

Web を利用した語彙習熟度テストの開発

山崎淳史*・中條清美**

Developing Vocabulary Proficiency Web-Tests

Atsushi YAMAZAKI and Kiyomi CHUJO***

In recent years, e-learning or online network-based language teaching has become more prevalent in higher education institutions in many parts of the world. The purpose of this study was to develop a cost-effective, easily administered and accurate online vocabulary proficiency test. To develop the e-test, a total of 90 test items were collected from both existing pencil and paper and floppy-disc vocabulary test programs and were screened for effectiveness using an 'item discrimination index'. Those items which did not discriminate test-takers' proficiency levels were eliminated, resulting in 63 items. Next, these 63 items were used as the basis of an online test program then given to 235 students in eight university classes in the fall of 2004, and both scores and feedback were obtained. It was noted that the test administration went smoothly and easily with no technical problems, resulting in economizing time and effort for both test-takers and teachers. Most students indicated they found the test to be easily understood and enjoyable to take, and that they were open to online testing. Test-takers' questionnaires provided useful suggestions for improvement. From the success of this trial, our next goal is to develop more multimedia programs into an online format to promote network-based language learning.

キーワード：Web テスト，語彙習熟度テスト，英語教育，e-Learning

1. はじめに

急速なインターネットの普及とともに，高等教育における e-Learning (=electronic learning) は世界的な広がりを見せている¹⁾。北米や北欧等の海外の e-Learning 先進国の大学だけでなく，日本の大学でも，近年，インターネットの活用が急速に進み²⁾，坂元(2003)によると，1999年から2002年にかけてインターネット上の情報を教材として配布した割合は，55.9%から76.9%へ，電子

メールや電子掲示板による学生からのレポート提出は52.6%から66.5%へと伸びている²⁾。

日本大学生産工学部の英語教育では，2001年度からCD-ROM等のメディアを用いた英語教育用コンテンツを開発して，いわゆるCBT(Computer-Based Training)に分類されるCALL(Computer-Assisted Language Learning)授業を実施しているが，国内外のネットワーク活用の拡大の動きを踏まえ，本学部の英語教育においても既存の学習コンテンツをスタンド・アローンからWebベースに移行する取り組みを開始した。本研

*日本大学大学院生産工学研究科博士前期課程数理工学専攻1年

**日本大学生産工学部教養・基礎科学系助教授

究では、その取り組みの第一弾として、テキスト情報のみからなる語彙習熟度テストの Web 教材への移行と、開発したテストの試用結果を報告する。

本研究の目的は、1) 従来使用してきた「英語語彙習熟度テスト」の問題をテスト理論等に照らして精選し、問題数の削減と測定精度の向上を図る、2) 精選した習熟度テスト問題を Web 教材化し、テストの実施、成績処理の効率化を図る、3) 開発した Web テストを試用し、受験者と指導者からテスト実施に関するフィードバックを得ることである。本研究の中心的な関心は、テスト項目精選の方法の有効性、一般英語教師による既存の CBT 教材の Web 化の可能性、Web 化がもたらす英語指導上の利点の確認、さらに、今後の CBT 教材の Web 化に関する示唆を得ることである。

以下では、まず 2 節で現状の語彙習熟度テストの問題点を明確にする。3 節でテスト問題の精選方法と精選結果について述べる。次に、4 節で精選したテストを Web 教材化した試作システムについて説明し、5 節で試用実践の結果を報告し、その有効性について考察する。6 節はまとめである。

2. 語彙習熟度テスト実施上の二つの問題点

現在、CALL 授業では、語彙力養成用 CD-ROM 教材「TOEIC 語彙 1」、「TOEIC 語彙 2」、「TOEIC 語彙 3」を³⁾⁴⁾、リスニング教材あるいは文法教材と併用して指導し、総合的な英語力の向上を目指している。そして毎学期、最初と最後の授業時に、語彙指導の効果を検証するため、語彙習熟度テストをペーパーテスト(図 1)、あるいは、フロッピーディスク (FD) を用いて実施している(図 2)。ペーパーテストと FD テストの内容は同一であるが、これらの語彙習熟度テストには以下に述べる二つの問題点がある。

