

「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 —

西 田 正

序 論

言語テスト法はそれが妥当で信頼できる測定法として確立するためには、長年の実験と実践による検証を受けなければならない。¹⁾ Cloze Test (以下 CT) もその例外ではない。従来 CT については、英語の proficiency test や読解テストなど、外部テストとの併存的妥当性が高く、また、受験者の将来の英語学力を予測する上で妥当なテストであると報告されているが (Oller & Conrad 1971, Wainman 1979, Shohamy 1982), 本来意図した英語能力をどの程度正確に測定できるかという CT の構成概念的妥当性に関しては意見が分かれるところである。

筆者は CT の妥当性を吟味する上で受験者に与えるテキストの種類と親密度が重要な要因であると仮定して、実験を継続して来た (西田 1985, 印刷中)。本稿では、この妥当性をめぐる対立点を挙げ、テキストと親密度を再考し、大学生を対象に新たに実施した実験の結果を報告する。

I. 構成概念的妥当性をめぐる対立点

明確な論争点は、CT が受験者の持つ英語の能力を文を越えた談話のレベルで捉えることができるか否かである。Oller (1973) (1976) は、英語運用能力の基底にある予測文法、つまり、異なるレベルの言語要素を瞬時に処理し必要によっては複数の技能を連動させる運用力こそが CT の測定対象であると主張している。また、Oller (1975) は、文中の語を削除する間隔を変えて、受験者が利用できる文脈を制御した CT を与えたところ、文脈の長さがテストの得点に反映したことから CT は受験者の文脈を利用する力が測れると述べている。

一方、Alderson (1983) は、クローズ得点は文法および語彙のテスト得点とそれぞれ相関が高いので、CT は文内の言語要素を処理する低次な

言語能力を測るにすぎず、文を越えた談話レベルで作動する読解力のような高次な運用力は扱えず、“Cloze is essentially sentence-bound.” (1983 : 211) と反論している。Porter (1983) も CT の構成概念的妥当性に疑義を挟んでいる。彼は、新聞、小説、手紙などから96の文を選び、ある特定の語を中心に左右に5語残すクローズから12語残すクローズまで8種類のCTを実施した。もし、Oller (1975) の主張のように読み手が談話上の制約を利用するのであれば、語の削除間隔がCTの成績に現れるはずである。しかし、予想は外れ、CTの受験者は削除語の復元時に文脈や談話のような、‘wide-ranging linguistic constraints’ (1983 : 74) を使うとは言えず、従って彼は、CTはこのような能力が測れないと判断している。

文脈を利用する程度を診る方法に選択したパッセージの文の配列を無作為に変えたクローズ (scrambled version) と原文のままのクローズ (sequential version) を受験者に与え、両クローズ間の得点を比較する方法がある。英語学力の異なる日本人学習者群にこの実験を試みたChihara *et al.* (1977) では、文脈の有無と英語学力の主効果と両要因間に交互作用が認められ、学力が高い受験者ほどCTの得点が高くなったが、これは文脈がある sequential version に限られることが判明した。また、英語の母国語話者に行ったRye (1984) の報告でも sequential version の得点が優位であることが認められている。しかし、Shanahan *et al.* (1982) では、sequential と scrambled version に見られる文脈の有無は得点の分散のごく僅かな部分を説明できただけで、CTの構成概念的妥当性に疑問を投げ掛ける結果となっている。

以上のような論争の中で、CT作成の出発点であるパッセージ、あるいは、テキストについて本質的な考察がなされていないのは意外である。Chavez-Oller *et al.* (1985) は、受験者の文脈利用を検出できなかったPorterが使用したCTの材料はテキストとしての integrity に欠け、また、文章理解を困難にする諸要因が入り混んでいると批判しているが、パッセージについて実質的な検討が必要であるという認識はCT研究においてはごく最近のことであると思われる²⁾。勿論、Alderson (1979) (1983)、Oller (1972) (1979)、Mullen (1979) 等が、選択するパッセージの難易度がCTの信頼性と妥当性に影響を与えると報告しているし、また、CT作成上の留意点も指摘しているが、パッセージやテキストに関する体系的な研究は不十分である。

II. CTの構成概念的妥当性とテキスト要因

1. テキストの種類 (物語文と説明文)

Brown (1983) は、語を削除する出発点を変えても、出来上がる CT に含まれる cohesive devices は一定であると指摘している。しかし、異なるジャンルから選ばれたテキスト間にその出現頻度の差が出ることを否定したわけではない。物語文 (narrative discourse) のように出来事が時間の経緯に沿って展開する文章では、cohesive devices の頻度が高いと予想できる。一方、文と文との結び付きが弱く単文あるいは二・三の文ごとに意味が完結するように文章が構成されている説明文 (expository discourse) においては、文間の意味的繋がりが薄い為に cohesive devices の頻度は低くなるであろう。テキスト内に現れる cohesive devices の出現の差は、それが直接、CT の得点に反映するというよりも、これによって文章構成が文レベルで独立しているか、あるいは、談話レベルで文相互が依存しているかを判断する目安となるであろう³⁾。従って、文章構成法が異なるテキストから CT を作成しテストを実施すると、受験者の文脈利用の程度が測れない場合がある。つまり、CT の構成概念的妥当性を検討する際に、物語文と説明文のいずれを使用するかによって、結果が異なる可能性が考えられる。物語文を原文とするクローズ (以下「物語文」クローズ) から作られた scrambled version と原文のままの sequential version の間には文脈の有無が CT 得点に深く関るが、説明文を原文とするクローズ (以下「説明文」クローズ) の両 version 間には CT 得点の大きな差は期待できない。

