

国立国語研究所学術情報リポジトリ

世界の言語研究所 (22) アジア・アフリカ言語文化研究所 Research Institute for Languages and Cultures of Asia and Africa(ILCAA) (日本)

メタデータ	言語: jpn 出版者: 公開日: 2019-03-25 キーワード (Ja): キーワード (En): 作成者: メールアドレス: 所属:
URL	https://repository.ninjal.ac.jp/records/2206

アジア・アフリカ言語文化研究所

Research Institute for Languages and Cultures of Asia and Africa(ILCAA)

(日本)

峰岸 真琴

(東京外国語大学アジア・アフリカ言語文化研究所)

1. 設置目的

アジア・アフリカ言語文化研究所(AA研と略称)は、1964(昭和39)年に東京外国語大学に附置された、わが国で初めての人文社会学系の共同利用研究所である。この設置は、第二次世界大戦後、アジア・アフリカ諸国との相互理解、相互協力を進めることが、日本の将来にとって重要であるとの認識が深まり、1961(昭和36)年に、これらの地域についての研究を進める共同利用研究所を設立すべきであるという、日本学術会議からの政府への勧告を承けてのものであった。共同利用研究所としてのAA研の目的は、国内外の研究者とともにアジア・アフリカ地域の言語文化に関する共同研究を行い、日本あるいは世界における人文・社会科学の研究の進展に寄与することである。

発足以来、AA研は言語学・歴史学・民族学の三つのディシプリンを柱として研究を進めてきた。その理由は、アジア・アフリカ諸地域を深く知るためには、まずは現地のことばの理解が大前提であるが、単にことばの理解に留まらず、その背景となる民族、地域の歴史・文化理解を進めることが不可欠であるという認識があったからである。これら三分野の共通点は、フィールドワークに基づく言語記述、文書調査、文化記述といった一次資料の収集と分析という研究方法にある。

2. 組織の沿革

設立以来、フィールドワークに基づく一次資料の収集というAA研の研究手法上の特色は堅持されてきたが、近年のマルチメディア技術に代表される情報学の急速な進展は、言語文化情報の統合による新たな研究上の展開を期待させるものでもあった。そこで1991年には研究体制の抜本的見直しをおこない、従来の小部門制から「言語文化基礎」、「言語文化情報」、「広域言語文化第I」、「広域言語文化第II」の四大部門制をとることとなった。

1995(平成7)年度には、文部省から「卓越した研究拠点(COE, Center Of Excellence)」に指定され、「中核的研究機関支援プログラム」のもとで、設備の充実、国際シンポジウムの開催、研究資料のデータベース化とその発信などにつとめてきた。

さらに、めざましい情報ネットワークの技術革新を活用したアジア・アフリカの言語文化資料

の情報資源化をめざし、文部省令によって1997(平成9)年度から附属情報資源利用研究センター(IRC, Information Resources Center)を設置している。同センターは、アジア・アフリカの言語文化に関する情報資源の蓄積・加工・公開と、それを活用した共同研究手法の開発、国際学術交流の推進を目的としている。ここで強調しておきたいことは、IRCは言語文化情報のコンテンツ化を推進するためのもので、情報工学的研究を目的とするコンピューティングセンターではない、ということである。文字処理や音声処理といった言語学に近い分野では、情報のコンテンツ化は比較的容易である。IRCでは、言語に関するコンテンツ化を牽引車として、歴史学、民族学分野で蓄積されてきた情報の資源化を促進し、さらにはこれらを統合する形での情報加工、公開法の研究を進めている。同センターの活動については、以下をご参照いただきたい。

<http://irc.aa.tufs.ac.jp/>

2004(平成16)年4月の国立大学法人化に際しては、国立大学法人東京外国語大学の附置研究所として位置づけられたが、法人の一部局として大学の学術研究を担い、博士後期課程を中心とした大学院教育に貢献するだけでなく、国際的な人文社会学研究をリードする共同研究拠点としての役割をさらに強化していくことが期待されている。そこで2005(平成17)年には、従来の大部門制を廃し、「情報資源戦略」、「言語動態」、「コーパス」、「文化動態」、「政治文化」の五つの研究ユニットからなるプロジェクト研究部を設置するとともに、フィールドサイエンス企画研究センター(FSC, Field Science Center)を新たに設置し(2006年度に正式発足)、臨地研究という視点から、国内外の一線の研究者を糾合した研究および研究企画を行っている。