2.1 テスト時間

語彙習熟度テスト実施に関する第一の課題は、90 分授業の約半分を占めるテスト時間の短縮である。語彙習熟度テストでは、図 1、図 2 に示すようなテスト問題が語彙教材 1 種について 40 問～50 問出題される。授業では

問題	選択肢
1. product 2. headquarters	1. 効率 2. 戦略 3. 輸出 4. 製品 5. 操作 6. 地域 7. 品質 8. 本社 9. 産業 0. 設備
3. career 4. auditor	1. 管理者 2. 職業 3. 顧問 4. 監査人 5. 設計技師 6. 会計士 7. 助手 8. 経営 9. 同僚 0. 人事
5. regulation 6. departure	1. 大会 2. 輸送 3. 会議 4. 準備 5. 期間 6. 規則 7. 資金 8. 要望 9. 到着 0. 出発

図 1 語彙習熟度ペーパーテスト



図 2 語彙習熟度 FD テスト

学習者のレベルと目標に応じて 3 種の教材から 1 種～3 種を選択して指導しているため、CALL クラスでは、2 種類以上の語彙習熟度テストを実施することになり、テストの解答に約 40 分以上を要することも多い。

2.2 採点、集計

現状の CALL では、習熟度テストにはペーパーや FD を使用しているため、その実施にテスト用紙、マークカード、FD 等のテスト資材の準備、配布、回収に、さらにテスト終了後には成績の採点や集計に膨大な時間と費用を要する。これらのテスト実施の簡易化と、採点・成績集計の省力化、効率化を図ることが解決すべき第二の課題である。

2.3 改善策

上述した語彙習熟度テスト実施上の二つの課題を解決する方法として、まず、テスト項目(問題)を精選し、問題数を削減すること、次に、テスト形態をペーパーや FD から Web ブラウザを利用したテストに移行することが考えられる。

3. テスト項目の精選

3.1 使用する得点データ

テスト項目の精選には、2004 年度前期に実施した「TOEIC 語彙 1」と「TOEIC 語彙 2」の語彙習熟度テスト計 90 問の 135 名分の得点データを用いる。

3.2 項目弁別力指数を使用する精選方法

テスト理論において「良いテスト」とは能力の高い(よく勉強した)と考えられる受験者と能力の低い(勉強していない)と考えられる受験者を正確に「弁別する」テストであり、その度合いの測定には「項目弁別力指数(item discrimination index)」が用いられる⁵⁾。項目弁別力指数は受験者の解答データに対して、設問が正答の場合を 1、誤答の場合を 0 とした場合の各設問の数値と全項目の総得点の数値との相関係数から算出される。

項目弁別力指数の高い問題とは、テストの総得点の高い受験者が正解し、総得点の低い受験者が不正解だったことを表す。逆に項目弁別力指数が低ければ、総得点とその問題の正誤に相関は無く、テスト項目として適していないことになる。本研究では表1に示した池田(1992)⁶⁾および静(2002)⁷⁾の分類基準を用いてテスト項目の精選を行なう。

項目弁別力指数に基づくテスト項目の精選には以下の方法を用いる。最初に、項目弁別力指数を計算するためのワークシートを静(2002:273)を参照して作成する⁸⁾(表2参照)。縦軸に受験者名、横軸にテスト問題項目をとり、受験者が正答した設問には1、誤答した設問には0を与えた場合の数値を作成する。「各問の数値」と網掛けで示した「正答数の数値」の相関係数を算出して最下欄に示す。下線で示した表2の問1と問5の項目弁別力指数の値は0.2以下であり、表1の基準から「不良問題」と判断されるので、問1と問5は削除する。

