

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 はじめに 4 6 21世紀COE 言語運用を基盤とする言語情報学拠点 7 言語情報学班 TUFS言語モジュール (TUFS Language Modules) 8 言語学班 14 言語教育学班 18 情報工学班 22 26 第1回言語情報学国際会議 出版物 32





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えました。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リティ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マルチメディアやインターネットを駆使し、多様な方法を動員した先進的な言語教育を目指しています。

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。

2004年3月

目次ページに戻る

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

7 VILCE 15 11 11		
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
 梶茂樹	音韻論	言語学
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



言語情報学拠点ホームページ http://www.coelang.tufs.ac.jp/



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)

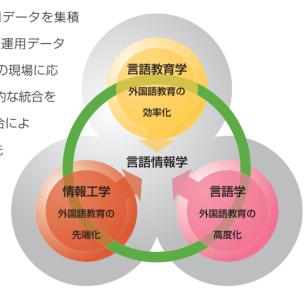
言語情報学とは

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言うことができます。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TUFS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TUFS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、吉冨朝子、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「TUFS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に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。



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



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、 フィリピノ語、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

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



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、フィリピノ語ベトナム語、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、それにかかわりなく「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して、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行なわれてきているだけに、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。

多言語による双方向モジュール

一つの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、英語とモンゴル語と中国語(繁体字)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。



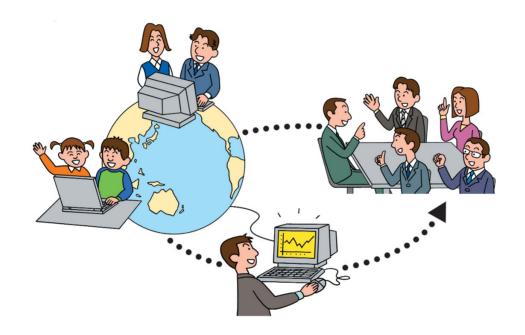
TUFS言語モジュール(多言語版) http://www.coelang.tufs.ac.jp/english/modules/

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。

生成する教材



インターネット上の素材は取りかえがとても簡単です。TUFS言語モジュールも同じで、そのプロセスは、まず言語素材が作成され、次にそれがウェブ化されます。そして実際に教材が使用され、評価が戻ってくると、もとの言語素材に修正や改訂が加えられ、新たなモジュール教材として生まれかわります。

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の授業で実際 に利用される予定です。今後も文法モジュールや語彙モジュールが随時公開されます。



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前期課程)

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。先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っ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



言語学班

言語学の研究

26の専攻語を含めると50もの言語が学べる東京外国語大学ではそれぞれの言語について長年、研究・教育がなされてきました。言語学の理論に基づいてことばの音声や語形(形態)や文法(統語)、それに語や文の意味、また、文と文のつながりを分析する談話研究などさまざまなレベルでの言語学の研究に取り組み、それを教育に応用してきました。このような成果の蓄積の上に私たちの語学教育が実践されてきたといえます。

ところで近年コンピュータ科学のめざましい発展によって語学教育への応用の可能性が急速に拡大しました。情報工学を最大限に利用して、いま述べた言語の分析と言語教育を組み合わせ、効率的で有機的なまとまりのある語学学習を実現しようとするのが私たちの「言語運用を基盤とする言語情報学拠点」の大きな目的なのです。そしてこのように情報工学を基盤にし、言語学と言語教育学を連携させて作り出すことをめざす新しい学問分野を「言語情報学」(Linguistic Informatics)と名づけることにしました。

言語学班の課題

ですから言語情報学の中で、言語学班の使命は当然、プロジェクト全体に対して理論的 基礎を提供することであるといえます。言語学班では現在次のような具体的課題に取り組 んでいます。

