

Ministry Of Higher Education
King Saud University
College of Computer and Information Sciences
Computer Engineering Department



Automatic Pronunciation Error Detection of Non-native Arabic Speech

A Thesis Submitted in Partial Fulfillment of the Requirements for the Master's Degree in the Department of Computer Engineering at the College of Computer and Information Sciences, King Saud University

By

Afnan Waheed Al Hindi

Supervised by

Dr. Mansour Alsulaiman

Dr. Ghulam Muhammad

Dr. Saad Al kahtani

1435AH/ 2014

Automatic Pronunciation Error Detection of Non-native Arabic Speech

By

Afnan Waheed Al Hindi

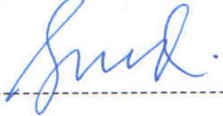
Approved on 07 SHAABAN, 1435 (05 JUNE, 2014)

Supervisor



Dr. Mansour Alsulaiman

Co-supervisor

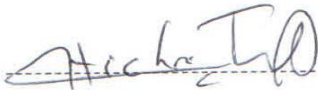


Dr. Ghulam Muhammad

Committee Members



Prof. Yousef A Alotaibi



Dr. Haikel S Hichri



Dr. Yakoub Bazi

ملخص الرسالة

القسم: قسم هندسة الحاسب

المسار: مسار شبكات

عنوان الرسالة: الكشف التلقائي عن الأخطاء في النطق لغير المتحدثين باللغة العربية

اسم الطالب: أفنان وحيد أحمد الهندي

اسم المشرف: الدكتور منصور محمد السليمان

اسم المشرف المساعد: الدكتور غلام محمد

اسم المشرف المساعد: الدكتور سعد علي القحطاني

الدرجة العلمية: الماجستير

تاريخ المناقشة: 7 شعبان 1435 هـ

الكلمات الدلالية للبحث: اللغة العربية، اللغات الأجنبية، التعرف الآلي على الكلام، تصحيح الأخطاء في

النطق، نموذج ماركوف الخفي، درجة صحة النطق.

الملخص:

بدأ اهتمام خبراء الحاسب والباحثين في التطبيقات المتعلقة بالصوت منذ عدة عقود. ومن أمثلة هذه

التطبيقات تعلم اللغة باستخدام الحاسوب. هذه الأنظمة توفر العديد من الفوائد لكل من المعلم والمتعلم

للغة، حيث أنها تسمح بتغذية راجعة مستمرة للطالب دون الحاجة إلى المتابعة الفردية من قبل المعلم،

وتسهل التعلم الذاتي في بيئة تفاعلية ومحفزة. وتعدّ عملية الكشف التلقائي عن الأخطاء في النطق من أهم العمليات التي تتميز بها التطبيقات الخاصة بتعلّم اللغة. و على الرغم من أن اللغة العربية تُعد واحدة من إحدى اللغات الأكثر انتشارًا في العالم، إلا أنها تفتقر نسبيًا إلى الأبحاث المتعلقة بالتعرف الآلي على الكلام العربي والكشف عن أخطاء النطق لغير المتحدثين بها مقارنة باللغات الأخرى مثل الصينية و الإنجليزية. هذا البحث يهتم بالكشف التلقائي عن أخطاء النطق لغير الناطقين باللغة العربية (الباكستانيين والهنود) بحيث يستطيع المتعلم معرفة الحرف الذي أخطأ في نطقه ليصحح نطقه. بعد دراسة قاعدة البيانات الخاصة بجامعة الملك سعود وجدنا أن أكثر الأخطاء شيوعًا هي في الحروف التالية: (ث، ص، ح، ض، ظ)، لذلك تمّ التركيز عليها في هذا البحث. التقنيات المستخدمة للكشف عن الأخطاء هي نماذج ماركوف الخفية وحساب درجة صحّة النطق. أوضحت لنا النتائج أن هذه الطريقة يمكن الاعتماد عليها بنسبة عالية، حيث أن درجة الصحة ممتازة، من 87% الى 100% ونسبة الرفض الخاطئ لكلام المتحدث من 0% الى 10%.