
クローズテストと予測文法能力

中 川 武

0. はじめに

本研究では、4年制大学に通う日本人学生を被験者としたクローズテストの実験を行い、クローズテストが大学生英語学習者の予測文法能力を反映するテストかを検証する。また、今後の学生の英語能力向上のためにどのような情報をフィードバックすることが可能かを考察する。

Oller (1979) は「予測理論 (expectancy theory)」を規定した際、この理論を適切に反映するテストとしてクローズテスト (cloze test) を挙げている。Oller は「言語使用は予測理論に基づいており、優れた言語使用者ほど高次の「予測文法 (expectancy grammar)」能力を保持し、運用できる」と主張した。一般的に高い「妥当性 (validity)^①」を保証するテストは英語学習者の予測文法を活性化させ、またそれがどの程度身についたのかを測定できるものとされるが、クローズテストは作成や実施が容易であることや、その成立背景として総合的言語能力の測定を視野に入れていることなどから多くのテスト研究者の関心を集め、これまでに様々な被験者 (native/non-native) を対象とした研究成果が報告されている。

クローズテストは Taylor (1953) によって発表され、元来は英語を母国語とする学生のための読書教材の難易度 (readability) を測定する手法に過ぎなかったが、諸研究の結果、英語を第二言語とする学習者の能力テストとしての実用的側面に注目が集まった。具体的な作成方法は後述するが、定率法の場合、基礎になる英文テキストから「何語おき」と機械的に語を削除して作成される。被験者はテキスト内の情報、文脈の前後関係、また被験者自身の持つ文法知識、語彙知識などを手がかりとして空欄に語を復元するが、この際に最も活性化されるのが「予測文法能力」である。

1. 調査目的

「クローズテストは予測文法能力を問うテストである」という説を検証することが本研究の目的である。クローズテストに加えてリーディング、リスニング各テストを外部基準として設定し、実施する。テスト間の相関関係に留意し、多角的視野からクローズテストの可能性について考察を試みる。クローズテストでは何が測定できて、その一方で何が測定できないのかについて検証する。

2. 調査方法

(1) 調査に用いられた教材

①クローズテスト

a. テキスト選定

クローズテストを作成・実施する際に配慮すべきこととして、テキスト（テスト作成の基礎になる英文）をどのように選定するかという問題がある。宗教、政治など被験者によって賛否が分かれるようなテキストは避け、また英文そのものの難易度が結果に大きく影響することをふまえる必要がある。高度の専門的知識を必要としたり、英文自体が余りに難しい場合、被験者は削除された語を推測することのみならず、文章全体のコンテキストを把握することができないからである。松村（1984: 63）は、被験者のテキストに対する親密さ、文化的背景、既知か未知かの違いが読みの予測活動に質量共に影響を与えると指摘している。今回の実験では、中川（1996）で日本人大学生を対象とした英語能力測定調査時に素材として用いられ、既にその信頼性及び実用性が保証されたテキストを用いた²⁾（巻末資料(a)1）。

b. 作成

テスト作成の段階で、語削除の方法について Oller は定率法（fixed-ratio method：5語から10語までの範囲で、一定の間隔を置き語削除をする方法）と可変率法（variable-ratio method：内容語や機能語など、テスト作成者が独自に設けた規準に基づき削除語を決定する方法）を挙げているが、本実験では「定率法」を採用した。可変率法を用いる場合、削除語決定の際にテスト作成者の主観が介入する（恣意に基づく項目設定がなされる）という問題を回避できない。一貫性を保ち、あくまで無作為な「ランダムサンプリング」にこだわること（佐藤：1988: 14）がクローズテスト本来の特質をより引き出すことを考慮すると、定率法を採用するのが妥当である。また空欄を設ける際に、最初と最後の英文をどのように扱うかは諸説あるが、被験者にテキストの内容について必要最小限の情報を与えるために元のまま残すという慣例に従った。同様の観点から、元テキストに付随するイラスト一点を合わせてテスト用紙に載せた。津田塾大学言語文化研究所読解研究グループ（1992）によれば、イラスト提示はテキストの話題に関する背景的知識を与える意味で、被験者の読解スキーマ、とりわけトップダウン処理の活性化につながるという事実が報告されているという。本実験でイラスト提示がテキストの内容理解の助けになったかどうかは、被験者の調査用紙に対する回答傾向を分析して検討する。

c. 採点

クローズテストの採点法は、正語法（exact-word method：削除された語のみ正答とする方法）と適語法（acceptable-word method：削除された語に加えて、文脈上適当と判断される解答も合わせて正答とする方法）があり、両採点法に関する比較研究も多い。内藤（1990: 30）は、「両採点法の間には大きな差はないとする報告もある」としているが、Brown（1980）、吉田（1984）、内藤（1990）らの見解を総合し、正確を期す上で適語法を採用することにした。適語法の難しさは、採点の一貫性をいかに確保するかという点であり、これはテストの信頼性（reliability）に直接関わる

問題でもある⁽⁴⁾。適語法では文脈によって正解が複数個出ることが多く、採点の際にテスト作成者が一定の採点基準を独自に設定しなければならない。よって採点に必要な時間的・精神的労力は正語法よりも圧倒的に大きい。本研究では Celia Hammond, Sharon Mathieson (共に豪州ビクトリア州ラトロブ大学言語センター元教員) および Mark Christensen (豪州クイーンズランド州立大学学生, 英語教員) の協力を得て、予め厳正な採点基準を設けた上で適語法を採用した。その結果いくつかの設問では、本来の正答 (巻末資料(a)1) に加えて、被験者の解答中に見られたその他の語も合わせて「正答」と処理された。また「綴り」の誤りが散見したが、これについては全て「誤答」とした (中でも顕著だったのは動詞の活用形の誤りで、問題18 *sayed (said), 20 *want (wants) などがある)。しかし問題3 And (and), 4 He (he), 12 in (In) のような大文字と小文字の区別についてはいずれも「正答」とみなし得点を与えた。慎重な検討の結果、最終的に採用された採点基準は以下の通りである (表1・コメントは上記の協力者3人によるものである)。