1) 音声研究

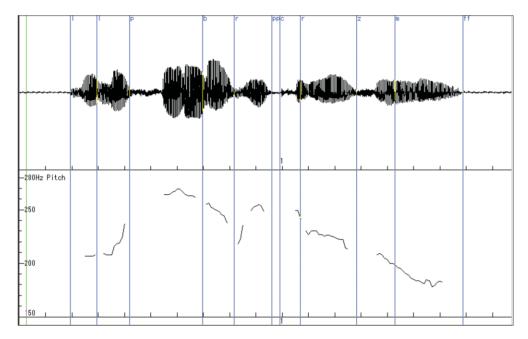
音声の研究ではまず、「IPAモジュール」をすでにウェブ上で見ることができます。これは世界中のすべての言語の音を同じ基準で記述できるように国際音声学会が定めた国際音声字母(IPA=International Phonetic Alphabet)を参照のために利用できる

ウェブ教材です。これを使って個々の母音、子音などの具体的音声を確かめ、学習することができます。

国際音声字母(International Phonetic Alphabet) ||PA home 音声器官 | 母音 | 子音(特気流) 子音(非筋気流) その他の記号 | 参考文献 m たたき音・はじき音 ş z θ δ | s z ł 5 φβfv č ј х Х Х в z | ∫ 3 ħΥ 岸接官 側面摩擦音 щ 記号が対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 (幕宗派)にな)子音を記述するための5つの基準 (1)発声の推撲(声帯の指統がある「有声音」が揺動がない「無声音」、息もれ声、きしみ声など) (2)開音の位置との部分で室次の消化を効害するか) (3) 宗派の滅路(宗派は口腔内の再ム中を過る「中線句」か、版を過る「角面的」か、厚煌音と接近音にのみ関係) (5)口蓋帆の状態(気流は鼻に抜ける「鼻音」か、抜けない「口音」か)

IPAモジュール 子音 (肺気流)

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ー統語構造と韻律ー」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。



音声工房を用いたフランス語の韻律分析

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整備を行っています。



中近東工科大学のトルコ語コーパス(METU Turkish Corpus)

原文

Biten bin yılın başlayan bin yıla devrettigi mesaj:

文法タグ記述

Biten\(\frac{2\(\)}{2\(\)}\)bit+Verb+Pos^DB+Adj+PresPart\(\frac{2}{2}\)Biten+Noun+Prop+A3sg+Pnon+Nom\(\)\(\)\<corr\(\)?resp=Ayg\(\)\in\(\)\>

 $bin \$3\$ bin + Num + Card \$bin + Num + Card \land DB + Noun + Zero + A3sg + Pnon + Nom \$bin + Verb + Pos \\ + Imp + A2sg \$$

 $y_1l_1 + 4 y_1 l + Verb + Pos + Imp + A_2 p_1 y_1 l + Verb + Pass + Pos + Imp + A_2 s_2 y_1 l + Noun + A_3 s_2 + Pon + Gen_4 y_1 l + Noun + A_3 s_2 + Pos + Imp + A_2 s_2 y_1 l + Noun + A_3 s_2 + Pos + Imp + A_2 s_2 y_1 l + Noun + A_3 s_2 + Pos + Imp + A_2 s_2 y_1 l + Noun + A_3 s_2 + Pos + Imp + A_2 s_2 y_1 l + Noun + A_3 s_2 h + Noun + A_3 s_3 h + Noun + A_3 s_3 h + Noun + A_3 h$

</corr>

başlayan¥1¥başla+Verb+Pos^DB+Adj+PresPart¥

<corr?sic=binyıla?resp=Aygün>

 $bin \$3\$ bin + Num + Card \$bin + Num + Card \land DB + Noun + Zero + A3sg + Pnon + Nom \$bin + Verb + Pos \\ + Imp + A2sg \$$

 $y_1la \\ ¥2 \\ ¥y_1l + Verb + Pos + Opt + A3sg \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + Dat \\ ¥y_1l + Noun + A3sg + Pnon + Dat \\ ¥y_1l + Noun + Dat \\ ¥y_1l + Noun + Dat \\ ¥y_1l + Noun + Dat \\ Y_1l + N$