3. 言語研究上の特色

ここではAA研の研究のうち、言語学分野を中心とした研究について紹介する。設立当初のAA研の具体的目標は、研究未開発の言語の記述に基づいて基礎語彙集や辞典を編纂し、さらにこれらの言語の教材を開発し、ネイティブスピーカーと日本人教師との共同による言語研修を行って、アジア・アフリカ地域に関わるさまざまな分野の研究者を養成することであった。言語の記述研究および基礎語彙集、辞典などの出版物については、以下をご参照いただきたい。

http://www.aa.tufs.ac.jp/editcom_j.html

3.1. 記述言語学と基礎語彙集

AA研の言語研究の特色は、設立以来一貫してフィールドワークに基づく個別言語の記述研究を主眼とする点にある。もちろんその時代時代の記述研究のあり方は、言語学が言語についてどのような関心を持ってきたかを反映している。例えば、AA研設立期の日本の言語学者の多くは、日本語の起源はどこにあるかという関心を共有していた。このためAA研においても、日本周辺の東アジア、東南アジアの言語を手始めに、共通の言語調査票に基づく語彙調査を行い、語彙比較に基づく系統論的な研究を行うための記述研究が進められてきた。1970年代以降には、橋本萬太郎所員が、東アジアの諸言語に系統を超えて存在する言語類型上の特徴から、この地域の諸言語の歴史的な変化の方向性を明らかにしようとする「言語類型地理論」を提唱し、共同研究

を組織した。言語類型地理論の提起する「地理的分布から言語史を考える」という方法論は、東アジアおよびその周辺の諸言語の記述研究を行う際の大きな刺激となった。さらに近年の研究は、これまでの記述研究の実績に基づいて、通言語的な比較・対照を行い、さらに言語類型論を始めとする理論的研究へと発展させる方向性を見せているが、それはあくまで個別言語事実に立脚した上での理論化、一般化を目指すものであり、特定理論に依拠して、そこから個別言語の現象をトップダウンに見渡すというものではない。

フィールドワークにおける個別言語の語彙収集・分析の成果をまとめたものが基礎語彙集のシリーズとして刊行されている。基礎語彙集の多くは、AA研で編纂された『アジア・アフリカ言語調査票』（上 1966, 下 1967）を手がかりとして用いている。AA研の設立当初、日本にはアフリカの言語を専門とする言語研究者は存在しなかった。本研究所在が若手の言語研究者をアフリカ各地に派遣し、一次資料を収集したことによって、日本におけるアフリカ諸言語の研究の基礎が築かれたと言っても過言ではない。特に、バントゥー諸語の音韻、形態、文法上の特色である声調の研究は、アクセント言語である日本語の話し手としての日本の言語学者が、その記述と分析に強みを発揮する分野である。

3.2. 言語研修

本研究所在設置当初からの事業である「言語研修」は、アジア・アフリカ諸言語の科学的記述研究に基づいて編纂された文法書、語彙集を用いて、話し手とじかに接する機会のまれな言語を短期集中で教授することで、多くの研究者を養成してきた。この事業を通じて育成された国内各地の研究者と共同してアジア・アフリカの総合的研究を進めてきたことは、本研究所在の実績である。これまでの研修言語と教材については、以下をご参照いただきたい。

http://www.aa.tufs.ac.jp/aa03_gengo_j.html

<http://www.aa.tufs.ac.jp/book/publist/training.pdf>

特に、アフリカの諸言語を始めとして、サハ（ヤクート）語、マダガスカル語、ネワール語、パシュトゥー語、ムンダ語の研修は、AA研で初めて開講され、あるいは他では開講されないという点で、教材開発と研修事業の両面で貴重なものである。

3.3. 辞典編纂

辞典の編纂は地味な研究分野であるが、言語文化知識の集大成としての辞典の価値は揺るぎないものである。とりわけ、坂本恭章(2001)『カンボジア語辞典』（全3巻）、坂本恭章(1993)『モン語辞典』は、AA研の東南アジア言語学における記念碑的な業績である。近年では、南アジア言語学における町田和彦(1997)『ヒンディー語動詞基礎語彙集』、星泉(2003)『現代チベット語動詞辞典（ラサ方言）』のように、基本動詞の用法に焦点を絞ったものが出版されている。

