

## ラーニングアナリティクスに基づくICT教育の次世代モデルに関する研究

ICTと教育研究部会

### 1. はじめに

(担当：湘南工科大学 梅澤克之)

教育ビッグデータとして得られた学習履歴データの分析を通じ、新しい学習・教育モデルの確立をめざし、反転学習などアクティブラーニングにおける教育モデルの検討・評価を行ってきた。また、そのために必要な脳波などの生体情報と端末の操作履歴などを統合し、学習者の理解度をより正確に判定する手法の検討・評価を行う。具体的には下記のような活動を行っている。

- 学習分析に基づく反転学習・協働学習のための学習効果の高い教育モデルの構築
- リアルタイム授業支援に有用な学習履歴分析方法
- これまでの学習履歴に生体情報を加えたマルチモーダルな学習分析手法の開発

本稿では、本研究部会の3名のメンバー（会津大学中澤真，早稲田大学中野美知子，湘南工科大学梅澤克之）より、これらのラーニングアナリティクスに基づくICT教育の次世代モデルに関する研究に位置づけられる4つの事例を紹介する。

### 2. 遠隔授業において学習状況を把握できる自学自習システムの研究

(担当：湘南工科大学 梅澤克之)

言語学習を統一的な枠組みで捉え、相手(学習者)を感じて助言を行う人工教師を搭載した自学自習システムを開発しその評価を行う研究を推進中である[1]。「相手を感じる」とは、学習行動の背後にある「学習者の理解度」や「学習者ごとの思考プロセスの差異」「学習時の集中度や退屈度」、学習者ごとの「問題解決の困難度」などの学習者の学習状況をシステム側が把握することを意味する。「相手を感じる」ための生体情報としての脳波を学習者から効率的に収集するためのシステムの構築を行った。図1に最終的な自学自習システムの全体像を示す。

本研究では、(i) ケアレスミス(脳波情報と解答時間から判定し、ケアレスミスをし易い問題を集中的に出題する機能)、(ii) 単語(英語における単語やプログラミング言語における予約語)の綴りや文法(構文)の誤りが多い学習者に対しては単語や文法の理解を向上させる(単語の綴り間違いや文法誤りを修正させる)問題を出題する機能、(iii) 文法的には正しいが論理的な誤りを含む問題を訂正させる問題を出題する機能、また、(iv) 上記(i)~(iii)について脳波や視線などの生体情報や学習履歴情報を組み合わせて学習に集中できていない状態、学習内容が簡単すぎる状態、学習内容が難しすぎて理解できていない状態、部分的に理解できない箇所がある状態を判別し、リアルタイムで課題の難易度を調整する機能を持った自学自習システムを目指している。

ケアレスミスに関しては、プログラミングの学習時の脳波情報を計測し、課題の回答時間と脳波の関係に着目し、注意力が欠けた際のケアレスミスを検出することを試みた[2]。図2に実験参加者1の各設問回答時の平均脳波を示す。丸印の箇所は $\beta/\alpha$ (脳波計の出力信号 $\alpha$ 波と $\beta$ 波の比)が下がっている箇所である。誤答で、かつ回答時間が短くて、かつ、統計的に有意に $\beta/\alpha$ が低ければ、そこがケアレスミスであると考えることが出来る。ケアレスミスを起こしやすい問題を集中的に訓練することでプロ

