初年次教育における G-TELP を用いた到達度測定の有効性

A Study on the Validity of Assessment with G-TELP of Academic Achievement in the First Year Experience

中島 彰子（Akiko NAKAJIMA）
CELL 教育研究所（CELL Institute for Educational Development）

本学では、年々深刻化している学生間の英語力格差に対応するため、初年次必修科目「英語表現」において習熟度別クラス編成、到達目標の設定、授業内容の統一化など、様々な英語教育の工夫改善を図っている。本稿においては、到達度測定テストとして採用している国際英検 G-TELP を取り上げ、初年次学生約 800 名のテスト結果をもとに、本学において G-TELP が適切に到達度を測定するテストであるか、有効性の観点から分析する。

キーワード：初年次教育, 必修科目, 到達度テスト, G-TELP

1. はじめに

学生間の英語力格差が拡大している昨今、これまで以上に学生の英語力を客観的に評価することが難しくなっており、成績評価が大きな課題と考えられている。英語力を測定するテストは数多く存在するが、本学では国際英検 G-TELP（General Tests of English Language Proficiency）を到達度測定テストとして採用し、様々な取組みと融合させることで、より客観的に学生の英語力を評価するよう工夫している。本稿では、初年次学生を対象に実施したG-TELPの結果をもとに、本学においてG-TELPが適切に到達度を測定するテストであるか、有効性の観点から分析する。

2. 「英語表現」の概要

大手前大学で初年次必修科目として設定されている「英語表現」は、複数教員により 48 クラスが開講されている。対象となる初年次学生は約 800 名で、クラスサイズは1クラス20名以下の少人数制を採用している。

コースは S、A、B、C の 4 コースがあり、S コースのみ入学前のプレイスメントテストで上位 10%を選出するが、ほとんどの学生は A、B、C コースから自分に合ったコースを選ぶことができる。このコース編成の特徴は、大学側が学生を習熟度別に振り分けるのではなく、学生本人が意欲度に応じてコースを選択することである。そのため学生は自分の選んだコースに責任を持ち、学習を進めていくことになる。各コースには到達目標が設定されており、学生は自分の所属するコース目標に向かって何度も繰り返し課題に取り組むという仕組みになっている。

2.1. 授業形態

「英語表現」の主な教材として、市販のe-Learning教材を採用しており、授業内だけでなく課外自己学習でも使用している。外部のテストを導入することにより授業がテスト対策に偏る危険性があるが、本学ではテスト対策に焦点を当てるのではなく、あくまでも学習の目標到達度を測るテストとしてG-TELPを取り上げ、より客観的に学生の英語力を評価している。また、e-Learning教材における学習項目に合わせ、G-TELPの数あるテストバージョンの中から、本学の学生の到達目標度を測定するのに適切と思われるバージョンを選んでいる。

授業内で各担当教員はリスニング、カンバセイション、リーディングの3セクションを中心に、学習内容の概要や文法事項等を説明し、その後は学生が各自e-Learning教材に取り組む。学生は、本学が独自に作成した「学習記録シート」に記入しながら学習を進める。このシートは、到達度目標を「～できる」というCan-do形式で提示したものである。学生が教材に取り組んでいる間、教員は机間巡回をしながら個別指導を行う。授業の最後に、特に学生の理解度が低い部分を

重点的に解説する。学期末にはG-TELPの模擬テストを行い、学生の自己採点により弱点を把握させた上で、学期末テストを実施している。

2.2. G-TELP 概要

TOEIC や TOEFL、一般的な入学試験を始めとする相対評価（Norm-Referenced Test）のテストに対し、G-TELP は個人の相対的なスコアではなく、テスト素材内容に対する個人の修得度を測定する絶対評価（Criterion-Referenced Test）のテスト（森田ほか 1998）である。相対評価のテストでは、受験者の中から相対的にスコアの高い人を選ぶことができるため、クラス編成などに有効利用できるが、絶対評価のテストでは、一人ひとりの到達度や習熟度がわかるため、受験者は自分の英語力がどのようなレベルであり、どの項目を修得しているのかがはっきりわかるようになっている。

