていきたい。

コロケーションと英語教育

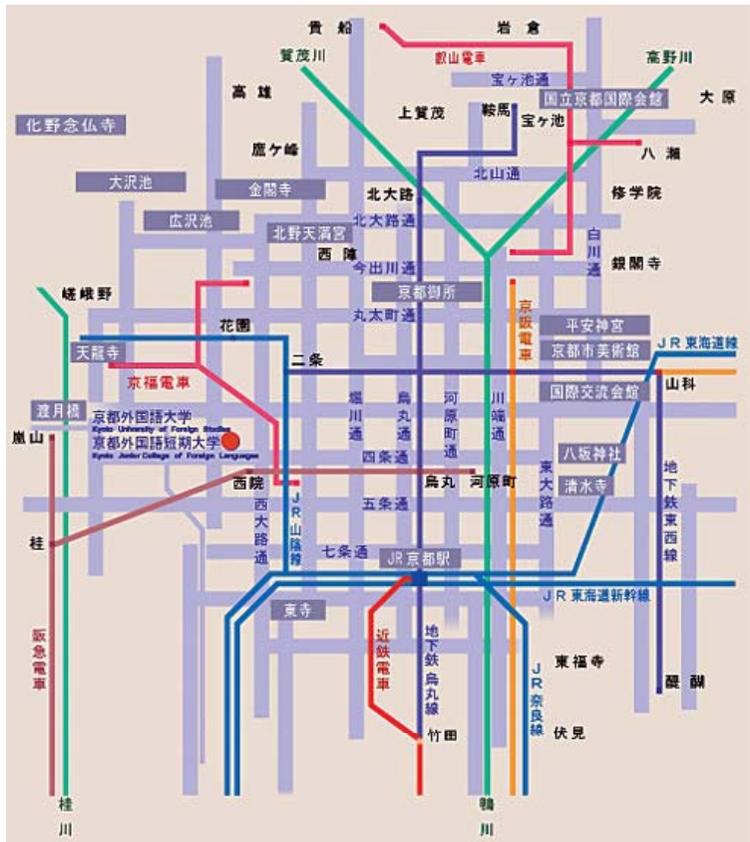
講師 Joseph Tomei (熊本学園大学)

Taking advantage of the revolution in personal computing, corpus linguistics has moved from a small subfield of linguistics to one which offers new and novel insights into the nature of language and grammar. The application of corpus linguistics to language teaching has been almost totally confined to use of traditional corpora, either to guide teachers (the COBUILD approach) or guide students (The Data Driven Learning (DDL) approach), especially through the notion of a concordance. In this presentation, a new but natural extension of teaching through corpora is discussed: the idea of using student generated corpora in order to permit students to discover patterns in their own language use. Previous uses of learner corpora have focused on a desire to make generalizations across student output and to compare that output to 'native' corpora. While this is undoubtedly valuable, equally valuable insights can be obtained from individual learner corpora with the added benefit of increased student interest and participation. In the first part of the presentation, I will discuss how a student corpus can be easily generated using My SQL and PHP and in the second part of the presentation, utilizing a liberal definition of collocation, I will present a sampling of the collocational patterns found in the student data and suggest possible classroom interventions. These collocational patterns provide an entry into discussing particular problems of more advanced Japanese students, including reference tracking, logical and rhetorical structure, grammatical complexity and vocabulary choice.

References

- Firth, J. R. (1957) *Papers in Linguistics, 1934-51*. London: OUP.
- Hori, Masahiro (2004) *Investigating Dickens' Style: A Collocational Analysis* Basingstoke: Palgrave, Macmillan.
- Hunston, Susan (2001) *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Hunston, Susan and Gill Francis (1999) *Pattern Grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- 井上永幸・赤野一郎編『ウィズダム英和辞典』(三省堂、2003年)
- Lakoff, George and Mark Johnson (1980) *Metaphors We Live By*. Chicago: University of Chicago Press.
- Lewis, Michael (ed.) (2001) *Teaching Collocation: Further Developments in the Lexical Approach*. Hove: Language Teaching Publications.
- Louw, Bill (1993) 'Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies,' in M. Baker et al. (eds.) *Text and Technology: In Honour of John Sinclair*. Amsterdam: John Benjamin, 157-76.
- Palmer, Harold (1938) *A Grammar of English Words* London: Longmans.
- Singleton, David (2000) *Language and the Lexicon*. London: Arnold.
- Stubbs Michael (1996) *Text and Corpus Analysis*. Oxford: Blackwell.