2. テキストの親密度

CT の受験者がテスト中にパッセージを読みその一部を理解しようとする限りにおいて、テキストが受験者にとって身近な内容であるか否かは無視できない要素である。最近提案されている interactive model は、読解過程における読み手の持つ先行知識 (prior knowledge) の重要性に注目している。読みの行為は著者と読者との interaction であり、書かれた内容に読み手の持つ知識が積極的に付加されて初めて内容の理解へと進む。それ故に、Bransford *et al.* (1984) は「理解は関連知識の活性化に左右される」(p. 28) と述べている。文章理解や内容理解が読み手の知識と結び付くのであれば、CT のテキストに現れた内容について受験者が所有している知識の質と量 (schema) が得点に影響することは十分予測できる。Johnson (1982) が受験者にとって文化的に疎遠なパッセージを選ぶとテ

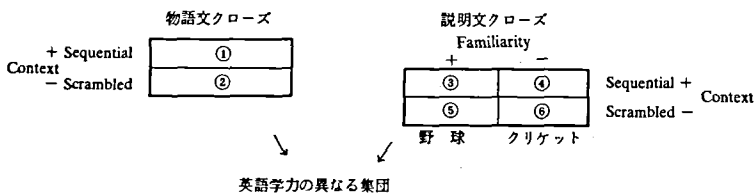
ストに文化的な bias が入ると忠告したり、Bernhardt (1984) が、CT の話題が受験者にとって不慣れな内容であれば、テストの不出来の原因は背景的知識の不足にあると指摘するのも、テキストの親密度が妥当性に深く関る点を見越しているからである。

Schema は長期記憶に貯えられた知識の体系であり、情報理解の外枠を成し、いわゆる上から下への理解を促進させると同時に下から上への理解も援助する。受験者にとって話題が身近で活性化できる豊かな schema がある場合（以下述べる「野球」クローズの受験時に当てはまる）、受験者は自己の知識体系を十分活用して削除語を復元できるであろう。しかし、関連知識の乏しい親密度の低い CT（「クリケット」クローズ）では、このようなことは期待できない。従って、受験者に親密度の異なる CT を実施すると、単に得点に差が生じるだけでなく、本来測定可能な能力が測れない事態が起こりうると思われる。

Ⅲ. CT の構成概念的妥当性とテキスト要因に関する実験

CT の妥当性とテキストの種類及び親密度との関係を検討するために、図 1 のような実験計画を立てた。

図 1



1) 材料 物語の展開するテキスト（総語数397）を8語おきに語を削除し46の復元項目を含む物語文クローズと、*The Encyclopedia Americana* (1958: 302-303) 及び *The Encyclopaedia Britannica* (1959: 688-689) から、野球とクリケットの試合に関する説明文を抜き出し、両テキストの readability (Flesch 1948) が等しくなるように修正を一部加えたパッセージ（野球の総語数360、クリケットの総語数364）を6語おきに語を消去し60の復元箇所のある説明文クローズを用意した。物語文の readability は88.33、野球クローズは70.65、クリケットは71.26で、前者

は easy に後者の 2 クローズは fairly easy にそれぞれ属する。各クローズから scrambled version を作り計 6 種の CT を準備した (図 1 と Appendix 1 を参照)。

2) 受験者 国公立大学 1 年生と 2 年生 156 名。受験者は事前に英語標準テスト CELT の文法テスト (S-A Form) を受け, A, B, C の 3 群に区分された。

3) 手順 物語文クローズと説明文クローズを A B C 3 群に実施した。前者のクローズでは各群を 2 分し, CT ①か②の何れか 1 つを与えた。2 週間後の後者のクローズでは, 各群を 4 分し, subgroup 1 は③→⑥, subgroup 2 は④→⑤, subgroup 3 は⑥→③, subgroup 4 は⑤→④の順序で受験させた。各受験者は CT の練習を行った後, CT の内容に関する親密度を 5 段階評定してから CT に取りかかった。各 CT の所要時間は 30 分。実施期間は昭和 61 年 6 月であった。採点法は EXACT 法を採用した。

結 果

物語文クローズの各群の平均得点, SD, 及び CELT の結果は表 1 にまとめた通りである。そのクローズ得点のプロフィールは図 2 に示した。各群の英語学力と文脈の有無 (sequential/scrambled version) を独立変数とし, CT の得点を従属変数と見なした 2 要因分散分析の結果, 表 2 の如く,

表 1 物語文クローズと英語学力

グループ	Sequential (46)	Scrambled (46)	CELT (100)
A (N=52)	$\bar{X}=20.96$	$\bar{X}=16.77$	$\bar{X}=74.08$
	SD = 6.00	SD = 3.99	SD = 9.72
B (N=52)	$\bar{X}=15.04$	$\bar{X}=13.65$	$\bar{X}=64.65$
	SD = 4.19	SD = 4.06	SD = 8.35
C (N=52)	$\bar{X}=13.77$	$\bar{X}=12.88$	$\bar{X}=56.33$
	SD = 4.17	SD = 3.43	SD = 9.22

英語学力 [$F(2, 150) = 23.55$ $p < .001$] と文脈 [$F(1, 150) = 9.43$ $p < .001$] の主効果が認められた。しかし, 英語学力と文脈の交互作用 [$F(1, 150) = 2.15$ $p > .05$] はなかった。テューキー法による群間の得点差は 5% レベルにおいて $A > C$, $A > B$ で B と C 間には有意差がなかった。これは sequential と scrambled の両テストにおいて確かめられた。また, 3 群を込みにした両 version の比較では sequential テストの優位が見られた。

