

東京外国語大学21世紀COEプログラム
言語運用を基盤とする言語情報学拠点

The 21st Century COE Program "Usage-Based Linguistic Informatics"
2005.10

項目またはページ番号をクリックで
各ページへ移動

C O N T E N T S

はじめに	4
21世紀COE 言語運用を基盤とする言語情報学拠点	6
言語情報学班	7
TUFS言語モジュール (TUFS Language Modules)	8
言語学班	16
言語教育学班	20
情報工学班	24
国際会議・国内会議・ワークショップ	28
出版物	33

はじめに

池端雪浦 (東京外国語大学学長)

2002年度から開始された文部科学省の「21世紀COEプログラム」は、我が国の大学に、世界最高水準の研究教育拠点 (Center of Excellence) を学問分野毎に形成し、研究水準のいっそうの向上と世界をリードする創造的な人材の育成をめざしています。本学は、「人文科学」と「学際・複合・新領域」の2つの学問分野にそれぞれ1件の申請を行い、そのいずれもが採択されるというすばらしい結果を得ました。本学大学院地域文化研究科の、個性ある研究教育のポテンシャルが高く評価されたことを嬉しく思います。本学ではヨーロッパ、南北アメリカ、オセアニア、アジア、アフリカと世界のほぼすべての地域にわたって言語学、文学、歴史学、哲学・思想、文化人類学、社会学、政治学、経済学などさまざまな学問分野のすぐれた専門家が協働して教育と研究にあたっています。

こうした地域研究や文化研究を行うにあたっての基盤となるのが、外国語能力です。50にのぼる言語と世界諸地域の文化・社会について教育研究を行っている本学は、言語と専門分野のダブルメジャー制の下で、高度な言語運用能力と、世界諸地域の文化と社会について深い知識を身につけて、異文化間の相互理解と地球社会における共生の実現に貢献できる人材の養成に努めてきました。2000年秋に移転を実現した府中キャンパスは情報環境が整備されており、インターネットや学内LANはもとより、パソコン台数やそれらを用いた情報リテラシー教育の水準は、国内の文科系大学としてはトップレベルにあると自負しております。このような素晴らしい情報インフラに支えられ本学は本年度よりe-learningシステムを利用した先進的な言語教育を実施しています。

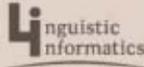
今回、21世紀COE拠点に採択された「言語運用を基盤とする言語情報学拠点」は、本学の将来計画の上に構想されたプロジェクトであります。推進メンバーの方々が精力的にプロジェクトに取り組み、実りある成果をあげ、COE拠点から次世代のわが国の外国語教育を担うべき人材が多数輩出されることを願ってやみません。21世紀COEプログラムの成功のために、本学の叡智を結集し、大学全体として協力してゆく所存です。

2005年9月

開始年度	2002年度
学問分野	人文科学
拠点のプログラム名称	言語運用を基盤とする言語情報学拠点
専攻等名	東京外国語大学地域文化研究科地域文化専攻

事業推進担当者

氏名	専門	役割分担
川口裕司	フランス語学・トルコ語学	拠点リーダー、言語情報学統括
在間進	ドイツ語学	言語教育学統括、言語情報学
富盛伸夫	理論言語学	言語学
高垣敏博	スペイン語学	言語学統括
敦賀陽一郎	フランス語学	言語学
亀山郁夫	ロシア文学	言語教育学
水林章	フランス文学・歴史学	言語教育学
野間秀樹	朝鮮語学	言語学
芝野耕司	情報学	情報工学統括・言語情報学
峰岸真琴	言語学	情報工学・言語情報学
宇佐美まゆみ	言語社会心理学・日本語教育学	言語教育学



21世紀COEプログラム 言語運用を基盤とする言語情報学拠点

[ホーム](#) | [サイトマップ](#) | [検索](#) | [学内連絡](#) | [English](#)

- ◆ 研究の概要
- ◆ 研究組織
- ◆ 国際会議・研究会
- ◆ TUPS言語モジュール
- ◆ TUPS e-learning system
- ◆ 多言語コーパス
- ◆ 出版物
- ◆ 研究成果一覧
- ◆ 若手研究者育成
- ◆ 学内連絡

言語運用を基盤とする言語情報学拠点とは

外国語教育に革新をもたらすような優れた外国語教材を開発しインターネットを通じて発信する。これがプロジェクトの目標である。そのために、これまで別々の道を歩んできた言語学、言語教育学、情報工学をまとめ言語情報学という新たな学問領域を創成する。3つの学問分野を連動させることで、従来の外国語教育の高度化、高率化、先端化が図られる。

PDF File

COE概要
パンフレット

最新ニュース [以前のニュースはこちら](#)

- 2005/7/1 [通言語文法モジュール](#)が公開されました。
- 2005/6/30 宇佐美まゆみ(監修)『BTSによる多言語話し言葉コーパス - 日本語会話』が刊行されました。
- 2005/6/2 サイトをリニューアルしました。
- 2005/4/19 雑誌『ふらんす』2005年4月号の記事
※この記事は白水社の許諾を得て転載しています。

21世紀COE 言語運用を基盤とする言語情報学拠点

Center of Usage-Based Linguistic Informatics (UBLI)

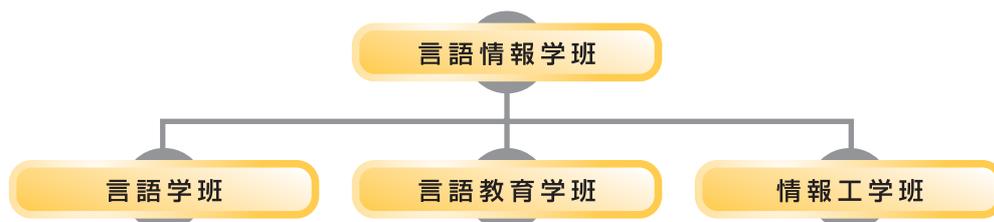
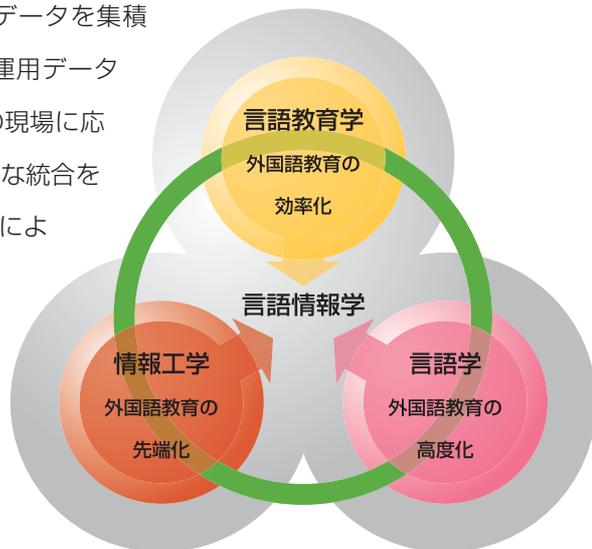
言語情報学とは

コンピュータ科学を基礎にして言語研究と言語教育を統合したらどうなるか、それを考えるのが「言語情報学」という分野です。このパンフレットでは私たちのCOE計画がどのようなものかを、できるだけわかり易く解説します。

東京外国語大学がこれまで行ってきた言語研究では、単に言語理論を追求するだけでなく、言語教育の実践を通して言語理論を見つめなおす眼差しの重要性が説かれてきました。こうした言語研究と言語教育の双方向的なフィードバックは、本学の個性であり、独自の学問的特色とすることができます。このCOE計画は、この特色を最大限に生かしつつ、世界の様々な言語について言語運用データを集積

し、情報工学の知見を利用して言語運用データの分析を行い、その成果を言語教育の現場に応用して、言語理論と言語教育の有機的な統合を図ります。言語学と言語教育学の統合によって、従来の外国語教育は、より先端的で効率的なものになることが期待されます。