G-TELP はレベル 1 から 5 までの 5 段階のレベルを設定している。各レベルの試験項目は、技能別に「グラマー（Grammar）」、「リスニング（Listening）」、「リーディング＆ボキャブラリー（Reading and Vocabulary）」の 3 セクションから成る。レベルごとにあらかじめ基準が設けられており、その基準にそってつくられた課題を「タスク」としている。タスクごとの正答率をパーセンテージ表示したスコアレポートが受験後に提供されるため、受験者に有意義な学習指針を示すことができる。

3. 調査

3.1. 調査の目的

上述した相対評価のテストを採用している大学は多いが、本学の「英語表現」のように複数教員が担当し、かつ同時開講している科目であれば、成績評価の統一尺度としてG-TELPをはじめとする絶対評価のテストを導入し、客観的な成績評価システムを構築していく必要がある。本調査ではG-TELPが適切に到達度を測定する試験であるか有効性の観点から分析し、今後G-TELPが日本の英語教育における成績評価の標準化について重要なテストの一つになるであろうことを示唆する。

3.2. 調査方法

本学では 2008 年度、2009 年度の春学期に同じバージョンの G-TELP を使い、学期末テストを実施した。対象者は初年次学生約 800 名である。レベルは「形式的な表現方法を用いて、ネイティブと簡単なコミュニケーションができる」と設定されているレベル 4 を使用した。過去 2 年間の全セクションにおけるタスク別（表 1）正答率及びリスニングとリーディング＆ボキャブラリーの 2 セクションにおける出題形式別（表 2）正答率を調査した。なお、参考のために本学の結果を全国平均と比較しているが、テスト実施は団体受験として各大学に任されているため、他大学における団体受験の結果及び個人受験による結果との正確な比較はできないことをお断りしておく。

4. 調査結果

4.1. タスク別正答率

グラマーセクションでは、2008 年度（図 1）、2009 年度（図 4）共に「SIMPLE PAST」、「SIMPLE PRESENT」が全国平均を約 20 点上回った。これは e-Learning 教材において何度も学習した文法項目である。2008 年度、2009 年度における 4 タスクの正答率もほとんど差がなく、グラマーセクションは本学において到達度を測るのに適しているテストだと言えよう。

表 1　セクション別タスク一覧

Skill Area	Task/Structure	問題数
GRM	YES/NO QUESTIONS	5
	PERSONAL PRONOUNS	5
	SIMPLE PAST	5
	SIMPLE PRESENT	5
LST	PICTORIALS	5
	SING STATEM/QUES	5
	TRANSACTIONS	5
	INSTRUCTIONS	5
RDG	APPLICATIONS	5
	ANNOUNCEMENTS	5
	FACTUAL ACCOUNTS	5
	PERS CORRESPOND	5

表2　セクション別出題形式一覧

Skill Area	Question Information Type	問題数
LST	Literal	15
	Inferential	5
RDG	Literal	10
	Inferential	6
	Vocabulary	4

