

原 著

## 英語歯科用語に関するテスト項目の研究(3)

—項目難易度表示の付与—

廣 瀬 浩 二

明倫短期大学 歯科衛生士学科

A Study of Test Items on English Terminology for Dentistry (3):  
Assigning Difficulty Index to Each Item

Koji Hirose

*Department of Dental Hygiene and Welfare, Meirin College*

本研究の目的は、英語歯科用語のテスト項目を作成分析し、項目難易度表示を付け、項目銀行を作成することであった。新たに作成したのは18項目である。このテストを64名の学習者に課した。古典的テスト理論と項目応答理論の双方によって分析した結果では、このテストは信頼に足るテストであった(信頼性係数は0.829)。また、本テストは測定しようとする内容を盛り込み妥当である。

テスト実用性の観点から、各項目に難易度表示を設定した。独自の分類基準、すなわち、項目難易度指数が0.700未満はAランク、0.700~0.799はBランク、0.800以上はCランクとした。これまでに分析済みの18項目と今回作成分析した18項目で36項目の項目銀行となった。36項目の中で、Aランクは3項目、Bランクは9項目、Cランクは24項目である。今後はバランス上、Aランク、Bランクの項目を充実したい。  
キーワード：英語歯科用語、言語テスト、項目応答理論、困難度表示、項目銀行

In this study, eighteen English language test problems for dentistry were newly composed and examined by classical test theory (CTT) and item response theory (IRT) analyses. Such a time-consuming but crucial process will lead to the completion of the item banking. Sixty-four subjects, who were all college students belonging to the department of dental hygiene, participated in this study. The test results were analysed by both CTT and IRT. The statistical information showed that the test was reliable due to the coefficient alpha 0.829 and valid since those test problems represented a sample of the course content (content validity).

In order to observe test practicality, original difficulty indices were provided for each test item. Each range of difficulty was as follows: items of less than 0.700 difficulty index fell into class A which expresses the greatest difficulty; items at 0.700~0.799 were designated as class B which shows the intermediate level; those over 0.800 comprised class C, which registers as easy. The item bank currently embraces thirty-six items, including eighteen items which had been already analysed by the time we commenced the study. Among those items, class A had only 3 items, B possessed 9, and C retained 24. In a further study, more class A or B items should be added for a better balance of the whole banking.

Key words : English for dentistry, Language test, Item response theory, Difficulty index, Item bank

## 緒 言

これまで本学歯科衛生士学科「英語Ⅱ」の教育プログラムに直結するテストの開発と改訂を行ってき

た。具体的には、英語歯科用語のテスト項目について基幹部分が英語、選択肢数3のテスト項目から構成されるテストタイプA(廣瀬<sup>1)</sup>)および基幹部分が日本語、選択肢数4から成るテストタイプB(廣瀬<sup>2)</sup>)

を作成した。さらに、これらテストタイプAとBを手直し改善するために、古典的テスト理論 (Classical Test Theory: 以下CTTと省略) および項目応答理論 (Item Response Theory: 以下IRTと省略) の両観点から詳細に比較分析した (廣瀬<sup>3)</sup>)。その結果、テストタイプBの方が適切な項目の多いことが統計的に判明した。そこで、今後はテストタイプBの形式に従ってテスト項目を作成し項目銀行を充実することにした。本稿では、新たなテスト項目を作成し、CTTとIRTの両面から検討を行った。CTTの欠点は、その分析結果が実施されたテストの受験者だけにしか適合しない、という点にある (Alderson et al.<sup>4)</sup>)。一方、IRTを用いた分析の特徴は、同一のテスト項目をたとえ異なった受験者集団に課しても、その項目の持つ特性、つまり項目困難度や項目弁別力の値に変化が生じないことにある。このことから、IRTによる分析を行うことによって、各テスト項目を難易度別に分類収集することが可能となるであろう。そこで、これまでに作成し分析済みのテスト項目と今回作成した項目に難易度表示を加えた。この表示を付け加えることによって、更に利用しやすい項目銀行になるはずである。

## 対象および方法

### 1. 対象

本研究の対象は、明倫短期大学歯科衛生士学科1年生64名である。テストは平成11年10月に通常の授業時間内で行った。

### 2. 方法

#### 1) テスト項目の作成

良い多肢選択式テスト (multiple-choice test) を開発するためには、3つの作業が必要となる。すなわち、基幹部分 (stem) を書くこと、選択肢 (options) を書くこと、および継続した項目の改訂である (Kehoe<sup>5)</sup>)。まず、基幹部分と選択肢の作成に着手した。

