



Cambridge Assessment
English

ケンブリッジ英語検定 4 技能 CBT
(Linguaskill リンガスキル)



リスニング・リーディングの

トライアルに 関する報告書

2016 年 4 月



研究チーム
Jing Xu
Trevor Benjamin

ケンブリッジ英語検定4技能CBT (Linguaskill リンガスキル)



目次

| | |
|--------------------------|----|
| Linguaskill リンガスキル とは | 3 |
| Linguaskill リンガスキルのトライアル | 3 |
| トライアルの主な調査結果 | 4 |
| トライアルの結果 | 5 |
| 提案 | 10 |
| 対応策 | 11 |
| 付録 | 12 |

複数レベルの 適応型英語能力テスト

ケンブリッジ英語検定 4 技能 CBT (Linguaskill リンガスキル) とは

ケンブリッジ英語検定 4 技能 CBT (Linguaskill リンガスキル) とは、複数のレベルに分かれているコンピュータ適応型の英語能力テストです。テストのスコアは、ヨーロッパ言語共通参照枠 (CEFR) 1 の A1-C2 レベルで報告されます¹。

Linguaskill リンガスキルは、企業とその従業員、教育機関など、英語学習者の英語コミュニケーション能力レベルを把握する必要のある、全世界にわたる個人および組織向けのテストです。迅速、簡単に実施できるうえ、コスト効果も信頼性も高い結果が得られます。

トライアル

Linguaskill リンガスキルのリスニングとリーディングテストは、2016 年 2 月下旬から 3 月 31 日にかけて試行されました。このトライアルに参加した英語学習者の数は、合計 248 名にのぼりました。

トライアルの目的：

- テストスコアの正確性と信頼性を調査する
- テストの公正性を調査する
- 受検者の受検体験を把握する
- テストが学習者のニーズをどの程度満たすことができるかを評価する
- テストの仕組みに改善の余地があるかどうかを突き止める

テスト結果に関する報告事項は、次のデータに基づきます。

- **トライアルテスト：**問題、解答およびテストスコア
- **オンラインアンケート：**77% の参加者がアンケートに回答し、テストに関する全体的な印象と意見を提供



参加者やデータの収集、分析など、研究で採用された詳しい方法については、付録 (12 ページ) を参照してください。

¹ ヨーロッパ言語共通参照枠 (CEFR) とは、欧州評議会 (2001) が策定しているガイドラインであり、外国語学習者の言語能力を多様なレベルで表します。

www.cambridgeenglish.org/cefr

トライアルの主な調査結果 ▶▶

Linguaskill リンガスキルのテスト スコアは信頼性が高く、正確です。

トライアルの結果：

信頼性：リスニング、リーディング、および全体的なテストの結果は、それぞれ 0.92、0.94、および 0.96 と推算されています。

0.90 を超える信頼性係数は、優良とみなされます。

正確性（精度）：トライアルテストの約 90%（リスニングテストの 91%、リーディングテストの 88%）が正確性の目標水準に達しました。

• コンピューター適応型テストの受検経験は、参加者のテスト結果に影響していないようです。

このことから、受検者は説明がなくても Linguaskill リンガスキルのインターフェースを理解できることが読み取れます。

• 参加者の大半は受検体験を好意的にとらえており、Linguaskill リンガスキルのリスニングとリーディングテストの受検を肯定的に感じている割合は 60% を超えました。参加者の約 3 分の 1 はテストに関して中立的な立場でした。

• 約 91% の参加者は、リスニングテストの指示が明確であったと答えています。

• 参加者は、テストのインターフェース、視覚教材での画像の使用、興味深い題材、日常的な言い回しと関連性の高いテストの内容について、肯定的にとらえています。自己採点と言語学習も高く評価しました。



トライアルの結果

テストのスコアの信頼性と正確性

主な調査結果



テストスコアの信頼性

Linguaskill リンガスキルは、コンピューター適応型テストです。受検者がそれまでの問題にどれだけ適切に答えたかに応じて、その後のテスト問題が選択されます（受検者のレベルに応じて出題が調整されます）。

受検者が異なる項目セット（問題）を解くことから、信頼性を計算する典型的な方法、たとえばクロンバックの α 係数などは、適応型テストには採用できません。その代わりに、同様の尺度であるラッシュモデルに基づく信頼性測定が使用されます。

