جامعة الموصل

كلية علوم الحاسوب والرياضيات

تصميم وتطبيق نظام لتحويل النص العربي المطبوع إلى كلام

عبد الوهاب فتحي شريف النعيمي

رسالة ماجستير

علوم الحاسوب

بإشراف

الدكتور رياض زغلول محمود

١٤٤٢ هـ ٢٠٢١ م

الملخـص

اللغة هي وســـيلة للتواصـــل بين أفراد المجتمع الواحد، فيعبِّرون عن أفكارهم وخواطرهم، وتحدد صفة مشتركة لذلك المجتمع.

تُقدم هذه الرســالة تطبيقاً حاســوبياً يعتمد على الإدخال البشــري لنطق أحرف اللغة العربية الغرض منه نطق الكلام المكتوب بالشكل المفهوم والصحيح. حيث يستفيد من هذا النظام شريحة واسعة من المجتمع وعلى سبيل المثال لا الحصر من ذوي الاحتياجات الخاصة وكذلك من غير الناطقين باللغة العربية.

ويتكون النظام المقترح من مرحلتين، المرحلة الأولى تتمثل بمحور تكوين قاعدة البيانات لأحرف اللغة العربية والمكونة من تســـجيل كافة الاحرف العربية بتشـــكيلاتها الأربعة (الفتحة، والضــمة، والكســرة، والســكون) ومواقع خزنها في الذاكرة وكذلك نوع الصــيغة المســتخدمة لتلك الحروف كانت (.WAV) إذ تم تكوينها لغرض إجراء المعالجات الأولية عليها.

اما المرحلة الثانية فتتمثل بعملية مقارنة الحرف المدخل من النص أو الكلمة بعد تجزئته الى أحرف مفردة مشكّلة مع الصوت المقابل لها ووضعها في مخزن حرفا تلو الاخر ليتسنى لنا بعد ذلك عملية نطقها ككلمة واحدة.

أُســتخدمنا في الجزء العملي من الرســالة لغة الماتلاب (R 2018a) لبناء النظام المقترح حيث تم اســـتخدام خوارزمية المقارنة (if-then) وتم تنفيذ النظام المقترح باســـتخدام حاســـوب محمول يعمل تحت بيئة نظام التشـغيل (Microsoft Windows 10) بمواصـفات Intel(R) Core(TM) i5-3320M CPU @ 2.60GHz 2.60 GHz وذاكرة عشوائية ٨.٠٠ GB.

University of Mosul
College of Computer Sciences
and Mathematics

Design and implementationof a text conversion system Arabic printed to speech

Abdulwahhab Fathe Shareef Al Nnaimi

M.Sc./Thesis

Computer Sciences

Supervised by

Dr. Riyadh Z.Mahmoud

Instructor Dr.

2021 A.D 1442 A.H

Abstract

Language is a means of communication between members of a single community, and they express their ideas and thoughts, and it defines a common characteristic of that community.

This thesis presents a computer application based on human input to pronounce the letters of the Arabic language for the purpose of pronouncing written speech in an understandable and correct form. A large segment of society benefits from this system, and they are people with special needs as well as non-Arabic speakers.

The proposed system consists of two phases, the first stage is represented by the axis of creating a database for Arabic language characters, which consists of recording all Arabic letters in their four formations (fatha, damma, kasrah, and sukoon) and their storage locations in the memory as well as the type of formulas for those letters were (.WAV) as they were created. For the purpose of performing initial treatments on it.

As for the second stage, it is the process of comparing the entered letter from the text after dividing it into single letters formed with the corresponding sound and placing them in a store letter by letter so that we can then process it as a single word.

In the practical part of the thesis we used Matlab (R 2018a) language to build the proposed system where the comparison algorithm was used (if-then) and the proposed system was implemented using a laptop running under the operating system environment (Microsoft Windows 10).