表1 採点基準

問題番号	正答	適語法による採点基準
1	with	taking も可
3	and	but もやや違和感がある (a bit strange) が可
9	farmer	man も可
10	want	need, take も可
12	In	After も可
14	still	all も可
15	very	so も可。much もやや古めかしい (old-fashioned) が可
18	answered	said, laughed も可
20	wants	hopes, wanted も可

d. ランダムサンプリング

テキストの2文目から機械的に5語置きに語削除を行った結果、問題1から20は以下のような範疇に区分された (数字は問題番号)。

- (1) 前置詞：1, 11, 12, 16
- (2) 接続詞：2, 3, 5
- (3) 主語 (主格)：4, 6, 9
- (4) 名詞：7
- (5) 動詞 (不定詞, 態を含む)：8, 10, 13, 18, 19, 20
- (6) 副詞：14, 15
- (7) 代名詞 (所有格)：17

冠詞を除いてバランスよく各言語成分 (前置詞, 接続詞に代表される機能語および名詞, 動詞に代表される内容語) が含まれた本テストは、「ランダムサンプリング」の概念に即したものであり、

作成段階においてクローズテストの利点を活用できたといえる。

②調査用紙

クローズテスト実施後、引き続き被験者は10項目の調査用紙に回答した(巻末資料(a)2)。テキストのあらすじを日本語で説明する項目9を除いて「そう思う(5)ーそう思わない(1)」の5段階で回答するものである。回答傾向の分析から、被験者がクローズテストに対してどこまで予測文法能力を意識した解答を試みているのか、また被験者のクローズテストの解答ストラテジーがその成績にどのような影響を与えるのかを検討することにした。

③リーディングテスト

佐藤敏子助教授(つくば国際大学)の協力を得て、本実験では被験者の予測文法運用を促す問題項目を設定した。これは従来の「長文読解」形式のリーディングテストとは大きく異なるものである。被験者はまず単語数10から20語ほどの短い英文(1文もしくは2文)を読み、その後にくる英文を4つの選択肢の中から選ぶ。各設問の文脈前後の相互関係や論理構成などを考慮した上で正答を選ぶ能力が問われる。問題数は15問である⁹⁾(巻末資料(b))。

④JACETリスニングテスト

大学英語教育学会(JACET)のテスト研究開発委員会により作成されたテストである。日本の大学に在籍する学生の基礎的な英語聴解力を測定するために開発されたものであるが、問題の詳細についてはJACETとの協定により公開できない。テストは4つのPartから構成されており、それぞれ10個の設問(計40問)からなる。

Part 1. 説明文を聞き、絵や写真の内容と一致するものを選ぶ

Part 2. 問題文を聞き、もっとも適当なものを選ぶ

Part 3. 対話文と質問を聞き、もっとも適当なものを選ぶ

Part 4. 問題文と質問を聞き、もっとも適当なものを選ぶ

(2) 被験者

4年制大学学生94名(英語専攻の学生なし)。

(3) 実施時期

(1) クローズテスト・20問(2000年7月)

(2) リーディングテスト・15問(2000年7月)

(3) JACET Basic Listening Comprehension Test (Form A)

リスニングテスト・40問(2000年4月)

(4) 分析方法

テストを実施・採点し、素点データをMicrosoft Excel 97に入力し、さらにSPSS for Windows 10.0.5Jを用いて基礎統計及び相関係数を検出した。検算としてMacintoshパソコン上で統計計算ソフトStatView J-4.5でも同様の作業を行い、両ソフト間の数値が一致することを確認した。

3. 結果

(1) 基礎統計

表2 テスト基礎統計

テスト	平均点	標準偏差	最小値	最大値	範囲	分散	中央値	合計
クローズ	7.511	4.345	0	18	18	18.876	7	706
リーディング	12.532	2.159	3	15	12	4.660	13	1178
JACET	17.426	6.209	8	39	31	38.548	16.5	1638

各テストの基礎統計 (表2) である。クローズテストの平均点は7.511 (正答率37.6%) であったが、高い分散値 (18.876) に示されるように得点分布が非常に大きく、0点から18点まで広範囲にわたっている。またヒストグラム (図1) でも明らかのように、本実験で用いられたクローズテストは、今回実施されたテストの中でも JACET リスニングテストと同様に正規分布に似た、比較的バランスの取れた分布を見せている。