上述の方法で実際に2004年前期に収集した「TOEIC語彙1」と「TOEIC語彙2」のテストの得点データ135人分を用いて90問の項目弁別力指数を算出し、表1の基準に基づいて分類した結果を表3に示す。90問のうち、「良好」が64問、「準良好」が8問、「要検討」が12問、「不良」が6問と分類される。その結果、90問中、項目弁別力指数が0.3未満であった「要検討問題」と「不良問題」の計18問が削除対象となった。

表1 項目弁別力指数を用いたテスト項目の分類基準

0.4~1.0	良好と思われる問題
0.3~0.4	準良好問題
0.2~0.3	要検討問題
~0.2	不良問題

表2 弁別力計算データシートの一部

	問1	問2	問3	問4	問5	問6	問7	正答数
受験者1	1	0	1	0	1	0	0	3
受験者2	1	1	0	0	1	0	1	4
受験者3	1	0	1	1	1	1	1	6
受験者4	1	1	1	1	0	1	1	6
受験者5	0	1	1	1	1	1	1	6
受験者6	0	0	0	0	1	0	1	2
受験者7	1	0	1	1	0	0	0	3
受験者8	1	0	0	0	1	0	0	2
受験者9	1	1	1	1	1	1	0	6
受験者10	1	0	0	0	1	1	0	3
受験者11	0	0	1	1	1	1	0	4
項目弁別力指数	0.0	0.7	0.6	0.8	-0.1	0.8	0.4	

表3 項目弁別力指数による分類結果

(90問中)	
良好と思われる問題	64問
準良好問題	8問
要検討問題	12問
不良問題	6問

3.3 得点上昇量を考慮した精選方法

テスト項目の精選には、上述の項目弁別力指数を算出する方法に加えて、もうひとつ独自の基準として、「プリテスト」と「ポストテスト」の得点上昇量を考慮した。上記の項目弁別力指数は1回のテストに基づいた基準であるが、本研究では同じテストをプリテストとポストテストの2回のテストに使用しているため、各項目についての受験者の得点上昇量(プリテストとポストテストの差)についてのデータも利用することができる。

項目弁別力指数の基準に基づいて「良好と思われる問題」と「準良好問題」と分類された項目の中にも、得点上昇量を観察してみると、プリテストとポストテストであまり得点差のない項目があることが判明した。このような項目の多くは「問題が易しすぎる」ために学習前から高い得点が得られてしまう。指導効果が得点に反映されないため、これらを除去する。除去する項目の得点上昇量の目安は、90問の各項目の得点上昇量と削除後の語数を考慮して、30ポイント未満とした。その結果、14項目が削除対象となった。

3.4 最終的な精選語数

上記の3.2の項目弁別力指数0.3未満の18問と、3.3の得点上昇量30ポイント未満の14問を合わせた32問のうち、7問は両者の基準が重複して適用されているので削除対象は25問となる。それらのうち、ペアになっている2つの問題の片方が削除対象になった設問が2問ある。このような場合はペアの片方を削除すると問題の難易度が変化するため、両者とも削除対象とする。最終的に、これら2問を加えた計27問を削除した結果、残った項目は63問となった。

3.5 精選の効果

テスト項目数を90問から63問に約2/3に精選したことによる数量的な変化を表4に示す。まず、約40分を要した解答時間は約28分に短縮することが可能となった。一方、テストの平均正答率は、精選前は83.5%、精選後は81.3%と大きな変化はない。正答率が精選後に2.2ポイント減じているのは、3.3で述べた、問題が易しすぎるために正しく習熟度を評価できていない項目を削除したことによる。次に、表4の標準偏差の変化を見ると、精選前に比べて、精選後の方が正答率のばらつきが大きくなっていることがわかる。この理由を表5で説明

表4 精選前と精選後の比較

	精選前	精選後
項目数	90 問	63 問
解答時間	40 分	28 分
平均正答率	83.5%	81.3%
標準偏差	14.7	17.5

表5 精選による受験者の正答率の変化

	精選前の正答率(%)	精選後の正答率(%)
受験者1	94.4	92.1
受験者2	94.4	92.1
受験者3	94.4	93.7
受験者4	94.4	93.7
受験者5	94.4	93.7
受験者6	94.4	93.7
受験者7	94.4	93.7
受験者8	94.4	95.2
受験者9	94.4	95.2
受験者10	94.4	95.2

