

## TOEIC® Tests スコアの CEFR レベルマッピング方法について

※TOEIC® Tests は「TOEIC® Listening & Reading Test」と「TOEIC® Speaking & Writing Tests」の総称です。

本資料は、ETS(2008) "Linking English-Language Test Scores onto the Common European Framework of Reference: An Application of Standard-Setting Methodology"を IIBC が概略としてまとめたものです。

### 【レポート内容要約】

2006年10月16日～20日の5日間で、22名のパネリストがベルリンに集められ TOEIC Tests および TOEIC Bridge®スコアと CEFR レベルとのマッピング作業を行った。

### パネリスト構成

CEFR に精通し、且つ TOEIC Tests が活用されているヨーロッパ 10ヶ国の言語スペシャリストー英語教授法、英語学習、英語試験の専門家等々の中から選出された。

性別：女性 13 名、男性 9 名

国籍：ベルギー、フランス、ドイツ、ギリシア、ハンガリー、イタリア、マルタ、ポーランド、ロシア、スロバキア

\* 22 名中 5 名は TOEFL iBT と CEFR とのマッピング作業時のパネリストでもある。

### Methodology

Listening, Reading は Modified Angoff method、Speaking, Writing は Modified examinee paper selection method が用いられた。

### 作業の進め方

#### 1. CEFR レベルの内容理解(pp.7-8)

カットスコアの設定作業にあたり、まずは CEFR が設定するレベルの内容を理解・精通するため以下の作業を行う。

- 1) 集合前ー各パネリストには課題(Homework Tasks)が与えられる
- 2) 集合時ーCEFR で表されている言語スキルに関する広範囲に渡るディスカッションをし、各レベル別で最低限必要とされる能力を定義する

#### ・事前課題ー集合前(p.7 および pp.36-40)

CEFR レベル概要・定義を十分理解したうえで、CEFR レベル A1 から C2 それぞれに当てはまる最低限の英語力(〇〇ができるから〇〇レベルである)について、主要な特徴、指標を個々に記述しておく。考察すべき CEFR 表は、例えばスピーキングに関しては、全体的な口頭表現(Overall Oral Production)、全体的な口頭コミュニケーション(Overall Spoken Interaction)、ネイティブスピーカー話者の理解度(Understanding a Native Speaker Interlocutor)、会話(Conversation)、友人との話し合い(Informal Discussion with Friends)。その後、例えば C2 の最低レベルにあてはまる受験者にできて C1 トップレベルにあてはまる受験者にできないことは何かなどについて事前に考察しておく。

・ベルリン集合後の作業(p.8)

パネリストは、4 技能それぞれの A2、B2、C2 レベルの最低限のスキルを定義し、表を作成(事前課題で作成した各自の表と CEFR 表を参照)。

ここで定義するのは各レベルに合致する最低レベル(the candidate who has just enough skills to be at a particular level)の言語使用者についてである。CEFR レベルで表記されているのは各レベルの典型的な言語使用者のレベルなので注意が必要である。その後、全体的なパネルディスカッションを通して A2、B2、C2 各レベルの定義について最終的にパネリスト間で決定する。この A2、B2、C2 レベルの定義を用いて、A1、B1、C1 レベルの最低限のスキルについてパネルディスカッションを通して定義する(パネリストが話し合い最終的に定義したそれぞれのスキルレベルについては Appendix B pp.46-49 を参照)。

### カットスコアの設定

パネリスト全体で各レベルの定義を決定したあとに、実際にスキル別のマッピング作業にはいる。

Reading, Listening セクション—Selected-response sections (pp.8-10)

Reading: 100 問 (Raw Score 0~100) →最終的に 5-495 へ換算

Listening: 100 問 (Raw Score 0~100) →最終的に 5-495 へ換算

・Modified Angoff Method を採用(※問題の設問ごとに、CEFR の各レベルに相当する受験者が 100 人いるとして、正解する者の割合を各パネリストが算出し、全体の平均を合議に基づき判定基準とする分析方法)