図2 物語文クローズ得点プロフィール

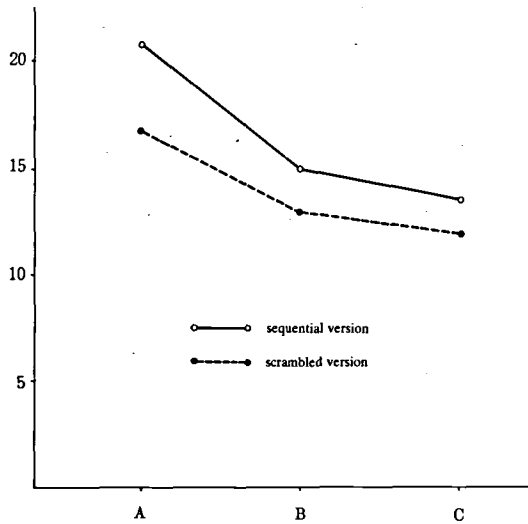


表2 物語文クローズの2要因分散分析

要因	M.S.	df	F
A	451.853	2	23.553***
B	180.923	1	9.431***
A×B	41.327	2	2.154
E	19.185	150	
T		155	

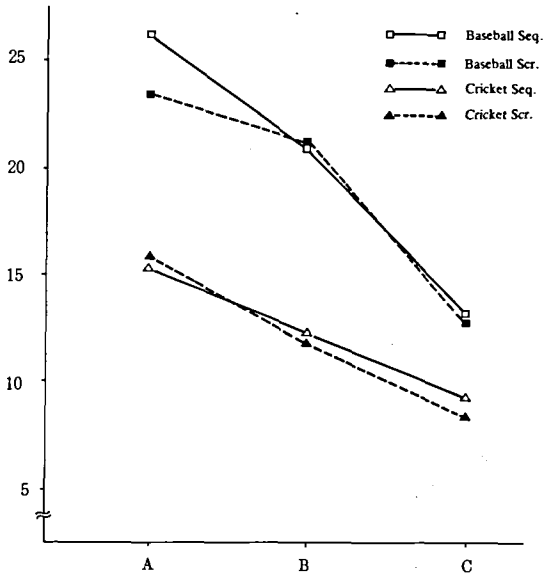
A=英語学力 B=文脈 *** $p < .001$

ABC3群に実施した説明文クローズの結果は表3に挙げた。そのプロフィールは図3に示した。この図からわかるように、説明文クローズでは、sequentialとscrambled versionの間の得点差はほとんどなく、物語文クローズのプロフィールと大きな対比を成している。英語学力、文脈、親密度の3要因分散分析の結果、表4にあるように、英語学力 [$F(2,300) = 99.35$ $p < .001$] と親密度 [$F(1,300) = 185.20$ $p < .001$] の主効果が認められたが、文脈 [$F(1,300) = 0.76$] の効果はなかった。各要因間の交

表3 説明文クローズ得点

グループ	野球クローズ (60)		クリケットクローズ (60)	
	sequential	scrambled	sequential	scrambled
A (N=52)	$\bar{X}=26.66$ SD = 5.27	$\bar{X}=23.73$ SD = 6.61	$\bar{X}=15.35$ SD = 4.96	$\bar{X}=16.85$ SD = 5.21
B (N=52)	$\bar{X}=21.00$ SD = 4.91	$\bar{X}=21.23$ SD = 4.28	$\bar{X}=12.58$ SD = 4.89	$\bar{X}=12.12$ SD = 3.65
C (N=52)	$\bar{X}=13.96$ SD = 5.23	$\bar{X}=13.31$ SD = 4.85	$\bar{X}= 9.23$ SD = 4.65	$\bar{X}= 8.38$ SD = 3.52

図3 説明文クローズ得点プロフィール



相互作用は英語学力×親密度 [$F(2,300)=6.57$ $p<.05$] 以外見られない。チューキー法の post hoc の結果、4 種類の CT を込みにした場合、 $A > B > C$ のように群間の英語学力差が出たが、英語学力と親密度の交互作用があるために、この関係は親密度の高い CT に限られる。以上 2 つの異なるクローズを使った実験結果をまとめると次のようになる。まず、英語力の要因は双方のクローズで検証され、英語の学力の高い群は低い群よりも

表4 物語文クローズの3要因分散分析

要因	M.S.	df	F
A	2365.734	2	99.348 ^{***}
B	18.029	1	0.757
C	4410.029	1	185.199 ^{***}
A×B	2.664	2	0.112
A×C	156.529	2	6.573 [*]
B×C	23.157	1	0.973
A×B×C	54.253	2	2.278
E	23.812	300	
T		311	

A=英語学力 B=文脈 C=親密度

^{***}p<.001 ^{*}p<.05

CT得点が高くなる傾向にある。しかし、この傾向は親密度の低いクローズでは全くその効果がない。A群の野球クローズを除いて sequential と scrambled version 間の得点はほぼ重なり合った。また、有意ではないが、A群のクリケットクローズとB群の野球クローズに見られるように、文脈を与えられないCTの得点が高くなっている。

物語文クローズと説明文クローズの項目分析

受験者156名の成績を基礎に各CTの項目分析を試みた結果は Appendix 2 に示した通りである。各CTの総得点と各テスト項目との相関からみたテスト項目の判別力(ID)に関しては、まず、物語文の sequential version では、30項目(全体の65.2%)、同 scrambled version では22項目(47.8%)が総得点と有意な正の相関がある。一方、野球クローズの sequential には33項目(55.0%)、同 scrambled には37項目(61.7%)、クリケットクローズの sequential には25項目(41.7%)、同 scrambled には39項目(65.0%)、それぞれ有意な相関を示す項目が含まれている。また UL 弁別指数に基づく項目判別力(UL)の.40以上の項目を挙げると、物語文クローズの sequential では18項目(39.1%)、同 scrambled では8項目(19.0%)、野球クローズの sequential では15項目(25.0%)、同 scrambled では20項目(33.3%)、クリケットクローズの sequential では16項目(26.7%)、同 scrambled では10項目(16.7%)含まれる。IDとULから判断すると、物語文クローズ内においては sequential version が項目の判別力が高くなっ