する。

表5には、精選前に正答率が94.4%であった受験者10人の精選後の成績（正答率）が示されている。精選後には、10人のうち、2人が92.1%、5人が93.7%と正答率が減少し、3人が95.2%に正答率が上昇していることから、精選前には同一得点であった10名が、精選後にはより正確に語彙力の差が弁別されるようになったことがわかる。同様の変化が残りの受験者にも生じたため、表4の正答率の標準偏差は精選後の方が精選前より大きくなっていると考えられる。このことから、不良項目の除去によってテストの測定精度が向上したと言えよう。

4. Web を利用した語彙習熟度テストの試作

本節では、テスト実施の簡易化と、採点・成績集計の省力化、効率化を図るために開発した、オンラインテストプログラムについて述べる。本プログラムは、前節で精選した63問のテスト問題を用いた語彙習熟度テストをオンラインで実施可能にしたものである。

4.1 試作プログラムのシステム構成

本プログラムはインターネット上に構築したサーバで動作する。各受験者はWebブラウザを使用してテストを受験する。すべての解答を終え、テストの解答データがサーバに送信されると、瞬時に採点が行われ、結果が表示される。同時に、成績データが各講義別に自動的にサーバに保存される。また、保存された成績データは、テスト問題プログラムと同じサーバ上に設置した成績集

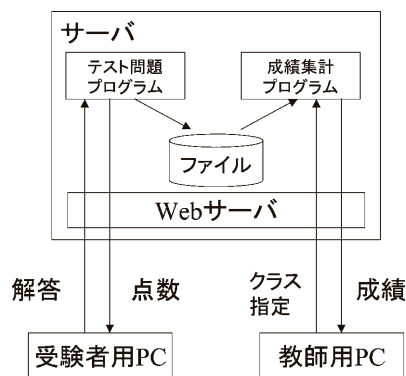


図3 システム構成

計プログラムにアクセスすることで取得できる。以下、図3に示したWebテストのシステム構成について説明する。

(1) サーバ：本プログラムはC言語を使用し、Webサーバ上でCGIとして動作する。その動作環境は以下のとおりである。

- CPU：Intel Celeron 1 GHz
- Memory：512 Mbyte
- OS：Fedora Core 2
- Webサーバ：Apache
- コンパイラ：gcc

なお、現在は教師のパソコンをサーバとして使用している。今後、他のCALL教材もWBT(Web-Based Training)に移行するならば、より処理能力の高いサーバを学内のネットワーク上に構築する必要が生じるであろうと考えられる。

(2) ネットワーク：ADSL回線(24 Mbps)を使用した。

(3) プログラム：認証プログラム、テスト問題プログラム、採点プログラムからなる。

- ① 認証プログラム：テストを受験するためのパスワード認証と受験者の情報の取得を行なう。
- ② テスト問題プログラム：HTMLで記述したテスト問題を提示する。1ページにつき2問ずつ問題が提示され、セレクトボックスから選択肢の番号を選んで解答する。「次の問題へ」ボタンを押すと、次のページの問題が読み込まれ、同様の方法で解答を選択する。
- ③ 採点プログラム：すべての解答が終了し、「採点」ボタンを押すとテスト問題ページから採点プログラムに解答データが送信される。採点プログラムは解答の採点を行ない、その結果を受験者のPCに送信する。また解答データと採点結果をサーバ内に保存する。

