語彙テストの形式が語彙知識と読解能力の測定に及ぼす影響

茨城県／筑波大学大学院在籍　森本 由子

概要　本研究は文脈が手がかりにならない言い換え形式のテスト（Type A），文脈が手がかりになる言い換え形式のテスト（Type B），そして文脈が手がかりになる空所補充形式のテスト（Type C）という3タイプの多肢選択式語彙テストと，単語の意味にかかわる知識，コロケーションの知識，読解能力との相関関係が異なるかどうかを調べた。この結果，言い換え形式のテストは空所補充形式のテスト（Type C）より単語の意味にかかわる知識を測定しており，文脈が手がかりになるテスト（Type B）は手がかりにならないテスト（Type A）よりも読解能力を測定している割合が大きかった。しかしコロケーションの知識については相関係数間に有意差が現れなかった。したがって，単語の意味にかかわる知識と読解能力については同様の文脈内語彙テストでも測定している能力の割合が異なることが示されたため，目的に応じて文脈内語彙テストを使い分ける必要があることが示唆された。

1　本研究の背景

語彙テストにおける目標語は，語彙サイズの測定や付随的語彙学習での成果を調べる際以外は文脈内で提示されることが望ましい，もしくは文脈内提示が必要である（Read, 2000, p.162）と言われている。また，Hughes（2003, p.182）によると，通常学習者や言語使用者が語彙に遭遇するのは文脈内であるため，語彙を文脈内で提示してテストすることによってテストの真正性を上げることができる。しかし，Laufer, Elder, Hill and Congdon（2004, p.204）は，語彙を文脈内でテストすることが最も妥当な方法かもしれないと述べていると同時に，文脈内語彙テストは語彙知識を測定しているのか，それとも推測の技能を測定しているのかあいまいになる時がある，と主張している。

文脈で提示する語彙テストの中でも，どのようなテストを用いるべきかという意見は研究者間で異なっている。例えば Nation（2001, p.354）によると，文脈内多肢選択式語彙テストの選択肢はすべて文脈に意味的に合致するもの，つまり各錯乱肢と目標語が入れ替わっても文の意味が通るものでなければ，目標語を知らなくても正解できてしまう可能性があるため，選択肢はすべて文脈と意味的に関連しているべきである。しかしその一方で Bachman（1986）によって，このようなテストに解答する際の方略は，文脈を無視して語彙知識のみを使用することかもしれないため，正解にたどり着く際に文脈にある情報を処理することは逆効果であるかもしれない，と指摘されている。

このように，語彙テストは常に文脈内で提示されるべきである（Read, 2000, p.101）と言われながらも，どのような形式のテストを用いるべきかについては一致した見解が得られていない。実際には文脈内語彙テストの中にも複数の形式があるため，それぞれの形式によって測定している能力が異なる可能性がある。また，Read（p.161）も目標語が文脈内で提示されるかどうかよりも，むしろテスト項目に答えるために受験者が文脈情報を必要とするかどうかが重要だと述べているため，文脈内で提示される語彙テストでも文脈情報の必要性に応じて区別する必要があると考えられる。したがって，本研究では文脈内語彙テストの中でも特に多肢選択式テストを取り上げ，文脈情報が手がかりとなる程度とテスト形式を変えた3タイプの多肢選択式テストの得点と，語彙の意味にかかわる知識，コロケーションの知識，

読解能力との関係を比較する。

2 先行研究

語彙テストは (a) 語彙サイズを測定するための語彙の広さのテスト，(b) 一般的な熟達度を測定するための語彙テスト，(c) 語彙の深さのテストの 3 つに分類される（Koda, 2005, p.56）。本研究で扱う文脈内多肢選択式語彙テスト（3.3.1参照）は語彙サイズを測定するものでも語彙の深さを測定するものでもないため，(b) に該当する。しかし，語彙テストは一般的に大きく広さと深さのテストに分けられている（e.g., Schmitt, 2000, pp.167-172）ことから，ここでは目標語の意味を尋ねるテストを語彙の広さのテストとして扱い，単語のそれ以外の特徴（形態素，コロケーションなど）を尋ねるテストを語彙の深さのテストとして表現した。

2.1 語彙の広さのテスト同士の比較

相澤（2005b）は (a) Vocabulary Levels Test（以下 VLT, Schmitt, Schmitt, & Clapham, 2001），(b) 語を知っているかどうか判断させる Yes/No テスト，(c) 目標語を単独で提示して意味を書かせるテスト，(d) 文脈内で目標語を提示して意味を書かせるテスト，(e) 目標語を単一提示する多肢選択式テスト，そして (f) 目標語を文脈内で提示する多肢選択式テストという 6 種類のテストの相関関係を，すべて同じ目標語を用いて比較した。

この結果，Yes/No テスト以外は r = .80前後という高い相関関係が得られた。Yes/No テストで相関が低くなった理由としては，Yes/No テストでは目標語以外にフィラーとして擬似語が含まれていたからだと考えられる。

つまり，実在する語のみを Yes/No テストで出題すると，知らない単語であってもすべての語に「Yes」と答える可能性があるため，実在しない語を含めていたからである。したがって，同じ目標語の意味を測定しているのであれば，形式が異なっていてもほぼ同じ能力を測定していると考えられる。

一方，Henning（1991）は多肢選択式語彙テストのみを扱い，8 種類のテストを比較している。この 8 種類とは，(a) 当時の TOEFL 語彙問題，(b) 単一提示問題，(c) 最小限の文脈内で提示した言い換え問題，(d) 最小限の文脈で提示した空所補充問題，(e) 短めの文脈内で提示した言い換え問題，(f) 短めの文脈で提示した空所補充問題，(g) 目標語は単一提示で，選択肢が文脈に埋め込まれている問題，(h) より長い文脈に目標語が埋め込まれている問題，である。(c) と (d)，そして (e) と (f) はそれぞれ全く同一の文脈から構成された問題であった。

この結果によると，上記の全テスト得点の総計と各テスト得点との相関係数は r = .66-.82と，比較的高くなっており，全体としてこれらのテストは似たような能力を測定していると考えられる。しかし，全く同一の文脈から構成されている (c) と (d) の相関係数が r = .49，(e) と (f) の相関係数が r = .42であった。

この結果から，同じ文脈を使用していても言い換え形式と空所補充形式のように問題形式が異なると，それぞれのテストが全く同じ能力を測定しているとは言えないかもしれない。(b) の単一提示問題とそれぞれの両形式のテストを比較すると，言い換え形式である (c) と (e) が空所補充形式の (d)，(f) よりも相関関係が高くなっているため（(c): r = .47，(d): r = .37，(e): r = .52，(f): r = .40），言い換え形式の方が空所補充形式より語彙の意味にかかわる知識そのものを測定している可能性がある。

