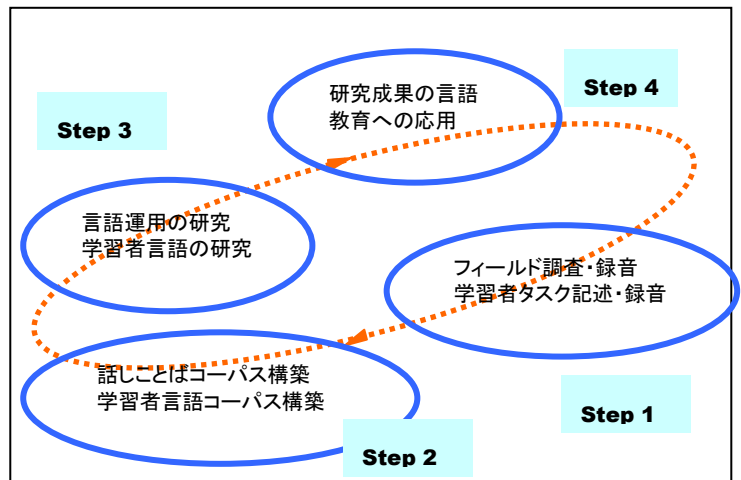


3. 研究実施計画

本研究全体のプロセスは右図のように4つの Step からなる。まず Step1 では、現地でフィールド調査を行い、話し言葉の録音を行う。書き言葉コーパスについては、デジタル化するデータを選定し、OCR 等で読み取りを行う。学習者コーパス構築では、e-learning 講義等の中で、学生にタスクを与え、学生がそれに回答する形で作文や録音を行う。次の Step2 では、臨地録音データの文字化を行い、読み取ったデータの確認を行いデジタル化する。一方、学習者タスクの回答はプログラムによりサーバーに自動的に蓄積され、研究者がそのデータベースから必要なデータを自由に取り出すことができるようにする。次の Step3 では、多言語による話し言葉および書き言葉コーパスの言語分析を行い、言語運用の実態とその特徴を研究する。また学習者言語コーパスのデータ解析を行う。最後の Step4 では、言語運用の研究と学習者言語データの分析成果をどのようにして言語教育に応用できるかを検討する。



平成 19 (2007) 年度

7月5日と10月30日に東京外国語大学において「フランス語教育の現状」「現代フランス語の中間言語音韻論」と題する講演会を開催し、ルーアン大学准教授シルヴァン・ドゥテ氏 Sylvain Detey を招き、川口とドゥテ氏がフランスの国際的プロジェクトである「現代フランス語の音韻論 PFC」の構築した音声コーパスの発音教育への応用について討論を行った。同時に PFC との連携の可能性についても話し合いをもった。このほか海野は、12月に台湾の淡江大学において「第二言語としての日本語の習得:学習者言語から見えてくるもの」の講演を行った。

コーパス研究班（日本語、中国語）では、海野が上級レベルの日本語学習者の作文データについて、構築する学習者言語コーパスの設計を行い、データの性質、タスク、データ提供者を定め、母語と学習歴の調査を行った上で、データ収集とコーパス化の作業を開始した。他方、望月は東京外国語大学日本課程の超上級学習者の作文データを学習者コーパスとして電子化し、誤用項目別に分類整理した。このデータをもとに北京大学での 2007 中日理論言語学研究国際フォーラムにおいて発表を行った。さらに、海野は『上級学習者の日本語作文データベース (2007 年度版)』を作成した。

コーパス研究班（英語）では、浦田が Written English と Spoken English における文法および語法の違いについて、コーパス言語学の成果に基づいた先行研究を渉猟し、最近の英語辞書での記述の実態について調査した。斎藤は英語学習者の語アクセントの知識と実際の発音との関係について、データ収集と分析を行い、Phonetics Teaching & Learning Conference 2007 (ロンドン大学、イギリス)で、日本人学習者