248人のトライアル参加者のデータに基づく、リスニング、リーディング、および総合テストのラッシュの信頼性の推算値は、すべて0.9を超えています。この推算値は、トライアル実施前のシミュレーションやケンブリッジの適応型テストの実績と一致しています。

表1. リスニング、リーディング、総合テストのラッシュ信頼性

| リスニング テストの信頼性 | リーディング テストの信頼性 | 総合テストの 信頼性 |
|------------------|-------------------|---------------|
| 0.92 | 0.94 | 0.96 |

テストスコアの正確性（精度）

受検者の語学能力の「真の実力」を完全に推算できるテストのスコアは存在せず、多少の統計的誤差を避けることはできません。こうした誤差は測定の標準誤差（SEM）と呼ばれます。ケンブリッジでは、その時点での受検者のテストスコアの68%が真のスコアから1 SEM以内、その時点での95%のテストスコアが真のスコアから2 SEM以内であると予想しています。

Linguaskill リンガスキルは、出題数でなく、精度が固定的なテストです。出題数は受検者ごとに異なりますが、精度（誤差）はほぼ一定であり、すべての得点のSEMは大体同じレベルに維持されます。

SEMの目標値は、0.44 ロジットに設定されています（リスニングおよびリーディングの両方）。ロジットは、受検者の能力を推算するために使用される統計単位です。受検者に結果を報告するためのCambridge English スケールと同じものではありません。

トライアルでは、リスニングテストの91%、リーディングテストの88%が目標SEMに達しました。

目標SEMに達しなかったトライアルテストは、次の2つのカテゴリに分けられました。

1. 最上位・最下位レベルの能力：このカテゴリの受検者は、全問（または、ほぼ全問）に正答または誤答したため、能力の推定値は、最上位または最下位階層に属します。こうしたケースは想定内であり、特に問題にはなりません。

- **リスニング：**3%の参加者が最上位または最下位層に属する（最上位層に属する参加者は8名）。
- **リーディング：**4%の参加者が最上位または最下位層に属する（最上位層に属する参加者は3名、最下位層に属する参加者は6名）。

2. 長さが最長：目標SEMに達する前に、出題数が最大に達した受検者カテゴリ。

- **リスニング：**目標SEMに達する前に、テストの6%が最大長に達しました。
- **リーディング：**目標SEMに達する前に、テストの8%が最大長に達しました。

このケースを詳細に分析すると、過半数は目標SEMを「かろうじて超える」精度（0.44というより0.5 ロジット未満）を持ち、CEFRの最上位または最下位レベル（C1以上、またはA1以下）で発生しています。

表2. 目標精度（SEM）に達しているテストの割合

| 目標SEMを達成？ | リスニング | リーディング |
|----------------|-------|--------|
| はい | 91% | 88% |
| いいえ – 最上位/最下位層 | 3% | 4% |
| いいえ – 最長 | 6% | 8% |

概説すると、このトライアルテストでは、目標レベルの精度は安定的に達成できる一方で、最上位/最下位レベル（CEFRのC2およびA1レベル）でのテストについては、改善のために一段と努力する余地があることが実証されました。

コンピューター適応型テストの受検経験はテストのスコアに影響するか？

主な調査結果

コンピューター適応型テストの受検経験がテスト結果に影響していないように見受けられることから、説明がなくとも、受検者はLinguaskill リンガスキルのインターフェースを理解できることがわかります。

コンピューター適応型テストの有効性を検討する際は、受検者のコンピューター知識・経験（コンピューターで言語テストを受検したことがあるなど）がテスト成績に影響するかどうか重要な観点となります。オンラインアンケートでは、Linguaskill リンガスキルのトライアル前にコンピューター適応型テストを受けたことがあるかどうかを参加者に質問しました。

アンケートの回答に基づき、参加者は（1）経験ありと（2）経験なしの2つのグループに分割されました。次に、この2つのグループのテストスコアが比較されました。以下の表3および図1では、2つのグループのリスニングおよびリーディングテストのスコアは、平均値と標準偏差がきわめて似ていることを示しています。

各グループのテストのスコアは、平均的に分布しています（図2を参照）。そのため、スコア分布の中心傾向を2つのグループ間で比較するために、ウィルコクソンの符号順位検定が実施されました。

