2大学における英語力到達度及び運用能力テストの分析

Analyses of English language achievement and proficiency tests at two universities

梅本　孝　　法月　健
Takashi UMEMOTO　　Ken NORIZUKI

（平成18年9月13日受理）

到達度評価（achievement assessment）の定期試験やクラス編成（placement）等で使用される運用能力（proficiency）テストが、大学の英語教育研究で客観的に分析されることはそれほど多くない。本研究では、2大学4クラス（A大学クラス1、2、B大学クラス1、2）計132人の学生に実施された一部共通項目を含む英語の定期試験と総合的な英語運用能力（general English language proficiency）を測定する外部基準（external criterion）テストの試験結果を検証した。分析の結果、クラスによってテストの相関関係は大きく異なり、一部のクラスでクラス編成や成績評定に加味する目的で使用した外部テストが適切に機能していない可能性が示された。直接比較が可能な定期試験の部門点でB大学よりも得点が有意に上回ったA大学の2クラス間では、いずれのテスト得点にも有意差は見られなかったが、能力別クラス編成が実施されていたB大学の2クラス間では、2種の外部基準テストの得点、定期試験の総合点に有意な差が確認された。さらに定期試験の部門点や定期試験時に実施した記述問題の文字数[1]などについても分散分析及び多重比較を行い、到達度、運用能力、日本語表現力の関係についてデータから検証されたことや検証できなかったことについて議論し、今後の研究課題について展望した。

1．はじめに

大学の英語教育において定期試験の得点が成績評定に占める割合は一般的に高いものと考えられる。筆者の一人が過去担当した科目を例に挙げると、同じクラスを2名の教員が週1時間ずつ担当し、成績を協議して決定する形態の科目の成績評定において、筆者とペアを組んだ教員が提示する成績データが、1回の定期試験の結果を決定的な基準に据えていることが少なくなかった。週1時間10数回実施された授業の成果が定期試験の結果に集約されるという論理だが、このように重要な意味合いを持つ定期試験が多角的な到達度評価（achievement assessment）の視点から議論されることはあまり多くない（Weir, 1993; Norizuki, 1998）。近年では、市販の標準化された（standardized）英語運用能力（proficiency）テストの得点がクラス編成（placement）の基準になったり、成績評定の一部に加味されることも少なくないが、教育プログラムの水準に適合しているのか疑問に

1 文字数は厳密に数えたわけではなく、原則として、横書きの一行目の文字数を数え、それと縦のラインの行数を掛けて算出した。その際に当時静岡産業大学の学生であった吉引博子の助力を得た。ここに記して、感謝したい。

感じながらも、このような外部作成のテストを長年にわたって使用し続けていることもある（Brown, 2004)。統計プログラムが普及した今日、テストの分析は言語テスト研究者だけの特権ではなく、英語教育に携わる者が自分の担当するクラスや教育機関内においても、率先して行っていくことが望まれる（清川等, 2003)。

本稿の筆者の一人である梅本は、２つの私立大学で授業を行う中で、学生の英語力や日本語表現力の違いに関心を持ち、客観的に比較分析するために定期試験と外部基準の英語運用テストのデータを収集した。もう一人の筆者の法月は梅本と研究計画を検討し、以下で述べる分析を行った。

２．研究方法

2.1　研究の目的

研究には２大学４クラス132人分のデータを使用したが、この研究の背景には、授業を遂行する一教師としての経験から梅本が日頃から追究してきた次のような教育的研究課題がある。

１）十分な時間をかけずに作成することの多い定期試験が、信頼性や妥当性が高いとされる外部基準の英語能力試験とどの程度の相関関係を有するか。

２）同じ教材を使って授業を行い、２大学の学習者の英語能力到達度や運用能力に顕著な違いを感じていたが、ほぼ同じ内容で実施した定期試験の結果に有意な差は存在するか。４つのクラス間でもそれぞれに有意な差は見られるか。

３）２大学の学習者には英語能力だけでなく、日本語の表現能力にも大きな違いがあると感じていたが、何らかの差を具現化し、英語能力との関係を探ることは可能か。その差をクラス間でも確認できるか。

４）複合的な性格の項目を含む定期試験のそれぞれの項目群で、相関やグループ間の比較の分析において顕著な特徴が見られるか。

2.2　使用テスト

2004年１月26日にＡ大学２クラスで、梅本が担当する１年生の必修英語クラスの定期試験が実施された。問題は単文で構成される４肢択一方式の単語補充問題が12問、英文和訳問題が５問で、A大学においては認知言語学的な解釈を記述させる問題が１問追加された。さらに自分が知っているおもしろい情報を記述する欄を設けたが、本研究の分析では定期試験の得点には含めず、文字数のみ分析した。この部分は定期試験の採点の際にボーナスポイントとして、A大学では20点満点で採点し、B大学では10点満点で成績に反映させることを宣告した。A大学を20点、B大学を10点にした理由は、この試験を行うまでに行った同様の試みにおいて、A大学においては読む者を引き込む大作が多く、B大学においてはそれほどの大作がなかったことによる。採点はすべて減点法で行った。A大学で外部基準となったテストはTOEFLで、2003年12月７日に実施された。実験に参加した被験者はクラス１が34名、クラス２が38名で、TOEFL未受験者の欠損値を除いた数はそれぞれ33名と36名だった。