3.4. 文字情報学

辞典の編纂には、多言語の処理と混植による組み版の技術が不可欠である。特に、アジア・ア

フリカ地域には、アラビア系、インド系、漢字といった固有の文字が用いられており、これらの言語の辞典を編纂するためには、単に複数の文字体系が混植可能であるばかりでなく、文字の構成法の分析、入力方法、各言語固有のソーティングのアルゴリズム、ハイフネーション、ジャスティフィケーションなどの組み版に関する知識と技術の蓄積が必要である。AA 研では1978年にメインフレームコンピュータを導入するなど、早くからインド系文字データの入力、処理、出力に取り組み、多言語処理システムを開発していた。上記のカンボジア語辞典、モン語辞典の刊行に見られるように、東アジアに留まらず、東南アジア、南アジア、西アジアの文字処理を行い、KWIC (Key Word In Context) 索引を用いての言語分析や印刷、組み版、出版へと応用してきた。

このような辞典や文法書の編纂によって蓄積された AA 研の言語学および文字学の実績を基に、中核的研究拠点形成プログラムとして「アジア書字コーパスに基づく文字情報学の創成」(GICAS, Grammatological Informatics based on Corpora of Asian Scripts, 2001(平成13)年度～2005(平成18)年度)が採用され、文字の運用コーパスに基づいた文字情報学の分野において先導的役割を果たすことになった。このように、現地からの一次資料の分析、データベース化から一般化、理論化を目指し、さらには辞典を始めとする具体的な成果物への応用を目指すことが、フィールドワークから情報化へという AA 研の研究戦略として定着しつつある。

4. グローバル COE プログラム

4.1. 大学院後期課程と21世紀 COE プログラム

AA 研が附置されている東京外国語大学には、大学院地域文化研究科博士後期課程が1992年に設置された。AA 研所員の多くが後期課程教育に加わり、次世代の研究者の育成にも取り組んでいる。2002(平成14)年度には、大学院地域文化研究科を主体とする21世紀 COE プログラムとして、人文科学の領域で「言語運用を基盤とする言語情報学拠点」が、複合領域で「史資料ハブ地域文化研究拠点」がそれぞれ採択され、本研究所の大学院担当教員も、これらの拠点の形成に参加・寄与してきた。

2006(平成18)年度に完了した21世紀 COE プログラムの実績を踏まえ、さらなる教育研究上の展開を図る拠点形成のための「グローバル COE プログラム」として、大学院地域文化研究科と AA 研との連携によって教育研究プログラム「コーパスに基づく言語学教育研究拠点」(拠点リーダー：峰岸真琴)を申請し、2007(平成19)年度からの5カ年の計画として採択された。

4.2. グローバル COE 拠点形成の目的

グローバル COE プログラムは、博士後期課程の教育研究のプログラムを充実させることで、各大学の特色となる研究分野において国際的かつ先端的な教育研究拠点を形成すると同時に、当該分野の教育研究における後継者の育成を目的とするものである。

東京外国語大学の最大の特色は、諸言語の基礎研究と第二言語教育の実践に基づいた応用研究・開発にあることは言うまでもない。そこで、本学のプログラムは、コーパスに代表される言

語運用データの収集、分析と、それに基づく実証的な研究、その成果の第二言語教育分野への応用開発に関する教育研究をいっそう充実させることを目的として、さまざまな教育研究プログラムを企画している。

本学には、日本語を含む26の専攻語がある。これに加えて、世界の広範な地域で話される数多くの言語に関する研究が行われている。これらの言語に対する実証的な研究手法は、現地調査による一次資料の収集と分析を行うフィールド言語学的アプローチ、文字や音声言語をデジタルデータ化し、一定の構造を与えられた言語データの集合としてのコーパスの構築とその分析によるコーパス言語学的アプローチ、および情報工学の手法によって、コーパス分析から得られた知見を教材開発、教育手法の開発など第二言語教育の高度化に応用する言語情報学的アプローチの三つのアプローチに大別される。三つのアプローチは互いに深い関連を持つが、とりわけ言語運用の実態を把握するために不可欠のコーパス構築とその分析は、今後本学の教育研究の基盤となるべきものとして位置づけられるため、「コーパスに基づく言語学」の拠点形成と銘打った。

グローバル COE プログラムの研究戦略策定に際しては、本学の教育研究対象である諸言語を、言語資料のデジタル化の進展の度合いを基準として、大言語、中言語、小言語の三つのグループに分類した。デジタル化された言語資料を代表するものは、大規模コーパスおよび電子辞書である。世界の諸言語は、これらの言語資料を豊富に持つかどうかを基準に分類することができる。またこれらデジタル化の進展の度合いによって、教育研究上の目標設定も異なってくる。