この検定で、リスニングテストのスコア ($Z = .92, p > .05$) とリーディングテストのスコア ($Z = -.49, p > .05$) は、2つのグループ間で有意差がないことが示されました。コンピューター適応型（CB）テストの受検経験がテストの結果を左右していることはないように見受けられます。

表3. 2つのグループのリスニングとリーディングテストのスコアの記述統計

| 変数 | グループ | n | 平均値 | 標準偏差 |
|-----------|--------------|----|--------|-------|
| リスニングスコア | CBテストの経験（なし） | 75 | 142.04 | 24.84 |
| | CBテストの経験（あり） | 82 | 139.33 | 24.64 |
| リーディングスコア | CBテストの経験（なし） | 75 | 205.00 | 27.97 |
| | CBテストの経験（あり） | 82 | 204.00 | 27.45 |

図1 2つのグループのリスニングとリーディングテストのスコアのボックスプロット
(CBテスト：コンピューター版テスト、CBT)

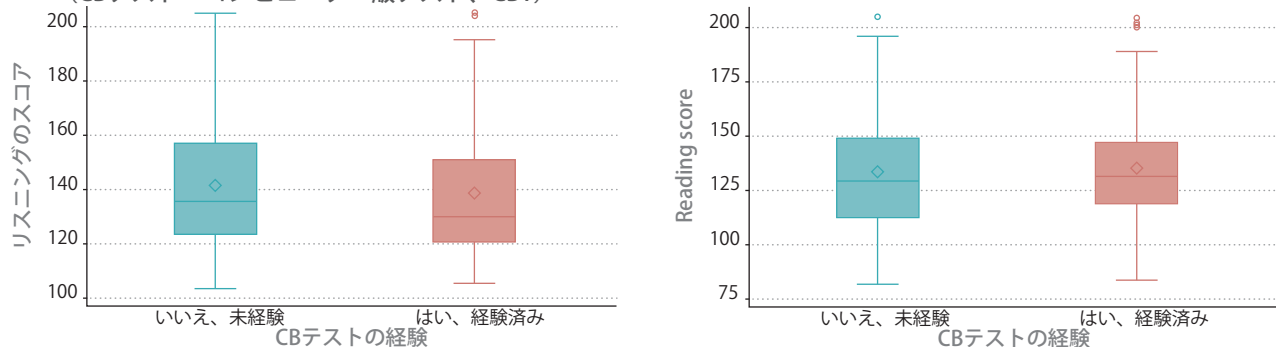
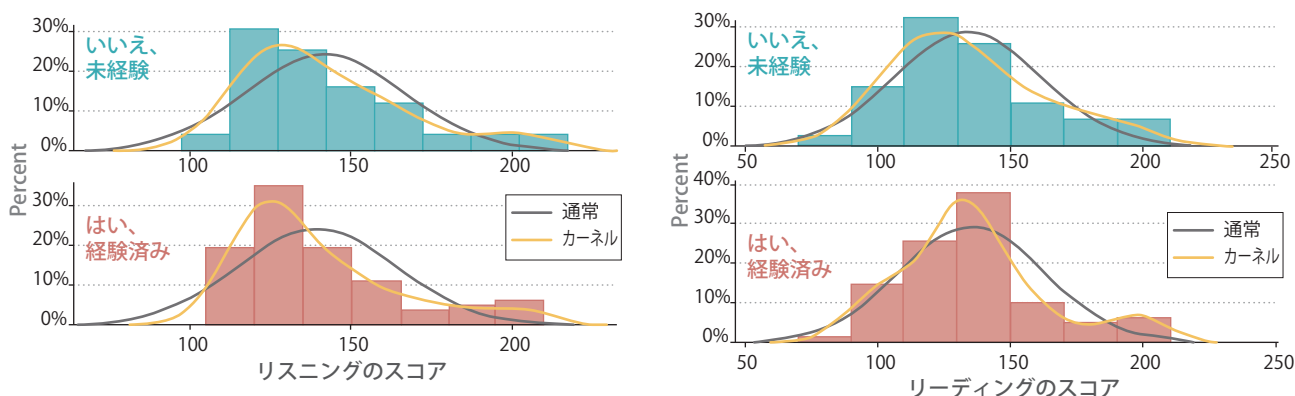