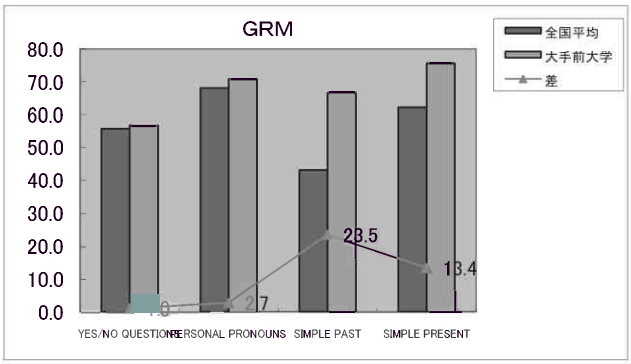

図 1　GRM タスク別正答率（2008 年度）

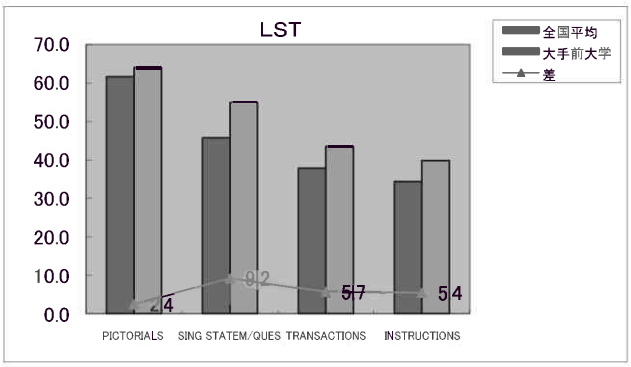

図 2　LST タスク別正答率（2008 年度）

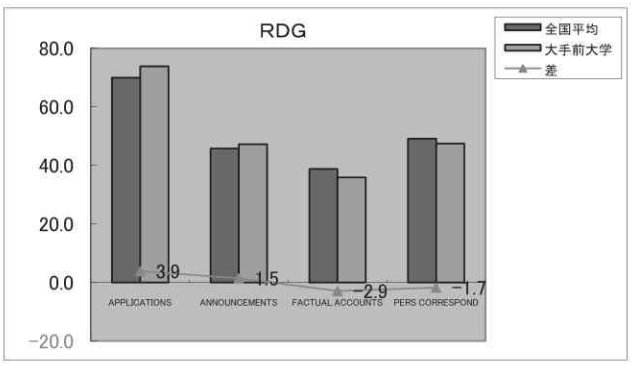

図 3　RDG タスク別正答率（2008 年度）

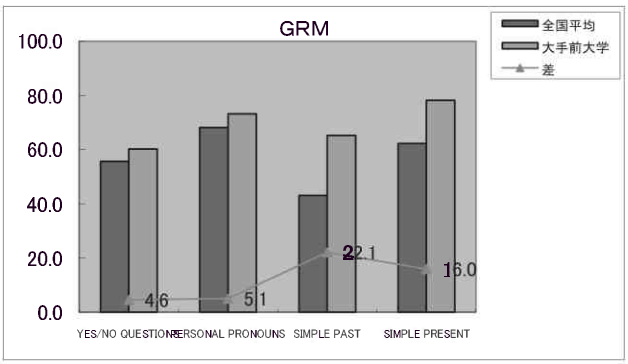

図 4　GRM タスク別正答率（2009 年度）

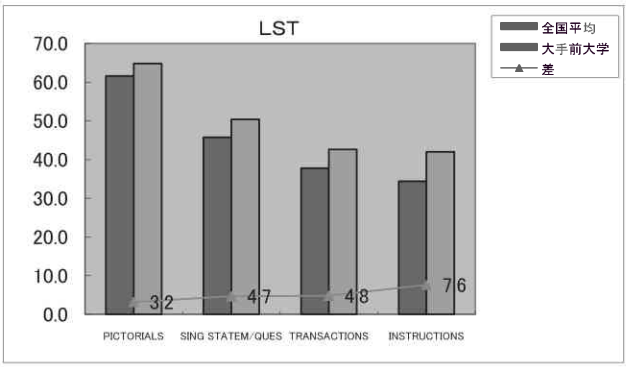

図 5　LST タスク別正答率（2009 年度）

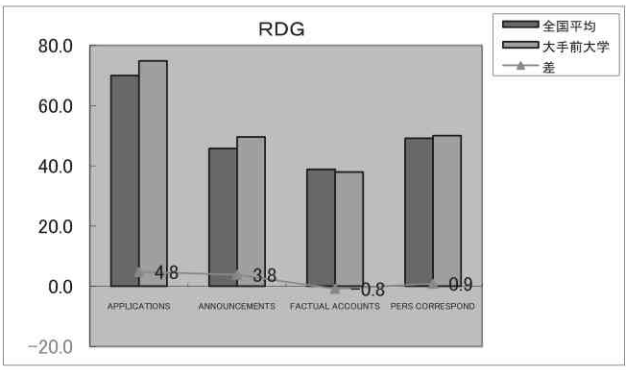

図 6　RDG タスク別正答率（2009 年度）

リスニングセクションは全てのタスクで全国平均を上回っているものの、グラマーセクションほど高い正答率ではなかった。もっとも全国平均との差が大きいタスクは、2008 年度（図 2）は「SINGLE STATEMENT/QUESTIONS」であり、2009 年度（図 5）は「INSTRUCTIONS」であった。なお、G-TELP のリスニングセクションは「TRANSACTIONS」と「INSTRUCTIONS」のタスクのみ問題の音声が 2 回流れるが、それ以外のタスクでは音声は一度しか流れず、非常に難しいセクションである。春学期は G-TELP のリスニングセクションに特化した授業を行ったわけではないが、それにも関わらずすべてのタスクで全国平均を上回ったのは、多くの学生が授業外にも e-Learning 教材を活用しながら学習を進めていたことの証左といえるだろう。

リーディング&ボキャブラリーセクションは 2008 年度（図 3）では「FACTUAL ACCOUNTS」と「PERSONAL CORRESPOND」、2009 年度（図 6）では「FACTUAL ACCOUNTS」のタスクが全国平均を下回っているものの、4 タスクの正答率もほぼ似通っていることから、適切に学生の到達度を測定できていると言える。

4.2. 出題形式別正答率

