

英文法診断テスト作成の試み

—— 設問形式と採点・診断方法の検討 ——

池 上 真 人

松 山 大 学
言語文化研究 第28巻第2号 (抜刷)
2009年3月

Matsuyama University
Studies in Language and Literature
Vol. 28 No. 2 March 2009

英文法診断テスト作成の試み

—— 設問形式と採点・診断方法の検討 ——

池 上 真 人

1. は じ め に

近年、大学生の中に英語の文法等の知識が曖昧な大学生が増加している印象を受ける。その理由の1つとして、文部科学省の公示している学習指導要領などで、聞くこと、話すことを中心とした「コミュニケーション能力」の育成が重視され、その結果、中学校、高等学校での英語教育において「コミュニケーション能力」育成を意識するがゆえに、英文の大まかな内容を読み取ることに重点をおいた教育が行われているからではないかと考えられる。そのため、一部の学生は、英文内の内容語を読み取る、あるいは聞き取ることで、なんとなく英文を理解し、英語を理解したつもりになっているのではないかと考えられるのである。また、別の理由として、入試等で多肢選択形式の英語問題が多用されていることも考えられる。近年はセンター試験や多くの私立大学の入学試験で多肢選択形式による出題が増えたため、学生も必然的に多肢選択形式の問題によって、英語を学んでいることが予想できる。その結果として、曖昧な理解のままでも何分の1の確率で正答にたどり着くことができるために、自分はこの項目の理解が不十分なのかを明確にしないまま、学習歴だけが進んでしまっているのではないかと考えられるのである。正しく英語を理解し使えるようになるためには、正しい文法知識は不可欠である。そして、このような知識が曖昧な学生たちに再度正しい知識を身につけさせるためには、彼らが何を正しく理解していて、何を理解していないのかを明らかにする必要がある。

学習者の知識を測る最も一般的な方法はテストを実施することである。テストの種類には、大別すると「集団基準準拠テスト」と「目的基準準拠テスト」がある。「集団基準準拠テスト」とは、主に総合的な言語能力を測ることを目的に、ある集団全体の中での受験者の位置づけを測るテストであり、相対評価を基本とする。例えば、順位や偏差値などによって、ある受験者が全体のどのあたりに位置するのかが示す実力テストや受験者の順位によって選抜を行う入学試験などは、代表的な集団基準準拠テストである。一方で「目標基準準拠テスト」とは、受験者がどの程度学習目標に到達できたかを測るテストであり、絶対評価によって測られる。代表的な目標基準準拠テストは、中学校や高等学校で実施されている中間・期末テスト、すなわち達成度判定テストである。達成度判定テストは、ある学習プログラムの学習内容をどれだけ理解しているかを判定するためのテストであり、一定のレベルを超えているかどうかで評価される。ある一定の学習事項において、受験者がどの段階まで理解ができているかを測るための診断テストも目標基準準拠テストの一つである。診断テストは、Brown (1996)によると、学習者がすでに習得している内容と、まだ習得には至っていない内容とを明らかにするためのテストであり、学習者が習得する必要がある学習事項を効果的に指導することを目的として実施される。冒頭に述べたような、知識が曖昧な状態にいる学生の現時点での弱点を明らかにし、どのように学習、指導すべきかを明らかにするためには、この診断テストを活用することが最も適していると考えられる。しかしながら、診断テストは学習者に対して非常に有用なテストであるにも拘らず、その作成方法に関する研究は多くない。それは、作成するにあたって様々な困難があるからである。青木他(2003:91)は、診断テスト作成の難しさについて、「診断テストは、他の要因が影響することなく、ある一つの能力を測定することが求められるという点、および、実際にどのように診断するのかという点とにおいて、その作成は困難である」と指摘している。また、Hughes (2003)は、文法項目の診断を純粋に行おうとすると膨大な数の問題が必要になり、それらを実施す

ることは日常的には非現実的であるため²⁾、診断テストはほとんど作られていないと述べている。

確かに学習者の理解度を診断することは簡単ではないが、彼らがどこで躓いているのかを明確にすることは、彼らに正確な知識を身につけさせるために不可欠である。そこで本研究では、文法知識の理解度を診断するための診断テストを作成することとした。特に本論文では、診断テスト作成の第一歩として、どのような設問形式と採点・診断方法が診断テストに適しているかを検討することを目的としている。

2. 先行研究の概観

理論的な研究の多い診断テストに関する研究の中で、具体的に診断テストを開発している数少ない研究に、青木他（2003）や金谷編（2006）がある。青木他（2003）では、仮定法を題材に診断テストを開発し、続く池上他（2004）において、実施結果を詳細に報告している。青木らの診断テストの設問形式は、仮定法を用いた文の条件節と帰結節それぞれに空欄を設け、4つの選択肢からそれぞれ正答を選ぶ空欄補充型多肢選択形式である。ただし、選択肢の中には複数の正答がある場合があり、どちらを選んでも正答と判断されている。青木らは、「仮定法過去」「仮定法過去完了」「仮定法未来」「時制のねじれ³⁾」「条件文」「名詞句・副詞句」の6つの用法について、それぞれ複数の問題を用意し各用法の正答率を算出するほか、条件節と帰結節の組み合わせに注目し誤答分析することで、誤答の傾向を診断結果として受験者に提供している。

