

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 5 月 31 日現在

機関番号：12603

研究種目：基盤研究(A)

研究期間：2010～2012

課題番号：22240078

研究課題名（和文） Ajax ログのラーニングマイニングによる評価とWeb 会話ラボの研究開発

研究課題名（英文） Study of Learning Mining of Ajax Log and Development of Web Conversation Laboratory

研究代表者

芝野 耕司(SHIBANO KOHJI)

東京外国語大学・アジア・アフリカ言語文化研究所・教授

研究者番号：50216024

研究成果の概要（和文）：

日本語 e-ラーニングシステム JPLANG のログは研究申請時点での 530 万件が年度末には 1830 万件を超え、登録ユーザは 2 万人と倍増した。

大規模なログデータの統計解析によって、ユーザの学習障害の特定を可能とすることができ、ラーニングマイニングの解析手法を確立することができた。

また、CEFR で要求される 5 技能教育を e-Learning 環境で完全に実現する Web 会話ラボ並びに関連する教授法の研究開発を行った。教材開発のために高校教科書を中心に 7000 万語の日本語コーパスを構築した。

研究成果の概要（英文）：

When we proposed the project, at that time, the number of logs of the Japanese e-Learning system JPLANG was 5.3 million and now it exceeded 18.3 million and its registered user exceeded 20,000.

We have developed a learning mining system based on large scale log data statistical processing algorithms.

We also developed Web conversational laboratory for completing web based five skill training required for CEFR. For the future teaching materials development, we have further developed 70 million words of Japanese textbook corpus.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010 年度	10,000,000	3,000,000	13,000,000
2011 年度	15,100,000	4,530,000	19,630,000
2012 年度	8,600,000	2,580,000	11,180,000
総計	33,700,000	10,110,000	43,810,000

研究分野：総合領域

科研費の分科・細目：科学教育・教育工学・教育工学

キーワード：e-ラーニング，日本語教育，CALL，LMS

## 1. 研究開始当初の背景

研究の学術的背景

ICT 技術の進化によって、情報爆発が起きている(UCB, How Much Information? 2003

<http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>)。時代は、メガからギガを超え、既にペタ、エクサバイトの時代に移行しようとしている。産業界では、Cloud Computing 環境を構築するとともに、

この新たな Computing 環境を利用し、新たに得られた情報を活用したデータマイニングやテキストマイニングが広く普及し、Amazon や楽天、Google など、こうした状況を積極的に利用し、ビジネスの革新が起きている。しかし、教育研究分野では、こうした状況は具現してはならず、また、対応する試みもほとんどないのが現状である。

1980年代までのメインフレームの時代は、ほとんどの情報が“紙”に記録されていた。その頃は、“紙”の情報の一部である金額だけがコンピュータに入力され処理されていた。しかし、現在では、町のレストランでもPOS 端末から注文が入力され、その情報の一部だけがPOS 端末から印字され、紙に記録さ、残りの情報は、データマイニングによって企業のビジネスに活かされる時代となった。しかし、教育分野では、こうした状況には全くない。それは、e-Learning がコンピュータ化の教員側へのメリットが大きい評価の面が中心であり、学習及び教授のためのe-Learning への試みが十分ではないことによる。

一方、言語教育では、ヨーロッパ標準の“ヨーロッパ共通言語教育参照枠(Common European Framework of Reference for Language Learning, Teaching, and Assessment, CEFR, Council of Europe 2001)”をもとに、聞き、話し、読み、書くという4技能を独話と会話の二つに分け、“話す”ことを強化したコミュニカティブアプローチに重点を移した。しかし、特殊機器としてのLL が姿を消し、PC がとって変わったが、企業の撤退によってLL の代替は実現されてはいるものの、新技術を用いて、Spoken Interaction を含む5技能への対応はなされていず、ICT の進化から語学教育が取り残されておるのが現状である。

着想に至った経緯

これまでこうした問題に、研究代表者を中心に継続的な研究開発を行ってきた。21世紀COE「言語運用を基盤とする言語情報学拠点」(H14~H18)では、17言語の発音・会話教材の開発を行い、基盤(A)「協調的ユビキタス多言語運用 e-learning 環境の研究」(H14~H16) (研究代表者：芝野)では、最も学内開発の教材が充実していた日本語に教材開発を限定し、教材開発を行うとともに、上記の問題を解決するため、独自のLMS (Learning Management System) JPLANG システムの開発を行った。