このCOE計画では、4つの班が組織化され、それぞれがある程度独立しながら、研究が進められます。



言語情報学班

言語情報学班は後に説明するTUFSS言語モジュールの開発と応用の全体を統括し、COE計画において中心的な役割を担う研究班です。その班のもとに言語学班、言語教育学班、情報工学班の3つの班が形成され、これらはTUFSS言語モジュールの開発・改良・応用に関わる基礎研究を行ないます。

研究組織全体を統括するのは、21世紀COEの事業推進担当者のうちの7名からなる統括班です。年次計画の遂行に関わる重要な意思決定は統括班会議でなされます。さらに各班の関係がより緊密になるように、連絡班も設置されています。連絡班の会議も頻繁に開かれ、プロジェクトの進捗がお互いに報告されます。

統括班：在間進、高垣敏博、敦賀陽一郎、芝野耕司、峰岸真琴、宇佐美まゆみ、川口裕司

連絡班：浦田和幸、黒澤直俊、海野多枝、吉富朝子、佐野洋、林俊成

COE計画によって創成される新たな学問分野の言語情報学は、情報工学を基礎にして言語研究と言語教育を統合した学問分野です。その成果として、いちばん目をひくのは、17の言語を対象とするインターネット上の言語学習システムです。このシステムは「TUFSS言語モジュール」と呼ばれ、これによって、いつでも、どこでも、インターネットさえつながっていれば、17の言語を学ぶことが可能です。

TUFS言語モジュール(TUFS Language Modules)

TUFS言語モジュールは2003年4月25日に内部公開が始まり、字句の修正や誤植の訂正が行なわれ、最初にIPA (International Phonetic Alphabet、国際音声字母) モジュールが6月9日に外部公開されました。IPAは世界のあらゆる言語の発音を同じ基準を用いて表記するとき用いられる音声記号です。IPAモジュールの中に出てくる母音や子音の発音記号をクリックし、オウムをクリックするとそれぞれの音を聞くことができます。

国際音声字母 (International Phonetic Alphabet)

IPA home	両唇音	唇歯音	歯音	歯茎音	歯後音	舌音	舌背音	舌根音	軟口蓋音	硬口蓋音	咽頭音	声門音
破裂音	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ			ʔ
鼻音	m	ɱ		n		ɳ	ɲ	ŋ		ɴ		
ふるえ音	ʙ			r						ʀ		
たたき音・はじき音				ɾ								
摩擦音	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ	
歯面摩擦音				ɸ β								
接近音		ʋ		ɹ			ɻ	ɰ				
歯面接近音				ɻ			ɻ	ɻ				

記号が枠になっているところは、右側のものが有声子音を表す。 : 発音が不可能であると考えられることを示す。

ʃ

説明

無声・後部歯茎・摩擦音 Voiceless postalveolar fricative

- (1) 声帯振動なし → 無聲音
- (2) 舌端と後部歯茎 → 後部歯茎音
- (3) 隙間が奥人中にある → 中線的
- (4) せはめを作り、空気を通りやすくする → 摩擦音
- (5) 上がって鼻腔への通路を閉鎖 → 口音

IPAモジュール <http://www.coelang.tufs.ac.jp/ipa/index.htm>

TUFS言語モジュール <http://www.coelang.tufs.ac.jp/modules/index.html>

TUFS言語モジュールの
17言語
英語、ドイツ語、フランス語、
スペイン語、ポルトガル語、
ロシア語、中国語、朝鮮語、
モンゴル語、インドネシア語、
フィリピン語、ラオス語、
カンボジア語、ベトナム語、
アラビア語、トルコ語、
日本語

東京外国語大学では26の専攻語が教えられています。現在は前ページの17言語から言語モジュールの開発が行なわれていますが、将来的にはすべての専攻語で言語モジュールの実現が期待されます。

次に公開されたのが発音モジュールで、2003年9月に、一挙にドイツ語、フランス語、スペイン語、ポルトガル語、ロシア語、中国語、朝鮮語、モンゴル語、フィリピン語、ベトナム語、日本語の11言語が公開されました。

モジュールと通言語的発想

TUFS言語モジュールは今までにない新しいタイプのウェブ教材です。その名のとおり、「モジュール的発想」に基づいて作られています。子供向けの雑誌に、よく付録の教材がついていました。それを組み立てるとき、あなたは設計図どおりに組み立てましたか。最後に同じものができるのであれば、皆、それぞれのやり方があっていいのではないのでしょうか。このように、ある程度まで互いに独立していながら、全体として一つのまとまりをもった言語学習教材を形成する。これがモジュール型言語教材の本質です。

TUFS言語モジュールは、発音、会話、文法、語彙の四つのモジュールからできており、これを「いつでもどこでも」見ることができ、設計図はついていますが、それにかかわらず「どの部分からでも」自由に好きなところから始めることができます。



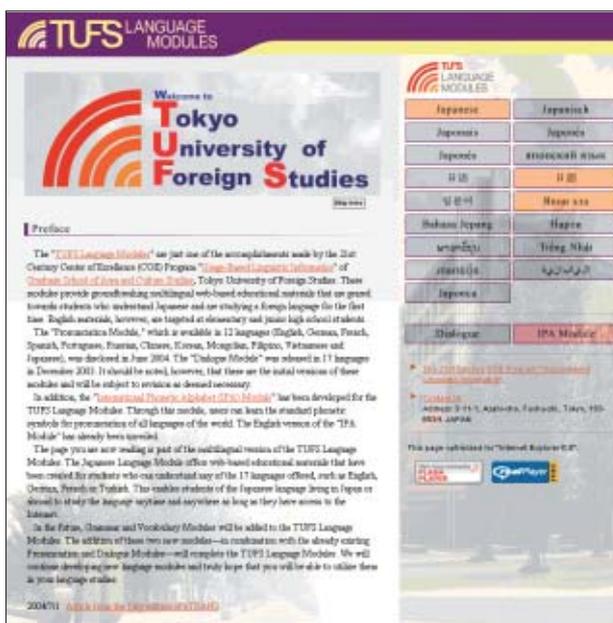
フランス語会話モジュール

TUFS言語モジュールによって、より自由な言語学習が可能になるわけですが、言語学習を評価するには、やはり何らかの統一的な物差しが必要です。モジュール教材は17の言語について、ある程度、内容が共通になるように工夫されています。このゆるやかな統一性のおかげで、一つの言語を中心に、17の言語を横断的にながめることも可能です。外国語を学ぶとき、母語との違いやことばの普遍性を知ることは、とても大切です。

このCOE計画では、さらに一歩進んで、TUFS言語モジュールを用いた言語教育における、東京外国語大学独自の言語能力記述モデルを追求し、将来的に一つのモデルを提案する予定です。17の言語にわたって、ある程度まで共通に言語能力を判定することが可能になれば、中等・高等教育における言語教育に一時代を画することになるかもしれません。ヨーロッパやアメリカでは多言語による言語能力記述の研究が着実に進んで来ているだけに、日本でも同様の取組みに、今、着手することは大変重要です。

2005年7月には通言語的発想を取り入れた通言語文法モジュールが公開されました。通言語文法モジュールとは、文法モジュールのさまざまな言語の文法項目を例にとり、これらの言語に共通する「文法一般の特徴」を通言語的に概観できるページです。文法とはいったい何なのか、人間の言語とは何なのかをじっくり考えてみるのも、語学学習の近道かもしれません。

多言語による双方向モジュール



TUFS言語モジュール（多言語版）
<http://www.coelang.tufs.ac.jp/english/modules/>

一つと同じ教材をいろいろな言語で学ぶことができたらどうでしょうか。それを実現しようというのがTUFS言語モジュールの多言語版です。今のところ、英語、フランス語、中国語（繁体字・簡体字）、朝鮮語、モンゴル語、トルコ語のわかる方が、日本語の発音モジュールと会話モジュールを学習できます。今後もさらにモジュールの多言語化を推し進めます。

TUFS言語モジュールは、最近のユビキタス環境や多文化・多言語主義の流れに対応するだけでなく、従来の日本における外国語教育のあり方そのものに対する一つの挑戦と言えます。TUFS言語モ