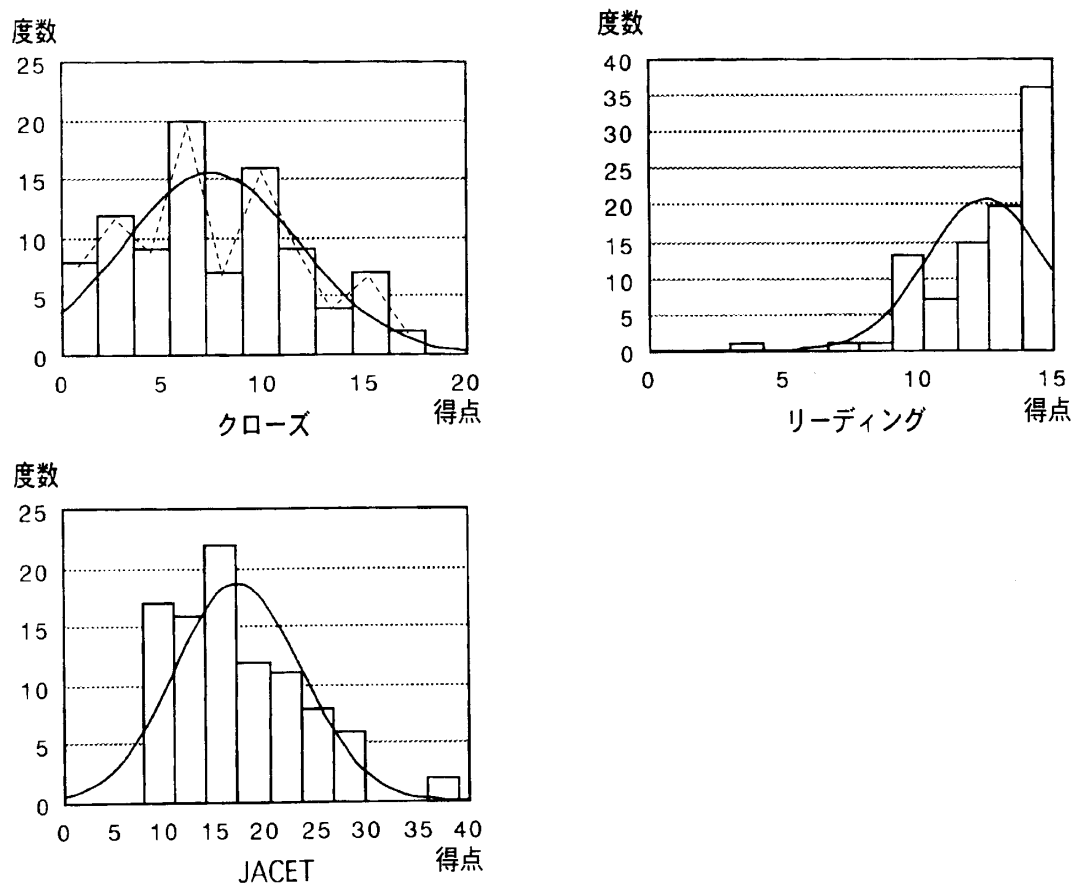


図1 ヒストグラム

クローズテストの得点分布には課題も残る。平均点近くの度数が落ち込む一方で、その両隣の度数が逆に大きくなっている（ヒストグラムの各度数の頂点を折れ線で結ぶと、2つの頂上ができる・図1点線の部分）。統計学上、平均点部分の度数が最大になるのが理想であることからすると、この特異な分布曲線は今回の被験者がクローズテストに対してまだ十分な反応をしていないか、あるいは単にテストそのものに慣れていないことによるものと推察される。低得点層に注目すると、0点の被験者が3名（3%）1点5名（5%）2点6名（6%）とここまでで全体の10%以上を占めており、正規分布曲線との比較からこれらの被験者は実際の能力以上に0点周辺に底上げされてしまったことになる。

さらにヒストグラムを検討し、各々のテストの問題点や改善点を指摘することにする。リーディングテストは平均点が12.532（正答率83.5%）となり、分布上でも満点とその近辺に被験者が集中していることから、今回のテストは若干易し過ぎた感がある。しかしこのことが「テスト項目の欠陥によるもの」と即断できる訳ではない。むしろ問題なのは項目数が少なかったことである。高得点層では満点（15点）の被験者が16名（17%）14点でも20名（21%）おり、全体で約4割の被験者が満点もしくはそれに近い点数を挙げているが、ここでは天井効果による頭打ち現象が見られ、テストが彼らの真の能力を反映しているかはやや疑わしい（クローズテストで見られた底上げ現象と反対である）。得点分布自体の均整（正規分布座標の左半分にあたる部分）は取れていることから、あとは座標の右半分に含まれる上位能力者を適切に評価できるような問題項目の開発が課題である。

JACETリスニングテストは、多くの実験研究を経て開発された「標準テスト」としての価値をここでも証明したといえよう。得点上は満点近くに突出した者が数名いるが、これらの学生は留学経験や海外滞在経験を持っており、英語学習歴に関しては他の学生とは異なった背景があるので、むしろそのことが忠実にヒストグラム上に示されたと解釈できる。被験者個人の学習履歴もテスト結果に大きく反映されるので、十分に配慮した上でデータ収集を行わなければならない。

(2) 相関係数

表3 相関係数

テスト	クローズ	リーディング	JACET
クローズ	1	0.687**	0.640**
リーディング	0.687**	1	0.546**
JACET	0.640**	0.546**	1

**相関係数は1%水準で有意であることを示す。

表3は各テスト相互の相関係数を算出したものである。どの係数も0.5（中程度の相関）以上を出力している。クローズテストを軸に値を見ると、リーディング・JACETリスニングの両テスト共に比較的高い相関係数を示している。一般的に相関係数は被験者数が多くなるほど上昇する傾向があるが、本研究では94名という割合小規模のデータサンプリングであったにも関わらず、総じて高い相関係数が得られたと分析できよう。とりわけクローズとリーディングの相関は0.687と高かった。