ている。しかし、説明文クローズにおいては、むしろ、scrambled versionの方が判別力が高くなっている。次に各クローズの通過率（IF）に目をやると、両 version 間の通過率の差は物語文クローズが最も大きく全体の65.2%に当る30項目で sequential version の通過率が高くなっているが、野球クローズでは43.3%の26項目、クリケットクローズでは38.3%の23項目で、それぞれ文脈の影響が現れただけである。この version 間に見られる通過率の差は実験の分散分析の結果と一致している。

考 察

CTの構成概念的妥当性はテキストの種類と親密度に影響されるという予測の下で6種のCTを大学生に実施したが、結果はほぼ予想通りであった。つまり、テスト用に選択するテキストが異なると文脈の有無が得点に関与しないこと、受験者にとって身近な内容を伝えるテキストとそうでないテキストでは受験者の文脈利用法が異なることなどが判明した。CTの項目分析の結果からは、物語文の sequential version の項目判別力が高く、また、説明文クローズと比較して物語文クローズにおいて受験者の文脈利用力をより良く判断できる点が明らかになった。

しかし、CTは受験者の‘wide-ranging linguistic constraints’や‘discourse constraints’を活用する能力が測定できるかどうかという疑問に対しては本実験結果からは明確な答えが出ていない。何故ならば、Chihara *et al.* (1977) と Chavez-Oller *et al.* (1985) で示された英語学力と文脈との交互作用が検出できなかったからである。この問題を考える上で受験者の英語学力を再考する必要があると思われる。Chihara *et al.* では被験者となった日本人英語学習者がどの程度の学力集団なのかテストで正確に測定されておらず、ただYMCAに通う advanced から basic course の学習者であるただけ報告されている。従って本実験の受験者よりもかなり英語力に差のある集団が実験に参加していると想像できる。また、Chihara *et al.* には英語の母国語話者もテストに加わっている点も忘れてはならない。英語学力にかなりの違いのある受験者をサンプリングした場合、文脈と英語学力間に交互作用が出やすいことは前回の実験結果（西田 印刷中）からも肯定できよう。前回では本実験で使用した材料と同一の物語文を高校3年生、大学1年及び2年生に与えた結果、英語学力と文脈の主効果と両要因の交互作用が検出され、sequential version に限って高校3年<大学1年<大学2年のように英語学力差が得点に反映した。CTは受験者の

intersentential な能力が測定可能かあるいは sentence-bound なのかの論争に終止符を打つ為にはテキストに含まれる要因の洗い出しと受験者の英語学力の実態を正確に把握することが特に必要である。

結 論

CT が今日のように広く英語教育の分野で利用されている理由のひとつはテストの経済性にある。作成に時間が取られず信頼性と妥当性の高いテストを実施でき、その結果受験者の学力を測定できるとする一種妄信に近い考え方が広く行き渡っている。しかし、CT ほど構成概念的妥当性の定まらないテストは他にないように思われる。Oller 達の主張する global language proficiency 自体その構造は依然として解明されていない。CT は 'grenade' (Oller 1976 : 276) であるから広い範囲に影響力を発揮するが、'rifle' のように標的が定まらず、一体何を測定できるのか明確でない。しかし、同時に、grenade を rifle の尺度でもって、その妥当性を否定することも筋違いと言わねばならない。テスト法の武器には両者が必要であるからである。要は grenade として機能を十分果たすためには、今回試みたようなテキストに関する要因を追実験し、CT の構成概念的妥当性を高める方法を開発せねばならない。

〔注〕

- 1) 本稿は全国英語教育学会福岡研究大会 (昭和61年8月4・5日) において「『野球』クローズと『クリケット』クローズ」と題して口頭発表した内容に一部加筆修正したものである。
- 2) Cziko (1983) は CT の構成概念的妥当性を検討する上での重要な要因として、パッセージの種類、難易度、削除語のタイプ、テストの所要時間など6つの要因を挙げている。
- 3) 本実験で用いた物語文クローズには46項目中12項目 (26.1%)、野球クローズには60項目中4項目 (6.7%) クリケットクローズには5項目 (8.3%) の cohesive devices が含まれており、テキストの構成にはっきりした相異がある。
- 4) 親密度の5段階評定は、5 = 大変よく知っている 4 = よく知っている 3 = 知っている 2 = よくは知らない 1 = 全然知らないである。受験者156名の評定平均は野球クローズが3.6、クリケットクローズが1.2であった。