(4) 成績集計プログラム：パスワードと講義についての情報を入力することで、該当するデータをサーバ内から取得し、表示する。またExcel等でデータ処理が行な

えるよう CSV 形式でデータをダウンロード可能にする。

4.2 語彙習熟度テストの使用実例

以下にテストの手順を実施画面とともに説明する。

①受験科目の選択

受験者は自分が受験するテストを画面の3種類のテストから選択する(図4)。CALL 授業では、各クラスの学習者の英語力レベルに合わせた授業を行なっているので、クラスによって授業の進度が異なる。そこで、習熟度テストには、「TOEIC 語彙 1 & 2」、「TOEIC 語彙 3」、「TOEIC 語彙 1 & 2 & 3」の3パターンを用意した。このうち、「TOEIC Vocabulary 1, 2 Test」のテスト項目63問は上述の3で述べた精選問題である。「TOEIC Vocabulary 3 Test」と「TOEIC Vocabulary 1, 2, 3 Test」については今後テスト項目の精選を行なう予定である。

②テストの指示画面

図5の「テスト指示画面」には選択したテストの出題数と解答時間が提示される。受験上の注意として、i) 一度次のページへ進むと前のページには戻れないこと、ま

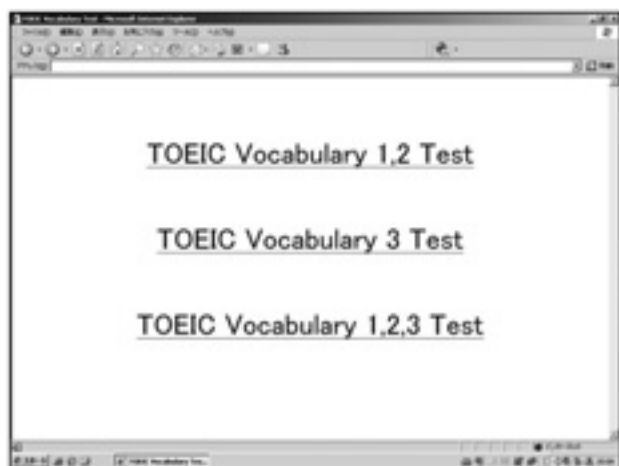


図4 テスト選択画面

た、ii) テストの時間配分が分かるように、現在何問中何問目を解答しているかが右上に表示されること、等を提示する。「テストを受験する」を選ぶと次の受験者情報の登録画面に進む。

③受験者情報の登録

パスワード、氏名、学生番号を入力し、受講している講義の曜日、時間を選択する(図6)。パスワードは教師が口頭で伝達する。受験者が「送信」をクリックすると入力情報を確認する画面が表示される。入力ミスがあった場合は再度入力を促すメッセージが表示される。

④テストの解答手順

受験者情報の登録が完了すると、テスト問題が提示される(図7)。テスト問題は1ページにつき2問ずつ出題される。受験者は提示された問題の解答を選択肢から選び、「次の問題へ」をクリックすると次ページの問題が提示される。

