

～GTEC と CEFR レベル関連付け調査～

平成 29 年 3 月 31 日

東京外国語大学大学院教授

根岸雅史

東京外国語大学大学院教授

投野由紀夫

株式会社ベネッセコーポレーション

鹿島田優子

一般財団法人 進学基準研究機構

岡部康子

1. GTEC とは

GTEC(Global Test of English Communication)とは、ベネッセコーポレーションによって開発されたテストである。中学・高校生向けの GTEC は、2016 年には、全国 1,770 校、約 93 万人に受験されており、多くの学校において、詳細なフィードバックをもとに指導改善や生徒への学習の動機づけのために活用されている。GTEC CBT は、大学入試での使用を目的としており、4 技能（リスニング、リーディング、スピーキング、ライティング）をコンピュータで受験する英語のテストである。GTEC CBT は、学習指導要領から想定される「日常的な言語使用場面」におけるタスクと大学での「アカデミックな言語使用場面」におけるタスクにより構成されている。

2. GTEC と CEFR レベル関連付け調査の背景

英語の 4 技能外部試験入試がさらに重要性を増し、注目されている。検定試験のスコアと CEFR レベルとの対応づけがスコア利用者（大学・受験者）にとって、大切な指標となっている。GTEC CBT については、実試験の開始前に CEFR レベル関連付け調査を行ったが、GTEC での CAN-DO 研究で過去長らく用いている CAN-DO アンケート結果に依拠するものであった。本調査では、実試験のデータに基づき、4 技能それぞれの閾値設定を行い、更にトータルスコアの閾値設定について再度検証を行うことを目的とした。

3. 調査手法

実施手順については、CEFR マニュアル（A Manual: Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment）の Standard Setting に基づいて行った。

調査手法としては受容技能のリーディングとリスニングでは、Bookmark Method（Council of Europe 2009）を用いた。IRT の項目パラメータ情報、困難度から各項目のスコアを算出し、0～350 点（満点）の範囲から 10 点刻みで項目を抽出した。その項目を易→難へと並べた小冊子（OID: Ordered Item Booklet）を作成して、CEFR の各レベルの閾値がどこにあたるかを判定した。スコアの高い問題については、GTEC CBT の問題を、低いものについては GTEC

の問題を使用した。

一方、発表技能のライティングとスピーキングの調査手法としては **Contrasting-Group Method** (Council of Europe 2009)を用い、ライティングの答案テキストデータ(GTEC は手書きでの解答のため、答案用紙のスキャンデータ)、スピーキングは録音された解答の書き起こしスクリプトを用意した。その解答データをスコアの低いものから高い順に並べたものについて、CEFR の各レベルの閾値に該当する受験者の解答結果を判定した。スピーキングについては、必要に応じて、解答音声も聞きながら判断を行った。

パネルは CEFR および英語の言語教育、教育測定に精通した研究者 6 名であった。またこの 6 名は全員、2015 年度の CEFR 関連付け調査にも参加している。調査は、事前課題と複数日にわたる集合型のワークショップ形式で行った。集合型のワークショップでは、① CEFR の閾値ごとに、CEFR 各レベルの閾値にいる人の英語力の目線合わせのための全体会、②その後、グループごとの分科会でそれぞれ閾値の想定を決定、③再度全体会にて3グループ合同での協議、という流れで行った。最初は CEFR の下位レベルの A1 と A2 の閾値を①～③の流れに沿って定め、その後順次、A2/B1, B1/B2, B2/C1 と上位の判定に移るというプロセスで、①～③の工程を繰り返した。

①目線合わせの全体会においては、閾値にいる人 (= Borderline Person) が2分の1の確率で回答できる英語力を Common European Framework of Reference for Languages: Learning, teaching, assessment の各技能および各 CEFR レベルのディスクリプターを読み、その英語力のイメージが6名のパネルの間でほぼ目線が合うまで議論を行った。②分科会においては、2名1組で、話し合いながら CEFR 閾値が合意に至るまで話し合った。迷う際には、その閾値前後の別の素材などもアイテムプールの中から参照し、慎重に閾値を決定した。③全体会においては、各グループの想定閾値をもとに、閾値決定の背景等と議論し合うことで、すり合わせていった。各問題の「タスク (設問内容) の難易度」と、タスク遂行のために読む、あるいは、聞く英語の「素材の難易度」とのバランスの中でのレベル判定に議論する場面もあったが、最終的に全パネルの合意のもと閾値を決定した。

結果は下記となった。

表 1. 技能別 CEFR レベル

	GTEC					GTEC CBT				
	Reading	Listening	Writing	Speaking	Total(4)	Reading	Listening	Writing	Speaking	Total(4)
C1						330	340	350	350	1370
B2	280	290	300	320	1190	280	290	290	300	1160
B1	220	220	240	280	960	220	220	220	220	880
A2	150	160	190	190	690	150	160	100	100	510
A1	上記A2未満					上記A2未満				

4. まとめ

2名で1組、3グループでの合議制という形で行われた **Bookmark Method**、**Contrasting-Group Method** は、CEFR レベルの感覚を共有するトレーニングをこなすことで、それぞれが大きくぶれずに判断できることがわかった。また各グループの想定閾値スコアがぶれた場合でも、前後のレベルの問題や解答パフォーマンスを参照することで、全員の認識がしだいに合ってくるなど、CEFR レベルに詳しい有識者の間での話し合いでテストの **Benchmarking** が有効に機能することも示された。

5. 参考文献

Council of Europe, (2009). A Manual: Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment

URL http://www.coe.int/t/dg4/linguistic/Source/ManualRevision-proofread-FINAL_en.pdf