相関係数が高いほど、2つのテストが測定するものが重複している度合いが高いと解釈できる。今回外部基準として実施されたリーディング、JACET リスニング両テストはいずれも「先読み」を必要とする。今回の結果から推測できるのは、次にどんな内容が続くかを予測して読む（聞く）作業と、テキスト全体の内容を予測しながら空欄を埋めるクローズテスト上の作業との類似性がかなり高いということである。また中でもリーディングテストとの相関が有力だったことは、元来クローズテストが英文テキストの難易度（読み易さ）を示す指標として開発されたことから考えても納得のいく結果である。

4. クローズテストについての考察

(1) 正答率に基づく分析

表4 クローズテスト正答率

問題番号	正答	正答率 (正解者/全被験者)	問題番号	正答	正答率 (正解者/全被験者)
1	with	21.3% (20/94)	11	of	54.3% (51/94)
2	and	9.6% (9/94)	12	in	21.3% (20/94)
3	and	54.3% (51/94)	13	be	9.6% (9/94)
4	be	72.3% (68/94)	14	still	6.4% (6/94)
5	and	25.5% (24/94)	15	very	24.5% (23/94)
6	he	61.7% (58/94)	16	for	5.3% (5/94)
7	end	31.9% (30/94)	17	his	26.6% (25/94)
8	to	84.0% (79/94)	18	answered	61.7% (58/94)
9	farmer	66.0% (62/94)	19	have	69.1% (65/94)
10	want	19.1% (18/94)	20	wants	26.6% (25/94)

①正答率が高かった項目

正答率（表4）を比較すると、問題項目によって高低の差が激しいことがわかる。最も高い正答率を示したのは問題8（84.0%）であるが、その理由として、項目が文脈に依存しないことが挙げられる。“...he wanted (8) pay the farmer...”とあるので、括弧を埋めるのは不定詞の文法知識の有無による。さらに“want to...”は被験者にとって想起しやすい語句である。問題4は“...when it got dark, (4) went back...”とあるが、“went”が続けば括弧に主語が入ることは容易に想像できることから高い正答率（72.3%）となった。「適語法」を用いたことで正答率が大幅に上昇したのが問題18“...smiled and (18)...”である。テキストでは“answered”が入るのだが、全被験者中この解答をした者は0名であり、正解を得た58名（61.7%）中、57名が“said”残り1名が“laughed”と解答した（いずれも別解として正答）。改めて採点法選択の難しさを思わせる項目である。もし正語法を採用すればこの問題項目は全員不正解となるからである。既に触れたが、先行研究では正語

法、適語法の両採点法間に有意差はないとする報告もある。しかしそれは十分な数の被験者と問題項目が確保される場合において成立するものであり、本研究のような20項目の比較的小さなテストを論じる場合、そのうち一項目の正答率が採点法の違いで62%になるか0%と処理されるかは、テスト全体の質に大きな影響を与えるであろう。明確な採点基準を設ければ適語法を採用すること自体に問題はなく、せっかくの被験者の解答への努力を無駄にしないためにも、項目数の少ないテストではむしろ適語法を積極的に採用すべきではないか。

②正答率が低かった項目

正答率が伸びなかった箇所は問題13, 14, 16に集中している。問題14 (6.4%)と16 (5.3%)はその質は異なるものの、いずれも被験者にとって普段あまり問われない語法や文法要素を含んだ項目であるが故の結果である。問題14では正解者6名のうち“still”と元テキスト通りに解答したのは1名であり、残り5名は“all”と別解での正答となった。先述の通り、この項目も採点法によっては議論の分かれる項目である。また副詞を問う項目は被験者には風変わりで見新しく映るのかも知れない。問題16は全項目中最も正答率が低かった。“...and thanked the farmer (16) saying such kind things...”という設問において必要となるのは“thank ... for”という語彙、熟語知識であり、括弧の後に動名詞“saying”があることからここに入れるべき品詞はある程度まで限定されるように思われたが、正答率は伸びなかった。“...thank you for...”という定型に慣れ親しんだ学習者は多いが“...thank (ed) the farmer...”では足下をすくわれた感があったのだろうか。

③言語成分別正答率

ランダムサンプリングによって抽出された各言語成分の平均正答率は以下の通りである（数字は問題番号）。

- (1) 前置詞：1, 11, 12, 16 (25.6%)
- (2) 接続詞：2, 3, 5 (29.8%)
- (3) 主語（主格）：4, 6, 9 (66.7%)
- (4) 名詞：7 (31.9%)
- (5) 動詞（不定詞, 態を含む）：8, 10, 13, 18, 19, 20 (45.0%)
- (6) 副詞：14, 15 (15.5%)
- (7) 代名詞（所有格）：17 (26.6%)

各要素に含まれる問題項目数が異なるので同列での比較はできないが、ランダムサンプリングの原則を貫いた結果、一部を除いては平均して20%から30%台の均一な正答率となった。副詞の正答率が低いのは頻出度の低さによるものであろう。

(2) 誤答分析

ここでは被験者の誤答分析を行う（表5）。クローズテストは自由応答質問（open-ended question）であるため、受験者は時に出題者の予想をはるかに超えた解答をすることがある。傾向を検討するために、ここでは2人以上の被験者から解答があった誤答のみを扱う。