一方、2004年1月29日にはB大学2クラスで、梅本が担当する1年生の必修英語クラスの定期試験が行われた。問題は認知言語学的な内容の問題を除いてはA大学と同一[2]で、外部基準としてはコース開始時の2003年4月11日に能力別クラス編成試験として実施されたベネッセコーポレーションの「英語コミュニケーション能力試験」とコース終了時の2004年1月28日に実施されたTOEIC Bridgeの2種類のデータを使用した。実験に参加した被験者はクラス1が31名、クラス2が29名で、TOEIC Bridge未受験者の欠損値を除いた数はそれぞれ27名、26名だった。なお、クラス1はクラス編成試験の結果の上位群、クラス2は中位群で構成されるグループだった。各グループに2、3名の中国人留学生が含まれていて、学習体験や言語使用環境においては日本人学生と異なっていたが、テスト数値的には日本人学生と大きく異なる特徴はなかったため、そのままのグループで分析を行った。

2.3 A大学とB大学の授業と試験問題の具体的な差

A大学では『TOIECテスト860点突破大特訓』(長本　吉斉著、ベレ出版　2001)のリーディング部門の単語補充問題(四肢択一)の部分をテキストとして授業中に使用した。半期で9ページ進んだ。授業中に解説を行った問題数は40問である。一つ一つの問題を懇切丁寧に説明した。具体的な中身としては問題の中で解答となっている単語、あるいは解答となっている単語を含む単語の固まりを見つけ出し、Googleで当該の単語、もしくは当該の単語を含む単語群を探し、例文を探し、その例文を含むより大きな文の固まりを黒板に板書し、説明した。原則として一つの問題に対して、2種類の異なった文章を例題として示し、説明した。また、Cognitive linguistics (David Lee著、Oxford 2001)の空間概念を認知言語学の観点から説明している部分を授業中に解説した。約15ページ進んだ。この2冊の授業における使用時間はおよそ半分ずつである。定期試験において、おもしろい丸秘情報を記述する、ということに対するボーナス点は20点であると宣言した。この記述欄には問題用紙(A4版)の1ページ全面を充てた。それ以上書いても良いということも述べた。試験時間は90分であった。

B大学でも同一のTOEICテキストを用い、同一の問題を解説した。説明の丁寧度はA大学よりもB大学の方がはるかに高かった。B大学でもA大学と同様に、問題の中で解答となっている単語、あるいは解答となっている単語を含む単語の固まりを見つけ出し、Googleで当該の単語、もしくは当該の単語を含む単語群を探し、それらを含む例文を探し、その例文を含むより大きな文の固まりを黒板に板書し、説明した。原則として一つの問題に対して、3種類の異なった文章を例題として示し、説明した。A大学では例題として板書した文章の訳は口頭で読み上げるだけであったが、B大学においては板書した英語の例文の訳は口頭で伝えるだけではなくすべて板書を行った。また、B大学ではA大学とは違い、上記の認知言語学のテキストを使用することは一切なかった。B大学では試験でのおもしろい情報を記述する、ということに対するボーナス点は10点であると宣言した。この欄はページの中ほどのところから始まっているが、裏面を使用しても良いとした。試

2 両大学ともに定期試験の後半部に6問の英文和訳問題を含んでいたが、大学間で配点が異なるうえに、A大学において部門点を集計できなかった。よって大学間の定期試験結果の比較においては、前半部の結果のみを扱うこととした。

験時間は90分であった。B大学ではA大学と違い、試験においてCognitive linguisticsに関連する試験問題はなかった。したがって、B大学の学生はA大学の学生よりも時間的に余裕があったはずである。総合的な学業習熟度に関してはA大学とB大学では大きな差があると考えられる。

単語補充問題は13問あったが、6番目の問題において(B)と(C)を同じ選択肢としてしまったために、この問題は両大学の学生共に解答させなかった。学生が解いた4択の単語補充問題は全部で12問ということになる。そのうち5番目の問題から13番目の問題（問題6を除く）はすべて、授業中に解説したテキストの中の問題とまったく同じ問題である。それに対して問題1から問題4までの問題は学生が試験のときに初めて見る問題である。英文和訳と認知言語学に関する問題は、個別に分析するデータが揃わなかったため、ここで詳しく触れることはしない。

2.4　分析方法

A大学、B大学各クラスにおける定期試験総得点、定期試験の共通項目である語彙補充問題の未習部門（4問）、既習事項（12問）の得点、外部基準テストとしてA大学はTOEFL、B大学はベネッセコーポレーション社の試験、TOEIC Bridgeの得点について、相関、t検定、分散分析、分散分析結果の多重比較を、Microsoft Excel(XP)とSPSS(11.5J)を使用して分析した。

3．結果

3.1　定期試験と外部基準の関係

表1は定期試験と外部基準相関係数を示したものである。前述の通り、A大学とB大学の定期試験は構成が若干異なる。A大学のクラス1、2では定期試験とTOEFLとの間に5%水準でやや低めの有意な相関を示している。A大学のクラス1とクラス2のデータを一つにすると、こちらも有意な相関（.387）が確認できた。

一方、B大学においては、上位群のクラス1については、クラス編成時のベネッセ社の試験とは中低位、定期試験とは1日違いでコース終了時に実施したTOEIC Bridgeとは中位の有意な相関値が示されているが、中位群のクラス2については、いずれも有意な相関に至っていない。B大学の総合データでは、中程度の有意な相関（ベネッセ試験で.442、TOEIC Bridgeで.538）が見られた。

表1　定期試験と外部基準テストの相関

外部テスト	TOEFL（A大学）		ベネッセ（B大学）		TOEIC Bridge（B大学）	
	クラス1	クラス2	クラス1	クラス2	クラス1	クラス2
定期試験	.400*	.386*	.424*	−.267	.576**	.107