図2 2つのグループのリスニングとリーディングテストのスコアの分布



参加者にとって、受検は有意義な体験でしたか？

主な調査結果

リンガスキル のリスニングとリーディングテストの受検に関して肯定的な感想を抱く参加者の割合は60%を超えました。参加者の約3分の1はテストに関して中立的でした。テストに関して否定的な感想を抱く参加者数はごくわずかでした。約91%の参加者は、リスニングテストの指示が明確であったと答えています。

オンラインアンケートにおいて、参加者はテストの全体的印象を評価（非常に肯定的から非常に否定的まで）するよう求められました。大半の参加者は次のように、2つのテストに対し全体的に非常に肯定的な印象を抱いています。

- **リスニングテスト**：肯定的または非常に肯定的という回答は64.1%、中立的という回答は28.6%、否定的または非常に否定的という回答は7.3%でした。
- **リーディングテスト**：肯定的または非常に肯定的という回答は63.9%、中立的という回答は32.4%、否定的または非常に否定的という回答は3.7%でした。

オンラインアンケートには、テストに関するいくつかの記述に賛同するか、賛同しないかという質問もありました。

アンケート結果によると、図3で示すとおり、受検者の大半は受検体験を肯定的にとらえています。

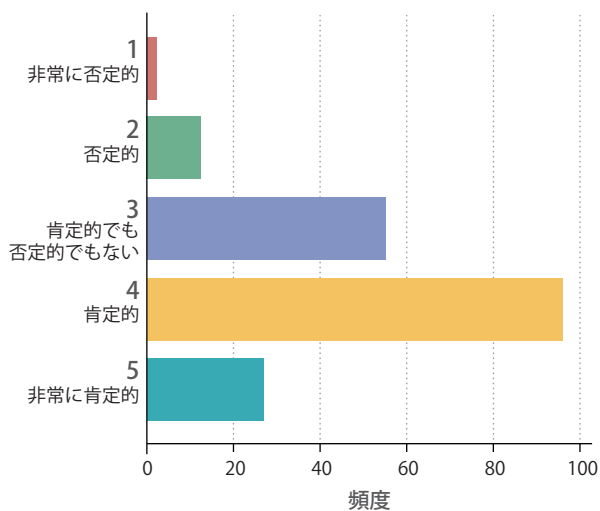
項目別では、表4および5で示すとおり、受検のしやすさ、出演者の会話の明瞭さ、テストでの英語能力の発揮しやすさといった点で、高評価が得られました。

対照的に、以下の分野での評価は低めでした。

- **リスニングテスト**：テストの指示の明瞭さ、2人の演者の声の区別しやすさ、演者のアクセントのわかりやすさ、テストの内容と日常生活や仕事との関連性
- **リーディングテスト**：テストの内容と日常生活や仕事との関連性

図3 リスニングとリーディングテストに対する参加者の全体的印象

リスニングの評価 (n=192)



リーディングの評価 (n=192)

