

テストメソッド効果の考察*

小林美代子

神田外語大学

言語能力の評価では、測定しようとしている能力を正確に測定する必要があるが、そのためには、言語能力以外の要素の介在を認識し、可能な限りその影響を抑えることが望ましい。本論は、Bachman (1990) の枠組みに基づき、英語読解力テストに用いられる文章の内容構成とテスト形式という2変数が、日本人学習者のテスト結果に及ぼす影響を考察する。日本人大学生総計754名を無作為に12のグループに分け、4種類の文章構成及び3種類のテスト形式のうち、各々一つを割り当て、読解力テストに解答してもらった。分散分析の結果、文章の内容構成とテスト形式には相互作用があることが判明し、その影響は統計的に有意であった。さらに、学生を英語能力別に3つのグループに分けて、それぞれの変数の影響及び相互作用を検証した結果、学習者の英語能力も、規則的な影響を与えることが判明した。最後に、日本の大学入学試験英語読解テストに使用された文章の内容構成を考察した追加研究の結果を報告する。

1. はじめに

本研究の主要な目的は、言語能力以外の要素が読解力テストの成績に及ぼす影響を検証することである。この研究では、文章の内容構成とテスト形式の2つの要素を検証する。さらに、検証結果を踏まえて、言語テストの実情を考察するために、日本の大学入学試験の読解力テストに用いられた文章の構成を調査する。

* 本稿は Kobayashi (2002a) 及び Kobayashi (2005) を基に、邦訳、加筆修正したものである。また、早稲田大学言語学勉強会定例会 (2004年7月) 及び JACET (大学英語教育学会) 月例研究会 (2004年7月) でも、同様の口頭発表を行った。

本研究の理論的枠組みは、ひとつには、Bachman (1990)の言語能力モデルとそれに付随するテストメソッド要素のモデル、そしてもうひとつは Meyer (1975, 1985)の文章の内容構成分析のモデルである。さらに、Kintsch & Yarbrough (1982)の研究を、テスト形式という変数を具現化する際に参照した。Bachman (1990)及びその発展である Bachman & Palmer (1996)は言語能力のモデルを提示しており、その議論の中で、テストの成績に影響を与える要素として、「テストメソッド効果」というものを取り上げ、その重要性に着目して、次のように述べている。

言語使用は、言語使用者の個々の様々な特性の間の複雑かつ複層的な相互作用を包含するだけでなく、これらの特性と言語使用及びテスト実施の状況の特性との間の相互作用をも包含するものである。この相互作用の複雑さ故に、言語能力というものは、言語使用の相互作用的枠組みの中で考えていかなければならないものと我々は考える。(Bachman & Palmer 1996: 62)

Bachman はテストメソッド効果を、1) テスト環境、2) 指示、3) インプットの特質、4) 回答の特質、5) インプットと回答間の関係、の5つに分類している。Bachmanによると、テスト作成者は、これらの言語能力以外の要素がテスト結果に及ぼしうる影響を認識し、可能な限りその影響を抑えることが肝要であるとしている。本研究は、これら5つの要素のうち、3番目と4番目、すなわち、インプットの特質(受験者に与えられる素材の特質)と回答の特質を、検証の主たる対象とする。というのも、読解力テストにおいては、読ませる文章というものは重要な役目を担っていると考えられるからである。例えば、文章の内容に対して、読み手が持っている背景的知識が読解に及ぼす影響については、広く検証されている(Alderson & Urquhart 1983, 1985a, 1985b; Bernhardt 1991; Carrell & Eisterhold 1983; Clapham 1996; Mohammed & Swales 1984; Salager-Meyer 1991; Steffensen & Joag-Dev 1984; Steffensen et al. 1979 等)。文章の特徴やリーダビリティに関する先行研究(Beck 及び共同研究者による一連の研究 1982, 1984, 1989, 1991, 1995; Britton et al. 1989; Davison & Kantor 1982; Duffy & Kabance 1982; Duffy et al. 1989; Graves et al. 1988, 1991; Klare 1985; Olsen & Johnson 1989; Reder & Anderson 1980; Urquhart,