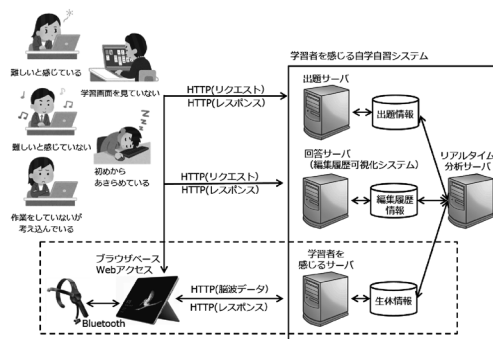


図1 自学自習システムの全体像

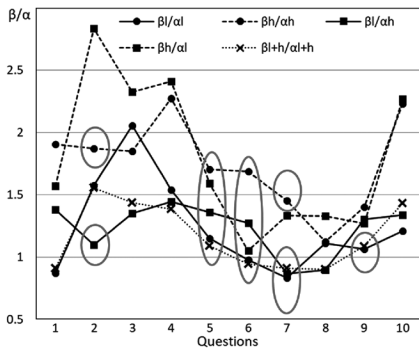


図2 実験参加者1の各設問回答時の平均脳波

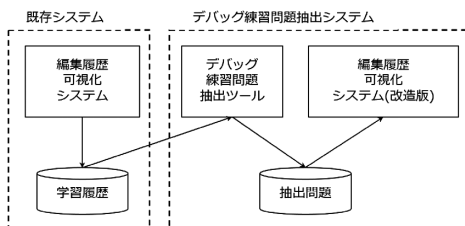


図3 デバッグ練習問題抽出システムの全体構成

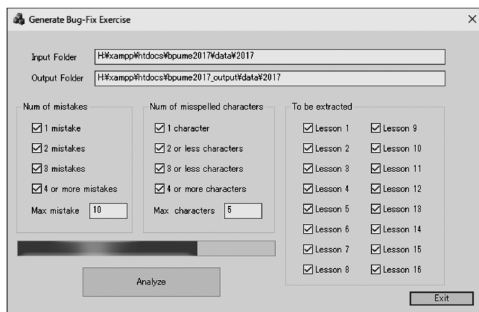


図4 デバッグ練習問題抽出ツール

プログラミングの技術向上につなげることができると考える。

文法誤りに関しては、プログラミングの授業を実際に行ったときにエラー表示がされているにもかかわらず教員にヘルプを求める学生が非常に多いという事例から、プログラミングの学習履歴の文法エラー情報をもとに、デバッグ練習用の問題を自動生成するデバッグ練習問題抽出システムについての研究を行った [3]。図3にデバッグ練習問題抽出システムの全体構成を、図4にデバッグ練習問題抽出ツールを示す。

なお、論理誤りに関しては、上記と同様にプログラミングの学習履歴を分析することにより、エラー情報が出力されない論理エラーを自動的に分析する方法に関して研究を推進中である。

### 3. グループ分け反転授業に関する研究 (担当：湘南工科大学 梅澤克之)

近年、授業と宿題の役割を「反転」させ、授業時間に先立って知識習得(自習)を済ませ、教室では知識確認や問題解決学習を行う授業形態をとる反転授業が注目されている。我々は、学生が自宅でe-learningにより自習を行うときに学習ログを取得し、学習時間と理解度の関係から学生を複数のグループに分類した上で教場での授業を行うグループ分け反転授業を提案し、提案方式と反転授業ではない従来の講義形式、および反転授業ではあるがグループ分けを行わない授業との比較評価を行い、本学会普及誌においても報告させていただいた [4]。

その後、分析対象を2年間分の学習履歴に拡大して統計的分析 [5]、アンケートによる評価 [6]、および自習時の理解度に関する評価 [7] などの研究を推進した。

### 4. 初等教育向けプログラミング授業支援システムの研究 (担当：会津大学 中澤 真)

2020年度から全面実施となる小学校でのプログラミング教育では、代表的なビジュアルプログラミング言語であるScratchが利用されることが多い。Scratchは初学者でも容易に学ぶことができるように設計されているが、それぞれの学習者の進捗や理解状況を教員が授業中に把握するには時間をかけて巡回指導する必要があり教員の負担は大きい。特に、一般的なテキスト型プログラミング言語と異なり、Scratchのようなビジュアルプログラミング言語では文法エラーは基本的に発生しないため、発見すべきは論理エラーのみとなるため対応に時間がかかることになる。このため、授業実施時間中に学習につまずいていたり、他の学習者と異なるタイプのプログラムを作成している者などを、教員がリアルタイムに状況把握できる授業支援システムが求めら

れることになる。

そこでまず、学習者のプログラムの編集過程を10秒ごとにサーバ上へ自動保存し、記録された編集履歴を学習者別、タイムライン別に確認できるプログラムの編集履歴可視化システムを構築し(図5参照)、授業終了後にバッチ処理で学習者のモデリングや理解度・学習効果などを分析することで授業改善に役立てられることを示した[8]。

