

## 読解効率に基づく英文読解能力測定テストの開発とその信頼性・妥当性の検証

小谷 克則<sup>†1</sup> 吉見 毅彦<sup>†2</sup> 井佐原 均<sup>†3</sup>

本稿では、理解度（テキストの内容に関する設問に対する解答の正答率）と読解速度の積である読解効率に基づいて第二言語学習者の読解能力を測定するテスト（読解効率テスト）の信頼性と妥当性を検証した。本研究の読解効率テストは、日本語を母語とする英語学習者 102 人を対象として、TOEIC 準拠テキストを用いて実施した。テストの信頼性は信頼性の程度を示す KR-21 信頼性係数により検証した。テストの妥当性は内容の妥当性、構成概念妥当性、基準関連妥当性の観点から検証した。その結果、本研究の読解効率テストには十分な信頼性と妥当性があることが分かった。このことから、読解効率テストを用いて英語学習者の読解能力の一部を測定することができると思われる。また、読解効率テストを実施した結果得られる学習者の理解度と読解時間は、英文読解教育・学習を支援するシステムを機械学習によって構築するための訓練データとして利用することができる。

## Development and Validation of English Reading Test Based on Effective Reading Speed

KATSUNORI KOTANI,<sup>†1</sup> TAKEHIKO YOSHIMI<sup>†2</sup>  
and HITOSHI ISAHARA<sup>†3</sup>

In this paper we develop and assess an English reading test for learners of English as a second language. Our reading test examines reading proficiency based on effective reading speed, which is the combined measure of reading speed and comprehension, and English texts were chosen from TOEIC reading section. 102 participants who took this test were learners of English as a second language. We assessed reliability of the test with KR-21 reliability coefficient that measured how reliable the test results were. Test validity was assessed by content validity, construct validity and criterion-related validity. This assessment demonstrated both the reliability and validity of the effective reading speed test. From this result, we found that an effective reading speed test could measure a part of reading proficiency of learners of English as a second language. In addition, an effective reading speed test provides us with

learner data of reading time and comprehension rate, which are applicable as training data of machine learning for constructing computer-assisted language learning/teaching systems.

### 1. はじめに

近年、インターネットの普及などにより、英文テキストを読む機会が急速に増えつつある。また、大学英語教育において、これまでの英語の教科書を読むだけでなく、新聞や雑誌などの一般的なテキストをはじめ、報告書や論文などの専門的なテキストをある程度の量読むことが必要とされている<sup>1)</sup>。このように実践的な英文読解を行う場合、内容の理解はもちろん、読解の速さも重要となる。したがって、実践的な英文読解能力の習得を目標として学習者の読解能力を測定する場合、理解度（テキストの内容に関する設問に対する解答の正答率）で示される能力だけでなく、読解速度で示される能力も考慮する必要がある。

TOEIC (Test of English for International Communications) や TOEFL (Test of English as a Foreign Language) などは理解度に基づいて読解能力を測定するテストであるが、これらは、制限時間を設けることによって、読解速度に関連する読解能力を間接的に測定しているといえる。2 章で概観するように、読解能力を測定する際に読解速度を評価指標として直接利用する研究<sup>2),3)</sup>がある。また、読解速度を、テキストの読みやすさを推定する指標として利用する研究<sup>4),5)</sup>もある。このように、読解速度あるいは読解効率（読解速度と理解度の積<sup>6)</sup>）は様々な点から重要な評価指標として利用されている。

しかしながら、読解効率が第二言語学習者の読解能力の評価指標として有効かどうかはこれまで十分には確認されていない。2 章で述べるように、第二言語読解能力の評価指標としての読解効率の信頼性と妥当性を検証した Shizuka の研究<sup>7)</sup>もあるが、Shizuka の検証実験は、1 文を読む設定であり、複数の文から構成されるテキストを読む設定ではない。日常生活においてテキストを読む場合の英文読解能力を測定するためには、1 文を読解対象とするのではなく、テキストを対象とする必要がある。また、Shizuka が検証した妥当性は基準

†1 関西外国語大学  
Kansai Gaidai University

†2 龍谷大学  
Ryukoku University

†3 豊橋技術科学大学  
Toyohashi University of Technology

関連妥当性であり、内容的妥当性と構成概念妥当性については検証されていない。

そこで、本研究では英文テキストを読解対象として、コンピュータ上の読解時間・解答記録ツール<sup>8)</sup>を用いた読解効率テスト(読解効率を評価指標とする読解能力測定テスト)を実施した。本稿では、この読解効率テストの信頼性と3つの妥当性(内容的妥当性、構成概念妥当性、基準関連妥当性)を検証し、その結果を報告する。テキストを読む場合、学習者はテキストを構成する文と文の間のつながり<sup>9)</sup>を理解することなどの文脈処理を行わなければならない<sup>\*1</sup>。本研究は、このような状況下で読解を行ったときの理解度と読解時間を測定したという点で、文脈処理が要求されない設定である Shizuka の研究と異なる。