JPLANG システムは、現代GP「e-日本語—インターネットで拓げる日本語の世界—」(H17~H19) (取組担当者：芝野)に引き継ぎ、日本語教育 e-Learning に的を絞った形での研究開発を行った。(http://jplang.tufs.ac.jp/参照)

基盤研究(A)、「AJAXを用いた直接操作・多言語・語学教育 e-Learning OSSの開発」(H18~H21)(研究代表者：芝野)では、ソフトウェア基盤をAjax, Ruby on Rails 及びFlash Media Server (FMS)に更新し、LL 機能を含むe-Learning システムへと大幅な改良を行った。また、FMS を利用したWeb 会議システムをJPLANG 環境に取り込むためのシステム、すなわち、Web 会話ラボの開発を行った。

教育GP「グローバル戦略としての日本語 e-Learning」(H20~H22)(事業担当者：芝野)では、現代GPの基盤に立って、海外への普及を推進するとともに、ブレンデッドラーニングと直接法による日本語教授法を「直接法で日本語を教える」(東京外国語大学出版会、2009年5月、491ページ)として、まとめ、出版した。

2006年4月の現行システム基盤への移行以降現在までで、JPLANG システムは、79か国374都市の81,325のIPアドレスから数万人以上が利用し、任意登録ユーザ数も8,000に達しようとしている。ログ件数は530万件を超え、ログは、1日約1万件増加しており、ユーザ登録は毎月約200名増えている。また、このログの中には、2006年からの4年間にわたる約250名の正規生への教育も含まれている。2009年度正規生の第1学期の利用は、約300時間の日本語授業時間に対して、学習者一人平均約65時間JPLANG を利用するようになり、成績優秀者は、JPLANG を多く利用していることが分かった。

これらの大規模なデータの蓄積とAjax による詳細なログの取得によって、ビジネス分野でGoogle やAmazon, 楽天などが活用しているデータマイニングやテキストマイニングを活用するための基盤研究を開始することが可能となった。ラーニングにおけるデータ・テキストマイニングに関する研究、すなわち、ラーニングマイニングの研究が可能となった。

また、前述のように、Web 会話ラボを開発し、これによって、CEFR で要求される5技能全部をJPLANG のe-Learning 環境で実施可能となった。しかし、まだシステム基盤は、開発されたものの、従来の一般教室やLL とは異なり、距離的な制約や時間的制約なしでの会話トレーニングが可能であり、国際共同教室も可能となった。具体的には、日本の中国語学習者と中国の日本語学習者とを組み合わせ日中共同での中国語・日本語の学習などが考えられる。また、Web 会話ラボ機能を用いた全く新しい教材開発も可能である。

## 2. 研究の目的

これまでに開発したAjax を利用した530万件以上の大量の情報(ログ)をいかに教育に

活かすのかをラーニングマイニングとして、理論、技術及び教育の側面から体系化を行うとともに、“ヨーロッパ共通言語教育参照枠(Common European Framework of Reference for Language Learning, Teaching, and Assessment, CEFR)”で要求される Listening, Reading, Spoken Production, Spoken Interaction, Writing の 5 技能教育を e-Learning 環境で完全に実現する Web 会話ラボ並びに関連する教授法及び新世代の教員養成方法の開発を行う。

また、既開発の日本語 e-Learning 環境である JPLANG に Web 会話ラボ教材の追加、LMS 拡張、海外普及及び海外サポートを強化する。

### 3. 研究の方法

研究組織は、代表者の芝野が全体を統括し、システム・評価班、日本語教材班、及び日本語教授法班の三つで構成する。各研究班の主な役割は、次の通り。

システム・評価班では、ラーニングマイニングの手法確立のため、530 万件のログを解析し、学習効果の評価及び統計解析を行うとともに、最近急速に統計手法等の環境が整ってきた R を前提に JPLANG システムへの統合を行う。