の英語発音の傾向について研究発表を行った。吉富は英語学習者の話し言葉データを収集し、文字化を行い、学習者コーパス構築のための作業を開始した。

コーパス研究班（ラオ語、クメール語（カンボジア語）、トルコ語）では、上田がクメール語の文学作品の電子化を開始した。鈴木は、現況ではラオ語コーパス資料が皆無に等しいため、2007年度はラオ語のコーパス資料の収集に努めた。川口はマルマラ大学の研究者と協力しつつ、話し言葉トルコ語のデータ収集を行った。

コーパス研究班（フランス語、スペイン語、ポルトガル語）では、高垣がスペイン語圏における語彙的バリエーションについて現地調査し、成果をマドリード自治大学で報告した。川口はディクテをタスクとするフランス語学習者言語コーパスを構築し分析を行った。分析結果は外国語教育学会 11 回大会で「フランス語ディクテーション・タスクにみられる中級フランス語学習者の音的誤答の分析」として報告された。また院生協力者の一人がフランス語学習者のスピーキングコーパスの構築に着手し、データの分析を開始した。ポルトガル語については、黒澤が学習者コーパスの構築を開始し、誤用データの収集を行った。また少数言語ミランダ語についても基礎的研究を行うとともに、話し言葉コーパスの分析を行った。

コーパス情報処理班では佐野と林が研究に着手し、「ネットワークを利用した学習者コーパス・学習者ポートフォリオ収集プログラム説明」と題する研究会を開催し、科研による今後の学習者コーパス収集の基本設計について述べると共に、フランス語をテストケースとして、オープンソースの e-learning システム Moodle の利用を検討し意見交換を行った。

平成 20（2008）年度

5 月に語学研究所において、「フランス語の言語分析と言語教育における変異」と題する講演会を開催し、セルジー・ポントワーズ大学のミュリエル モリニエ Muriel Molinié 准教授、パリ第三大学の マリ・アニック モレル Mary-Annick Morel 教授を招き、講演と討論を行った。12 月 3 日にはメーヌ大学より、ジャン・フランソワ ブルデ Jean-François Bourdet 准教授を招き、「オンライン学習と教育、新たな挑戦、その解決策とは？」と題する講演会を実施した。さらに 12 月 19 日には、トルコ共和国よりトルコ語コーパス研究の専門家 4 名、中近東工科大学より シュクリエ・ルヒ Şükriye Ruhi 教授とデリヤ・チョカル・カラダシュ研究員 Derya Çokal Karadaş、メルシン大学からはイエシム・アクサン Yeşim Aksan 准教授とムスタファ・アクサン Mustafa Aksan 教授を招聘し、トルコ語コーパス研究に関する国際ワークショップを語学研究所で開催した。第1部はトルコ語における談話標識をめぐってであり、第2部はトルコ語の言語コーパスがテーマであった。このワークショップの成果論文は、Global COE プログラム「コーパスに基づく言語学教育研究拠点」の研究報告集である『コーパスに基づく言語学教育研究報告 3』（2009）に掲載された。

2008 年度の各研究班の研究成果は以下のものであった。

コーパス研究班（日本語、中国語）では、海野が上級、超級レベルの日本語学習者の作文データを収集し、学習者コーパスとして電子化した。他方、海野は、単独あるいは他の研究者と共同で、釜山外国語大学で開催され

た第7回日本語教育国際研究大会、トルコで開催された第13回ヨーロッパ日本語教育シンポジウム、さらにシンガポールで開催された Center for Language Studies International Conference において、積極的に研究発表を行った。望月は、2008年度日本課程留学生1年生の文章表現および留学生2年生の口頭表現Ⅱの授業において、一年を通じて収集された日本語学習者コーパスを全て電子化し、1) 執筆者別、2) 誤用項目別のデータベースを作成した。このデータベースをもとに、「中国語を母語とする上級日本語学習者によるヴォイスの誤用分析ー中国語との対照からー」を執筆し、東京外国語大学論集 78号に掲載した。また上級日本語学習者のための日本語作文副教材も製作し始めた。中国語については、分担者の林が中心になり、院生協力者と共に台湾の淡江大学外国語学部日本語学科の研究者らの協力を得て、2008年12月に学部・大学院生たちの自然談話を録音した。