References

Alderson, J. Charles (1979) "Scoring Procedures for Use on Cloze Tests," in

- Yorio, Carlos A., Kyle Perkins, and Jacquelyn Schachter (eds.) *On TESOL '79: The Learner in Focus*. Washington, D.C.: TESOL, 193–205.
- _____ (1983) "The Cloze Procedure and Proficiency in English as a Foreign Language," in Oller, John W. Jr. (ed.) *Issues in Language Testing Research*. Rowley: Newbury House Publishers, Inc., 205–217.
- Bernhardt, Elizabeth Buchter (1984) "Toward an Information Processing Perspective in Foreign Language Reading," *MLJ*, 68, 4, 322–331.
- Bransford, John D., Barry S. Stein, and Tommie Shelton (1984) "Learning from the Perspective of the Comprehender," in Alderson, J. Charles and A.H. Urquhart (eds.) *Reading in a Foreign Language*. London: Longman, 28–47.
- Brown, James Dean (1983) "A Closer Look at Cloze: Validity and Reliability," in Oller, John W. Jr. (ed.) (1983), 237–250.
- 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.
- Cziko, Gary A. (1983) "Another Responses to Shanahan, Kamil, and Tobin: Further Reasons to Keep the Cloze Case Open," *RR Q*, 18, 3, 361–365.
- Johnson, Robert K. (1982) "Questioning Some Assumptions About Cloze Testing," in Heaton, J. B. (ed.) *Language Testing*. London: Modern English Publications Limited, 59–72.
- Mullen, Karen A. (1979) "More on Cloze Tests of Proficiency English as a Second Language," in Brière, Eugène J. and Frances Butler Hinofotis (eds.) *Concepts in Language Testing: Some Recent Studies*. Washington, D.C.: TESOL, 20–32.
- 西田 正 (1985) 「クローズテストにおけるテキストと削除変数について」『中国地区英語教育学会研究紀要』No. 15, 47–54.
- _____ (印刷中) 「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)
- Oller, John W. Jr. (1972) "Scoring Methods and Difficulty Levels for Cloze Tests of Proficiency in English as a Second Language," *MLJ*, 56, 3, 151–158.
- _____ (1973) "Cloze Tests of Second Language Proficiency and What They Measure," *LL*, 23, 1, 105–118.
- _____ (1975) "Cloze, Discourse, and Approximation to English", in Burt, Marina K. and Heidi C. Dulay (eds.) *On TESOL '75: New Directions in Second Language Learning, Teaching and Bilingual Education*. Washington, D.C.: TESOL, 345–355.

- _____ (1976) "Language Testing," in Wardhaugh, Ronald and H. Douglas Brown (eds.) *A Survey of Applied Linguistics*. Ann Arbor: The University of Michigan Press, 275-300.
- _____ (1979) *Language Tests at School*. London: Longman.
- _____ and Christine A. Conrad (1971) "The Cloze Technique and ESL Proficiency," *LL*, 21, 2, 183-195.
- 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*. London: Academic Press, 63-74.
- Rye, James (1984) "Cloze and Intersentential Constraint," *J. of Research in Reading*, 7, 2, 113-122.
- Shanahan, Timothy, Michael L. Kamil, and Aileen Webb Tobin (1982) "Cloze as a Measure of Intersentential Comprehension," *RR Q*, 17, 2, 229-255.
- Shohamy, Elana (1982) "Predicting Speaking Proficiency from Cloze Tests: Theoretical and Practical Considerations for Test Substitution," *AL*, 3, 2, 161-171.
- Wainman, H. (1979) "Cloze Testing of Second Language Learners," *ELTJ*, 33, 2, 126-132.

APPENDIX 1.

(紙面の都合上各 CT は sequential version を実例として挙げる)

- 1) 物語文クローズ No. _____ Name _____

After a long tiring walk, Tyll was hungry, so he stopped at an inn for dinner. He ordered beef and vegetables. Thirty long (1) MINUTES passed. There was now some life in (2) THE kitchen, it seemed, but there was still (3) NOTHING in front of him, not even a (4) NAPKIN. "The cow must be under the butcher's (5) KNIFE," he said for himself. He started humming (6) AN old tune and beating the table in (7) RHYTHM. "What's all that noise about?" cried the (8) INNKEEPER. "Oh, it's just my stomach that's complaining!" (9) SAID Tyll calmly. "Well, if you can't wait, (10) GO and find another place." said the man (11) ANGRILY.

Tyll stopped humming. Picking up his dusty (12) BAG, he took out a roll of bread (13) LEFT over from a meal at another inn (14) THE day before. Just to kill time, he (15) BEGAN to eat it. With this and the (16) FIRST smells of the meat and the other (17) FOODS from the kitchen, his stomach began feeling (18) BETTER. "See, this is a meal that won't (19) COST me much!" thought Tyll.

Another half hour (20) PASSED before the innkeeper began to serve

the (21) GUESTS . When at last Tyll's turn came, he (22) SET a place and a wine glass in (23) FRONT of him, but Tyll pushed them away. "(24) WHAT'S the matter?" asked the innkeeper. "I'm no (25) LONGER hungry. My stomach is full of the (26) SMELLS from the kitchen," answered Tyll. The innkeeper (27) GAVE Tyll an angry look and walked away.

(28) AS Tyll had a long way to go, (29) HE rested a little more. Then he stood (30) UP and made for the door. The innkeeper (31) STOPPED before him at the door and stuck (32) A wooden tray under his nose. "What's this (33) FOR ?" asked Tyll. "Those who sit at one (34) OF my tables for such a long time (35) MUST pay. And you said the smells from (36) THE kitchen were enough for you, didn't you? (37) NOW pay, or I warn you I'll call (38) THE police."

Tyll took out of his pocket (39) ALL the coins he had. As he dropped (40) THEM without haste one by one on the (41) WOODEN tray, he asked. "How do you like (42) THE sound?" "It's great!" came the answer. "Like (43) THE kitchen smells that filled my stomach, let (44) THE sound of these coins fill your pocketbook." (45) ADDED Tyll, and he quickly picked up all (46) THE coins. Then pushing the innkeeper aside, he opened the door and went away.

2) 野球クローズ

No. _____ Name _____

Baseball is a game played 1) WITH a bat and ball on 2) A large field, a section of 3) WHICH, the "diamond," contains four bases 4) WHICH form the corners of a 5) 90-FOOT square. It is played between 6) TWO teams of nine players each, 7) ONE of which takes defensive positions 8) IN the field while the other 9) TAKES turns batting. The ball is 10) THROWN or "pitched" from approximately the 11) MIDDLE of the diamond toward the 12) MOST important base, home plate.

The 13) BATTER, standing beside home plate, attempts 14) TO hit or bat it away 15) FROM defensive players so that he 16) CAN reach first base and eventually 17) COMPLETE the circuit of the bases 18) AND score a run by reaching 19) HOME plate. In this he may 20) BE assisted by succeeding batters.

