

VALIS: 英語学習者のプロフィールと発話データの収集

原田 康也¹ 前坊 香菜子² 河村 まゆみ³ 鈴木 正紀⁴

¹早稲田大学法学学術院 〒169-8050 東京都新宿区西早稲田 1-6-1 harada@waseda.jp

²早稲田大学日本語教育研究センター 〒169-8050 東京都新宿区西早稲田 1-7-14 xiangcai_2@suou.waseda.jp

³言語アノテータ kawamuras@pat.hi-ho.ne.jp

⁴Ordnate Corporation 800 El Camino Real Suite 400, Menlo Park, CA 94025 msuzuki@ordinate.com

あらまし 大学における英語教育において、英語を用いた自己表現能力と対人折衝能力の涵養を達成課題としつつ、学習者の自己学習・相互学習の向上を目指した授業実践が求められている。本発表では第一著者を研究代表者とする科学研究費補助金（2006年4月-2009年3月）基盤研究（B）：課題番号18320093『学習者プロファイリングに基づく日本人英語学習者音声コーパスの構築と分析』などにより進めている英語学習者のプロフィールならびに発話データの収集についての現状を報告する。第一著者は、担当する早稲田大学法学部一年生の授業で、あらかじめマルチカードに印刷して用意した質問を利用した対面での応答練習、応答練習に基づく少人数グループでの相互プレゼンテーション、応答練習と相互プレゼンテーションに基づく文章作成と相互チェック・相互評価を中心とした初年度導入教育を進め、これに続く二年次には少人数グループによる『調べと発表とまとめ』の授業を実践し、学生の自己学習・相互学習について一定の手ごたえを得ているが、これらを実証的に研究する基礎的な資料の構築と分析も必要である。こうした観点から、比較的自発的な発話において学習者が使用できる語彙・表現について調査し、学習者の自律的相互学習を実証的に研究する上での基礎的な資料とすることを目指して、授業時の学生の発話をマルチトラック・ハードディスク・レコーダを中心に試作した装置を利用してデジタル収録するとともに、Bluetooth・ワイヤレス・マイクとハードディスク・ビデオカメラを用いて音声画像をデジタル収録している。収録している音声データがきわめて多量であるため、その大部分を対象とすることは期間と予算の制約から不可能であるが、将来的な作業の見通しを付けるために2007年度から発話データの書き起こしを試行している。

キーワード 学習者プロファイリング、デジタル音声録音、デジタルビデオ、書き起こしツール

VALIS: Collection of English Learners' Profiles and Utterances

Yasunari HARADA¹ Kanako MAEBO² Mayumi KAWAMURA³ Masanori SUZUKI⁴

¹Faculty of Law, Waseda University 1-6-1 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050 Japan: harada@waseda.jp

²Center for Japanese Language, Waseda University 1-7-14 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050 Japan:
xiangcai_2@suou.waseda.jp

³Language Annotator: kawamuras@pat.hi-ho.ne.jp

⁴Ordnate Corporation 800 El Camino Real Suite 400, Menlo Park, CA 94025: msuzuki@ordinate.com

Abstract The first author has implemented college English classes emphasizing face-to-face oral interactions within small groups of students in class, presupposing and expecting further cultivation of learners' ability to learn for themselves, by themselves and among themselves. Previous experiences confirm such expectations, and the authors are currently working on compiling spoken corpora of learners' interactions with relatively high-quality audio and video recordings of those interactions. Students participating in those English classes take English proficiency tests such as TOEIC® and Versant for English so that segments of recorded utterances may come with tags designating proficiency of the learner who is responsible for the segment. It is interesting to note, in passing, how introduction of such digital recording devices into the classroom environment positively affect students' motivations and performances in those practices. In this interim report, we also touch on our transcription procedures and possible tools, while detailing our recording procedures and project goals.

