

## 読解過程検定ツール OPERATION2 の評価

## Evaluation of a Reading Process Monitoring Tool OPERATOIN2

小谷 克則† 吉見 毅彦‡ 九津見 毅†† 佐田 いち子†† 井佐原 均†

Katsunori Kotani Takehiko Yoshimi Takeshi Kutsumi Ichiko Sata Hitoshi Isahara

## 1. はじめに

本稿は、筆者らが開発した読解過程検定ツール OPERATION2 を紹介し、検定ツールを用いた読解能力測定実験結果を報告する。筆者らは、テキスト読解処理における局地的な処理速度変化の検出を目的とするツール OPERATION2 を開発した。この OPERATION2 の評価を行うために、第二言語(英語)による読解実験を行った。その結果、本ツールが収集したデータを基に、第二言語学習者の読解能力を適切に推定できることを確認した。

人間の読解過程を検定する手段として、読解速度の計測、眼球運動の計測、脳波や脳電位の計測などがある。これらの計測手段の中でも、眼球運動や脳波・脳電位による計測法[Vitu and Regan 1995; Tan et al. 2001]は、運用のコストなどの問題から、これまで第二言語教育の場において用いられることはほとんどなかった。一方、読解速度計測法は、他の計測法と異なり、第二言語教育において採用されることもある。しかし、読解速度計測法だけでは、テキスト内に表出する未習得箇所を特定することが困難などの理由から、理解度テストと共に用いられることが多い。

従来の読解速度計測は、テキスト全域における処理過程が等速度で行われると想定する必要がある。しかし、実際のテキストには、学習者の未習得の単語・構文などが含まれるため一定速度を保つことはないと思われる。このような問題点を解決するために、これまでは理解度テストや穴埋め問題などを用いて、テキスト内の局所的な理解度を推定してきた。これらの理解度テストは、テキスト全域に渡って設定することが困難であることから、テキスト内に表出する全ての未習得部分を特定することは困難である。したがって、学習者の未習得部分など処理時間に影響を与える箇所をテキスト全域に渡って記録する必要がある。

このような考えに基づき、筆者らはテキスト読解中の局所的な処理時間変化を記録するツール OPERATION2 を開発した。本稿が提案する読解過程検定ツール OPERATION2 は、第二言語学習者にとって指導が必要な箇所を特定することができる。さらに、いわゆる理解度テストをすることなく、読解中に生じる問題を特定できることから、自律学習にも用いることができるなど、第二言語教育に貢献するものであると考えられる。

## 2. 読解過程検定ツール OPERATION2

本ツールの基本的なシステム構成は、図1に示すとおりである。システムの基本構成部分である読解過程検定モジュールが収集した読解時間データを読解過程表示モジュールにおいてグラフ表示することが可能である。指導者はこのグラフ表示機能を用いることによって、学習者の問題を特定することが可能となる。

想定されるシステム使用者は、学習者とその指導者である<sup>1</sup>。学習者は指導者が用意したテキストをツールが表示する読解範囲毎に読解を行う<sup>2</sup>。学習者がシステム上で読解を行うと、システムは、(1)読解位置、(2)読解時間、非読解時間(例えば、辞書引きに要した時間など)を計測・記録する。

図2は読解時間測定モジュールの画面である。ユーザは、まず、開始ボタン(ツールバーの下の左から二番目の「Go!」ボタン)をクリックし、テキストの読解を開始することをシステムに伝える。開始ボタンをクリックした後、ユーザが表示用ボタンの上にマウスポインタを移動させる(表示用ボタンをクリックする必要はない)と、割り当てられた文が画面に表示される(図2の状態)。ユーザは、第1文を読み終えると、その表示用ボタンの上からマウスポインタをはずし、次の第2文の表示用ボタンへマウスポインタを移動させる。すると、画面上から第1文が消え、代わって第2文が画面上に表示される。このような操作をテキストの最終文まで繰り返すことで、テキストの読解が進んでいく。テキスト全体を読み終えると、ユーザは、終了ボタン(開始ボタンの右隣のボタン)をクリックし、読解終了をシステムに伝える。

ユーザは、第*i*文を読んだ後、第*i+1*文を読む必要はなく、次に読む文を自由に選択することもできる。また、テキスト全体を読み終えた後、読み直したい文を選んで再度表示させることもできる。なお、読解実験の種類によってはユーザのこのような自由な振る舞いを制限したい場合もありえるが、現状のシステムには文の選択順序を制限する機能は備わっていない。