1984 等) を幅広く考察した結果、本研究は、文章の特質の中でも、文章の内容構成に着目することとした。

2. 文章の内容構成

文章の内容構成という概念を、実際にどのような形で測定可能なものにするかを決定するために、さらに先行研究を探索していく中で、Meyer (1975, 1985) の内容構成分析モデルというものに遭遇した。「意味的内容構造分析(semantic content structure analysis)」と呼ばれるこの分析体系によると、各概念の単位は他の概念単位との関連において、ある一定の役割を担っているとされ、同時に各々の概念間の修辞上の関係に基づいて、概念の各単位は階層構造を成して配列されるとしている。この階層構造の最も上に現れる修辞上の関係は最上層修辞構造と呼ばれ、これが文章全体の特徴を決定する。最上層修辞構造には、「集合(collection)」「因果関係(causation)」「応答(response)」「描写(description)」「比較(comparison)」の 5 つがあり、そのうちのひとつとして認識される。(Meyer は後に「応答(response)」を「問題解決(problem-solution)」と命名し直している。) これらの 5 つのタイプの最上層修辞構造が我々の思考様式を表していると考えられる(Meyer 1985: 20)。

これら 5 種類の修辞構造のうち、「集合」の関係では、各概念間のつながりが最も弱く、ある共通のトピックに関して、ゆるやかな関連性があるに過ぎない。「時間の経過」は「集合」関係の一形態であり、出来事が時間の経過に沿って配列されている。「因果関係」では、概念は時間の経過に沿って描写されるだけでなく、ひとつの概念と別の概念が因果関係によって関連づけられている。最後に、「応答(問題解決)」の関係では、因果関係に応答する形で、ある解決が示されるため、各概念間の相互関係はさらに密なものになると考えられる。「比較」と「描写」は、上述の 3 つの修辞構造とは異なり、概念が階層構造的な関係を持つとされる。「描写」の構造では、「ある概念がもう一つの概念の上位関係にあり、そのもう一つの概念は、前の上位概念を修飾説明する (Meyer 1985: 20)」。

「比較」の構造では、比較対象の要素によって関連する階層的概念が二つ以上存在する。ということは、「比較」構造の方が、「描写」構造よりも、概念間の相互関係が複雑なものとなる。

一連の実証研究(Meyer & Freedle 1984; Meyer et al. 1980, 1993 等)により、Meyer は、かっちりした構成の文章の方が、概念間の関係が蜜であるため、文章内の概念が記憶に残りやすいと主張している。Meyer の主張は、他の研究 (Carrell 1984; Goh 1990; McGee 1982; Richgels et al. 1987 等) によっても、追認されている。本研究は、これらの先行研究の結果が第二言語としての英語の読解力にも適用されるかどうかを検証しようとするものである。予備調査の結果を踏まえて、「連想 (Meyer の「集合」を応用したもの)」「描写」「因果関係」「問題解決」の4種類の文章構成を検証する。

3. テスト形式

本研究では、もう一つの変数として、テスト形式を検証する。Meyer と共同研究者たちは、学習者が文章を読んだあとに、理解したことを思い出させる「想起」という手法で読解を測定しているが、「想起」は第二言語の読解力評価では、ほとんど見かけることのない測定形式である。より一般的なテスト手法を用いた方が、実際の読解力評価に対する示唆の汎用性が高いと考えられる。テスト形式がテスト結果に及ぼす影響についてはかなり研究がされている (Graesser et al. 1980; Graves et al. 1991; Kintsch & Yarbrough 1982; Lewkowicz 1983; Reder & Anderson 1980; Shohamy 1984; Shohamy & Inbar 1991 等)。その中でも、Kintsch & Yarbrough(1982)は、自由記述式問題とクローズテストという二つの手法が読解力に及ぼす影響を検証し、その結果、自由記述式の方が、文章の主要な概念を理解する能力を測定するのに優れているのに対し、クローズテストは、局所的な理解に触れるのみで、大局的な理解は測定し難いことを実証している。Kintsch & Yarbrough のこの知見は、文章の内容構成を研究の主眼とする本研究にとって、とりわけ有用であると判断し、

この手法を応用し、自由記述式とクローズテストを検証することとした。

予備調査と先行研究 (Bensoussan & Kreidler 1990 等) のさらなる考察の後、要約問題というテスト形式の方が、自由記述問題よりも、文章内容の全体構成の理解をさらに効果的に反映するものと判断し、本研究では、上の二つのテスト形式に加え、要約問題も採用することとした。自由記述式問題の場合、問題の内容によって、様々な種類やレベルの理解を測定することになり、必ずしも大局的な理解を測定するとは限らない。例えば、ある設問が局所的な情報についての理解のみで解答可能であった場合、その設問では、読み手が文章全体を理解したかどうかを測定することは不可能である。たとえ、設問が大局的な理解を要するようなものである場合でも、設問自体が文章の主要部分を示唆してしまう可能性もある。その点、文章を要約する際には、読者自身が、主要な概念と付随的な概念とを識別し、文章の全体的な構造を掌握しなければならず、そのため、文章の内容構造が大きく関与することが考えられる。