読解効率テストを実施した結果得られる、あるテキストについての学習者の理解度と読解時間(テキストを構成する文ごとの読解時間の和)は、英文読解教育・学習を支援するシステムを機械学習<sup>10)</sup>によって構築するための訓練データとして利用できる。実際、これまでに、英語学習者の未習得の語彙や文法の推定に利用することを目的とした文読解時間予測システム<sup>11)</sup>などが報告されている。このシステムは、ある英文読解能力を持った学習者がある文を読むときにかかる標準的な時間を予測する。この標準的読解時間と学習者が実際に文を読むのに要した読解時間とを比較し、標準的読解時間よりも実際の読解時間のほうが一定以上長ければ、その文にはその学習者の未習得の語彙や文法が含まれている可能性が高いと判断する。標準的読解時間よりも実際の読解時間のほうが一定以上長い文に対して詳細な分析を加えることによって、学習者の未習得の語彙や文法の推定が容易に行え、さらに、その学習者に適した読解支援につなげることができる。このようなシステムを構築するために、自然言語処理ツールを利用してテキストから抽出した語彙的、統語的、談話的特徴と、学習者の TOEIC のリーディングセクションの得点とを素性として機械学習が行われている。

## 2. 関連研究

### 2.1 読解能力と読解速度テスト

読解能力を測定する際、テキストの理解度を考慮したテストがこれまで数多く用いられてきた<sup>12)</sup>。このような理解度に基づくテストに加え、最近では、読解速度に基づくテストが用いられるようになってきた<sup>13),14)</sup>。Bell<sup>2)</sup>は、英語学習者約30人を対象に、精読による指導の効果と多読による指導の効果と比較する指標として、理解度に加え、読解速度を用

いた。Taguchi<sup>ら</sup><sup>3)</sup>は、英語学習者約30人を対象に、多読による指導の効果と繰り返し読みによる指導の効果と比較する際に理解度と読解速度を利用した。また、Tan<sup>ら</sup><sup>15)</sup>によると、学習者に読解速度を意識した読みを行わせるためにチャンキングの指導を行い、その後の訓練を通じて、読解速度が上がる事が確認されている。

### 2.2 読みやすさと読解速度テスト

テキストの読みやすさの評価指標として、文長や単語長などテキストの言語的特徴に基づくリーダビリティスコアがこれまで用いられてきた<sup>16)</sup>。これに対して、読解速度を指標として用いる研究もみられる。Urano<sup>4)</sup>は、英語学習者40人を対象に、第二言語(英語)における語彙の簡略化の効果を評価する指標として理解度と読解速度を用いた。飯野<sup>5)</sup>は、英語学習者約90人を対象に、リーダビリティの異なるテキストを読むのに必要な読解時間に基づいて繰り返し読みの効果を調査した。

### 2.3 読解効率テストの信頼性と妥当性

理解度に基づく読解能力測定テストはこれまでにその信頼性と妥当性が様々な研究によって検証されている<sup>17)</sup>。しかし、読解効率テストの信頼性と妥当性は、理解度によるテストの場合と比較して、十分に研究されてきたとはいえない。これまでに、第二言語学習者を対象とした読解効率テストの信頼性と基準関連妥当性を検証した研究として Shizuka の研究<sup>7)</sup>があげられる。

Shizuka は、英語学習者約40人を対象に、1文を読む設定で読解能力測定テストの信頼性と基準関連妥当性を検証した。具体的には、次のような、文表示方式と読解後に与える課題と評価指標の各選択肢を組み合わせた読解能力測定テストを実施し、どの組合せの場合に信頼性と基準関連妥当性が高くなるかを検証した。文表示方式の選択肢は、文の全単語を1度に表示する方式と、1単語ごとに表示する方式であり、読解後に与える課題の選択肢は、学習者の第二言語(英語)で書かれた4つの選択肢の中から正しくないものを1つ選び出すという課題と、学習者の第一言語(日本語)で書かれた4つの選択肢の中から正しいものを1つ選び出すという課題であり、評価指標の選択肢は、理解度、読解速度、読解効率である。信頼性の検証にはクロンバック  $\alpha$  信頼性係数<sup>18)</sup>を用い、基準関連妥当性の検証には、学習者に対して別途実施した3種類のテストの平均点と当該読解能力測定テスト結果との相関係数を用いた。

## 3. 読解効率テストの実施

本研究では、次のような手続きで読解効率テストを作成し、実施した。

\*1 3.3 節で述べるように、この読解時間・解答記録ツールのコンピュータ画面への表示単位を今回は1文に設定したため、学習者はテキスト全体を一括して読むのではなく1文ずつ読んでいくことになる。しかし、たとえ表示単位がテキストではなく文であっても、学習者は文脈処理を行わなければならないと考える。

### 3.1 読解効率テストの受験者

読解効率テストの受験者は日本語母語話者 107 人であった。全受験者 107 人のうち 5 人についてはテスト結果に不備が発見されたため分析対象外とし、残りの 102 人を有効受験者とした。性別は、女性が 94 人、男性が 8 人であった。平均年齢は 30.4 歳 (S.D. 9.2) であった (9 人は年齢不詳)。受験者の募集は、TOEIC を受験した経験があることを条件とし、参加報酬を支払うことを明示して行った。TOEIC 受験時期に関しては、読解効率テスト実施日より 1 年以内に受験した受験者が 62 人、1 年以上前に受験した受験者が 34 人、不明が 6 人であった。