日本語教材班では、既存教材の拡充及び Web 会話ラボ教材の開発を行うとともに、日中間でのパイロット授業を実施し、その成果を教材開発に反映させる。

日本語教授法班では、「直接法で日本語を教える」で開発した初級日本語指導書に加え、中級及び上級日本語の指導書の開発を行う。

また、海外への普及活動を継続するとともに、海外へのサポート体制を拡充する。

これまでの研究開発では、合計数十人の大学院生の協力を得て開発を行ってきたが、今後も大学院言語応用専攻日本語教育学専修コースを中心に大学院前期及び後期課程の学生の参加を積極的に促すとともに、中国、タイ、イタリアなどの海外の協力大学との共同も強化する。

### 4. 研究成果

(1)これまでに開発した Ajax を利用した 530 万件以上の大量の情報(ログ)をいかに教育に活かすのかをラーニングマイニングとして、理論、技術及び教育の側面から体系化を行うとともに、“ヨーロッパ共通言語教育参照枠(Common European Framework of Reference for Language Learning, Teaching, and Assessment, CEFR)”で要求される Listening, Reading, Spoken Production, Spoken Interaction, Writing の 5 技能教育を e-Learning 環境で完全に実現する Web 会話ラボ並びに関連する教授法

及び新世代の教員養成方法の開発を行った。また、既開発の日本語 e-Learning 環境である JPLANG に Web 会話ラボ教材の追加、LMS 拡張、海外普及及び海外サポートを強化した。

【システム・評価班】 システム・評価班は、JPLANG システムの保守及び拡張を行う。ラーニングマイニング関連では、MySQL のデータベースに格納されている 530 万件のログデータを R で利用可能な形に変換した。また、日本語教材班と共同でドリルの練習問題を現在の固定型ではなくルールベースに変更し、毎回別の問題で学習者が練習を行えるように拡張するためのルール抽出を行った。Web 会話ラボについては、実際の利用に合わせて、システムの拡張及び改良を行った。

【日本語教材班】 日本語教材班は、まず Web 会話ラボのパイロット教材の開発を行った。初級日本語、中級日本語、上級日本語の改版のためシステム・評価班と共同でデータの整理を行った。海外共同授業のパイロット授業の実施の環境整備を行った。

【日本語教授法班】 中級日本語の指導書テスト版を作成した。Web 会話ラボを用いてのパイロット授業を実施した。

(2)これまでに開発した Ajax を利用した 530 万件以上の大量の情報(ログ)をいかに教育に活かすのかをラーニングマイニングとして、理論、技術及び教育の側面から体系化を行うとともに、“ヨーロッパ共通言語教育参照枠(Common European Framework of Reference for Language Learning, Teaching, and Assessment, CEFR)”で要求される Listening, Reading, Spoken Production, Spoken Interaction, Writing の 5 技能教育を e-Learning 環境で完全に実現する Web 会話ラボ並びに関連する教授法及び新世代の教員養成方法の開発を行う。また、既開発の日本語 e-Learning 環境である JPLANG に Web 会話ラボ教材の追加、LMS 拡張、海外普及及び海外サポートを強化する。

【システム・評価班】 申請時点の 530 万件のログはその後の 1000 万件増え、1500 万件を超えた。このログの分析を行うとともに、ルールベースのダイナミックなドリルの生成についての検討を行った。また、教材内容をダイナミックに変更することを可能にし、上級教材を拡充する基盤として、日本の高校教科書指導書データを収集し、大規模な教科書コーパスの構築を行なっている。

【日本語教材班】 日本語教材班は、まず Web 会話ラボの教材の開発を行った。

【日本語教授法班】 教材班、評価班と共同で海外での講習会の実施に関する検討を行った。特に、3 月にはセルビアのベオグラード大学を訪問し、セルビアで近隣のヨーロッパ及び中東地域の日本語教育関係者を集

