

Cloze test と C-test の比較¹⁾

西 田 正

1. Cloze test の略歴と C-test の提唱

1953年に Taylor が提案した cloze procedure が、外国語教育の評価の分野で初めて検討されたのは1959年である (Klein-Braley 1985: 81)。その後の外国語教育における同テストは、概略、次のような3つの段階を経て現在に至っていると思われる。第一の段階は、英語の母国語教育における cloze test の研究の影響を受けて、外国語、あるいは、第二言語としての英語の教授にこのテストを積極的に導入する時期であった。外国語教育全体が、アメリカ構造言語学と行動主義心理学の因習から、生成言語学、あるいは、認知主義心理学へと、新たな転換を模索した1970年代には、cloze test は、従来の部分的測定法に代わる新しい総合的測定法であると主張された (Oller 1976)。当時の研究は、cloze test の得点と外部基準テストの得点との相関を求め、このテストが、統合された (integrative)、あるいは、総合的な (overall) 言語能力を測定する妥当で、信頼できる試みであることを実証することに多くの努力が払われた (Oller and Conrad 1971, Oller 1973, Aitken 1977, Caufield and Smith 1981)。また、配置テストと cloze test の相関から判断して、このテストが学習者の将来の英語力を十分予測でき (Stubbs and Tucker 1974)、また、読解テストとしても有効である (Anderson 1971, 1972) など、と報告され、cloze test がいろいろな分野の測定と評価に応用できるという可能性が強調された。

第二の段階は、このような cloze test の推進論に反対論が加わった時期である。反対論は、まず、このテストが測定できる対象は、主張されているような global な言語技能ではなく、極めて低次元の言語技能であるという批判 (Alderson 1979 a, b, 1980) に始まった。この批判は、Oller (1975) や Chihara, *et al.* (1977) などで支持された cloze test の構成概念的妥当性に向けられたものであり、Porter (1983) も、cloze test は談話の言語的制約を利用する学習者の能力が測れないと反論を展開した。このような

批判に対して、Chavez-Oller, *et al.* (1985) は、Chihara, *et al.* (1977) のデータを項目分析して、Alderson 達の主張する cloze test は sentence-bound であるとする考えを否定し、従来の意見を擁護した。

第三の段階は、cloze test への慎重論と C-test の提案である。慎重論の背後には、テキスト内の語の削除法、採点法、選択するテキストなどのテスト作成に含まれる要因や、受験者の集団的特性、などが変化すると、cloze test は、本来対象とする言語能力が十分測定できない場合があるという認識があった。そして、このテストの構成概念的妥当性を高める研究として、例えば、機械的な削除に代わる意図的削除法 (rational deletion)

(Bachman 1982, 1985), 主題、題述などの談話における伝達単位を削除して作成する discourse cloze (Deyes 1984), 選択するテキストの種類と内容の親密度の判定 (西田 1986a, b) などの試みが出て来た。さらには、現在ではテスト形式を変更した C-test (Raatz and Klein-Braley 1981) が提唱されている。

このテストは、cloze test と同様に、テキスト内の言語の redundancy を人為的に制限したテストである。cloze test では、文中の語を一定の間隔で削除したり、限定した語をある意図でもって消去する方法を取るのに対して、C-test では、語の削除間隔を一語おきとし、削除部分の前半部を残す方法を採用する。Klein-Braley (1985: 80) は、cloze test や C-test のような pragmatic な言語テストは、受験者の言語運用能力を無作為に測定することが目標であり、その為には、総体としての言語を無作為に抽出する必要があると主張している。テスト作成に関するこのテストの特徴は、1) 異なる 5 から 6 種のテキストを使う、2) 削除箇所は最低 100 とする、3) テストの anchor point を確保するために母語話者が満点を取れるテストである、4) 採点は正語法に従うなどである (Klein-Braley and Raatz 1984: 136)。

2. Cloze test と C-test の比較

2. 1. 比較の問題点と目的

形式が異なる二つのテストを比較し、その優劣を決定することは、かなりの困難さが伴う。cloze test と C-test は、テキスト内の言語の redundancy を縮小するというテスト作成の大原則を共有しても、出来上がった形式が異なる。例えば、cloze test は、ひとつのテキストから作成された