Keyword learner profiling, digital audio recording, digital video recording, annotation tools, transcription,

1. はじめに

今日の大学教育では英語教育においても自己表現能力と対人折衝能力の涵養を達成課題としつつ学習者の自己学習・相互学習の向上を目指した授業実践が求められている。第一著者が担当する早稲田大学法学部一年生の授業では、あらかじめマルチカードに印刷して用意した質問を利用した対面での応答練習、応答練習に基づく少人数グループでの相互プレゼンテーション、応答練習と相互プレゼンテーションに基づく文章作成と相互チェック・相互評価を中心とした初年度導入教育を進め、これに続く二年次には少人数グループによる『調べと発表とまとめ』の授業を実践し、学生の自己学習・相互学習ならびに英語運用能力の向上について一定の手ごたえを得ている。本稿では、上記を実証的に研究する基礎的な資料の構築と分析を目指して現在進めている共同研究のうち、学生の発話を無圧縮でデジタル収録するマルチトラック・ハードディスク・レコーダを中心に試作した装置の利用状況、Bluetooth・ワイヤレス・マイクとハードディスク・ビデオカメラを用いた音声画像の収録ならびに各種試験を利用した学習者のプロファイリングについて紹介し、将来的な作業の見通しを付けるために試行的に進めている発話データの書き起こしの途中経過を報告する。

2. データに基づく学習活動の分析と検証

近年の外国語学習においてはコミュニケーション活動を重視し、教員による知識の提示に基づく学習者における知識の獲得を目的とした活動より学習者個人の外国語運用ならびにペアまたはグループでの外国語コミュニケーション活動に比重を置く傾向が見られる。このような学習デザイン的前提として、運用経験ならびにグループ活動が学習者（相互）による自律的外国語学習に有効に機能すると仮定されているものと考えられる。¹ しかし、日本の大学英語学習環境においてこのような仮定が成立することを直接的に証明する客観的なデータが十分に提示されているとは言えず、この点について今後とも詳細な検証が必要とされている。

著者たちの研究グループでは、アカデミック・リテラシーの習得において学習者の相互的活動が重要な位置を占めていること²、最も広い意味での学習履歴の蓄積と分析が学習活動の効率化と学習資源の改善のために必要であること³、外国語の運用経験が外国語の学習に不可欠であること⁴を前提にこれまでの研究を進

めてきた。2004年度には早稲田大学特定課題研究助成費（一般助成）課題番号 2004A-033『大学英語教育高度化のための外部試験を活用した学習者プロファイリングの研究』、2005年度には早稲田大学特定課題研究助成費（一般助成）課題番号 2005B-022『英語教育高度化に向けた学習者プロファイリングとマルチモーダル学習者コーパスの研究』、2006年度からは、科学研究費補助金（2006年4月-2009年3月）基盤研究（B）：課題番号 18320093『学習者プロファイリングに基づく日本人英語学習者音声コーパスの構築と分析』の交付を受けて、大学生英語学習者が比較的自発的な発話や文章作成においてどのような語彙・表現を用いるか、各種外部試験のスコアなどとの関連においてデータに基づいた分析を可能とするための基礎的な資料作成に取り組んでいる。付随的に収録されている学生の相互交流の様子から、学生の自律的相互学習の状況が明らかになる可能性も期待できる。

3. 統合的言語活動を中心とした授業実践

第一著者が担当する法学部1年春学期の英語授業Bridgeでは、2週間を一つのサイクルとして、マルチカードを利用した応答練習⁵を導入として、応答練習の内容を思い出しつつ400語を目標にWordで複数のパラグラフからなる文章にまとめ、6人の学生がプリントアウトに相互にコメントを付して採点して返却し、これに基づいて修正した文章を提出するという流れで作業を進める。1年秋学期のGateでは、文章作成に先立って応答練習の内容を思い出しながらPowerPointで2,3枚のスライドを作成し、4名程度の少人数グループの中で相互に発表した後に文章作成の作業に入る。2年Themeの授業では学期のはじめに上記と同様の練習を行ったのち、学生をそれぞれの関心に基づいて3名前後の人数からなるグループにわけ、一週間という短い時間的制約の中でグループごとに調べてまとめた結果をグループ間で相互に発表し、さらに一週間の見直し期間の後にクラス全体に発表した後、その内容を文章化して整理するという作業を行う。クラス全体の発表に際してはその内容や形式について全員で相互評価を行い、文章化してまとめた資料については引用の形式や情報の出処の表示等について相互チェックを行うことを通じて、情報の取り扱いについての基本的な作法を身に付けることを目指している。