⑤採点

すべての解答が終了した後、最終ページにある「採点」ボタンをクリックすると正答数が表示され、結果が瞬時にフィードバックできる(図8)。

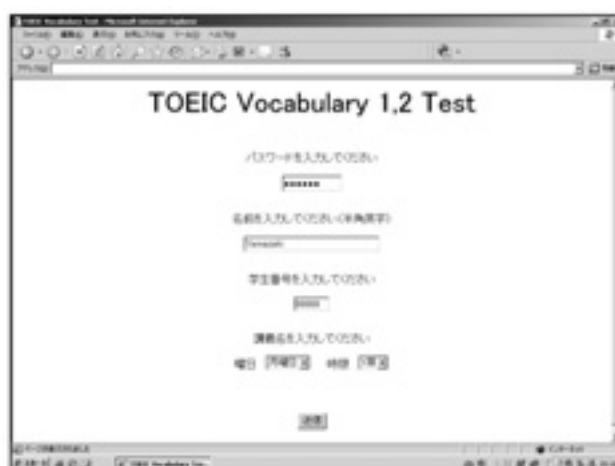


図6 受験者情報の登録



図5 テストの指示画面

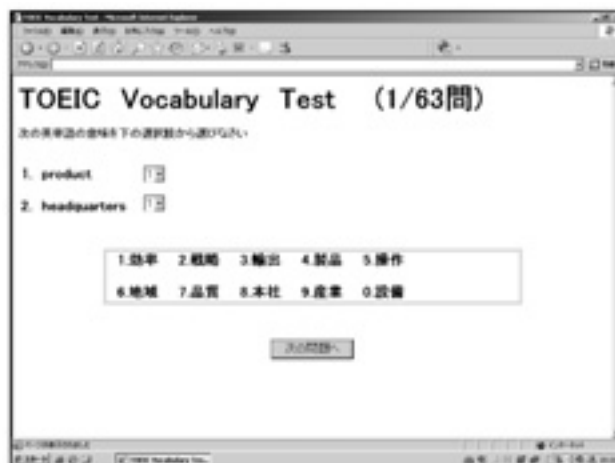


図7 テスト画面



図 8 採点画面



図 10 成績閲覧



図 9 講義名の入力画面

4.3 成績集計

サーバに保存された成績情報は以下の手順で集計，ダウンロードすることができる。

①ユーザ認証

ID とパスワードでユーザ認証を行なう。

②講義名の入力

成績を表示したい講義の曜日，時間を入力する(図 9)。

③成績の閲覧

「成績閲覧」をクリックすると，該当する成績データをサーバ内から取得し，表示する(図 10)。「ダウンロード」をクリックすると CSV 形式でデータをダウンロード可能である。

5. Web テスト実施に対する評価

5.1 試用結果

本研究で試作したオンラインテストは，平成 16 年度後期「コミュニケーション II」の授業 8 クラスの計 235 名が試用した。担当教師は 2 名であった。本システムは，受験者 20 名から 70 名の同時使用で支障なく動作するこ

とが確認された。

5.2 教師の感想

オンラインテストを実施した教師からは，「省力化と時間短縮が有難い。日常の授業で実施する小テストにも使えると非常に有益であろう。」という感想を得た。Web を利用した WBT 方式では，自動採点，集計システムにより，多人数の受験者の採点を瞬時に行ない，結果を瞬時にフィードバックすることが可能である。今回の試用で，Web テストはテスト実施にかかる教師の負担を飛躍的に軽減してくれる強力なツールであることを確認することができた。

5.3 受験者の感想

本研究で試作したオンラインテストを試用した受験者に，意見や感想を自由に書いてもらった。表 6 に，受験者の感想を肯定的，否定的，改善要望に分けて代表的意見を示す。

表 6 の「肯定的な感想」に見られるように，受験者の多くが Web テストを歓迎する意見を述べている。

表 6 の数少ない「否定的な感想」のうち，①の「前のページに戻れないこと」については後に述べた改善点についてのコメントを参照されたい。②の意見は 1 人のみから出されたが，今後 Web を利用したテストや教材に移行していく上で懸念される事項であるので，サーバにかかる負荷等を検討したいと考える。③については，4.2 ④で述べたように既に現在何問目かを表示しているが，より分かりやすいように工夫したい。④の「データが消える心配」については現在，解答の最後に一括してデータを処理している方式を，例えば 10 問ごとに保存するような方法も検討したい。

最後に改善点について述べる。改善要望の①については単語が隠れないよう画面のレイアウトを工夫したい。②については現在の入力方式は入力ミスが防げるという利点があるので現状を維持する。③については，テスト終了後，どこを間違ったかという情報を表示するように