一方、金谷編（2006）は、文法（名詞句）、語彙、音声に関する診断テストを開発している。金谷らの作成した診断テストも多肢選択形式であるが、特に本研究に深い関連があると考えられる文法に関するテストを見ると、テストには「内部構造の把握を問う問題」「句の境界の把握を問う問題」「内部構造の把握と句の境界の把握の両方を問う問題」の3つのタイプの問題があることが示

されている。「内部構造の把握を問う問題」とは、日本語訳の一部に対応する英語の名詞句（あるいは、前置詞句+名詞句）に関して、語順を変えた4つの選択肢から適切な修飾関係にある1つの正答を選ぶものである。また「句の境界の把握を問う問題」とは、be動詞を抜いた文を4～6つに区切って、どの位置に挿入するのが適切かを問う問題、また1文に4～6つの区切りを入れて、意味のまとまりで切れている箇所を選ぶ問題である。「内部構造の把握と句の境界の把握の両方を問う問題」とは、被修飾語である名詞をその単語を抜いた英文のどの箇所に入れるのが適切かを問う問題や、ある肯定文に対応した疑問文を4つの語順を入れ替えた英文から1つ選ぶ問題である。金谷らは、名詞句を構造的に「限定詞（冠詞、所有格、指示語、数量詞、数詞）+名詞」「（限定詞）+前置修飾語（形容詞、名詞、名詞+'s、疑問詞）+名詞」「名詞+前置詞句」「(a)名詞+to不定詞 (b)名詞+分詞（現在分詞・過去分詞）」「名詞+関係詞」の5つのグループに分け、上記のような設問形式を用いてそれぞれの用法のグループごとに複数の問題を設定し、その総合正答率を算出することでその用法の理解ができていないか否かを診断している。また誤答分析として、間違っただけで選ばれた選択肢の傾向を探ることで、教師と受験者にさらなる情報を与えるとしている。

診断テストを具体的に作成している以上の研究を設問形式、採点・診断方法（誤答分析）の2つの観点からみていきたい。

まず設問形式で共通しているのは、空欄補充型か訳対応型か区切り選択型などの違いはあるが、どちらも多肢選択形式を採用している点である。多肢選択形式は、採点のしやすさや客観性の面で利点が大きく、また単位時間あたりに実施できる問題量が多いため、大量の問題の実施が必要となる診断テストに適していると言える。しかしながら、一方で多肢選択形式には問題点も多く存在している。例えば、Hughes (2003) は、多肢選択形式の問題点として、当て推量による解答が容易なこと、効果的な錯乱肢の作成が困難であることなどを指摘している。また、Cheng (2004) のように、リスニング問題において、

多肢選択形式を使った問題の得点結果が記述式のそれよりも有意に高かったとする研究結果もある。実際、多肢選択形式の場合は、選択肢そのものがヒントになっていたり、消去法などのテスト・テクニックを使って正答を選べたりする可能性があり、「正答が分からなくても正答を選べる」という大きな問題点があることには留意しなければならない。テストは、しばしば「妥当性」「信頼性」「実用性」の3つ側面から検討しなければならないと言われているが⁴⁾、多肢選択形式はきちんと学習者の能力が測れているかという妥当性の面での問題点を抱えていると言える。

それぞれの研究が多肢選択形式をどのような工夫して用いているかを見てみると、まず青木らは、多肢選択形式の設問として一般的な空欄補充を用いているが、一文に空欄を2箇所空けることにより、英文の空欄以外の部分がヒントとなることを防ぐと同時に、より細かな診断を可能としている。また、正答を複数にすることで、一見して明らかに間違いだと分かるような選択肢を防ぐ試みも行っている。また、金谷らは英文の理解力に焦点を絞り、正しく理解できていないにも拘らず語句等の意味から正答を推測できてしまうことを回避するために、英文の区切りを選ばせる問題や、特定の語が英文のどこに入るかを問う問題など、一つの診断項目に対して複数の問題形式を用いることで、より確実に診断ができるようにしている。

2つの研究の採点・診断方法についてみていくと、共通している点は診断しようとしている文法の用法ごとに複数の問題を用意し、それらの問題の正答率を算出することでその項目の診断をしていること、さらに受験生が選んだ選択肢を記録することで誤答を分析し、それによって診断を行っている点である。採点・診断方法については、青木らが詳細に解説している。青木らの診断テストでは、単に正答率の算出のみではなく、解答パターンを分析することで誤答に段階を設定し、採点によっても受験者の理解の段階を診断している。具体的には、まず条件節の空欄と帰結節の空欄の双方に正しい選択肢を選べた場合を「正答」、仮定法過去の問題を仮定法過去完了の正しい組み合わせで答えた場合

のように、時制は間違っているが、条件節と帰結節の組み合わせは正しい場合を「時制判断の不完全」、条件節と帰結節のどちらかが正しい場合を「組み合わせの不完全」、どちらも間違っている場合を「誤答」と分析している。それによって、正誤という2つの段階ではなく、複数の段階で誤答を分析でき、採点の段階から診断を可能としている。

