

# 詳細な学習ログを用いた英語リーディング過程の分析

## ～(3)リーディング過程における学習者モデル～

中澤 真 <sup>†</sup> 会津大学 <sup>†</sup> 短期大学部	梅澤 克之 <sup>‡</sup> (株)日立製作所 <sup>‡</sup> IT ビジネスサービス本部	小林 学 <sup>††</sup> 湘南工科大学 <sup>††</sup> 情報工学科
小泉 大城 <sup>†††</sup> 青山学院大学 <sup>†††</sup> 理工学部	後藤 正幸 <sup>††††</sup> 早稲田大学 <sup>††††</sup> 理工学術院	平澤 茂一 <sup>†††††</sup> 早稲田大学 <sup>†††††</sup> 理工学術院 総合研究所

### 1. はじめに

学習者に対する適応的な学習支援を e ラーニングシステムで実現するためには、学習者モデルの構築が不可欠である。また、インストラクショナル・デザインに基づく教材や授業構成の見直しの際にも学習者モデルは重要な役割を担う。しかし、従来の Learning Analytics では学習状況やそのプロセスを把握するために用いる学習ログの粒度が粗く、正確なモデリングが難しかった[1]。

そこで本研究では、早稲田大学の授業『Discussion Tutorial English』で実際に使用している英語教材を電子化したものを学生に事前学習させ、学習時の操作履歴を細かい粒度で記録した。このデータに基づき、英語のリーディング過程における学習者の行動特性や学習のアプローチに関する学習者モデルを示す。

### 2. 学習ログの収集と前処理

粒度の細かい学習ログの収集と分析のために、筆者らはページ単位の閲覧行動やウィンドウのアクティブ・非アクティブを検出できる Web ベースの電子教材を用いてきた[2]。今回は教材内に選択問題や記述問題を組み込み、この解答時の行動も詳細に記録することができる荒本らのシステムを用いてデータを収集し、これを分析した[3]。学習者は全 28 ページの教材で英語のリーディングをしつつ、この教材内に組み込んだ選択問題 6 問(ページ 11～16)と記述問題 10 問(ページ 17～20, 22～26, 28)に取り組むことになる(英語学習の具体的内容や条件およびログの収集システムについては文献[3][4]を参照)。

Learning Analytics においてデータの前処理は重要な過程であり[1]、今回のように粒度が細かいログを用いる場合は、さらに慎重にクリーニングする必要がある。とくに、本研究の電子教材はシーケンシャルに構成されているため、閲覧行為ではなく単なるページ遷移もログとして記録されてしまう可能性を考慮する必要がある。実際、一つのページの 1 回の閲覧行為で 1 秒未満の

ログの頻度が極端に多いことを図 1 で確認できる。

このため、本研究では 1 秒未満の閲覧行為のログは解析対象から除外する。なお、クリーニングの対象を 2 秒未満とすると、タイトルや指示内容のみが記載された短い閲覧時間で内容を把握できるページ(図 2 のページ 10, 21, 27)の履歴がほとんど除外されてしまうため適当でないと判断した。図 2 は 1 秒以上の閲覧行為のログのみを集計した結果であるが、ページ 17 以降の記述問題の平均閲覧時間が総じて長いことや、記述問題の中で最も平均点が高かったページ 18 の閲覧時間が短いことなど、特性がよく現れていることが確認できる。