独立したテストであるのに対して、C-test は、数種のテキストから作成されたサブテストの総体である。また、削除個所を100以上にするという C-test の条件を cloze test に当て嵌めて、5語から7語の削除の間隔で cloze test を作ると、それは相当長いテストになる。従って、両テストを直接比較するためには次のような方法で比較の接点を見つける必要がある。1) 削除方法は各テストの作成法に従うが、削除個所の数を cloze test に合わせる。2) テキストは複数選択し、各テキストから作成したテストをそれぞれ独立したテストとする。このようなテスト形式の一部修正によって比較する対象は、この二つのテストが、それぞれの支持者達が主張するように、受験者の運用力を測定できるテストであるならば、どちらのテストが信頼性と妥当性に関してより優位であるかを検討することである。具体的には次のことを比較検証する。つまり、1) いずれのテストが選択したテキストの影響を受け易いか、2) 信頼性はいずれのテストが高いか、3) 基準テストとの相関から判断していずれのテストが併存的、あるいは、予測的妥当性が高いかである。

2. 2. 実験方法

1) 受験者 広島大学で開講する外国語科目の英語を受講する1, 2年生, 83名。

2) 材 料 Cox, *et al.* (1976) から抜粋した3種類のテキストを基に作成した3つの cloze test と3つの C-test を使用した。テキスト A, B, C の readability (Flesch 1948) の指標は、それぞれ, 55.2 (fairly difficult), 60.0 (standard), 60.9 (standard) であり、3種のテキスト間には読み易さの点で大きな違いはない。各 cloze test には7語間隔で語を削除した24の復元個所があり、各 C-test にも同数の復元個所がある。外部基準テストは CELT の Form S-A (文法テスト) と Form L-A (聴取テスト) を利用した。

3) 手 順 同一テキストから両テストを作成したために、テスト内容が受験者に記憶されるのを避ける必要上、夏休み前に C-test を、休み明けに cloze test を実施した。両テスト (Appendix I と II を参照)²⁾ は、それぞれ、3枚1組にして与え、各テストの提示順序は counter balance した。テストの所要時間は、C-test は各々15分、cloze test は各々25分であった。外部テストの CELT は5月に実施した。文法テストの所要時間は45分、聴取テストの所要時間は40分であった。

4) 採点法 C-test には正語法を, cloze test には英語の母語話者の協力を得て適語法を採用した。³⁾

2. 3. 実験結果

表1 各テストの得点と信頼性 (N=83)

		\bar{X}	SD	MAX	MIN	Range	KR 20
C-test	A	7.69	3.01	14.0	2.0	12.0	.66
	B	10.25	4.19	20.0	3.0	17.0	.77
	C	7.34	2.94	15.0	2.0	13.0	.69
	T	25.29	8.50	47.0	9.0	38.0	.78
Cloze	A	6.88	2.51	13.0	0.0	13.0	.55
	B	5.28	2.22	11.0	0.0	11.0	.40
	C	6.45	2.68	13.0	0.0	13.0	.56
	T	18.71	5.97	34.0	3.0	31.0	.67
CELT	S-A	65.50	12.16	91.0	29.0	62.0	.86
	L-A	45.81	12.48	90.0	24.0	66.0	.75
	T	111.06	21.30	171.0	53.0	118.0	

各テストの平均得点, 標準偏差, 得点の広がり, 信頼性係数は表1に示した通りである。C-test A, B, C間の平均得点の差は $F(2,246) = 17.99$ $p < .01$ で有意であった。テスト A と B の平均得点差は $t = 2.58$ ($p > .05$), B と C の差は $t = 2.93$ ($p < .01$), A と C との差は $t = .53$ であった。従って, C-test では, B が最も易しく, A と C とはテストの難易度が同程度であった。cloze test, A, B, C間の平均得点差も $F(2,246) = 6.13$ $p < .01$ で有意であった。A と B の差が $t = 4.36$, B と C の差が $t = 3.06$ であり, それぞれ 1%水準で有意な差となった。しかし, A と C との差は $t = 1.08$ ($p > .05$) であった。従って, cloze test では, B が最も困難なテストであり, A と C のテストの難易度は同程度である。同一のテキストから作られた cloze test と C-test の難易度は必ずしも同じではないことがわかる。