次に学習者の状況を授業中にリアルタイムに把握するために、編集履歴からプログラム中で使用されている構成要素を分析し、その使用率を特徴量として学習分析することに取り組んだ[9]。学習者が特定の課題のプログラム作成を目指す場合、使用するブロックの種類には偏りが生じるため、プログラムを詳細に構文解析せずとも、この特徴量を比較することで他の学習者の取組状況から逸脱している者を検出できる。例えば、図6に示したように作業開始の60秒経過時点で10%程度の学習者はキャラクタを前進させる構成要素「forward」を早くも使用しているが、半分以上の学習者は最初に使用するべきイベントハンドラ構成要素である「whenGreenFlag」

さえ使用できておらず、進捗にかなり差が生じていることを把握できる。また、クラス全体の進捗把握としては、180秒経過時点で「whenGreenFlag」は全員が設置済みとなり、300秒時点では、「turnLeft」を設置するところまで半数の学習者が到達しているというように、状況を正確に把握することが可能になる。また、本来使用しない構成要素「turnRight」を使ってしまっている学習者を検出しており、誤りの傾向を把握してリアルタイムに注意を促すことも可能になっている。

さらに、時点ごとの学習履歴の比較に加えて、学習者別の履歴を時系列的に分析する研究も現在推進中である。

## 5. ラーニング・アナリティックを用いた高大一貫英語教育教材の開発 (担当：早稲田大学 中野美知子)

高等学校での教材開発実験を開始するにあたり、以下の3点に注意する必要がある。

- 1) 生体情報を得る実験をするには、生徒本人の同意の他に、保護者の同意を得る必要がある。
- 2) 教員や学校の許可を得るには、単なるデータ収集を目的とすることは許されない。学習上の有用性を確認し、将来、反転学習や協調学習に使用できる質の高いものでなければならない。
- 3) 生徒たちの教育の向上を目指しつつ、生徒たち自身が教育効果を感じられる実験でなければならない。

1) については、時間的な制約があったので、生徒たちの同意を受けるだけにとどめ、このパイロットでは、生体情報を収集しないことにした。2)と3)を満たすために、多くの大学には、系属・付属の高等学校があるが、高大一貫教育を本格的に導入している所は少ないため、このパイロット研究では、高大一貫教育に適した英語教材を開発し、将来、反転学習や協調学習に使用できる教材を目指した。また、近年、大学入試に記述問題やスピーキング問題を取り入れようとしているので、本研究では500語の英文6本をクリティカルに読み、記述問題解答能力の育成に焦点を当てた。反転学習や協調学習を目指す教材なので、事前・事後テストの他に、1)主張(Claim)の見つけ方、主張を支持する内容

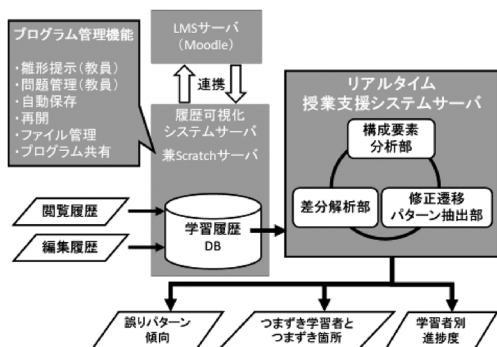


図5 編集履歴可視化システム構成図

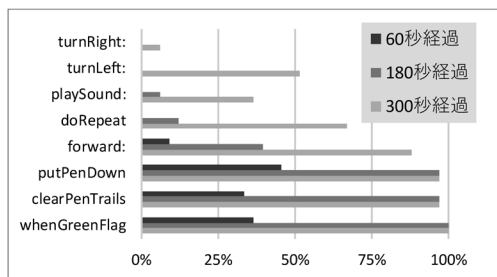


図6 プログラム構成要素の使用率の推移

(Support) の読み取り, 主張を展開していくためのサブテーマへの気づき, 想定される主張への反論 (Counter-argument) や反駁 (Rebuttle) への気づきを高める課題; 2) 移民政策と動物実験に関する 500 語のエッセイを読み, 適切な反論の仕方と証拠の出し方を学び; 3) 対面授業では, 読解過程の可視化グラフを提示しながら, 読解ストラテジーの解説, 反論, 立証の仕方を復習した. 事前・事後テストを比較することで, 学習成果を検証した.