しかし，Henning（1991）は受験者のテスト形式への親密度やテストの妥当性，信頼性を研究の焦点に当てていたため，この点についてはこれ以上詳しく検証されていない。

2.2 語彙の広さと深さの関係

語彙の広さと深さの関係について調べた研究において，語彙知識の深さは主に (a) Word Associates Test（以下 WAT）や Depth of Vocabulary Knowledge test（DVK）などを用いて単語のパラディグマティックな知識とシンタグマティックな知識を測定している場合，(b) 各研究者が開発した形態素のテストを使用している場合，の 2 つに大別できる。ここでのパラディグマティックな知識とは，目標語と同様の意味，もしくは目標語の意味の一部を表している単語を正しく選択できる知識であり，シンタグマティックな知識とは目標語とコロケーションを成す単語を正しく選択できる知識である。本研究ではコロケーションの知識を語彙知識の深さとして扱うため，ここでは (a) の先行研究について扱う。

■表1：語彙の広さと深さの関係を調べた先行研究

	語彙の広さのテスト	語彙の深さのテスト	相関係数
Akase（2005）	Nation（1990）の VLT	Read（1996, 2000）の WAT	.55, .60, .41 （全員・上位群・下位群の順）
Noro（2002）	Nation（1990）の VLT	Read（1993, 2000）の WAT	.69, .36, .51 （全員・上位群・下位群の順）
Nurweni & Read（1999）	Translation Test	Read（1993）の WAT	.79
Qian（1999）	Nation（1983, 1990）の VLT	DVK	.82
Qian（2002）	Nation（1983）の VLT, 1995年以前の TOEFL 語彙問題	Read（1993, 1995）の WAT を32問と，Qian（1998）が作った 8 問を加えた計40問	VLT との相関：.70, TOEFL 語彙問題との相関：.68
Qian & Schedl（2004）	TOEFL 語彙問題	DVK	.75
Shimamoto（2005）	Nation（2001）の VLT	Productive Word Knowledge Test	語の意味の知識との相関：.62, パラディグマティックな知識の解答した数との相関：.34, パラディグマティックな知識の正答数との相関：.47, シンタグマティックな知識の解答した数との相関：.28, シンタグマティックな知識の正答数との相関：.32

表1は，語彙の広さと深さの関係を調べた先行研究をまとめたものである。この表から，語彙の広さのテストにおいて目標語を単独で提示した Nurweni and Read（1999）の Translation Test や，Nation（1983, 1990, 2001）の VLT を使用した研究と，目標語を文脈内提示している TOEFL 語彙問題と語彙の深さを調べた研究の間には，相関係数の差があまりないことがわかる。しかし，TOEFL 語彙問題は目標語が提示されている文脈を参照しなくとも正解にたどり着くことができるという点で，単一提示形式のテストとあまり変わらない特徴を持っていると考えることもできる。英検や TOEIC の語彙問題のように，文脈情報を利用して答えなければならない場合は，文脈内の単語と目標語や選択肢との間のパラディグマティックな知識やシンタグマティックな知識との相関関係がより高くなるかもしれない。

また，Schmitt（1999）は学習者の TOEFL 語彙問題 6 問における正答率や学習者が答えを推測して解答した割合と，それぞれの目標語の多義語としての意味・語連想・品詞・コロケーションについての知識を，学習者にインタビューすることによって比較した。その結果，TOEFL で正解を選べた多義語については，学習者が知っている意味もあれば知らない意味もあり，また TOEFL で正解を選べなかった多義語についても全く多義語としての知識がない，ということはほとんどなかった。語連想については英語母語話者の非典型的な連想と類似していたが，典型的な連想とは類似していなかった。最後に，参加者は目標語を含んだ文を産出する際に，常に正しいコロケーションを使えているわけではなかった。

したがって，Schmitt によると TOEFL の語彙問題に正解することは語彙知識のほんの一部を測っているにすぎないため，このようなテストが本当に測っているものは何かを詳しく検証する必要がある（p.209）。本研究では TOEFL の語彙問題形式だけではなく，文脈を参照することが必要な形式のテストも用いて，それぞれの語彙テストが何を測定しているのかを検証する。

2.3　語彙の広さと読解能力との関係

表2は，語彙の広さと読解能力との相関係数を表している。これによると，目標語を単独で提示した VLT（Nation, 1983, 1990, 2001）を使用した研究が多く，相関係数は .44 ～.78とさまざまである。また，

■表 2：語彙の広さと読解能力の関係を調べた先行研究

	語彙テスト	読解テスト	相関係数
Aizawa（2005a）	Schmitt et al.（2001）の VLT	TOEIC もしくは mini-TOEIC	.59（TOEIC，mini-TOEIC いずれも）
Akase（2005）	Nation（1990）の VLT	英検準 1 級の問題 2 題	.44（上位群のみでは.35，下位群のみでは.18）
Noro（2002）	Nation（1990）の VLT	TOEFL の workbook 1から 1 題，SLEP テストから 2 題	.74（上位群のみでは.46，下位群のみでは.59）
Qian（1999）	Nation（1983, 1990）の VLT	TOEFL リーディング問題（1995年以前の問題であるため，語彙問題は含まれていない）	.78
Qian（2002）	Nation（1983）の VLT，1995年以前の TOEFL 語彙問題	TOEFL リーディング問題（1995年以前の問題であるため，語彙問題は含まれていない）	.74（VLT との相関），.73（TOEFL 語彙問題との相関）
Qian & Schedl（2004）	TOEFL 語彙問題	1995年以降の TOEFL リーディング問題（語彙問題は除外）	.75
Shimamoto（2005）	Nation（2001）の VLT	TOEIC IP	.58

TOEFL 語彙問題と読解能力との相関を調べた Qian & Schedl（2004）では，相関係数は .75であり，VLT との相関係数と大きく変わるものではなかった。

Read（1997, p.307）によると，一般的には語彙テストが文脈化（contextualized）されているほど，読解能力が語彙テストのパフォーマンスにおいて役割を果たすだろうと言われているが，表 2 を見る限りではこの主張はいまだ理論的に支持されていない。

この理由としては，文脈内語彙テストを使用している研究（Qian, 2002; Qian & Schedl, 2004）は TOEFL の語彙問題を使用しているが，TOEFL の語彙問題は文脈を参照せずに解答することが可能であるため（Bachman, 1986），文脈の影響がほとんど現れなかった可能性がある。そこで，本研究では文脈が解答の手がかりになる度合いを変えて，読解能力との関係が変化するかどうかを調べる。