最も多かった誤答は問題10“have”（40.4%）である。文脈に沿った適切な予測があればここに

表5 クローズテスト誤答傾向

問題番号	正答	誤答(人数/割合%)		
1	with	of(18/19.1)	to(18/19.1)	for(7/7.4)
2	and	were(19/20.2)	are(11/11.7)	to(7/7.4)
3	and	at*(5/5.3)	so*(5/5.3)	for/after*(4/4.3)
4	he	and(4/4.3)	artist(4/4.3)	that(2/2.1)
5	and	he(20/21.3)	have**(14/14.9)	to(12/12.8)
6	he	had**(6/6.4)	for(3/3.2)	chance/from...(2/2.1)
7	end	time(12/12.8)	last(9/9.6)	one/part(6/6.4)
8	to	なし		
9	farmer	he*(18/19.1)	his/out/to(2/2.1)	
10	want	have(38/40.4)	pay(8/8.5)	no(5/5.3)
11	of	more(5/5.3)	the(3/3.2)	chance/from...(2/2.1)
12	in	For*(21/22.3)	Next*(7/7.4)	At(4/4.3)
13	be	pictures*** (17/18.1)	of(14/14.9)	have**(12/12.8)
14	still	not(27/28.7)	to(13/13.8)	have(5/5.3)
15	very	not(17/18.1)	had/pictures(3/3.2)	a/being...(2/2.1)
16	for	to(17/18.1)	is(14/14.9)	was(14/14.9)
17	his	to(13/13.8)	of(7/7.4)	some(5/5.3)
18	answered	happy/pleased(3/3.2)	smiled(2/2.1)	
19	have	am(13/13.8)		
20	wants	want(28/29.8)	is(12/12.8)	went(6/6.4)

* 語頭が大文字、小文字いずれかの解答を合算したもの

** have, has, hadなどの各活用形による解答を合算したもの

*** 単数形、複数形いずれかの解答を合算したもの

... それ以外の解答が複数個あるもの

“have” はまずあり得ないが、過半数に近い被験者がこのように解答したのは ‘No, I do not (10) money, but...’ とあるために、多くの被験者はここで文脈を全く無視し、語彙や文法の知識に過度に頼ることで “I do not have...” と「英文が格段に作りやすい表現」へと迂回したことがその原因と思われる。この類の「語法重視・文脈軽視」の傾向は問題11と19でも顕著であって ‘...give me one (11) your pictures.’ でも、話の展開に全く関わらないところで “one more...” の語連鎖だけが強く想起されたものと考えられるし ‘I (19) a son in London.’ も “I am a (an)...” といった、これまで過去に何度となく繰り返されたパターン・プラクティスの弊害として出現していると分析できる。問題19を解くにあたって「私自身が息子である (I am...)」のか「私に息子がいる (I have...)」のかという選択は、テキスト全体の内容理解に決定的な影響をもたらす大きな差異であるが、むしろこのような現象は被験者が文法・語法知識のみを局部集中させて活用した結果、半ば偶

然発生的・単発的に出現してきたものと思われる。

解釈が分かれる項目が問題14と15である。どちらも“not”が最も多い誤答なのだが、これらは先述の問題10, 11, 19とはやや異なる含みを持っている。“not”が文意を根底から覆してしまうことからして、これらの項目の正誤がテキスト全体の内容理解度に影響すると予測されるからである。問題14と15は「あなたの描いた絵がずっとここにあることに価値があるのだ」と自身の作品を称賛された画家が喜ぶ」というごく自然な文脈にあたるのだが、両方に“not”を入れると‘...your painting will not be here.’ “The artist was not pleased...”と正しい文脈からはことごとく遠ざかってしまう。このことから“not”挿入の有無がテキスト内容把握の失敗や読解ストラテジーの欠如に大きく関わると考えられるが、しかしその一方で“will”“was”といった空欄の周辺にある単語が、文法事項習得の段階で「肯定文↔否定文」の句型練習を繰り返し行ったことをここでも被験者に強く想起させ、それが半ば機械的な“not”の挿入につながったことも依然として否定できない。前節と同様の「語法重視・文脈軽視」の一端という可能性である。結局、これについてはクローズテストの解答からだけでは解明できず、テスト自体の限界であり弱点ともいえよう。また問題14は前節で指摘の通り副詞“still”が正解で、設問自体の珍しさが影響したのか正答率が極端に伸びなかったことが、被験者の混乱にさらに拍車をかけたとも思われる。

音声的な誤りとも解釈できるのは問題20である。正答“wants”に対して単純な文法的誤り“want”が多いことは容易に想像できるが、“went”という誤答が6名(6.4%)いる。“He (20) to become an artist.”なのだが“want”と“went”の綴りおよび音声の相似による誤りに加えて“go to...”という熟語的な連鎖から“went to...”を導いたとも考えられる。

問題12は今回の協力者のひとり Mark Christensen が私信の中で「極めて典型的な日本人学習者による誤り」と指摘している。学習段階で「時間的な幅には“for”を用いる」と一般化され易いので、即座に“for a week...”とつながってしまうのであろうが、彼によれば“in a week (time)... (1週間もすれば)” “in six months (time)... (これから6カ月のうちに)”などは「重要な語法にも関わらず、日本人英語教師の指導案から完全に忘れられている(同私信)」というべき頻出・必須事項であるという。さらにこのことから“I will be back in a minute. (すぐに戻ります)”といった文が日本人学習者には想起されにくいことや“You will be a good teacher in three years (time). (3年もすればいい教師になるよ)”とするとところを“*You will be a good teacher three years later.”と不自然な時間軸に基づいて作文してしまう弊害を合わせて指摘している。

(3) 調査用紙

クローズテスト解答傾向を調べるために合わせて配布した調査用紙(資料(a)2参照)は全被験者94名のうち91名から有効回答を得た(表6・質問紙の一部で回答が欠落した被験者3名も含む)。回答率1は5段階自己評価のうち5か4のいずれか(質問に対して「そう思う」寄りの回答)を選んだ被験者集団の結果であり、回答率2は5段階のうち2か1のいずれか(「そう思わない」寄りの回答)を選んだ被験者集団を対象としている(質問9を除く)。「3(どちらにも当てはまらない)」を選んだ被験者は分析から除外した。平均1は回答率1に挙げられた回答者全員のクローズテスト平