Each 21) TEAM is given an equal number 22) OF times to bat, and the 23) GAME is divided into innings or, 24) MORE properly, half

innings. A half 25) INNING terminates when three members of 26) THE batting team are “out”, i.e., 27) HAVE fallen victims to defensive play 28) AS determined by the rules. A 29) NORMAL game consists of nine full 30) INNINGS and is won by the 31) TEAM which has scored the most 32) RUNS.

When a game starts the 33) HOME team takes the field. The 34) CATCHER, clad in protective mask, chest 35) PROTECTOR, and shinguards and wearing a 36) HEAVY padded glove, must stand in 37) THE triangular catcher’s box extending 8 38) FEET behind the plate. The pitcher, 39) WHO with the catcher comprises the “40) BATTERY,” must have one foot on 41) THE pitcher’s rubber when he pitches. 42) THE positions of other players are 43) NOT restricted, but commonly the first 44) BASEMAN stands near first base, the 45) THIRD baseman near third. The second 46) BASEMAN covers the first base side 47) OF second base, the shortstop the 48) THIRD base side of it.

With 49) THE first pitch the game begins. 50) IF the batter swings and misses 51) IT is a strike. It is 52) A strike if he fails to 53) SWING at a ball which passes 54) THE “strike zone” — the space directly 55) OVER the plate and between the 56) BATTER’S knees and armpits when he 57) TAKES his natural stance. An umpire, 58) IN chest protector and mask, stands 59) BEHIND the catcher and judges each 60) PITCH.

3) クリケットクローズ

No. _____ Name _____

Cricket is normally played between 1) TWO sides of 11 players each, 2) ONE of whom is captain. There 3) ARE two umpires to control the 4) GAME according to the laws, and 5) TWO scorers to record its progress.

6) AT the start of a match 7) THE captain who wins the toss 8) OF a coin decides whether his 9) OWN or the other side shall 10) TAKE first innings. All cricket matches 11) LASTING two or more days are 12) DECIDED by the aggregate of runs 13) MADE by the two sides in 14) THE two innings. One-day matches may 15) BE decided on a single-innings basis.

16) AN innings is completed when ten 17) BATSMEN have been dismissed. But the 18) CAPTAIN of the batting side may 19) DECLARE

his innings closed if he 20) THINKS his side have made enough 21) RUNS. Results are recorded by the 22) MARGIN of runs, or if the 23) SIDE batting last passes the other 24) SIDE'S total before all their batsmen 25) HAVE been dismissed, by the number 26) OF their wickets outstanding.

The ball 27) IS bowled from each wicket alternately 28) IN over of six balls. For 29) A fair delivery the ball must 30) BE bowled, not thrown or jerked. 31) AT the moment of delivery a 32) BOWLER must have some part of 33) ONE foot behind the bowling crease 34) AND within the return crease; the 35) BOWLER'S hand may be at any 36) HEIGHT. A breach of these conditions 37) CONSTITUTES a non-ball, called by the 38) UMPIRE and counting one run to 39) THE opposing side. Similarly a ball 40) WHICH passes out of reach of 41) THE striker constitutes a wide and 42) COUNTS one run. When six balls 43) HAVE been bowled, the umpire calls "44) OVER." And a new over is 45) BEGUN by a different bowler at 46) THE opposite end, with a corresponding 47) ADJUSTMENT of the field. Subject to 48) THE alternation of overs, any member 49) OF the fielding side may bowl 50) AT either end as many overs 51) AS his captain decides, but the 52) SAME bowler may not bowl two 53) CONSECUTIVE overs.

The object of the 54) BOWLER is primarily to get the 55) BATSMAN out, and only secondarily to 56) PREVENT him from getting runs. The 57) OBJECT of the batsman is to 58) MAKE runs, for only runs can 59) WIN a match, but to make 60) RUNS he must stay in.

APPENDIX 2. クローズテストの項目分析
1) 物語文クローズ

Line No.	sequential			scrambled			Line No.	sequential			scrambled			IF Diff
	ID	IF	UL	ID	IF	UL		ID	IF	UL	ID	IF	UL	
1	.34	.46	.48	.24	.32	.29	24	.22*	.91	.14	.17*	.82	-.19	+.09
2	.13*	.90	.09	.48	.68	.52	25	.06*	.14	.10	.10*	.14	-.04	±.00
3	.29	.60	.33	.24	.09	.14	26	.42	.72	.57	.17*	.49	.24	+.23
4	.31	.15	.19	.13*	.08	.05	27	.31	.12	.29	.21*	.10	.10	+.02
5	.18*	.13	.10	.23	.09	.14	28	.20*	.13	.19	-.02*	.21	.05	-.08
6	-.01*	.17	.00	.32	.21	.38	29	.44	.67	.57	.42	.60	.52	+.07
7	.20*	.05	.14	.37	.01	.05	30	.38	.71	.43	.19*	.76	.24	-.05
8	.44	.46	.62	.11*	.78	.05	31	.43	.55	.57	.40	.45	.57	+.10
9	.27	.73	.24	.15*	.87	.10	32	.22*	.06	.14	.15*	.05	.14	+.01
10	.28	.69	.29	.30	.69	.24	33	.34	.03	.10	-.00	.00	.00	+.03
11	.30	.31	.23	.33	.18	.29	34	.42	.31	.52	.39	.32	.48	-.01
12	.47	.24	.48	.29	.09	.19	35	.48	.24	.48	.39	.12	.29	+.12
13	.22*	.03	.10	-.00	.00	+.03	36	.23	.78	.19	.36	.74	.39	+.04
14	.28	.67	.24	.32	.64	.57	37	.20*	.10	.19	.00*	.08	.00	+.02
15	.24	.35	.33	.24	.33	.14	38	.41	.18	.42	.18*	.21	.19	-.03
16	.40	.53	.48	.34	.40	.33	39	.23	.42	.19	.44	.46	.57	-.04
17	.03*	.05	.05	.22*	.09	.14	40	.56	.22	.57	.10*	.01	.00	+.19
18	-.00	.00	.00	-.00	.00	+.00	41	.52	.64	.71	.20*	.33	.29	+.31
19	-.00	.00	.00	.15*	.01	.05	42	.05*	.08	.05	.10*	.08	.09	+.00
20	.37	.55	.57	.42	.42	.48	43	.54	.49	.71	.43	.41	.52	+.08
21	-.00	.00	.00	.01*	.03	.05	44	.58	.19	.57	.48	.12	.33	+.07
22	.03	.03	.00	.01	.05	+.02	45	.58	.50	.76	.20*	.71	.24	-.21
23	.47	.59	.66	.33	.50	.38	46	.32	.73	.38	.21*	.72	.29	+.01

