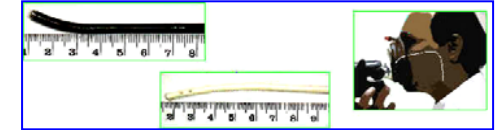
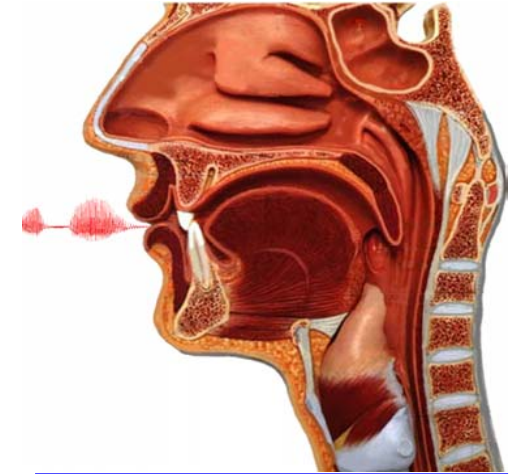


قاعدة الصوتيات العربية



معهد بحوث الحاسب والإلكترونيات
مدينة الملك عبد العزيز للعلوم والتقنية
المملكة العربية السعودية

١٤٢٦ هـ

ص ب ٦٠٨٦ الرياض ١١٤٤٢

هاتف: ٤٨١٣٧٦٥ - ١ - ٩٦٦

فاكس: ٤٨١٣٧٦٤ - ١ - ٩٦٦

sstc@kaest.edu.sa
www.kaest.edu.sa

التعريف

تعتبر اللغة العربية أول لغة وضع لأصواتها وصفا شاملا ودقيقا قبل أكثر من اثني عشر قرنا. إلا أن ذلك الوصف ليس كافيا للاستفادة منه في خدمة اللغة العربية في عصر تقنية المعلومات وتطبيقاتها المختلفة.

ونظرا للقصور القائم في هذا المجال فقد بدأت مدينة الملك عبد العزيز للعلوم والتقنية في منتصف التسعينيات العمل على إنشاء قاعدة بيانات عن الصوتيات العربية لتكون أساسا لمشاريع بحثية ودراسات صوتية متعددة، وفي نهاية العقد نفسه نتج عن هذه الجهود "قاعدة الصوتيات العربية". وقد استخدمت أحدث التقنيات والأساليب العلمية لجمع بيانات القاعدة حتى يتمكن الباحثون من توظيفها في مشاريعهم الصوتية واللغوية مختصرة بذلك الوقت والجهد والتكلفة العالية التي يتطلبها جمع بيانات من هذا النوع.

المكونات

تتكون "قاعدة الصوتيات العربية" من:
أكثر من ١٢,٠٠٠ ملف صوتي لجميع أصوات العربية في جميع مواضع الكلمة.
٧ متطوعين في التجارب.

بيانات عن حجم الهواء الخارج مع كل صوت عربي وضغطه داخل الجهاز الصوتي.
حركة اللسان والتقائه بالحنك.
حركة أعضاء النطق في الوجه.
صور لوضع الرقيقتين الصوتيتين ولسان المزمار والفتحة الحنك حلقيه أثناء نطق الأصوات العربية.
كمية الصوت الخارجة من الأنف مع كل صوت عربي.

المستفيدون

يمكن الاستفادة من القاعدة من قبل:

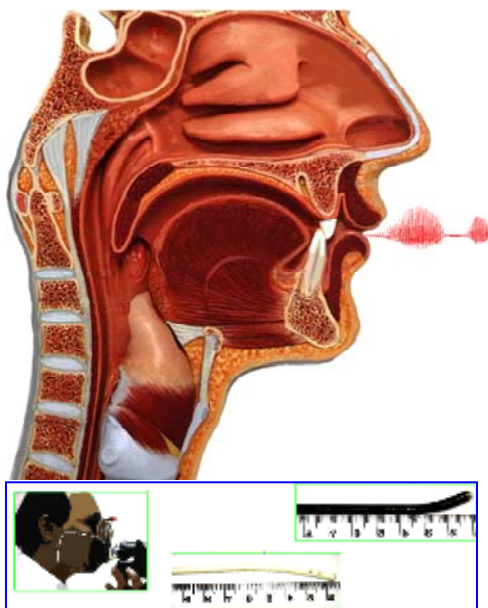
العاملين في نظم محاكاة الجهاز الصوتي.
العاملين في نظم التعرف الآلي على الكلام.
معالجي عيوب التخاطب.
العاملين في مجال البصمة الصوتية.

التطبيقات

تستخدم القاعدة في عدة تطبيقات منها:

التعرف الآلي على الكلام العربي.
التوليد الآلي للكلام العربي.
البصمة الصوتية.
علاج عيوب التخاطب

Arabic Phonetic Database



Computer and Electronic Research Institute
King Abdulaziz City for Science and Technology

© 2005

P. O. Box 6086
Riyadh 11442

Kingdom of Saudi Arabia

Telephone: 966-1-4813765

Fax: 966-1-4813764

sstc@kacst.edu.sa

www.kacst.edu.sa

Introduction

Arabic is the first language which its sounds were fully categorized and described more than 12 centuries ago. Unfortunately, this information is not suitable for the modern information technology.

Therefore, KACST started in the med 90's to build a phonetic database on Arabic. By the turn of the century, Kacst Arabic Phonetic Database (KAPD) was available for researchers and research centers. The most sophisticated equipments were used to collect the database to provide data that can easily be utilized by researchers and those who are interested in speech and phonetics.

Contents

KAPD contains the followings:

- More than 12,000 wave files for all Arabic sounds in different word positions.
- 7 subjects participated in KAPD experiments.
- Aerodynamic data during the production of Arabic sounds.
- Electropalatographic data.

- Facial movements.
- Images of the vocal folds, epiglottis and velopharyngeal port.
- Degree of nasalance.

Beneficiaries

The following specialists need KAPD:

- Speech therapists.
- Researchers on Speech Synthesis and Speech Recognition.
- Researchers on speaker identification and verification.

Applications

KAPD can be used in the following applications:

- Vocal tract modeling.
- Speech recognition systems.
- Speech synthesis systems.
- Speech Therapy.
- Voiceprint.