ジュールによって、中学や高校の段階から多言語に興味を抱く学生たちが増えてくるでしょう。また、TUFs言語モジュールの多言語版の充実によって、外国語大学は日本人が外国語を学ぶための大学という旧来の殻から脱皮し、海外の方や留学生が日本語や他の言語を学ぶこともできる大学へと生まれ変わる可能性を秘めているのです。

2004年度からは、複数の言語の発音・会話モジュールが東京外国語大学や他大学の授業で実際に利用され、内部・外部評価が行われ、その結果は『言語情報学研究報告5 第二言語の教育・評価・習得』の中に収められています。また2005年末には文法モジュールがドイツ語、フランス語、ロシア語、中国語、フィリピン語、ベトナム語、カンボジア語、トルコ語、日本語で公開されます。2006年度には語彙モジュールも随時公開されます。

言語文化ポータルサイトと言語コーパス

東京外国語大学では26言語が専攻語として教えられています。2005年度にこの大学の特色である言語文化関連のポータルサイトと少数言語のコーパスを構築し、ウェブ上に公開しました。 http://www.coelang.tufs.ac.jp/multilingual_corpus/portal_site.html

言語文化ポータルサイトには26言語について、言語概要、その言語を学習するための参考図書、さらに大使館や新聞やテレビに関するリンク集などの情報が記載されています。

2005年には少数言語コーパスとして、イラン語派東イラン語・パミール諸語に属するワヒー語(Wakhi)の基礎語彙とオーストロネシア語族の西部語派に属する北部ブヌン語(Bunun)のテキストを公開します。さらに2004年度にマレーシア語、ロシア語、スペイン語、2005年度にフランス語、トルコ語、イタリア語で臨地による日常会話の録音が行われました。今後これらの言語において話ことばコーパスが構築され、ウェブ上に公開されます。

The screenshot shows a website interface for the Wakhi language. On the left is a navigation menu with options like '言語文化ポータルサイト' and '言語コーパス'. The main content area is titled 'ワヒー語(Wakhi)' and features four images: a person in traditional dress, a landscape, a group of people, and a building. Below the images is a section titled '収録内容の引続きについて' with detailed text about the corpus and its sources.

ワヒー語の基礎語彙

http://www.coelang.tufs.ac.jp/multilingual_corpus/wakhi

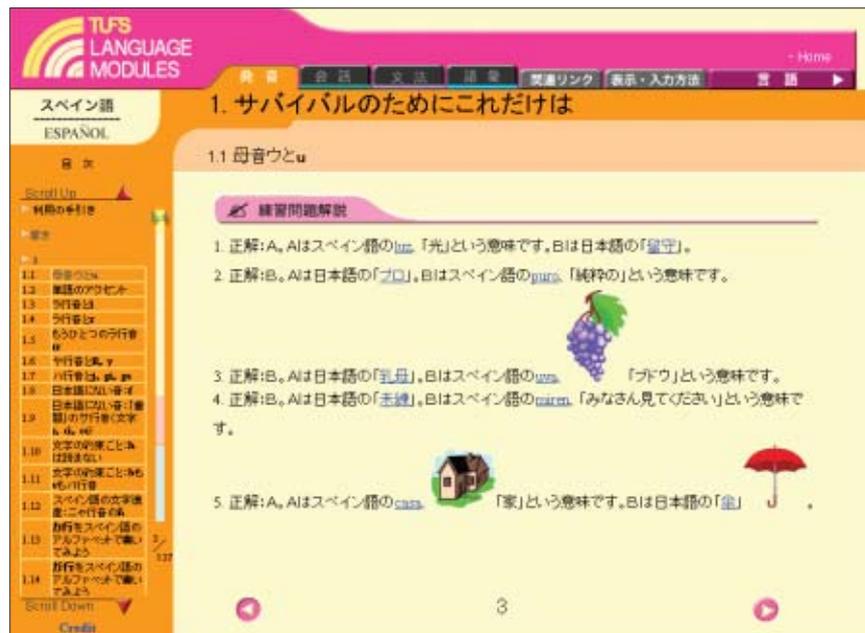
おもしろいぞ発音学習・感動編

木越勉（東京外国語大学大学院博士後期課程）

発音ってとても大切なのに、発音練習というと単調な訓練という感じで敬遠されがちです。発音なんかどうでもいいからとにかく通じればいいんだよとか、なかみが勝負だからねという考えの人もいます。でも、発音があまり悪いとことばは通じませんし、発音のせいで肝心のなかみどころか人間性まで過小評価されてしまうとしたら、うかうかしてはいられません。

本学大学院生の叡智を結集し、この大切な発音学習を楽しめるユーザーフレンドリーな教材として開発したのがTUFS発音モジュールです。まずは詩などの朗読を聞き、初めてのことばがどんな響きかを知ることから始めます。サバイバル、円滑なコミュニケーション、ネイティブ並の発音という3つのパートに分けられ、コミュニケーションのための外国語学習を目的に作られたこの教材は、日常的な語彙や表現を通して正しい発音、リズム、イントネーションが学習者のニーズに応じて身につけられるよう、練習問題を中心にビジュアルに分り易く説明してあります。発音学習のイメージがきつとがらりと変わります。

マウスクリックひとつで何度も繰り返してネイティブのモデル発音を聞き真似ることができるインターネットならではの利点を最大限に活かしたこの教材で、世界のことばを是非身近なものにしてください。



スペイン語発音モジュール

使って満足会話モジュール

阿部一哉（東京外国語大学大学院博士後期課程）

会話モジュールのモジュールは、「～集」「～専門」「部品」という意味です。

会話モジュールではフランス語・ベトナム語・ロシア語などの各国語ごとに40ダイアログの会話集を用意しています。

それぞれの会話は、会話練習専門プログラムを使って練習します。これには、文字を読み、動画を見て、音声を聴くといった基本機能に加え、「今度は日本語を見ないで聞いてみる」「ジェーンの声だけ再生してマイケル役になって話してみる」「単語の意味を調べる」など、より細かい便利な機能が用意してあります。

なお個々の会話は、「お願いする」「意見を言う」「妥協する」など、「言葉で何をしたいのか」に応じた表現を中心に作られています。学習したらすぐ使えるような表現がたくさん用意してあるのです。

ただし会話モジュールは会話を練習するための「部品」。発音や文字の読み方は発音モジュールという「部品」を使って練習します。発音は発音モジュール、表現は会話モジュール。このセットでどうぞネイティブ並みの表現力を身につけてください。

ドイツ語会話モジュール

出典：東外大ニュースNo114

フランス語の扉をひらいてみましょう

杉山香織 (東京外国語大学大学院博士前期課程)

フランス語は「発音が難しそう」だとか、「綴りのとおりに発音しないから大変そう」といったイメージが先行してしまっていて、フランス語はその発音のために、敷居の高い言語だという印象を持たれてしまいがちです。しかし、実は一定のルールさえ覚えてしまえば、むしろ英語よりも発音は単純明快で、どんな単語や文でも発音することができてしまいます。

このフランス語発音モジュール理論編では、個々の音からイントネーションまで、フランス語の発音習得に必要な情報が全てそろっています。インターネットの特性により、ネイティブの発音といった、日本にはなかなか耳にする機会がない音声や、その発音時の口の動きといった動画へ、どこからでも簡単にアクセスできます。また、初歩的な解説から、フランス語音声学といった専門分野もあつかうことで、よりいっそう広い学習者のニーズにこたえることができ、そのため自習や大学の授業での使用など様々な学習形態にも対応できるようになりました。それぞれの項目において、学習者が楽しみながら発音や知識を確認できる練習問題が豊富におさめられている点も、このモジュールの特長です。

発音の習得は、言語習得の第一歩です。このフランス語発音モジュール理論編で、フランス語の扉をひらいてみませんか。

Practical Phonetics & Phonology

TIPS LANGUE MODULES
発音モジュール 理論編

Francis

理論編トップページ

目次

1. 言語概説

2. イントネーション, リズムグループ, 音節の基礎

3. フランス語の文字

4. 個々の音の習得

5. 綴りと発音

5.1. 最終子音字, 母音字の読み方

5.2. アクセント記号のない e の綴りの読み方

5.3. 例外的な発音

5.4. エリジョン

5.5. アンシェスマン

5.6. リエゾン

5.7. 無音の h と有声の h

6. 音節, リズムグループ, アクセント, イントネーション

さらに学びたい人のために

6. 音節, リズムグループ, アクセント, イントネーション > イントネーション > 2. イントネーションの機能 > 2. 疑問文 > 倒置疑問文

疑問文 ① Est-ce que... 次ページ

6.3.2.2. 疑問文 ② 倒置疑問文

◎ 倒置疑問文 (主語と動詞が入れ替わっているもの) のイントネーション

倒置していることから疑問文であることが分かるので、イントネーションは下降することもあります。しかし多くは、答えを強く望む場合や強調する場合は、上昇します。

Êtes-vous d'accord?
賛成ですか?