* = not significant at $p < .05$

2) 説明文(野球) クロース

Time No.	sequential				scrambled				Time No.	sequential				scrambled				
	ID	IF	UL	UD	IF	UL	UD	IFD		ID	IF	UL	UD	ID	IF	UL	UD	IFD
1	.10*	.85	.28	.56	.67	.67	.48	+.18	31	.42	.55	.48	.29	.51	.19	+.04		
2	.25	.37	.33	.40	.43	.43	.01	-.03	32	.11*	.01	.05	.14*	.01	.05	±.00		
3	.28	.03	.10	.00*	.05	.05	.09*	-.02	33	.09*	.03	.05	.09*	.05	.10	-.02		
4	.42	.23	.38	.20*	.21	.29	.38	+.02	34	.38	.79	.33	.26	.94	.10	-.15		
5	—	.00	—	-.01*	.01	.00	.44	-.01	35	.39	.44	.38	.33	.72	.29	-.28		
6	-.15*	.92	-.14	.21*	.86	.24	.37	+.17	36	—	.00	—	-.01*	.01	.00	-.01		
7	.17*	.82	.28	.30	.65	.48	.33	+.09	37	.38	.40	.57	.52	.40	.67	±.00		
8	.18*	.24	.19	.27	.33	.33	.05*	-.09	38	.05*	.13	.24	.17*	.18	.24	-.05		
9	.21*	.04	.05	.20*	.06	.10	.06	-.02	39	.38	.15	.33	.08*	.19	.05	-.04		
10	.26	.40	.29	.52	.32	.71	.19	+.08	40	.19*	.12	.14	-.02*	.21	-.05	-.09		
11	.11*	.06	.09	.32	.06	.14	.00	±.00	41	.55	.46	.71	.59	.51	.67	-.05		
12	.46	.69	.62	.24	.67	.29	.29	+.02	42	.21*	.23	.24	.26	.21	.24	+.02		
13	.41	.81	.38	.36	.90	.29	.38	-.09	43	.38	.21	.33	.47	.24	.48	-.03		
14	.54	.59	.90	.53	.65	.71	.06	-.06	44	.39	.73	.29	.45	.56	.48	+.17		
15	.04*	.23	.04	.05*	.29	.05	.29	-.06	45	.40	.86	.38	.41	.68	.38	+.18		
16	.46	.50	.67	.26	.44	.33	.06	±.06	46	.62	.74	.48	.45	.31	.48	+.43		
17	.22*	.01	.05	.16*	.01	.05	.00	-.01	47	.43	.28	.48	.32	.33	.43	-.05		
18	.18*	.04	.09	.49	.05	.14	.01	-.01	48	.48	.35	.62	.72	.33	.52	+.02		
19	.17*	.88	.19	.41	.81	.38	.07	+.07	49	.17*	.51	.48	.29	.53	.48	-.02		
20	-.26	.67	.48	.40	.59	.43	.08	+.08	50	.28	.18	.29	.33	.19	.19	-.01		
21	.33	.47	.43	.35	.33	.48	.14	+.14	51	.30	.28	.29	.22*	.18	.24	+.10		
22	.28	.33	.38	.47	.33	.62	.00	±.00	52	.31	.29	.33	.21*	.28	.32	+.01		
23	.45	.31	.40	.33	.17	.33	.14	+.14	53	.25	.15	.24	-.13*	.03	.00	+.12		
24	.28	.09	.19	.29	.04	.10	.07	+.07	54	-.00*	.06	.10	-.10*	.09	.05	-.03		
25	.14*	.31	.29	.42	.24	.48	.05	+.05	55	.30	.09	.19	.26	.15	.19	-.06		
26	.46	.49	.62	.59	.50	.81	.01	-.01	56	.32	.51	.19	.40	.54	.52	-.03		
27	.22*	.03	.10	-.10*	.05	-.05	-.02	-.02	57	.19*	.09	.14	.10*	.10	.05	-.01		
28	.18*	.03	.10	—	.00	—	.03	+.03	58	.11*	.05	.00	-.10*	.04	-.05	+.01		
29	.22*	.01	.05	.10*	.03	.05	.02	+.02	59	.34	.73	.29	.24	.81	.19	-.08		
30	.50	.58	.48	.36	.35	.47	.23	+.23	60	.12*	.05	.14	.09*	.01	.00	+.04		