3. 設問形式と採点・診断方法の検討

前節では、先行研究として2つの研究を取り上げ、その設問形式と採点・診断方式を見てきた。設問形式については、両研究ともに多肢選択形式を用いていた。妥当性の観点から考えると、多肢選択形式の代わりに記述式や口頭式でのテストを用いれば良いと考えることもできるが、記述式、口頭式は採点時間や一度に実施可能な問題数などの実用性、また採点に採点者の主観が入りやすいという信頼性の面で問題がある。本研究では、教師に過度の負担をかけることなく学生たちの診断をするために、可能な限り採点・診断を自動化することを考えている。そのため、実用性の観点から、記述式や口頭式の問題ではなく、両研究同様に多肢選択形式の問題を検討することにした。また、本研究では、多肢選択形式に加えて、並び替え形式も検討する。並び替え形式は、英文の中で並び替えを行う部分を制限し、用いることができる語句を精選すれば、かなり解答数を制限することができる。そのため、多肢選択形式に代わる診断テストに有用な形式になる可能性があるからである。

次節からは、2つの形式をどのように用いるならば、学生の英文法力を効果的に診断できるかを検討する。

3.1. 多肢選択穴埋め形式の検討

「はじめに」でも述べたように、近年、多くの大学生が文法知識を曖昧にしたまま大学に入学してくる。また、授業を行っていても、マークシート形式の

問題、つまり多肢選択形式の問題であれば正答を選べるのだが、記述式だと正解にたどり着けない学生も多く見られる。これらの学生は、多肢選択形式の問題を解く際に、自分がなぜ特定の選択肢を選んだのか、その理由を説明できない。これらの学生に理由を問うてみると、多くの学生が言う言葉が「なんとなく、これだと思った」である。もちろん彼らは英語母語話者ではないので、「なんとなく選んだ」ということは、正答を選べる場合もあれば、誤答を選ぶ場合もあることを意味している。また選択式であれば正答を選べるが、その選択肢が正答である理由はわからないということは、実際は理解できていないということの意味しており、その知識は不十分であると言える。このような学生には、単なる多肢選択形式のテストを実施しても、本当の習熟度を測ることは難しく、学生自身も自分が何を理解できていないのかを知ることは難しい。それでは、このような学生に対して、どのような多肢選択形式の問題を用いれば、より彼らの理解を正確に診断することができるだろうか。

多肢選択穴埋め形式には、いくつかのタイプがある。一般的に用いられているのは、選択肢が3～4つあり、正答が1つのタイプである。しかしながら、このタイプの問題点はすでに指摘してきたように、正答以外の選択肢がヒントになってしまったり、消去法で問題が解けるほか、正答が1つであるため明らかに間違いであるような錯乱肢になりやすく、錯乱肢の役割を果たさなくなる場合があることである。

〔設問 A〕

The next meeting _____ on Saturday.

(次の会議は土曜日です。)

(1) was (2) be (3) will be (4) would be

〔設問 B〕

The exam _____ at 8:30 on Monday.

(そのテストは月曜日の8時半開始である。)

- (1) begins (2) began (3) begin (4) begun

※下線が正答

例えば、上記の2つの設問は、それぞれ未来時制の表し方を問うている設問である。

〔設問A〕〔設問B〕のような単一の正解を持つ設問の場合、〔設問A〕であれば、正解以外の選択肢は過去形か原形であり、正誤の判断が容易な錯乱肢になってしまっている。この場合、現在形や進行形は正答となるため、錯乱肢に用いることができない。そのため、弁別力をあまり持たない選択肢だと分かっているでも入れざるを得ないのである。また、〔設問B〕の場合は、現在形で未来を表すことを知らなくても、(2)が過去形、(4)が過去分詞形ということが分かれば、この2つの選択肢は明らかに不自然であるため、選択肢が4つあっても実際には(1) begins と(3) begin の2択となる。そのため、この設問に正解しているからといって、現在形で未来を表す用法が分かっているとは判断できない。

正答は1つであるが、「正答なし」という選択肢があるタイプの設問も時々見られる。このタイプの設問は、必ず正しい選択肢があるというヒントをなくすることができるため、正答が必ず1つあるタイプよりも、受験者の理解度を正確に測ることができると考えられるが、やはり「正答なし」を選んだ学生が何を知っていたのかが不明という問題点が残る。例えば、下記の〔設問C〕はこのタイプの設問で「正答がある」場合の例、〔設問D〕は「正答がない」場合の例である。〔設問C〕の場合、受験者は「正答なし」を選んでいなければ、正答が4つの中にあると考えていると推測できる。その上で、(1)を選んでいれば、現在形で未来を表すことが分かっていると言えるだろう。しかしながら、〔設問D〕のように「正答なし」が正解の場合は、この問題から診断できるこ

とは、選択肢(1)~(4)が未来のことを表す際には用いることができないということを受験者が判断した、ということであり、未来の表現を知っているかどうかは不明である。そのため、「診断」ということを考えた場合、「正答なし」を選択肢に入れることは避けたほうが良いだろう。

〔設問 C〕

The library _____ at 9:00 a.m. tomorrow.

(明日は、図書館は9時開館です。)

(1) opens (2) open (3) opened (4) has opened

(5) 正答なし

〔設問 D〕

My father _____ next week.

(私の父は来週退職する。)

(1) retired (2) retire (3) would retire (4) has retired

(5) 正答なし

※下線が正答

上記の問題をある程度解決するための方法としては、正答を複数にすることが考えられる。具体的には、受験者に正答数を教えることなく、設問に解答させるのである。そうすることによって、受験者はすべての選択肢について、正誤の判断をしなければならない。また正答数を複数とすることで、他の選択肢が別の選択肢のヒントとなることも、ある程度は避けることができると考えられる。例えば、下記の〔設問 E〕の場合は、確定した予定などの場合は、未来の事を現在形で表すことができるという文法規則を知らなければ、(1)(3)の複数正答を選ぶことはできない。まず、このように文法知識の有無がより分かると

いう点において、他のタイプの多肢選択形式の設問よりも、診断テストに適していると考えられる。

〔設問 E〕

The next train _____ at 12:00.