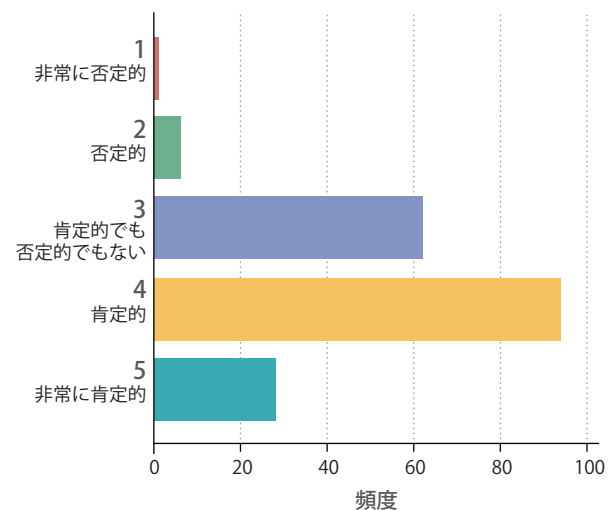


表4 リスニングテストについてのコメントに対する参加者の賛否

| リスニングテストについてのコメント | 同意・強く同意 (4または5) | 中立 (3) | 反対・強く反対 (1または2) | 回答数 | 平均値 | 標準偏差 |
|-------------------------------------|--------------------|-----------|--------------------|-----|------|------|
| コンピューターで試験に簡単に取り組めた | 81.3% | 15.0% | 3.7% | 187 | 3.89 | .66 |
| テストでの演者の声は、十分な音量で明瞭だった | 77.5% | 17.7% | 4.8% | 186 | 3.96 | .78 |
| テストでやらなければならないことが明確にわかった | 59.2% | 32.1% | 8.7% | 184 | 3.70 | .88 |
| テストの各タスクを完了するのに十分な時間があった | 73.6% | 19.8% | 6.6% | 182 | 3.94 | .89 |
| 演者が複数居た場合、それぞれを簡単に区別できた | 66.1% | 30.1% | 3.8% | 186 | 3.84 | .81 |
| 演者のさまざまなアクセントを簡単に理解できた | 62.4% | 30.1% | 7.5% | 186 | 3.66 | .83 |
| リスニングの各タスクには、日常生活や仕事での英語の用法が反映されていた | 66.4% | 27.8% | 5.8% | 187 | 3.76 | .86 |
| リスニングテストで全力を尽くした | 71.1% | 20.9% | 8.0% | 187 | 3.86 | .86 |
| 英語のリスニング能力をテストで発揮できた | 75.3% | 23.1% | 1.6% | 186 | 4.00 | .75 |

表5 リーディングテストについてのコメントに対する参加者の賛否

| リーディングテストについてのコメント | 同意・強く同意 (4または5) | 中立 (3) | 反対・強く反対 (1または2) | 回答数 | 平均値 | 標準偏差 |
|---------------------------------|--------------------|-----------|--------------------|-----|------|------|
| コンピューターで試験に簡単に取り組めた | 83.9% | 15.0% | 1.1% | 180 | 4.11 | .68 |
| テストでやらなければならないことが明確にわかった | 76.7% | 22.2% | 1.1% | 180 | 4.01 | .72 |
| テストの各タスクを完了するのに十分な時間があった | 85.6% | 13.3% | 1.1% | 180 | 4.13 | .67 |
| テストの文章には、日常生活や仕事での英語の用法が反映されていた | 68.9% | 27.8% | 3.3% | 180 | 3.87 | .82 |
| リーディングテストで全力を尽くした | 78.7% | 17.9% | 3.4% | 179 | 4.07 | .79 |
| 英語のリーディング能力をテストで発揮できた | 84.5% | 14.4% | 1.1% | 180 | 4.12 | .68 |

参加者が考える、テストの長所および限界とは？

主な調査結果



参加者は、テストのインターフェース、視覚教材での画像の使用、興味深い題材、日常的な表現と関連性の高いテストの内容について、好意的にとらえています。また、自己採点と言語学習も高く評価しました。約91%の参加者は、リスニングテストの指示が明確であったと答えています。

オンラインアンケートには自由解答形式の質問を3つ設け、テストに関して参加者の意見を募りました。

回答は、肯定的または否定的のいずれかに分類され、コードが付けられました²。参加者の意見は、表6と表7に例示されています。

表6 リスニング・リーディングテストについての参加者の肯定的意見

| | リスニングテストについての肯定的意見 | リーディングテストについての肯定的意見 |
|----------|---|---|
| インターフェース | <ul style="list-style-type: none"> 設計が優れたテストのインターフェース | <ul style="list-style-type: none"> インターフェースのレイアウト 画像を使用した視覚教材 明確なテストの指示 コンピューターでのリーディング体験 |
| 内容 | <ul style="list-style-type: none"> 興味深い話題 多彩な話題 テストの内容と日常生活との関連性 | <ul style="list-style-type: none"> 興味深い話題 タスクの難易度 短文のリーディングタスク |
| ラーニング | <ul style="list-style-type: none"> 自己採点、練習、言語学習の機会 | <ul style="list-style-type: none"> 自己採点、練習、言語学習の機会 |

表7 リスニング・リーディングテストについての参加者の否定的意見

| | リスニングテストについての否定的意見 | リーディングテストについての否定的意見 |
|----------|--|---|
| インターフェース | <ul style="list-style-type: none"> 音声開始時期の、操作手段や指示の不足 問題を読む時間が不十分 | <ul style="list-style-type: none"> 文章にマーカーを引いたり、メモを付けたりできない 画像の質 非表示のオプション |
| 内容 | <ul style="list-style-type: none"> 問題を読むために高度の読解力が必要 リスニングタスクの難易度（発声速度、アクセント、話題、語彙等） テスト時間の長さ | <ul style="list-style-type: none"> 語彙と頭字語の難易度 時間的制約 |
| 学習環境 | <ul style="list-style-type: none"> テスト環境での騒音 | <ul style="list-style-type: none"> インターネット接続の不良で起こるエラー |