平成11年4月から平成11年10月までに指導した165の英語歯科用語の中から18のテスト項目を新たに作成した。作成にあたっては、多肢選択式テスト作成の基本的ルールに従った。各項目は表1に示すように、最初に日本語の用語を挙げ、次に日本語に対応する4つの選択肢 (正しい選択肢1, 錯乱肢3) を配列した。受験者には、各項目において日本語に対応する英語用語をア～エの中から選び、解答欄に記入させた。

表1. 作成したテスト項目

(1)骨膜剥離子 ( )	ア, peripatetic elevator ウ, periosteum elevator	イ, peripheral elevator エ, periosteal elevator
(2)骨膜刀 ( )	ア, periosteum knife ウ, peripatetic knife	イ, periosteal knife エ, peripheral knife
(3)骨やすり ( )	ア, bone fille ウ, bone fill	イ, bone file エ, bone file
(4)根管 ( )	ア, root canal エ, root canaliculi	イ, root canaille ウ, root canary
(5)根管充填 ( )	ア, root cannel filling ウ, root canaille filling	イ, root canal filling エ, root canary filling
(6)根管清掃 ( )	ア, cleaning of root canaille ウ, cleaning of root canal	イ, cleaning of root canary エ, cleaning of root cannel
(7)口腔診査 ( )	ア, aural examination ウ, aural examinant	イ, oral examinant エ, oral examination
(8)個人トレー ( )	ア, specialty tray ウ, specific tray	イ, special tray エ, species tray
(9)近心隣接面 ( )	ア, mesial proximal surface ウ, mesial proximate surface	イ, mesic proximal surfeit エ, mesic proximate surface
(10)頬側面 ( )	ア, buccal surface ウ, buccal surfeit	イ, buckeye surface エ, buckeye surfeit
(11)抗生物質 ( )	ア, antibiosis エ, antibiotic	イ, antibaryon ウ, antibacterial
(12)咬翼型フィルム ( )	ア, bitingwing film ウ, bitewing firm	イ, bitewing film エ, bitingwing film
(13)咬合面 ( )	ア, occlusal surface ウ, occlusive surface	イ, occlusion surfeit エ, occult surfeit
(14)酸性フッ素溶液 ( )	ア, acidulated fluoride solution イ, acidulous fluoride solution ウ, acidulated fluorine solution エ, acidulous fluorine solution	
(15)暫間披覆冠 ( )	ア, contemporary crown ウ, temporary crown	イ, temporal crown エ, temporality crown
(16)楔状欠損 ( )	ア, weed-shaped defect ウ, web-shaped defect	イ, wedge-shaped defect エ, weeny-shaped defect
(17)3才児歯科検診 ( )	ア, dental examinant for 3-year old イ, dental examination for 3-year old ウ, dentil examination for 3-year old エ, dentil examinant for 3-year old	
(18)酸蝕症 ( )	ア, erosion エ, erose	イ, elocution ウ, erosive

#### 2) 分析方法

今回作成したテスト結果のデータ分析に際し、大友賢二と中村洋一<sup>6)</sup> 作成 (Copyright: Kenji Ohtomo & Youichi Nakamura) の Test Data Analysis Program: TDAP ver. 1.0 (1996) を使用した。統計的に処理されたデータを項目困難度・項目弁別力・モデルとの適合度の3点から解析した。

## 結果と考察

### 1. 基礎統計量

CTTによる項目分析から判明した基礎統計量は、表2に示す通りである。それによると、得点は2点(低得点)から18点(高得点)まで散らばってはいる。標準偏差(SD)は3.559で、平均値からの距離はある程度確保されている。歪み(Skewness)が負の数値を得たことから、得点分布は負の方向に歪んでいる。平均値が14.156になったのもそのためである。また、尖度(Kurtosis)は2.117と低い値であったため得点分布は最頻値に集中してはいない。基礎統計量をみる限り、本テスト得点になんら問題はみられない。そもそも、本研究で作成しているテストは目標規準準拠テスト(criterion-referenced test: CRT)で、授業目標の習得度を判定するものであるから、テスト得点が正規分布曲線を描く必要は全くない。むしろCRTで正規分布が現われるのはテストそのものの、または、カリキュラムや指導のどこかに間違いがあることを示している(Brown<sup>7)</sup>)。