分析対象の受験者 102 人の読解能力を把握するために TOEIC のリーディングセクションの得点 (以下、TOEIC 読解スコアと呼ぶ) について調査した。分析対象の受験者の TOEIC 読解スコアの平均点は 311.1 点 (S.D. 99.8)、最低点は 105 点、最高点は 470 点であった。TOEIC 公式資料 ([http://www.toeic.or.jp/toeic/data/data\\_avelist.html](http://www.toeic.or.jp/toeic/data/data_avelist.html)) によると、日本における TOEIC 受験者 (2010 年 3 月) の平均点は 266.4 点 (S.D. 96.2)、最低点は 5 点、最高点は 495 点である。分析対象の受験者と日本における TOEIC 受験者の読解能力を比較すると、分析対象の受験者のほうがより高得点域を占めていることが分かる。

読解効率テストは、受験者の読解能力を適正に反映している必要がある。本研究では読解効率テストを集団基準準拠テストと位置づける。このため、受験者の読解能力の分布は正規分布であることが望まれる。そこで、分析対象者の TOEIC 読解スコアの分布を調査した。分析対象者の TOEIC 読解スコアの分布は図 1 のとおりであった。図 1 の分布が正規分布であるかどうかを調べるために、コルモゴロフ・スミルノフ検定を行った。その結果、分布は正規分布と異なるといえないことが確認できた ( $K = 0.99$ ,  $p = 0.13$ )。

### 3.2 読解対象テキスト

読解効率テストにおける読解対象テキスト (付録 A.1 参照) として、TOEIC のリーディングセクションのパートに即した 84 テキスト (その内容に関する設問を含む) を TOEIC 準拠の問題集<sup>19),20)</sup> から収集した。本実験の読解対象テキストとして、TOEIC を用いた理由はテキストの種類にある。TOEIC で用いられているテキストはビジネスにおいて日常的に用いられているテキストであり、本研究が対象とする新聞や雑誌の記事や報告書や論文などのテキストとの類似性が高い。このため、TOEIC のテキストを理解するのに必要とされる英文読解能力 (構成概念) と、本研究が対象とする英文読解能力の間にも共通する部分があると考えられる。さらに、TOEIC は英語テストとして、その信頼性と妥当性が広く認められていることから、英語学習者の読解能力を測定するのに適しているテキストであると

受験者(人)

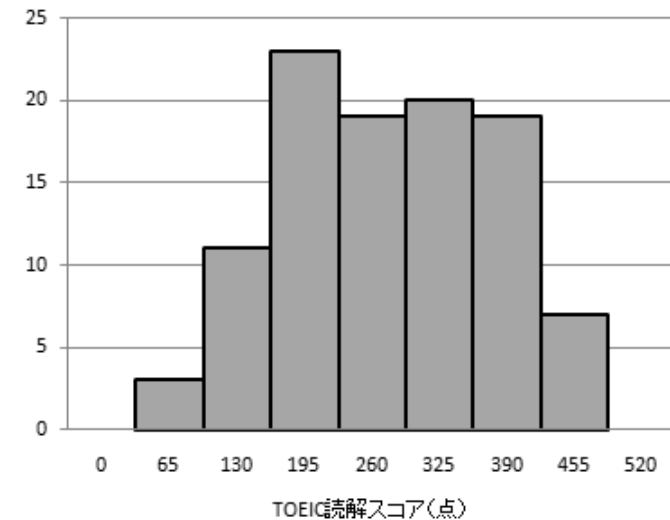


図 1 受験者の TOEIC 読解スコアのヒストグラム

Fig.1 Histogram of participants' TOEIC reading scores.

考えられる。

収集したテキストの読みやすさを調査するために、84 テキストについてリーダビリティスコア (Flesch Reading Ease)<sup>21)</sup> を算出した。その平均スコアは 58.3 (S.D. 16.1) であった。また、最低スコアは 15.0、最高スコアは 92.8 であった。TOEIC で使用される英文テキストのリーダビリティスコアは公表されていないため、収集したテキストが TOEIC のテキストと同程度の読みやすさであるかは確認できない。しかし、高校生・大学生向けテキストの Flesch Reading Ease が 30 から 60 程度であることから、収集したテキストの読みやすさは、高校生・大学生向けテキストの読みやすさと同程度であると想定できる。

### 3.3 読解効率テストの実施方法

読解効率は読解時間と理解度に基づいて算出される。このため、読解効率の算出に必要なテキストの読解時間と設問の解答を記録する必要がある。

84 テキストを 7 テキストからなる 12 のテキストセットに分け、各受験者に 1 つあるいは 2 つのテキストセットを割り当てた。この結果、各テキストセットには 10 人前後の受験