以下に紹介する装置での音声収録については、応答練習ならびに少人数での相互プレゼンテーションが主な対象となるが、クラス全体への発表も収録している。

¹ 詳細については[1]などを参照。

² 詳細については[3]などを参照。

³ 詳細については[6]などを参照。

⁴ 詳細については[4]などを参照。

⁵ 詳細については[5], [7]-[9]などを参照。

4. 発話収録装置

4.1. システム使用環境と要求仕様

学生の音声を事後の研究利用に対して可能な限り制約を与えない形式で蓄積することを目指し、以下の点を勘案してシステムの仕様を検討した。

- 36名（もしくは3人のグループが12）までのクラスで音声を一斉に収録できること
- 複数の教室で移動して利用できること、比較的フロアスペースの広いコンピュータ教室においても一般の教室でも使用できること
- 教員による取り扱いと操作が簡単であること
- 将来的にさまざまな蓄積・解析・検索を可能にするため、無圧縮デジタルデータにて記録すること

4.2. システムの主な構成

本システムは、アレシス製ハードディスクレコーダ1台、同社製マイク8chフェーダ2台、ソニー製マイクロホン12本、マイクケーブル12本、可動式機器保管庫によって構成される。マイクロホンをマイクケーブル経由にてマイク8chフェーダに接続し、増幅した音声をハードディスクDATレコーダに収録する。

- アレシス製ハードディスクレコーダ
12トラック同時録音（88.2/96kHz）
記録メディア：標準IDEハードディスク（5400rpm以上）、ホットスワップ
10GB録音時間：
12トラック（96kHz）にて45分
量子化ビット数：
24ビットリニアPCMエンコーディング
音響特性
周波数特性：22Hz～44kHz±0.5dB
歪率：0.003%以下
S/N比：112dB以下、A-weighted
- アレシス製マイク8chフェーダ（アンプ）
周波数特性：10Hz～65kHz、+0/-1dB（ノミナルレベル時の入出力）
- ソニー製マイクロホンECM-360
周波数特性：50Hz～16kHz
正面感度：-46dB±3dB
S/N比：68dB以上

5. ビデオカメラによる音声・画像収録

機材の選定に当たっては、多数の媒体から該当する物理的実体を検索する必要がないようにファイルとしてサーバに蓄積することを前提として、デジタルでデータを処理できることを重視した。SONYのビデオカメラ用付属品でHCM-1というBluetoothを使うワイヤレスマイク（とカメラに接続する受信部）が

あり、これを試用することを想定してDCR-SR100を選定した⁶。

ビデオカメラの内蔵ハードディスクは30GBであるが、週6コマの英語の授業でフルに使うと10GB前後を使用する可能性がある。学生が手持ちで撮影できるようにバッテリーで作動させているため交換が必要であり、バッテリーの交換に合わせてファイルを吸い上げてハードディスクに蓄積することとした。USB 2.0接続の外付けハードディスクへ10GB転送するのに10分程度、12台の転送に2時間程度、バックアップに2時間程度という計算になるが、実際に授業後または週末に合計5,6時間ほど費やすと、転送とバックアップとバッテリーの交換と充電を行うことができる。12台のビデオカメラそれぞれで一週間に撮影するファイルの容量にはばらつきがあるが、全体として30GBとして一学期14週間で520GBとなり、バックアップも含めて1テラバイトの外付けハードディスクのドライブひとつでは収まりきれないぐらいの容量となる。

6. 試験データ

各年度の対象学生に外部試験の受験を促すとともに、英語学習経歴に関する簡単なアンケートを実施して集計している。

2006年度の春学期に第一著者は法学部1年必修（自動登録）の英語3クラス約100名、2年必修選択の英語2クラス約70名、3年選択の英語1クラス4名を担当した。Harcourt Assessmentの全自動口頭英語試験Versant for English⁷の試験用紙を4月中旬授業と6月下旬とに配布して受験を促すとともに、5月中旬に受講生から希望を募り6月23日実施の第123回TOEIC公開テストの団体受験を申し込んだ。秋学期には1年必修（自動登録）英語3クラスの約100名、2年必修選択英語2クラス30名弱、3年選択英語1クラス11名を担当した。春学期より引き続いて受講している1年生約100名を除いて、後期開始の10月にVersant for Englishの試験用紙を配布して受験を促した。11月中旬より希望者を募り、1月14日実施の第128回TOEIC公開テストの団体受験を申し込み、12月中旬にVersant for Englishの試験用紙を配布しTOEIC公開テストの受験とあわせて、2007年1月14日前後に受験するように指示した。⁸