また，語彙の広さと読解能力だけではなく，語彙の深さと読解能力の間にも高い相関があることがわかっているが（e.g., Qian, 1999），語彙の深さと読解能力との相関については本研究の直接の焦点ではないため扱わなかった。

2.4 本研究の目的

以上の先行研究をまとめ，本研究では単語の意味にかかわる知識，コロケーションの知識，読解能力の 3 点と，形式の異なる文脈内語彙テストとの関係がどのように変化するかを明らかにする。テスト形式以外の要因を変化させることなく比較することで，精緻に各テスト形式の違いを検証し，どのような場合に特定のテスト形式を用いることが望ましいかについての示唆を得ることができる。

3 研究方法

以上の先行研究と目的から，以下 3 点の仮説を設定した。

3.1 仮説

仮説 1　目標語が正解の手がかりになる度合いが高くなるほど，多肢選択式語彙テストと短答式語彙テストとの相関は高くなる（3.3.1の Type A > B > C）

仮説2　文脈が正解の手がかりになる度合いが高くなるほど，多肢選択式語彙テストとコロケーションテストとの相関は高くなる（Type C > B > A）

仮説3　文脈が正解の手がかりになる度合いが高くなるほど，多肢選択式語彙テストと英語読解テストとの相関は高くなる（Type C > B > A）

仮説1での目標語とは，語彙テストにおける下線部の単語のことである。言い換え形式のテストでは下線部の語と同じ意味の単語を選択肢から選ぶようになっているが，空所補充形式のテストでは下線が付いている単語がなく，その部分が空所になっている。

したがって，言い換え形式のテストでは目標語の意味を知っていれば選択肢から正解を選ぶことができるが，空所補充形式のテストではその目標語がないために，文脈を読み取らなければ正しい選択肢を選ぶことが不可能である。

また，仮説2，仮説3における文脈が手がかりになる度合いについては，空所補充形式のテストでは目標語の部分が空所になっているため文脈を参照しなければならないが，言い換え形式のテストでは目標語があるため，目標語の意味を知っていれば必ずしも文脈を参照する必要がない。

そこで，言い換え形式の中でも文脈が手がかりになる場合とならない場合を変化させることによって，コロケーションの知識や読解能力との関係が異なるかどうかを調べることにする。

3.2　参加者

大学生207名が参加した。専攻は文系・理系の多岐にわたっていた。しかし，207名から母語が日本語の学習者のみを選出したため，最終的には199名のデータを分析対象とした。この199名をランダムな3群に分けた結果，Type A を解いた受験者が65名，Type B が68名，Type C が66名となった。

3.3　使用テスト

3タイプの文脈内多肢選択式語彙テスト，単一提示短答式語彙テスト，Word Associates Test（Read, 1993, 1998, 2000），TOEIC リーディングセクション Part VII（以下 TOEIC, ETS, 2002）の4種類のテストを使用した。

3.3.1　文脈内多肢選択式語彙テスト

3タイプの多肢選択式語彙テストを作成した。問題例は以下のとおりである。

Type A：言い換え形式で，目標語は正解する手がかりになるが文脈は手がかりになり得ない，もしくはなりにくい問題（資料1参照）

The man in the portrait is a direct ancestor of mine.

a. book　b. movie　c. paper　d. picture*

（*が正解）

Type B：言い換え形式で，目標語も文脈も手がかりになり得る問題（Type C と選択肢は等しい，資料2参照）

The man in the portrait is a direct ancestor of mine.

a. break　b. cost　c. energy　d. picture*

Type C：空所補充形式で，目標語はないが文脈が正解する手がかりになる問題（英検，TOEIC 形式，資料3参照）

The man in the（　　）is a direct ancestor of mine.

a. break　b. cost　c. energy　d. picture*

Type A は言い換え問題でかつ文脈が手がかりにならない問題である。これは，Type A の選択肢が目標語と入れ替わった場合でも文の意味が通るため，参加者が目標語の意味を知らなかった場合，文脈と選択肢を照らし合わせて正答を選ぶことが不可能なためである。Type B も同じく言い換え問題であるが，Type A と違って文中の目標語が正答以外の選択肢と入れ替わった場合，文の意味が通らなくなる。

したがって，解答者が目標語の意味を知らなくても，文と選択肢の意味が理解できたならば文の意味から正答を推測できる可能性がある問題となっている。Type A のすべての選択肢が目標語と入れ替わっても文の意味が通じるか，そして Type B においては正答以外の選択肢が目標語と入れ替わると意味が通じなくなるかについては，英語母語話者に確認した。

Type C は空所補充問題である。目標語がないため，文脈を使用して解答しなければならない。Type B と C は目標語の有無以外はすべて等しい形式になっている。

以上 3 タイプの多肢選択式語彙テストの特徴をまとめたものが，表 3 である。ここでの目標語は言い換え形式における下線部の語のことであり，文脈は目標語や空所とともに提示されている文のことである。Prince（1996）によると，第二言語の語彙研究では一文の文脈が最も広く使われているため，本研究での文脈は一文とする。

■ 表 3：各テストタイプの特徴

	形式	目標語	文脈
Type A	言い換え	手がかりになる	手がかりにならない
Type B	言い換え	手がかりになる	手がかりになる
Type C	空所補充	手がかりにならない	手がかりになる

■ 表 4：語彙評価の観点
（Adapted from Read, 2000, p.9）

Discrete	⇔	Embedded
語彙知識や語彙の使用を，独立した構成概念として測定している場合		他の，より大きな構成概念の中で語彙を測定している場合
Selective	⇔	Comprehensive
特定の語彙が評価の焦点になっている場合		インプットとして使われたマテリアル（リーディング・リスニング教材）やテスト受験者の反応（ライティング・スピーキングのタスク）全体を測定している場合
Context-independent	⇔	Context-dependent
文脈を参照することなしに期待された解答をすることができる場合		文脈の情報に注意して期待された解答をする，という能力を測る場合

以上 3 つのテストを，表 4 に挙げた Read（2000）による語彙テストの 3 観点から描写すると，Type A，B，C はすべて Discrete と Selective であり，Type A は Context-independent（文脈独立），そして Type B，C は Context-dependent（文脈依存）という特徴を持っている。

これらの多肢選択式語彙テスト作成には以下の手順を踏んだ。まず，JACET8000（大学英語教育学会基本語改訂委員会, 2003）から最も高頻度な1,000語を除いて，目標語を Microsoft Excel の乱数関数（望月・相澤・投野, 2003, pp.203-206）からランダムに選出した。高頻度な1,000語を除いた理由は 2 点ある。まず，大学生は1,000語レベルの単語はすべて知っている，と先行研究（八島, 2002）から予想されているため，これらの項目弁別力は低くなると考えられることである。次に1,000語レベルの単語には機能語が多く含まれているが，機能語の同義語はほとんどないために選択肢の作成が難しいことである。また，本研究の目標語は内容語のみとしたため，eh のような感嘆詞，cm のような短縮形が除かれた。同時に，内容語の中でも Guinea のような固有名詞や，nutritional のように 1 語で表せる同義語がないために正解の選択肢の作成には適さない単語も除かれた。