コーパス研究班(英語)では、浦田は Written English と Spoken English における文法および語法の違いについて、既成のコーパスを用いて分析に着手した。斎藤は、①英語学習者の語アクセントの知識と実際の発音との関係を調査すべく、データ収集と分析を行った。また International Journal of Speech-Language Pathology に Isao Ueda との共著論文 "On the production and knowledge of tonic misplacement by Japanese learners of English" を投稿した。さらに 2009年1月に神戸大学で開催された「第4回プロソディーと情報構造に関するワークショップ」に参加し知見を得た。吉富は 2007年度に引き続き、英語学習者の話し言葉データの収集・文字化を継続し、学習者コーパス構築のための作業を行った。学習者話し言葉データには、英語圏からの帰国子女の story telling データとインタビューデータ、海外在住経験のない高校生英語学習者のインタビューデータ、大学生の story telling データと自由会話データおよび内省データが含まれる。これらのデータの一部を分析した学習者言語の基礎研究にも着手した。

コーパス研究班(ラオ語、クメール語、トルコ語)は以下のような研究を行った。現況では、ラオ語コーパス資料が皆無に等しいため、鈴木は 2007年度と同様、今年度もラオ語のコーパス資料の収集に努めた。具体的にはラオ語の現代小説を数編データベース化した。またラオスへ赴き、話し言葉コーパス収集のための現地での研究協力体制の確立と次年度以降のフィールド調査の内容について、ラオス側と具体的な協議を行った。上田は固有の文字を用いるクメール語のコーパス収集の可能性を探るため、資料の整備として、文字化された文学作品の一部を電算化した。川口も前年度に続いてマルマラ大学の研究協力者に臨地録音および文字化を依頼し、トルコ語話し言葉コーパスの拡充を図った。

コーパス研究班(フランス語、スペイン語、ポルトガル語)では、高垣が前年度に引き続き、スペイン語圏における語彙的バリエーションについて現地調査を行った。川口は、ディクテをタスクとするフランス語学習者言語コーパスの拡充を継続するとともに、ルーアン大学から早稲田大学の客員准教授となったドゥテ氏とともに、フランス語学習者の中間言語音韻論研究のための枠組みの検討を開始し、3月に最初の検討報告書を作成した。これを受けて、パリで12月に開催された Journée IPFC2008 Paris にて、ドゥテ氏とともに国際プロジェクト「現代フランス語の中間言語音韻論(IPFC)」を立ち上げた。黒澤は、ポルトガル語学習者コーパスを整備し、誤用データの収集を行うとともに、少数言語ミランダ語についても基礎的研究を継続した。

さらにコーパス研究班は、本科研により収集した話し言葉データを Global COE に提供し、Tree Tagger を用いて品詞タグを付与し、最終的にフランス語およびスペイン語話し言葉コーパスの品詞検索ページを開発することに

した。品詞タグの校正には多くの学生アルバイトが動員された。

コーパス情報処理班は、林を中心にオープンソース e-learning システム Moodle を利用した学習者言語コーパスの構築システム Vers. 1.0 を開発した。6 月には語学研究所において、Moodle を用いたコーパス収集と音声録音方法に関するデモンストレーションを行い、研究協力者と当該システムの利用可能性について議論した。

平成 21 (2009) 年度

10 月に *Multilingualism and Codeswitching in Netherlands and Canada* のテーマでワークショップを開催した、会議にはオランダのティルバーグ大学よりアド・バクス Ad Backus 教授を招聘した。バクス氏は多言語併用とコードスイッチングの研究において、話し言葉コーパス資料に基づいた研究の役割について報告した。他にも神田外語大学から矢頭典枝専任講師、神奈川大学から時田朋子助手が参加し、カナダにおけるバイリンガル公務員のメール資料、バイリンガルの子供の言語コーパスについて興味深い発表が行われた。このワークショップの成果論文は、Global COE プログラム「コーパスに基づく言語学教育研究拠点」の研究報告集である『コーパスに基づく言語学教育研究報告 5』(2010)に掲載された。