均点，平均2は回答率2に挙げられた回答者全員の同平均点であり，差は平均1から平均2を引いた数値である。

表6 調査用紙回答傾向とクローズテスト結果の関係

問題番号	回答率1 (回答者/全被験者)	回答率2 (回答者/全被験者)	平均1	平均2	差
1	20.9% (19/91)	46.2% (42/91)	6.2	7.6	-1.4
2	52.7% (48/91)	17.6% (16/91)	7.3	7.5	-0.2
3	73.6% (67/91)	5.5% (5/91)	7.0	9.8	-2.8
4	45.1% (41/91)	26.4% (24/91)	8.6	6.9	1.7
5	33.0% (30/91)	34.1% (31/91)	8.6	6.9	1.7
6	28.1% (25/89)	23.6% (21/89)	7.8	6.1	1.7
7	22.7% (20/88)	35.2% (31/88)	10.0	5.6	4.4
8	36.7% (33/90)	32.2% (29/90)	5.7	9.5	-3.8
9*	29.7% (27/91)*	70.3% (64/91)*	11.4*	5.8*	5.6*

* 質問9は日本語でテキストの内容を記述するものであり、回答率1と平均1が正しく記述できた被験者集団、回答率2と平均2が記述できなかった被験者集団のデータである。

回答率に注目すると，質問3（空所を埋めるのに最も活用するのは「語い（単語や熟語）」の知識である）で7割を越える被験者が「そう思う」としているのが目立つ。質問2（空所を埋めるのに最も活用するのは「文法（英語を用いる際のきまりごと）」の知識である）でも肯定が5割を超えた。これらの数値から，少なくとも「語彙・文法」がクローズテストを解く上で不可欠な要素であることを半数以上の被験者が認めていることが分かる。それにも関わらず，この認識がテスト結果にさほど反映されていないのは意外である。これらの項目では質問に対して「そう思わない」と回答した平均2集団の点数が，若干ではあるが「そう思う」と回答した平均1を上回っている。同様に「イラスト提示が読解スキーマ・トップダウン処理の活性化を促し，テキスト理解およびテスト得点に肯定的に作用する」と予測した質問1（空所を埋めるのに，イラストが参考になる）も，「そう思う」と回答した者は20.9%（平均点6.2）にとどまり「そう思わない」と回答した46.2%（平均点7.6）を回答率および平均点共に下回るという予想外の結果となった。

一方，回答傾向と平均点の差が明確であったのは質問4（前から順番に空所を埋めるうちに，だんだん話全体のあらすじが見えてくる），同7（何度か読み返して空所を埋めることができる（問題番号の後ろから前に戻って回答する））である。これらの質問はクローズテストの解法の習熟度・熟練性に直接関わるものである。クローズテストを解くには，予測文法能力をふんだんに活用しながら，次にどんな内容が続くかを常に考慮しつつ，またその予測が正しいかを確認・修整しながら読みの作業を進めていくことが必要となるが，この過程を強く意識している被験者ほどテスト点数が高いことが証明された。このテスト解答ストラテジーの修整に関する質問7に対して「そう思う」と回答した者の平均点は10.0点となり，全体の平均点（7.511）を大きく上回った。テキスト内容理解に関する読みの予測について「何度か読み返す」ことでその正誤を判断できる被験者ほど高得点

を得ていることから、今回の結果から改めてクローズテストの高得点が効果的な読みの成果と直結していることが明らかになった。

質問8（最後まで目を通して、話全体の内容（あらすじ）が見えてこない）でも「そう思わない（＝あらすじが見えた）」と回答した者の平均点が大きく伸びている。内容理解に成功した被験者ほどクローズテストの得点率が高くなっている。質問9の「この話の「オチ（おもしろみ）」を説明してください（日本語で）」という問いでは、正しく説明できた被験者の平均点（11.4）が、説明できなかった者の平均（5.8）の約2倍という大差となった。ただしこの項目については空欄の者も多く、最後までテキスト理解が不十分なままクローズテストを解いた被験者が予想以上に多かったことを意味する。またこれは余談だが、この項目は「日本語」で説明する形式であるにもかかわらず、被験者の日本語能力の不足を思わせるような、意味や論旨が明確でない回答が少なからずあった。英語教育と日本語教育の是非が問われて久しく、また近年では「日本人の英語能力向上もさることながら、乱れつつある日本語を矯正し、母国語である日本語教育の充実を一層推進するべきだ」といった風潮もある。日本人学生の手による日本語の回答を見ながら、改めて今後の課題が浮き彫りにされた感を得た。

5. 結論

「クローズテストは予測文法能力を問うテストである」という説は、今回の実験研究においてクローズテストが「先読み」を必要とするリーディング、JACETリスニング各テストと有意な相関係数を出力したことから証明された。またテスト得点の統計処理結果に加えて、調査用紙に対する回答の傾向からも、被験者が予測理論に基づいた予測文法能力を活用しており、その成果がクローズテスト得点上に反映されていることが明らかになった。