その後，ジーニアス英和大辞典（小西・南出, 2002）など複数の辞書の例文を参考にして文脈が作られ，JACET8000（大学英語教育学会基本語改訂委員会, 2003）から錯乱肢が選ばれた。この際，錯乱肢は正答と同じ品詞であり，さらに正答とほぼ同じ頻度であることを条件とした。ここでのほぼ同じ頻度とは，同一問題に付随する選択肢の頻度が JACET8000で1,000語の差を超えないようにすることである。例えば，ある問題の選択肢はすべて2,200語から3,200語の頻度の範囲に含まれるようにする，もしくは1,500語から2,500語の範囲に含まれるようにするなどである。これによって，少なくとも頻度の点からは学習者にとってすべての選択肢がほぼ等しくなるように統制した。最後に，選択肢をアルファベット順に並べ替えることで，選択肢の配列が意図的にならないようにした。

これらのテストは各100問ずつ作成し，計98名の大学生を対象に行った予備実験の結果によって，それぞれ項目弁別力の高い問題を選出した。項目弁別力は点双列相関係数（r_{pb}）を用い，Henning（1987）に基づき r_{pb} = .25以上を弁別力が高い項目とした。まず，Type A，B，C すべてにおいて r_{pb} = .25以上であった11問を選出し，その後，3 つのテストタイプのうち 2 つで r_{pb} = .25以上であり，残りの 1 つが r_{pb} = .00以上であった29問を選出することで，計40問を選んだ。この結果，項目弁別力の平均値は，Type A が r_{pb} = .35，Type B が r_{pb} = .48，Type C が r_{pb} = .44となった。相関係数の場合にはそのまま値を平均することができないため，Fisher の Z変換を用いて算出した（Glass & Hopkins, 1996, p.362参照）。

3.3.2　単一提示短答式語彙テスト

このテストは，目標語の意味を日本語で書くものである。この目標語もJACET8000（大学英語教育学会基本語改訂委員会, 2003）から乱数関数を用いてランダムに30問を選出した。ただし，この際には目標語が外来語であっても，日本語で語彙化されているものを選んだ。この場合の語彙化とは，学習者のL1においてL2の目標語と同義語（単語，複合語，語彙的決まり文句）があることを指す（Paribakht, 2005, p.705）。なぜなら，もし語彙化されていない単語を出題した場合，テスト受験者がその単語の意味を知っていたのか，それとも発音から推測して解答したのか不明であるためである。目標語が日本語で語彙化されているかどうかは，４名の大学院生によって決定された。ここで，radioなどは語彙化されていないとしてテスト問題からは排除されたが，theory は「理論」という日本語を当てはめることができるとして，そのまま残すこととした。問題例は以下のとおりである。答えをカタカナで書いた場合には得点にならないことを明記した。下に例題を挙げ，資料４に全項目を示した。

1. publicly　（　　）
2. tool　（　　）

3.3.3　Word Associates Test（WAT）

このテストはRead（1993, 1998, 2000）によって開発された，40項目から成るテストである。以下はその問題例である。（*が正解）

sudden

beautiful	quick*	change*	doctor
surprising*	thirsty	noise*	school

左側にある４つの単語はパラディグマティックな知識を問うための選択肢であり，右側にある４つの単語はシンタグマティックな知識を問うための選択肢である。パラディグマティックな選択肢からは目標語と同義語，または目標語の意味の一部になる単語を選び，シンタグマティックな選択肢からは目標語とコロケーションを成す単語を選ぶ。

テスト受験者は８つの選択肢の中から全部で４つの単語を選択することになっており，左側から１単語と右側から３単語，左右から２単語ずつ，そして左側から３単語と右側から１単語を選ぶという３つの解答パターンがある。この理由は，受験者に答えを推測させにくくするためである。また，WATの目標語はすべて形容詞となっており，パラディグマティックな知識を問う選択肢も同様にすべて形容詞，そしてシンタグマティックな知識を問う選択肢はすべて名詞である。したがって，同義語としては形容詞同士の単語を，そしてコロケーションとしては形容詞と名詞の結び付きを測っている。

Shimamoto（2005）によるとWATは，(a) 形式が複雑で目標語は初級者にとって難易度が高い，(b) 目標語がすべて形容詞であるため，１つの品詞の意味しか測定できない，(c) WATの形式は推測を誘発する，という３つの改善点が述べられている。しかし，語彙の深さに関する多くの研究でWATが使用されている点，そして既に妥当性が確認されている（Read, 2000）という点から，本研究はWATをコロケーションテストとして用いた。

WATの元の指示文は英語で書かれていたが，参加者の母語が日本語であることを考慮し，指示文を日本語に訳したものを使用した（資料５参照）。

3.3.4　TOEIC リーディングセクション

旧形式のTOEICのリーディングセクションにはPart Ⅴ，Ⅵ，Ⅶの３種類が存在した。そのうち，Part Ⅴは文法・語彙問題，Part Ⅵは誤文訂正問題，そしてPart Ⅶは読解問題であった（ETS, 2002）。本研究では読解テストとしてTOEICを用いるため，Part Ⅶのみを使用した。しかし，より項目弁別力が高い問題を選ぶため，同じくTOEICのPart Ⅶを行ったUshiro et al.（2006）を参照し，40問の中からより項目弁別力が高い19問を選出した。この際，より従来型の読解能力を測る問題を選んだため，弁別力が高くなっていたとしても文章が短い招待状のようなテキストから読み取る問題は除くこととした。

3.4　手順

各グループの参加者が多肢選択式語彙テストのType A，B，Cのいずれかを15分で解き，その後，参加者全員が短答式語彙テスト（８分），WAT（25分），TOEIC（19分）を解いた。時間制限は予備実験などを経て決定した。

3.5　採点方法

多肢選択式語彙テストとTOEICは（0, 1）で採点し，それぞれ40点満点，19点満点であった。短答式

語彙テストは２通りの採点方法が考えられる。１つは品詞の誤りを不正解とする厳しい採点方法であり，もう１つは品詞の誤りを正解とする緩やかな採点方法である。両方の採点方法を行った結果，信頼性係数は表５のようになった。この結果，いずれの群においても厳しい採点方法の方が信頼性係数は高かったため，本研究では厳しい採点方法を採用し，30点満点とした。同形の単語で複数の品詞を持つ単語の場合は，辞書に掲載されている品詞であればすべて正解とした。