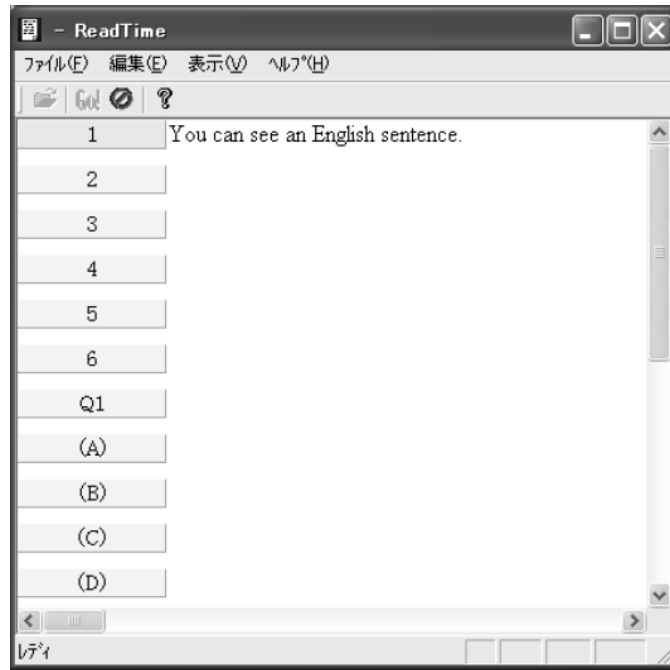


図2 読解時間・解答記録ツール

Fig.2 Tool for recording of reading time and answers.

者が割り当てられた。

読解時間と解答の記録には、コンピュータ上の読解時間・解答記録ツール(図2)<sup>8)</sup>を使用した。受験者はマウスポインタを操作し、画面上に表示された読解範囲ごとに読み進めていく。本研究の読解効率テストでは、1文を読解範囲とした。

図2の画面左にある読解アイコン(1~6)上にマウスポインタを移動させると、そこに格納されている読解範囲が表示される。マウスポインタを読解アイコンからはずすと読解範囲が表示されなくなる。受験者がマウスポインタを読解アイコン上に置いている時間がその読解範囲の読解時間として秒単位で記録される。

設問への解答にもこのツールを使用した。設問が格納されている設問アイコン(Q1)と解答の選択肢が格納されている選択肢アイコン((A)~(D))を用意した。受験者が順次マ

ウスポインタを移動させて設問と選択肢を読み、正答と考えた選択肢アイコン上でマウスをクリックすると、その選択肢アイコンが受験者の解答として記録される。このツールの入力ファイルと出力ファイルの形式を付録A.2に示す。

読解効率テスト開始前に受験者に対して次の注意を与えたが、読解時間を計測していることは伝えなかった。(1)設問に解答できる程度にテキストの内容を理解する,(2)制限時間は設けないので、自分のペースでテキストを読み進める,(3)コンピュータの操作に慣れるため、事前に練習用テストを用いて練習する。

上記の方法により実施した読解効率テストにおける受験者の読解効率を求めた。まず、テキストの読解時間と単語数(カンマやピリオドを除く)とから1分間あたりの読解量を読解速度(Words Per Minute(WPM))として算出した。読解速度の算出において、設問の読解とその解答の選択に要した時間は対象外とした。次に、テキストの設問に対する解答と正解を比較し、理解度を算出した。そして、読解速度に理解度を乗じて読解効率(effective Words Per Minute(eWPM))を求めた。最終的に、受験者が読んだ7テキストからなる1つあるいは2つのテキストセット(7テキストあるいは14テキスト)における読解効率の平均値をその受験者の読解効率データとした。

#### 4. 読解効率テストの信頼性と妥当性の検証

言語テストの検証方法<sup>22)</sup>に従って、読解効率テストの信頼性と妥当性を検証する。

##### 4.1 信頼性の検証方法

テストの信頼性は、ある受験者が同一のテストを受験したとき同程度の結果になるかどうかによって検証される。信頼性を検証する方法として、再テスト法や内部一貫性法などがある。再テスト法では、同じ受験者に同一のテストを複数回行い、同一評価を与えるかどうかによって信頼性が検証される。読解効率テストの場合、再テスト法では前回のテストの記憶や学習効果<sup>23)</sup>などにより適切ではないと考えられる。言語的情報を排除したテキストの場合であっても、学習効果により読解速度が約3%程度上昇することが報告されている<sup>24)</sup>。

そこで、本研究では1回のテストを複数に分割し、分割されたテスト間で信頼性を検証する内部一貫性法を利用し、KR-21<sup>25)</sup>による信頼性係数を算出する。KR-21信頼性係数は式(1)で算出され、0から1の範囲をとる。式(1)において、 $k$ はテストの項目数、 $X$ はテスト結果の平均値、 $S$ はテスト結果の標準偏差である。

$$KR-21 = \frac{k}{k-1} \left( 1 - \frac{X(k-X)}{kS^2} \right) \quad (1)$$

本研究の読解効率テストでは読解対象テキストをテストの項目と見なす。したがって、 $k$  はテキスト数 84 である。 $X$  は、3.3 節で述べた、受験者へのテキストの割当てで読解効率テストを実施して得られる読解効率データの平均値であり、 $S$  は同読解効率データの標準偏差である。

#### 4.2 妥当性の検証方法

テストの妥当性は、あるテストが測定対象とする受験者の能力をどの程度測定できているかによって検証される。テストの妥当性は着眼点により、内容的妥当性、構成概念妥当性、基準関連妥当性に分けられる。

##### 4.2.1 内容的妥当性の検証方法

内容的妥当性は、測定対象が測定したい内容の代表的標本であるかどうかによって検証される。本研究の読解効率テストでは、テキストについての理解度と読解速度に基づいて読解能力を測定する。このため、テキストとそれに関する設問が内容的妥当性を持てば、読解効率テストも内容的妥当性を持つといえる。そこで、読解対象テキストとその設問の内容的妥当性を検証する。