</corr>

 $\label{eq:continuity} devrettigi & 2 & devrett & Pos^DB + Noun + PastPart + A & 3 & sg + P & 3 & sg + Nom & & & & & & \\ & + A & dj + PastPart + P & 3 & sg & & & & & \\ & + A & dj + PastPart + P & 3 & sg & & & & & \\ & + A & dj + PastPart + P & 3 & sg & & & & & \\ & + A & dj + PastPart + P & 3 & sg & & & & \\ & + A & dj + PastPart + P & 3 & sg & & & \\ & + A & dj + PastPart + P & 3 & sg & & & \\ & + A & dj + PastPart + P & 3 & sg & & & \\ & + A & dj + PastPart + P & 3 & sg & & & \\ & + A & dj + PastPart + P & 3 & sg & & & \\ & + A & dj + PastPart + P & 3 & sg & & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + PastPart + P & 3 & sg & & \\ & + A & dj + P & 3 & sg & & \\ & + A & d$

mesaj¥1¥mesaj+Noun+A3sg+Pnon+Nom¥

:¥1¥:+Punc¥

また同時に、そのようなコーパスを活用し、動詞分類、動詞の相・叙法、格あるいは文型、構文、談話などの個別的な文法テーマに取り組んだり、言語教育、比較言語学、対照言語学などへの具体的応用を考えることも課題としています。

3)機能別・目的別多言語コーパスの構築

さらに言語運用研究には他にも、言語の機能に注目する「目的別の多言語コーパス」 (願望、質問、謝罪など、日常表現に関する談話コーパス)や内容に焦点をあてた「分野別の多言語コーパス」(文化研究、歴史研究、言語研究、経済学に関する基礎語彙を網羅した分野別コーパス)などの可能性があり、このような特定の目的に合ったコーパス構築にも現在取り組みつつあります。インターネットを利用したり、若手研究者が現地に赴いて実際の発話を収集することになります。これにより、外国語教材の開発のみならず、異文化理解や教養教育に大きな貢献ができるものと期待されます。

4) TUFS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、「TUFS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TUFS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。



カンボジア語会話モジュール



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。 TUFS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、TUFS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うと,,[言わせて もらうと、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	あの一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?。	Р	Р
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— 人暮らし始めたんですけどね。	Р	Р



第二言語習得班(日本語)

プロジェクト名:第二言語の学習者言語研究・第二言語学習者研究

2003年度は、第二言語の学習者言語研究と、第二言語学習者研究プロジェクトを遂行し、その成果を以下の3つにまとめました。

(1)『上級学習者の日本語作文データベース』

東京外国語大学外国語学部日本課程の外国人留学生約150名が5つの異なるトピックについて執筆した日本語作文資料を電子化し、データベース化したものです。作文は、執筆された年度及びトピックごとに配列されています。また、学習者の母語別に分類した母語別バージョンも用意しました。日本語の学習者言語研究の資料として、また、上級者用日本語作文教材開発の資料としても広く活用できます。

- (2) 『第二言語学習ストラテジーハンドブック』
- (3) 『第二言語学習ビリーフハンドブック』

上記の二つは、第二言語学習ストラテジーと第二言語学習ビリーフそれぞれに関する、 先行研究の解説と具体的研究事例を含むハンドブックです。外国語教育、特に日本語教育 における先行研究を広く集め、そこに示された諸概念の定義、分類、研究方法論を整理し て解説し、研究事例として、本学の学生が専攻語を生かして集めた、様々な地域からの学 習者のデータを示しました。巻末には各国語によるアンケートサンプル集と参考文献表を つけ、これらの分野について新たに研究を行う人々への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)

プロジェクト名:(1) 学習者言語研究——英語学習者言語コーパス基礎調査

- (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
- (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・資料集にまとめました。このような調査結果を踏まえて、2004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ました。英語会話モジュール 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、また英語会

目次ページに戻る

話モジュール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。以下が内容の詳細です。

- ①ダイアログの日本語訳(意訳、口語体)と英語の原文の間の橋渡しをする逐語訳の提示、 及び逐語訳と意訳の違いに関する解説。
- ②ダイアログの中心となっている機能(target expressions)に関する解説。
- ③ダイアログ付属の練習問題、解答、及びその和訳と解説。
- ④会話モジュール教材を使った教室での活動案の紹介。
- ⑤ダイアログのターゲット機能を表す表現のさまざまな言い換え表現の紹介。

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 **Sentencestart:2.1**

評価班

(1) 発音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利用者に対してアンケートを実施し、その結果を分析し、今後の改善点を模索しています。