KR20から判断できる各テストの信頼性を比較すると, C-test が cloze test に勝っている。C-test の信頼性は.77から.66であるのに対して, cloze test は.56から.40である。CELTの文法テストと聴取テストはともに高い信頼性を示しており, 外部テストとして信頼できるテストである。

表2 C-test 間の相関と妥当性

C-test	A	B	C	T	L-A	CELT T
A	—	.52***	.60***	.82 (.63)***		
B		—	.53***	.86 (.58)***		
C			—	.82 (.64)***		
T				—		
S-A	.55***	.54***	.47***	.62***	.52***	.87 (.49)***
L-A	.51***	.47***	.44***	.56***	—	.86 (.47)***
T	.60***	.58***	.51***	.68***		()=修正値

*** $p < .001$

表3 Cloze test 間の相関と妥当性

Cloze	A	B	C	T	()=修正値
A	—	.28*	.49***	.77 (.48)***	
B		—	.42***	.70 (.42)***	
C			—	.84 (.58)***	
T				—	()=修正値
S-A	.57***	.52***	.58***	.72***	
L-A	.37***	.30***	.43***	.47***	
T	.54***	.46***	.59***	.69***	

* $p < .05$ ** $p < .01$ *** $p < .001$

表2と表3は、C-testと cloze testに関するそれぞれのテスト間の相関係数と外部テストとの相関係数をまとめた結果である。なお、()内の数値はテストの総合点と各テストの得点との相関を修正した値である (cf. Henning 1987:69)。表2と3の相関値は全て0.1%、あるいは、5%以下の危険率で有意相関である。テスト間の相関を見ると、cloze testにおけるよりもC-testにおいてその相関は一定している。C-testにおいては相関値が.5から.6台であるのに対して、cloze testにおいては.4台が多く、有意な相関であるが、.28という低い相関がある。cloze testはC-testと異なり選択するテキストに影響を受け易いようである。つまり、受験者は、いずれのC-testにおいても同じような得点を取る傾向が強いが、cloze testでは、どのテストを受けたかによって、得点が変わり易いと言えよう。外

表4 C-test と Cloze test 間の相関

C-test	A	B	C	T
Cloze A	.42***	.45***	.41***	.50***
B	.28*	.28*	.40***	.38***
C	.49***	.49***	.49***	.58***
T	.52***	.53***	.59***	.64***

* $p < .05$ *** $p < .001$

部テストとの相関に関しては、ここでは cloze test と聴取テストとの相関が低いことを指摘するに留める。

表4は、3種のC-testとcloze test間の相関を示したものである。cloze Bと各C-testとの相関は低いが、両テスト間の相関は似かよっている。しかし、同一テキストから出来ているテスト間の相関が他のテストとの相関より高いとは限らない。このことは既に指摘したように同一テキストに基づくcloze testとC-testが同程度の難易度にならなかったことと関連している。

表5 C-test 間の相関差の検定

C-test	A	B	C	x^2
S-A	.55	.54	.47	.55
L-A	.51	.47	.44	.06
T	.60	.58	.51	.74
t	.35	.81	.31	

表6 Cloze 間の相関差の検定

Cloze	A	B	C	x^2
S-A	.57	.52	.58	.37
L-A	.37	.30	.43	.91
T	.54	.46	.59	1.32
t	2.27*	2.30*	1.78	