めた講習会の実施について、ベオグラード大学と共同開催の可能性の検討を行った。

(3) 研究申請時点で 530 万件であったログは研究終了時点では 1830 万件を超え、1300 万件増加した。また、登録ユーザは 2 万人と倍増した。

これらの利用実績及びラーニングデータの蓄積を元に、大規模なログデータの統計解析によって、ユーザの学習障害の特定を可能とすることができ、ライニングマイニングの解析手法を確立することができた。今後の課題は、ラーニングマイニングの結果を学習者にフィードバックするために、学習者が必要とする追加教材を提供する手法を確立することである。幸いなことに、言語学習では、実際の言語使用データを教材とすることができる。国際標準である“ヨーロッパ共通言語教育参照枠(Common European Framework of Reference for Language Learning, Teaching, and Assessment, CEFR)”では、場面と言語機能をもとに必要とする各言語での表現を教授することによって言語教育を行うことを求め、結果として、学習者のできる能力(Can Do)を涵養すること求めている。すなわち、適切にコーパスを構築し、このコーパスから上記の CEFR で要求されるシラバスに沿った言語使用データを学習者に提示出来れば、効果的な学習が可能となる。このため、今後は、各学習者に対応したテーラーメイドエデュケーションの研究を行う必要がある。

また、CEFR で要求される Listening, Reading, Spoken Production, Spoken Interaction, Writing の 5 技能教育を e-Learning 環境で完全に実現する Web 会話ラボ並びに関連する教授法及び新世代の教員養成方法の開発を行った。

既開発の日本語 e-Learning 環境である JPLANG に Web 会話ラボ教材の追加、LMS 拡張、海外普及及び海外サポートを強化した。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 12 件)

- ① Kohji Shibano, Constructing a Corpus of Japanese Language for a Basis for Japanese Language Study and Education, Proceedings of the 5th International Conference on Computer Assisted Systems For Teaching & Learning Japanese (CASTEL-J), 査読有, Vol. 5, 2012, 1-4

- ② Hajime Mochizuki, Yuhei Nakamura, Kohji Shibano, Indonesian Shallow Stemmer for Text Reading Support System, Lecture Notes in Electrical Engineering vol.157: Recent Progress in Data Engineering and Internet Technolog, 査読有, Vol. 2, 2012, 244-255

- ③ 藤森弘子, 異文化理解を目指したインターアクション教育に関する一考察—遠隔授業での相互作用に着目して—, 東京外国語大学留学生日本語教育センター論集, 38 号, 2012, 19-31

- ④ 鈴木美加、藤森弘子、藤村知子、鈴木智美、中村彰、坂本恵、花蘭悟、伊集院郁子, 日本語学習における目標記述をめぐって—全学日本語プログラムの can-do リスト作成に向けて—, 東京外国語大学留学生日本語教育センター論集, 査読有り, 38 号, 2012, 155-166

- ⑤ 藤森弘子, 異文化理解を目指したインターアクション教育に関する一考察—遠隔授業での相互作用に着目して—, 東京外国語大学留学生日本語教育センター論集, 査読有, 38 号, 2012, 19-31

- ⑥ 伊集院郁子・大津友美, JLC1年コース作文データベースの構築, 東京外国語大学留学生日本語教育センター論, 査読有, 38 号, 2012, 125-133

- ⑦ 鈴木美加、藤森弘子、藤村知子、鈴木智美、中村彰、坂本恵、花蘭悟、伊集院郁子, 日本語学習における目標記述をめぐって—全学日本語プログラムの can-do リスト作成に向けて—, 東京外国語大学留学生日本語教育センター論集, 査読有, 38 号, 2012, 155-166

- ⑧ 大津友美, 新聞記事読解のためのストラテジーの分析&#8211;より良い新聞記事読解をめざして—, 日本 NIE 学会誌, 査読有, Vol.