##### 4.2.2 構成概念妥当性の検証方法

構成概念妥当性は、あるテストが測定対象である構成概念を示しているかどうかによって検証される。本研究の読解効率テストでは構成概念は英文読解能力であるが、上級の受験者と初級の受験者を比べると、上級の受験者ほど読解効率テストの構成概念を明らかに持っていると考えられる。そこで、読解効率テストの結果に基づいて受験者を習熟度別に弁別できるかどうかを検証する<sup>\*1</sup>。具体的には、TOEIC 読解スコアの高い受験者から順に、25%にあたる 26 人を上位、50%にあたる 50 人を中位、25%にあたる 26 人を下位に分けたとき、上位、中位、下位の間で読解効率テストの結果に統計的有意差が認められるかどうかをみる。

##### 4.2.3 基準関連妥当性の検証方法

基準関連妥当性は、構成概念妥当性の一部であり、あるテストの結果と構成概念妥当性が確立されたテストの結果とが類似しているかどうかによって検証される。本研究の読解効率テストは TOEIC のリーディングセクションと同様に読解能力を測定するテストである。したがって、TOEIC 読解スコアによって示される読解能力と類似した結果になれば、読解効率テストに基準関連妥当性があると見なせる。実際に、TOEIC の信頼性と妥当性は調査<sup>26)</sup>

により確認されている。ただし、TOEIC 読解スコアによって示される読解能力が多肢選択式により測定される読解能力である。そのため、読解効率テストによって測定される読解能力と完全には相容れないが、TOEIC のリーディングセクションの構成を制限時間とテキスト量から勘案すると、読解速度に対応する能力も含むと考えられる。そこで、読解効率テストの結果が正規分布であることを確認した後、TOEIC 読解スコアとどの程度の強さの相関関係にあるかを調べる。

## 5. 読解効率テストの検証結果

### 5.1 読解効率テストの結果

3 章の方法で読解効率テストを実施した結果、割り当てた読解対象テキストについての受験者の読解効率データが 102 件得られた。この読解効率テスト結果の平均値、中央値、最小値、最大値、標準偏差は表 1 のとおりである。

### 5.2 信頼性の検証結果

読解効率テストの KR-21 信頼性係数は 0.96 と算出された。KR-21 信頼性係数によりテスト結果の標準誤差が示される。したがって、読解効率テストの分散のうち 96%が一貫していると判断され、4%が誤差であると解釈できる。

TOEIC の信頼性を調査した結果<sup>26),27)</sup>によると、TOEIC 全体に対する KR-20 信頼性係数<sup>25)</sup>は 0.95 であり、TOEIC のリーディングセクションに対する KR-20 信頼性係数は 0.93 である<sup>\*2</sup>。KR-20 は解答が正誤のように二値の場合に用いられる指標であり、KR-21 と KR-20 との比較であるため厳密ではないが、本研究の読解効率テストの信頼性は、TOEIC 全体、あるいは TOEIC のリーディングセクションの信頼性と同程度であると見なせる。このことから、読解効率テストの信頼性が確認できたといえる。

### 5.3 内容的妥当性の検証結果

本研究では、テキストと設問を TOEIC 準拠の問題集から抜粋している。一般的に TOEIC には十分な内容的妥当性があると見なせる<sup>26)</sup>ため、本研究の読解効率テストは内容的妥当性が確保されていると考えられる。

\*2 KR-20 信頼性係数は式 (2) で算出される。式 (2) において、 $k$  はテスト項目数、 $IV$  は項目分散、 $S$  はテスト全体の標準偏差である。

$$KR-20 = \frac{k}{k-1} \left( 1 - \frac{\sum IV}{S^2} \right) \quad (2)$$

\*1 文献 8) では、読解効率と異なり理解度を考慮しない読解速度テストの結果に基づいて 22 人の受験者を習熟度別に弁別できるかどうかを検証されている。

表 1 読解効率テストの結果  
Table 1 Result of effective reading speed test.

	読解効率 (eWPM)	理解度	読解速度 (WPM)	読解時間 (秒)
平均値	55.8	0.78	69.5	117.4
中央値	54.0	0.82	64.6	105.6
最小値	16.1	0.33	29.0	39.6
最大値	135.7	1.00	153.6	375.0
標準偏差	23.3	0.14	24.2	57.3

表 2 読解効率テストの結果に対する順位和検定の結果  
Table 2 Rank-sum test result for effective reading speed test result.

習熟度 \ 読解効率	平均値	標準偏差	順位和	平均順位
上位グループ	79.7	23.0	2088.0	80.31
中位グループ	54.0	14.9	2569.0	51.38
下位グループ	35.4	12.5	596.0	22.92

#### 5.4 構成概念妥当性の検証結果

受験者の習熟度 (TOEIC 読解スコアによる上位, 中位, 下位) 間において読解効率テストの結果 (平均値) に統計的に有意差があるかどうかをクラスカル・ウォリスの順位和検定を用いて検定した。読解効率テストの結果に対する順位和検定の結果を表 2 に示す。

検定の結果, 習熟度間における読解効率テストの結果 (平均値) に統計的に有意な差が確認できた ( $H(2, N = 102) = 48.9, p < 0.01$ )。さらに多重比較 (シェッフェの方法) を行った結果, すべての習熟度の間に統計的に有意な差 ( $p < 0.01$ ) が確認できた。このことから, 読解効率テストには構成概念妥当性があることが示唆される。

