

博士論文審査及び最終試験の結果

審査委員（主査） 投野 由紀夫 印

学位申請者 檜村 真由 （かしむら まゆ）

論文名 Development and Validation of English CAN-DO Descriptors for Students Majoring
in Engineering in Japanese Higher Education Institutions

【審査結果】

審査委員会は投野由紀夫教授を主査とし、主任指導教員の吉富朝子教授、副指導教員の大津友美教授、工藤洋路教授、そして外部委員の根岸雅史名誉教授の5名で構成された。上記論文の公開審査ならびに最終試験は2025年6月25日に行われ、審査委員会は全員一致で、申請者に対し博士（学術）の学位を授与するのに相応しいとの結論に達した。

【論文概要】

本論文は、実証的なニーズ分析に基づき、工学を専攻する日本の高等教育機関で学ぶ学生達のための英語能力を評価する CAN-DO ディスクリプタ（能力記述文）を開発し、その妥当性を検証することを目的としている。

工学を専攻する日本の高等教育機関で学ぶ学生のニーズ分析において、実施頻度の高かったタスクを洗い出し、その情報を元に、4技能5領域、つまり、聴くこと、読むこと、書くこと、話すこと（発表）、話すこと（やりとり）それぞれに分けて CAN-DO ディスクリプタを開発している。そして、それに続く CAN-DO ディスクリプタの検証研究では、項目応答理論の1つで、心理計量学の手法であるラッシュモデルを用いて、開発した CAN-DO ディスクリプタの信頼性と妥当性の検証を行っている。

本研究の第1章では、本研究の主旨と主な流れのほか、本研究の着想に至った経緯として、著者が過去に行った調査研究が紹介されている。檜村(2013)は、海外インターンシップ及び研究留学が、工学を専攻する学生たちの、英語を使用することへの不安を減らしているということを明らかにした。また、TOEIC®形式のリスニング、リーディング問題の事前事後の正答率の変化から、海外インターンシップ・研究留学が、リスニング、リーディングの英語運用能力を引き上げる効果を持つ可能性を示した。さらに、著者独自の質問項目から、海外経験からの

帰国後、英語を使いたい・学びたいという意欲は高まり、海外経験が動機付けになり、帰国後も意欲は高いままであるが、実際の英語学習や使用に結びつかないということを明らかにした。これらのことから、海外経験等の事前・事後に、自分の英語運用能力を把握できる指標が必要だとしている。

第2章は、本研究テーマに関連する English for Specific Purposes (ESP)、工学専攻の学生の英語のニーズ調査、学習者による自己評価、CAN-DO ディスクリプタ及び工学専攻の学生のための CAN-DO ディスクリプタ開発などの先行研究、テストや尺度の妥当性検証の研究をレビューしている。また、*The Common European Framework of Reference for Languages* (2001)をレビューし、その中心概念である、行動指向主義、social agentとしての言語使用者、複言語主義、4つのモード、European Language Portfolio や言語運用能力テストやカリキュラムとの関連付けについてもまとめている。

第3章においては、3つのニーズ分析の結果を報告している。ニーズ分析研究1では、工学を専攻する学生達が参加する海外インターンシップ、研究留学および、所属機関における英語の授業以外での英語を使用する作業について実証的な調査を行うことで、工学を専攻とする学生達が海外経験の中で、どのようなタスクを何の言語を使って行っているのかを浮き彫りにしている。ニーズ分析研究2では、調査協力者の工学専攻の学生が行う頻度の高いタスクをリスト化した。ニーズ分析研究3では、工学専攻の学生が英語の授業以外の学生生活において行う英語を使った活動を明らかにした。これら3つのニーズ分析から、工学専攻の学生が行う頻度の高いタスクをまとめることで、CAN-DO ディスクリプタ作成のための準備を行った。

第4章においては、第3章で得た結果を元に、技能ごとに、工学専攻の学生のための CAN-DO ディスクリプタ開発がまとめられている。開発した CAN-DO ディスクリプタの構成は、Listening が 36 項目、Reading が 32 項目、Writing が 36 個、Spoken Production が 31 個、Spoken Interaction が 38 個、合計 173 個である。これらは既存の CAN-DO ディスクリプタを参考にしながら作成された。開発した CAN-DO ディスクリプタは、海外での滞在において必要となるサバイバル・イングリッシュや現地の人と親しくなる際に起こるもの、新聞を読むことやテレビを見ること等、より高いレベルの英語運用能力が必要となるもの、学校生活や授業において必要となるもの、そして、工学に関連した活動（研究、海外インターンシップでの就労体験）において必要となるものというカテゴリーに分類された。