2008 年度の各研究班の研究成果は以下のようであった。

2010 年 3 月に東京外国語大学で行われた外国語教育学会の 2009 年度シンポジウムにおいて、本科研の分担者 3 名が、それぞれ「英語の発音一意識と無意識」、「ポルトガル語教育における発音指導」、「ラオス語教育における発音指導要領の紹介」と題する報告を行い、本科研の成果の言語教育への応用について考察した。

コーパス研究班(日本語、中国語)では、海野は日本語学習者言語コーパスの拡張を行うとともに、オーストラリアで行われた第 8 回日本語教育国際研究大会において、e-ラーニングによる日本語学習における初級学習者への意識調査と外国人児童生徒の日本語支援に関する研究報告を行った。また院生協力者が、産出テストを用いて中級日本語学習者コーパスに見られる語彙的コロケーションの分析を行った。望月は、日本語学習者言語コーパスの拡張を行うとともに、中国語母語話者の誤用分析を実施し、その成果を台湾の東吾大学で行われた 2009 年日語教学国際会議の基調講演において研究発表した。12 月には台湾での第九屆世界華語文教學研討會において、中国語母語話者の誤用と効果的な教授法に関する講演を行った。中国語の話し言葉については、前年度に林が院生協力者と共に台湾の淡江大学外国語学部日本語学科の協力を得て収録した自然談話の文字化を行った。この台湾中国語話し言葉コーパスは約 13 万語に達する。

コーパス研究班(英語)では、斎藤がこれまで収集したデータをもとにして積極的に国際会議発表を行った。4 月にトルコで開催された国際会議 *Mind-Context Divide: Language Acquisition and Interfaces of Cognitive-Linguistic Modules* で、8 月にイギリスで行われた *Phonetic Teaching and Learning Conference* で、さらに 2010 年 3 月、セルビアでの *The 2nd Belgrade International Meeting of English Phoneticians* において、日本人学習者に特有の発音に関して研究報告を行った。

コーパス研究班(ラオ語、クメール語、トルコ語)では、鈴木が前年度に引き続き、ラオ語コーパス資料の収集を行った。また現地協力者の援助を得て、フィールド録音を行い、データベース化して話しことばコーパスも構築した。また鈴木は 7 月に、語学研究所でラオ語の方向動詞について報告し、12 月のラオス出張においても本

科研成果の一部をセミナーで発表した。上田は、2007-08 年度に電算化した資料について、キーワードの指定による検索可能性および信頼度等を検討した。

川口は 12 月にイスタンブールに臨地録音に出かけ、8 時間 51 分の自然談話を収録し、文字化を行い、約 76000 語のコーパスを作成した。同時にボアジチ大学において、言語規範と言語変異に関する招待講演を行った。7 月にはトルコ語の話し言葉コーパスのページが、Global COE「コーパスに基づく言語学教育研究拠点」のサイト内に公開された(http://cblle.tufts.ac.jp/multilingual_corpus/tr/)。公開された話し言葉コーパスは、21 世紀 COE「言語運用を基盤とする言語情報学拠点」から継続して構築されてきた言語コーパス(2005 年と 2007 年)であり、総語数は約 16 万語である。その後も 2008 年から 2010 年まで着実にコーパスが拡充され、現在の総語数は約 33 万語に達している。

コーパス研究班(フランス語、スペイン語、ポルトガル語)では、4 月にフランス語およびスペイン語の話し言葉コーパスの品詞検索ページが、Global COE「コーパスに基づく言語学教育研究拠点」のサイト内に公開された(<http://cblle.tufts.ac.jp/tag/fr/index.php?menulang=ja>, <http://cblle.tufts.ac.jp/tag/es/index.php?menulang=ja>)。公開された話し言葉コーパスは、21 世紀 COE「言語運用を基盤とする言語情報学拠点」から継続して構築されてきた言語コーパスであり、フランス語は約 22 万語、スペイン語は約 16 万語を擁する。また品詞タグに見られる誤りは、その後も随時修正され続けている。