⁶ DCR-SR100は内蔵ハードディスクがWindows形式でフォーマットされ、撮影を終えた時点でファイルの日時が書き込まれるので、ファイルのプロパティを見るだけでどの授業のどの時点でのファイルか判明する点が極めて便利である。
⁷ これは[4]、[5]などにてOrdinate CorporationのPhonePass SET-10として言及した試験に相当する。

⁸ TOEIC公開テストについて、春学期は119名が受験を申し込み81名が受験した。秋学期については74名が受験を申し込み53名が受験した。

2007年度春学期には法学部1年必修（自動登録）の英語3クラス（登録受講者計約75名）、2年必修選択の英語2クラス（登録受講者計約70名）、3年選択の英語1クラス（登録受講者11名）を担当した。授業開始時に受講者全員を対象として英語学習経歴に関する簡単なアンケートを実施するとともに Versant for English の試験用紙を4月中旬授業と6月下旬とに配布して受験を促した。秋学期については1年必修（自動登録）英語3クラスの登録受講者が約75名、2年必修選択英語2クラスの登録受講者が約65名、3年選択英語1クラスの登録受講者が10名となった。後期開始の10月ならびに学年終了に近い12月に Versant for English の試験用紙を配布して受験を促す予定である。春学期には麻疹休講の影響もあり TOEIC の受験を見送ったが、秋学期には10月中に受講生から希望を募り、1月中旬に TOEIC 公開テストの受験を促す予定である。

7. 音声データの収録と書き起こしの方向性

7.1. 書き起こし対象データ

本研究プロジェクトで集積を目指している主要なデータは大学生英語学習者の英語による発話と、それぞれの発話を行っている大学生の学習経歴や英語学習到達度の指標となるような外部試験のスコアである。発話としては、応答練習においては担当教員があらかじめ用意した質問を（2回）読み上げる音声とこの質問を聞いて（あらかじめ用意していない）回答をする音声を中心とする。また、4名ないし6名程度のグループ内で行う少人数での相互プレゼンテーション活動における英語による発話は、発表者による（短い準備時間の後の、比較的自発性の高い）英語による発話を中心で、場合によってプレゼンテーションの途中またはプレゼンテーションのあとでその場で思いついた簡単な質問と、これに対する回答が得られる場合もある。また、グループで調査してきた内容を発表する場合には、収録対象となるのはあらかじめ用意してきた内容が中心になる。

音声収録を行っている学習活動中、特に少人数で同時に活動を進めている場合、受講生は相互にさまざまな発言を行っているが、その大部分は日本語によるものであり、ある意味で当初想定していた主たる分析対象のデータではない。しかしながら、本研究が長期的に目指すところは単なる発話データの収集だけではなく、これを通じて学習者の自律的相互学習を根拠付ける実証的なデータを得るところにもあるため、日本語による相互交流についても可能な範囲で分析が可能となるようにデータを整備することが望ましい。

インタラクションの書き起こしと分析については、

榎本美香・伝康晴による一連の研究報告がある。⁹しかし、本研究プロジェクトでは、授業に限定して音声収録の設置と撤収を完了しなければならないという運用面での制約ならびに予算上の制約などから、3人の学習者に1つのトラックを割り当てている。¹⁰このことが各種ツールの選定に影響している。

7.2. 音声データに対する書き起こし

transcription {書き起こし|転記}といっても、対象とするデータに対する検索を可能ないし容易にするために言語的記号に転換する作業を本質とするところをのぞくと、目的とする成果物や想定する検索手法によって作業内容が異なってくる。たとえば、動画像を含むデータに対して何らかの目的で transcription {書き起こし|転記}を用意しようとする場合、動画像に撮影された対象の名称、画像中の位置、その移動と変化などを記述する必要が生じる。この場合、言語記号への転記が作業の中心となるが、対象とする動画像に撮影されていることが想定される事物やその相互関係、時間的変化にともない生起するイベントなどについての何らかのオントロジーとそのラベルを設定して、これに基づいて動画像の内容を同定することが求められるであろう。

