

Author	Licence
Daniel Tweddle	CC BY 4.0

New Vistas

Volume 12, Issue 1, 2026, UWL Annual Doctoral Students' Conference 2025

Abstract

The development of mispronunciation detection and diagnosis systems is crucial for enhancing effective communication and pronunciation learning for non-native speakers. While existing Mispronunciation Detection and Diagnosis (MDD) studies predominantly focus on American English, the need for dialect-specific systems becomes evident with the growing demand for accurate pronunciation tools across diverse linguistic contexts.

This study addresses a significant gap in British English MDD research by introducing a novel, dialect specific dataset. The creation of the Modern Received Pronunciation (MRP) dataset involved compiling phonetically rich sentences and passages from public domain texts, recruiting and recording the speech of 29 participants. These recordings were processed, segmented and transcribed in preparation for their use in a deep learning system. A baseline model was established using Convolution Neural Networks, Long Short-Term Memory Networks, and Connectionist Temporal Classification.

Experiments were conducted to compare baseline model performance when trained on MRP data against an industry standard corpus, TIMIT. Results show that MRP outperformed TIMIT in phoneme error rate, correct diagnosis rate and F1 score, when tested against utterances from L2 Arctic. TIMIT achieved a lower false rejection rate, demonstrating the need for accessible L2 corpora containing British English mispronunciation annotations. This research lays the foundation for further improving dialect specific MDD systems through techniques like language models and attention mechanisms.

The novel dataset not only facilitates testing of CTC models on British English but also contributes to the development of digital tools for pronunciation error detection, aiding language learners in producing consistent speech. Furthermore, it bridges a critical gap in UK-based MDD research, offering potential for broader applications in dialect learning and therapeutic contexts

Keywords

Mispronunciation Detection and Diagnosis