注2：文法が整っていない回答は言い換えまたは修正され、反復的意見はまとめられ、さらに一部の回答はスペイン語から翻訳されました。

提案

この報告書の調査結果を基に、次の提案が出されています。

- CEFR レベル A1 および C2 向けに、より多くのテスト問題を作成する。
- テストに親しむ機会（チュートリアルや演習テストなど）を設けて、受検者の不安を軽減させる。
- リスニングテストの進め方に関して、受検者が自分でもっと管理できるようにするか、より多くの情報を提供する。
- 英語学習の初心者には難易度の高いタスクを課さない（強いアクセント、難しい語彙、一般的でない話題など）。
- 受検者が文章を読みながらマーカーを引いたり、メモを書いたりできるようにする。
- より質の高い画像を使用する。
- テスト環境を改善する（騒音、インターネット接続など）。



対応策

上記の提案を基に、ケンブリッジ大学英語検定機構では次の対応策を講じています。

- 特にレベル A1 および C2 に向けて、より多くのテスト問題を作成しています。
- デジタル式チュートリアルと練習用教材の開発プランを立て、受検者がテストのフォーマットやインターフェースに親しめるようにしています。
- 問題を読むために音声を停止する時間に関して、情報伝達を徹底するため指示を統一化して、受検者がリスニングテストをさらに自律的に進められるようにしました。
- トライアル後、電子アイテムバンクを2度見直し、不適切なテスト問題（強いアクセント、難易度の高い語彙、一般的でない話題など）を特定し、変更しました。
- すべてのタスクを見直し、不明瞭な画像は解像度を上げた画像に変更しました。
- ケンブリッジ大学英語検定機構のテスト管理者向けガイドラインへの準拠を試験センターに指示する一方で、テストの設定とインターネット接続の点検に関して代理機関に役立つようにユーザーガイドを作成しました。

付録

データ収集

Linguaskill リンガスキル のリスニング・リーディングテストは、「Metrica」と呼ばれるオンラインテスト・プラットフォームを通じて配信されます。

次のデータは、Metrica を通じて収集されました。

- 受検者のテスト解答
- 受検者が取り組んだ問題
- 受検者のテストのスコア

テスト終了時に、SurveyMonkey で実施されるオンラインアンケートへの参加を受検者にお願ひしました。アンケートには、192 名 (77.4%) の受験者にご協力いただきました。

参加者

このトライアルに参加した英語学習者の数は、合計 248 名にのびりました。

| Participants | | | | | | | |
|----------------------------------|---------|-------|--------------------------------------|---------------------------|--------------------------------------|-----------------|-------|
| A：性別 | 男性： | 38.0% | E：英語能力 (CEFRのレベル) | リスニング | A1：Breakthrough | 16.3% | |
| | 女性： | 55.5% | | | A2：Waystage | 40.0% | |
| | 不明： | 6.5% | | | B1：Threshold | 21.2% | |
| B：年齢 | 16歳以下： | 1.2% | | | B2：Vantage | 13.9% | |
| | 17～24歳： | 73.9% | | | C1：Effective Operational Proficiency | 3.7% | |
| | 25～39歳： | 15.5% | | | C2：Mastery | 4.9% | |
| | 40～59歳： | 2.9% | | | リーディング | Below A1 level | 7.8% |
| | 60歳以上： | 0.8% | | | | A1：Breakthrough | 19.2% |
| | 不明： | 5.7% | | | | A2：Waystage | 31.4% |
| C：第一言語 | 中国語： | 2.9% | | | | B1：Threshold | 20.8% |
| | フランス語： | 0.4% | B2：Vantage | 11.8% | | | |
| | ハンガリー語： | 0.4% | C1：Effective Operational Proficiency | 4.9% | | | |
| | イタリア語： | 35.1% | C2：Mastery | 3.7% | | | |
| | 日本語： | 0.4% | F：コンピューター 適応型テストの 受検経験 | コンピューター適応型テストの 受検経験なし： | 49.3% | | |
| | クメール語： | 0.4% | | コンピューター適応型テストの 受検経験あり： | 50.7% | | |
| | マレー語： | 5.7% | | | | | |
| | スペイン語： | 15.1% | | | | | |
| | タミル語： | 0.4% | | | | | |
| | タイ語： | 34.7% | | | | | |
| その他： | 4.5% | | | | | | |
| D: Reason for taking the test | 会社の要求： | 6.9% | | | | | |
| | 学習継続： | 6.9% | | | | | |
| | 昇進： | 10.2% | | | | | |
| | 移民： | 0.8% | | | | | |
| | 自己研鑽： | 37.1% | | | | | |
| | 学校の要求： | 29.3% | | | | | |
| | その他： | 8.8% | | | | | |