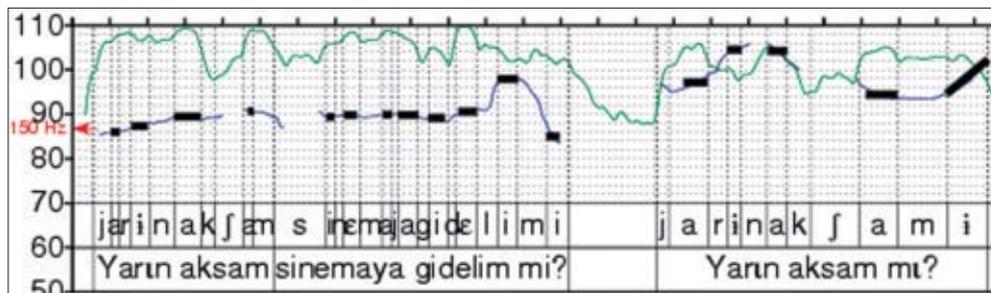
▶ Êtes-vous d'accord? 賛成ですか?

Question 1

1. 賛成
2. 賛成
3. 賛成
4. 賛成

Copyright © 2007 Tokyo University of Foreign Studies. All Rights Reserved.

さらに個別研究として進めている「通言語音声研究プロジェクト—統語構造と韻律—」は、さまざまな言語をこえて、アクセントやイントネーションなどいわゆる韻律の構造がどのようになっているのかを分析し、成果がすでに公開されている発音モジュールなどに応用し、より現実の発話に接近させようとする試みです。このような音声学的プロジェクトはいずれも言語情報学の基礎研究となっています。通言語音声研究では2004年10月に『言語情報学研究報告4 通言語音声研究 音声概説・韻律分析』を出版しました。報告集にはインドネシア語、朝鮮語、フィリピン語、ベトナム語、ラオス語、モンゴル語、アラビア語など11言語の詳細な音声概説が収められています。このほか海外からの投稿を含む、アクセントおよびイントネーションと統語構造に関する論集も進行中です。



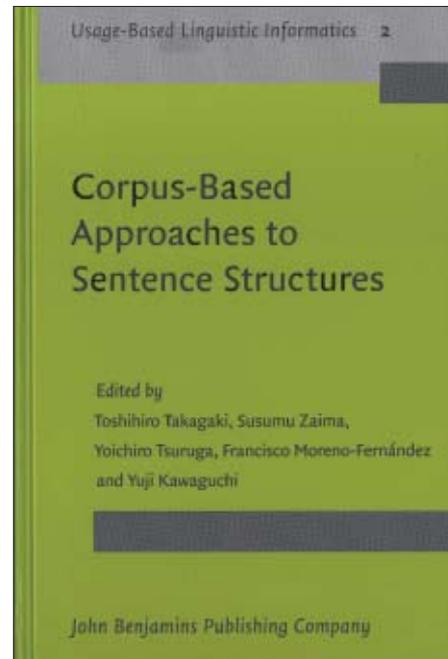
Prosogramを用いたトルコ語イントネーションの分析

2) コーパス言語学

言語学はもともとことばの内部構造の記述をしたり理論的に分析したりすることに専念し、ことばが現実に使われている場面での実態を科学的に分析することには十分な配慮がなされてきたとはいえないかもしれません。実際に用いられている談話やテキストの運用に基づいた理論を構築し、それを語学学習の現場に生かすことに大きな意義があると考えます。

そこで、英語やフランス語、スペイン語、ロシア語、中国語、マレーシア語など諸言語のコーパス（大量の資料体）を作り、これを操作するための関連ツールを開発することにより、言語運用研究を飛躍的に進めるための作業を行っています。

また、こうして利用できるようになった個別言語のコーパスに基づいて言語学研究がなされ、『コーパス言語学における構文分析』（言語情報学研究報告 3）、『コーパス言語学における語彙と文法』（言語情報学研究報告 7）や Corpus-Based Analyses on Sentence Structures 「コーパスに基づく文構造の分析」(Linguistic Informatics II)、Corpus-Based Approaches to Sentence Structures 「コーパスに基づく文構造へのアプローチ (John Benjamins 社) など次々と成果が発表されています。



3) 多言語コーパスの構築

言語の機能、言語の運用、文法研究、音声研究など多目的に利用できる個別言語の自然な発話コーパスを収集する目的で、若手研究者が現地に赴いています。すでに、スペイン語、ロシア語、マレーシア語、トルコ語、フランス語、イタリア語などでフィールド調査を行い、成果からそれらの言語のコーパスが構築されています。また、フィールド調査に関して『フィールド調査による口語資料の収集と分析』（言語情報学研究報告 8）も刊行されています。



4) TUFSS言語モジュールへの応用

このような多言語コーパスは最終的には私たちのプロジェクトの最大の課題である、「TUFSS言語モジュール」に応用されることになります。機能別・目的別多言語コーパスを用いることにより、言語情報学班が開発するTUFSS言語モジュール、その中でもとくに会話モジュールや語彙モジュール、さらには多言語文化ポータルサイトなどのコンテンツを提供し、現在開発中のモジュールを一層充実させることになります。

savoirとconnaîtreの現在形

savoirとconnaîtreの違いを簡単に説明しておきましょう。
savoirには後ろにくる目的語にいくつかの制限があります。たとえば、人や具体的な名詞が目的語になることはふつうありません。これに対してconnaîtreは、生物や具体的な名詞が目的語になります。

この解説では、言語運用データを参照することができます。
言語運用データ

http://www.coelang.tufs.ac.jp - 文法モジュール - Microsoft Internet Explorer

	目的語なし	名詞 je 動詞	ce qui	ça	無生物名詞	生物名詞	不明	出現数
je sais 1	16							32
je sais 2	64	1						128
合計	80	1						160
%	50	1						100%
je connais	30	3	1	3	53	75	1	166
%	19	2	1	2	32	45	1	100%

(je sais 1はトゥール・コーパス、je sais 2はオーヴェルニュ・コーパスを表す)

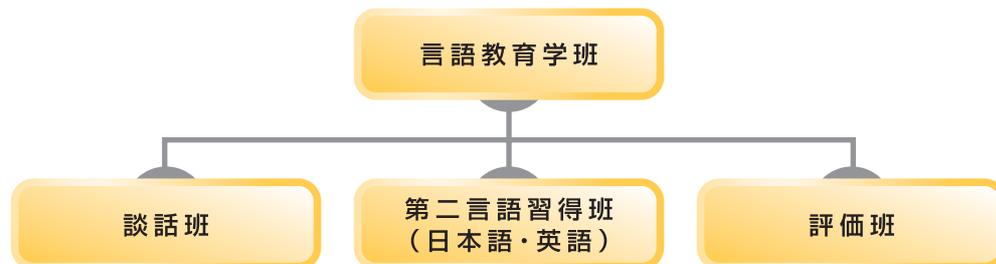
表1から je saisは従属節 si, que, ce queなどを目的語要素とし、je connaisは主に名詞を目的語にしていることが分かります。ただし話しことばでは je sais、je connaisともに目的語なしで、独立して用いられることが多い点に注意が必要です。これを図式化したのが図1です。

je sais / je connais の目的語要素 (図1)