表2. CTTの項目分析による基礎統計量

受験者数 (Number of examinees)	64
素点の総和 (Sum of the raw scores)	906
最小得点 (minimum score)	2
最大得点 (maximum score)	18
中央値 (Median)	16
範囲 (Range)	16
平均値 (Mean)	14.156
分散 (Variance)	12.663
標準偏差 (Standard deviation)	3.559
歪み (Skewness)	-1.391
尖度 (Kurtosis)	2.117

### 2. CTTによる項目分析結果

表3は、CTTにより、困難度・点双列相関係数による項目弁別力・実質選択肢数等の数値を求めた結果である。標準適切度合計による順位(Rank in SATOT order)で配列したものである。

CTTによる項目分析では、特に、項目困難度(DIFF)と項目弁別力指数(DISC)に注目した。項目困難度は0.500が最も適切である、といわれる。項目困難度の観点からいえば、項目番号(14)が最も適切なテスト項目であると考えられる。項目(4)(6)はかなり易しい項目である、と判断されるし、項目(1)(2)(5)

(7)(8)(12)(15)(16)も易しい部類に属すると考えられる。しかしながら、これら易しい項目も配列の仕方によっては有効に活用できる。

また、項目弁別力指数の観点からみると、項目番号(17)が弁別力にやや欠けると思われる。しかしながら、ほとんどの項目は+.400以上の数値を示し、満足できる結果となった。

表3. CTTによる項目分析結果: 標準適切度合計順位

RANK	NO	DIFF	DISC	AENO	ADIF	ADIS	AAEN	SATOT
1	11	0.703	0.730	2.537	0.844	1.000	0.994	2.059
2	16	0.813	0.651	1.969	0.625	0.736	0.978	1.746
3	14	0.500	0.544	3.452	0.750	0.421	0.995	1.711
4	9	0.766	0.553	2.184	0.719	0.440	0.963	1.647
5	13	0.781	0.576	2.084	0.688	0.495	0.942	1.620
6	18	0.750	0.522	2.244	0.750	0.375	0.950	1.619
7	3	0.766	0.543	2.154	0.719	0.417	0.938	1.598
8	10	0.734	0.404	2.328	0.781	0.195	0.957	1.564
9	2	0.875	0.574	1.646	0.500	0.492	0.962	1.512
10	12	0.813	0.404	1.965	0.625	0.195	0.974	1.477
11	8	0.875	0.468	1.660	0.500	0.280	0.982	1.440
12	15	0.844	0.551	1.748	0.563	0.436	0.900	1.435
13	5	0.875	0.543	1.598	0.500	0.428	0.890	1.369
14	17	0.484	0.212	3.224	0.719	0.047	0.882	1.330
15	6	0.906	0.387	1.424	0.438	0.524	0.826	1.272
16	1	0.813	0.527	1.763	0.625	0.385	0.770	1.254
17	4	0.969	0.462	1.174	0.313	0.272	0.921	1.205
18	7	0.891	0.466	1.477	0.469	0.277	0.805	1.142

#### (NOTES)

RANK = 標準適切度合計による順位 (Rank in SATOT order)  
 NO = 項目番号 (Item number)  
 DIFF = 項目困難度指数 (Item difficulty index)  
 DISC = 項目弁別力指数 (Discrimination power index)  
 AENO = 実質選択肢数 (Actual equivalent number of options)  
 ADIF = 項目困難度適切度 (Appropriateness of difficulty)  
 ADIS = 項目弁別力適切度 (Appropriateness of discrimination power index)  
 AAEN = 実質選択肢数適切度 (Appropriateness of actual equivalent number of options)  
 SATOT = 標準適切度合計 (Standard appropriateness total)

### 3. 信頼性

表4は、信頼性係数に関する数値を示したものである。アルファ係数は比較的高い数値(0.829)を表示し、テストの一貫性・信頼性は高い、と考えられる。また、測定の標準誤差は1.472と非常に小さく、最大得点が18であることを考慮に入れても得点の偶然による分散はかなり少ない。

表 4. 信頼性係数

アルファ係数 (Coefficient alpha)	0.829
測定標準誤差 (Standard error of measurement)	1.472
得点の総和 (Sum of the raw scores)	906
平均値 (Mean of the raw scores)	14.156
標準偏差 (Standard deviation of the test)	3.559
分散 (Variance of the test)	12.663
項目分散の総和 (Sum of item variance)	2.750
平均通過率 (Average proportion of passing)	0.786
平均不通過率 (Average proportion of failing)	0.214