表示用ボタン上にマウスポインタを置いたときに表示させる単位は、文である必要はなく、文よりも大きな単位である段落などであっても、逆に、小さな単位である句などであってもよい。

読解時間の測定は次のように行なわれる。読解時間の測定は、ユーザが開始ボタンをクリックした時点から始まり、終了ボタンをクリックした時点で終了する。表示用ボタン上にマウスポインタが置かれたとき、システムはその時刻を記録する。表示用ボタン上からマウスポインタがはずされたときにもその時刻を記録する。そして、両時刻の差をその文の読解時間とする。また、ある文の表示用ボタン上からマウスポインタがはずされ、次の文の表示用ボタン上にマウスポインタが置かれるまでの時間も同時に記録される。

†情報通信研究機構, ‡龍谷大学, ††SHARP(株)

<sup>1</sup> 学習者と指導者が同一者の場合、自律学習支援ツールとして使用することが可能である。<sup>2</sup> 入力テキストは、XML形式ファイルである。指導者は、この入力テキストへのタグ付けによって読解範囲を指定することが可能である。

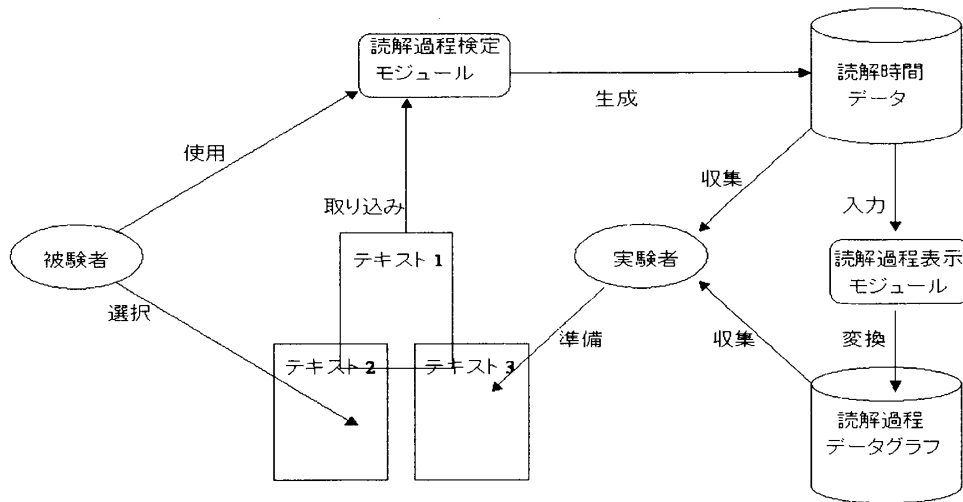


図 1. システム概略図

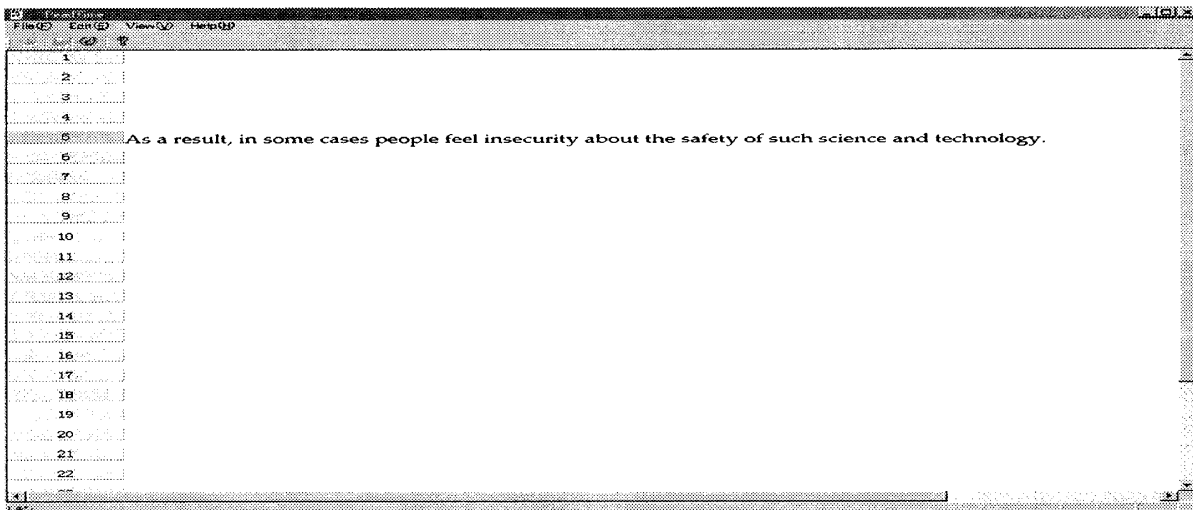


図 2. 計測画面

さらに、システムは記録・収集したデータから読解過程をグラフ表示化する。指導者は、読解時間データグラフを基に読解処理時間が変化している箇所を簡便に検出することができる<sup>3</sup>。読解範囲毎の累積的読解時間を基にグラフを作成すると図3の通りである。