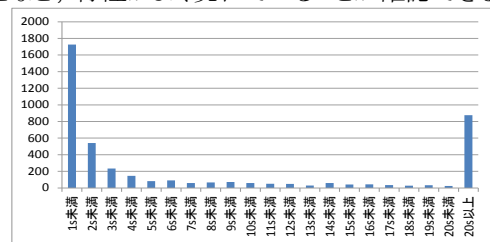


図 1: 1 回の閲覧行動あたりの閲覧時間の度数分布

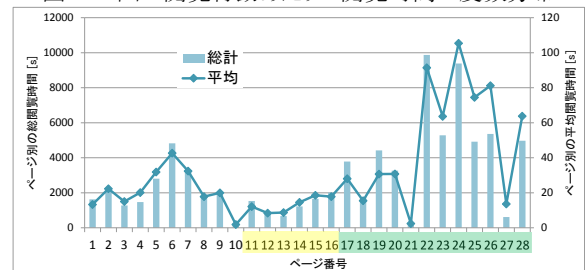


図 2: ページ別の総閲覧時間と 1 回あたりの平均閲覧時間

### 3. 教材の遷移モデル

先に示した図 2 のように、収集した学習ログを集計するだけでも、学習環境の特性をある程度確認することができるが、今回はログの時系列性に注目して閲覧ページの遷移について解析をした。図 3 は全学習者の閲覧行動に基づき、ページ  $i$  からページ  $j$  に遷移する確率を  $i$  行  $j$  列の要素として示した結果である。グラデーションが濃い部分ほど値が大きいことを表している。当然のことながらページ  $i$  の後にページ  $i+1$  に遷移する確率が高いが、前のページに戻る遷移や、同一のページでの自己遷移の確率が高いページの存在も確認できる。

この遷移確率行列に基づき、現在閲覧中のページより前のページに戻った場合(Back)、直後のページに遷移した場合(Next)、2 ページ以上先のページに遷移した場合(Forward)、他の作業をしてから同じページに自

An investigation into reading process in English, by making use of detailed learning logs. - (3) Learner model in the reading process -

<sup>†</sup>Makoto Nakazawa, University of Aizu

<sup>‡</sup>Katsuyuki Umezawa, Hitachi, Ltd.

<sup>††</sup>Manabu Kobayashi, Shonan Institute of Technology

<sup>†††</sup>Daiki Koizumi, Aoyama Gakuin University

<sup>††††</sup>Masayuki Goto, Waseda University

<sup>†††††</sup>Shigeichi Hirasawa, Waseda University

己遷移した場合のそれぞれの確率を教材のページ単位で算出した. この4つの値のリストを各ページの特徴ベクトルと考え, これを k-means 法でクラスタリングすることにより, 教材の遷移モデルを求めた. なお, クラスタ数は4, 距離はユークリッド距離を用いて計算している.

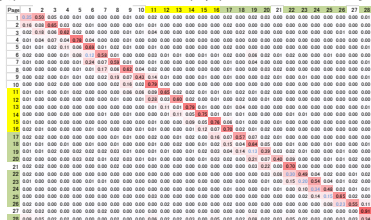


図 3: 閲覧回数に基づくページ別の遷移確率

図 4 に示したように<sup>1</sup>, Back への状態遷移確率が最も高いクラスタ P4 は, 平均得点が低かった記述問題のページと, リーディング対象の英文の最終ページで構成されており, 難易度の高いページ群のモデルと考えられる. 一方, Next へ遷移する確率が最も高いクラスタ P3 は, 選択問題のページやタイトルページなどで構成されており, 情報量が少ないあるいは難易度が低いページ群のモデルとなっている. また, クラスタ P2 の特徴は自己遷移確率が他よりも高い点であり, 単語を調べるなど教材の閲覧以外の操作をしながら取り組む内容のページ群で構成されている.

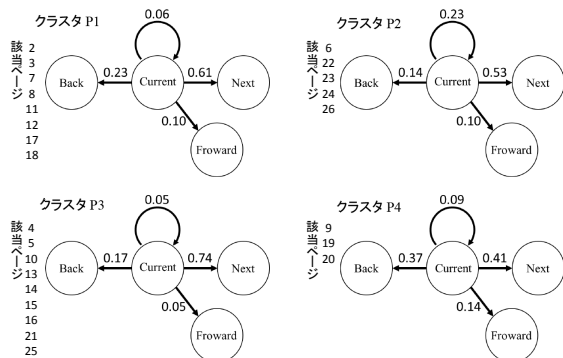


図 4: 教材のページ遷移モデル

#### 4. 学習者の遷移モデル

次に学習者の遷移モデルを考える. 4つの状態の遷移確率を成分として特徴ベクトルを考える点は先と同じであり, これを学習者ごとに選択問題の全ページでの平均遷移確率を用いたモデルと, 記述問題の全ページでの平均遷移確率を用いたモデルの二つを考える. 図 5, 図 6 は二つのケースをそれぞれ k-means 法でクラスタリングした結果である. 記述問題のケースでは, 図 4 の遷移モデルとほぼ同じタイプのモデルで構成されており, Back への状態遷移確率が最も高いクラスタ UE4 は, 英語力の低い学習者の遷移モデルとなっていることがわかる. 一方, 英語力の高い学習者のモデルは UE2 であり, 自己遷移確率が高い点の特徴である. 英文を読み返すことよりも, 英作文の推敲に力点を置いていると考えられる.