本テストに関しては、比較的高い信頼性係数 (0.829) が得られた。データ上、もし項目数を表 5 のように増やしていくならば、更に高い信頼性係数を得ることができるであろう。

表 5. アルファ係数とテスト項目数の関係

If you want the following coefficient alpha:	Number of your new test items should be:
0.800	15
0.820	17
0.840	20
0.860	23
0.880	27
0.900	33
0.920	43
0.940	58
0.960	89
0.980	182
0.829	⇐ 18
0.964	⇐ 100

#### 4. 妥当性

本研究のテストは目標規準準拠テスト (CRT) であることから、本テストの妥当性を調べるには、内容的妥当性 (content validity) と構成概念妥当性 (construct validity) の分析法が適用できる。

本学歯科衛生士学科「英語Ⅱ」の授業内容は、一般英語と歯科英語である。一般英語は、国際的コミュニケーションのための一般英会話である。具体的には、初級者用に作成されたスピーキングとリスニング技能養成のテキストに基づいた指導である。歯科英語は、語彙と診療所での会話用である。

内容的妥当性を測る統計的測定法はないため、テ

スト開発の過程で綿密な計画を立てておく必要がある (Hatch and Lazaraton<sup>8)</sup>)。本テストで測ろうとしているのは何か。また、テストは測ろうとしている事項を測っているか。本テストの意図は、「英語Ⅱ」の指導内容の一部、つまり歯科の要素の中でも専門的語彙に関わる部分を評価することである。したがって、専門的語彙の習得状況を測定すればよいことになる。その結果が表 1 のようなテスト形式となった。

#### 5. 構成概念妥当性

構成概念妥当性を測るには、いくつかの統計的方法があり、異集団法 (differential-group experiment) や介入法 (intervention experiment) が代表的である (Brown<sup>9)</sup>)。いずれも実験法であり、本研究では、これら実験法に適する研究計画を設定していないため、構成概念妥当性について断定できない。

#### 6. IRT による項目分析結果

表 6 は、IRT の 1 パラメーター・ロジスティック・モデルにより、項目困難度パラメーター・モデルとの適合度を算出した結果である。

表 6. IRT による項目分析結果

Item No.	Final Calib.	Standard Error (d)	Fit (t)
1	-0.025	0.370	-1.779
2	-0.659	0.433	-2.793
3	0.353	0.342	0.370
4	-2.596	0.817	-5.677
5	-0.659	0.433	-0.599
6	-1.082	0.490	-3.327
7	-0.858	0.459	1.627
8	-0.659	0.433	-1.265
9	0.353	0.342	-1.231
10	0.579	0.330	2.291
11	0.790	0.320	-2.598
12	-0.025	0.370	1.494
13	0.233	0.350	-2.078
14	2.018	0.306	-0.097
15	-0.317	0.396	-2.032
16	-0.025	0.370	-2.713
17	2.113	0.308	3.889
18	0.468	0.336	-1.056

IRT による項目分析では、特に、モデルとの適合度に注目した。モデルとの適合度は、2.00 が危険域である。2.00 あるいはそれ以上の場合は、応答妥当性を欠

きモデルとは適合しないと考えられる (Henning<sup>10)</sup>)。モデルとの適合度では、項目番号 (10) は 2.291, (17) は 3.889 といずれも 2.00 以上を示している。この 2 項目はモデルと適合せず、テスト項目から除外したほうが良いと思われる。一方、負の適合度を持つ項目が多かった。-2.00 あるいはそれ以下の場合、モデルと過剰適合していると考えられるが、テスト項目から除外する理由にはならない。

ここで問題となった 2 項目は CTT の分析では、項目 (10) は項目困難度及び項目弁別力いずれにおいても適正であった。ところが、項目 (17) だけは項目困難度で問題なかったが、項目弁別力では弁別力に欠けると判断された。したがって、ここで問題とすべきは項目 (17) で、改善に値するだろう。