クローズテストが習熟度別クラス編成を目的としたプレイスメントテストとして活用可能であることがその得点分布から明らかになったが、能力の劣る学生を適切に評価していないという課題が残る。実際にはマイナス能力の学生がテストの得点上においてゼロ（0点）に底上げされるといった過大評価がなされてしまう傾向は、素点を基にする古典的テスト理論の限界でもあるが、さらに改良の手を加えながらテスト項目自体をより洗練させていくことが必要に思われる。その一方で今回用いたリーディングテストについては、クローズテストの場合とは全く逆の問題がある。はるかそれ以上の能力を持つと思われる多くの被験者が、名目上の「満点」に押し込まれてしまった。これを防ぐには、同種の問題項目をさらに追加してテスト全体の精度を高め、分布の極端な偏りをなくすことが有効であろう。また相関係数の検討からは、クローズテストが「読解能力」を測定するテストとして利用可能であることが改めて示唆された。

クローズテストにイラストを添付したことがテスト得点増に結びつかなかったのは予想外の結果であったが、これが被験者側の読解スキーマ不足（イラストから想像をふくらませ、トップダウン処理による読解作業を進めていくというスキーマ自体の欠落、欠損・もしくはスキーマを保持していてもそれを実行できなかったこと）による現象なのか、それとも単に偶発的なものなのかを検証

することを今後の研究課題としたい。

「リーディングテスト」と聞いてまずイメージされるのが「長文読解問題」ではないだろうか。英文を読み、内容に関する設問に答えるといった長文読解形式のテストは、指導する側のみならず学習する側にとっても極めて利便性の高いものである。テスト項目を設定する際にも、多肢選択、記述、要約、パラフレーズなど様々な種類の設問を作ることができ、この拡張性の高さが長文読解テストの隆盛につながった一因とも思われる。しかしその一方で、テスト実施の段階で多くの問題があることもまた事実である。(1)どんなテキストをテスト材料として用いるか (2)設問の内容や数は適切か (3)被験者能力とテスト難易度のバランスは保たれているか、といった課題の大半は、現在もお現場教員の経験や勘に頼る以外に有力な解決法が見つかっていない。確かに長文読解テストには、複数の要素からなる多種多様な設問を盛り込めるが、その反面「被験者のどんな能力を測定するためのテストなのか」「測定されるべきことが正確に測定されているか」といったテストの妥当性が保証されなくなってしまう危険性もまた大きいのではないだろうか。

質量を備えたリーディングテストを教師自身が作成するのは非常に難しいし、その時間的労力もはかり知れないものがある。結局は何度も試行錯誤を繰り返す以外にこれといった方法が見あたらないが、本研究において、クローズテストが日本人大学生英語学習者を対象とした場合に妥当に機能し得るという調査結果を得られたことで、今後さらに同種の被験者を対象とした、クローズテストによる読解能力測定の可能性を模索してもよいのではないだろうか。クローズテストが開発されてからじき半世紀になろうとしている。予測理論に基づき、総合的な言語能力をも測定でき、しかも作成と実施が容易な点から一時は盛んにもてはやされ、多様な被験者を対象とした実験報告がなされたクローズテストであるが、賛成・反対の両派が入り乱れて活発な議論が展開された後、その成果が十分に整理されないうちに、研究そのものが行き詰まってしまった感がある。とりわけ日本人学習者を対象とした実験研究報告の数は依然として少なく、研究の余地が残されていると思われる。今後も引き続き調査を継続することで、クローズテストに再び正当な評価を与えるための方向付けを行いたい。

(なかがわ・たけし 産業情報学科)

注

- (1) 測定されるべき能力を正確に測定できるテストほど妥当性が高く、また何回実施しても同じ結果が保証されるテストほど信頼性が高い。テストの信頼性と妥当性については、大友 (1996) が詳しい。
- (2) Hill (1965), *Elementary Stories for Reproduction* 1, 1st series, p. 12. (見出し1,000語レベルのテキストである)
- (3) Mikulecky and Jeffries (1986), *Reading Power* を材料に問題を作成した。

資料

(a) 1. クローズテスト

次の英文を読んで、それぞれの空所に入れるのに適する語を記入してください。1つの空所には必ず1つの語が入ります。I'm や isn't などの短縮形は、以下の英文中では全て I am や is not のように2つの語に分けて表記してあります。



An artist went to a beautiful part of the country for a holiday, and stayed with a farmer. Every day he went out (1) his paints and his brushes (2) painted from morning to evening,

(3) then when it got dark, (4) went back to the farm (5) had a good dinner before (6) went to bed.

At the (7) of his holiday he wanted (8) pay the farmer, but the (9) said, 'No, I do not (10) money, but give me one (11) your pictures. What is money? (12) a week it will all (13) finished, but your painting will (14) be here.'

The artist was (15) pleased and thanked the farmer (16) saying such kind things about (17) paintings.

The farmer smiled and (18), 'It is not that. I (19) a son in London. He (20) to become an artist. When he comes here next month, I will show him your picture, and then he will not want to be an artist any more, I think.'

解答 クローズテスト

1. with	2. and	3. and	4. he	5. and
6. he	7. end	8. to	9. farmer	10. want
11. of	12. in	13. be	14. still	15. very
16. of	17. his	18. answered	19. have	20. wants

(a) 2. 調査用紙 (9. を除き全て「そう思う (5) - そう思わない (1)」の5段階で回答する)

1. 空所を埋めるのに、イラストが参考になる
2. 空所を埋めるのに最も活用するのは「文法 (英語を用いる際のきまりごと)」の知識である
3. 空所を埋めるのに最も活用するのは「語い (単語や熟語)」の知識である
4. 前から順番に空所を埋めるうちに、だんだん話全体のあらすじが見えてくる
5. 次にどんな内容が続くかを予測、想像しながら空所を埋める
6. 日本語では答えが浮かぶのに、それを英語でどう書いたらいいのかわからない
「そう思う」場合には、その問題番号と「日本語で浮かぶ答え」を以下に挙げてください
()
7. 何度か読み返して空所を埋めることができる (問題番号の後ろから前に戻って回答する)
「そう思う」場合には、その問題番号を以下に挙げてください
()
8. 最後まで目を通して、話全体の内容 (あらすじ) が見えてこない
9. この話の「オチ (おもしろみ)」を説明してください (日本語で)

(b). リーディングテスト

and

Read the sentence(s), and tell which comes next.