(次の電車は 12 時に出発します。)

- (1) leaves (2) leave (3) will leave (4) has left

〔設問 F〕

They _____ that we could win the match.

(彼らはその試合に勝つことを疑いもしませんでした。)

- (1) seldom doubt (2) had no doubt (3) scarcely doubted
(4) never doubted

※下線が正答

複数正答の問題では、複数の正答を設定することで、より広い文法知識を一つの設問で問うことが可能となる。例えば、〔設問 F〕は正答が(2)(4)の2つ存在している複数正答の設問であるが、そのため正解が一つの多肢選択式問題とは異なり、消去法は使うことはできず、設問を解く際にはそれぞれの選択肢が正答かどうかを1つ1つ吟味する必要がある。これによって、教師は学習者の“seldom” “no” “scarcely” “never” それぞれの理解を問うことができるのである。また同時に、設問の数を減らすこともできる。このように多肢選択式設問に複数正答を設定することによって、このような選択形式の設問をより診断テストに有益な形で用いることができると考えられるのである。

表1は、4年制私立大学1年生26名2年生25名を対象に、〔設問 A〕～〔設問 E〕を解答させた調査結果である。背景色が濃くなっているのは正答を意味

表 1：多肢選択形式のそれぞれのタイプの設問の解答

通し番号	設問 A	設問 B	設問 C	設問 D	設問 E
1	2	4	2	4	3
2	3	1	1	3	1
3	2	2	5	3	1
4	3	1	2	3	2, 3
5	3	3	5	2	2, 3
6	3	1	2	5	1
7	3	4	2	5	3
8	3	3	1	4	3
9	2	3	5	4	1, 3
10	3	3	4	3	1, 4
11	3	2	3	3	3
12	3	1	1	4	1, 3
13	3	3	5	2	1, 4
14	3	3	5	3	3
15	3	2	2	1	2
16	3	3	1	4	1
17	3	1	1	4	2
18	3	3	5	5	4
19	3	3	5	3	1, 4
20	3	3	1	3	3
21	3	1	1	5	1, 3
22	3	1	3	3	4
23	3	1	1	5	3
24	3	3	5	5	3
25	2	3	4	3	1
26	3	3	2	4	3
27	3	1	1	5	3
28	3	1	1	5	1, 4
29	3	1	2	3	1, 3
30	3	1	1	5	1, 3
31	3	1	1	5	2, 3
32	3	1	1	5	2, 3
33	3	1	1	5	1
34	3	1	5	3	1, 3
35	3	1	1	5	1, 3
36	3	1	1	5	1
37	3	1	2	5	2
38	3	1	1	5	1, 3
39	3	1	1	5	1
40	3	1	1	2	1, 3
41	3	1	1	5	3
42	3	1	1	3	1
43	3	1	1	5	1, 3
44	3	1	1	3	1
45	3	1	1		1, 3
46	2	3	5	2	1
47	3	1	1	5	1, 3
48	3	1	2	3	2
49	3	1	1	5	1, 3
50	3	1	1	5	1, 3
51	3	1	1	5	1
正答率	90.2%	62.7%	54.9%	45.1%	27.5%

する。調査では、〔設問 A〕～〔設問 E〕は、時制に関する他の問題と合わせて提示され、複数正答の問題に関しては、問題文の後に「ただし、正答が複数ある場合もある」との指示がなされている。また調査参加者は単数正答の問題に解答したのち、複数正答問題を解答している。

分析の結果、各設問の正答率はそれぞれ、〔設問 A〕が90.2%、〔設問 B〕が62.7%、〔設問 C〕が54.9%、〔設問 D〕が45.1%、〔設問 E〕が27.5%であった。ただし、〔設問 E〕に関しては、複数正答の全てに正答を選べた参加者のみの割合である。

それぞれの問題形式について検討していきたい。まず基本的な未来時制が分かっている学習者として、〔設問 A〕に正答している参加者だけを抽出した。その結果、51名中46名が正答である選択肢を選んでいた。次に、その中で、未来時制を現在形で表すことを問うている〔設問 B〕の正答者を抽出した。その結果、正答しているのは32名となった。そして、この現在形を使って未来時制を表していることを別の問題形式で問うている〔設問 C〕〔設問 D〕〔設問 E〕それぞれの正答率を算出した（表2）。

表2を見ると、正答が1つのタイプの空欄補充型多肢選択問題で正答を選んだ参加者のうち、「正答なし」を選択肢に加えた問題で正答を選んでいるのは78.1%である。また、同様の形式で「正答なし」を正解にした場合は、約6割の参加者のみが正解している。複数正答の問題〔設問 E〕については、複数ある正答を正しく選べたのは約4割であり、少なくとも選択肢の中で現在形を選べた参加者は7割であった。