### 3. 読解能力測定実験

本実験は、OPERATION2 による読解過程検定の有効性を確認することを目的に行った。実験参加者 22 名は、全員、日本語母語話者である。この参加者は第二言語である英語運用能力に応じて二群にグループ分けした。第一群 11 名 (A-グループ) が英語運用能力の高いグループであり、第二群 11 名 (B-グループ) が運用能力の低いグルー

プである<sup>4</sup>。読解対象テキストは、TOEIC 準拠の市販教材を用いた。

実験課題は、OPERATION2 が計測する読解過程データに基づき、実験参加者の読解能力を適切に推定できるかであった。OPERATION2 の開発にあたり、筆者らは読解に必要な処理過程を言語処理部分と情報処理部分に区別すべきだと考えた。ここで言う言語処理部分とは、テキスト内の言語情報を理解するための過程である。情報処理部分とは、テキスト内の重要箇所の抽出や類似箇所の特定を行うなど背景知識情報など言語情報以外の情報を用いて処理する過程である。このように読解過程を区別することにより母語と第二言語における読解の特徴がうまく説明できる。

<sup>3</sup> 読解時間データは、TXT 形式ファイルによって作成される。そのため、時間変化量の測定などをテキストエディタや Perl 等のプログラムにより検出が可能となる。したがって、大量の読解時間データであっても、読解時間変化箇所を特定することが可能である。

<sup>4</sup> A-グループ参加者が英語教員などで構成され、B-グループが英語学習者 (大学生) で構成される。グループ分けは、参加者の自己申告に基づいて行った。今回の実験参加者は、教員とその学習者であることから運用能力差を自明のものであるとした。今後、実験前に能力測定などを行い、更に詳細な能力差測定の可能性を検討する。

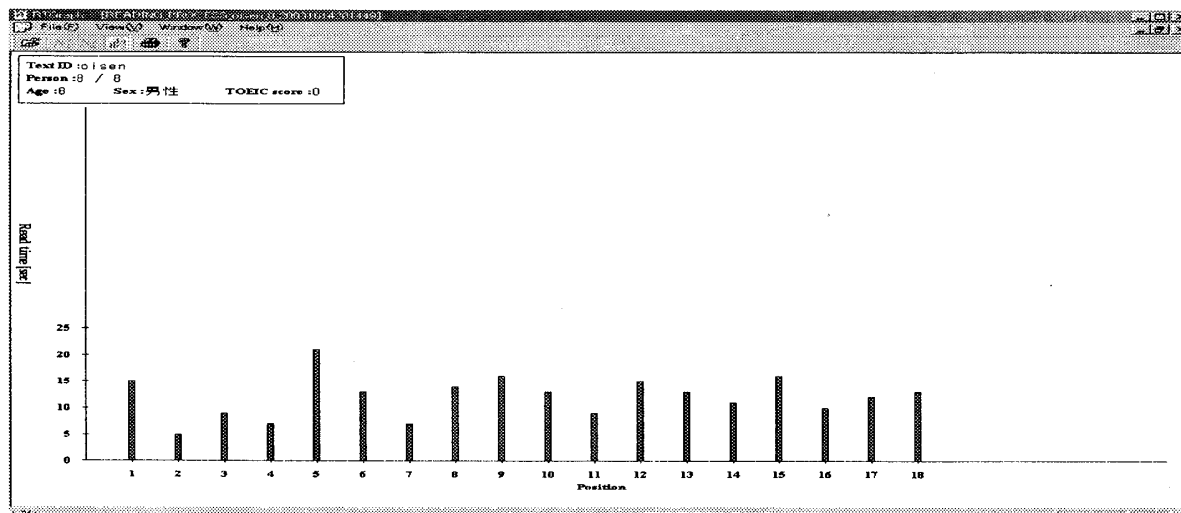


図3. グラフ表示画面

母語言語による読解の場合、言語処理過程の多くの部分が自動化されていると考えられる。したがって、同一言語の母語話者間において、その処理過程は均一であると考えられる。この均一性は、一般的な言語能力の均一性からも支持される。一方、情報処理過程は、言語能力以外の要因に負うところが大きく、同一言語話者間においても不均一な処理過程とであると考えることができる。[Ganong 1980; Goodman et al. 1981; Healy 1980; Kennedy et al. 2000; Seidenberg et al. 1984; Wright & Garrett 1984]