文章構造とテスト形式という2つの主要な変数に加え、予備調査の結果を鑑みて、本研究では、第3番目の変数として、学習者の言語能力レベルの影響も検証することとした。そのために、研究協力者は、英語能力テストの結果をもとに、3グループに分けられた(詳細は下記を参照)。

4. 研究の目的

本研究の質問は以下の通りである。

1. 読解力テストに用いられる文章の内容構成は読解力テストの結果に影響を及ぼすか。
2. テスト形式は読解力テストの結果に影響を及ぼすか。
3. 学習者の言語能力レベルは、文章構成及びテスト形式と相互作用があるか。

よって、以下の帰無仮説を設定した。

H01: 読解力テストに用いられる文章の内容構成 またはテスト形式と読解力テストの結果には何の相互関係もない。

H02: 読解力テストに用いられる文章の内容構成 またはテスト形式、学習者の言語能力レベル及び読解力テストの結果には何の相互関係もない。

5. 研究の方法

5.1 研究協力者

日本各地の 8 大学に在籍する総計 754 名の大学生 1 年生と 2 年生に研究協力してもらった。協力者の英語能力は、日本の大学生の平均的英語能力レベルということから、初中級から中級程度と考えられる。研究協力者は無作為に 12 のグループに分けられ、上述の 2 変数 (4 種類の文章構成と 3 種類のテスト形式) を検証するために勘案された 12 種類の英語読解力テスト (詳細は下記参照) のうちの 1 つを割り当てられた。

5.2 テスト材料

次の 2 つのテストを実施した。

1. 英語能力テスト (主に文法・語彙知識を測定するテスト)

このテストの目的は次の 2 つである。1 つは、研究協力者の 12 のグループの互換性を確認すること、2 つ目は言語能力レベル別の比較をするために、研究協力者を 3 グループに分ける基準を設定することである。このテストの記述統計は、平均 (50 点満点中) 29.7、標準偏差 8.07、信頼係数 (α) .82 であった。

2. 読解力テスト

読解力テストに用いた文章は、予備調査で特定された阻害要因を排除するため、本研究用に特別に準備した。背景的知識の関与も考慮した後、応用言語学の修士号以上の学位を有する 27 名の識者の判断を仰ぎ、先に挙げた 4 つの文章構成タイプ (連想、描写、因果関係、問題解決) を代表すると思われるものとして「国際援助」と「海運の安全」に関する 8 つの文章 (2 つのトピックに各々 4 種類の内容構成) を選定した。文章の平均語数は 369.3 語、リーダビリティの平均はフレッシュ指数 64.4 であった。文章選定後、各々の文章に対して、3 つのテスト形式 (自由記述、

クローズテスト、要約) でテスト問題を作成した。テストの項目数は、クローズテストが 25 項目、自由記述式が 5 項目、要約問題が 10 項目であった。自由記述問題と要約問題は、読解力以外の能力 (つまり書く能力) の関与を避けるため、研究協力者の母語である日本語で解答させた。

4 種類の文章構成タイプに対する設問が、できる限り等質となるよう最大限の努力を払った。たとえば、クローズテストでは、予備調査の結果と補充すべき単語の詳細な分析をもとに、空所の間隔と空所の始まる位置を決定した。また、応用言語学の専門家の意見も活用した。

5.3 手順

変数の影響を検証するためには、すべての学生がすべてのテストに解答すれば厳密な比較が可能となるはずであるが、それには 2 つの観点から問題が生じる。一つは、実施上の困難さであり、もう一つは、研究の妥当性の問題である。本研究で使用するテスト数は 24 で (2 つのトピックに対して、4 種類の内容構成タイプを代表する文章 4 つずつ、各々に 3 つのテスト形式)、全部を実施するには膨大な時間を要する。また、3 つの異なるテスト形式に答えるためには、同じ文章を 3 度読むことになり、提示する順番を入れ替えて統制したとしても問題が残る。そのため、本研究では、各学生に、24 のテストのうち、内容構成タイプとテスト形式の 12 の組合せのうち 1 つのみ、異なるトピックの文章を 1 つずつ、計 2 つのテストを与えることとした。その結果、学生たちを内容構成とテスト形式の異なる 12 のグループに無作為に分けることとなった。分散分析の結果、英語能力テストの成績において 12 グループ間に有意な差はないことが立証され ($F(11, 723) = 0.39, n.s.$)、英語の能力において、ほぼ等質とみなすことができることが確認された。