*5%水準（両側）で有意　**1%水準（両側）で有意

3.2　外部基準で見たクラス間の差

表2には各大学の2クラス間で外部基準テストの得点分布が比較されている。A大学2

クラスにおいては、TOEFLの得点に統計的に有意な差はなく、両クラスはこのテスト結果で見る限り、英語運用能力において均衡した水準にあると考えられる。

一方、B大学2クラスについては、クラス編成の基準となったベネッセ社の試験結果では当然ながら両クラス間の差が大きなt値で示されているが、クラス編成から半年以上経て受験したTOEIC Bridgeの結果にも有意な差が確認できる。

なお、外部基準として用いた試験が異なるため、A、B大学の英語運用能力を直接比較するデータはなかった。

表2　外部基準テストの得点分布比較

外部テスト	TOEFL（A大学）		ベネッセ（B大学）		TOEIC Bridge（B大学）	
	クラス1	クラス2	クラス1	クラス2	クラス1	クラス2
平均点	449.15	451.03	287.58	224.45	117.3	104.92
分散	1878.57	2085.40	1857.59	194.89	316.92	60.87
被験者数	33	36	31	29	27	26
t値	−0.175		7.733**		3.307**	

** 1%水準（両側）で有意

3.3　定期試験の部門別得点及び記述欄文字数と外部基準との関係

表3はA大学クラス1におけるさまざまな相関値をまとめたものである。定期試験の単語補充問題は授業時に扱っていない未習項目が4問と授業時で扱った既習項目8問で構成される。このクラスでは未習項目と外部基準TOEFLの得点間で中低位の有意な相関（.422）が見られるが、既習項目とTOEFLの得点間には有意な相関は見られない。定期試験の後半は英文和訳と認知言語学的な説明に関する内容であったが、B大学の試験に含まれない問題もあるため、分析から省略した。右端下の数値（−.021）は自分の知っているおもしろい丸秘情報を記述した文章の文字数（以後、文字数）とTOEFLの得点との相関係数である。この記述項目はA大学とB大学の試験問題で扱いが異なり正確な比較は不可能だが、参考として分析に加えた。しかしながら、他のクラスも含めて外部基準テストとはほとんど相関が無かった。

A大学クラス2の相関値は表4にまとめられているが、特筆すべきことは、クラス1と対照的に未習項目においては相関がなく（−.016）、既習項目において中低位の有意な相関（.411）が見られることである。

B大学クラス1の相関係数を表5で見ると、未習問題が2つの外部基準テストと中程度の有意な相関（.466、.511）を示しているが、既習項目、文字数と外部基準との相関はほとんど無いと言える。実施に半年以上の間隔がある2つの外部基準テストでかなり高い相関値（.713）を示したことは、興味深い結果と言える。

対照的に、表6のB大学クラス2においては、すべての相関が低く、有意な結果を示していない。2つの外部基準テスト間の低いマイナスの相関値（−.132）からは、クラス編成時とコース終了時でクラス内の学習者の英語能力の位置づけがまったく変わってしまったか、テストが適切に機能していなかったことが示唆される。

表3　定期試験の部門別得点、記述部文字数、TOEFLの相関（A大学クラス1）

	未習項目	既習項目	未習＋既習	文字数
未習項目	−	−	−	−
既習項目	−.035	−	−	−
未習＋既習	.638**	.747**	−	−
文字数	−.135	.052	−.050	−
TOEFL	.422**	.066	.312	−.021

*5％水準（両側）で有意　**1％水準（両側）で有意

表4　定期試験の部門別得点、記述部文字数、TOEFLの相関（A大学クラス2）

	未習項目	既習項目	未習＋既習	文字数
未習項目	−	−	−	−
既習項目	.275	−	−	−
未習＋既習	.624**	.923**	−	−
文字数	.028	−.071	−.046	−
TOEFL	−.016	.411*	.350*	−.021

*5％水準（両側）で有意　**1％水準（両側）で有意

表5　定期試験の部門別得点、記述部文字数、ベネッセ、TOEICの相関（B大学クラス1）

	未習項目	既習項目	未習＋既習	文字数	ベネッセ
未習項目	−	−	−	−	−
既習項目	−.155	−	−	−	−
未習＋既習	.270	.909**	−	−	−
文字数	.082	.137	.168	−	−
ベネッセ	.466**	.092	.286	−.062	−
TOEIC Bridge	.511**	.265	.460*	.134	.713**

*5％水準（両側）で有意　**1％水準（両側）で有意

表6　定期試験の部門別得点、記述部文字数、ベネッセ、TOEICの相関（B大学クラス2）

	未習項目	既習項目	未習＋既習	文字数	ベネッセ
未習項目	−	−	−	−	−
既習項目	−.058	−	−	−	−
未習＋既習	.306	.932**	−	−	−
文字数	.163	−.122	−.058	−	−
ベネッセ	−.269	−.087	−.180	−.258	−
TOEIC Bridge	.016	−.019	−.012	.137	−.132

*5％水準（両側）で有意　**1％水準（両側）で有意

3.4　定期試験で見た学習者の到達度の違い

外部基準テスト結果は2大学で比較ができないため、まず定期試験のうち2大学4クラスに共通に実施した問題について、学習者の到達度の違いを分析することとした。

表7は、一元配置の分散分析の結果をまとめたものである。分析した2つの定期試験の部門別得点及びその合計得点、文字数において、サンプル全体で有意な差が生じていることが示されている。

表7　定期試験結果の一元配置分散分析

		平方和	自由度	平均平方	F値
未習項目	グループ間	49.220	3	16.407	23.488**
	グループ内	89.409	128	.699	
	合計	138.629	131		
既習項目	グループ間	356.826	3	118.942	45.610**
	グループ内	333.803	128	2.608	
	合計	690.629	131		
未習＋既習	グループ間	16614.290	3	5538.097	68.832**
	グループ内	10298.589	128	80.458	
	合計	26912.879	131		
文字数	グループ間	10850484	3	3616827.9	30.578**
	グループ内	15140104	128	118282.061	
	合計	25990587	131		