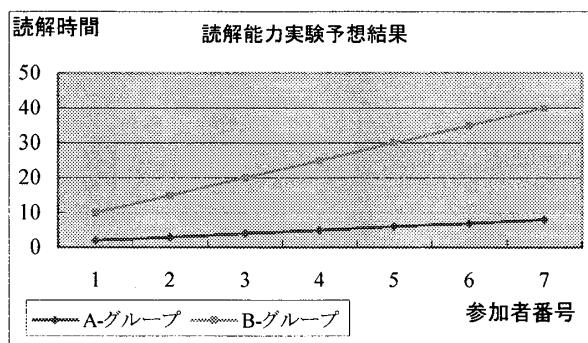
第二言語の場合、情報処理過程だけでなく、言語処理過程も、不均一な処理過程であると考えられる。この言語処理過程における個人差は、第二言語能力の不均一性が示すとおりである。先ほどの母語の場合と比較すると、第二言語における言語処理過程の不均一性は、処理過程が自動的に行われない部分があることに起因すると考えられる<sup>5</sup>。非自動化部分は、学習者が学習項目などを参照に処理すると考えられる。筆者らは、この非自動化部分の存在が読解過程において読解時間の増加として表出すると考えた。

本実験は、第二言語処理過程における非自動化部分の抽出が目的である。したがって、言語処理過程部分だけが起動し、情報処理過程が起動しないようにテキストの統制を試みた。そこで、実験課題として用意したテキストは、何の文脈も持たない単文の集合とした。読解テキストが単文集合であるため、読解実験に一般に用いられる重要箇所抽出や要約等のタスクを実験参加者に課さなかった。今回の実験で参加者に課したタスクは、読み進めることであった。これらの想定を基に、本実験は言語処理過程において非自動化部分が多ければ処理時間の増大として表出すると考え、

<sup>5</sup> 筆者らは、読解を、視覚性の情報である文字およびその系列から音韻や意味など視覚以外の言語性内容を喚起する過程の総体と想定する。そして、読解時間の計測により推定可能なタスクとして①文字(系列)が既知のものであるとわかる過程(視覚性認知)、②文字(系列)の意味を理解する過程(意味理解)、③文字(系列)の読み方を喚起する音韻過程の三つを想定する。[河村、溝淵 1998; Friderici & Lachamnn 2002 など]

読解能力に応じて処理時間が異なると仮定した。したがって、実験によって確認すべきことは、実験参加者グループ A, B の間の読解時間に統計的な有意差があることである。予測される読解パターンをグラフ 1 に示す。

グラフ 1 は、読解能力が異なる A-グループと B-グループの読解過程を検定ツール OPERATION2 により測定した場合に、予測される結果である。グラフ縦軸が読解に要した時間を示す。グラフ横軸が参加者番号を示す。ここでは読解時間による順位によって両グループの参加者に番号を割り当てた。読解能力が高い A-グループに属する参加者全員が B-グループに属する参加者よりも処理に要する時間が短いと予想される。



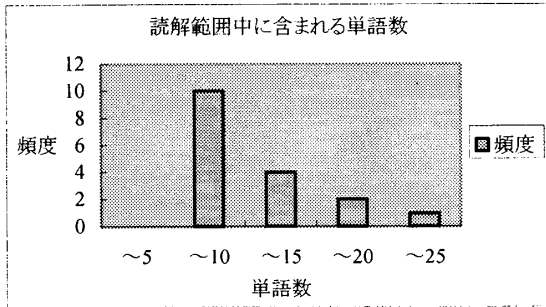
グラフ 1. 実験結果予想グラフ

実験参加者は、A-グループ 11 名と B-グループ 11 名であった。A-グループ参加者は読解能力が B-グループより高いと考えられる。両グループ共に市販されている TOEIC 準拠テキストから抜き出された英文 18 文を読解した。読解の際に、参加者らは全員、検定ツール OPERATION2 を用いて読解した。

検定ツール OPERATION2 は、1~2 文を一つの読解範囲として参加者らに提示する<sup>6</sup>。1 読解範囲に含まれる単語数

<sup>6</sup> 教材中に例文として提示されるものを用いた。その際に、2 文で一つの例文と提示されている場合はそのまま用いた。結果、読解範囲中に示される文数が 1~2 文となった。