この結果から、すぐに全ての形式のうち、複数正答型の多肢選択問題が最も学習者の知識を正しく測っているとの結論を得ることはできないが、すくなく

表2：〔設問 A〕〔設問 B〕正答者の〔設問 C～E〕の正答率

設問 C	設問 D	設問 E (完全解答)	設問 E (1のみ正解)
78.1%	62.5%	40.6%	68.8%

とも〔設問 B〕のように唯一の正答がある場合よりも正しく学習者の知識を測っているのではないかと推測できる。

前述の通り、「正答なし」という選択肢は、それが正答の場合は受験者の知識を診断することが難しいため、本論文で提示した3つのタイプの多肢選択形式の問題の中では、複数正答型が最も妥当性があると考えられるだろう。

それでは、次に採点・診断方法についての検討に移りたい。

採点方法に関しては、複数正答の空欄補充型多肢選択問題も従来の方法と大きくは異ならない。むしろ、選択肢一つ一つに対しての正誤を判定し、診断をするため、正誤問題に近い診断方法を用いることができる。例えば、先ほどの〔設問 E〕を例にすると、表3のような診断が可能である。

〔設問 E〕の診断では、それぞれの選択肢ごとに正答誤答を判断し、誤答の場合はそれぞれにコメントが示される。また、すべての選択肢を正解した場合のみ、「未来時制の問題に正解した」と判断され、「正解です」というコメントが表示される。さらに、それぞれの選択肢にタグを付けておくことで——例えば、選択肢(1)であれば「現在形」、(2)は「原形」、(3)は「will」、(4)は「完了形」——未来を表す場合に受験者がどれを選んだかが誤答情報として蓄積することができ、教師や受験者にフィードバックすることも可能である。

〔設問 F〕の場合は、表4のような診断となる。〔設問 F〕は、否定表現に関

表3：多肢選択形式〔設問 E〕の診断例

選択肢	解 説
(1)を正答とした場合	コメントなし
(1)を誤答とした場合	確定した予定などは未来のことでも現在形で表すことができます。
(2)を正答とした場合	原形で未来のことを表すことはできません。
(2)を誤答とした場合	コメントなし
(3)を正答とした場合	コメントなし
(3)を誤答とした場合	未来のことを表現するときには will を用いることができます。
(4)を正答とした場合	完了形で未来のことを表すことはできません。
(4)を誤答とした場合	コメントなし

表4：多肢選択形式〔設問F〕の診断例

選択肢	解説
(1)を正答とした場合	“seldom”は「めったに～ない」という意味で、頻度が少ないことを意味します。
(1)を誤答とした場合	コメントなし
(2)を正答とした場合	コメントなし
(2)を誤答とした場合	まず、選択肢(2)の“doubt”は他の選択肢と異なり、名詞です。そして、“no”は名詞と共に用いて「全く～ない」という否定の意味を持ちますので、“have no doubt”は、「全くない疑いを持っている」つまり「疑いもしない」という意味になります。
(3)を正答とした場合	“scarcely”は「ほとんど～ない」という意味で、程度が少ないことを意味します。
(3)を誤答とした場合	コメントなし
(4)を正答とした場合	コメントなし
(4)を誤答とした場合	“never”動詞の前に用いると、「決して（絶対）～ない」という強い否定を表します。そのため、“never doubted”は「決して疑わない」→「疑いもしない」という意味になります。なお、“never”の後ろを原形にしまわないように気をつけましょう。

する設問であり、それぞれの選択肢について正誤の判断がされることで、4つの否定表現の理解が診断できる。ただし、この1問だけでは、たとえこの設問に正答していても、間違いと判断された選択肢——この場合は(1)と(3)——の正確な使い方が分かっていたのかどうかの判断ができないため、別の問題で正しいと判断される選択肢として再度問うことが必要となるだろう。

以上のように、多肢選択形式の設問に関しては、複数正答を導入することである程度妥当性に配慮ができるのではないかということが明らかとなった。次節では、並び替え形式について、検討していく。

3.2. 並び替え形式の検討

並び替え形式の最大の利点は、正答が明示されていないことである。また受験者がある程度自分自身で語句を組み立てなければならぬことも利点である。しかしながら、それゆえに生じる並び替え形式の設問における課題は、可能な解答パターンが非常に多いことだといえる。そのため、診断するために必

要な採点を自動化することが非常に難しい。解答される可能性のあるパターンが多すぎるために、並び替え問題で受験者の理解度を診断しようとする、記述式同様に教師が一つ一つを判定しなければならない。しかし、すでに多肢選択形式の設問を検討した際にも述べたが、実用性を考えると教師に過度の負担がかかる採点方法は現実的ではない。そこで、本研究では解答パターンを減らすために、まず文の中で並べ替える部分を小さくすることで解答パターンの数を減少させることにした。また誤答分析を行って受験者の理解度を診断するために、余分な語句を並び替え要素に加えた。これによって、並び替え形式の問題を診断テストに用いることができると考えられる。具体的にみていきたい。

〔設問 G〕

訳：ジョンがいつ帰ってくるか教えてください。

Please let me know _____ back.

{will / when / comes / came / come / has / John}

〔設問 H〕

訳：もし明日雨が降ったら、映画に行くことにします。

If _____ ① tomorrow, I _____ ② to the movie.