** 1％水準（両側）で有意

表8は等分散を仮定しないTamhaneのT2でクラス別に多重比較した結果をまとめたものである。重複するグループ比較の結果（例、A大学クラス1とクラス2、A大学クラス2とクラス1の比較）は、正負の関係が逆になるだけで絶対値の大きさは同じであるため、提示していない。最小有意差法などによる他の多重比較の統計値でも同様の結果を示したが、スペースの関係で割愛する。すべての比較において、異なる大学間での学習者の得点や文字数に有意な差は見られたが、同一大学のクラス間では、限られた項目にしか有意差は確認できなかった。B大学においては未習問題と既習問題の正解数合計と定期試験全体の得点において5％水準で1と2のクラス間に有意な差は確認できたが、未習、既習項目別の部門点や文字数で有意差は確認できなかった。A大学においては、部門点ではいっさい有意差は生じず、相関分析で定期試験とも外部基準テストとも有意な関係が見られなかった文字数においてのみ、1％水準でクラス間に有意な差が見られた。

表8　等分散を仮定しない多重比較（TamhaneのＴ２）

従属変数	比較クラス		平均値の差	標準誤差	有意確率
未習項目（４問）	Ａ１	Ａ２	－.3080	.20616	.599
	Ａ１	Ｂ１	.8719**	.24694	.005
	Ａ１	Ｂ２	1.1856**	.22345	.000
	Ａ２	Ｂ１	1.1800**	.19169	.000
	Ａ２	Ｂ２	1.4936**	.16030	.000
	Ｂ１	Ｂ２	.3137**	.21017	.599
既習項目（８問）	Ａ１	Ａ２	.0433	.28698	1.000
	Ａ１	Ｂ１	2.5408**	.43887	.000
	Ａ１	Ｂ２	3.8945**	.39950	.000
	Ａ２	Ｂ１	2.4975**	.43060	.000
	Ａ２	Ｂ２	3.8512**	.39040	.000
	Ｂ１	Ｂ２	1.3537	.51251	.062
未習＋既習（12問）	Ａ１	Ａ２	－.2647	.36375	.978
	Ａ１	Ｂ１	3.4127**	.48072	.000
	Ａ１	Ｂ２	5.0801**	.44912	.000
	Ａ２	Ｂ１	3.6774**	.46105	.000
	Ａ２	Ｂ２	5.3448**	.42801	.000
	Ｂ１	Ｂ２	1.6674*	.53101	.016
自由記述問題の文字数	Ａ１	Ａ２	343.7183**	104.20273	.010
	Ａ１	Ｂ１	672.2268**	92.17292	.000
	Ａ１	Ｂ２	722.9787**	88.61492	.000
	Ａ２	Ｂ１	328.5085**	71.65831	.000
	Ａ２	Ｂ２	379.2604**	67.01993	.000
	Ｂ１	Ｂ２	50.7519	46.14443	.856

*５％水準（両側）で有意　　**１％水準（両側）で有意

表９は、各大学の２クラス間で定期試験全体の得点分布を比較した結果を示すものである。表２の外部基準テストの結果と同様にＡ大学の２クラスには有意差は出ず、Ｂ大学においてのみ有意差が生じた。定期試験の後半の得点構成、内容が異なるため、４クラス間のテスト全体のデータを使った直接比較はできなかった。

表9　定期試験の得点分布の差

	Ａ大学定期試験		Ｂ大学定期試験	
	クラス１	クラス２	クラス１	クラス２
平均点	81.765	84.237	59.226	49.034
分散	182.064	98.564	213.847	126.392
被験者数	34	38	31	29
ｔ値	－0.876		3.038**	

**１％水準（両側）で有意

4．考察

1）定期試験と外部基準テストとの相関

今回の分析では、B大学のクラス2を除いて.4の中低位から.5の中位の有意な相関値が観測された。外部基準との相関による併存的妥当性は.5〜.7の数値が最も一般的とAlderson, *et al.*（1995）は述べているので、B大学のクラス1を除いては、併存的妥当性が十分な水準にあったとは言えない。A大学の場合、定期試験の総合点に、認知言語学的な見地から語法を説明する問題が含まれるが、Purpura（2004）やAlderson（2000）が言語学的用語の知識（metalinguistic knowledge)は言語能力とは切り離して考えるべきだと主張するように、このような問題は純粋に英語能力の到達度を測定する内容とは言えず、併存的妥当性を下げる要因となった可能性がある。B大学のクラス2の場合は、2つの外部基準テスト間での相関も低く、外部基準が適切に機能していなかった可能性もある。今回は、テストの内部均質性（internal consistency）や再テスト法（test-retest reliability）による信頼性検証を行うためのデータも揃わなかった。定期試験の信頼性、妥当性検証のデータ収集と同時に、クラス編成や成績認定の目的などで使用する外部基準テストは、テストの知名度に惑わされず、学習者の習熟度に適応した水準のものを選定することが重要である。

2）定期試験の成績の差

2大学4クラスで共通実施した定期試験の項目群においてA大学とB大学の間で有意差が確認された。少なくとも分析した部分においては両大学で到達度評価に成果の差が見られたことになる。A大学2クラスでは分析した定期試験項目群と定期試験全体で有意差は出ず、到達度評価においてこれら2クラスに成果の差は認められなかった。B大学2クラスでは、定期試験全体と前半部の単語補充選択問題の既習・未習問題の合計得点の分布において到達度に差が生じた。