川口は前年度から始まった IPFC プロジェクトの仲間とともに、スイスでの Colloque AFLS 2009、フランスでの Colloque Variétés, Variations & Formes du Français において共同研究の報告を行った。また 12 月の Journée IPFC 2009-Paris において、日本人学習者によるフランス語の鼻母音に関する分析を発表したが、同発表は英文で Global COE プログラム「コーパスに基づく言語学教育研究拠点」の研究報告集である『コーパスに基づく言語学教育研究報告 5』(2010)に掲載された。さらにフランス語学習者言語コーパスとポルトガル語の習者言語コーパスについては、11 月に東京学芸大学で行われた外国語教育学会第 13 回大会において、院生協力者による 2 つの研究発表がなされた。

コーパス情報処理班の林は、法政大学で開催された日本 e-Learning 学会の国際シンポジウムにおいて、川口と共に Web 教材開発に関する研究発表を行った。また林と佐野は、インターネット環境において言語の研究や教育に利用可能となる音声の録音および再生を行い、それをサーバー上にアップロード可能となるツールを開発し、Moodle と併用できるようにした。このツールによって、第二言語の音声コーパス構築が可能となり、繰り返しタスクによる録音、語学教材利用時の繰り返し録音・再生等を行うことができるようになった。

平成 22(2010)年度

最終年度となるため、3 年間の研究実績の整備と最終成果公開を中心に研究を行った。話し言葉コーパスおよび書き言葉コーパスに関しては、過去 3 年間は、電子コーパス構築を継続しながら、それぞれの言語がそれぞれ固有のテーマで研究を行い、国内外の学会・研究会で発表を行ってきた。もちろん最終年度もそうした研究活動は継続するが、研究分担者が共通するテーマで最終成果を残すことが重要であるとの認識から、「語順」をテーマとして言語研究論文をまとめることになり、2011 年 5 月の段階で、

英語、スペイン語、トルコ語、ラオ語、クメール語、中国語、日本語について研究論文が執筆され、Global COE プログラム「コーパスに基づく言語学教育研究拠点」の研究報告集である『コーパスに基づく言語学教育研究報告 7』(2011)に掲載された。

学習者言語コーパスの最終成果としては、2011年3月16日と17日に、東京外国語大学において国際会議 The TUFs G-COE International Symposium on Learner Corpora を予定していたが、3月11日に発生した東日本大震災と東京電力福島第一原子力発電所の事故によって、国際会議は止むおえず中止となった。しかしながら、この国際会議で発表される予定であった諸研究を論文としてまとめ、John Benjamins 社から2012年の春に刊行予定である。

最終年度の各研究班の研究活動は以下の如くであった。

コーパス研究班(日本語、中国語)では、上級学習者の新たな日本語作文データを収集し、電子化した。また2009年度までに収集した作文データを電子化し、学習者番号と作文番号からなるデータ番号を付与し、コーパス化する作業を継続して行った。これらのデータをまとめて『上級日本語学習者の作文データベース 2010年度版』としてCD化した。さらに、これまで収集したデータから、上級学習者にみられる誤用の分析を試みた。院生協力者は、とくに中級日本語学習者コーパスに見られる動詞「する」のコロケーションを分析し、海野と共著で、Global COE プログラム「コーパスに基づく言語学教育研究拠点」の研究報告集、『コーパスに基づく言語学教育研究報告 7』に論文を投稿した。他方、望月は、日本語学習者言語コーパスの拡張を行うとともに、中国語母語話者の誤用分析を実施し、5月にアメリカで開催された International Association of Chinese Linguistics において研究発表し、中国でも院生協力者と共に研究報告を行った。望月はまた、院生協力者の助けを得て、オンラインの日本語誤用コーパス辞典の研究も行い、7月に台湾で行われた世界日本語教育大会で成果報告をした。