会場へのアクセス



① JR「京都」駅より

- 烏丸口 (のりば D3) より市バス 28 に乗車、「京都外大前」で下車
- 八条口 (のりば 2) より市バス 71 に乗車、「京都外大前」で下車 (ともに所要乗車時間 約 30 分)
- 地下鉄「京都」より「四条」で下車 (所要乗車時間約 4 分)、③へ接続



② 阪急京都線「梅田」駅より

- 特急に乗車、「桂」で急行・普通に乗換、「西院」下車 (所要乗車時間約 40 分)、④へ接続

③ 地下鉄「四条」駅より

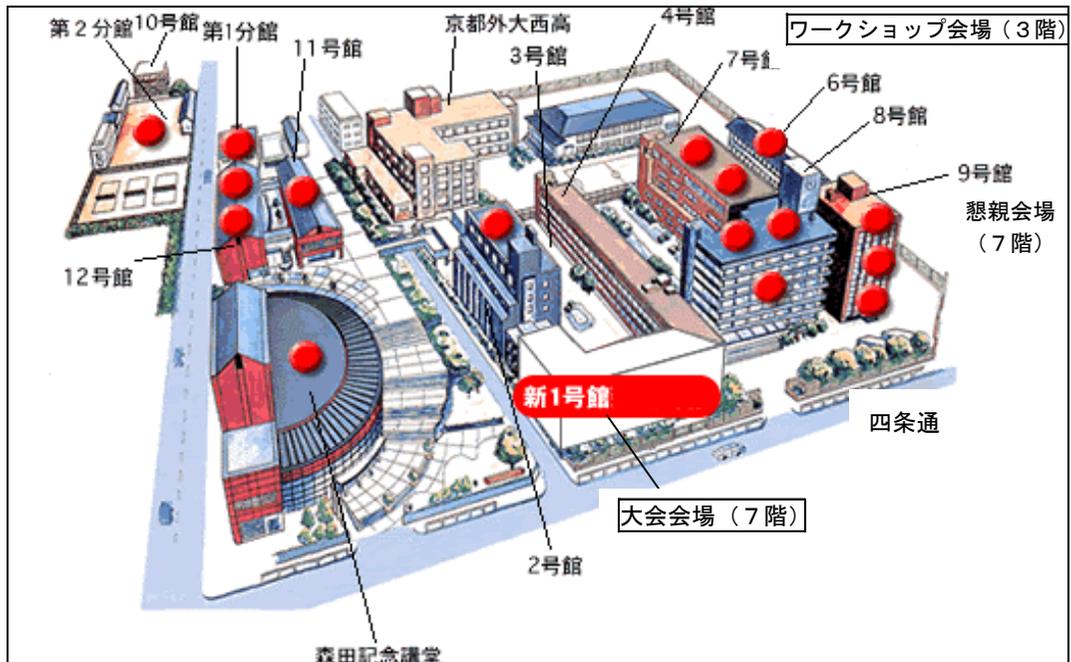
- 「四条烏丸」より市バス 3, 29 に乗車、「京都外大前」で下車 (所要乗車時間 約 15 分)
- 阪急「烏丸」より「西院」下車 (所要乗車時間 約 4 分)、④へ接続



④ 阪急電車京都線「西院」駅より

- 徒歩西 (左) 方向へ約 15 分
- 西大路四条 (西院) より市バス 3, 28, 29, 67, 69, 71 に乗車、「京都外大前」で下車 (所要乗車時間 約 5 分)

キャンパスマップ



《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 733 教室前で午前 10 時から行います。
- 大会の受付は 1 号館 7 階ホール前で正午から行います。
- 大会当日の受付は混雑が予想されますので、2004 年度会費(一般 5000 円、学生 3000 円)はできれば郵便振替で 4 月 15 日までにお納めください

昼食のご案内

- 1 号館地下にカフェレストラン(軽食と飲料)、12 号館に学生食堂があります。
- 9 号館1階の「カフェ太郎」でも食事がとれます。
- 大学の向かいには、ラーメン専門店(2 軒)、すき家、コンビニ、モスバーガーがあります。
- 正門を出て左(東)方向へ 3 分のところに、「ジャスコ」があり、3 階は食堂街になっています。
- 阪急西院駅周辺にも飲食店があります。

その他

購買部 11 号館 1 階南側
書籍部 (丸善) 11 号館 1 階北側

2004年3月 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>