* $p < .05$

表5と6は、C-testとcloze test内の各々のテストと基準テストとの相関の差を検定した結果である。C-test A, B, Cと文法テストとのそれぞれの相関は、.55, .54, .47であり、三つの相関の差は有意ではない ($x^2 = .55$)。また、聴取テストとの相関差も有意でない ($x^2 = .06$)。さらに、文法と聴取テストのCELT総合点との相関差も有意となっていない ($x^2 = .74$)。C-test A, B, C, それぞれに関する文法と聴取テストとの相関の差を検定すると、C-test Aと文法テストとの相関が.55、聴取テストとの相関が.51あり、その差は $t = .35$ となり有意ではない。同様に、C-test BとCについても2つの基準テスト間の相関の差は有意でない ($t = .81$ 及び $t = .31$)。このような相関差は表6におけるcloze testにもほぼ当て嵌ま

る。cloze A, B, C 間の文法テストとの相関差は有意でなく ($x^2=.37$), また, 聴取テストとの相関差も有意でない ($x^2=.91$), さらに, 基準テストの総合点との相関差も同様である ($x^2=1.32$)。cloze test と C-test との唯一の違いは, 2つの基準テストとの相関差が cloze A と B には認めら

表7 C-test と Cloze test 間の相関差の検定

	C-T A	Cloze A	t	C-T B	Cloze B	t	C-T C	Cloze C	t
S-A	.55	.57	.23	.54	.52	.20	.47	.43	.40
L-A	.51	.37	1.50	.47	.30	1.25	.44	.43	.10
T	.60	.54	.71	.58	.46	1.16	.51	.53	.22

れた点である。cloze A と文法テストとの相関と cloze A と聴取テストとの相関の差は有意であり ($t=2.27 p<.05$), B についてもその差は有意 ($t=2.30 p<.05$) ある。従って, cloze A と B は聴取テストよりも文法テストとの相関が高くなった。

表7は, 同一テキストから作成した C-test と cloze test に関する基準テストとの相関値の差を検定した結果である。文法テストと C-test との相関値と文法テストと cloze A との相関値の差は有意でない ($t=.23$)。また, 聴取テストと C-test との相関と cloze A と聴取テストとの相関との差も有意でない ($t=1.50$)。さらに, 基準テストの総合点についてもこの2つのテスト間の相関差は有意でない ($t=.71$)。表から明らかのように, 他の C-test と cloze test においても基準テストとの相関の差は検定されなかった。以上のことから同一のテキストから作られた C-test と cloze test は外部基準テストと同程度に相関することがわかる。

本実験の結果は次のようにまとめることができるであろう。1) C-test においても, cloze test においても, テキスト間の難易度が見られた。2) KR20の信頼係数から判断すると, C-test は cloze test よりも信頼性が高い。3) C-test (A, B, C) 内の相関は cloze test (A, B, C) 内の相関に比べて安定している。つまり, C-test の得点は選択するテキストに影響を受けにくい。4) C-test A, B, C と文法テスト及び, 聴取テストとの相関は差がなく, このことは cloze test A, B, C についても言えることである。5) C-test は文法テストと聴取テストに相関する程度が同じであるが, cloze test は文法テストとの相関が高く出るテストである。5) 基準テストとの相関から C-test の併存的妥当性は cloze test のそれと殆ど変

わらない。

2. 4. 考察

本実験の結果を一般化するためには、次の2つの要因を吟味する必要がある。まず第一の要因は受験者集団の特性である。Klein-Braley and Raatz (1984) は、cloze test を等質集団に実施すると、テストの信頼性と妥当性は満足の行くレベルに達しにくいと指摘している。また、中国人英語学習者に cloze test の実験を行った Brown* (1984) も、受験者の英語学力の等質性が同テストの信頼性と併存的妥当性を低めることを認めている。本実験に参加した大学生の英語学力は、昭和62年度入学の広島大学1年生1678名に実施した CELT の文法テストの成績 ($\bar{X}=65.54, SD=11.69$) と殆ど変わりなく、受験者は広島大学の学生の平均的な英語力を持った集団であった。従って、もし、英語力に幅のある学習者の集団が本実験に参加しておれば、C-test と cloze test の信頼性と妥当性に何らかの有意な差が出たかもしれない。しかし、その場合は、双方のテストの信頼性も妥当性も高くなるので、今回の実験結果と同様に両テスト間に大きな差が出るとは予想できない。事実、英語力の分布が広い大学生群に、cloze test と C-test を実施した高梨 (1988) では、テストの信頼性も妥当性も本実験結果より高くなっているが、テスト間には信頼性と妥当性に有意な差がないと報告されている。それ故に、両テストの信頼性と併存的妥当性がこのよ