(2) TUFS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TUFS言語能力記述モデルの作成を目指しています。



情報工学班

TUFS言語モジュールのe-learning化

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」の研究成果の一つであるTUFS言語モジュールのe-learning化を推進しています。TUFS言語モジュールでは、17の言語教材が提供されます。17言語の教材のうち英語は、小中学生を学習対象者としています。日本語は、日本語を学ぶ外国人の学生を対象者としています。残りの15言語の教材は、大学で第二外国語を学ぶ大学生を対象者にしています。

私たちは、本年度までに、これらの17言語のモジュールをウェブページ化し、インターネット上に公開しました。

17の言語モジュールのそれぞれは、発音モジュール、会話モジュール、文法モジュール、語彙モジュールの4つのモジュールから構成されています。このように1つの言語は、4つの教材モジュールから構成されています。それぞれのモジュールは共通の記述の枠組みで作られていることに特徴があります。この特徴によって、言語が違っても各モジュールに記載される項目は同じになっています。例えば、会話モジュールは、発話の状況と機能を決めて、各言語で機能を実現する発話を記述しています。また、発話モジュールはIPAに基づいて各言語の発音教材が作られています。



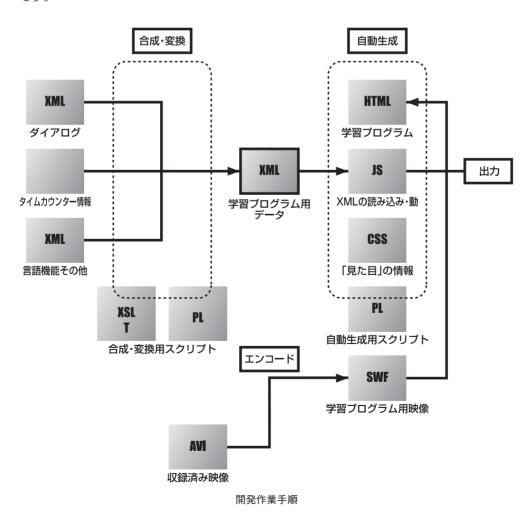
アラビア語会話モジュール



こうした17言語全てに適用できる普遍的な枠組みによる会話教材や文法教材の作成は、 今までにはない新しい語学教材の開発手法です。こうした手法を確立することによって、 まだ語学教育の教科書がない言語でも速やかに教材を作成することができるようになりま した。この方法論の開発は、教材作成の効率化にとって大変に意味のあることです。

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、文法項目などです。同じ基準は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は汎用な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ライズする(教材を視覚的に見せる)ことが可能になっています。2004年度までに17言語のモジュールは、一つの教材スタイルを決め、ウェブページ化し、インターネット上に公開しました。

今後は、本格的なe-learning化を目指すことになります。e-learningプラットホームを導入して、既に開発されたウェブページを運用することによって、学習者の学習履歴を管理したり、学習進度を計測したり、また、テスト結果を自動集計したりすることができるようになります。より高度な学習環境が提供されるでしょう。

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の形成です。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どのことがらを自動的に調べます。そして、いつも文章中で使われる単語や、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。

自動的な分析を英語や日本語以外の幾つかの言語で可能にすることも行っています。例えば、現在、ロシア語やスペイン語に取り組んでいます。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、TUFS言語モジュール教材開発につながっています。





フィリピノ語会話モジュール





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)

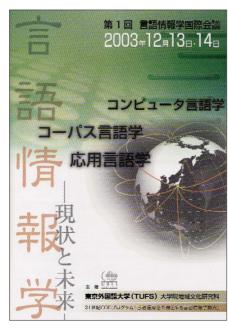
場 所:東京外国語大学研究講義棟 マルチメディアホール

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もご存じの通りです。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。

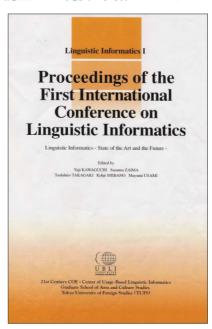
会議では、言語情報学という新しい学問領域がどのようなものなのか、その現状を認識するとともに、将来の可能性が議論されました。3つのセッションがあり、会議は報告から質疑応答にいたるまで、すべて英語で行なわれました。