大半の参加者のリスニングおよびリーディング能力は、CEFR スケールで A1 ～ B1 の間にあります。CEFR スケールで、C1 ～ C2 レベルの参加者の人数はごくわずかです（図 6 を参照）。

図4 参加者の年齢

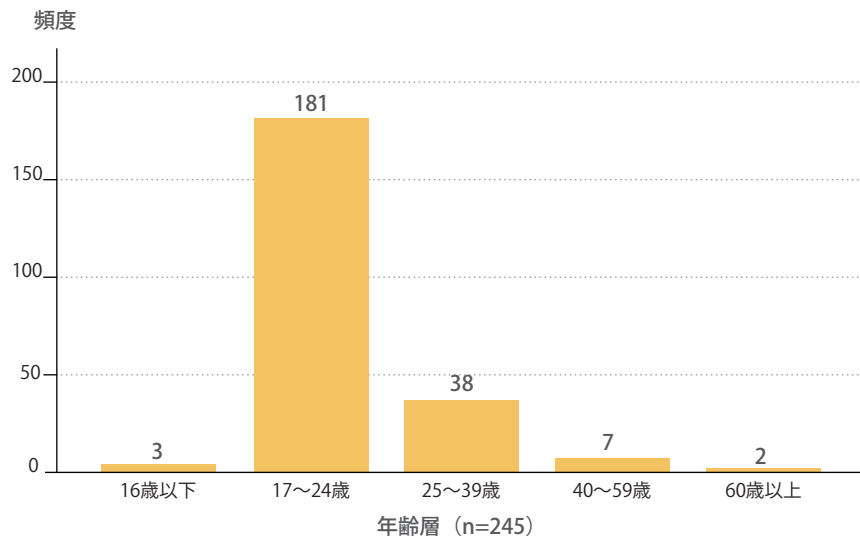


図5 参加者の第一言語

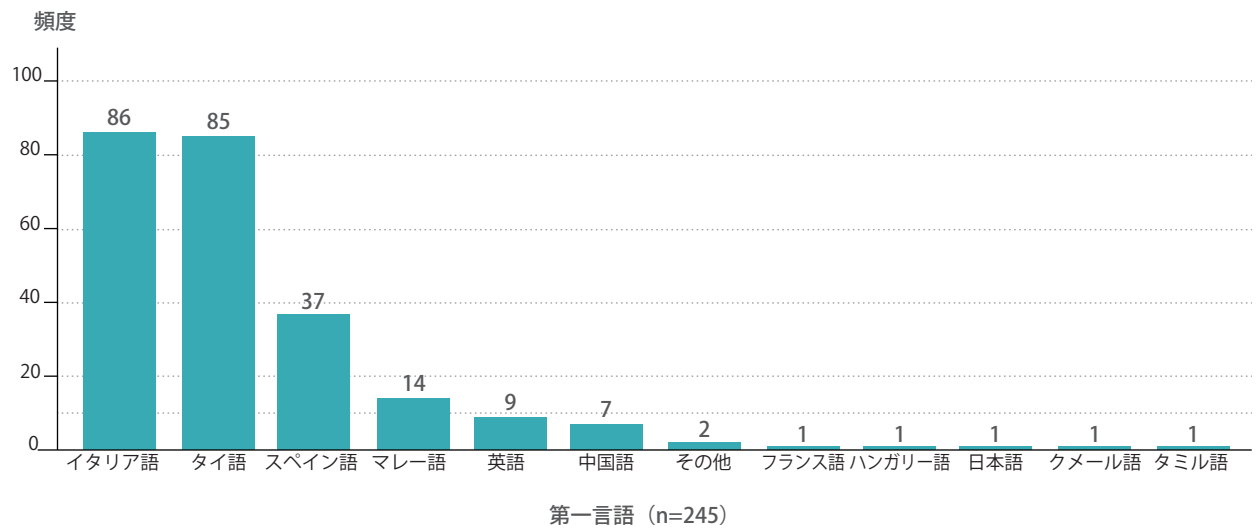
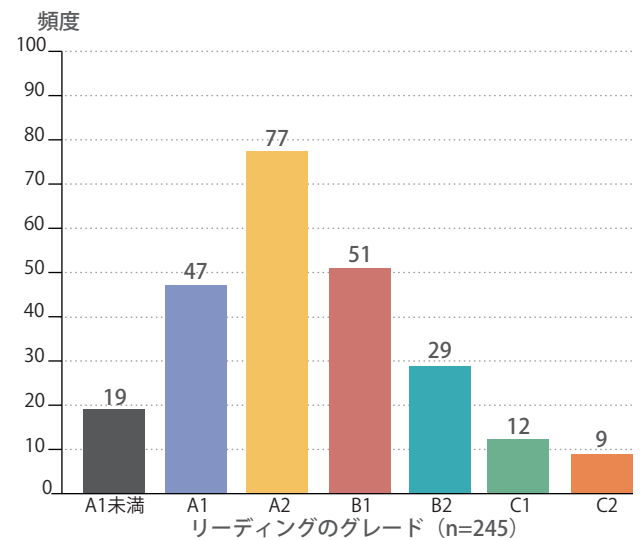
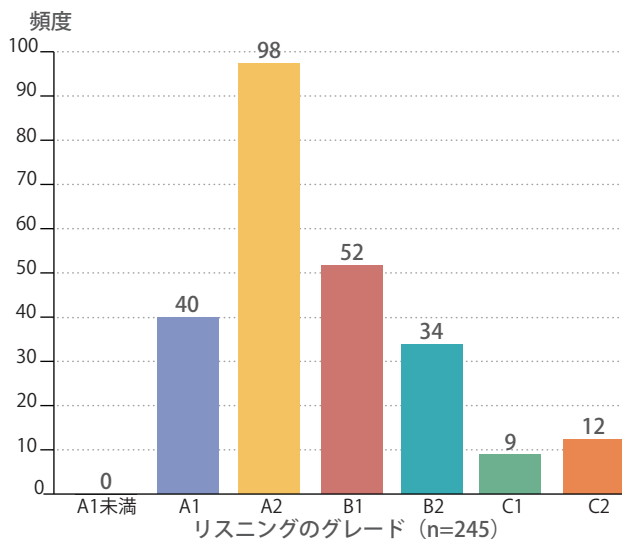


図6 参加者のリスニングおよびリーディング能力レベル



データ分析

ケンブリッジ英語検定 4 技能 CBT (Linguaskill リンガスキル) のテストスコアの正確性および信頼性は、古典的テスト理論と項目応答理論から発展した手法を使って推算されました。

コンピューター適応型テスト受検経験の (テストのグレードへの) 影響、テストの全体的印象、テストの長所と限界に対する意見が、アンケートとトライアル受検者の結果を組み合わせて、調査されました。

当トライアルは、2016 年 2 月下旬から 3 月 31 日にかけて実施されたテストで、この締切日後に受領されたテストデータは分析の対象外となりました。

当トライアルの参加者は、今後の Linguaskill リンガスキル受検者の母集団をなす代表的サンプルであると仮定されます。

参考文献


欧州評議会 (2001) Common European Framework of Reference for Languages: Learning, teaching, assessment. Cambridge: Cambridge University Press.


お問い合わせ先

ケンブリッジ大学英語検定機構 試験開発部門 日本統括
Email: InfoJapan@cambridgeenglishreps.org
〒101-0054 東京都千代田区神田錦町1-21-1
ヒューリック神田橋ビル9階 ケンブリッジ大学出版株式会社
M: 080-5545-8969、T: 03-3518-8276、F: 03-3518-8274
www.cambridgeenglish.org (English)、
www.cambridgeenglish.org/jp (日本語)
cambridgeenglish.org/helpdesk


上述の内容はすべて、2016年4月時点での正確な情報に基づいています。

Copyright © UCLES 2017 | CER/6099/7Y12

 cambridgeenglish.org/linguaskill

 [/CambridgeEnglish](https://www.facebook.com/CambridgeEnglish)

 [/CambridgeEnglishTV](https://www.youtube.com/CambridgeEnglishTV)

 [/CambridgeEng](https://twitter.com/CambridgeEng)