音声データに対する transcription {書き起こし|転記}の annotation {注記|タグ付け}についてはここまで厳密な事前の準備は必要ないように思えるかもしれないが、何が音声的に発声されているかの annotation {注記|タグ付け}と何が言語的に発話されているかの annotation {注記|タグ付け}とでは、transcription {書き起こし|転記}に用いる記号も異なるし、そのために必要な事前の準備や作業に対する訓練も異なってくる。音声的書き起こしは、母語話者の中での地域的発音のゆれ、習熟した非母語話者による発音の特徴などを研究するために音声データを収集整理したり、収集した音声データを用いて音声認識エンジンを開発しようというような場合に必要となることが想定される。こうした書き起こしの結果は英語や日本語などの文字列ではなく、IPA (International Phonetic Alphabet: 国際音声字母) のような記号を用いた音素列といった音声的記号単位の列となるであろう。目的に応じて自然言語の一般的標記に近いアルファベット列に若干の補助記号を加えて標記することも可能であるが、音声と前後し、あるいは重なって生起するノイズの記述も最小限必要となるはずである。書

⁹ 発話単位の認定基準などについては [2] の紹介がわかりやすい。

¹⁰ ビデオカメラを使用して音声・画像も収録しているが、操作を学生に任せるため、学習活動の記録として重要なインタラクションが欠落することも多く、音声データに比して補助的な記録手段と考える必要がある。

き起こしの目的に応じて、ピッチパターンやアクセントなどの supra-segmental な情報を記号的に記述することが必要になる場合もある。

音声的書き起こしを実施するためには、さまざまな発音の微細な異同を知覚できるような書き起こし者が必須であり、さらに目的に応じて作業者を訓練し、プロジェクト・マネージャが方針を常に確認して作業の均一性を確保する必要がある。このため、書き起こし作業は対象となる音声データに対して多大な時間を要することが一般的であり、こうした目的で音声データを収集する際には、あらかじめ入力データの長さや発声すべき単語（列）を正規化ないし標準化することが必須である。

今回の作業では、応答練習や少人数でのプレゼンテーションなど比較的自発的な英語による発話の中で学生がどのような語彙や文型パターンなどを使用できるか観察し、日本語でのインタラクションから自律的相互学習の証拠を抽出しようというのが当面の目標である。英語の応答については大まかな時間枠の設定があるだけで内容については学生の自発的な発想に任されている上、日本語インタラクションについては教室内の英語学習活動の一貫であるという点を除いてほとんど何の制約もない。研究目的からもデータの長さから言っても、今回のデータに対してまず行うべき作業は言語的な書き起こしであり、データ内容の概観が用意できてから音声的な研究資料の切り出しを考えるのが順当であろうと考える。アノテーションツールが広く利用可能となりつつある現状では、音声そのものの研究・検討には音声データを直接操作することが可能で、これを記号化した書き起こしは必ずしも必要でないと言える。

7.3. 音声データアノテーションツールの選定

Windows で利用できる音声データ処理用ソフトウェアのうち、ライセンスや使用料も含めて利用しやすいものとしては WaveSurfer や Praat などがあり、早稲田大学では MultiTrans とあわせて各学部コンピュータ教室の PC にインストールして教育研究用に広く活用されている。今回の音声データについては、将来的に音響音声学的な分析を進める可能性を残しているものの、当面はデータの内容を一覧したり検索したりし易くするための書き起こしが作業の中心となる。

2006 年度の後半になってから、LDC¹¹ の WaveSurfer, MultiTrans, TableTrans など AGTK¹² ツールキットを SourceForge.net よりダウンロードし、作業用 PC にイン

ストールして動作状況を確認することから作業を始めた。96kHz/24bit サンプルングで最長 90 分 1 トラックのデータをそのまま読み込んで作業できるとは想定していなかったが、上記 3 ソフトの場合この程度のデータについて、読み込みには支障がなかった。