Category	je sais (%)	je connais (%)
従属節	~20	~5
φ	~50	~10
名詞	~5	~45

フランス語文法モジュールにおける言語運用データの応用

言語教育学班



談話班

プロジェクト名：談話データ分析に基づく言語運用についての基礎的研究

談話班の研究目的は、主に、以下の3つです。①言語運用の研究のために自然会話をデータとして用いるため、まずは、自然会話の文字化システムを整備し、談話分析の方法論を洗練させていくこと。②多様な場面の自然会話データを収集すること。③精緻化された方法論に基づいて会話・談話を分析した結果を、将来的には、TUFS会話モジュールの開発に応用していくことです。2003年度は、その成果を以下の5つにまとめました。

(1) 『BTSによる多言語話し言葉コーパス2－日本語（機能別の自然談話集）』

(1は、2002年度の報告書)

創作会話と自然会話の特徴を比較研究するために、TUFS会話モジュールで提示されている40の「見出し機能」が現れている談話を、日本語母語話者同士による自然談話全186会話（約463分）の中から抽出してまとめる作業を行いました。このコーパスを用いることによって、自然談話が多角的に研究できるとともに、創作会話と自然会話の中で、

同じような機能がいかに実現されているかという類似点や相違点を比較研究することが可能になりました。

(2) 『BTSによる多言語話し言葉コーパス3－日本語（日本語母語話者同士の会話）』

(3) 『BTSによる多言語話し言葉コーパス4－日本語（日本語母語話者と学習者の会話）』

2003年度は、雑談場面を中心として、「日本語母語話者同士」の会話と「日本語母語話者と日本語学習者」の会話の収集を行いま



した。これらは、いずれも、会話の対話者の年齢、性別、話題などが、統制された形で集められていますので、様々な観点から比較・対照研究ができるようになっています。

(4) 『自然会話分析研究方法論ハンドブック』

2003年度COE作業の一環として行ってきた自然会話分析の方法論的考察をもとに、『自然会話分析研究方法論ハンドブック』をまとめました。この分野の基本図書・参考文献リストをつけ、今後、広い意味で、「言語運用」を基盤とする「言語情報学」の分野で研究する学生や研究者にとって参考となるよう工夫されています。

(5) 『自然会話分析と会話教育——統合モジュールへの模索』

成果のひとつである『話し言葉コーパス 2』をもとに、機能ごとに自然会話の分析を行うとともに、その会話教育への貢献に関して論じた研究のいくつかをまとめました。TUFSS会話モジュールで扱われる機能が自然会話ではどのように実現されているかを分析することによって得られる観点は、TUFSS会話モジュールの開発や改善に貢献できるとともに、より統合的な教材を開発していくための一つの糸口になっています。

文字化システム (BTS-J) とコーディング (分類) の例
『BTSによる多言語話し言葉コーパス——日本語 (機能別の談話資料集) 2』より

ライン番号	発話文番号	発話文終了	話者	発話内容	発話文全体POL	文末POL
45	Jan.42	*	JBM03	あーそう、出身地で言わせてもらおうと,,[言わせてもらおうと、は聞き取りにくい発話]	/	/
46	43	*	YM01	はい。	NM	NM
47	Feb.42	/	JBM03	僕自身は生まれも育ちも東京なんで,,	/	/
48	44	*	YM01	はいはいはい。	NM	NM
49	45	*	JBM03	あーなんも面白いことがない<笑い>っていう…。	NM	NM
50	46	*	YM01	どのあた、どのあたりですか?。	P	P
51	Jan.47	/	JBM03	えーと生れたのが文京区で、(はいはい)で一幼稚園の(うん)年長の(うん)ときに板橋区高島平に,,	/	/
52	48	*	YM01	はいはいはい。	NM	NM
53	Feb.47	/	JBM03	えー引っ越しをして、でずーっとそこで、暮らしてて,,	/	/
54	49	*	YM01	うんうん。	NM	NM
55	Mar.47	*	JBM03	でーまあ稼ぐようになってからはやっぱりひとり暮らし始めたんですけどね。	P	P

第二言語習得班（日本語）

第二言語習得班（日本語）では、第二言語としての日本語習得に関わる2つのプロジェクトを行いました。

一つめは、東京外国語大学の外国人留学生約150名が執筆した作文資料を、学習者の母語やトピックごとに整理した、『上級学習者の日本語作文データベース』の作成です。日本語の学習者言語や習得の研究、また、教材開発の資料として今後幅広く活用していく予定です。



二つめは、第二言語の学習ストラテジーと学習ビリーフそれぞれに関する二冊のハンドブック、『第二言語学習ストラテジーハンドブック』と『第二言語学習ビリーフハンドブック』の執筆です。これらは先行研究の解説に、本学の学生が各自の専攻語を活かして収集してきた世界各地の学習者についてのデータ事例を合わせ、注釈付き参考文献表や各国語によるアンケートサンプル集をつけました。今後新たにこの分野の研究を始める人への総合的な入門書となるように工夫されています。

第二言語習得班（英語）

2003年度は「学習者言語研究 - 英語学習者言語コーパス基礎調査」と「TUFS言語モジュール英語会話教材付属教師用マニュアル作成」の2つのプロジェクトに取り組み、その成果を以下のようにまとめました。

(1) 学習者言語研究 - 英語学習者言語コーパス基礎調査

英語学習者言語コーパスの実態を把握するための基礎調査として、既存データの調査、分類、評価を行ない、調査結果を報告・資料集にまとめます。このような調査結果を踏まえて、来年度以降、コーパスを利用した英語学習者言語分析研究を実施する予定です。

(2) TUFS言語モジュール英語会話教材付属教師用マニュアル作成

英語会話モジュール教材使用者向けのマニュアルを作成しています。英語会話教材は小学生向けに開発されているため、現場の教師が指導の際に利用するため、また英語D教材を使って独学を行なう大人の初級学習者のために、できるだけ平易な言葉で解説を提供するものです。このマニュアルは英語会話モジュールとリンクを張って自由にアクセス、ダウンロード可能な形で公開することになります。





英語会話モジュール

評価班

(1) 発音モジュールと会話モジュールの評価アンケート分析

先に完成している発音モジュールのうちドイツ語、フランス語、スペイン語、ベトナム語、日本語の利用者に対して教材及び学習評価に関するアンケート調査を実施し、その結果を分析し、今後の改善点を模索しました。分析結果は『言語情報学研究報告5 第二言語の教育・評価・習得』に掲載されています。2005年には会話モジュールの教材および授業評価を行い、その結果を報告書にまとめる予定です。