■ 表５：２つの採点方法による信頼性係数

	人数	厳しい採点方法	緩やかな採点方法
Type A に解答した群	65	.78	.73
Type B に解答した群	68	.79	.76
Type C に解答した群	66	.72	.71

また，WAT は１つの項目につき４つの答えが存在するため，正答を選んだ数をそのまま得点とした。40項目のテストであるため，満点は160点となった。

4 結果

4.1 テストの信頼性と妥当性

4.1.1 信頼性

クロンバックαを用いて信頼性を算出した。表６がその結果である。多肢選択式語彙テスト，短答式語彙テスト，WAT ではおおむね高い信頼性が得られているが，TOEIC では十分な信頼性が得られなかった。したがって，ピアソンの相関係数を算出する際に希薄化の修正を行うこととした。

■ 表６：各テストの信頼性係数

	問題数	Type A（65名）	Type B（68名）	Type C（66名）
多肢選択式語彙テスト	40	.83	.89	.85
短答式語彙テスト	30	.78	.79	.72
WAT	40	.89	.89	.87
TOEIC	19	.58	.49	.62

4.1.2 表面妥当性

TOEIC 以外のテストについては，それぞれの受験者にテストの表面妥当性があるかどうかを５段階で判断させた。この際の指示は「このテストは，あなたの語彙力を測っていると思いますか？　数字に丸をつけてください」とし，５が「とても語彙力を測っている」であり，１が「全く語彙力を測っていない」という意味である。それぞれのテストの表面妥当性の平均は表７のようになった。すべてのテストは５点満点中平均3.5点以上であるため，高い表面妥当性があると言える。

■ 表７：各テストの平均表面妥当性

	人数	平均表面妥当性
多肢選択式語彙テスト Type A	65	3.75
多肢選択式語彙テスト Type B	68	3.67
多肢選択式語彙テスト Type C	66	3.78
短答式語彙テスト	199	4.06
WAT	199	3.93

4.2 相関分析

表８に，各テストの記述統計を示す。参加者全員に共通して行った短答式語彙テスト，WAT，TOEIC は各グループ間でほぼ等しい平均点であった。一方，各グループが異なるタイプの問題を解いた多肢選択

■ 表８：記述統計

		Mean	*SD*	*Min*	*Max*	Full
Type A に解答した群（65名）	多肢選択式（Type A）	22.94	6.79	8	36	40
	短答式	10.62	4.06	2	19	30
	WAT	102.89	15.71	43	128	160
	TOEIC	10.92	2.88	5	17	19
Type B に解答した群（68名）	多肢選択式（Type B）	29.28	7.57	11	40	40
	短答式	11.59	4.05	2	19	30
	WAT	103.94	15.38	56	128	160
	TOEIC	11.82	2.57	7	18	19
Type C に解答した群（66名）	多肢選択式（Type C）	23.58	7.08	7	36	40
	短答式	11.29	3.60	0	20	30
	WAT	105.15	14.01	56	130	160
	TOEIC	11.47	2.94	4	18	19

（注）多肢選択式＝多肢選択式語彙テスト；短答式＝短答式語彙テスト

式語彙テストにおいては，Type A と C の平均点はほぼ等しいが Type B の平均点が高くなった。これは，Type A と C は目標語，もしくは文脈のどちらかしか正解への手がかりにならないのに対し，Type B は目標語と文脈共に手がかりになり得たからだと考えられる。

4.2.1　ピアソンの相関分析

各グループにおいて，テスト間の相関係数を算出した。その結果が表 9，10，11である。この際，青木（2003）を使用して95%信頼区間も算出した。また，4.1.1より TOEIC の信頼性が十分ではなかったため，希薄化の修正を行った値も算出した。Glass and Hopkins（1996）によると希薄化の修正とは「両変数に測定誤差がないと仮定したときの，２変数間の相関（p.126）」である。希薄化の修正を行うことによって，信頼性が低いために相関係数が低くなってしまっている現象を修正することができると考えられる。ただし，この値は推定値にすぎないため，本研究では希薄化の修正値と同時に実際の相関係数の値も考察する。

Type A，B，C における多肢選択式語彙テストと短答式テストの95%信頼区間，多肢選択式語彙テストと WAT の95%信頼区間，そして多肢選択式語彙テストと TOEIC の95%信頼区間はそれぞれ重なり合っている。一方，希薄化の修正後の数値では，多肢選択式語彙テストと短答式語彙テストにおける相関係数の Type B と C のグループ間において，95%信頼区間が重なり合っていない。したがって希薄化の修正後では，両者の相関係数は少なくとも95%の確率で異なると言える。しかしその他の希薄化の修正後の相関係数間は重複しているため，相関係数の差があるとは言えない。

次に，以下の Glass and Hopkins（1996，pp.359-360）の数式に従って，各相関係数間に有意差があるかどうかを調べた。

$$z = \frac{Z_1 - Z_2}{\sigma_{z_2 - z_1}}$$

（Z_1 は r_1 の変換値，Z_2 は r_2 の変換値であり，

$$\sigma_{z_1 - z_2} = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}$$

である）

その結果，多肢選択式語彙テストと短答式語彙テストにおける相関係数の Type B と C のグループ間において有意差が見られ（$z = 2.02 > 1.96$），この結

■ 表 9：Type A グループにおけるピアソンの相関係数

	多肢選択式 Type A	短答式	WAT	TOEIC
多肢選択式 Type A	1.00	.72** (.58-83)	.60** (.43-.74)	.46** (.24-.63)
短答式	.90 (.83-.94)	1.00	.62** (.55-.80)	.55** (.36-.70)
WAT	.70 (.55-.80)	.73 (.59-.83)	1.00	.37** (.14-.56)
TOEIC	.66 (.50-.78)	.82 (.72-.89)	.52 (.31-.67)	1.00

（注）左下は希薄化の修正後の数値。カッコ内は95%信頼区間。** $p < .01$.

■ 表10：Type B グループにおけるピアソンの相関係数

	多肢選択式 Type B	短答式	WAT	TOEIC
多肢選択式 Type B	1.00	.80** (.70-.87)	.74** (.60-.83)	.67** (.52-.79)
短答式	.96 (.93-.97)	1.00	.67** (.51-.78)	.53** (.33-.68)
WAT	.83 (.73-.89)	.80 (.69-.87)	1.00	.59** (.40-.72)
TOEIC	1.00	.85 (.77-.91)	.89 (.82-.93)	1.00

（注）左下は希薄化の修正後の数値。カッコ内は95%信頼区間。** $p < .01$.