TableTrans はひとつのトラックの音声を時間的な小部分に領域分割しつつ、開始点・終了点情報とあわせて、複数の属性を定義してこれを表形式で入力・表示できるため、話者・使用言語・発話内容を最小限の情報として書き起こししたい今回の作業についてはほどよいツールと思われた。また、試用に向けてその機能を確認しているうちに、TableTrans が WaveSurfer の上位互換ソフトとして作成されているため、フォルマント・ピッチ・そのほかの音声分析ツールをそのまま使用できることもわかり、将来的に詳細な音声分析を行う可能性を考えた場合、データのインポート・エクスポートなどの負荷を考えずに作業を進められる可能性が期待できることもわかった。

7.4. 書き起こし作業の経過

2007 年度になってから、まず新学期の授業 1 週間分（2 年生 Theme 2 クラスと 1 年生 Bridge 3 クラス）の音声データの書き起こしを始めたが、書き起こしの詳細度や手順を検討するための準備段階として位置づけ、いくつかの方法を試みている。さまざまな組み合わせを仔細に検討する時間的余裕がないが、1 クラス 20 分 × 12 トラックの概略の書き起こしが完了し、2 クラス目の作業を進めているところである。¹³

英語部分については、開始点と終了点の確定も比較的容易で、エラータグについて見送る限り、ある程度の時間で一通りの作業が完了する。しかし、日本語部分については話者交代が明瞭でない場合が多く、しかも初回の授業であるため学生も収録作業に慣れていない場合もあり、マイクの使い方が悪いため、音声が小さく聞き取りにくく、波形で分割の検討をつけるということがうまくできないため、切り分けの作業に異様に時間がかかる場合もある。ひとつのグループが他のグループの受講生に進め方を問い合わせることもあり、そうすると話者の識別がますます困難になる。話者の交代が多く、うなずきが多いと、これまた切り分けが難しくなる。こうしたことから、あるトラックに関しては切り分けと書き起こしで 7 時間以上も必要となった。別のグループに関しては 2 時間から 3 時間程度で

¹¹ LDC は Linguistic Data Consortium
<http://www ldc.upenn.edu/> に詳細。

¹² AGTK は Annotation Graph ToolKit
<http://sourceforge.net/projects/agtk/>
よりダウンロード可能。

¹³ 書き起こし作業担当グループにアノテーションツール利用経験者がいないため、TableTrans の使い方や機能についてオンラインで提供されているマニュアルを眺めながら手探りで試行錯誤している状態であるため、その本来の機能を十分に有効活用できていない可能性がある。TableTrans を先行して活用している国内の研究グループを探しているが、いまのところははっきりしたことがわかっていない。

概略の書き起こしが終了するが、それにしても1クラス分の書き起こしに対して36時間かかるとすると、専従的な作業員1週間分の作業量に近く、1週間5クラスの概略の書き起こしを完了するには、専従的な作業員の1月分の業務量を要するという概算になる。現時点では、主として英語で発話している質問の読み上げと応答の部分だけを切り出して書き起こしを加える作業にどの程度の時間がかかるか、概算見積もりの前提となる作業を行っているところである。

音声ファイルをTableTransに読み込み、開始点と終了点を設定してその区間の話者・主要な言語・書き起こしテキスト・発話タイプを注記している。書き起こし作業の開始時点では、発話タイプとして主に英語による自己同定(氏名と出席番号)・質問(読み上げ)・応答(自由発話)などに加え、主に日本語による進行調整・相互扶助・雑談などの詳細を記号として注記することを検討したが、実際に試してみるとこの粒度のタイプ分けでも書き起こし作業員には分類のための判断の負担が大きく、作業時間がかなり膨らむことがわかった。書き起こしが終了すれば進行調整・相互扶助・雑談などの詳細な発話タイプは比較的容易に同定できる可能性が高いので、現在の時点では自己同定・質問(読み上げ)・応答・その他というような分類で記号により注記している。

このほか、英語で話している途中で言葉に詰まって日本語で話し始める可能性や、日本語でのやり取りの中に質問と応答に出てきた英語の表現が引用される可能性を考えて、使用言語の切り替えに特定の記号を挿入することを検討したが、書き起こしを見れば半角文字列と全角文字列の区別が容易に可能であるため、これまた見送ることとした。関連して、英単語をカタカナ風に発音しているときの書き起こしを日本語として扱うか英語として扱うかの判断が悩ましくなる。同様に、fillerの一部は日本語として扱うか英語として扱うか若干悩ましい場合がある。こうした問題点について作業を進めながらさらに検討を進めている。

