

論文の和文要旨	
論文題目	<p>Criterial Features in L2 English of Japanese Learners Based on Complexity, Accuracy, and Fluency</p> <p>日本人英語学習者の基準特性研究：複雑性、正確性、流暢性の観点から</p>
氏名	<p>フ シャオリン</p> <p>胡 晓琳</p>

ヨーロッパ言語共通参照枠（CEFR）は、学習、教授、評価のための極めて影響力のある汎用枠として広く認識され利用が拡大している。CEFR は、第二言語（L2）学習者を6つの習熟度レベルに分類し、それぞれのレベルで言語学習者が社会的エージェントとして言語をどのように使うことができるかを描写するための能力記述子を使用している。CEFR の特性として能力記述子により汎用的な枠組の提案を行っているため、文法・語彙などの個別言語特有の詳細を避けており、理論的にはさまざまな外国語への適用が可能である。しかし、この特徴は、CEFR を具体的に個別言語に実装する際に L2 学習者の具体的な言語特性を理解していないと、関係者にとってレベルと言語材料を結びつける際に問題を生じてくる。CEFR スケールの汎用性と適応性をさまざまな文脈で向上させるために、多くの研究が CEFR レベルを識別するための「基準特性（criterial feature）」を特定しようと試みてきた。しかし、特に複雑さ（complexity）、正確さ（accuracy）、流暢さ（fluency）（総じて CAF という）の観点からの、英語学習者のスピーキング能力の基準特性についての具体的な研究はまだ限られている。

このギャップを埋めるために、本研究では、L2 学習者の口頭パフォーマンスの

CEFR レベルを区別するための基準となる CAF 特性を調査した。さらに、本研究では、CAF 特性を用いて CEFR 予測モデルを確立する実現可能性を経験的に調査し、このモデルがダイアログとモノログのタスクにおける L2 口頭パフォーマンスの CEFR レベルをどの程度正確に予測できるかを検証することを目的とする。

本研究では、定量的・コーパスベースのアプローチを用いて、NICT JLE コーパスからランダムに 170 件の 15 分間の英語スピーキングテストのトランスクリプトを選択した。このコーパスの分析に基づいて、CAF 特性を用いた CEFR 予測モデルを構築した。その後、別のデータベースである TSST コーパスの 45 のトランスクリプト (TSST コーパスは、各トランスクリプトに 10 のモノログタスクが含まれている) を用いて、予測モデルの信頼性と妥当性を確認した。

まず、本研究では、テキスト指標の自動解析ツールと人手の特徴抽出を併用して NICT JLE コーパスから 36 の CAF 特性の結果を取り出した。これらの CAF 特性は、異なる CEFR レベルでの発生状況を観察し、特徴量としての有用性を判断するためであった。結果は、CAF 特性の大部分が CEFR レベル間で有意な差を示し、高い習熟度を持つスピーカーがより複雑な言語構造をより少ないエラーで使用し、低い習熟度を持つ学習者に比べて高い流暢性を示すことが明らかになった。

次に、これらの CAF 特性を変数として使用して CEFR 予測モデルを構築した。クラスカル-ウォリス検定の結果に基づき、CEFR レベルを区別するのに有意に寄与しなかった 9 つの CAF 特性を後続の分析から除外した。残りの 27 の CAF 特性は、3 つの CEFR グループ (A1、A2、B1) の少なくとも 1 組の間で有意差を示した。その後、最適な変数の組み合わせを決定するために判別分析を用いた。変数間の 2 変量相関に基づいて、2 つの CAF 特性が多重共線性の問題を解決するためにモデルから除外された。最終的に、複雑さの指標 3 つ (文字数/単語、節/AS 単位、節の平均長)、全体的な正確さの指標 1 つ (エラーのない節の比率)、3 つの非流暢さの指標 (繰り返し/100 単語、短い一時停止/100 単語、節末の長い一時停止/100 単語) を用いて CEFR 予測モデルを構築した。これらの中で、節/AS 単位が L2 学習者の 3 つの CEFR レベル (A1 から B1 まで) を予測する際に、最も中心的な役割を果たすことがわかった。結果は、これら 7 つの CAF 特性を用いて 3 レベルが 86.7%の交差検証ケースで正しく分類できた。

第三に、CEFR 予測モデルの頑健性を検証するために、TSST コーパスを含む新たなデータセットにモデルを再適用することによるモデル評価分析が行われた。2つのコーパス間でのタスクタイプとトランスクリプトの長さの違いにもかかわらず、結果は、73.3%の L2 学習者が適切な習熟度のグループに正しく分類された。ただし、タスクタイプの違いから、TSST コーパスの中の低習熟度の学習者は、モノログのトランスクリプトで高い流暢さを示し、その結果、習熟度レベルが過大評価され、A1 のスピーカーの 60%しか正確に予測されなかったことを指摘する必要がある。対照的に、A2 と B1 レベルの予測精度はそれぞれ 80%と 100%と精度が高かった。

本論文の調査結果は、異なる CEFR レベルの L2 学習者の間で CAF 特性に明確な違いが存在することを実証したと言える。さらに、これらの CAF 特性の中には、CEFR レベルを効果的に区別できる可能性のある重要な基準特性として特定できる特性が存在することもわかった。本研究で確立された CEFR 予測モデルを用いて、7つの主要な CAF 特性に基づいて L2 学習者の CEFR レベル (A1 から B1 まで) を高確率で予測することが可能であることが示された。さらに、このモデルの頑健性が検証され、その適用範囲がダイアログとモノログの両方のスピーチに拡張できることを示した。しかし、TSST コーパスの一部で観察された予測精度の低下は、異なるタスクタイプのデータにこのモデルを適用する際に、より高い予測精度を達成するためにさらなる調整が必要かもしれないことを示唆している。

本研究の限界点としては、主に以下の側面が挙げられよう。(1) CAF 特性が網羅的ではないこと、(2) 2つのコーパス間のデータタイプの違い、(3) 避けられない手作業による CAF 特性の膨大な計算時間、および (4) サンプルサイズの小規模なこと、である。これらの問題は、将来の研究においてさらなる改善が求められる。

それらの限界にもかかわらず、本研究は基準特徴としての CAF 特性の特定は、CEFR レベルにコーパス準拠の客観的な言語確率情報をもってそのレベル特徴記述を補完する価値がある。この補完は、異なる熟達度での L2 学習者の口頭発表が CAF においてどのように発達するかについての理解も深めることが可能である。さらに、この研究の結果は、理論的、教育的、そして方法論的な意義を持つものであることが示される。理論的には、本研究は、CEFR レベルの記述と特徴を、CAF 特性に基づく音声言語の特徴という観点から掘り下げている。教育学的に言えば、本研究で提案され

た基準特性と CEFR 予測モデルは、特に CEFR に基づく教材の開発とその実施、自己評価練習の推進や会話能力の自動評価など、実用的な応用の可能性を秘めている。方法論の観点からは、本研究は、学習者コーパスを用いて、既存の CAF 特性研究を CEFR 基準特性研究につなげる新たな方法論的道筋に貢献するものである。