■ 表11：Type C グループにおけるピアソンの相関係数

	多肢選択式 Type C	短答式	WAT	TOEIC
多肢選択式 Type C	1.00	.63** (.46-.76)	.68** (.53-.80)	.62** (.45-.75)
短答式	.81 (.71-.88)	1.00	.65** (.49-.77)	.49** (.29-.66)
WAT	.80 (.69-.87)	.83 (.73-.89)	1.00	.44** (.23-.62)
TOEIC	.85 (.77-.91)	.74 (.60-.83)	.60 (.42-.74)	1.00

（注）左下は希薄化の修正後の数値。カッコ内は95%信頼区間。** $p < .01$.

果は95%信頼区間を使用した解釈と一致している。また，多肢選択式語彙テストと TOEIC における相関係数では，Type A と B のグループ間において有意傾向が見られた（$z = 1.81 > 1.65$）。しかし，その他のグループ間や，多肢選択式語彙テストと WAT の相関係数におけるグループ間には有意差・有意傾向共に見られなかった。

4.2.2 偏相関分析

2.3の先行研究より，読解能力と語彙知識は中程度から高い相関関係がある傾向となっている。また，表 9，10，11より短答式語彙テスト，WAT，TOEIC 間の相関係数は .37から .67であるため，多肢選択式語彙テスト以外の各テスト得点が互いに影響を及ぼしていることが見てとれるため，擬似相関が起こっている可能性がある。つまり，2 変数間の間に直接的な関係があるわけではなく，第 3 の要因が 2 つの変数に共に影響を及ぼしているために見られる，見かけ上の相関が起こっているかもしれない。そこで，短答式語彙テスト，WAT，TOEIC の純粋な影響を調べるために偏相関分析を行った。それぞれ 2 つのテスト結果を制御する 2 次の偏相関分析を行い，その結果は表12，13，14に示した。

表12より，多肢選択式語彙テストと短答式語彙テストの偏相関係数は，Type B において最も高く，次いで Type A，C であった。また表13，14より，多肢選択式テストと WAT，多肢選択式語彙テストと TOEIC の偏相関係数は Type C で最も高く，次いで Type B，A であった。4.2.1で挙げた Glass and Hopkins（1996）の数式を用い，各偏相関係数間に有意差があるかどうかを調べたところ，多肢選択式語彙テストと短答式語彙テストの間の偏相関係数において，Type B と C の間に有意差が見られ（$z = 2.47 > 1.96$），また Type A と C の間には有意傾向が見られた（$z = 1.73 > 1.65$）。さらに，多肢選択式語彙テストと TOEIC の間の偏相関係数においては，Type A と C の間に有意差が見られ（$z = 1.97 > 1.96$），Type A と B の間には有意傾向が見られた（$z = 1.73 > 1.65$）。しかし，多肢選択式語彙テストと WAT の間の偏相関係数では，いずれのグループ間にも有意差・有意傾向は見られなかった。したがって，Type A，B のような言い換え形式のテストの方が空所補充形式よりも単語の意味にかかわる知識との偏相関係数が高く，また文脈が手がかりになる Type

■ 表12：WAT と TOEIC の得点を統制した偏相関分析の結果

	Type A		Type B		Type C	
	選択	短答	選択	短答	選択	短答
選択	1.00	.50**	1.00	.59**	1.00	.23
短答	1.00		1.00		1.00	

（注）選択 = 多肢選択式語彙テスト；短答 = 短答式語彙テスト。** $p < .01$.

■ 表13：短答式テストと TOEIC の得点を統制した偏相関分析の結果

	Type A		Type B		Type C	
	選択	WAT	選択	WAT	選択	WAT
選択	.00	.29*	1.00	.33**	1.00	.43**
WAT	1.00		1.00		1.00	

（注）選択 = 多肢選択式語彙テスト。* $p < .05$, ** $p < .01$.

■ 表14：短答式テストと WAT の得点を統制した偏相関分析の結果

	Type A		Type B		Type C	
	選択	TOEIC	選択	TOEIC	選択	TOEIC
選択	1.00	.10	1.00	.39**	1.00	.43**
TOEIC	1.00		1.00		1.00	

（注）選択 = 多肢選択式語彙テスト。** $p < .01$.

B，C の方が，文脈が手がかりにならない Type A よりも読解能力との偏相関係数が高かった。しかし，コロケーションの知識との偏相関係数に関しては，各テストタイプ間で有意差が見られなかった。

5 考察

5.1 仮説 1 の検証

仮説 1「目標語が正解の手がかりになる度合いが高くなるほど，多肢選択式語彙テストと短答式語彙テストとの相関は高くなる」は一部支持された。多肢選択式語彙テストと短答式語彙テストの関係については，ピアソンの相関係数，偏相関係数共に Type B の方が Type C よりも有意に相関係数が高く，また偏相関係数では Type A が Type C よりも相関係数が高いという有意傾向が現れた。Type A と B が共に言い換え形式のテストであり，Type C が空所補充形式のテストであることを考慮すると，言い換え形式の多肢選択式語彙テストの方が空所補充形式のテストよりも単語の意味にかかわる知識をより強く測定しているテストであると言える。この理由としては，言い換え形式のテストは目標語と選択肢の意味を合致させる形式である一方，空所補充形式のテストは目標語がないために直接単語の意味にかかわる知識を測定しているのではないかもしれないからだと考えられる。

一方，言い換え形式の Type A と B を比較すると，相関係数，偏相関係数共に有意差は現れなかった。この結果によると，言い換え形式の問題であれば，文脈が正解への手がかりになる場合もならない場合も，ほぼ同程度の強さで単語の意味にかかわる知識を測定していると考えられる。

5.2 仮説 2 の検証

相関分析，偏相関分析の結果共に多肢選択式語彙テストと WAT の相関係数においてテストタイプ間に有意差が見られなかったため，仮説 2「文脈が正解の手がかりになる度合いが高くなるほど，多肢選択式語彙テストとコロケーションテストとの相関は高くなる」は支持されなかった。したがって，本研究で扱った形式の異なる多肢選択式語彙テストが，コロケーションの知識を測定している程度はほぼ等しいと考えられる。しかし，本研究でコロケーションテストとして使用した WAT は，実際にはパラディグマティックな知識とシンタグマティックな知識の両方を測定している（3.3.3参照）。そのため，パラディグマティックな知識を測定している部分とシンタグマティックな知識を測定している部分に分け，それぞれが 3 タイプの多肢選択式テストとどのような相関関係にあるのかを調べた。その結果は表15において示されている。