第5章では、開発した CAN-DO ディスクリプタの妥当性検証について論じている。妥当性検証の調査では、日本の高等教育機関（高等専門学校、大学、大学

院)で工学を専攻している学生達を対象に行った。調査協力者(総計 1,990 名)は、各技能・領域の開発した合計 173 項目の CAN-DO ディスクリプタに対して、英語で行うとしたらどの程度できるかについて、5 件法のリッカート形式で自己評価をさせた。そのほか、学年、専攻の情報、TOEIC® Program Test を自己評価調査実施時点から半年以内に受験した者については、受験した TOEIC®テストのスコアも収集した。

収集された自己評価は、先行研究と同様にラッシュモデルによる分析を行った。実際の分析には、統計ソフト WINSTEPS (Version 5.2.5.0)を用い、一次元性が担保されているかを解析するための値、各ディスクリプタの項目困難度、フィット値、調査協力者個人の能力値を求めた。

これらラッシュモデルでの分析と並行して、CAN-DO ディスクリプタの妥当性の検証を行う際に、Chapelle et al. (2008) and Chapelle and Lee (2021)に示されている論証に基づく妥当性検証(argument-based validation)の手法が用いられた。この論証方法では、推論(Inference)を支える論拠(Warrant)を明確にし、論拠を積み重ねていくことで、妥当性を高めていく。第 5 章のリサーチ・クエスチョンと直接的に関連がある Inference は評価推論 (Evaluation Inference) , 一般化推論 (Generalization Inference) , 説明推論 (Explanation Inference) , 外挿推論 (Extrapolation Inference) である。これらの推論に対して、それぞれ論拠を設定した。

評価推論には、「各言語スキルに対して作成された CAN-DO ディスクリプタは信頼性が高く、英語運用能力において異なる熟達度レベルの学生を効果的に区別することができる。」を論拠として示した。この推論を支える前提条件として、以下の 3 つを設定し、これらは第 5 章のリサーチ・クエスチョンの一部に対する仮説と合致する。(a)収集したデータは、ラッシュモデルの適用条件である一次元性を保持している。(b)各言語スキルに対して作成された CAN-DO ディスクリプタは、非常に高い信頼性を示す。(c) 各言語スキルに対して作成された CAN-DO ディスクリプタは、ラッシュモデルでの分析の結果、異なる熟達度レベルの学生を効果的に区別できることが示される。(a)に関しては、各スキルの収集データにおいて、標準的な手法である Eigenvalue の値では、いずれのスキルにおいても、2.0 以上となり、一次元性が担保されない結果になったが、PTMEA (point measure)相関の値が各技能内でどの項目においても高く(聴くこと: 0.38-0.82, 読むこと: 0.57-0.85, 書くこと: 0.57-0.85, 話すこと(発表): 0.56-0.81, 話すこと(やりとり): 0.62-0.83)、この結果をもって一次元性が担保されていることの証明とした。(b)に関しては、信頼性係数の値が、いずれの技能においても 0.97 以上で高い値を示したことから、証明された。(c)を論証する際には、person

reliability と item reliability の値を用いた。Bond et al. (2021) によれば、高い person reliability は、テストの実施によって、受験者のスコアが低いものから高いものまで広く分布していることを示し、高い item reliability は、ラッシュモデルにおいて、尺度上の項目が易しいものから難しいものまで広く分布していることを示す。各技能の CAN-DO ディスクリプタ群は、いずれの値も高く、異なる熟達度レベルの学生を効果的に区別できることが示された。

一般化推論では、「作成された CAN-DO ディスクリプタは、妥当性検証に参加した対象者および他のサンプルにも適切である。」を Warrant として設定し、以下の 2 つを前提条件とした。(a) 作成された CAN-DO ディスクリプタは、項目の難易度の観点から、参加者の熟達度レベルをカバーできる。(b) 作成された CAN-DO ディスクリプタは、他のサンプルに対しても適切である。(a) は先に述べた評価推論の論拠の前提条件 (b) と同じであるため、論証されており、(b) はラッシュモデルで出力される Item-Person map で確認し、多くの CAN-DO ディスクリプタの項目困難度が調査協力者たちの個人能力値をカバーしていることが確認できたが、高い個人能力値を持つ調査協力者のための項目が一部ないことがわかった。

説明推論では、論拠として、「作成された CAN-DO ディスクリプタの項目難易度は、大規模調査を基に作成された CAN-DO ディスクリプタの難易度と近い。」ということを設定し、作成された CAN-DO ディスクリプタの項目難易度は、CEFR における類似の CAN-DO ディスクリプタの CEFR レベルと対応しているということを前提条件とした。ラッシュモデルを用いた分析で示された各 CAN-DO ディスクリプタの項目困難度と CEFR のレベルの整合性がうまくいった項目も多かったが、うまくいかなかったものもあった。