ここでいう大言語とは、国家規模でのコーパス構築が行われ、市販の電子辞書が手軽に入手できる言語で、英語、フランス語、ドイツ語、朝鮮語、それに日本語がここに加えられる。中言語とは、国家規模でのコーパスあるいは市販の電子辞書のいずれかが存在しない言語である。本学の専攻語で言うと、中国語、ロシア語、ヒンディー語といった使用人口の多い言語から、西アジア、南アジア、東南アジア、東北アジアの諸国語がここに含まれる。小言語とは、国家規模でのコーパスも市販の電子辞書も存在しない言語で、少数民族の言語がその典型である。

4.3. 拠点の人材育成の目標

大学院教育における人材育成の観点から見ると、大言語については言語研究者および言語教育の専門家の両面において人材の需要が高い。本学大学院の院生の多くは、ここでいう大言語の研究、教育を専攻している。中言語についても、言語研究者および言語教育の専門家に対する一定の需要が存在する。小言語については、言語研究者の需要は存在するが、言語教育の専門家への社会的需要は現在のところ少ない。

言語研究上の価値の観点から見ると、大言語については、その言語そのものの研究対象としての価値に加えて、英語教育や外国語としての日本語教育の場合のように、教育法の向上を目的とした、第二言語習得研究にも大きな価値がある。中言語については、例えば上述のカンボジア語のように、文字レベルでデジタル化されて刊行された辞典が存在する場合もあるが、構造化された電子辞書の開発には至っていない。このような言語の電子辞書を開発すれば、教育上、研究上の強い潜在的需要が存在する。小言語については、言語の多様性や普遍性の研究に寄与する点に

において、その研究上の価値は高く、少数民族の言語文化の継承の意味で、コーパスを作ることに文化的意義が認められる。しかし、大規模コーパスを作ることは現実的ではない。

このような社会状況においては、一見して小言語の専門研究者に対する社会的需要に期待できないのではと危惧されようが、必ずしもそうとは言えない。小言語は、その地域の共通語である大言語と使い分けられる場合が多く、小言語の研究者は、少なくとも初期の段階では当該地域の大言語を媒介として研究を開始し、その研究成果の公開も、英語ならびに地域の大言語を通じて行う必要がある。また近年の情報化の進展は、フィールドでの成果をどのようにコーパス化して言語分析を行うかといった研究手法の高度化をもたらしただけでなく、音声コーパス、電子辞典の開発を通じて、現地での少数民族文化の維持、発展に貢献するなど、研究成果の地域社会への還元が強く求められている。

この意味で、フィールド言語学、コーパス言語学、言語情報学の三分野の手法をバランス良く習得した人材の育成が社会的に期待されている。先行する21世紀 COE プログラムにおいても、ある地域の大言語の教育職に就きながら、研究においては当該地域の大言語の諸方言や少数民族の言語文化を対象とし、その成果を教材化して現地社会に還元し、また研究成果を国際学会で英語で発表する人材が実際に育成されたという実績がある。グローバル化の進展に伴う地域社会と文化の存続が危惧される現代社会においてこそ、言語に代表される文化の多様性を深く理解し、教育研究に携わることのできる人材が求められているのである。

また、英語や中国語といった国際的に使われる大言語は、実際は均質のものではなく、多様な地域・社会方言からなり、運用の現場ではそれぞれが使い分けられ、それぞれの地域の歴史と文化を反映している。研究者として英語、地域共通語、方言、少数民族語に通じ、フィールド言語学、コーパス言語学、言語情報学の手法を身につけた人材を輩出することが、本プログラムの人材育成戦略の大きな目標である。とりあえずモノリンガルでも暮らしていける日本にあっては、このようなマルチリンガル、マルチディシプリンの人材育成は困難に思えるが、本学の大学院生は、元来外国語に関心を持って入学し、在学中に複数の言語を習得する資質を持ち、そのための努力を惜しまない。このような才能に恵まれた人材をいっそう充実した情報環境に置き、従来専攻言語ごとに蓄積されてきた、多様な言語の処理方法についての研究成果と独自のノウハウを共有させれば、ある面では工学の専門研究者よりも有利な立場で研究を進めることも可能である。グローバル COE プログラムによって、大学院生の能力を余すところなく開発する教育研究プログラムを充実させることで、さらに有能な教育研究の人材を育成できると期待される。