1. It used to take months to cross the Atlantic Ocean in a ship. But now a plane can cross the Atlantic in

- a. a few months.
- b. a long time.
- c. a few hours.
- d. more time.

2. The easiest way to travel is to walk. You don't need anything special. All you need is

- a. two feet.
- b. to hurry.
- c. a car.
- d. gasoline.

3. Cars are a problem in big cities. There are too many cars in the streets. Most city streets are noisy

a. safe.

b. large.

c. parking.

d. crowded.

4. Henry went to the doctor because he could not see well. The doctor told him he should get a new pair of

a. shoes.

b. pants.

c. gloves.

d. glasses.

5. Some plants grow in very dry places. These desert plants can live for a long time with no

a. water.

b. sun.

c. food.

d. flowers.

6. Penguins are unusual birds. They can swim very well. But they have very small wings, so they can not
- a. sing.
 - b. fly.
 - c. fish.
 - d. talk.
7. On special days, the whole family has dinner at our house. Everyone came last week for my father's
- a. parents.
 - b. family.
 - c. birthday.
 - d. house.
8. The piano is a popular instrument because it is easy to play. The violin is less popular because it is more
- a. fun.
 - b. beautiful.
 - c. difficult.
 - d. boring.
9. Many farmers in Florida grow oranges. They also grow lemons, limes, and grapefruit. Florida is famous for its
- a. apples.
 - b. vegetables.
 - c. weather.
 - d. fruit.
10. In very hot countries, the sun can hurt your eyes. It is a good idea to wear a hat when you go outside. You may also need
- a. shoes.
 - b. sun glasses.
 - c. a swimsuit.
 - d. gloves.
11. The police stopped a man. They said he took some money from a store. they found the money in his pocket. So they took him to
- a. the hospital
 - b. the police station.
 - c. the bank.
 - d. his home.
12. In every country there is a different kind of money. For example, you may plan to go from Japan to the United States. Then you must change yen to
- a. dollars.
 - b. cash.
 - c. money.
 - d. airplanes.
13. Some people enjoy the winter. They like winter sports such as skiing and skating. They enjoy ice and
- a. weather.
 - b. snow.
 - c. sun.
 - d. rain.
14. The weather report says rain is coming tonight. Before we go out, we should
- a. watch the news.
 - b. lock the door.
 - c. close the windows.
 - d. read the newspaper.
15. Every country has its favorite beverage. Italians like to drink coffee after dinner. The English prefer

tea. Americans often drink

- a. ideas.
- b. soda.
- c. food.
- d. dinner.

解答 リーディングテスト

1. c	2. a	3. d	4. d	5. a
6. b	7. c	8. c	9. d	10. b
11. b	12. a	13. b	14. c	15. b

参考文献

- Brown, J. D. 1980 "Relative Merits of Four Methods for Scoring Cloze Tests," *Modern Language Journal*, 64 (3), 311-317.
- Hill, L. A. 1965 *Elementary Stories for Reproduction 1* (1st series), Oxford University Press.
- 松村幹男編 1984『英語のリーディング』大修館書店。
- Mikulecky, B. S. and Jeffries, L. 1986 *Reading Power*, Addison-Wesley Publishing.
- 内藤徹 1990「Cloze Testの可能性：Reading Cloze から Aural Cloze へ」*STEP Bulletin*, 2, 29-41.
- 中川武 1996「C-test についての一考察」卒業論文 (筑波大学)。
- Nakagawa, T. 1999 *Validation of the C-test and the M-C Cloze test among Japanese EFL Learners*.
修士論文 (筑波大学)。
- 大友賢二 1996『項目応答理論入門』大修館書店。
- Oller, J. W. Jr. 1979 *Language Tests at School*, Longman.
- 佐藤史郎 1988『クローズテストと英語教育』南雲堂。
- Taylor, W.L. 1953 "Cloze Procedure: a New Tool for Measuring Readability," *Journalism Quarterly*, 30, 414-438.
- 津田塾大学言語文化研究所読解研究グループ 1992『学習者中心の英語読解指導』大修館書店。
- 吉田一衛 1984『英語のリスニング』大修館書店。

Interaction between the Cloze Test and Expectancy Grammar of Japanese EFL Learners

Takeshi Nakagawa

Countless previous studies have suggested a cloze test to be correlated favorably with learners' knowledge of expectancy grammar. Focusing on the Japanese undergraduate-level students, one research experiment is conducted to clarify what the cloze test measures and to estimate how the test score and expectancy grammar skills overlap.

The procedure of the survey is as follows;

- (1) constructing a research plan (three types of tests and questionnaire),
- (2) collecting the data through the test administrations,
- (3) computing all the raw score data into statistics, and
- (4) eliciting some remarks from the figures.

The following conclusions are drawn by the results:

- (1) The cloze test shows a moderate correlation with the listening and the reading test respectively.
- (2) The cloze test requires the test-takers to utilize their expectancy grammar skills.
- (3) The cloze test can be used as a valid tool for a placement test.

Key words: cloze test, correlation, expectancy grammar