外挿推論は、「作成された CAN-DO ディスクリプタで評価される英語熟達度の構成概念は、他の英語運用能力試験で評価される英語熟達度を説明しうる。」を論拠とし、以下、2 つを前提条件とした。(a) 工学系学生が自己評価で回答した CAN-DO ディスクリプタへの反応は、ラッシュモデルの期待に適合していることが示される。(b) 作成された CAN-DO ディスクリプタによる自己評価結果（ラッシュ分析から得られた個人能力値）は、英語運用能力試験の成績と正の関連を示す。(a) はラッシュモデルで出力されるフィット値を使用し、論証を試みた。その結果、どの技能の CAN-DO ディスクリプタ群においても、項目困難度が最も低い項目の幾つかがラッシュモデルに適合していないことがわかり、一部の項目は CAN-DO リストから除外した方が良いことが統計解析では示された。また、(b) については、TOEIC Bridge™ を受験した調査協力者の Listening 及び Reading のテストスコアと個人能力値はほぼ無相関であることが示されたが、

TOEIC® のテストスコアと個人能力値の相関は、各技能において、中程度の相関があることがわかった(聴くこと: 0.46, 読むこと: 0.50, 書くこと: 0.51, 話すこと(やりとり): 0.56, 話すこと(発表): 0.50)。この結果には、TOEIC Bridge™ が CEFR B1 レベルまでしか測定できないことや、TOEIC Bridge™を受験し、本調査に参加した学生の多くが、専門知識を用いて行うタスクをまだあまりやることがない高専低学年の学生が多かったことが要因として考えられる。

工学を専攻する学生達のニーズ分析を元に CAN-DO ディスクリプタの開発と妥当性検証までの研究を行い、教育的示唆として得たもののうち、重要なものとして以下を挙げる。まず、作成した CAN-DO ディスクリプタを CEFR と関連付けるためには、作成段階において各ディスクリプタに詳細な条件や基準を明記することが重要である。ここでいう「条件」とは、タスクが実施される文脈を指し、「基準」とは期待される成果を示すものである。これらの要素をディスクリプタに明確に示すことで、各 CEFR レベルにおいてどの程度のパフォーマンスが期待されるかを明確に示すことができる。

次に、CEFR レベルとの関連付けの精度を高めるためには、各熟達度レベルに対して「アンカー項目」を設定することも重要である。「アンカー項目」とは、基準点となる項目のことであり、ディスクリプタ同士が異なる言語レベル間で比較可能かつ一貫性を保つための指標となる。このように CEFR と体系的に関連付けることで、ディスクリプタの妥当性と信頼性が向上し、より広い言語教育の枠組みに統合しやすくなる。

さらに、CAN-DO ディスクリプタを妥当性検証のためだけでなく、ポートフォリオの一部として主に活用することを想定する場合、今回の妥当性検証研究で Pre-A1 レベルと判断された最も基礎的な CAN-DO ディスクリプタも保持することが重要であるとしている。これにより、特に技術系高専の1年生など、熟達度が下位に位置する学生にとっても、自分に合った CAN-DO ディスクリプタが含まれていることが保証される。特に、CEFR A1 レベルの基本的なタスクすら習得に苦労している学生にとっては、このような初歩レベルの CAN-DO ディスクリプタが含まれていることで、学習から取り残されているという感覚を防ぐことができる。低いレベルであっても成功体験を積むことができれば、英語教師は学生の自信や自己効力感を高め、英語学習への動機づけを促すことができる。これにより、より前向きで持続可能な学習体験の実現につながるとしている。

【講評】

本論文は以下の点において非常に優れているとの評価がなされた。

- ① 約 500 頁におよぶ大作であり、その中で計 63 名の工学専攻の英語学習者に対して 3 種類のニーズ分析を行い、それによって開発された CAN-DO ディスクリプタの妥当性検証を 1990 名の英語学習者を対象に実施した大規模な実証研究である。ニーズ分析に基づく CAN-DO ディスクリプタの開発のみ、あるいは妥当性検証のみであっても、学術論文として成立しうるところ、両方を行っている点が高く評価できる。
- ② 工学専攻の学生に対する英語教育を実践してきた教育者ならではの、実践への目配りができており、カリキュラム策定や教材開発など、英語教育現場への応用可能性が高い成果を出していることから、教育的意義も高い論文である。
- ③ 特定分野のための英語教育 (ESP) についての言語評価研究、とりわけ工学分野についての実証研究はまだ多くはないため、貴重な研究である。
- ④ 社会人として日頃英語教育に従事しながら、大量のデータを収集し、自律的に詳細な統計分析をおこない、通常であれば研究チームで取り組む規模の調査を一人で行った。さらに調査結果を大部な論文の形にまでまとめ上げた労力は大変なものであったと推察される。