表8 C-test における内容語と機能語の割合

	Text A	Text B	Text C
Content Words	11	13	13
Function Words	13	11	11

表9 Cloze test における内容語と機能語の割合

	Text A	Text B	Text C
Content Words	12	8	11
Function Words	12	16	13

うに似かよる結果となったのは、受験者の特性ではなく、テスト自体に原因があると思われる。

第2の要因はテキストである。この要因には、テキストの難易度、テキスト内容の得点へのバイアス、削除語の特性、などが含まれる。本実験で採用した3種のテキストは、readabilityがほぼ等しく、内容は受験者が特別な知識を必要とするほどの偏りはない。本実験で使用したテストに含まれる削除箇所を内容語と機能語に分け、その分布をまとめたのが表8と9である。表8から判明するように、C-testでは内容語と機能語の分布に関してテキスト間で差がない ($x^2 = .44 p > .05$)。一方 cloze test ではテキストBにおける内容語と機能語の出現頻度が他のテキストより異なっているように思われるが、その分布の検定をすると有意差がない ($x^2 = 1.47 p > .05$)。また、C-testと cloze test 全体を通しての内容語と機能語の分布の差も有意ではなかった ($x^2 = 1.92 p > .05$)。従って、このような両テストの削除部分の似かよりが実験結果に影響したかも知れない。

3. 結論

従来の cloze test の modification として登場した C-test は、cloze test に比べてテストの信頼性がある程度高いが、基準テストとの併存的妥当性に関しては必ずしも cloze test に対して優位であるとは言えない。

[注]

- 1) 本論は、第5回大学英語教育学会中国・四国支部大会(昭和63年6月5日)における「クローズテストの信頼性と妥当性をめぐって」と第19回中国地区英語教育学会(昭和63年10月8日)における「C-testについて」と題した口頭発表内容の一部に加筆、修正したものである。
- 2) cloze test の実例は紙面の都合上一部のみ掲載した。
- 3) Ms Carol Rennart に感謝する次第である。

References

- Aitken, Kenneth G. (1977) "Using Cloze Procedure as an Overall Language Proficiency Test," *TESOL Q*, 11, 1, 59-67.
- Alderson, J. Charles (1979a) "The Cloze Procedure and Proficiency in English as a Foreign Language," *TESOL Q*, 13, 2, 219-227.
- (1979b) "Scoring Procedures for Use on Cloze Tests," in Yorio, Carlos A., Kyle Perkins, and Jacquelyn Schachter (eds.) *On TESOL '79: The Learner in Focus* (TESOL) pp. 193-205.
- (1980) "Native and Nonnative Speaker Performance on Cloze Tests," *LL*, 30, 1, 59-76.
- Anderson, J. (1971) "A Technique for Measuring Reading Comprehension and Readability," *ELT*, 25, 2, 179-182.
- (1972) "The Application of Cloze Procedure to English Learned as a Foreign Language in Papua and New Guinea," *ELT*, 27, 1, 66-72.
- Bachman, Lyle F. (1982) "The Trait Structure of Cloze Test Scores," *TESOL Q*, 16, 1, 61-70.
- (1985) "Performance on Cloze Tests with Fixed Ratio and Rational Deletions," *TESOL Q*, 19, 3, 535-556.
- Brown, James Dean (1984) "A Cloze is a Cloze is a Cloze?" in Jean Handscombe, Richard A. Orem, and Barry P. Taylor (eds.) *On TESOL '83: The Question of Control* (TESOL) pp. 109-119.
- Caufield, Joan and William C. Smith (1981) "The Reduced Redundancy Test and the Cloze Procedure as Measures of Global Language Proficiency," *MLJ*, 65, 1, 54-58.
- Chavez-Oller, Mary Anne, Tetsuro Chihara, Kelley A. Weaver, and John W. Oller, Jr. (1985) "When Are Cloze Items Sensitive to Constraints Across Sentences," *LL*, 35, 2, 118-125.
- Chihara, Tetsuro, John Oller, Kelley Weaver, and Mary Anne Chavez-Oller (1977) "Are Cloze Items Sensitive to Constraints Across Sentences?" *LL*, 27, 1, 63-73.
- Cox, Alwyn, Neville Grant, and Helen O'Neill (1976) *English Examined* (Thomas Nelson and Sons Ltd.).
- Deyes, Tony (1984) "Towards an Authentic 'Discourse Cloze'," *AL*, 5, 2, 128-137.
- Henning, Grant (1987) *A Guide to Language Testing: Development · Evaluation · Research* (Newbury House Publishers).
- Klein-Braley, Christine (1985) "A Cloze-up on the C-Test: A Study in the Construct Validation of Authentic Tests," *Language Testing*, 2, 1, 76-104.
- and Ulrich Raatz (1984) "A Survey of Research on the C-Test," *Language*