- 7, 2012, 33-42
- ⑨ Hajime Mochizuki, Yuhei Nakamura, Kohji Shibano, Indonesian Shallow Stemmer for Text Reading Support System, 査読有, 2011, 1102-1107
- ⑩ Hirose, K., Kawano, S., Konishi, S. and Ichikawa, M., Bayesian Information Criterion and Selection of the Number of Factors in Factor Analysis Models, Journal of Data Science, 査読有, Vol. 9, No. 2, 2011, 243-259
- ⑪ Minako Nakamura, Yukito Kado, and Kohji Shibano, Developing a Digital Archive of the Works of Tatsue Sata, a Leading Creator of Japanese Modern Ballet, Journal for the Anthropological Study of Human Movement (JASHM), 査読有, Vol. 16, No. 1&2, 2010
- ⑫ 宮武かおり, 芝野耕司, 日本語話し言葉コーパス(CSJ)の学会講演データにおけるソーシャルダイクシス分析, 2010 世界日本語教育大会論文集, 査読有, 2010, 1362-0, 1362-9
- [学会発表] (計 10 件)
- ① 芝野耕司, Usage Based Japanese Language Education, Fource Conference on Japanese Linguistics and Language Teaching (4AIDLG)(招待講演), 2013 年 03 月 21 日~2013 年 03 月 23 日, Naples, Italy
- ② Kaori Miyatake, Kohji Shibano, Corpus-based verification of a comprehensive Japanese honorific model using social deixis, politeness, and speech act, AAAL (American Association for Applied Linguistics) 2012 (査読有), March 24-27, 2012, Boston, USA
- ③ 芝野耕司, 日本語教科書コーパス, Google N-gram, そして日本語均衡コーパス, 13th international conference of the EAJS (European Association for Japanese Studies), (査読有), 24-27 August, 2011, Tallinn, Estonia
- ④ Kaori Miyatake, Kohji Shibano, How to make claims persuasively --- corpus-based analysis of Japanese academic presentations, 12th International Pragmatics Conference, (査読有), 3-8 July 2011, Manchester, UK
- ⑤ Kaori Miyatake, Kohji Shibano, Usage-based analysis of situational functional Japanese honorifics based on large Japanese corpora, American Association for Applied Linguistics 2011 Conference in Chicago, Illinois (AAAL) (査読有), 2011-03-26, Chicago, USA
- ⑥ Kaori Miyatake, Kohji Shibano, Capturing Requirements -toward an application of an e-learning system to Japanese language educations, EuroCall2010 (査読有), 2010-09-10, Bordeaux/University of Bordeaux, France
- ⑦ 藤村知子, 芝野耕司, e ラーニングシステム JPLANG を採り入れた日本語コースの運営を考える—授業時間の不足を補い、日本語運用能力を伸ばすための e ラーニングの活用—, 第 15 回ヨーロッパ日本語教育シンポジウム(査読有), 2010-08-26, ブカレスト, ルーマニア
- ⑧ 芝野耕司, 日本語教育基盤としての教科書及び均衡コーパスの開発, 第 15 回ヨーロッパ日本語教育シンポジウム(査読有), 2010-08-26, ブレスト, ルーマニア
- ⑨ Kaori Miyatake, Kohji Shibano, Indexed gender difference: the use of Japanese social deixis, The 18th International Conference on Pragmatics and Language

Learning (PLL18) (査読有), 2010-07-18,  
Kobe University, Japan(神戸)

- ⑩ Xiao Tingting, Kohji Shibano, The Japan  
That Does Say No -FTAs in Japanese  
Parliamentary Debate, The 18th  
International Conference on Pragmatics and  
Language Learning (PLL18) (査読有),  
2010-07-18, Kobe University, Japan(神戸)

[図書] (計1件)

- ① Kohji Shibano ed, ISO, ISO/IEC TS 13249-7  
Information technology — Database  
languages — SQL multimedia and  
application packages — Part 7: History,  
2013, 116

[その他]

ホームページ等

<http://jplang.tufs.ac.jp/>

## 6. 研究組織

### (1) 研究代表者

芝野 耕司 (SHIBANO KOHJI)  
東京外国語大学アジア・アフリカ言語文化  
研究所・教授  
研究者番号：50216024

### (2) 研究分担者

市川 雅教 (ICHIKAWA MASANORI)  
東京外国語大学大学院総合国際学研究  
院・教授  
研究者番号：20168313

藤村 知子 (FUJIMURA TOMOKO)  
東京外国語大学留学生日本語教育センタ  
ー・教授  
研究者番号：20229040

大津 友美 (OHTSU TOMOMI)  
東京外国語大学留学生日本語教育センタ  
ー・教授  
研究者番号：20437073

佐野 洋 (SANO HIROSHI)  
東京外国語大学大学院総合国際学研究  
院・教授  
研究者番号：30282776

金子 比呂子 (KANEKO HIROKO)  
東京外国語大学留学生日本語教育センタ  
ー・教授  
研究者番号：50224605

藤森 弘子 (FUJIMORI HIROKO)  
東京外国語大学留学生日本語教育センタ  
ー・教授  
研究者番号：50282778

鈴木 美加 (SUZUKI MIKA)  
東京外国語大学留学生日本語教育センタ  
ー・教授  
研究者番号：90226556