今回の研究の限界の一つに、得点構成や問題内容の違いのため、定期試験後半部を分析できなかったことがある。しかしながら異なる能力層の集団に完全に同一の定期試験問題を実施したり、自由記述式の問題を排除して、客観的分析が容易な‘all or nothing’的な多肢選択式の問題だけで試験問題を構成するのは、教育的に有益な方策とは言えない。

今後の研究では、むしろ異なる種類の自由記述式項目群を含む複数のグループを利用することを考慮すべきであろう。それには、異なるテストの等化（test equating）手続き（Norizuki, 2004）や、部分点をつけた項目に対するステップ難易度境界値（step difficulty thresholds）の算出（Norizuki, 2000)を可能にする項目応答理論の応用も検討に値する。

3） 日本語表現力と英語能力の関係

4クラスの授業を担当した梅本が2大学の学生に感じた違いは、英語力だけでなく日本語表現力の違いや表現することに対する意欲であったと言える。今回の研究では、テストのボーナス点に含める内容として、おもしろい情報を記述する文章の文字数を唯一の変数として分析に含めたが、2大学の定期試験は問題や得点構成、与えられた筆記スペースに

よる心理的な影響の差など、厳密には大学間で直接比較できない要素を多く含んでいた。このような制約を踏まえた上で、大学間のグループに文字数に有意な差が出たことを認識する必要があるが、最も注目される結果は、分析したすべての英語能力テストの変数に見られなかった有意差が、この文字数においてA大学2クラス間で確認できたことであろう。文字数がすべてのクラスにおいて英語能力変数と有意な相関になかったことから、今回の研究では日本語表現力や表現することへの意欲と英語能力との関係は確立できなかったが、今後より明確な実験デザインを構築して、より広範な角度から、言語能力の構造や言語学習の発達的プロセスについて分析を進めていくことが重要である。

4）複合的な性格の英語テスト

かつて言語能力は単一の基底能力に集約されると主張した言語学者たちがクローズやディクテーション等が最も効果的なテスト法だと位置づけていた時期があったが、現在では、言語能力は複合的な技能で単一なテストで測定することはできないという考え方が言語テスト学者たちの共通の認識と言ってよい。テスト問題の形式やテストの目的、受験者の志向など、様々な要因が言語テストのあり方を巡って議論される時代になった。

たとえば今回の研究で、定期試験の後半部に位置づけられながら、大学間で得点構成が異なるうえに、個別項目や項目群の応答データが集計できなかったため、分析から除外したものに英文和訳の問題がある。この種の問題形式の信頼性・妥当性に一定の支持をする研究（Ito, 2004）がある一方で、訳読問題への依存が教育に及ぼす負の波及効果（negative backwash effect）を危惧する考えもある。教育の目的に応じて、最も適切な問題形式を使って、更なる学習継続の意欲を喚起させるテストが作成されることが望まれる。

今回の研究で注目された現象の一つに、未習項目群、既習項目群と外部基準テストとの相関があった。定期試験の主目的が到達度評価であると考えたときに、到達度の評価領域が運用能力を測定する外部基準テストの目標対象とは独立するものであれば、非常に高い相関でなくても問題はない。Hughes（2003）は、到達度テストの目標がコース目標に基づくものであって、実際のコース内容に拘束されないことを勧めるが、これは未習項目をより多く使用することを支持するものである。完全な未習事項でなくても文脈や表現を少しだけ変えるだけでも十分に意味がある。

英語テストの大半の研究が、教育現場から遠く離れたところで行われている印象を持っている教師も少なくないが、教師の現場であるクラスにおけるテストの研究開発に、もっと関心が注がれる必要がある。今回の研究結果から提案できる将来の研究指針の一つは、既習問題を授業で扱った内容を一部修正する形で提示し、より応用度の高い問題を未習問題として提示し、さまざまな問題形式と、相応の問題数で外部基準テストと比較することで、クラステストや到達度のあり方に識見を与えることである。

5．まとめ

今回の研究は授業プログラム終了後に残された評価データを使用しての分析であり、明示的な言語能力やテスト理論に立脚して、計画的に収集したデータに基づく分析ではない。同じ教育施設で指導していても、毎年変化していく学生の特徴に戸惑いながら、新たな教

育や評価方法を模索することは、教師なら誰でも経験することだが、異なる大学で異なる気質の学生に接し、彼らの英語能力やその他の能力、学習への取り組みにおける差異を強く感じるとき、より効果的な教育指導を導くために、その現象の本質を少しでも解明したいと願望するのは教師の本能とも言える。

本稿の冒頭付近で掲げた研究課題の多くには、明確な回答が得られなかったが、直感で感じたこととそうでないことがさまざまな数字として現れ、更なる探究心に駆られる結果となった。データを参考にして、教材の選定や指導法の改善についても柔軟に対応し、研究の成果を教育に還元させていく努力が大切である。

今後も英語力の分析を通じて、進んで自らの教育実践の向上に努めるだけでなく、偏見や客観的な根拠に欠ける類型化された解釈に流されない英語教育の確立を目指して、クラステスト分析の意義を実証していきたい。

参考文献

Alderson, J.C. 2000: *Assessing Reading*. Cambridge University Press.

Alderson, J.C., Clapham, C. and Wall、 D. 1995: *Language test construction and evaluation*. Cambridge University Press.

Brown, H.D. 2004: *Language assessment: principles and classroom practices*. Pearson Education.

Hughes, A. 2003: *Testing for language teachers* (*Second Edition*). *Cambridge* University Press.