■ 表15：WAT をパラディグマティックとシンタグマティックに分けた時の多肢選択式語彙テストと WAT の相関係数と偏相関係数

	ピアソンの相関係数			偏相関係数		
	Type A	Type B	Type C	Type A	Type B	Type C
多肢選択式テストと WATpara との相関係数	.64**	.71**	.61**	.36**	.41**	.31*
多肢選択式テストと WATsyn との相関係数	.49**	.61**	.62**	.18	.17	.41**

（注）WATpara = パラディグマティックな知識；
WATsyn = シンタグマティックな知識。*p < .05; **p < .01

表15より，パラディグマティックな知識は相関係数，偏相関係数共にテストタイプ間で大きな差は見られなかった。また，4.2.1の数式を用いて相関係数間の差を調べたが相関係数間，偏相関係数間に有意差，有意傾向共に見られなかった。一方，シンタグマティックな知識においては偏相関係数で Type C が Type A，B を大きく上回っている。しかし，4.2.1 の数式を用いた結果，相関係数間，偏相関係数間に有意差，有意傾向共に見られなかった。

つまり，WAT をパラディグマティックな知識，シンタグマティックな知識それぞれを測定している部分に分割して相関係数，偏相関係数を算出したとしても，各テストタイプ間に有意差が見られなかった。したがって，文脈内語彙テストの形式を変化させたとしても，コロケーションの知識を測定する度合いは変わらないと言える。ただし，WAT はパラディグマティックな選択肢とシンタグマティックな選択肢の両者から合わせて 4 つの正答を選ぶ形式になっているため，表15はテストの局所独立性の侵していることを考慮にして，結果を解釈する必要がある。

5.3 仮説3の検証

相関分析の結果から，多肢選択式テストとTOEICの得点の相関係数においてはType A，B間の差に有意傾向が見られ，また偏相関分析の結果からはType AとCの間に有意差，そしてType A，B間に有意傾向が見られた。表3より，Type Aは文脈が手がかりにならない問題であり，Type B，Cは文脈が手がかりになる問題であるため，文脈が手がかりになりうる問題の方が，文脈が手がかりにならない問題よりも読解能力との相関係数が高くなることが示された。しかし，Type B，Cの間には有意差，有意傾向共に見られなかったため，仮説3「文脈が正解の手がかりになる度合いが高くなるほど，多肢選択式語彙テストと英語読解テストとの相関は高くなる」は一部支持にとどまった。この結果によると，文脈が正解の手がかりになるのであれば，言い換え形式のように目標語が提示されている場合も空所補充形式のように目標語が空欄になっている場合も学習者の読解能力をより反映したテストになる。一方，文脈が手がかりにならないテストでは，学習者の読解能力とはより独立した知識を測定できるということを示唆している。

6 結論

本研究の結果から，以下の3点が結論付けられる。まず，言い換え形式のテストであれば文脈が手がかりになるか否かにかかわらず同程度に単語の意味にかかわる知識を測っているのに対し，空所補充形式のテストは言い換え形式よりも単語の意味にかかわる知識を測っている度合いが弱い。次に，今回用いた3タイプの文脈内多肢選択式語彙テストにおいては，コロケーションを測定している度合いに違いが見られなかった。そして，空所補充形式のテストや，言い換え形式の中でも文脈が手がかりになる方が，文脈が手がかりにならないテストよりも読解能力を強く測定している。

今後の研究においては，より多くの項目を用いて結果を一般化していくことが必要であろう。また，本研究でコロケーションテストとして扱ったWATは形容詞と名詞の結び付きのみを測定するテストであった（3.3.3参照）。したがって，今後の研究ではより精緻なコロケーションテストを作成し使用することで，包括的なコロケーションの知識と語彙テストとの関係を調べることができるであろう。また，今回は大学生を対象としたため，幅広い熟達度の学習者を対象に研究を行うことが望ましい。

最後に教育的示唆であるが，本研究の結果より，同じ文脈内多肢選択式語彙テストであっても形式によって測定している技能が異なることが解明された。したがって，テスト作成者，テスト使用者共に，測定したい技能に応じて文脈内語彙テストを使い分けることが必要となろう。例えば，より単語の意味にかかわる知識を測定したい場合は言い換え形式のテストを用い，語彙知識と共にその語を有する文を理解する能力も含めた，より総合的な知識を測定したい場合には，空所補充形式のテストや，言い換え形式の中でも文脈依存形式のテストを使用することが望ましいと思われる。このように使い分けることによって，何を測定しているのかが不明確なことがある（Laufer et al., 2004）と言われている文脈内語彙テストの測定対象を明確にすることができるであろう。これによって，単一提示形式よりも望ましいとされている文脈内提示形式の語彙テストの欠点を改善していくことが可能になると考えられる。

謝　辞

本研究の実施の機会を与えてくださった（財）日本英語検定協会の皆様，特に池田央先生に心より感謝申し上げます。また，本研究に関して数々のご助言をくださった筑波大学の卯城祐司先生，そして草稿に目を通してくださった同大学院博士課程の土方裕子さんにも厚く御礼申し上げます。

参考文献（＊は引用文献）

＊Aizawa, K.（2005a）. Assessing the depth of receptive vocabulary knowledge. *JACET Summer Seminar Proceedings*, 4, 18-28.

＊相澤一美.（2005b）.『英語語彙習得研究の疑問点と今後の課題』. 大学英語教育学会第44回全国大会口頭発表資料.

＊Akase, M.（2005）. The roles of breadth and depth of vocabulary knowledge in EFL reading comprehension: with a focus on English major students. *Annual Review of English Language Education in Japan*, 16, 141-150.

＊青木繁信.（2003）.『検定および推定』. Retrieved April 9,

2006, from http://aoki2.si.gunma-u.ac.jp/lecture/stats-by-excel/work-sheets/tests.html

＊Bachman, L.(1986). The test of English as a foreign language as a measure of communicative competence. In W. Stansfield(Ed.), *Toward communicative competence testing: Proceedings of the second TOEFL invitational conference* (pp.69-88). Princeton, NJ: Educational Testing Service.

＊大学英語教育学会基本語改訂委員会.(2003).『大学英語教育学会基本語リストJACET8000』. 東京：大学英語教育学会.

＊Educational Testing Service [ETS].(2002).『TOEIC公式ガイド＆問題集』. 東京：国際ビジネスコミュニケーション協会.

＊Glass, G.V., & Hopkins, K.D.(1996). *Statistical methods in education and psychology.* MA: Allyn & Macon.

＊Henning, G.(1987). *A guide to language testing*. New York: Newbury.

＊Henning, G.(1991). *A study of the effects of contextualization and familiarization on responses to the TOEFL vocabulary test items*. TOEFL Research Report, Educational Testing Services: Princeton.