### 5.1 パイロット実験の具体的な目的

- 1) 事前テストと事後テストが平行 (パラレル) テストになっているかを, 検証する必要がある. 平行テストとは, 二つのテストの形式が同じで, 平均値, 分散が等しく, かつ相関が高いという条件を満たすものである.
- 2) 平行テストであれば, 事前・事後テストのスコアを比較し, 教育効果を確認することができる.
- 3) 記述問題は 7 題あったが, それぞれの問題の難易度を主観的に評価させ, 問題が適切であったか, 難しすぎていたかを吟味した. これは, 次回の実験で, 生体情報により, 課題の難易度を確かめていくときの基礎資料になると考えた.

### 5.2 実験方法

41 名の A クラスと 44 名の B クラスが実験に参加した. すべての演習は PC 教室で行われた. 高大一貫教育の英語教材を開発するという実験の目的と実験内容の説明の他, 成績に反映されないこと, 授業として実験を行うが, 自由意思を尊重することが明記された同意書を配布した. 演習内容は, 事前・事後テストの他に, 3 回の演習を行った. 1) 主張, 支持する内容, 主張を展開していくためのサブテーマへの気づき, 想定される主張への反論, 反駁への気づきを高める課題; 2) テレビゲーム視聴の功罪, 移民政策, 動物実験に関する各 500 語のエッセイを読み, 適切な反論の仕方と証拠の出し方を学び; 3) 対面授業では, 読解過程の可視化グラフを提示しながら, 読解ストラテジーの解説, 反論, 立証の仕方を復習した. 11 月 18 日から 12 月 2 日で実験が終了する予定であったが, B クラスは期末試験の都合で, 事後テストは 12 月 9 日に実施された. 事前テストと事後テストの両方の結果をサーバに送付した

ものは 20 名と 19 名であった.

### 5.3 実験結果 1

難解な単語には内部辞書をつけ, どの単語を参照したかを記録したが, 内部辞書はほとんど利用されなかった. 事前タスクは 500 語の留学についての英文を読み, 内容理解確認問題 (多肢選択) 6 問に解答し, 7 題の記述問題は留学について多角的に検討する課題であった. 事後タスクは, 国際化する日本の観光産業についての 500 語の英文を読み, 内容理解確認問題 (多肢選択) 6 問に解答し, 7 題の記述問題は日本の観光産業を多角的に考える課題であった. 表 1 は, 二つのクラスの記述統計を示し, 表 2 は対応ある  $t$ -検定の結果を示す.

事前テストと事後テストの平均値には差が見られず, 分散にも差がないが, 事前・事後テストの問題間の相関は 0.263 で, 弱い関係性しか見られなかった. したがって, 事前テストと事後テストは, 完璧な平行テストであったとは, 言い難い. 従来から言われているように, 平行テストを作成することは困難であるので, 次回のパイロット実験では, 事前テストを事後テストは同じテストを用い, 記憶の効果を最小にするため, 少なくとも 3 か月の期間をおいて, 実験すべきことが示唆された. しかし, 平均値と  $SD$  の値が近似していることから, 事前・事後テストは, 今回は平行テストとみなし, 何らかの教育効果の示唆を得ることにした.

表 1 A・B クラスの 1 回目と 5 回目の記述統計

	$M$	$n$	$SD$
1 回目	23.59	39	12.65
5 回目	25.72	39	10.44

表 2 対応のある  $t$ -検定の結果

	1 回目と 5 回目の差
平均値	-2.13
標準偏差	9.78
平均値の標準誤差	1.57
差の 95% 信頼区間 (上限)	1.04
(下限)	-5.30
$t$ 値	-1.36
自由度	38
有意確率 (両側)	0.182

## 5.4 実験結果2：教育効果

表3および表4に示すようにAクラスでは、事前テストと事後テストの成績に有意差があり、教育効果が認められる。また、表5はBクラスの記述統計をしめし、表6は対応あるt検定の結果を示す。