#### 5.5 基準関連妥当性の検証結果

読解効率テストの結果の分布を図 3 に示す。この分布に対してコルモゴロフ・スミルノフ検定を行った。その結果, この分布は正規分布と異なるということが確認できた ( $K = 0.90, p = 0.16$ )。このことから, 読解効率テストの結果は正規分布であると見なす。

読解効率テストの結果と TOEIC 読解スコアの散点図を図 4 に示す。読解効率テストの結果と TOEIC 読解スコアの相関をみるために, スピアマンの順位相関係数の有意性検定を行った。

受験者(人)

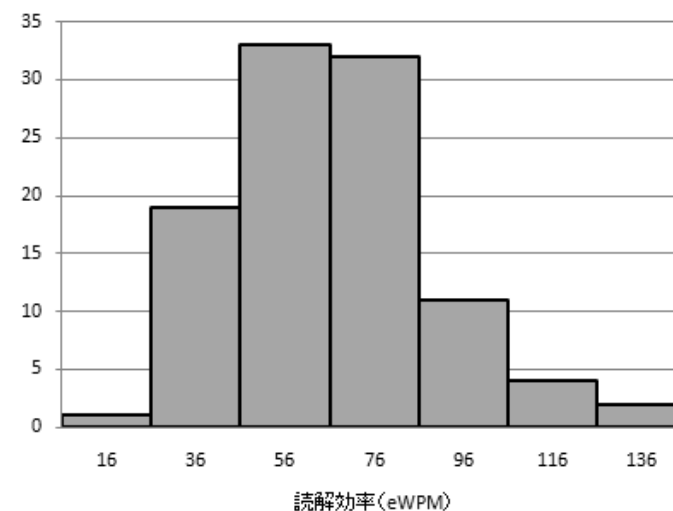


図 3 読解効率テストの結果における読解効率のヒストグラム

Fig. 3 Histogram of effective reading speed in effective reading speed test.

その結果, 相関係数が 0.73 であり, 統計的に有意な相関が確認できた ( $F(1, 100) = 103.4, p < 0.01$ )。このことから, 読解効率テストと TOEIC 読解スコアはやや強い相関関係にあるといえる。

この相関係数から読解効率テストは TOEIC のリーディングセクションに対して 53.3% の説明率を有することが分かる。したがって, 読解効率テストは TOEIC のリーディングセクションにおいて測定されている読解能力の一部を測定していると考えられる。読解効率テストの結果と TOEIC 読解スコアの間に見られる相関係数 0.73 は決して高いとはいえない。しかし, TOEIC の基準関連妥当性を検証する際に求められた, TOEIC 読解スコアと面接による英語能力評価との相関係数が 0.79 (説明率 62.4%) であったこと<sup>26)</sup> を考慮すると, 読解効率テストの相関関係は著しく低いとはいえない。このことから, 読解効率テストの基準関連妥当性が確認できたといえる。

## 6. おわりに

本稿では, 第二言語の読解能力を測定するテストとして, 読解速度と理解度に基づく読解

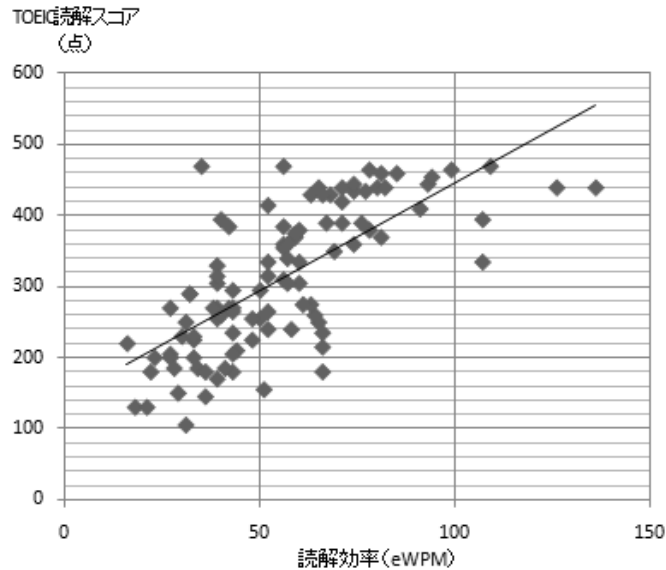


図4 TOEIC 読解スコアと読解効率テストの結果の散布図

Fig. 4 Scatter Plot of TOEIC reading score and effective reading speed test result.