4.4. 研究計画の概要

本プログラムでは、コーパス化の進展という尺度による諸言語の分類と、教育研究者の社会的需要についての現状分析を踏まえて、大言語、中言語、小言語という分類に応じた研究戦略を策定した。まず、三つの研究アプローチに従って教育研究事業を行う班として、主として大言語を対象とする言語情報学班、中言語を対象とするコーパス言語学班、小言語を対象とするフィール

ド言語学班をそれぞれ組織した。

言語情報学班は、英語、日本語、フランス語、スペイン語、ドイツ語を始めとした大言語の研究と第二言語教育研究を目的とする。これらの言語では既存の大規模コーパスが容易に利用可能であり、若手研究者の関心は、主として既存のコーパスを利用しての、それぞれの研究目的に合った分析用ツールの効果的な利用法を習得し、高度な言語研究を行うこと、および分析ツールの開発、機能別コーパスや学習者コーパスの構築、これらのコーパスを利用した言語教育法や教材の開発といった応用研究に向けられている。

コーパス言語学班は、中国語、ロシア語、タイ語などの中言語を対象に、コーパスの構築及びコーパス分析のツール開発、さらには電子辞書の開発のための研究を目的とする。これら中言語の場合は、小規模のデジタル化されたテキストデータが世界各地に散在しているが、それぞれの言語固有の形態統語上の特徴に依存する分析ツールを利用する環境が整っていない。とりわけ、テキストをスキャンして形態素の解析や品詞タグの付与を行うための機械可読辞書（MRD, Machine Readable Dictionary）が存在しない言語については、形態統語論研究の基盤となるMRDの開発は必須である。現在のXML技術を活用すれば、MRDを基にして一般が利用可能な電子辞書の開発を行うことも重要な研究開発テーマとなりうる。

フィールド言語学班は、主として音韻、形態、統語類型論上のテーマを念頭に置きつつ、国内外でのフィールド調査を行い、フィールドで収集した一次資料をコーパス化する際に、どのようなデータ構造を与えるかについての研究を行う。

これら三つの班による研究は、それぞれの成果を互いに活かしながら進められる。フィールド言語学から得られる言語類型上の情報は、言語コーパスの構築においてどのような品詞や形態に関するタグ付けが必要となるかについての貴重な情報をコーパス言語学に提供する。こうして構築される多様な言語のコーパスを用いての言語研究は、言語の運用の実態を実証的に解明する手段として、言語情報学における言語教育法の開発研究へと応用される。とりわけ、非母語話者の「学習者話し言葉コーパス」の構築と分析は、言語教育法の開発研究の実証的基盤となるものとして、今後の第二言語教育法の高度化のために発展が期待される分野である。また、言語情報学の研究開発の成果は、少数民族の言語文化の保持、再生のためのマルチメディア教材や辞書の開発による現地コミュニティへの成果還元の形でフィールド言語学にフィードバックされることが期待される。

4.5. 教育プログラム

以上述べてきたように、グローバルCOEプログラムは、国際的な先端的研究拠点を形成するだけでなく、教育プログラムの充実による後継者の育成に重点を置く点にその特徴がある。本学大学院においても、これまで述べてきたような、コーパスに代表される言語運用データの実証的な分析・研究を進めると同時に、大学院博士後期課程を中心とした、さまざまな教育プログラムを企画している。

既に本学21世紀COEプログラムでも実績を上げてきたことであるが、博士後期課程の院生の

グループを中心とした自主的な研究プロジェクトの公募をいっそう充実する。フィールド言語学の実習、国内外での実地調査による言語データの収集と、その分析に基づくコーパスの構築実習、分析ツールを用いた言語分析実習、外国語教育の現場における学習者コーパスの構築を大学院生自身の企画で行わせることなどによって、調査や教育現場での実践経験を積ませることを計画している。このようにして得られた多様な言語のデータを用いた研究成果と、その分析上のノウハウが「コーパス言語学拠点」としての本学に蓄積されることで、多様な言語の個性と運用実態に応じた分析ツールの開発が可能になる。

さらに、これら若手研究者による研究成果を、国内外の学会で発表させるための支援も行う。また、webを介した論文執筆支援システムを開発して、現行の複数教員による博士論文指導体制を技術的に支援する計画もある。

大学院後期課程においては、コーパスに関わる言語学教育プログラムを充実し、「フィールドからコーパス構築へ、さらに教室での応用へ」を目指して、三つの分野の一線の研究者と大学院生が共同して講演・研究発表を行い、その成果をe-Learning化してweb上で発信することで、今後拠点の教育研究活動の全貌を見やすい形にまとめていく予定である。