Bクラスの事後テストでは、事前テストより点数が悪くなっている。内容確認問題の事前テストでは、平均4.8, *SD*(0.951)と事後平均4.4, *SD*(1.095)で、平均値に差はない。差をもたらしめているのは記述問題である。記述事前平均23.05 *SD*11.01, 事後平均21.95 *SD*11.05であった。

表3 Aクラスの1回目と5回目の記述統計

	<i>M</i>	<i>n</i>	<i>SD</i>
1回目	18.47	19	12.08
5回目	25.16	19	9.33

表4 Aクラスの対応のあるt検定の結果

	1回目と5回目の差
平均値	-6.68
標準偏差	7.81
平均値の標準誤差	1.79
差の95%信頼区間 (上限)	-2.92
(下限)	-10.45
t値	-3.73
自由度	18
有意確率(両側)	0.002

表5 Bクラスの1回目と5回目の記述統計

	<i>M</i>	<i>n</i>	<i>SD</i>
1回目	28.45	20	11.44
5回目	26.25	20	11.62

表6 Bクラスの対応のあるt検定の結果

	1回目と5回目の差
平均値	2.2
標準偏差	9.64
平均値の標準誤差	2.16
差の95%信頼区間 (上限)	6.71
(下限)	-2.31
t値	1.02
自由度	19
有意確率(両側)	0.320

Aクラスでは、教育効果が認められ、Bクラスでは認められない理由は、Aクラスは10日の演習直後に事後テストを行ったが、Bクラスは期末試験の後に1週間遅れで受験したことが影響しているとも考えられる。次のパイロットでは、2クラスとも試験時期を同じにすべきことが示唆された。さらに、記述問題に対する主観的難易度判定に差があるかを考察してみた。AクラスとBクラスを比較してみると、結果は次の表ようになる。逆転項目なので、1に近いほど「非常に難しい」5に近いほど「非常に簡単」を示す。

表7と表8はAクラスとBクラスの事前・事後テストにおける主観的困難度の記述統計結果であり、次の表9と表10は、成績と主観的困難度の相関を示している。Bクラスの事後テストではQ4以外、点数と難易度の主観的評価の相関は高く、難易度平均も「とても難しい」か「やや難しい」と判定されているが、正の相関なので、記述の成績の良い生

表7 Aクラスの事前・事後テストにおける主観的困難度の記述統計

	Aクラス			
	事前		事後	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Q1	2.63	0.94	2.33	1.19
Q2	2.32	0.82	2.17	1.25
Q3	1.84	0.76	2.06	1.00
Q4	1.89	0.83	2.17	1.10
Q5	1.58	0.93	2.28	1.07
Q6	1.84	0.84	2.28	1.07
Q7	1.68	0.85	2.22	1.06

表8 Bクラスの事前・事後テストにおける主観的困難度の記述統計

	Bクラス			
	事前		事後	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Q1	2.73	0.83	2.53	1.12
Q2	2.57	0.82	2.37	1.12
Q3	2.33	0.69	1.89	1.05
Q4	2.17	0.88	2.11	1.05
Q5	2.03	0.69	1.95	0.91
Q6	2.30	0.83	2.21	1.13
Q7	2.03	0.89	2.05	0.78

表9 Aクラスの記述問題事後テストに対する主観的困難度と点数の相関

	M	SD	主観的困難度と点数の相関
Q1	2.53	1.12	0.40
Q2	2.37	1.12	-0.07
Q3	1.89	1.05	-0.01
Q4	2.11	1.05	-0.32
Q5	1.95	0.91	-0.06
Q6	2.21	1.13	-0.07
Q7	2.05	0.78	0.49

表10 Bクラスの記述問題事後テストに対する主観的困難度と点数の相関

	M	SD	主観的困難度と点数の相関
Q1	2.33	1.19	0.42
Q2	2.17	1.25	0.43
Q3	2.06	1.00	0.48
Q4	2.17	1.10	0.14
Q5	2.28	1.07	0.31
Q6	2.28	1.07	0.28
Q7	2.22	1.06	0.06