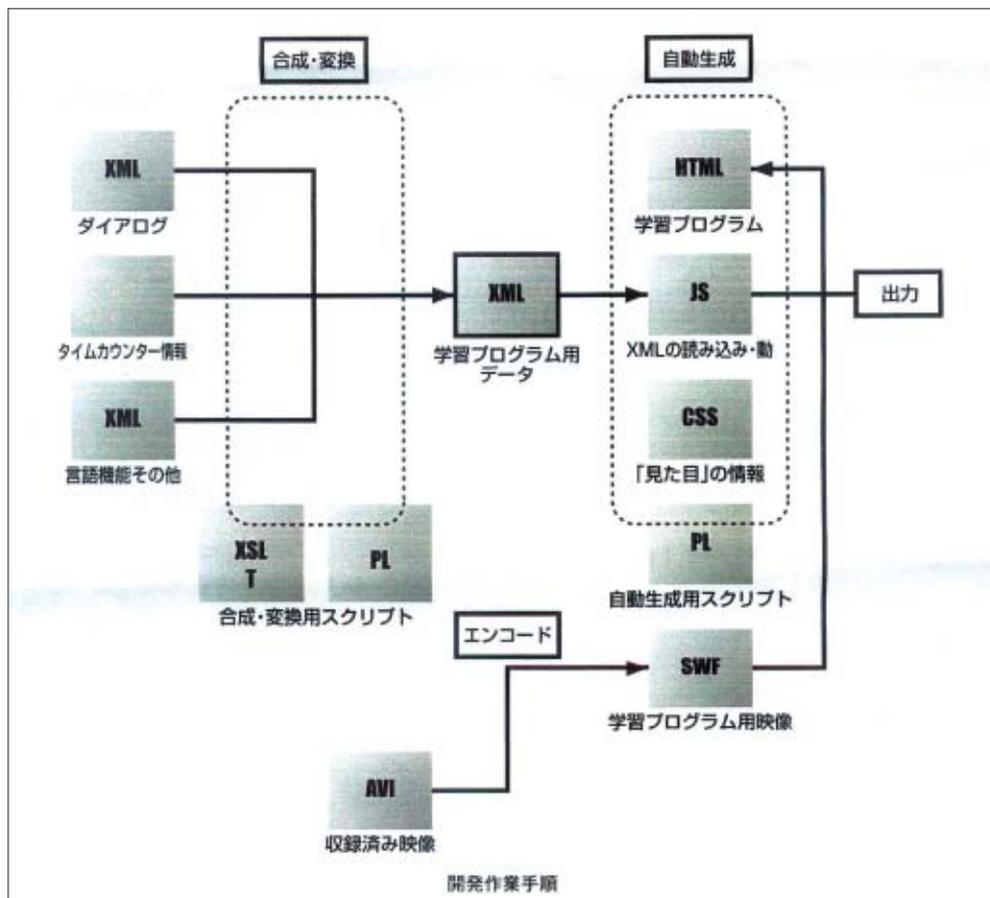
(2) TUFFS言語能力記述モデル作成

外国語能力のパフォーマンス記述を行ったcan-do statement研究を概観し、東京外国語大学独自のTUFFS can-do listの作成を目指します。2005年にはCommon European Frameworkに関するアンケート調査を実施し、その結果を公開します。

教材の電子化

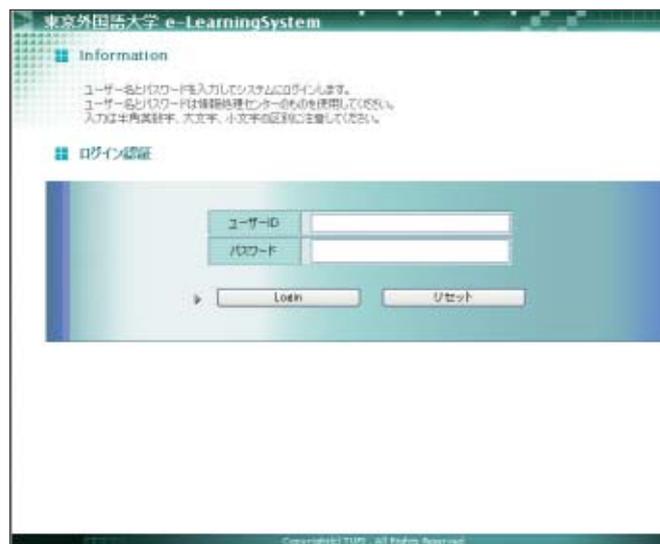
この言語教育的な記述の方法論と枠組みを教材の電子化にも応用しました。言語モジュールを構成するそれぞれの素材は、同じ基準によって記述されています。単語、文、会話、音声、映像、文法項目など、同じ基準でラベルとして表現できます。このラベルをXMLタグとして、教材素材をデータベース化しました。XMLは汎用名マーク付けが可能で、独自のタグセットを持つことができます。テキストを自己記述できるわけで、本言語モジュールでは、自己記述の枠組みが先に述べた言語教育的な方法論に支えられているのです。

また、XMLを利用することによってテキスト内容とスタイルを分離することもできます。これによって、単語、文、会話、音声、映像、文法項目などの教育素材を、様々な教育目的に応じて、あるいは学習対象者の学習型に合わせてビジュアル化する(教材を視覚的に見せる)ことが可能になっています。



e-LearningシステムによるTUFS言語モジュールの利用

TUFS言語モジュールは2003年度に外部公開が始まりました。2005年度からはWeb上の公開だけでなく、一般の学習者にも自習できるように、e-Learningシステム上でも公開されています。TUFS言語モジュールで開発された全てのコンテンツを「講義」としてe-Learningシステム上に登録し、利用者が自由に各講義に参加して、利用できるようになっています。このe-Learningシステムを利用することによって、学習者の学習履歴を管理したり、学習進捗を計測したり、テスト結果を記録することができるようになっています。また、デジタルポートフォリオデータとして学習者の教材のアクセスログも細かく記録されています。これらのデータを解析して、その結果をより良い教材作成に反映する予定です。



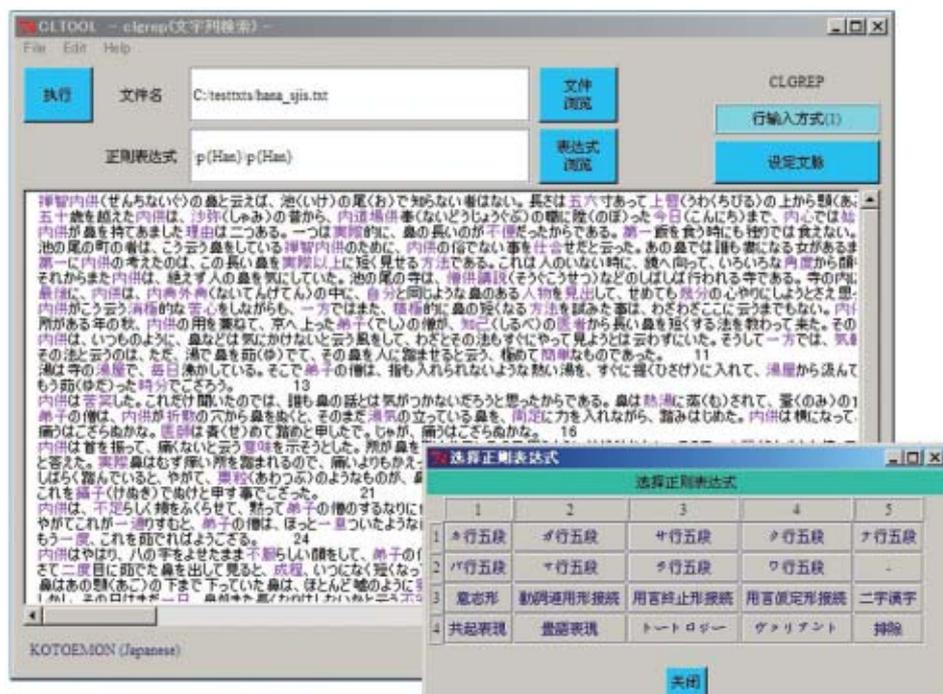
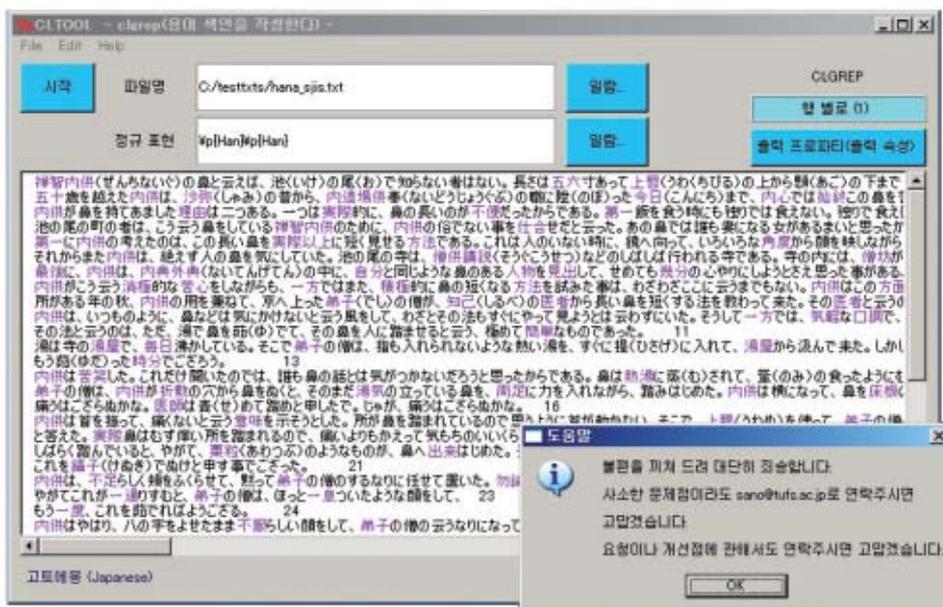
多言語情報処理