コーパス研究班(英語)では、浦田がイギリスの英語学者デイヴィッド・クリスタル David Crystal の一連の「パフォーマンス講演」を話しことば資料の一種とみなして、関係代名詞の用法について調査した。その結果、クリスタルの何種類かの文章における場合を原稿を読み上げるタイプの講演の場合と比較したところ、同一人であっても、使用域によって関係代名詞の用法に相違が見られる様子を実証することができた。斎藤は、日本人英語学習者の英語イントネーション核の誤配置に関して、様々な角度から分析と考察を行い、成果を学会発表すると同時に、2007年より行ってきた個々の研究をまとめて全体像を把握する論文を執筆した。吉富は、日本語母語話者による英語学習者言語データ、特にナラティブの音声データを分析した。その結果、英語母語話者と英語上級学習者との間の formulaic sequences の使用や談話特性の相違が、言語使用の正確さ・流暢さ・自然さの差に影響していることを検証し、結果を論文にまとめた。また、この結果を踏まえて英語教育において、コミュニケーション能力を高めるために考慮すべき点を検討した。

コーパス研究班(ラオ語、クメール語、トルコ語)では、2010年度、鈴木が新たに話し言葉に近い映画やテレビドラマのシナリオを入手し、実際の映像と比較した後にコーパス資料のデータベース化を行った。その後、タイやラオスの言語学者と協議して有効なテキスト検索の方法を検討した。また、蓄積されたコーパス資料を使用してアスペクト辞や動詞連続の特徴などについて口頭発表や論文を執筆した。上田は、クメール語のコーパス収集の可能性を探るため前年度までに電子化したクメール語の文学作品をもとに、テンスとアスペクトを示すとされる語に

ついて用例を収集し、検討を行い、成果を論文にまとめた。また、キーワード指定による検索可能性、信頼度等の検討も継続して行った。川口は前年度に収録し電子化したトルコ語話し言葉コーパスについて、コーパス内のタグ情報を頼りに、述辞に後続する要素を網羅的に調査し、結果を論文にまとめた。

コーパス研究班(フランス語, スペイン語, ポルトガル語)では、高垣はスペイン語の動詞 **decir** を含む文の語順を分析し論文にまとめた。黒澤はポルトガル語教育における発音指導の問題を学会報告し、学会誌にその内容が掲載された。川口はフランスの PFC の推進メンバーと話し合いをもち、IPFC のサイトを東京外国語大学のサーバー内に開設し、これを中心にして世界各地のフランス語学習言語研究を行っていくことにした。こうして 10 月に IPFC のサイト <http://cbllle.tufts.ac.jp/ipfc/> が公開された。IPFC のプロジェクトはフランスの PFC プロジェクトを中心にして、本科学研究費と Global COE の双方が協力していくことになった。こうして 12 月の *Journée IPFC2010-Paris* において、院生協力者が Moodle を利用した学習者言語データの収集方法を海外研究者を前に説明し、川口が IPFC のサイトについて解説した。IPFC は 2011 年 3 月 15 日に *Journée IPFC 2011-Tokyo* を東京外国語大学で行う予定であったが、これも 3 月 11 日に発生した東日本大震災と東京電力福島第一原子力発電所の事故によって中止となった。ただ発表の PPT 原稿は IPFC サイト内に公開された。

フランス語話し言葉コーパスを利用した語彙およびリエゾンに関する 2 つの研究が、5 月の日本ロマンス語学会第 48 回大会において、院生協力者によって発表された。さらにフランス語学習者コーパスの母音を分析し、その結果を 11 月の外国語教育学会第 14 回大会で院生協力者が発表した。川口は 12 月の招待発表で、フランス語と日本語を例に挙げて、言語の史的研究における言語コーパス分析の重要性について考察した。

コーパス情報処理班は、前年度に構築された Moodle を利用した学習者言語の録音システムを運営し、コーパス構築と管理を進めた。林は本システムの概略と研究現状を *California Language Teachers Association Conference 2011* で報告した。