徒は、課題が簡単と感じ、成績の悪い生徒は、課題が難しいと感じていると解釈できる。この点は、次回、生体情報の実験で確認すべき点である。

Aクラスも主観評定の平均値は1か2なので、Bクラスと同様の主観的難易度を示しているが、相関は低く、しかも負の相関が多い。Q1とQ7以外は負の相関になっている。このことは、記述問題の成績が良い生徒ほど、難易度を高く感じ、成績が悪い生徒ほど、課題の難易度を簡単だと回答していることになる。この点も次回の生体情報の実験で精査すべき点である。

Aクラスでは、成績と主観的困難度評価が正の相関を示し、Bクラスでは、逆に、負の相関を示す傾向があった。すなわち、Aクラスの成績の良い生徒は課題が簡単と答えているのに対し、Bクラスの成績上位の生徒は課題が難しいと判断していることになる。この差はどこから来るのか？ 次回の調査では、インタビューをするか、自分の記述解答に対する満足度も調査すべきと考える。

### 5.5 本章のまとめ

今回のパイロットでは、Aクラスでは教育効果が

あり、Bクラスでは教育効果の証拠は示されなかった。クラスの雰囲気にも差があったし、事後テストの時期に期末テストがあり、期末テストの前に事後テストを受けたAクラスでは、実験効果があった。事後テストの時期や事前と事後のテストも完全な平行テストではなかったという欠陥もあった。また、主観的な難易度の調査でもAとBでは、極端な差が観察された。この点は、生体情報の調査で、確認すべきであるし、インタビューか記述解答に対する満足度や確信度も調査するべきである。

## 6. おわりに

(担当：湘南工科大学 梅澤克之)

本研究部会の活動が実り豊かなものになりましたのは、研究部会の皆様、会員の皆様のご支援によるものに他なりません。部会設立をお認めくださった経営情報学会をはじめ皆様に改めて感謝申し上げます。

### 参考文献

- [1] 梅澤克之, 中澤真, 小林学, 石井雄隆, 中野美知子, 平澤茂一「言語学習を対象とした時空を越えて相手を感じられる自学自習システムの開発の概要」『2019年電子情報通信学会ソサイエティ大会予稿集』Vol. 2, 2019年9月, p. 276.
- [2] 梅澤克之, 中澤真, 石井雄隆, 小林学, 中野美知子, 平澤茂一「簡易脳波計を用いたプログラミング学習時のケアレスミスの検出」『電子情報通信学会教育工学研究会 (ET) 予稿集』2020年3月, pp. 135-140.
- [3] 梅澤克之, 中澤真, 後藤正幸, 平澤茂一「学習履歴を活用したデバッグ練習問題抽出システムの開発」『第18回情報科学技術フォーラム (FIT2019) 予稿集』Vol. 3, 2019年9月, pp. 331-334.
- [4] 後藤裕介, 石田崇, 梅澤克之, 齋藤友彦, 中澤真「プログラミング教育におけるラーニングアナリティクスの事例」『経営情報学会誌』第28巻, 第1号, 2019年6月, pp. 49-52.
- [5] 梅澤克之, 石田崇, 中澤真, 平澤茂一「グループ分け反転授業の2年間の実授業への適用と評価」『電子情報通信学会教育工学研究会 (ET) 予稿集』2019年6月, pp. 7-12.
- [6] 梅澤克之, 石田崇, 中澤真, 平澤茂一「グルー

ブ分け反転授業の2年間の実授業への適用とアンケート評価」『情報処理学会コンピュータと教育研究会（CE）予稿集』2019年6月，pp. 1-8.

- [7] 梅澤克之，石田崇，中澤真，平澤茂一「グループ分け反転授業における自習時の理解度について」『電子情報通信学会教育工学研究会（ET）予稿集』2019年7月，pp. 21-24.
- [8] 中澤真，荒本道隆，後藤正幸，平澤茂一「編集履歴可視化システムを用いた Learning Analytics ～

Scratch を用いた初等教育向けプログラミング教育における学習者の思考パターン分析」『情報処理学会第78回全国大会講演論文集』2016年3月，pp. 4-531-4-532.

- [9] 中澤真，梅澤克之「授業支援のためのビジュアルプログラミングの編集履歴に基づく学習分析」『情報処理学会第82回全国大会講演論文集』2020年3月，pp. 299-300.