Testing, 1, 2, 134-146.

西田 正 (1986a) 「クローズテストの妥当性とテキストの選択」『英語教育学研究——垣田直巳先生退官記念論文集』(大修館書店) pp. 433-444.

—— (1986b) 「『物語文』クローズと『説明文』クローズ——テキストの選択とクローズテスト——」『言語文化研究』(広島大学総合科学部紀要V) 12巻, pp. 124-143.

Oller, John W. Jr. (1973) "Cloze Tests of Second Language Proficiency and What They Measure," *LL*, 23, 1, 104-118.

—— (1975) "Cloze, Discourse, and Approximation to English," in Burt, Maria K. and Heidi C. Dulay (eds.) *On TESOL '75: New Directions in Second Language Learning, Teaching and Bilingual Education* (TESOL) pp. 345-355.

—— (1976) "Language Testing," in Wardhaugh, Ronald and Douglas Brown (eds.) *A Survey of Applied Linguistics* (The University of Michigan Press) pp. 275-300.

—— (ed.) (1983) *Issues in Language Testing Research* (Newbury House Publishers, Inc.).

—— and Christine A. Conrad (1971) "The Cloze Technique and ESL Proficiency," *LL*, 21, 2, 185-195.

Porter, Don (1983) "The Effect of Quantity of Context on the Ability to Make Linguistic Predictions: A Flaw in a Measure of General Proficiency," in Hughes, Arthur and Don Porter (eds.) *Current Developments in Language Testing* (Academic Press) pp. 63-74.

Raatz, Ulrich and Christine Klein-Braley (1981) "The C-Test — A Modification of the Cloze Procedure," in Culhane, T., C. Klein-Braley, and D.K. Stevenson (eds.) *Practice and Problems in Language Testing* (University of Essex Department of Language and Linguistics Occasional Papers No. 26) pp. 113-138.

Stubbs, Joseph Bartow and Richard Tucker (1974) "The Cloze Test as a Measure of English Proficiency," *MLJ*, 58, 5-6, 239-241.

高梨芳郎 (1988) 「クローズ法と C-test について」第14回全国英語教育学会京都研究大会口頭発表資料。

Taylor, Wilson L. (1953) "Cloze Procedure: A New Tool for Measuring Readability," *Journalism Quarterly*, 30, 415-433.

Appendix I

C-Test (A)

The fight to eradicate malaria from the world is gradually being lost. A rec ___ (1) report b ___ (2) the Wo ___ (3) Health Organi ___ (4) speaks o ___ (5) a resur ___ (6) of mal ___ (7) in so ___ (8) areas fr ___ (9) which i ___ (10) had ear ___ (11) been elimi ___ (12). Coming a ___ (13) it do ___ (14) at t ___ (15) very ti

___ (16) when ano ___ (17) killer dis ___ (18) is be ___ (19) finally sta ___ (20) out, th ___ (21) news i ___ (22) particularly disapp ___ (23). But t ___ (24) success of the anti-smallpox campaign illustrates, by contrast, the central problem facing any attempt to eradicate malaria.