下の表 1 は 12 グループの構成を要約したものである。たとえば、グループ 2 は「因果関係」型の文章に基づくクローズテストに解答し、グループ 9 は「連想」型の文章を要約することがわかる。

表 1 研究協力者 12 グループの構成

		文章構成			
		連想	因果関係	描写	問題解決
テスト形式	クローズ	グループ 1 (N=63)	グループ 2 (N=61)	グループ 3 (N=66)	グループ 4 (N=65)
	自由記述	グループ 5 (N=59)	グループ 6 (N=57)	グループ 7 (N=54)	グループ 8 (N=57)
	要約	グループ 9 (N=66)	グループ 10 (N=63)	グループ 11 (N=62)	グループ 12 (N=62)

5.4 専門家による判断

本研究中、文章の選定、テスト項目の作成・分析、採点者間の信頼性確認等のため、応用言語学に精通した専門家の協力を仰いだ。たとえば、異なる文章間のテスト項目の一貫性を確保するために詳細な項目分析を行った。

5.5 データの分析

統計用コンピュータプログラム SPSS/PC を用いて、データ分析を行った。記述統計の他に、分散分析を実施し、仮説の検証を行った。有意水準は $p < .05$ とした。また、採点者間の信頼性を確認するため、自由記述問題と要約問題では、全体の 15% の解答を 3 人の経験ある採点者が個別に採点を行い、相関係数を計算した結果、.85 から .92 という係数が得られ、満足のいく水準と判断した。

6. 結果の考察

下の図 1 は、4 種類の内容構成と 3 つのテスト形式による読解力テストの結果を表したものである。

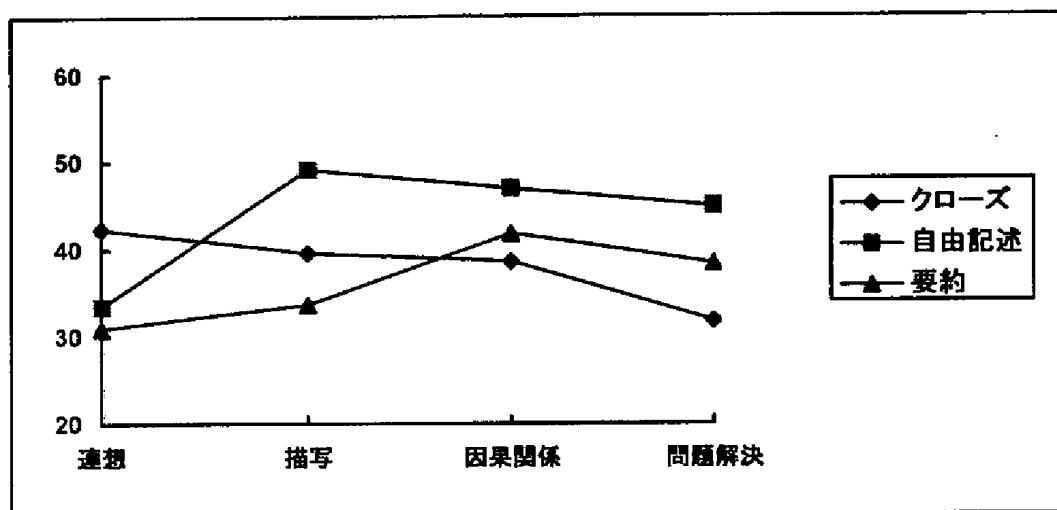


図1: 読解力テスト結果 (%換算)