{ ① _____

{will / it / rain / rains / rained / have / has}

{ ② _____

{has / go / will / went / gone / have}

〔設問 G〕の正答は“when John will come”である。この問題で問うているのは「時制」であるため、現在形、過去形、未来表現、完了形、さまざまな時制の表現が作れるように語句が用意されている。並び替えの語句は全部で7つあるが、“come”の活用形が2つ含まれているので、実質的には5つである。

〔設問 H〕は一文の中に2つの並び替え問題を挿入している。並び替え問題はそれぞれ7つの語句と6つの語句の並び替えであるが、〔設問 G〕同様に活用形を含むため、解答パターンはそれほど膨大にはならない。〔設問 H〕の場合、下線①と下線②はそれぞれの解答がヒントになってしまう可能性を持っているので、両方を問題として問う必要がある。しかしながら、両方を一度に問おうとすると解答パターンが膨大になってしまうので、2つに分割することで解答を制限している。このように英文の長さや問いたい項目によって、並び替えを問う箇所を短くし、解答を制限することで採点や診断を自動的に行うことを可能とする。

次に、並び替え形式の採点に関する先行研究を概観した後、診断テストに用いることのできる採点方法について検討したい。

並び替え形式の採点には、通常、正解・不正解の判定が用いられることが多い。しかし、不正解の中には限りなく正解に近い不正解や全くの不正解が混在している。そのため、並び替え形式の問題に関する研究の多くは、部分点をどのようにつけるかという点に焦点を当てている。採点方法に関しては、Alderson et al. (2000)、静他 (2002)、安間 (2006) などの研究がある。Alderson et al. (2000) は、文の並び替えについて Exact 法、Previous 法、Next 法、Edges 法の4つの採点法を比較している。Exact 法とは、正答の絶対位置にある要素ごとに1点を与える方法である。例えば、正解の並びが「ABCDE」の場合、「ABECD」であれば、「AB」が正答と同じ場所にあるので2点が与えられ、また「EABCD」であれば、正答と同じ場所にある要素がないので得点は与えられない。Previous 法と Next 法は、絶対位置ではなく、前後のペアの組み合わせに対して得点を与える方法である。例えば、「ABCDE」が正答の場合に「ABECD」という解答であれば、Previous 法で採点すると、(1)「B」の前に「A」、(2)「D」の前に「C」、この2箇所が正答と同じなので2点が与えられ、Next 法で採点すると、(1)「A」の次に「B」、(2)「C」の次に「D」、この2箇所が正答と合致するので2点が与えられる。また Edges 法では、最初と最後の絶対位置だけを見るの

で、解答が「ABCDE」の場合、「A」だけが得点に値し、1点が与えられる。静他（2002）は Alderson et al.（2000）の採点法を元に、Binary 採点法、Exact 採点法、Pair 採点法、Combined 採点法の4つの採点方法を提案し比較している。Binary 採点法とは絶対位置も語順もすべてが正しい場合に1点を与える方法であり、Exact 採点法は絶対位置に対してそれぞれ1点、Pair 採点法は2つの隣接するペアが正しい場合に1点ずつ、また加えて最初と最後の絶対位置が正しければそれぞれ1点ずつ与える。Combined 採点法はExact 採点法とPair 採点法を混ぜたものである（表5参照）。

安間（2006）は、語句の並び替え問題の採点方法として、MRS 採点法を提案している。この採点法は並び替えに使った各語句を何回動かせば完全に正しい文、つまり正答になるかによって判定する採点法である。例えば、正答が「ABCDE」だった場合、「ABCED」や「EABCD」であれば、「E」を1度動かせば正答になるので、-1点、「ACBED」や「EACBD」であれば、「B」と「E」を1度ずつ動かす必要があるため、-2点である。換言すれば、この採点方法は、並び替え形式の問題において、不正解と判定された解答がどの程度正解から離れているのかを測る方法であると言える。

本研究が目的としているのは、診断的採点であるため、先行研究のような採点方法をそのまま取り入れることは難しいが、これらの先行研究の採点方法を応用することで、診断に用いることができると考えられる。例えば、静他（2002）のPair 採点法を応用して、ペアではなくグループとすれば、並び替えられた

表5：4つの採点法による4種類の得点の例 [正答が“ABCDE”の場合]

応答パターン	Binary 得点	Exact 得点	Pair 得点	Combined 得点
ABCDE	1	5	6	11
ABCED	0	3	3	6
EABCD	0	0	3	3
DBAEC	0	1	0	1
ABDEC	0	2	3	5

(静他, 2002 : 61)

語句の中の主語と動詞の順序や形容詞と名詞の順序、そのほかの語順に関する診断が可能となるだろう。並べ替え部分を短くすることで解答パターンを制限するため、ある程度の誤答の想定ができるので、このグループ化によって誤答に応じて解説を用意することも可能である。また絶対位置と組み合わせることによって、より細かい診断ができると考えられる。例えば、“ABCDEF”が正答の場合に、“EFD”や“BAC”という誤答をグループ化しておく、解答の中にそれが登場した際に、自動的に解説を与えることができる。さらに“A”や“F”の絶対位置にそれ以外の語句がきた場合に解説を与えることも考えられる。