- 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。



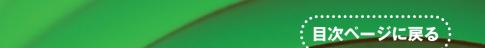
国際会議プログラム



予稿集

国際会議プログラム

Dec.13 (Sat.)		
	総合司会 吉冨朝子・矢頭典枝	
学長挨拶	池端雪浦(東京外国語大学学長)	10:00-10:10
プロジェクト説明	川口裕司(COE拠点リーダー)	10:10-10:20
招待講演 I	司会 在間進、LAM Patrick Pieter van REENEN, Anke JONGKIND (アムステルダム自由大学、オランダ) One Or Two Phonemes: /ø/ - /u/ In Old French, /s/ - /z/ In Dutch And Frisian. New Solutions To An Old Problem	10:20-11:00
	Christian LECLÈRE (マルヌ・ラ・ヴァレ大学 LADL、フランス) The Lexicon-Grammar of French Verbs: a	11:00-11:40
	Syntactic Database 在間進、阿部一哉、松井健吾(東京外国語大学) TUFS IPAモジュール、TUFS言語モジュール	11:40-12:20
	seigneur au sereur honneur pa cle ur va cle ur emplur pa cur la corem la co	
昼食		12:30-13:30
招待講演Ⅱ	司会 高垣敏博 Francisco MORENO-FERNÁNDEZ (アルカラー・デ・エナレス大学、スペイン) Corpus of spoken Spanish language	13:30-14:10
	The representativeness issue 上田博人(東京大学) Methods of 'Hand-made' Corpus Linguistics - A Bilingual Database and the Programming of Analyzers -	14:10-14:50
	y for Spanish from Spain. Terms and concepts related to the guistic variation used with a markable lack of r. The languar antitative to seems more antitative to seems	





セッション1

コンピュータ 言語学/

コーパス言語学

司会 馬場彰

室井禎之(早稲田大学)

Multilateral interpretation of corpus-based

semantic analysis

Antonio RUIZ TINOCO(上智大学)

Tools for Creating Online Dictionaries

- Judeospanish, a Case Study 宮本正美(神戸市外国語大学)

A formal analysis of Spanish adjective position





司会 富盛伸夫

黒澤直俊 (東京外国語大学)

On the language of Portuguese «Estoria do

Muy Nobre Vespesiano»

中村卓也(マルヌ・ラ・ヴァレ大学博士課程)

Analysing texts in a specific domain with local grammars – The case of stock

exchange market reports –

鑓水兼貴、川口裕司、市川雅教(東京外国語大学) Multivariate Analysis in Dialectology A Case Study of the Standardization

in the Environs of Paris

15:00-15:30

15:30-16:00

16:00-16:30

16:40-17:10

17:10-17:40

17:40-18:10

10:00-10:40

10:40-11:20

Dec.14 (Sun.)

招待講演Ⅲ

司会 宇佐美まゆみ

Janet HOLMES

(ヴィクトリア大学、ニュージーランド)

Socio-pragmatic aspects of workplace talk David BLOCK (ロンドン大学、イギリス) What do we mean by 'second' in Second

