

Automatic Pronunciation Assessment Systems for English Students from Argentina

Jazmin Vidal^{1,2}, Cyntia Bonomi¹, Pablo Riera^{1,2}, and Luciana Ferrer¹

¹ Instituto de Investigación en Ciencias de la Computación (ICC), CONICET-UBA, Argentina

² Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (UBA), Argentina
jvidal@dc.uba.ar

Abstract. English proficiency is crucial for education, work, social mobility, and global engagement. Despite government efforts across Latin America to expand language learning opportunities, teacher shortages create persistent disparities, particularly affecting low-income and rural students who cannot supplement public education with private lessons. Consequently, many graduate without basic English conversational skills. Computer-Assisted Language Learning (CALL) has improved language education by offering remote learning solutions, reducing teacher workloads, and providing stress-free practice opportunities. However, these systems remain suboptimal for pronunciation learning due to poor performance in detecting errors in short speech segments. Additionally, they historically emphasized native-like pronunciation rather than intelligibility understood as those mispronunciations that cause communication breakdowns. Our long-term goal is to develop a free mobile and web application tailored to the needs of Argentinian English learners. We focus on segmental-level pronunciation (phones or syllables), which research shows is more effective for novice learners than phrase-level evaluation. We prioritize errors that are most impacting intelligibility. To address these challenges, we created EpaDB, a database of non-native English speech by Argentinian speakers for developing phone-level pronunciation scoring systems. We then explored two strategies for handling extreme data scarcity: first, a transfer learning approach that demonstrated significant improvements over standard methods. Second, we compared different self-supervised learning speech models for the task. All our code is available for research purposes in an opensource repository.

Keywords: automatic system, pronunciation, student, english..

Sistemas automáticos de evaluación de la pronunciación para estudiantes de inglés de Argentina

Resumen. El dominio del inglés es crucial para la educación, trabajo y participación global. Pese a esfuerzos gubernamentales en América Latina, la escasez de profesores genera disparidades que afectan principalmente a estudiantes de bajos recursos y zonas rurales, quienes no pueden complementar su educación con clases particulares. Consecuentemente, muchos se gradúan sin habilidades conversacionales básicas. La Asistencia Computarizada de Aprendizaje de Idiomas (ACAI) ha mejorado la educación del lenguaje mediante soluciones remotas que reducen la carga docente y ofrecen práctica personalizada. Sin embargo, estos sistemas son subóptimos para la pronunciación en especial cuando se trata de detectar errores segmentales. Además, suelen estar centrados en la natividad y no en la inteligibilidad, es decir, en captar errores que puedan afectar la comunicación. Nuestro objetivo es desarrollar una aplicación gratuita para estudiantes argentinos de inglés, enfocándonos en la pronunciación a nivel segmental (fonos o sílabas), método más efectivo para principiantes que la evaluación por frases completas. Priorizamos los errores que más afectan la comunicación. Para esto, creamos EpaDB, una base de datos de habla inglesa de argentinos para desarrollar sistemas de calificación de pronunciación. Exploramos dos estrategias ante la escasez de datos: un enfoque de aprendizaje por transferencia que demostró mejoras significativas sobre métodos estándar, y la implementación comparativa de modelos de aprendizaje auto supervisado, con todo nuestro código disponible abiertamente.

Palabras clave: sistema automático, pronunciación, estudiante, inglés.

Publicado en: Communications of the ACM, Volumen 68, Issue 8, 2024. Factor de impacto: 1.146 (Q1).S DOI: <https://doi.org/10.1145/3653326>