図1によると、クローズテストの結果は「連想」型の文章構成の場合、平均点が高く、「問題解決」型の場合、低いことがわかる。つまり、クローズテストは、構成が甘い文章を用いると得点が高くなり、構成がしっかりした文章を用いると低くなる傾向がある。それに対し、自由記述式及び要約問題の場合には、「連想」型の文章を用いた場合、最も得点が低い。自由記述式問題では「描写」型の文章で最も得点が高いが、さらに構成がきっちりしている「因果関係」及び「問題解決」型においても高い得点となっている。要約問題では「因果関係」型で最も得点が高いが、「問題解決」型でも高い得点となっている。分散分析の結果、この差は有意であった(補遺参照)。また、それぞれの変数が独立して影響を与えるだけでなく、2つの変数間には相互作用があり、これも統計的に有意($F(6, 723) = 6.149^{**}, p < .005$)であった。このことから、読解力テストの成績は使用する文章の内容構成とテスト形式の影響を受けることが判明したと言える。クローズテストにおいて、明確な文章構成が高得点に結びつかなかった点は、興味深い。この点に関しては先行研究がないため、その理由の解明には今後の研究が待たれるが、文章中の情報の密度が関与していることも考えられる。「問題解決」型のような文章の場合、概念間の関連を蜜にして議論を展開する必要があるため、限られたスペースの中により多くの異なる概念が圧縮され、語句の繰返しが少ない可能性がある(Kintsch & Keenan 1973 参照)。単語の出現頻度はクローズテスト

の得点に大きく関与することが実証されている(Kobayashi 2002b 参照)。今後の更なる研究結果を待ちたい。

学習者の言語能力という観点から結果を考察すると、さらに興味深い結果が得られた。次の図 2-4 は 3 つの言語能力レベル別グループの得点を示したものである。

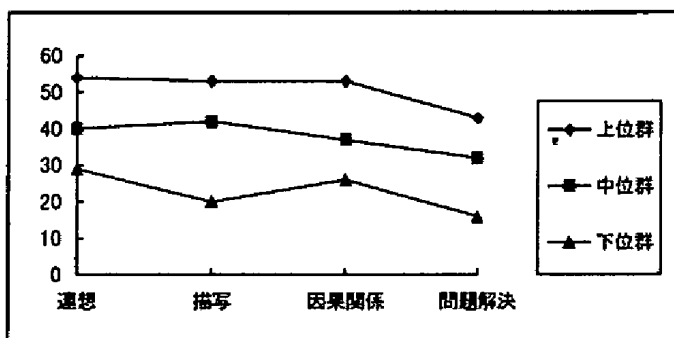


図 2: クローズテストの結果 (%換算)

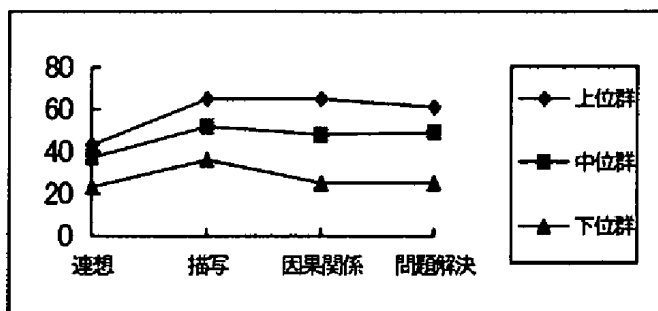


図 3: 自由記述式問題の結果 (%換算)

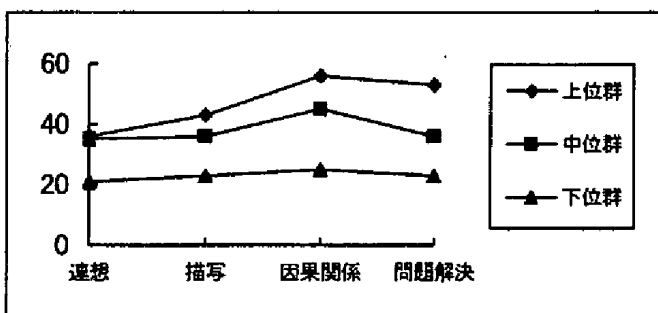


図 4: 要約問題の結果 (%換算)

概観して、文章の内容構成の影響は上位群の学習者において、より顕著であることがわかる。特に、自由記述問題と要約問題において、その傾向が顕著である。その反面、下位群の学習者の場合には、テスト形式

及び文章の構成、いずれの変数にもさほど影響を受けていないことがわかる。この差異は、統計的に有意であった（補遺参照）。この結果から、文章の構成とテスト形式という2変数の影響が学習者の言語能力レベルによって異なることが判明した。とりわけ、自由記述式と要約問題の2つのテスト形式の場合、内容構成の甘い文章を用いると、学習者の言語能力を正確に反映せず、特に上位群の学習者が本来の能力を発揮することができず、不利になる恐れがあることを示唆する。