効率テストを開発、実施し、テストの信頼性と妥当性を検証した。検証実験では、日本語を母語とする英語学習者 102 人を対象に TOEIC 準拠のテキストを用いて読解効率テストを実施した。その結果、テスト結果に信頼性と妥当性（内容的妥当性、構成概念妥当性、基準関連妥当性）が十分にあることが分かった。このことから、本研究の読解効率テストを用いて英語学習者の読解能力の一部を測定できると考えられる。

今回の読解効率テストの受験者を募集する際、TOEIC 受験時期の制限を設けなかった。このため、3.1 節で述べたように、読解効率テスト実施時期と TOEIC 受験時期が 1 年以上離れている受験者も含まれている。このような受験者については、読解効率テスト実施時期と TOEIC 受験時期とで英文読解能力が変化している可能性もあるため、その TOEIC 読解スコアを利用することに問題があることは否めない。今後、読解効率テストの実施と同時に受験者に TOEIC（の模擬テスト）を受けてもらい、その TOEIC 読解スコアを利用するという設定での追加実験が必要である。

本研究では、読解対象テキストとして TOEIC 準拠テキストを用いた。実際に読解効率テ

ストを英語授業などで利用する場合、テキストは TOEIC に限らず新聞記事など様々なテキストが対象となる。読解効率テストの信頼性と妥当性がテキストの種類やテキスト長などによって影響を受けるかどうかを今後調査する必要がある。

謝辞 本研究は科研費（22300299）の助成を受けたものである。

## 参考文献

- 1) Alderson, J.C.: *Assessing Reading*, Cambridge University Press, Cambridge (2000).
- 2) Bell, T.: Extensive Reading: Speed and Comprehension, *The Reading Matrix*, Vol.1, No.1 (2001).
- 3) Taguchi, E., Takayasu-Maass, M. and Gorsuch, G.J.: Developing Reading Fluency in EFL: How Assisted Repeated Reading and Extensive Reading Affect Fluency Development, *Reading in a Foreign Language*, Vol.16, No.2, pp.70-96 (2004).
- 4) Urano, K.: *Lexical Simplification and Elaboration: Sentence Comprehension and Incidental Vocabulary Acquisition*, Master's thesis, University of Hawaii (2000).
- 5) 飯野 厚：教師の音読をともなった繰り返し読みが高校生の英文読解に及ぼす効果，*STEP BULLETIN*, Vol.16, pp.27-38 (2004).
- 6) Jackson, M.D. and McClelland, J.L.: Processing Determinants of Reading Speed, *Journal of Verbal Learning and Verbal Behavior*, Vol.108, pp.151-181 (1979).
- 7) Shizuka, T.: The Effects of Stimulus Presentation Mode, Question Type and Reading Speed Incorporation on the Reliability/Validity of a Computer-based Sentence Reading Test, *大学英語教育学会紀要*, Vol.29, pp.155-172 (1998).
- 8) 吉見毅彦, 小谷克則, 九津見毅, 佐田いち子, 井佐原均：英語学習者の読解能力推定のための読解時間測定法, *教育システム情報学会誌*, Vol.22, No.1, pp.24-29 (2005).
- 9) Halliday, M.A.K. and Hasan, R.: *Cohesion in English*, Longman, London (1976).
- 10) Vapnik, V.: *Statistical Learning Theory*, Wiley-Interscience, New York (1998).
- 11) Kotani, K., Yoshimi, T. and Isahara H.: Automatic Evaluation of Foreign Language Reading Proficiency Based on Reading Time and Linguistic Features, *Proc. PACLING*, pp.35-40 (2009).
- 12) 高梨庸雄, 高橋正夫：英語リーディング指導の基礎, 研究社出版, 東京 (1997).
- 13) Alderson, J.C. and Windeatt, S.: Computers and Innovation in Language Testing, *Language Testing in the 1990s: The Communicative Legacy*, Alderson, J.C. and North, B. (Eds.), pp.226-235, Macmillan, London (1991).
- 14) Day, R.R. and Bamford, J.: *Extensive Reading in the Second Language Classroom*, Cambridge University Press, Cambridge (1998).
- 15) Tan, A. and Nicholson, T.: Flashcards Revisited: Training Poor Readers to Read Words Faster Improves Their Comprehension Text, *Journal of Educational Psychology*, Vol.89, pp.276-288 (1997).

- 16) DuBay, W.H.: *The Principles of Readability*, Impact Information, California (2004).
- 17) Norris, J. and Ortega, L.: Defining and Measuring SLA, *The Handbook of Second Language Acquisition*, Doughty, C.J. and Long, M.H. (Eds.), pp.717–761, Blackwell, Oxford (2003).
- 18) Cronbach, L.J.: Coefficient Alpha and the Internal Structure of Tests, *Psychometrika*, Vol.16, No.3, pp.297–334 (1951).
- 19) Lougheed, L.: *How to Prepare for the TOEIC Test: Test of English for International Communication*, Barron's Educational Series, Inc., New York (2003).
- 20) Arbogast, B.: *TOEIC Official Test Prep Guide 2nd edition*, Peterson's, Lawrenceville (2001).
- 21) Flesch, R.: A New Readability Yardstick, *Journal of Applied Psychology*, Vol.32, pp.221–233 (1948).
- 22) Brown, J.D.: *Testing in Language Programs*, Prentice-Hall, Upper Saddle River (1996).
- 23) Carver, R.P.: Optimal Rate of Reading Prose, *Reading Research Quarterly*, Vol.18, pp.56–88 (1982).
- 24) Wilkins, A.J., Jeanes, R.J., Pumfrey, P.D. and Laskier, M.: Rate of Reading Test: Its Reliability, and its Validity in the Assessment of the Effects of Colored Overlays, *Ophthalm Physiology*, Vol.16, pp.491–497 (1996).
- 25) Kuder, G.F. and Richardson, M.W.: The Theory of the Estimation of Test Reliability, *Psychometrika*, Vol.2, pp.151–160 (1937).
- 26) The Chauncery Group International, Ltd.: *TOEIC Technical Manual*, The Chauncery Group International, Ltd., Princeton (1998).
- 27) Woodford, P.E.: The Test of English for International Communication (TOEIC), *English for International Communication*, Moffit, C. (Ed.), pp.61–72, Pergamon Press, New York (1982).