＊Hughes, A.(2003). *Testing for language teachers. Second Edition.* Cambridge University Press.

＊Koda, K.(2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

＊小西友七・南出康世.(2002).『ジーニアス英和大辞典』. 東京：大修館書店.

＊Laufer, B., Elder, C., Hill, K., & Congdon, P.(2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing,* 21, 202-226.

＊望月正道・相澤一美・投野由紀夫.(2003).『英語語彙の指導マニュアル』. 東京：大修館書店.

＊Nation, I.S.P.(1983). Testing and teaching vocabulary. *Guidelines,* 5, 12-25.

＊Nation, I.S.P.(1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle.

＊Nation, I.S.P.(2001). *Learning vocabulary in another language*. Cambridge University Press.

＊Noro, T.(2002). The roles of depth and breadth of vocabulary knowledge in reading comprehension in EFL. *Annual Review of English Language Education in Japan,* 13, 71-80.

＊Nurweni, A., & Read, J.(1999). The English vocabulary knowledge of Indonesian university students. *English for Specific Purposes,* 18, 161-175.

＊Paribakht, T.S.,(2005). The influence of first language lexicalization on second language lexical inferencing: A study of Farsi-speaking learners of English as a foreign language. *Language Learning, 55*, 701-748.

＊Prince, P.(1996). Second language vocabulary learning: The role of context versus translations as a function of proficiency. *The Modern Language Journal,* 80, 478-493.

＊Qian, D.D.(1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review,* 56, 282-307.

＊Qian, D.D.(2002). Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning,* 52, 513-536.

＊Qian, D.D., & Schedl. M.(2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing,* 21, 28-52.

＊Read, J.(1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing,* 10, 355-371.

＊Read, J.(1997). Vocabulary and testing. In Schmitt, N. & M. McCarthy(Eds.), *Vocabulary, Description, Acquisition and Pedagogy* (pp.303-320). Cambridge University Press.

＊Read, J.(1998). Validating a test to measure depth of vocabulary knowledge. In A. Kunnan(Ed.), *Validation in language assessment*(pp.41-60). Mahwah, NJ: Lawrence Erlbaum.

＊Read, J.(2000). *Assessing vocabulary*. Cambridge University Press.

＊Schmitt, N.(1999). The relationship between TOEFL vocabulary items and meaning, association, collocation and word-class knowledge. *Language Testing,* 16, 189-216.

＊Schmitt, N.(2000). *Vocabulary in language teaching.* Cambridge University Press.

＊Schmitt, N., Schmitt, O., & Clapham, C.(2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18, 55-88.

＊Shimamoto, T.(2005). Exploring lexical network systems of Japanese EFL learners through depth and breadth of word knowledge. *Annual Review of English Language Education in Japan,* 16, 121-130.

＊Ushiro, Y., Shimizu, M., Hijikata, Y., Kasahara, K., Nakagawa, C., Morimoto, Y. et al.(2006). The effects of visual materials on reading comprehension tests. *JACET Bulletin,* 42, 97-112.

＊八島等.(2002).「日本人高校生の語彙サイズ」.『関東甲信越英語教育学会紀要』, 16, 29-41.

資　料

資料 1：多肢選択式語彙テスト（Type A）の例

Tom needed solitude for devoting himself to studying to pass the high school's entrance examination.
a. assistance　b. competitors　c. discretion　d. isolation

The victim exclaimed in delight at the sight of the rescue party.
a. danced　b. jumped　c. shouted　d. smiled

The security guard cleared the passage for the famous singer through a crowd of admirers.
a. area　b. room　c. stage　d. way

資料 2：多肢選択式語彙テスト（Type B）の例

Tom needed solitude for devoting himself to studying to pass the entrance examination of the high school.
a. chapel　b. entity　c. isolation　d. refusal

The victim exclaimed in delight at the sight of the rescue party.
a. affected　b. depended　c. shouted　d. tended

The security guard cleared the passage for the famous singer through a crowd of admirers.
a. hope　b. tree　c. way　d. world

資料 3：多肢選択式語彙テスト（Type C）の例

Tom needed (　　) for devoting himself to studying to pass the entrance examination of the high school.
a. chapel　b. entity　c. isolation　d. refusal

The victim (　　) in delight at the sight of the rescue party.
a. affected　b. depended　c. shouted　d. tended

The security guard cleared the (　　) for the famous singer through a crowd of admirers.
a. hope　b. tree　c. way　d. world

資料 4：単一提示短答式語彙テスト

以下の単語の意味を書いてください。
外来語があっても，日本語（カタカナではなく平仮名・漢字）で書いてください。
カタカナで書いた場合，得点にはならないので注意してください（例：dance をダンスと書く，など）。
制限時間は 8 分です。

1. publicly　(　　)
2. tool　(　　)
3. shabby　(　　)
4. planet　(　　)
5. negative　(　　)
6. compartment　(　　)
7. rhyme　(　　)
8. registration　(　　)
9. pray　(　　)
10. false　(　　)
11. grove　(　　)
12. equilibrium　(　　)
13. referral　(　　)
14. reaction　(　　)
15. root　(　　)
16. dig　(　　)
17. rescue　(　　)
18. persecution　(　　)
19. fascinate　(　　)
20. theory　(　　)
21. prose　(　　)
22. imaginative　(　　)
23. sergeant　(　　)
24. importantly　(　　)
25. offensive　(　　)
26. auxiliary　(　　)
27. presence　(　　)
28. insurance　(　　)
29. respect　(　　)
30. variability　(　　)

資料 5：WAT の指示文

これは，英語でよく使用されている形容詞の意味について，あなたがどの程度知っているかを調べるテストです。問題形式は以下のようになっています。

A　**sudden**

beautiful　quick surprising　thirsty	change　doctor noise　school

左側にある単語は，“sudden” の意味を説明するものです。	右側にある単語は名詞で，“sudden” の後に続いてフレーズや文を作るものです。

全部で 8 つの選択肢がありますが，その中で正解は 4 つだけです。4 つの正しい語を選ばなければなりません。

“sudden” とは，「素早く，そして予測していないときに起こること」であるため，左側の正解は “quick” と “surprising” になります。	通常，“a sudden doctor” や “a sudden school” という言い回しはしませんが，“a sudden change” や “a sudden noise” という言葉は使います。よって，“change” と “noise” が右側の正解になります。

以下のように正解を書きます。

A　**sudden**

beautiful　(quick) (surprising)　thirsty	(change)　doctor (noise)　school

注意：この例では左側と右側に 2 つずつ答えがあります。しかし，実際の問題では左側に 1 つと右側に 3 つ正解があったり，左側に 3 つと右側に 1 つ正解がある問題があります。