情報工学を技術基盤とする言語学と言語教育学の有機的な統合、これが私たちのCOE計画の成果目標である言語情報学拠点形成です。このCOE計画では、世界の様々な言語について膨大な言語運用データを蓄積しようとしています。そして情報工学班は、情報工学の知見を利用して、これらの言語運用データの分析を行います。

コンピュータを使ったことばの分析では、文中にどのような単語が現れているのか、その単語の品詞は何か、どのような種類の活用語尾か、単数形であるのか複数形であるのかななどのことがらを自動的に調べます。そして、いつも文章中に使われる単語や、ある段落にしか出ない単語、共起して出現する単語なども調べます。調査の方法は、コンピュータが自動的に調べることもあります。また、文章中の特定の単語を人間が容易に調べられる

ようわかりやすく表示したりすることもあります。

現在、英語、日本語、韓国語、中国語(簡体字・繁体字)、ロシア語の5カ国語について、データ分析を自動的に行うソフトウェアを開発しています。最終的には、TUFUS言語モジュールの対象となる、その他12言語についても自動的な分析が可能になるでしょう。言語運用データ分析は言語学班によって行われ、その直接の成果が、言語教育学班に引き継がれ、TUFUS言語モジュール教材開発につながっています。



国際会議・国内会議・ワークショップ

第1回言語情報学国際会議

2003年12月に東京外国語大学において、第1回言語情報学国際会議が開催されました。

日 時：2003年12月13日（土）、14日（日）

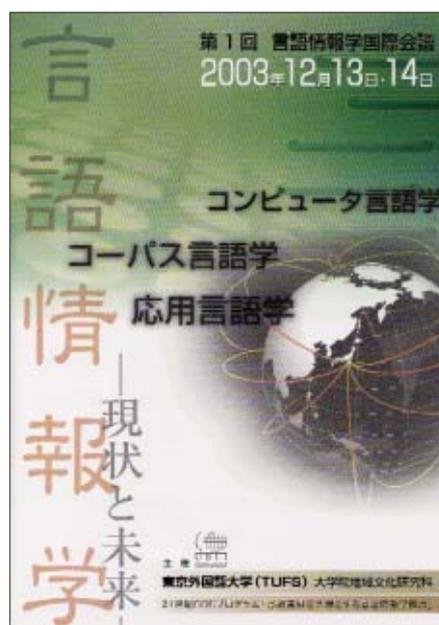
場 所：東京外国語大学研究講義棟 マルチメディアホール

言語理論とコンピュータ科学がこれまで言語教育に大きな影響を与えてきたことはみなさんご存じの通りです。しかしながら、言語学、言語教育学、情報工学の連携は、必ずしも行われてきませんでした。COE計画が目指すのは、その三つの学問領域の間の協働と統合です。この統合的学問領域は言語情報学（Linguistic Informatics）と呼ばれます。

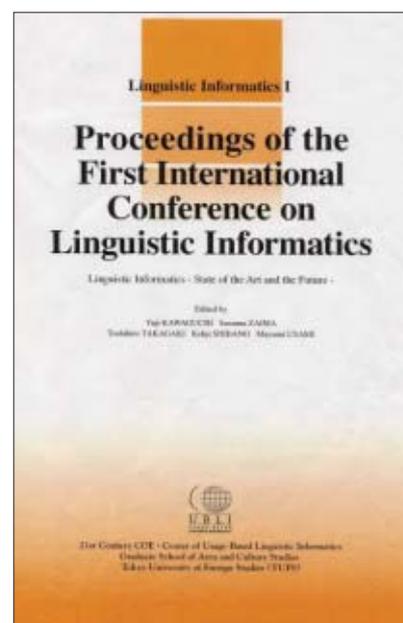
会議では、言語情報学という新しい学問領域がどのようなものなのか、その現状を認識するとともに、将来の可能性が議論されました。3つのセッションがあり、会議は報告から質疑応答にいたるまで、すべて英語で行なわれました。

1. コンピュータと言語学・・・コンピュータと言語学の協働の可能性をさぐる。
2. コーパス言語学・・・言語コーパスを用いた言語研究の現状を報告する。
3. 応用言語学・・・外国語学習と言語理論の関連性をさぐる。

国内外の研究者が講演や報告を行い、大学院生たちの発表も多数ありました。



国際会議プログラム



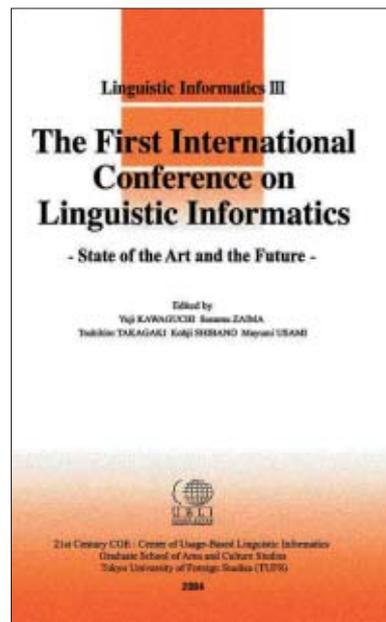
予稿集

言語情報学の領域は、コンピュータ言語学、コーパス言語学、文献学、方言学、談話研究、言語教育学、e-learningなど、広い分野におよんでいます。このため会議に先立って、発表内容を予稿集として出版し、参加者は会議の当日、予稿集を手にしながら研究報告を聞きました。第1回言語情報学国際会議には、2日間で、延べ300名の出席者があり、会議後のレセプションも盛況のうちに終了しました。

この会議の報告集は、「言語運用を基盤とする情報情報学拠点」の研究論文集である『Linguistic Informatics』のシリーズの第3巻「The First International Conference on Linguistic Informatics」として刊行されたほか、内容と装丁を新たにオランダのJohn Benjamins社からもシリーズUsage-Based Linguistic Informaticsの第1巻として出版されました。



国際会議ポスター



21世紀COE研究論文集

21世紀COE統合シンポジウム

2005年1月、2月に東京外国語大学において、21世紀COE統合シンポジウム「グローバル化と多文化的想像力」が開催されました。

日 時：2005年1月27日（木）、28日（金）、2月10日（木）

会 場：東京外国語大学 マルチメディアホール等



主催：東京外国語大学21世紀COEプログラム「言語運用を基盤とする言語情報学拠点」
東京外国語大学21世紀COEプログラム「史資料ハブ地域文化研究拠点」

企画責任者 亀山 郁夫

言語情報学国内会議

2005年10月に東京外国語大学において、言語情報学の国内会議が開催されました。

日 時：2005年10月4日（火）、5日（水）

会 場：東京外国語大学 マルチメディアホール等

近年、言語研究においてコンピュータを利用したコーパス分析や統計解析などが盛んに行われるようになり、従来にはない新たな知見が見いだされるようになりました。一方、自然な談話を語用論の観点から分析し、それをウェブ教材に取り入れる等の試みも始まっています。さらに学習者の言語コーパスを分析し、その成果を外国語学習に応用する努力も続けられています。このように言語学と言語教育学とコンピュータ科学の距離は、ますます近づきつつあると言えます。言語運用を基盤とする言語情報学拠点では、コンピュータ科学の基盤の上に言語学と言語教育学を統合した学問分野として言語情報学(Linguistic Informatics)を定義しています。

言語情報学国内会議

言語情報学とは何か

新たな若手研究者の育成に向けて

10月4日(火) 13:10~18:30

会場 419号館(419号)

【招待講演】 佐野 史子 (京都府立大学准教授)
少人数英語の授業と教育 ―活用フィルムシアタの事例から―

Ⅰ. 言語学・言語教育学への寄与
佐々木 肇彦 (高野 三枝子) (言語学)