### 7. 難易度表示の付与

#### 分析済みのテスト項目

これまでに基幹部分が日本語、選択肢が英語で構成された 18 のテスト項目を作成した。これらの項目では選択肢数が 4 であるから、理想困難度 0.500, 当て推量確率 0.125, 最適困難度 0.625 となる。また、IRT の分析からこれら各項目の最終項目困難度は判明している。ここでは、最適困難度を参考に独自の分類基準を設定し、3 段階の難易度表示 (A, B, C) にした。すなわち、項目困難度指数が 0.700 未満は A (難), 0.700~0.799 は B (普), 0.800 以上は C (易) である。その結果、A ランクは 1 項目、B ランクは 3 項目、C ランクは 14 項目となった。

表 7. 分析済テスト項目の難易度表示

(1)彫刻刀 ( )	ア, carvone knife	イ, carvacrol knife	ウ, carven knife	エ, carving knife	⇒C
(2)槌 ( )	ア, mallet	イ, mallee	ウ, malleolus	エ, malleus	⇒C
(3)中切歯 ( )	ア, centric incisor	イ, centroclinal incisor	ウ, centripetal incisor	エ, central incisor	⇒C
(4)側切歯 ( )	ア, latent incisor	イ, lateral incisor	ウ, later incisor	エ, latter incisor	⇒C
(5)定期検査 ( )	ア, recap	イ, recall	ウ, recast	エ, recce	⇒C
(6)伝達麻酔 ( )	ア, block anaesthesiology	イ, block anaesthetic	ウ, block anaesthesia	エ, block anesthetics	⇒C
(7)他覚症状 ( )	ア, objective symptom	イ, objective symptomatology	ウ, objective symptomology	エ, objective symphysis	⇒C
(8)倒立円錐バー ( )	ア, inverted comma bur	イ, inverted con bur	ウ, inverted cone bur	エ, inverted cone bur	⇒B

(9)鍍銀法 ( )	ア, silver impregnant method	イ, silver imprescriptibility method	ウ, silver impregnation method	エ, silver imprecation method	⇒A
(10)ドライソケット ( )	ア, dry sock	イ, dory socket	ウ, dry socket	エ, drory sock	⇒C
(11)塗蠟絹糸 ( )	ア, floss silk	イ, flossy silk	ウ, florin silk	エ, florist silk	⇒C
(12)乳歯 ( )	ア, mill teeth	イ, milk teeth	ウ, milky teeth	エ, millcake teeth	⇒B
(13)嚢胞 ( )	ア, cycle	イ, cyst	ウ, cyclist	エ, cystine	⇒C
(14)膿瘍切開 ( )	ア, incision of abscess	イ, incisor of abscess	ウ, incisor of absence	エ, incision of absence	⇒B
(15)唾器 ( )	ア, salivary ejector	イ, salivary ejection	ウ, saliva ejection	エ, saliva ejector	⇒C
(16)麦粒鉗子 ( )	ア, dressed forceps	イ, dressing forceps	ウ, dressing forcite	エ, dressed forcite	⇒C
(17)抜歯 ( )	ア, extraction of the tooth	イ, extractor of the tooth	ウ, extradition of the tooth	エ, extraneousness of the tooth	⇒C
(18)抜歯鉗子 ( )	ア, dental extortionary forceps	イ, dental extortionary forcite	ウ, dental extracting forceps	エ, dental extracting forcite	⇒C

### 8. 新たに作成したテスト項目

表 8 は、今回作成し分析したテスト項目に難易度表示を付け加えたものである。これらの項目についても、IRT で分析した項目困難度指数をもとに 3 段階の難易度を表示した。即ち、項目困難度指数が 0.700 未満は A (難), 0.700~0.799 は B (普), 0.800 以上は C (易) である。その結果、A ランクは 2 項目、B ランクは 6 項目、C ランクは 10 項目となった。ここでは、表 7 より A ランク B ランクの項目数が増え、バランスを保つ上で好ましい結果となった。

表 8. 新たに分析したテスト項目の難易度表示

(1)骨膜剥離子 ( )	ア, peripatetic elevator	イ, peripheral elevator	ウ, periosteum elevator	エ, periosteal elevator	⇒C
(2)骨膜刀 ( )	ア, periosteum knife	イ, periosteal knife	ウ, peripatetic knife	エ, peripheral knife	⇒C
(3)骨やすり ( )	ア, bone fille	イ, bone filet	ウ, bone fill	エ, bone file	⇒B
(4)根管 ( )	ア, root canal	イ, root canaille	ウ, root canary	エ, root canaliculi	⇒C