Ito, A. 2004: Two types of translation tests: their reliability and validity. *System* 32(3)、 395-405.

清川英男・鈴木純子・浜岡美郎. 2003: 『英語教師のためのExcel活用法』 大修館

Norizuki, K. 1998: Research issues on university English language testing.『環境と経営（静岡産業大学経営学会)』第4巻第2号、 199-208.

— 2000: The nature of vocabulary knowledge in EFL assessment. 『環境と経営（静岡産業大学経営学会)』第6巻第2号、 301-310.

— 2004: In search of new dimensions for readability for Japanese learners of English.『静岡産業大学国際情報学部紀要』第6巻、 167-180.

Purpura, J.E. 2004: *Assessing Grammar*. Cambridge University Press.

Weir, C. 1993: *Understanding and developing tests*. Prentice Hall.

梅本　孝　法月　健

資料 1

英語I 和 総合英語B23　　A大学の代表的な解答例

-0　-9　-0

学籍番号：[redacted]　氏名：[redacted]　得点：91

空欄に入る適切な語を選びその記号を下線の上に書きなさい。

1. All visitors must go through a security ---B--- before entering the factory.
   (A) balance
   (B) check
   (C) mark
   (D) view

2. ---D--- the researchers all used special instruments to measure the samples, they could not agreed on the results.
   (A) Despite
   (B) Regarding
   (C) In spite
   (D) Although

3. Even before it was issued, Mr. Lee ---B--- responsibility for any inaccuracies found in the report.
   (A) claims
   (B) claimed
   (C) claiming
   (D) will claim

4/4

4. Due to his unconventional approach to leadership, the mayor of Guelph is ---A--- described by the local media as a nonconformist.
   (A) frequently
   (B) frequented
   (C) frequency
   (D) frequent

—

5. The international airport is located about sixteen miles ---C--- from the shopping district.
   (A) far
   (B) beside
   (C) away
   (D) remote

8/8

6. Each invoice should be labeled clearly and placed in the ------- drawer for easy reference.
(A) accurate
(B) approximate
(C) approximate
(D) agreeable

7. For two companies to be productive, when negotiating both groups involved must ------ a mutually favorable solution.
(A) understand
(B) solve
(C) reach
(D) arrive

8. Many employees, especially mid-level white-collar workers, have been ------ by widespread downsizing.
(A) decentralized
(B) deregulated
(C) demolished
(D) demoralized

9. Mr. Cabella is supervising the installation of the computer-controlled machines that will ------ the assembly line and reduce labor costs.
(A) conform
(B) automate
(C) appoint
(D) uplift

10. Because of the ------ economic outlook, the outlay for expansion will be smaller this year.
(A) disappoint
(B) disappointing
(C) disappointment
(D) disappointed

11. Angel Investments has announced that all of its clients are ---- for a complimentary membership at a nearby fitness club.
(A) possible
(B) dependable
(C) eligible
(D) allowable

C

12. The data analysis revealed various ---- that called for additional research.
(A) confidences
(B) inconsistencies
(C) ineptitudes
(D) endorsements

B

13. Quark Report is a journal ---- monthly by the maker of Super Computers and its accompanying software products.
(A) distributes
(B) will distribute
(C) are distributed
(D) distributed

D

— 0

訳しなさい。

14. When the manager went over the accounting figures, he found several serious discrepancies.
マネージャーが会計帳簿を調べた時、いくつかの重大な矛盾をみつけた。

15. The owner of the car left in the parking lot cannot be located, so it will have to be towed away.
駐車場に残された車の持ち主が見つからなかったので、それはけん引されることになるだろう。 —2

16. Employees are to notify their immediate supervisors of all absences.
雇われている者は自分の直属の上司に全ての休みのことを知らせる予定だ。 —1

3

17. Any redistribution of any information in this Website is expressly prohibited unless written authorization is granted by the publisher of Indicant.Net.

このウェブサイト上の全ての情報の権利は インディカントネット出版社により　4

全体の意味を意訳して述べなさい。（１８か１９かどちらかを選択し、選択した問題のみ、解答しなさい。）

18. Is it possible to know what the future holds for us? The Creator of mankind does have a plan for us, and He reveals it to us through an annual cycle of festivals described in the Scriptures.

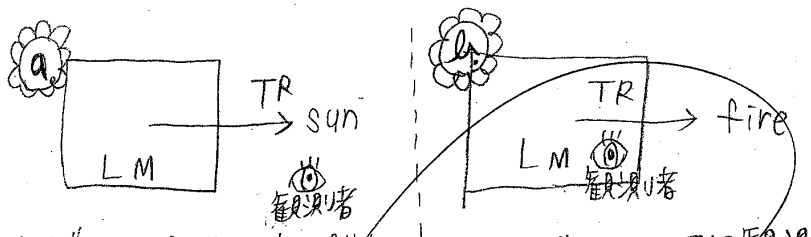
19. Dr. Anderson received his bachelor's and master's degrees from Moody Bible Institute and his Doctor of Philosophy in Sociology and Religion at Oxford Graduate School and has advanced to the level of fellow in the Oxford Society of Scholars.