8. 付記

本稿で報告する共同研究においては、楠元範明と前野譲二がデータ処理環境の構築と機材の選定に関わる情報提供と意思決定を行っている。デジタル音声収録装置の仕様検討ならびに構築を鈴木陽一郎が、口頭英語自動試験のスコア分析を鈴木正紀が行っている。デジタル音声収録装置の運用支援ならびに試行的アノテーションについては河村まゆみが、音声書き起こしならびに分析用各種ソフトウェアの検討・試用については前坊香菜子が主に担当している。

9. 謝辞

本稿の著者たちを中心とする共同研究は科学研究費補助金(2006年4月-2009年3月)基盤研究(B):課題番号18320093『学習者プロファイリングに基づく日本人英語学習者音声コーパスの構築と分析』の助成を受けている。本稿で報告した発話収録装置の試作と試用にあたって早稲田大学特定課題研究助成費(一般助成)課題番号2004A-033『大学英語教育高度化のための外部試験を活用した学習者プロファイリングの研究』(研究代表者:原田康也)ならびに課題番号2005B-022『英語教育高度化に向けた学習者プロファイリングとマルチモーダル学習者コーパスの研究』(研究代表者:原田康也)による助成を受けている。

文 献

- [1] 和泉絵美, 内元清貴, 井佐 原 均 “学習者コーパスからの表現バリエーションの抽出と置き換えストラテジー指導への利用,” 自然言語処理, Vol. 12 No. 4, pp. 193-210, 言語処理学会, 2005年8月.
- [2] 榎本美香・石崎雅人・小磯花絵・伝康晴・水上悦雄・矢野博之, “相互行為分析のための単位に関する検討,” 人工知能学会研究会資料, SIG-SLUD-A402, pp. 45-50, 2004.
- [3] 原田康也, 辰己丈夫・楠元範明, “『情報教育』の情報化,” 情報処理学会研究報告, Vol.2000, No.20, CE 55-6, pp.41-48, 情報処理学会, 2000年2月18日.
- [4] 原田康也, 「外国語学習における知的情報処理と言語処理技術の応用」, 2001年情報学シンポジウム講演論文集, pp.25-32, 社団法人情報処理学会発行, 2001年1月18日.
- [5] 原田康也, “エーワンのマルチカードを用いた英語応答練習,” 情報処理学会研究報告 CE-69-3 pp.17-22, 情報処理学会, 2003年5月16日.
- [6] 原田康也・前野譲二・楠元範明・辰己丈夫, 「学習履歴の双対性: 学習履歴を活用した e-learning 高度化の数理的基礎を目指して」, 情報処理学会研究報告 IPSJ SIG Technical Reports 2003-CE-70 (1), 学術刊行物 情処研報 Vol. 2003, pp.1-8, 社団法人 情報処理学会, 2003年7月12日, ISSN 0919-6072.
- [7] 原田康也・辰己丈夫・前野譲二・楠元範明・鈴木陽一郎, “対面での応答を重視した英語学習活動と発話収録装置の試作と試用,” IPSJ SIG Technical Reports 2005-CE-80 (4), 学術刊行物 情処研報 Vol. 2005, pp.25-32, 社団法人 情報処理学会, 2005年6月18日, ISSN 0919-6072.
- [8] 原田康也・前坊香菜子・河村まゆみ・前野譲二・楠元範明・鈴木陽一郎・鈴木正紀, “VALIS: 学習者プロファイルに基づく学習者音声コーパス構築を目指して,” IPSJ SIG Technical Reports 2006-CE-88 (24), 学術刊行物 情処研報 Vol. 2007, No. 12, pp.169-176, 社団法人 情報処理学会, 2007年2月16日, ISSN 0919-6072.
- [9] 原田康也・前坊香菜子・河村まゆみ, “VALIS: 英語学習者発話データの書き起こし,” IPSJ SIG Technical Reports 2007-CE-90 (1), 学術刊行物 情処研報 Vol. 2007, No. 69, pp.1-8, 社団法人 情報処理学会, 2007年7月7日, ISSN 0919-6072.