・設問データは、100,000 人以上の受験者が受けているものを使用

・パネリスト(22 人)は事前にカットスコア判断過程について訓練を受け、過程を理解しているかを確認するために実際に判断する機会を与えられる。

・A2、B2、C2 を計 3 回、A1、B1、C1 を 1 回実施

1 回目の判断では、パネリストは、設問ごとに CEFR レベル別の受験者が正答できる確率を 0、5、10、20、30、40、50、60、70、80、90、95、100 の評価基準で出す。確率が高ければ、簡単な問題だと考えられる。この結果、各パネリストが推測するカットスコアが出される。パネルの平均カットスコアと最高、最低のカットスコアが出され、それについてパネリストはディスカッションをし、決定の論拠を共有する。

フィードバックとディスカッションの過程で、Item performance information (P+ values: あるテストフォームの受験者の平均スコア)の指数を共有し、ある設問が受験者の全体的な言語能力に反して比較できるほど難しかったり、異なるレベルの受験者にとって特に難しかったり、簡単だったりするということが分かる。

2 回目の判断をする前に、パネリストは他のパネリストの論拠と基準となる情報について考察する。2 回目の判断では、設問のレベルではなく、セクション全体のレベルについて決定する。2 回目の判断の A2、B2、C2 レベルの平均は、すでに作成されたセクションのスコア分布にあてはめられ、それぞれのレベルに分けられた受験者の割合と照らし合わせて議論する。その後パネリストは、セクションレベルの推奨カットスコアを変更する最後の機会を与えられる(3 回目の判断)。

A2、B2、C2 レベルの最終判断が発表され、その後パネリストは A1、B1、C1 レベルを組み込むよう指示される。

Speaking, Writing セクションーConstructed-response sections (pp.11)

Speaking (Raw Score 0~24)→最終的に 0~200 へ換算

Writing (Raw Score 0~26)→最終的に 0~200 へ換算

・Modified examinee paper selection method を採用

(※受験者の実際の解答を使用してカットスコアの設定を行う)

・Modified Angoff 法と同様、フィードバック、ディスカッション、データの提供などを行いながら、3 回で判断をする。

パネリストは採点基準表(TOEIC Speaking & Writing Scoring Rubrics)を確認し、さまざまなスコアの受験者の 11 の解答を聞くまたは読む。低いスコアから高いスコアの解答の順に並べられている。パネリストには 11 人の受験者の設問ごとのスコアをまとめた表が渡され、各 CEFR レベルの最低カットスコアを判断する。A2、B2、C2 レベルについて 3 回の判断をした後、A1、B1、C1 レベルを組み込む。2 回目以降の判断では、各レベルにとって難しすぎるあるいは簡単すぎると判断した場合は、スコアではなく N/A(該当なし)と回答することも可能。最後にセクションごとに、A1 から C2 の最大 6 レベルのカットスコアを提示する(該当スコアのないレベルも有り)。

#### TOEIC® Tests スコアにおける CEFR マッピングの修正

上記 2006 年 10 月に行ったマッピング作業の際に、Reading において C1 レベルにおけるカットスコアの合意がパネル間で得られず N/A とされた。但しこれは TOEIC Reading セクションで出題される問題が簡単すぎて C1 レベルまでの測定ができないと判断された訳ではなく、各パネリストが示したカットスコアが最終的に一つのスコアとして一定の合意に至らなかった。言い換えれば、TOEIC Reading セクションは C1 レベルまで測定できることについては、パネリスト間で合意されていた。

加えて TOEFL iBT において CEFR とのマッピングが修正されることとなり、TOEIC Tests についても再検証が行われた。その結果として、TOEIC Reading の C1 レベルにおけるカットスコアとして 455 と設定し、また Speaking および Writing の C1 レベルカットスコアを当初の 200 から 180 へ引き下げるといふ措置が取られた。

上記の作業を経て、現在以下のような TOEIC Tests スコアと CEFR レベルとの Mapping 結果が ETS より公表されている。

	Score Range	A1	A2	B1	B2	C1	C2
Listening	5-495	60	110	275	400	490	N/A
Reading	5-495	60	115	275	385	455	N/A
Speaking	0-200	50	90	120	160	180	N/A
Writing	0-200	30	70	120	150	180	N/A

以上