表6 オンラインテストの感想

肯定的な感想
① 次の問題に進む際、時間がかかることなく次のページに円滑に進めてよかった。 ② マークミスはなくなるので良いと思う。 ③ インターネットでのテストで良い点はテスト時間を短縮させること。 ④ このテストはやりやすく問題数もちょうど良かった。 ⑤ 簡単にはじめられて良かった。 ⑥ 1問1問が大きく表示されるので見やすかった。 ⑦ 普段のテストもこれを使えたらいいと思います。 ⑧ Webテストは初めての体験だったが楽しみながらテストが出来たし、やりやすかった。 ⑨ パソコンでやるテストはスムーズにいくしとっても良いと思う。これからもパソコンでやるテストを増やしていったほうがいいと思う。
否定的な感想
① 1度次の問題に行ったら戻れないのがちょっと痛いです。 ② 途中で1回ページの移りが遅いところがあった。 ③ どういった問題が最後のほうにあるか分からなかったのでペース配分が分からなかった。 ④ コンピュータを使うので点数が消えたりするのが心配。
改善要望
① 解答番号を選ぶ時に単語が隠れてしまい少し見にくい場合が何度かあった。 ② 数字はキーボードで打っていくほうがやりやすいと思います。 ③ 自分がどこが間違っているか採点一覧結果を表示して欲しい。 ④ 前の画面に戻ることが出来なくなるので見直しができるようにして欲しい。

改善したい。④については多数の要望があったが、テスト時間の短縮がWebテスト採用の理由の一つであったので現状の形態を維持したいと考える。

なお、受験者の自由筆記の感想はシステムに関する評価が大部分で、テストの内容について特筆すべき感想は述べられていなかった。おそらく受験者は前期に同形式のペーパー/FDテストを受験しているため当該テスト内容に慣れているため、システムに関心が向いたと考えられる。今後、質問紙を用いてテスト内容についても意見を収集したいと考える。

6. まとめ

本研究では、①従前の語彙習熟度テスト問題を用いて、テスト理論に基づいて精選する方法の有効性を確認したこと、②既存のCBT教材のWeb化の可能性を確認できたこと、③試作したWebテストの実用性を確認できたこと、以上の3点について報告した。

本研究において、テスト問題の精選作業は自動抽出でなく手作業で行なった。自動化の前に、テスト理論だけではうまく機能しない問題を発見し、「使えるテスト」の作成に必要なステップを洗い出すことが必要と思われたからである。その結果、項目弁別力指数だけでは現実のテストへの応用には十分でなく、独自に付加した「プリ

テストとポストテストの得点上昇量」も考慮した精選が必要であることが判明した。このような知見に基づいてテスト項目の抽出を自動化するのが今後の課題である。

Webテストの開発は、今後の英語CALL教材のWeb化への示唆を得るために行なったパイロット・スタディである。オンラインテストは省力化と時間短縮の点で教師、受験者の両方に大きなメリットが認められたことから、今後も研究を継続していきたいと考える。

謝辞

本稿をまとめるにあたり、千葉大学の西垣知佳子氏にご協力いただきました。また、査読者より有益なコメントをいただきました。心より感謝いたします。

参考文献

- 1) 松本青也, 「e-Learningの可能性」, 『Step英語情報』, 7(3), 2004, 56-59.
- 2) 坂元昂, 「国際化・情報化の変革の中での高等教育の近未来像」, 『平成15年度情報処理教育研究集会講演論文集』, 文部科学省・北海道大学, 2003, 3-10.
- 3) 中條清美, 「英語初級者向け『TOEIC語彙1, 2』の選定とその効果」, 『日本大学生産工学部研究報

- 告], 36, 2003, 1-16.
- 4) 中條清美, 牛田貴啓, 山崎淳史, マイケル・ジナン
グ, 内堀朝子, 西垣知佳子, 「ビジュアルベーシック
による TOEIC 用語彙力養成ソフトウェアの試作
III」, 『日本大学生産工学部研究報告』, 37, 2004,
29-43.
- 5) 静哲人, 『英語テスト作成の達人マニュアル』, 大修
館書店, 東京, 2002.
- 6) 池田央, 『テストの科学』, 研究社, 東京, 1992.

7) 静哲人, 前掲書.

8) 静哲人, 前掲書.

注1) 本研究における e-learning とはネットワークを
利用してすべてがオンラインで提供される WBT
(Web-Based Training) を指す。CD-ROM 等を
使ってオフラインで教材が提供される CBT
(Computer-Based Training) は含まない。

(H 17. 2 .10 受理)