アンダーソン博士は 学士とマスターの地位をムーディバイブル十字会から、そして オックスフォード大学院で社会学と宗教学の博士号を授かり、オックスフォード学会 のフェローの地位になった。　−2

20. 次の二つの文はどちらにも out が使われているが、意味は逆になる。どうしてそうなるのかを分り易く説明せよ。その際に必ず何らかの絵を描き図示しなさい。

(1) a. The sun is out.

b. The fire is out.



aの例では、LMの外に観測者の視点があり、LMからTR・Sunが out(出ていく)ことで、TR・Sunが観測者の視界・認識領域に入ってくる。観測者にとっては、TRが"現れた"ことになる。

bの例ではLMの中に観測者の視点があり、LMからTR・fireが out(出ていく)することで、TR・fireが観測者の視界・認識領域から出ていくのである。観測者にとっては、TRが"消滅した"ことになる。

4

~~33 × 78 = 2574~~
~~27 × 44 = 1188~~
1743

21. 自分が有する丸秘情報を分り易く順序だてて論述しなさい。 /8

今回は、豚・鳥・牛のお肉特集です★☆★
よく運動し、よくタンパク質を摂り、冬の寒さにも風邪にも負けない体で、
春を待ちましょい♡

秘伝の角煮♡レシピ♡

所要時間 30時間半 でも放置するのみ♪

冬、寒いですね。角煮、食べたくなりますよね。
というわけで角煮の作り方伝授しちゃいます☺ 時間はかかるけど
手間はかからないので是非作ってみて下さい!!! 之 こう見えて、小里子
なにげ料理好きなのです♪ 美味しいですよ〜♡

材料(4人分)　塩 — 適量
- 豚バラ肉かたまり — 600g
- ねぎの青い部分 — 2〜3本ぶん
- しょうが — 2かけ
- ゆで卵 — 4個
- しょうゆ — 80cc
- 酒 — 大さじ3
- 蜂蜜(砂糖もOK) — 大さじ3〜5

① 塩をすりこんだ豚バラ肉を脂が下になるように、強火で熱したフライパンに入れる。脂がとけて出てくるのでキッチンペーパーでふきとりつつ、まんべんなく焦げめをつける。
身と脂がきれいに層になってしまるのがいい
強火

② 焦げめをつけた豚バラかたまりを熱湯につける(→余分な脂をとる)
HOT湯

③ ねぎの青い部分、しょうが薄切り、肉をひたの水で煮る。強火。沸騰したら弱火にする。アクを取る。コトコト2〜3時間煮る。弱火!!(水がなくなったら足す)
強火 → 弱火

④ 肉がやわらかくなったら火を消し、そのまま放置。冷めると水面に脂のかたまりが浮いてくるのでさっぱりめにしたかったら取り除く。しょうが・ねぎを取り除く。肉を好きなサイズに切る。
※ゆで汁は捨てちゃダメ!!

⑤ ゆで汁にしょうゆ・酒・蜂蜜・ゆで卵・切り分けた肉を入れ、落とし蓋をして中〜弱火でコトコト煮る。(私は1時間くらい煮る)
しょうゆ 酒 蜂蜜 落とし蓋
中〜弱火

⑥ 煮汁が煮つまってきたらで仕上がり♡☺
※いったん冷ましてまた火を通すと一層味がしみこみます。

上に白髪ネギのせるとおいしい♡
パンに材料いろいろと一緒にはさんでもおいしい♡♡♡
パン

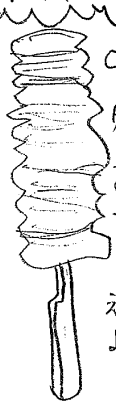
次はバイトの経験をふまえたメニューの紹介です☺

とりでん

焼き鳥と釜飯のお店

かれこれ1年勤めになるでしょうか…
牛角系列のチェーン店です。
炭を使っているお店と残念ながら
ベルトコンベアのお店があるのですが
天のン池袋店、町田店は美味!!!

オススメ・串物



せせり

(800)

肩ロス・すなみ

首の部分の肉で
ジューシー♪
味としてはももに
近いの私はもも
よりも好き♪

ぼんちり

(900)

お尻の部分の
肉だけあって
脂がのってる☺
1羽から1つしか
とれない稀少
品なんですよ!
みそ味オススメ。

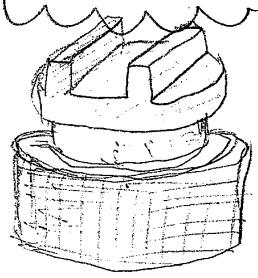
(1000)

もちもちチーズのベーコン巻

(1000)

焼き鳥としては
邪道ですが…
チーズともちのとろける
食感とベーコンは
逆らえません!!!
おいしいっっ!!

オススメ・釜飯



1 魚圭といくらの釜飯

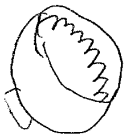
(1100)

定番?でもキングでずトップです。
とりでんのいくら、おいしいんです!だから
いくらの釜飯も美味しいけど、どうせなら
魚圭も食べちゃって下さいな☺

2 パリパリごぼうの五目釜飯

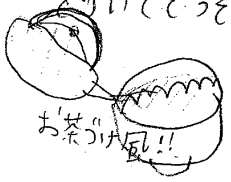
ふつうの五目釜飯より段ちがいで美味
しい!ふつうの五目釜飯には煮ごぼうが
入るのですが、パリパリごぼうには
揚げごぼうが入るのです!これが、ほんと
に美味しいッ!!!ほんとにほんとに!
ごぼう嫌いの私の友も絶賛!!

(1200)



よそって
1. そのまま
どうぞ

2. 上からかつおだし
かけてどうぞ



3 イカ明太バター

とりでんの明太子もおいしい! あとバターとの
競演はどやばいですのうんうん。

(1300)

オススメ・その他メニュー

・アボガドサラダ・
揚げたワンタンの
皮の中に具が入って
るのですが、皮をく
ずくして具と一緒
どーぞ!!!

