

学位論文内容の要旨

報告番号	先端科学技術甲第 191 号	氏名	ヤング ジュリオ クリスティアン
論文題目	Evaluation of Speech Recognition, Text-to-Speech, and Generative Text Artificial Intelligence for English as Foreign Language Learning Speaking Practices		

個別スピーキング練習のための実行可能なソリューションとしてチャットボットが活用されている。チャットボットは、学習者の不安を軽減し、サポート的な学習環境を育む、ネイティブ・スピーカーに代わる有望な代替手段を提供する。しかし、チャットボット・アプリケーションの開発には、特に学習教材の作成とその中で使用される音声に関連する多大なコストがかかる。そこで本研究では、音声制作の問題を解決するための Text-to-Speech 技術である WaveNet と、対話練習教材を作成するためのオフライン生成 AI 技術、特に ChatGPT の性能評価に着目した。

本研究では、スピーキング用アプリケーションにおける AI 技術の統合を調査し、音声制作の問題を解決するための WaveNet と対話練習教材を作成するためのオフライン生成 AI 技術である ChatGPT の性能評価に着目した。さらに、英語学習領域におけるオフライン音声認識 (SR) の実装が限定的であることを考慮し、オフライン SR としての Vosk の適用可能性についても研究する。実験を実施し、WaveNet の TTS によって作成された教材は、ネイティブ・スピーカーの音声よりも自然ではないと認識されるが、学習者にとっては、同等に理解しやすく、活用できることを明らかにする。論文は全 5 章からなっている。

第 1 章は序章として、研究の背景、動機、目的を明記している。この章では、英語学習プロセスを支援するスピーキング練習用アプリケーションの重要性を強調し、これらのコンテンツ開発にかかる高いコストに関する課題を強調する。続いて、EFL 学習におけるテキスト生成 AI と音声合成技術の導入が進行中であることを論じ、特定された課題に対する解決策としての可能性を強調する。さらに、オフライン音声認識 (SR) を英語学習に取り入れるという、あまり研究されていない分野についても触れ、この分野における既存研究の少なさを強調する。その後、TTS と Text Generative AI をコンテンツに統合し、学生の音声入力を自動分析するオフライン SR 技術を搭載したスピーキング練習アプリを提案・開発する方向性を確立した。提案するアプリは、学生が自主的にスピーキングを練習するための補助教材として機能する。この方向性に基づき、システムに組み込まれた既存の AI 技術の実装、分析、評価の概要を示す一連の実験について詳述する。

第 2 章では、外国語としての英語学習のための音声合成技術によって生成される音声品質を評価する実験の実施に主眼を置いている。この実験では、英語の教科書から派生したリスニング練習用の補助ダイアログを使用して構成されている。各ダイアログのセリフは

ネイティブ・スピーカーによって作成され、ダイアログ内の対応するテキストと照合される。WaveNet TTS を用いて、各対応テキストの TTS バージョンを生成する。各音声グループの品質を評価するために、音声教材をランダムにシャッフルし、各音声の発音の正確さ、理解しやすさ、分かりやすさ、自然さなど、様々な知覚基準について学生の判断を求める。その後、学生に各音声の書き写しをさせ、書き写しエラー率を計算する。その後、各音声グループの知覚品質と書き取りエラー率の両方を分析することで、結論を導く。このプロセスを通じて、英会話練習にネイティブ・スピーカーが参加することの代替となりうる TTS 生成音声教材の妥当性を見極めることを目的とする。

第 3 章では、開発した音声コンテンツに WaveNet の TTS を使用した音読練習からなるスピーキング練習アプリケーションを紹介している。このアプリケーションは、オフライン音声認識を統合し、学生の音声入力を効率的に認識し、修正フィードバックを提供する。このため、SR システムによって認識された学生の音声入力結果と、ダイアログ内の参照テキストを比較した。その後 SR による書き取りの誤り率を測定するために、アプリを利用した実験を行った。その他、オフライン SR システムを学習アプリに組み込むことに関する参加者の経験や見解を収集し、評価プロセスにおけるアプリの質的側面を分析した。

第 4 章では、スピーキング練習アプリケーションの開発において、教育者に多大な時間と労力を要求するダイアログ作成プロセスの課題を取り上げる。この課題に対処するため、本章では ChatGPT を音読活動用に調整された参照用ダイアログを生成する強力なツールとして活用し、学生にスピーキングスキルを向上させるプラットフォームを提供する。一連の実験を通して、ChatGPT によって生成された教材の適切性が評価される。各実験では、ChatGPT によって生成された教材に最も適したターゲットオーディエンスを決定するために、一連の読みやすさのメトリクスを採用している。ChatGPT の出力品質に様々なプロンプトが影響することを認識し、この章ではプロンプトエンジニアリング技術の違いが ChatGPT から得られる結果の品質にどのような影響を与えるかを分析した。

第 5 章では、ChatGPT で作成した教材を既存のスピーキング練習アプリケーションに統合する。学生のスピーキング練習をサポートするアプリケーションの有効性を評価するために、実験が行われる。この章では、事前・事後テストを利用し、学生がアプリの使用により顕著な影響を経験するかどうかを検証する。さらに、アプリケーションを使用した生徒の経験を積極的に求め、ユーザーエクスペリエンスに関する貴重な定性的洞察を提供する。

最後に、第 6 章では、各 AI 技術の個別実験から得られた知見を、アプリ利用の実験結果から、効果や学生の視点と統合する。本章は、AI 技術を EFL のスピーキング練習に取り入れることの長所と短所の分析で締めくくられる。さらに、アプリ使用後に学生が認識したニーズについて議論し、改善点を強調する。これらの洞察に基づき、より効果的な AI 対応スピーキング練習アプリを開発するための今後の研究の方向性を提案する。