

UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
DEPARTAMENTO DE COMPUTACIÓN

Análisis de features dinámicas para sistemas de calificación de pronunciación de lengua extranjera

Alumno: Leandro Matayoshi
Libreta Universitaria: 79/11
Correo electrónico: leandro.matayoshi@gmail.com
Directora: Dra. Luciana Ferrer

November 16, 2017



1 Resumen

La evolución de la tecnología durante estos últimos años ha permitido el desarrollo de herramientas automáticas que asisten al estudiante durante el aprendizaje de una lengua extranjera, siendo Duolingo uno de los ejemplos más conocidos de este tipo.

La pronunciación es un aspecto importante a tener en cuenta durante la enseñanza. En la actualidad, es posible evaluar de forma acertada la pronunciación a nivel párrafo u oración, disminuyendo la precisión al realizar el análisis para unidades de tiempo más cortas (palabras y fonos). La correcta calificación le permitirá al estudiante conocer acerca de sus potenciales errores sobre fonos específicos, de forma tal que pueda trabajar en ellos para corregirlos.

El proyecto dirigido por la Dra. Ferrer, cuyo fin es el desarrollo de un sistema de asistencia computarizada para el aprendizaje de inglés de niños Argentinos, constituye un marco ideal para explorar variantes que permitan mejorar los sistemas de calificación de pronunciación a nivel fono.

El objetivo principal de la tesis será evaluar la utilización de features que modelan las dependencias temporales durante la pronunciación de un fono ya que pueden contener información no utilizada hasta el momento.

2 Motivación

El aprendizaje de una lengua extranjera es un proceso complejo que involucra muchos aspectos, entre los cuales se encuentra la pronunciación.

La Dra. Ferrer trabaja en un proyecto cuyo objetivo es desarrollar un sistema ACAI (Asistencia Computarizada para el Aprendizaje de Idiomas), para niños argentinos en proceso de aprendizaje del idioma Inglés. El proyecto cuenta con 3 etapas principales: Recolección de las grabaciones de los infantes y construcción de la base de datos, el desarrollo de un sistema de reconocimiento automático del habla para generar los alineamientos temporales y finalmente el desarrollo del sistema automático de puntuación. La tesis se focalizará en este último punto.


Si bien actualmente es posible obtener estimaciones confiables de la calidad de pronunciación a nivel párrafo y oración, cuando la meta es generar puntajes para unidades más pequeñas como palabras y específicamente fonos, los sistemas del estado del arte están lejos de alcanzar su máximo rendimiento.

La generación de buenos sistemas de puntaje a nivel fono es un aspecto clave en el marco del proyecto ya que muchos niños que recién comienzan a estudiar no están en condiciones de enunciar párrafos u oraciones largas. En suma, identificar errores a este nivel permite al estudiante focalizar sus esfuerzos en corregir aspectos concretos y específicos.

3 Propuesta de tesis


En este trabajo se continuará con la exploración de soluciones al problema de calificación de la pronunciación en el marco del proyecto dirigido por la Dra. Ferrer. Nuevamente la unidad central de análisis será el fono y los modelos se entrenarán a partir de pronunciaciones de personas no nativas, ya que se han logrado buenos resultados al adoptar esta estrategia.

Replicando y posteriormente utilizando como baseline los sistemas referentes de los trabajos


previos  desarrollarán clasificadores basados en features dinámicos, no habiéndose probado este tipo de features con anterioridad.

4 Objetivo

El objetivo es analizar la incorporación de features dinámicos que permiten capturar las dependencias temporales durante la pronunciación de cada fono. Los sistemas previos han utilizado en su mayoría características espectrales estándar en la literatura del procesamiento del habla: *Mel Frequency Cepstral Coefficients*.

Independientemente de si los resultados de los clasificadores  ~~mejoren o empeoren~~, los experimentos permitirán saber si los features utilizados previamente y los introducidos en este trabajo proveen información complementaria, lo cual puede ser de gran utilidad para futuras implementaciones.

5 Plan de trabajo

1. Revisión de la implementación actual del sistema para calificación de la pronunciación y lectura de la literatura relacionada
2. Replicación de experimentos basados , cuyos features asumen independencia temporal
3. Exploración de nuevos sistemas basados en features dinámicas
4. Evaluación de complementariedad de features mediante la fusión de baselines con los nuevos sistemas
5. Análisis de resultados y propuesta de mejoras