選択問題のケースではクラスタ UM3 のモデルのように, 英語力の低い学習者は戻ることも調べることもほとんどせずに解答し, 結果として正答率も低くなっていることが明らかになった.

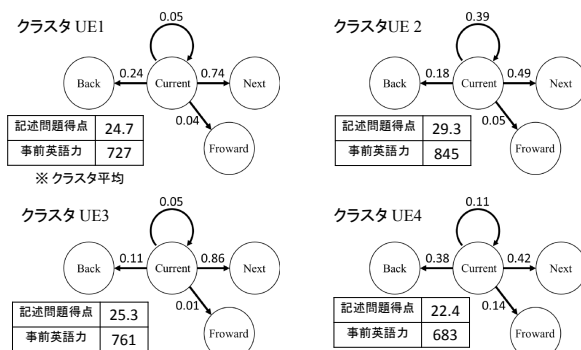


図 5: 記述問題に対する学習者の遷移モデル<sup>2</sup>

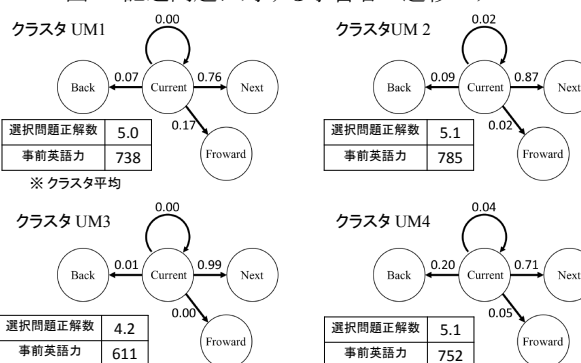


図 6: 選択問題に対する学習者の遷移モデル

#### 5. まとめと今後の課題

本研究で示した教材のページ遷移モデルを用いることで, ページ単位で教材の改善や再設計を容易に行うことが可能になる. また, 学習者の遷移モデルは, 学習プロセスに問題のある者を抽出し, 学習指導の支援機能としての可能性を示している.

今回の解析では閲覧回数に基づき遷移確率を算出したが, 遷移後のページ閲覧時間に基づく遷移モデルについても今後解析する予定である.

#### 謝辞

本研究の一部は, 会津大学短期大学部競争的研究費および独立行政法人日本学術振興会学術研究助成基金助成金基盤研究(C) 23501178の助成による.

#### 参考文献

- [1] J.A. Larusson, B. White (Eds.), Learning Analytics: From Research to Practice, Springer, 2014.
- [2] 中澤真, 小泉大城, 後藤正幸, 平澤茂一, “詳細な学習履歴を活用した学習者行動の分析,” 情報処理学会第76回全国大会 講演論文集, pp.4-357-4-358, 2014年3月.
- [3] 荒本道隆, 佐藤一裕, 中原歌織, 平澤茂一, “詳細な学習ログを用いた英語リーディング過程の分析～(1) リーディング過程の可視化～,” 情報処理学会第77回全国大会 講演論文集, 2015年3月.
- [4] 中野美知子, 吉田諭史 ほか, “詳細な学習ログを用いた英語リーディング過程の分析～(2) ログデータから見た成績との関係～,” 情報処理学会第77回全国大会 講演論文集, 2015年3月.

<sup>1</sup> 1 ページ目と最終ページ, 最終ページの前ページは, Back, Forward など遷移が起こりえないケースがあるため除外してある.

<sup>2</sup> 事前英語力は学習前のアンケートから TOEIC, WeTEC などの点数を換算して算出[4].