7. 結論

本研究は、Bachmanの言語能力とテストメソッド効果のモデルを理論的枠組みとしたが、本研究で考察した2つの要因が学習者の読解力テストの得点に規則的な影響を与えることが判明し、Bachmanの提唱するテストメソッド効果の一部を実証することができた。ここで判明した結果をさらに確証するために、今後の研究が待たれるが、本研究が言語テスト及び第二言語習得研究に与える示唆は明確であろう。

言語テストの結果の多くは、入学試験・資格試験等、人生のさまざまな節目で、重要な判断の基準として利用される。その重要性を考えると、言語テストは、測定しようとしている言語能力を正確に測定する必要がある。そのためには、テスト作成者は、本研究で検証したような言語能力以外の要素の介在を認識し、可能な限りその影響を抑え、より妥当性・信頼性の高いテストをめざすことが望まれる。

8. 追研究

本研究の追調査として、日本の大学入学試験の読解力テストに用いられた文章を検証するため、過去7年間にわたる大学入試センター試験の読解問題用文章を分析した。本試験と追試験の両方から大問4と大問6に用いられた文章を抽出し、各年4つずつ、総計で28の文章の内容構成を分析した。そのうち半数は、主として、道徳的教訓の入った心温まる物語文で、時間の経過という緩やかな構成以外、明確な文章構成を持た

ないものである。残りの半分は、表やグラフを伴うものであるが、分析の結果、その大半は「描写」型の構成で、「因果関係」や「問題解決」を主たる内容構造に持つ文章は皆無であった。本論で上述したとおり、「描写」型文章構成は、「因果関係」や「問題解決」に比べ、構造がかっちりしていない。センター試験で用いられた実際の読解用の文章の大半が、このように構成の甘いものであると、受験者の能力を厳密に弁別することが難しく、特に、能力の高い受験者が不利になることが考えられる。

さらに、センター試験の読解問題の設問タイプにも問題が見られた。Kobayashi (1995, 2004) によると、局所的な理解を求める設問や推測的理解を要する設問は受験者間の弁別度が低いことが判明しているが、センター試験の大多数の設問が、このような設問タイプであり、早急の改善が求められる。また、文章の内容ではなく、表やグラフの情報を理解するだけで解答できる設問が多々見受けられたが、これでは、読解問題とは言えない。

もちろん、この試行的分析結果を一般化することには慎重でなければならないが、少なくとも、この分析でわかった範囲内で言えることは、センター試験の読解力テストに用いる文章の選定及び内容理解問題の作成には改善の余地がありそうである。センター試験は、その結果の持つ意味が大きく、多くの受験生の人生を左右するような責任の重い試験である。それ故、できる限り、理想に近いテストを作成することが望まれる。センター試験を将来改善する際に配慮すべき点をこの分析は提言していると思われる。

近年、妥当性研究の中でも「結果的妥当性(Consequential validity)」という概念が注目されるようになってきているが、これは、テストが教育ひいては社会に及ぼす影響を考慮し、それが望ましいものとなっているか確認することを妥当性検証の大切な要素とするものである。テスト作成者及び利用者は、テストが言語教育現場及び社会全体に及ぼす影響を認識して、妥当なテストを作成し、目的に合ったテスト結果の解釈及び利用をする責任があることを忘れてはならない。