Language Acquisition





12:30-13:30

昼食

目次ページに戻る

招待講演Ⅳ	司会 宇佐美まゆみ 西原鈴子(東京女子大学) Integrating applied linguistics research findings with Japanese language pedagogy: a challenge in contrastive pragmatics In developing dialomodules and conveteaching materials: USAMI Mayumi, Ed Professor Tokyo University of For Email: usamima@tufs.s	12:30-13:10
セッション2 応用言語学 談話研究/ 言語教育	司会 海野多枝 宇佐美まゆみ(東京外国語大学) Why do we need to analyze natural conversation data in developing dialogue module and conversation teaching materials? 鈴木卓(フェリス女学院大学)、松本剛次、宇佐美まゆみ(東京外国語大学) An analysis of teaching materials based on New Zealand English conversation in natural settings: Implications for the development of conversation teaching materials 木林理恵、金庚芬、謝韞、関崎博紀、木山幸子、施信余、李恩美、松本剛次、宇佐美まゆみ(東京外国語大学) An examination of the process of developing a Corpus of Spoken Japanese: Implications for the development of conversation teaching materials 謝韞、木林理恵、木山幸子、施信余、李恩美、金庚芬、松本剛次、宇佐美まゆみ(東京外国語大学) A comparative analysis of discourse behaviors in Japanese natural conversation and the Japanese skits of TUFS dialogue module: Implications for the development of conversation teaching materials	13:20-13:50 13:50-14:20 14:20-14:50
セッション3 応用言語学 e-learning/ シラバス研究	司会 根岸雅史 Mark PETERSON(東京外国語大学) Computer Assisted Language Learning (CALL) Moving into the networked future Malcolm H. FIELD(はこだて未来大学) Beyond the Novelty: Providing Meaning in CALL 藤原愛(東京外国語大学) Pronunciation Practice Based on a Skill Syllabus	15:20-15:50 15:50-16:20 16:20-16:50





セッション4 応用言語学 TUFS 言語モジュール	司会 吉冨朝子 阿部一哉(東京外国語大学) Development and Assesment of the TUFS Language Module – Multilingual e-learning Application – 木越勉(東京外国語大学) Creation of TUFS Language Pronunciation Module 結城健太郎、阿部一哉、林俊成(東京外国語大学) Development and Evaluation of TUFS Dialogue Module – Multilingual and Functional Syllabus	17:00-17:30 17:30-18:00 18:00-18:30
閉会の辞		
レセプション		19:00-20:30

目次ページに戻る

プログラムからもわかるように、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無事終了しました。

この会議の報告集は、「言語運用を基盤とする言語情報学拠点」の研究論文集である 『言語情報学』のシリーズの第3巻として刊行されるほか、内容と装丁を新たにしてオランダのJohn Benjamins社からも出版される予定です。



国際会議ポスター





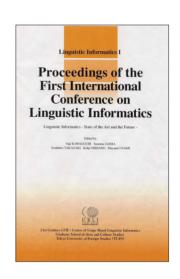
出版物

研究論文集

言語情報学 I

Proceedings of the First International Conference on Linguistic Informatics

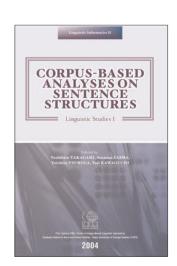
Edited by Yuji KAWAGUCHI, Susumu ZAIMA, Toshihiro TAKAGAKI, Kohji SHIBANO, Mayumi USAMI 2003年10月刊行



言語情報学 I

Corpus-Based Analyses on Sentence Structures

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

- 1. IPAモジュール
- 2. 発音モジュール
- 3. 会話モジュール
- 4. 文法モジュール
- 5. 語彙モジュール



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

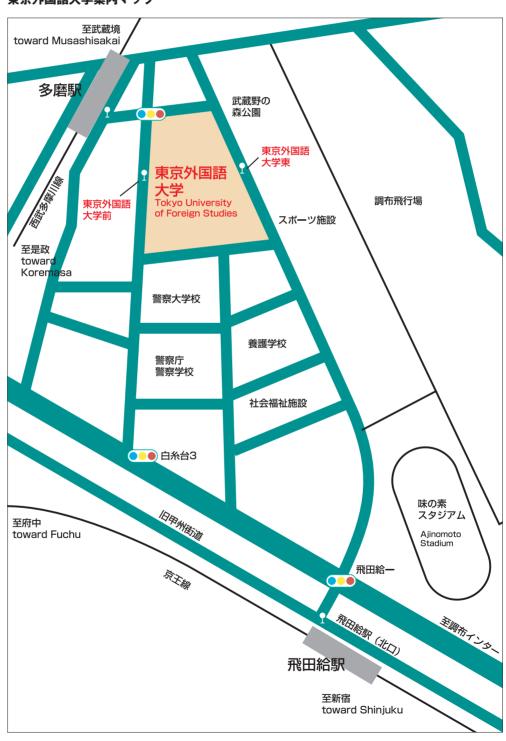
- 1. 言語学
- 2. 応用言語学
- 3. 情報工学







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)