* = not significant at $p < .05$

3) 説明文(クリケット)クローズ

line no.	sequential				scrambled				line no.	sequential				scrambled			
	ID	IF	UL	IFDiff	ID	IF	UL	IFDiff		ID	IF	UL	IFDiff	ID	IF	UL	IFDiff
1	.41	.65	.48	.68	.47	.88	.71	-.03	31	.34	.44	.43	.30	.45	.43	-.01	
2	.00*	.74	.05	.37	.56	.10	+.18	.32	—	.00	.00	.00	.15*	.01	.05	-.01	
3	.16*	.97	.10	.28	.96	.10	+.01	.33	.05*	.08	-.05	.00*	.05	.05	-.05	+.03	
4	.37	.55	.43	.28	.35	.33	+.20	.34	.10*	.01	.00	.32	.01	.00	.05	±.00	
5	.14*	.32	.33	.27	.38	.33	-.02	.35	.10*	.01	.05	—	—	.00	—	+.01	
6	.36	.42	.23	.34	.44	.38	-.06	.36	—	.00	—	—	—	.00	—	±.00	
7	.59	.44	.81	.59	.44	.76	±.00	.37	.07*	.01	.05	—	.00	—	—	+.01	
8	.25	.15	.20	.36	.23	.29	-.08	.38	.45	.19	.33	.30	.18	.33	.30	+.01	
9	—	.00	—	—	.00	—	±.00	.39	.56	.45	.71	.32	.40	.52	.40	+.05	
10	.15*	.08	.14	.17*	.13	.05	-.05	.40	.32	.13	.33	.50	.05	.19	.19	+.06	
11	—	.00	.00	—	.00	—	±.00	.41	.50	.27	.52	.51	.23	.62	.62	+.04	
12	.16*	.04	.10	.00*	.04	.05	±.00	.42	.29	.01	.05	—	.00	—	—	+.01	
13	—	.00	—	.41	.03	.10	-.03	.43	.49	.74	.52	.35	.73	.38	.38	+.01	
14	.50	.22	.62	.08*	.23	.05	-.01	.44	.14*	.12	.19	.19*	.08	.10	.10	+.04	
15	.50	.45	.62	.26	.60	.29	-.15	.45	.21*	.05	.10	.27	.03	.10	.10	+.02	
16	.13*	.24	.10	.01*	.37	-.05	-.13	.46	.63	.47	.86	.53	.45	.67	.67	+.02	
17	.11*	.03	.00	.33	.24	.38	-.21	.47	—	.00	—	—	.00	—	—	±.00	
18	.08*	.09	.05	.28	.06	.19	+.03	.48	.48	.13	.24	.15*	.18	.33	.33	-.05	
19	—	.00	—	—	.00	—	±.00	.49	.44	.50	.71	.47	.50	.67	.67	±.00	
20	.21*	.05	.10	.37	.00	.14	+.01	.50	—	.00	—	.17*	.04	.10	.10	-.04	
21	.08*	.01	.00	.40	.05	.19	-.04	.51	.27	.21	.62	.39	.12	.33	.33	+.09	
22	—	.00	—	—	.00	—	±.00	.52	—	.00	—	.26	.03	.10	.10	-.03	
23	.16*	.01	.05	.29	.03	.10	-.02	.53	—	.00	—	.26	.01	.05	.05	-.01	
24	.35	.05	.14	.40	.06	.19	-.01	.54	.32	.03	.05	—	.00	—	—	+.03	
25	.51	.53	.67	.39	.54	.52	-.01	.55	.07*	.01	.00	.48	.06	.14	.14	-.05	
26	.52	.83	.52	.46	.69	.52	+.14	.56	.42	.08	.29	.45	.15	.29	.29	-.07	
27	.40	.41	.52	.26	.44	.38	-.03	.57	.14*	.08	.19	.34	.10	.19	.19	-.02	
28	.14*	.03	.05	—	.00	—	+.03	.58	.17*	.08	.14	.36	.14	.33	.33	-.06	
29	.16*	.13	.14	.24	.03	.05	+.10	.59	.09*	.13	.10	.14*	.10	.05	.10	+.03	
30	.57	.67	.71	.40	.67	.38	±.00	.60	.20*	.05	.10	.27	.06	.19	.19	+.01	

* = not significant at p < .05

Narrative Cloze versus Expository Cloze

Tadashi NISHIDA

Recent research studies on cloze tests have focused on the question of construct validity. Oller (1973) (1976) argues that cloze tests measure some sort of global language proficiency. Chihara *et al.* (1977) and Chavez-Oller (1985) maintain that cloze items are sensitive to constraints across sentences. Alderson (1983) and Porter (1983), however, provide some evidence to suggest that cloze is essentially sentence-bound. These conflicting conclusions seem to have failed to take into account important factors in the test-making process.

The study reported here was designed to examine whether construct validity is affected by the type of passage used and familiarity with the topic. One narrative passage with 395 words in length was taken from a short story. Two expository passages of approximately 360 words each were taken from two encyclopedias. One of these was on a familiar topic concerning a baseball game and the other described a cricket game. Every 8th word was deleted from the narrative passage and every 6th word from the expository passages. The sentences of these cloze tests were scrambled to produce one more test over each passage. Six cloze tests were administered to 156 Japanese undergraduate students who consisted of three different English proficiency groups.

To determine whether there were significant differences between the mean scores for the narrative cloze tests across the groups, a two-way ANOVA was performed. Two main effects were significant at $p < .001$; context (sequential vs. scrambled), $F(1,150) = 9.43$; level of English proficiency, $F(2,150) = 23.55$. However, the interaction was not significant. A three-way ANOVA was also employed for the scores of the expository cloze tests. The main effects for level of English proficiency, $F(2,300) =$

99.35, and familiarity (baseball vs. cricket), $F(1,300) = 185.20$, were significant at $p < .001$ but the context effect was found not significant. No interactions were found significant except that between familiarity and level of proficiency, $F(2,300) = 6.57$. Item analysis performed on each cloze test by means of item discrimination and facility indices showed that the sequential version of the narrative cloze test included more items whose scores were consistent with the total ones.

The conclusion to be drawn from these experimental results is that different passages selected for cloze tests and different degree of the topic familiarity have effects on the construct validity which cannot be neglected.