Ⅱ. 応用言語学・言語教育学への寄与
矢野 典雄 (丸 藤社)

Ⅲ. 情報工学・言語情報学への寄与
野田 一也 (成之 久美子)

【閉会式】 卒、青年研究会に賞められるものは何か?
司会 高野 三枝子 (東京外国語大学)

言語学・応用言語学・情報工学の言語情報学への寄与

10月5日(水) 12:50~18:00

会場 101マルチメディアホール (12:50-14:40)

【招待講演】 高野 三枝子 (東京外国語大学准教授・言語学)
言語のカタチ性について
赤坂 成司 (東京工業大学准教授)
情報工学と言語学習実践システム

会場 115教室 (15:00-18:00)

Ⅰ. 言語学: 言語情報学への寄与
佐野 史子 (京都府立大学准教授)
言語のカタチ性について
言語運用、検証、そして応用

Ⅱ. 応用言語学: 言語情報学への寄与
宇佐美 まゆみ (本学大学院准教授)
語法研究におけるローカル分析とグローバル分析の意義

Ⅲ. 情報工学: 言語情報学への寄与
林 啓成 (本学大学院准教授)
TEFLS言語モジュールの設計およびe-Learningシステム上の利便

Ⅳ. JASAL討論: 言語情報学とは何か

◎本場
東京外国語大学21世紀COEプログラム
言語運用を基盤とする言語情報学拠点

【お問い合わせ先】
東京外国語大学 21世紀COEプログラム 言語運用を基盤とする言語情報学拠点事務局
〒153-8504 東京都目黒区目黒3-11-1 Tel. 042-352-9541 E-mail: coe@lang@fu.ac.jp



ワークショップ

2005年12月に東京外国語大学において、ワークショップ「話ことばコーパス—その重要性と応用—」が開催されます。

日 時：2005年12月9日（金）

会 場：東京外国語大学 中会議室

C-ORAL-ROMはイタリアのフィレンツェ大学、スペインのマドリード自治大学、フランスのエクス大学、ポルトガルのリスボン大学の4大学による、「ロマンス諸語の話ことばコーパス」に関する共同プロジェクトです。このワークショップはC-ORAL-ROMと「言語運用を基盤とする言語情報学拠点（UBLI）」の共催で開催されます。自由討論では、C-ORAL-ROMとUBLIの協働の可能性や話ことばコーパスを用いた言語研究の将来などについて議論がなされます。

第2回言語情報学国際会議

2005年12月に東京外国語大学において、第2回言語情報学国際会議が開催されます。

日 時：2005年12月10日（土）

会 場：マルチメディアホール

第2回国際会議では、話ことば分析、コーパス言語学、語用論に関する講演を通して、言語情報学の理論的基盤を明確にし、この新しい分野の学問的な位置づけを行います。

自由討論では、3つの観点から言語情報学の現状が議論されます。

1. 言語分析におけるコーパスの意義
2. 言語コーパスの言語教育への応用
3. 情報工学の言語学と言語教育への寄与

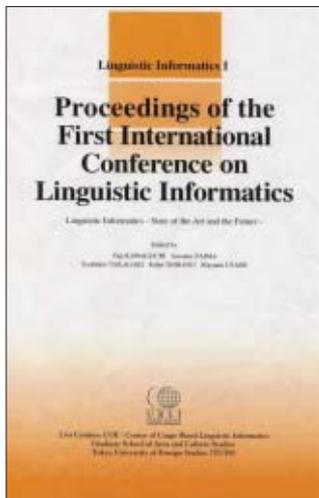
この第2回国際会議の報告集は国内会議の研究報告とあわせて、John Benjamins社から出版される予定です。

出版物

論文集

Linguistic Informatics I

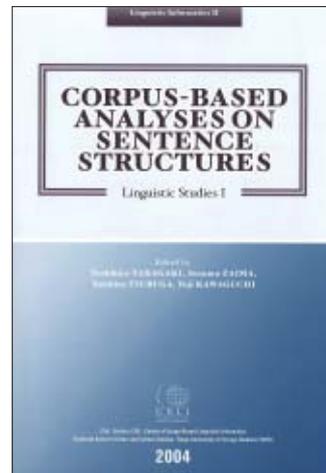
Proceedings of the First International Conference on Linguistic Informatics



Edited by
Yuji KAWAGUCHI, Susumu ZAIMA, Toshihiro TAKAGAKI,
Kohji SHIBANO, Mayumi USAMI
2003年10月刊行

Linguistic Informatics II

Corpus-Based Analyses on Sentence Structures



Edited by
Toshihiro TAKAGAKI, Susumu ZAIMA, Yoichiro TSURUGA,
Yuji KAWAGUCHI
2004年4月刊行

Linguistic Informatics III

The First International Conference on Linguistic Informatics



Edited by
Yuji KAWAGUCHI, Susumu ZAIMA, Toshihiro TAKAGAKI,
Kohji SHIBANO, Mayumi USAMI
2004年10月刊行

研究報告集

言語情報学研究報告 1
TUFS言語モジュール



川口裕司、芝野耕司、峰岸真琴（編）
2004年3月刊行

言語情報学研究報告 2
言語学・応用言語学・情報工学



川口裕司、峰岸真琴（編）
2004年3月刊行

言語情報学研究報告 3
コーパス言語学における構文分析



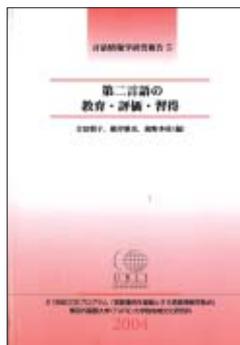
敦賀陽一郎、黒澤直俊、浦田和幸（編）
2004年8月発行

言語情報学研究報告 4
通言語音声研究—音声概説・韻律分析—



川口裕司、森口恒一、斉藤純男（編）
2004年11月発行

言語情報学研究報告 5
第二言語の教育・評価・習得



吉富朝子、根岸雅史、海野多枝（編）
2004年12月発行

言語情報学研究報告 6
自然会話分析と会話教育—総合的モジュール作成への挑戦—



宇佐美まゆみ（編）
2005年4月発行

言語情報学研究報告 7
コーパス言語学における語彙と文法



敦賀陽一郎、高垣敏博、浦田和幸（編）
2005年10月発行

言語情報学研究報告 8
フィールド調査による口語資料の収集及びその分析



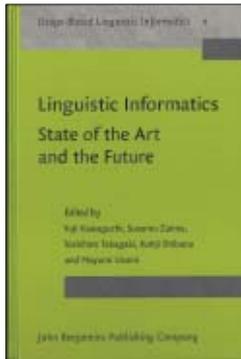
風間伸次郎、川口裕司（編）
2005年10月発行

John Benjamins Publishing Company

Series : Usage-Based Linguistic Informatics

Usage-Based Linguistic Informatics 1

Linguistic Informatics State of the Art and the Future

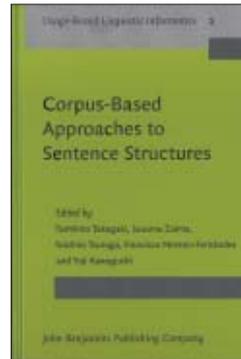


Edited by
Yuji Kawaguchi,
Susumu Zaima,
Toshihiro Takagaki,
Kohji Shibano,
Mayumi Usami

2005年発行

Usage-Based Linguistic Informatics 2

Corpus-Based Approaches to Sentence Structures



Edited by
Toshihiro Takagaki,
Susumu Zaima,
Yoichiro Tsuruga,
Francisco Moreno-
Fernández,
Yuji Kawaguchi

2005年発行

東京外国語大学案内マップ



183-8534 府中市朝日町 3-11-1 東京外国語大学
21世紀COE 言語運用を基盤とする言語情報学拠点
TEL. 042-330-5541 (事務局)



183-8534 府中市朝日町 3-11-1
東京外国語大学21世紀COE
言語運用を基盤とする言語情報学拠点
TEL. 042-330-5541 (事務局)