参考文献

- Alderson, J.C. & Urquhart, A.H. (1983). The effect of student background discipline on comprehension: a pilot study. In A. Hughes, & D. Porter, (Eds.) *Current Developments in Language Testing*, (pp.121-127). London: Academic Press.
- Alderson, J.C. & Urquhart, A.H. (1985a). This test is unfair: I'm not an economist. In P.L.Carrell, J.Devine, & D.Eskey (Eds.), *Interactive Approaches to Second Language Reading* (pp.168-182). Cambridge: Cambridge University Press.
- Alderson, J.C. & Urquhart, A.H. (1985b). The effect of students' academic discipline on their performance on ESP reading tests. *Language Testing*, 2, 192-204.
- Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L.F. & Palmer, A.S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Beck, I.L., McKeown, M.G. & Gromoll, E.W. (1989). Learning from social studies texts. *Cognition and Instruction*, 6, 99-158.
- Beck, I.L., McKeown, M.G., Omanson, R.C. & Pople, M. (1984). Improving the comprehensibility of stories: The effects of revisions that improve coherence. *Reading Research Quarterly*, 19, 263-277.
- Beck, I.L., McKeown, M.G., Sinatra, G.M., & Loxterman, J.A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly*, 26, 251-276.
- Beck, I.L., McKeown, M.G. & Worthy, J. (1995). Giving a text voice can improve students' understanding. *Reading Research Quarterly*, 30, 220-238.
- Beck, I.L., Omanson, R.C. & McKeown, M.G. (1982). An instructional redesign of reading lessons: Effects on comprehension. *Reading Research Quarterly*, 17, 462-481.
- Bensoussan, M. & Kreidler, I. (1990). Improving advanced reading comprehension in a foreign language: summaries vs. short-answer questions. *Journal of Research in Reading*, 13, 55-68.

- Bernhardt, F.B. (1991). *Reading Development in Second Language: Theoretical, Empirical and Classroom Perspectives*. New Jersey: Ablex Publishing Corporation.
- Britton, B.K., Van Dusen, L., Gulgoz, S. & Glynn, S.M. (1989). Instructional texts rewritten by five expert teams: Revisions and retention improvements. *Journal of Educational Psychology*, 81, 226-239.
- Carrell, P.L. (1984). The effects of rhetorical organization on ESL readers. *TESOL Quarterly*, 18, 441-469.
- Carrell, P.L. & Eisterhold, J.C. (1983). Schema theory and ESL reading pedagogy. *TESOL Quarterly*, 17, 553-573.
- Clapham, C. (1996). *The Development of IELTS: A study of the effect of background knowledge on reading comprehension*. Cambridge: Cambridge University Press.
- Davison, A. & Kantor, R. (1982). On the failure of readability formulas to define readable texts: A case study from adaptations. *Reading Research Quarterly*, 17, 187-209.
- Duffy, T.M. & Kabance, P. (1982). Testing a readable writing approach to text revision. *Journal of Educational Psychology*, 74, 733-748.
- Duffy, T.M., Higgins, L., Mehlenbacher, B., Cochran, C., Wallace, D., Hill, C., Haugen, D., McCaffrey, M., Burnett, R., Sloane, S., & Smith, S. (1989). Models for the design of instructional text. *Reading Research Quarterly*, 24, 434-457.
- Goh, S. T. (1990). The effects of rhetorical organization in expository prose on ESL readers in Singapore. *RELC Journal*, 21, 1-13.
- Graesser, A.C., Hoffman, N.L. & Clark, L.F. (1980). Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior*, 19, 135-151.
- Graves, M.F., Slater, W.H., Roen, D., Redd-Boyd, T., Duin, A.H., Furniss, D.W. & Hazeltine, P. (1988). Some characteristics of memorable expository writing: Effects of revisions by writers with different backgrounds. *Research in the Teaching of English*, 22, 242-280.
- Graves, M.F., Prenn, M.C., Earle, J., Thompson, M., Johnson, V. & Slater, W.H. (1991). Improving instructional texts: Some lessons learned. *Reading Research*

- Quarterly*, 26, 110-132.
- Kintsch, W. & Keenan, J. (1973). Reading rate as a function of number of propositions in the base structure of sentences. *Cognitive Psychology*, 6, 257-274.
- Kintsch, W. & Yarbrough, J.C. (1982). Role of rhetorical structure in text comprehension. *Journal of Educational Psychology*, 74, 828-834.
- Klare, G.R. (1985). *How to Write Readable English*. London: Hutchinson.
- Kobayashi, M. (1995). Effects of text organization and test format on reading comprehension test performance. Unpublished Ph.D Thesis. Thames Valley University, London.
- Kobayashi, M. (2002a). Method effects on reading comprehension test performance: Text organization and response format. *Language Testing*, 19, 191-218.
- Kobayashi, M. (2002b). Cloze tests revisited: Exploring item characteristics with special attention to scoring methods. *Modern Language Journal*, 86, 571-586.
- Kobayashi, M. (2004). Reading comprehension assessment: From text perspectives. *Scientific Approaches to Language* (Center for Language Sciences, Kanda University of International Studies), 3, 129-157.
- Kobayashi, M. (2005). An investigation of method effects on reading comprehension test performance. *2004 JALT Pan SIG Conference Proceedings*: 59-68.
- Lewkowicz, J. A. (1983). Method effect on testing reading comprehension: A comparison of three methods. Unpublished MA Thesis. University of Lancaster.
- McGee, L.M. (1982). Awareness of text structure: Effects on children's recall of expository text. *Reading Research Quarterly*, 17, 581-590.
- Meyer, B.J.F. (1975). *The organization of prose and its effects on memory*. Amsterdam: North-Holland Publishing Company.
- Meyer, B.J.F. (1985). Prose analysis: Purpose, procedures, and problems: Parts I and II. In B.K. Britton & J.B. Black (Eds.), *Understanding Expository Text* (pp. 11-64 & pp. 269-304). Hillsdale, N.J.: Lawrence Erlbaum.
- Meyer, B.J.F., Brandt, D.M. & Bluth, G.J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly*, 16, 72-103.