- (5)根管充填( )ア, root canal filling イ, root canal filling  
ウ, root canal filling エ, root canal filling⇒C
- (6)根管清掃( )ア, cleaning of root canal イ, cleaning of root canal  
ウ, cleaning of root canal エ, cleaning of root canal ⇒C
- (7)口腔診査( )ア, oral examination イ, oral examination  
ウ, oral examination エ, oral examination ⇒C
- (8)個人トレー( )ア, specialty tray イ, special tray  
ウ, specific tray エ, species tray ⇒C
- (9)近心隣接面( )ア, mesial proximal surface  
イ, mesial proximal surface  
ウ, mesial proximal surface エ, mesial proximal surface ⇒B
- (10)頬側面( )ア, buccal surface イ, buccal surface  
ウ, buccal surface エ, buccal surface ⇒B
- (11)抗生物質( )ア, antibiotic イ, antibiotic ウ, antibacterial  
エ, antibiotic ⇒B
- (12)咬翼型フィルム( )  
ア, bitewing film イ, bitewing film  
ウ, bitewing film エ, bitewing film ⇒C
- (13)咬合面( )ア, occlusal surface イ, occlusion surface  
ウ, occlusive surface エ, occult surface ⇒B
- (14)酸性フッ素溶液( )  
ア, acidulated fluoride solution  
イ, acidulous fluoride solution  
ウ, acidulated fluorine solution  
エ, acidulous fluorine solution ⇒A
- (15)暫間被覆冠( )ア, contemporary crown イ, temporal crown  
ウ, temporary crown エ, temporality crown ⇒C
- (16)楔状欠損( )ア, weed-shaped defect  
イ, wedge-shaped defect  
ウ, web-shaped defect  
エ, weeny-shaped defect ⇒C
- (17)3才児歯科検診( )  
ア, dental examination for 3-year old  
イ, dental examination for 3-year old  
ウ, dentil examination for 3-year old  
エ, dentil examination for 3-year old ⇒A
- (18)酸蝕症( )ア, erosion イ, elocution  
ウ, erosive エ, erose ⇒B

の1項目で項目弁別力の欠如がみられ改良することとなった。これまで行ってきたテストタイプの分析・テスト項目の分析を通し、改良項目数は遥かに減少した。

また、これまでに作成した項目も含め各々に項目困難度表示を追加した。これによって、実用性は高まった。

## 文 献

- 1) 廣瀬浩二：英語歯科用語に関するテスト項目の研究 (1)一項目困難度, 項目弁別力指数, モデルとデータの適合度の検討一. 明倫歯誌, 1: 39-43, 1998
- 2) 廣瀬浩二：英語歯科用語に関するテスト項目の研究 (2)一オプションの数と種類の違い一. 明倫歯誌, 2: 51-56, 1999
- 3) 廣瀬浩二：語彙テスト項目の分析. 外国語教育評価学会研究紀要, 2: 55-64, 1999
- 4) Alderson J C, Clapham C, and Wall A D: Language Test Construction and Evaluation. Cambridge University Press., Cambridge, U.K., 1995
- 5) Kehoe J: Writing Multiple-Choice Test Items. ERIC., EDO - TM - 95 - 3, 1995
- 6) Ohtomo K and Nakamura Y: Test Data Analysis Program. Vers.1.0. Computer software. MS-DOS, 1.2MB, disk. 1996
- 7) Brown J D: Differences Between Norm-Referenced and Criterion-Referenced Tests. in Brown, J. D. and Yamashita S O (eds.), Language Testing in Japan. The Japan Association for Language Teaching., Tokyo, Japan, pp.12-19, 1995
- 8) Hatch E and Lazaraton A: The Research Manual. Heinle and Heinle Publishers., Boston, MA, U.S.A., 1991
- 9) Brown J D: Understanding Research in Second Language Learning. Cambridge University Press., Cambridge, U.K., 1988
- 10) Henning G: A Guide to Language Testing. Heinle & Heinle Publishers., Boston, MA, U.S.A., 1987

## 結 語

本研究では、英語歯科用語の習得状況を評価するためのテスト項目を作成するとともに困難度表示を付加し、項目銀行の実用性に注目した。

今回新たに作成したテストの得点を統計的に処理した結果、IRTによる項目分析では、モデルとの適合度に関し2項目に不適合がみられた。一方、CTTの分析では、これら2項目の中1項目は項目困難度及び項目弁別力いずれにおいても適正であった。別