一方で、以下のような指摘と質疑がなされた。

- ① ニーズ分析の結果を具体的な CAN-DO ディスクリプタに落とし込む工程の説明がやや不足している。また、ニーズ分析を基に CAN-DO ディスクリプタを開発することが目的だったが、CEFR の CAN-DO ディスクリプタに、ニーズ分析でわかったことを当てはめたというふうにも見える。大規模な妥当性検証をする前に、開発した CAN-DO ディスクリプタを質的に分析し検証していれば、より精度の高い CAN-DO ディスクリプタとなりえたのではないか。
- ② 難易度のレベル設定をする際、CEFR のレベルと、今回開発された CAN-DO ディスクリプタのレベルが一致せず逆転している項目についての扱いや解釈について考察が足りていない。スタンダード設定としてブックマーク・メソッドを使ってアンカー項目を立てなかったのはなぜか。
- ③ これから技術者となる工学専攻の英語学習者を対象とした調査としては大規模で価値があるが、すでにエンジニアとして英語を使用している者を対象としたニーズ分析をしなかったのはなぜか。
- ④ 工学分野の英語学習者が必要とする特定分野に特化した英語力に加え、海外インターンとして生活上必要な英語力についてもニーズ分析の対象としたの

はなぜか。生活面における英語力も対象とした場合、複言語使用が見られたとのことだったが、そうであれば複言語を加味した CAN-DO ディスクリプタとする可能性はなかったのか。またインターンシップでの留學生活の実態に関する情報収集はしたのか。

- ⑤ ニーズ分析を行った際、うまく「できなかった」ことの理由が、言語スキルの不足・専門技術分野に関する知識不足・文化的違いによるコミュニケーションの困難のいずれによるものであるか、回答者は正確に区別できていたと思うか。経験のない項目については CAN-DO の判断が不正確になりがちであることは先行研究で明らかにされているが、経験の有無について確認しなかったのはなぜか。より高い英語力をもつ学習者を対象とすることで、より妥当性の高い CAN-DO ディスクリプタが取り出せるのではないか。
- ⑥ 外挿推論の妥当性保証として、開発した CAN-DO ディスクリプタと TOEIC スコアを比較しているが、外挿推論の妥当性保証の考え方はこれで正しいのか。また工学分野に求められる英語力と TOEIC で求められる英語力が異なることについては考慮したのか。

これらの指摘や質疑に対して、樫村氏からは以下のような回答があった。

・ニーズ分析の結果を CAN-DO ディスクリプタに落とし込むプロセスは、一般英語用の CAN-DO 項目と工学英語のニーズ分析で抽出された項目の間に大きな違いもあったため、その扱いについて試行錯誤をした側面があり、検証方法に至らぬ点があった。より多くの既存の CAN-DO 枠組みについて検討を重ねておくべきだった。

・各熟達度レベルに対応したアンカー項目を入れるべきであったと今は認識している。

・今回は工学者については調査できなかったが、その必要性は感じている。また留學中の学習者の生活やインターンシップの実態についての聞き取り調査は行いにくかったが、情報が得られていればより深い考察が可能であった。

・専門分野の英語力だけでなく、生活に必要な英語力も重要であると考え、汎用性のためにも一般英語の側面も含めた言語能力を評価対象に含める判断をした。

・CAN-DO ディスクリプタにあるような言語経験の有無について、本論文では取り上げられなかったが、現在そのことについて調査を実施している。

・今後の研究の中で、複言語能力を含めた CAN-DO ディスクリプタの開発ができたら良いと思う。近年注目されている mediation を加味した評価枠組みになるだろう。

審査委員から出された指摘や質疑は、いずれも今後の研究の発展を期待しての提言であり、本論文が、工学分野の英語教育に大いに貢献する学術的・実践的成果を否定するものではなかった。いずれの指摘や質疑に対しても檜村氏は誠意ある回答をし、自らの研究の課題についても十分に自覚していた。また今後の研究における課題解決の理論上・方法論上の具体的な方策や展望も示された。

【総合評価】

以上の論文評価および最終試験での質疑応答の内容から、本論文は英語教育学、とりわけ特定分野のための英語教育(ESP)に貢献する学術的意義および教育実践的意義の高い優れた実証研究であると評価され、学位申請者が研究者としての資質を十分に有していると判断された。よって審査委員会は全員一致で、学位申請者が博士(学術)の学位を授与するにふさわしいとの結論に達した。