・デザート・
パリパリ最中・
最安値のデザート
ですが、味はいち
ばん!!パリっパリ
最中の中にアイス・
あんこ・白玉入りで

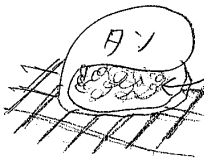
(400)

←中身をひろ
がらとじてめし
が出

やめたけど…

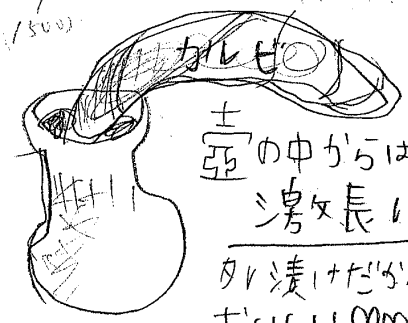
余々々苑 モー 2,000円です!!!

・ねぎタン塩・



味つけした
ねぎみじんを
蒸し焼きに
するのです

・つぼ漬カルビ・



自分で
チョキン
して下さい。

壺の中からは
激長いカルビ!!!

タレ漬けだから味もしみこんでて
おいしい♡♡♡♡

玉ねぎ、ピーマン、えび等も一緒に
漬けてあるから宝探しみたいで
たのしい。
なんと栗もとってる

※余々々苑は
デザートに
アイスをサービス♪

最後に…

★ 英語でinにあたるものは
ドイツ語でもinです★

たくみが「梅と飲んで楽しかったー」とゆってたので
主たさそってあげて下さいね!!
1年間、お世話になりました☺ありがとうございます!

(1793)

資料2

英語IB 木1　1/29/'04　　B大学の代表的な解答例

−35　−17　−52　[redacted]

学籍番号：[redacted]　氏名：[redacted]　得点：

空欄に入る適切な語を選びその記号を下線の上に書きなさい。

1. All visitors must go through a security ------ before entering the factory. B
   (A) balance
   (B) check
   (C) mark
   (D) view

2. ------ the researchers all used special instruments to measure the samples, they could not agreed on the results. D
   (A) Despite
   (B) Regarding
   (C) In spite
   (D) Although

3. Even before it was issued, Mr. Lee ------ responsibility for any inaccuracies found in the report. A
   (A) claims
   (B) claimed
   (C) claiming
   (D) will claim

4. Due to his unconventional approach to leadership, the mayor of Guelph is ------ described by the local media as a nonconformist. B
   (A) frequently
   (B) frequented
   (C) frequency
   (D) frequent

2/4

5. The international airport is located about sixteen miles ------ from the shopping district. C
   (A) far
   (B) beside
   (C) away
   (D) remote

3/8

1

6. Each invoice should be labeled clearly and placed in the ---A--- drawer for easy reference.
   (A) accurate
   (B) approximate
   (C) approximate
   (D) agreeable

7. For two companies to be productive, when negotiating both groups involved must ---B--- a mutually favorable solution.
   (A) understand
   (B) solve
   (C) reach
   (D) arrive

8. Many employees, especially mid-level white-collar workers, have been ---A--- by widespread downsizing.
   (A) decentralized
   (B) deregulated
   (C) demolished
   (D) demoralized

9. Mr. Cabella is supervising the installation of the computer-controlled machines that will ---B--- the assembly line and reduce labor costs.
   (A) conform
   (B) automate
   (C) appoint
   (D) uplift

10. Because of the ---C--- economic outlook, the outlay for expansion will be smaller this year.
   (A) disappoint
   (B) disappointing
   (C) disappointment
   (D) disappointed

11. Angel Investments has announced that all of its clients are ------- for a complimentary membership at a nearby fitness club.
(A) possible
(B) dependable
(C) eligible
(D) allowable

12. The data analysis revealed various ------- that called for additional research.
(A) confidences
(B) inconsistencies
(C) ineptitudes
(D) endorsements

13. Quark Report is a journal ------- monthly by the maker of Super Computers and its accompanying software products.
(A) distributes
(B) will distribute
(C) are distributed
(D) distributed

訳しなさい。

14. When the manager went over the accounting figures, he found several serious discrepancies.

15. The owner of the car left in the parking lot cannot be located, so it will have to be towed away.

16. Employees are to notify their immediate supervisors of all absences.

17. Any redistribution of any information in this Website is expressly prohibited unless written authorization is granted by the publisher of Indicant.Net.

このウェブサイトの管理人の文書による正式な許可がない限り、どの情報のいかなる再分配もよく禁止されている。

全体の意味を意訳して述べなさい。（１８か１９かどちらかを選択し、選択した問題のみ、解答しなさい。）

18. Is it possible to know what the future holds for us? The Creator of mankind does have a plan for us, and He reveals it to us through an annual cycle of festivals described in the Scriptures.

私たちの未来がどうなるのか知ることができるのでしょうか？人間を造った創造主は私たちにとって計画がありまする。神は私たちにそれをはっきり露見されます。 −2

19. Dr. Anderson received his bachelor's and master's degrees from Moody Bible Institute and his Doctor of Philosophy in Sociology and Religion at Oxford Graduate School and has advanced to the level of fellow in the Oxford Society of Scholars.

20. 自分が有する丸秘情報を順序だてて論述しなさい。前期と同じ情報は認められません。（１０点満点）

4

藤枝郵便局の向いにある「まし美」というラーメン屋の一番人気の「ベトコンラーメン」は辛くてすごくおいしくてTVにもよくとりあげられます。「ベトコンラーメン」の名前の由来は、おやっさんがベトナム戦争の時、傭兵としてベトナムにいた時に、現地の人に食わしてもらったものから味のヒントをえて、名前はベトナムの通称からきている。

あと奥さんはベトナムからうめてきたベトナム人。

4

29 × 6 = 174