C-Test (B)

I was not at all happy at the prospect of the 700-mile drive from Dar-es-Salaam to Nairobi. It w ___ (1) not th ___ (2) I disl ___ (3) driving b ___ (4) I susp ___ (5) that wh ___ (6) is a ve ___ (7) pleasant tr ___ (8) in t ___ (9) dry sea ___ (10) could pr ___ (11) disastrous dur ___ (12) the lo ___ (13) rains, a ___ (14) the mons ___ (15) had arr ___ (16) the prev ___ (17) week. I w ___ (18) fully aw ___ (19) of t ___ (20) possibility o ___ (21) a brea ___ (22), of hit ___ (23) large ani ___ (24) as they stopped, dazzled by my headlamps, or even of skidding off the road.

C-Test (C)

The rat has many enemies and its greatest enemy is man. Nevertheless, ra ___ (1) are surv ___ (2). Although m ___ (3) has wa ___ (4) an unen ___ (5) war aga ___ (6) them, ra ___ (7) have method ___ (8) and systema ___ (9) colonized t ___ (10) world. I ___ (11) almost ev ___ (12) country, fr ___ (13) the free ___ (14) cold o ___ (15) the Arc ___ (16) regions t ___ (17) the enerv ___ (18) heat o ___ (19) the tro ___ (20), rats a ___ (21) to b ___ (22) found, thri ___ (23), increasing a ___ (24) multiplying.

Appendix II

Cloze Test (A)

The fight to eradicate malaria from the world is gradually being lost. A recent report by the World 1) _____ Organization speaks of a resurgence of 2) _____ in some areas from which it 3) _____ earlier been eliminated. Coming as it 4) _____ at the very time when another 5) _____ disease is being finally stamped out, 6) _____ news is particularly disappointing. But the 7) _____ of the anti-small-pox campaign illustrates, by 8) _____, the central problem facing any attempt 9) _____ eradicate malaria. (後略)

Cloze Test (B)

I was not at all happy at the prospect of the 700-mile drive from Dar-es-Salaam to Nairobi. It was not that I disliked 1) _____ but I suspected that what is 2) _____ very pleasant trip in the dry 3) _____ could prove disastrous during the long 4) _____, and the monsoons had arrived the 5) _____ week. I

was fully aware of 6) _____ possibility of a breakdown, of hitting 7) _____ animals as they stopped, dazzled by 8) _____ headlamps, or even of skidding off 9) _____ road (後略)

Cloze Test (C)

The rat has many enemies and its greatest enemy is man. Nevertheless, rats are survivors. Although man 1) _____ waged an unending war against them, 2) _____ have methodically and systematically colonized the 3) _____. In almost every country, from the 4) _____ cold of the Arctic regions to 5) _____ enervating heat of the tropics, rats 6) _____ to be found, thriving, increasing and 7) _____.

Apart from their ability to adapt 8) _____ to varying climatic conditions, rats possess 9) _____ quality that (後略)

A Comparison of the Cloze and C-test as a Measure of English Proficiency

Tadashi NISHIDA

This paper reports the results of an experiment conducted to compare the reliability and validity of the cloze test and the C-test as a device for measuring global English language proficiency. Three different passages with approximately the same levels of readability were chosen from an ESL textbook for the construction of cloze tests A, B, and C and C-tests A, B, and C. For each cloze test every seventh word was deleted up to a total of 24 blanks. For the C-tests the second half of every second word was deleted until the same number of mutilations was reached. A group of 83 Japanese university students of English participated in the experiment.

Standardized structure and listening subtests of *CELT* were used for the criterion measures against which the two reduced redundancy tests were evaluated. Reliability coefficients computed by the Kuder Richardson formula 20 were obtained for each test. Pearson product-moment correlations were also calculated between the tests and the criterion tests.

The following research questions were addressed: 1) Are the C-tests more likely to be affected by the selection of passage? 2) To what extent are the cloze and C-tests different in reliability? 3) Is there any difference between the tests in predicting the student's success in learning?

The statistics computed in the experiment showed no significant differences between the tests except that the cloze tests are slightly more affected by the different passages.