- Meyer, B.J.F. & Freedle, R.O. (1984). Effects of discourse type on recall. *American Educational Research Journal*, 21, 121-143.
- Meyer, B.J.F., Marsiske, M. & Willis, S.L. (1993). Text processing variables predict the readability of everyday documents read by older adults. *Reading Research Quarterly*, 28, 234-249.
- Mohammed, M.A.H. & Swales, J.M. (1984). Factors affecting the successful reading of technical instructions. *Reading in a Foreign Language*, 2, 206-217.
- Olsen, L.A. & Johnson, R. (1989). A Discourse-based approach to the assessment of readability. *Linguistics and Education*, 1, 207-231.
- Reder, L.M. & Anderson, J.R. (1980). A comparison of texts and their summaries: Memorial consequences. *Journal of Verbal Learning and Verbal Behavior*, 19, 121-134.
- Richgels, D.J., McGee, L.M., Loman, R.G., & Sheard, C. (1987). Awareness of four text structures: Effects on recall of expository text. *Reading Research Quarterly*, 22, 177-196.
- Salager-Meyer, F. (1991). Reading expository prose at the post-secondary level: the influence of textual variables on L2 reading comprehension (a genre-based approach). *Reading in a Foreign Language*, 8, 645-662.
- Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language Testing*, 1, 147-170.
- Shohamy, E. & Inbar, O. (1991). Validation of listening comprehension tests: the effect of text and question type. *Language Testing*, 8, 23-40.
- Steffensen, M.S. & Joag-Dev, C. (1984). Cultural knowledge and reading. In J.C. Alderson, & A.H. Urquhart, (Eds.), *Reading in a Foreign Language* (pp. 48-61). London: Longman.
- Steffensen, M.S., Joag-Dev, C. & Anderson, R.C. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly*, 15, 10-29.
- Urquhart, A. H. (1984). The effect of rhetorical ordering on readability. In J.C. Alderson & A.H. Urquhart (Eds.), *Reading in a Foreign Language* (pp. 160-175). London: Longman.

補遺 仮説検証の要約

	仮説	実証 (<i>F</i> 値)
1	読解力テスト結果と文章構成の間には相互関係はない	4.846** (<i>df</i> = 3, 731)
2	読解力テスト結果とテスト形式の間には相互関係はない	10.395** (<i>df</i> = 2, 732)
3	読解力テスト結果、文章構成、テスト形式の間には相互関係はない	6.149** (<i>df</i> = 6, 723)

** $p < .005$

	仮説	実証 (<i>F</i> 値)		
4	読解力テスト結果、文章構成、と言語能力の間には相互関係はない	下位群 n.s. (<i>df</i> = 3, 235)	中位群 3.82* (<i>df</i> = 3, 234)	上位群 8.87** (<i>df</i> = 3, 254)
5	読解力テスト結果、テスト形式と言語能力の間には相互関係はない	下位群 n.s. (<i>df</i> = 2, 236)	中位群 8.57** (<i>df</i> = 2, 235)	上位群 14.28** (<i>df</i> = 2, 255)
6	読解力テスト結果、文章構成、テスト形式と言語能力の間には相互関係はない	下位群 2.63* (<i>df</i> = 6, 227)	中位群 3.06* (<i>df</i> = 6, 226)	上位群 6.61** (<i>df</i> = 6, 246)

* $p < .05$, ** $p < .005$

261-0014

千葉県千葉市美浜区若葉 1-4-1

神田外語大学

言語科学研究科

m.kobay@kanda.kuis.ac.jp