では、〔設問 G〕を用いて、具体的に説明していく。表6は〔設問 G〕の解説例である。グループ化した語句については、パイロット的に実施した調査結果の誤答を基に作成している。例えば、“when John comes”という解答をある受験者がしたとすると、解答欄の中にグループ化された“John comes”が含まれているので、「**下線部分は、名詞節になります。そのため、comesという現在形で表すことができません。時や条件を表す副詞節と混同しないようにしましょう。**」という解説を自動的に与えることができる。特に、この“when John comes”という誤答は、学習者にしばしば見られる誤答であるため、他よりも詳しい解説がついている。また、“when John will comes”というように、助動詞“will”の後ろに原形以外の形の動詞を持ってきてしまった場合は、時制とは関係のない誤りではあるが、「**助動詞の後ろの動詞は原形です。**」という解説を与えることもできる。絶対位置を用いた診断としては、解答の冒頭に“**when John**”以外が入れた場合は「**語順に誤りがあります。**」という解説が与えられる。これは例えば“when will John come”といった誤答の場合に与えられる。誤答のパターンとしては、そのほかにも動詞を二つ並べたり、語順がばらばらであったりする解答が出てくる可能性もあるが、その場合は、受験者はこの設問に関して全く理解ができていないと考え、解説の付けられた誤答グループがない場合の誤答に関しては「その他」という処理によって、一律同じ解説

を与える。このように「その他」を設定することで、すべての誤答に解説をいれる必要性がなくなる。「その他」を設定している理由であるが、誤答グループ化している語句は、教師がある程度予測できる誤答である。これ以外の誤答については、かなり不完全な理解ということが言えるため、解説を付けたとしても教師の直接的な介入なしには解説の理解が難しいと考えられるからである。

これらの情報は受験者にフィードバックされると同時に、教師にも集計結果としてフィードバックされるため、教師の指導にも役立つと考えられる。

本節で検討している並び替え問題は、並び替え問題としては語句をかなり絞っているため、通常の並び替え問題に多肢選択を加えたような形式であるとも言える。簡単に言うと、選択肢が多くなった多肢選択形式のような設問になっている。しかしながら、それぞれの語句は最初から組み合わせられて提示されていないため、多肢選択形式の設問よりもヒントとなる箇所が少ない。ま

表 6：並び替え問題〔設問 G〕の解説例

グループ化した語句	解 説 例
John comes	時制を誤っています。下線部分は、名詞節になります。そのため、comes という現在形で表すことができません。時や条件を表す副詞節と混同しないようにしましょう。
John came	時制を誤っています。ここで過去形を用いることはできません。
John come	時制を誤っています。ここで原形を用いることはできません。
John has come	時制を誤っています。ここで完了形を用いることはできません。
John will comes, John will came, John will has	助動詞の後ろの動詞は原形です。
絶対位置	解説
冒頭に when John 以外が入れられた場合	語順に誤りがあります。
その他	ここでは、「ジョンがいつ帰ってくるのか」という未来のことについて述べているので、正解は“when John will come”となります。

た、並べ替える語句の中に余分な語句を加えることで、語順と用法の両方の知識を問うことができるため、より学習者の英語を産出する力を測ることができると同時に、総設問数を減らすことができるのではないかと考えている。

以上のように本節では多肢選択形式と並べ替え形式の設問について、それぞれの形式で診断テストに用いることのできる方法を提案した。2つの形式には、それぞれに利点、欠点がある。多肢選択形式の利点は、採点のしやすさと大量の問題を処理できることであり、欠点は、常に正解が明示されていることによって受験者に大きなヒントを与えてしまう点やテスト・テクニックが使いやすいという点、当て推量を排除できない点である。一方、並び替え問題の利点は、正答を選択肢中に明示しなくてよいため、受験者にヒントになる情報が少なく、より正確に理解度を測ることが出来る点であるが、同時に並べ替え形式では、採点や誤答分析のための準備に相当の労力がかかるという点や誤答分析をどこまで詳細にするかという課題が残っている。そのため、どのような文法項目、用法にどちらの形式が適切なのかを十分に吟味する必要があると考えられる。

4. お わ り に

本論文の目的は、英文法に関する診断テストを作成するために、実用的な設問形式と採点・診断方法を検討することにあった。そして、空欄補充型多肢選択形式の設問においては、従来の1つのみの正答を持つ形式を改めて複数正答を用意することで、なるべくテクニックによる正解をなくし、その結果として、より正確な診断をすることを提案した。また並び替え形式に関しては、解答パターン数を制限するために、多肢選択形式に近い並び替え問題を提案し、余分な誤答となる語句を加えることで、誤答パターンの分類を可能とし、その結果として診断機能を持たせるという提案を行った。

今後の課題としては、まずパイロット調査の実施結果をもとに本調査を行

い、本研究で作成している診断テストの妥当性や信頼性を検討することが挙げられる。また、多肢選択形式と並び替え形式の比較を行い、それぞれの形式が様々な文法項目のどんな用法に対してより効果的なのかを検討することも課題である。

(本稿は、2006年に交付を受けた松山大学特別研究助成による研究成果の一部である。)