リスニング及びリーディング&ボキャブラリーセクションは質問のタイプを「Literal Information Question」と「Inferential Information Question」に分類している。「Literal Information Question」とは質問の音声や文字から、内容や情報が的確に把握されたかを問うものである。これに対し「Inferential Information Question」は、音声や文字から推測して答える問題であり、出題比率は表 2 の通りである。2008 年度、2009 年度（図 7、8）共にほぼ同じ正答率であり、全国平均ともほとんど差はなかったので 2008 年度の図は省略する。

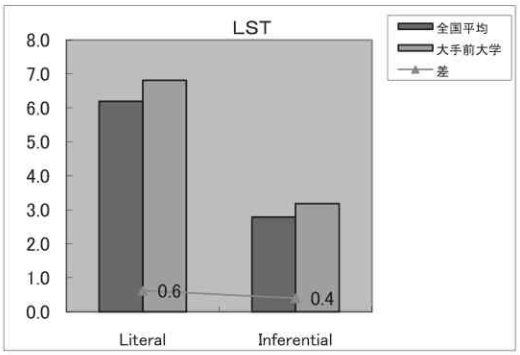

図 7　LST 出題形式別正答率（2009 年度）

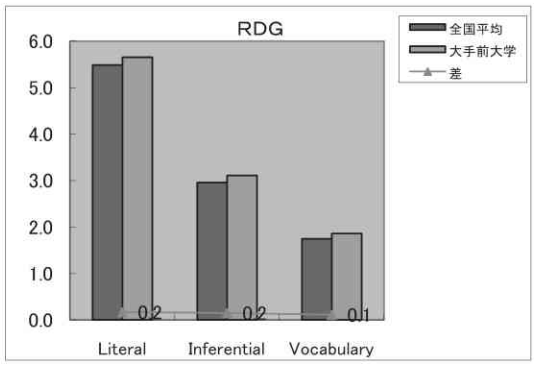

図 8　RDG 出題形式別正答率（2009 年度）

5. まとめ

本学で実施した過去 2 年間の調査結果がほぼ似通っていることから、本学において G-TELP は学生の到達度を適切に測定するテストだと考えられるが、今回の調査結果のみで判断できるものではない。本学のみならず日本の英語教育の発展のためにも、今後も G-TELP についてさらなる分析を進めていきたい。

謝辞

本稿における調査に御協力頂いたG-TELP日本事務局の福井拓氏、本学の初年次必修科目に関わっておられる全ての方々に心より感謝申し上げます。

参考文献

森田勝之, 野中泉, 江藤裕之（1998）『国際英検 G-TELP 受験のための公式ガイドブック』, 金星堂, 東京.

SUMMARY

In the present paper we focus on G-TELP, an international English proficiency test, which we employ in our university for assessment of achievement levels of students and analyze the results of G-TELP conducted in ca 800 first year students to figure out whether this test is an appropriate measure, in respect of validity of measurement, to assess achievement of English study in our university.

KEYWORDS: FIRST YEAR EXPERIENCE, COMPULSORY SUBJECT, ACHIEVEMENT TEST, G-TELP