は6~23単語であった。一範囲あたりの単語数の分布はグラフ2に示す通りである<sup>7</sup>。



グラフ2. 実験対象テキスト情報

読解過程検定ツール OPERATION2により、全22名の読解時間データが収集された。収集した読解時間データをA、Bグループ毎に読解範囲あたりの平均読解所要時間とその標準偏差を算出した結果は表1に示される通りである。

	A-グループ	B-グループ
最小値	2.8300	8.6100
最大値	8.4400	28.2200
平均値	4.7718	14.2318
中間値	4.2800	12.1700
レンジ	5.6100	19.6100
分散	3.1141	30.9747
標準偏差	1.7647	5.5655

表1. 読解実験結果

(秒)

各グループの読解平均時間データに対して Welch の t 検定を行った結果、両群間の平均値に有意水準 5% で有意差が認められた<sup>8,9,10</sup>。つまり、読解過程検定ツール OPERATION2 が収集したデータに基づき、読解能力の推定が可能であることが確認できた<sup>11</sup>。

#### 4. まとめ

読解過程検定ツール OPERATION2 が収集した読解時間データに基づいて、読解能力の異なるグループ間の差を推定できることを確認した。本実験結果が示すのは、テキス

ト内の局所的な読解時間に基づくデータから読解能力を推定することが可能であるということである。

今後、読解能力別に分類した多数のグループ間において同様の実験を行い、OPERATION2 の性能をさらに検証する。また、グループ間の有意差だけでなく、テキストに含まれる文間の可読性判定などの課題も残される。

#### 参考文献

- A. D. Friderici, & T. Lachmann (2002) From Language to Reading and Reading Disability. In E. Wirtunk, A. D. Friderici & T. Lachmann editors, *Basic Functions of Language, Reading, and Reading Disability*, 9-21. Kluwer Academic Publishers, Massachusetts.
- W. F. Ganong. 1980. Phonetic Categorization in Auditory Word Perception. *Journal of Experimental Psychology*, 6: 110-125.
- G. O. Goodman, J. L. McClelland, and R. W. Gibbs. 1981. The Role of Syntactic Context in Word Recognition. *Memory and Cognition*, 9: 580-586.
- A. F. Healy. 1980. Proofreading Errors on the Word *the*: New Evidence on Reading Units. *Journal of Experimental Psychology*, 6: 45-57.
- 河村満, 溝渕淳. 1998. 読みの脳内機構, 荻阪直行(編) 読み—脳と心の情報処理. 朝倉書店, 東京.
- A. Kennedy, R. Radach, D. Heller, and J. Pynte.. 2000. *Reading as a Perceptual Process*. Elsevier, Oxford.
- 小谷克則, 吉見毅彦, 九津見毅, 佐田いち子, 井佐原均. 2004. 読解過程システムの構築, 言語処理学会第10回年次大会発表論文集, 393-396.
- M. S. Seidenberg, G. S. Waters, M. Sanders, and P. Langer. 1984. Pre- and Post-lexical Loci of Contextual Effects on Word Recognition. *Memory and Cognition*, 12: 315-328.
- L. H. Tan, H.-L. Liu, C. A. Perfetti, J. A. Spinks, P. T. Fox, J.-H. Gao. 2001. The Neural System Underlying Chinese Logograph Reading. *Neurolinguage*, 13: 836-846.
- F. Vitu and J. K. O'Regan. 1995. A Challenge to Current Theories of Eye Movements in Reading. In J. M. Findlay, R. Walker, and R. W. Kentridge, editors, *Eye Movement Research: Mechanisms, Processes and Applications*, pages 381-392. Elsevier Science B. V., Amsterdam.
- B. Wright and M. Garrett 1984. Lexical Decision in Sentences. *Memory and Cognition*, 12: 31-45.

<sup>7</sup> 本実験では、実験参加者の読解能力差の推定を行うため、文中の単語数や単語難易度などの単語情報等が読解時間に与える影響は無視できるものとした。

<sup>8</sup> K 統計値と片側確率 P は、次の通りである。K 統計値: 0.56780(A-グループ); 0.88659(B-グループ), 片側確率 P: 0.50000(A-グループ); 0.25000(B-グループ)。

<sup>9</sup> 有意水準  $\alpha=0.05$  とした。自由度(10, 10)とし、F 統計値が 9.9467 であった。算出された片側確率 P は、0.00058 であった。

<sup>10</sup> 自由度を 11.9906 とし、t 統計値が -5.12372 であった。算出された両側確率 P は、0.00033 であった。

<sup>11</sup> 難易度の異なるテキストを用いた読解能力推定実験に関しては、小谷他(2004)参照。