注

- 1) もちろん中学校や高等学校においては、成績評価の上では、他の学習者と得点を比較することもあるだろうが、テストの基本的な目的は相対評価をすることではない。
- 2) 例えば、Hughes (2003) は、冠詞を例に挙げ、Collins Cobuild English Usage (1992) に取り上げられている用法は20以上もあり、それぞれに3～4問ずつの問題を作ると、冠詞だけで膨大な数の問題を実施しなければならなくなる、と指摘している。
- 3) 「ねじれ」とは、条件節は仮定法過去完了、帰結節は仮定法過去というように、条件節と帰結節の時制が異なっているタイプのものを指している。
- 4) この3要素に「波及効果」を加えて4要素という場合もある。また、3要素と言った場合に「実用性」と「波及効果」が入れ替わる場合もある。各要素について説明すると、「妥当性」とは、そのテストが目的としている能力を正しく測っているかどうかを示すものであり、例えば、スピーキングの能力を測ることを目的に、リスニング問題で構成されたテストを実施するような場合は、正しく目的とするスピーキング能力が測れているのかどうかの「妥当性」を検討する必要がある。「信頼性」とは、そのテストが安定的に目標としている能力を測っているかどうかを示すものである。例えば、能力が変わらない限り、何度受験しても同じ得点を出すテストを「信頼性」が高いテストとすることができる。逆に、同じ能力を持った学生に対して、測るたびに違う結果が出るようなテストは「信頼性」が低いと言える。この妥当性と信頼性は、しばしば綱引きの状態になり、片方だけを重視してしまうと、もう片方がないがしろになりかねない。例えば、ライティング能力を測るテストを実施しようとする場合、最も妥当なのは実際に自由に文章を書かせ、評価することである。しかしながら、そうすると採点の際には、採点者の主観が得点を左右し兼ねず、信頼性に不安が生じる。一方で、採点者の主観を排除し、客観的な採点をするために、語句並び換え問題を採用した場合は、信頼性は高まるかもしれないが、今度は語句並び換え問題で本当にライティング能力を測れているのかどうかを吟味する必要がある出てき、妥当性に疑問が残ることとなる。このように、妥当性と信頼性とはうまくバランスを取る

ことが必要である。「実用性」とは、実施のしやすさを示すものである。実際にテストを実施しようと考へ、妥当性も信頼性も高いテストを考へ出しても、実施に莫大な費用がかかったり、特殊な人材が必要であったり、あるいは採点が非常に難しかったりすると、「実用性」に問題があることになる。つまり、容易に実行可能で、できる限り妥当性と信頼性の高いテスト方法を探る必要があるのである。「波及効果」とは、そのテストを実施することが学習に及ぼす影響を指す言葉である。例えば、入学試験の英語の問題に音声を使ったリスニング問題が出されることが分かっているならば、受験生は積極的にリスニングに関する勉強をし、音声を聞き取る訓練をするようになるだろう。しかし、もし試験で音声が用いられず、発音記号を選んだり、アクセントの場所を選ぶ問題のみが出題されることが分かっているならば、受験生は音声を聴く訓練をしなくなるかもしれない。このように「波及効果」には、プラスに働く面とマイナスに働く面の両方がある。「はじめに」で述べたように、入学試験に多肢選択形式のみしか出されないことが分かっているならば、受験生は選択形式の問題を解くことを英語の勉強の中心にしてしまうかもしれない。これはマイナスの波及効果と言えるだろう。

参 考 文 献

- Alderson, J. C., and Percsich, R., and Szabo, G. (2000). Sequencing as an item type. *Language Testing*, 17(4), 422-447.
- Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River, N. J.: Prentice Hall Regents. (訳書：ブラウン, J. D. (1999). 『言語テストの基礎知識』大修館書店.)
- Cheng, H. (2004). A comparison of multiple-choice and open-ended response formats for the assessment of listening proficiency in English. *Foreign Language Annals*, 37(4), 544-555.
- Herman, J., and Winter, L. (1985). *Test design manual: guidelines for developing diagnostic tests. Diagnostic Testing Project*. California University, LA, Center for the Study of Evaluation.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge, UK: Cambridge University Press.
- 青木昭六, 田中正道, 山岡俊比古, 萬谷隆一 (1988). 『英語のテストインナー実践的アプローチ』開隆堂出版.
- 青木信之, 加納亜弥, 田辺千尋, 池上真人, 永堀瞳, 樋口慎一 (2003). 「診断テスト作成の試みⅠ - 仮定法を題材に -」『中国地区英語教育学会研究紀要』, 33, 91-100.
- 安間一雄 (2006). 「並べ替え問題の部分採点法」第10回言語テスト学会全国研究大会発表資料.
- 池上真人, 加納亜弥, 青木信之 (2004). 「診断テスト作成の試み - 仮定法診断テストの実施結果とフィードバック -」『教育学研究ジャーナル』創刊号, 11-20.
- SLA 研究会 (編) (1994). 『第二言語習得研究に基づく最新の英語教育』大修館書店.
- 金谷憲 (編) (1994). 『定着重視の英語テスト法 - 長期的視野に立った中学校英語評価』河

源社.

- 金谷憲 (編), 英語診断テスト開発グループ (2006). 『英語診断テスト開発への道』 英語運用能力評価協会.
- 静哲人, 竹内理, 矢島智子, 吉澤清美 (2002). 「並べ替え問題における異なる採点法の信頼性と妥当性への影響」『第 28 回全国英語教育学会神戸研究大会発表要綱』, 61-64.
- 静哲人, 竹内理, 吉澤清美 (2002). 『外国語教育リサーチとテストングの基礎概念』 関西大学出版部.
- JACET 教育問題研究会 (2005). 『新英語科教育の基礎と実践 — 授業力のさらなる向上を目指して—』 三修社.
- 田中正道 (編著) (1999). 『伝達意欲を高めるテストと評価: 実践的コミュニケーション能力を育てる授業』 教育出版.
- 中村洋一 (大友賢二監修) (2002). 『テストで言語能力は測れるか — 言語テストデータ分析入門—』 桐原書店.
- 西村則久, 明関賢太郎, 安村通晃 (1999). 「英作文における自動添削システムの構築と評価」『情報処理学会論文誌』 40 (12), 4388-4395.