## 付 録

### A.1 英文テキストとその内容に関する設問例

[Text]

This week Merrymaker Cruise Lines has filed a \$700 million lawsuit against Wambaugh Marine Industries, a shipyard that filed for bankruptcy last November before completing a three-ship contract for Merrymaker.

The suit contends that Merrymaker suffered \$400 million in additional construction costs and lost passenger bookings and also seeks \$300 million in punitive damages.

In 1987, Merrymaker signed the \$600 million contract with Wambaugh for three 2,600-passenger ships, but the bankruptcy halted all work at the shipyard.

The suit also contends that the yard misrepresented its financial condition in order to get the Merrymaker contract.

[Question]

1. What had Wambaugh agreed to do?
  - (A) Provide 2,600 passengers
  - (B) Build three ships
  - (C) File suit against Merrymaker
  - (D) Revise the Merrymaker contract
2. What was the value of the Wambaugh-Merrymaker agreement?
  - (A) \$300 million
  - (B) \$400 million
  - (C) \$600 million
  - (D) \$700 million
3. Why did work stop at the shipyard?
  - (A) Costs of materials were too high.
  - (B) Passenger bookings had decreased.
  - (C) The shipyard filed for bankruptcy.
  - (D) The ships were completed.

### A.2 読解時間・解答記録ツールの入出力ファイル

読解時間・解答記録ツールへの入力である読解対象テキストのファイル形式は、XML形式である。図5に例を示す。図5のようなXMLファイルツールに入力すると、ツールによってタグが解釈され、図2のように画面に表示される。<PARAG> タグに囲まれた部分が読解範囲であり、読解時間が測定される単位である。

ツールからの出力である読解時間データは各読解範囲の読解時間と空白時間（あるアイコン上からマウスポインタがはずされ、次のアイコン上にマウスポインタが置かれるまでの時間）からなる。このデータは図6のようなTEXT形式のファイルとして保存される。このため、読解時間データファイルは、テキストエディタで編集したり、たとえばPerlプログラムなどで加工したりすることが容易にできる。図6の1行目はユーザ情報である。2行目の「POS」はマウスポインタが置かれたアイコン、「TIME」は読解測定開始からそのア



```
<?xml version="1.0" encoding="SHIFT_JIS" standalone="yes"?>
<DOCUMENT>
<PARAG>
<CAPTION> 1 </CAPTION>
<TEXT> You can see an English sentence. </TEXT>
</PARAG>
<PARAG>
<CAPTION> 2 </CAPTION>
<TEXT> ... </TEXT>
</PARAG>
...
</DOCUMENT>
```

図 5 読解時間・解答記録ツールへの入力ファイルの例  
Fig. 5 Example of input file to tool for recording of reading time and answers.

001	guest1	20	男性	780
POS	TIME	DIFF	BLANK	
START	0	0:00:00	0:00:00	
1	0:00:00	0:00:07	0:00:00	
2	0:00:07	0:00:06	0:00:00	
3	0:00:13	0:00:05	0:00:00	
4	0:00:18	0:00:04	0:00:00	
...				
15	0:01:09	0:00:04	0:00:00	
16	0:01:13	0:00:03	0:00:00	
17	0:01:16	0:00:04	0:00:00	
18	0:01:20	0:00:04	0:00:00	

図 6 読解時間・解答記録ツールからの出力ファイルの例  
Fig. 6 Example of output file from tool for recording of reading time and answers.

アイコンにマウスポインタが置かれるまでの経過時間、「DIFF」はそのアイコンにマウスポインタが置かれていた時間、「BLANK」は空白時間をそれぞれ意味する。このデータから、たとえば、最後のアイコン 18 に到達するまでに 1 分 20 秒かかり、このアイコンに表示された読解範囲の読解に 4 秒かかったことが分かる。

(平成 22 年 5 月 7 日受付)  
(平成 23 年 1 月 14 日採録)



小谷 克則  
2004 年関西外国語大学大学院外国語学研究科博士課程修了。博士（英語学）。2002 年より 2008 年まで情報通信研究機構特別研究員。2003 年よりアジア太平洋機械翻訳協会課題調査委員会委員。2003 年より電子情報技術産業協会言語資源分科会委員。2004 年より関西外国語大学外国語学部勤務。



吉見 毅彦（正会員）  
1987 年電気通信大学大学院計算機科学専攻修士課程修了。1999 年神戸大学大学院自然科学研究科博士課程修了。（財）計量計画研究所（非常勤）、シャープ（株）を経て、2003 年より龍谷大学理工学部勤務。2004 年より 2008 年まで（独）情報通信研究機構専攻研究員を兼任。



井佐原 均（正会員）  
1980 年通商産業省電子技術総合研究所入所。1995 年郵政省通信総合研究所関西支所知的機能研究室室長。情報通信研究機構上席研究員を経て、2010 年より、豊橋技術科学大学情報メディア基盤センター教授。