الحوسبة الآلية في العلاج الآلي للكلام العربي

Automated computing in the treatment of arabic speech mechanisms

راي علي، [1]

جامعة الجزائر2 قسم علوم اللسان الإيميل: ms.ali.rai@gmail.com

تاريخ الاستلام: 30 /2020/10 تاريخ القبول: 18 /2021/5 تاريخ النشر: 2021/12/31

الملخص الهدف من عملنا هو انجاز برنامج للمساعدة على تعلم الأصوات الخاصة باللغة العربية من طرف متكلمين غير ناطقين بالعربية وخصصنا في دراستنا هذه المتكلم إنجليزي، وضعنا مجموعة من الكلمات من الأصوات الخاصة باللغة العربية والتي لا توجد في اللغة الإنجليزية[ح خ ق ض ع] في مواقف مختلفة (في البداية، الوسط والنهاية) وقمنا باختيار مجموعتين من الأشخاص: المجموعة الأولي وهي مجموعة من الأشخاص الناطقين بالعربية والمجموعة الثانية من الإنجليزيين نتخذ أصواتهم للتدريب، ثم قمنا باستخراج العوامل LPC لهذه الاصوات بعد للتحليل الصوتي وفي النهاية قمنا بحساب معامل التعرف بواسطة نظام يتعرف على الأصوات الخاصة باللغة العربية حيث يوضح فيها تقنية التعرف على أصوات المتمركزة أساسا على البرمجة الديناميكية، ونلاحظ أن معدل الاعتراف لتعليم أصوات اللغة الإنجليزية وصلت إلى 33.65% ويمكننا أن نقول أن المتعلمين لا ينطقون الفونيمات بشكل صحيح وعلى وجه التحديد الفونيمات ح و ق.

كلمات المفاتيح: العلاج الآلي للكلام، التعرف الآلي للكلام، التحليل الصوتي، اللغة

Abstract: (This work is under the framework of the Automatic Speech Recognition .The aim of our work is developing a system that helps to learn the sounds specific to Arabic speakers by standard non-arabophones., we developed a corpus of sentences containing all these sounds in different positions that exist in Arabic (initial, medial and final)..

To calculate the recognition rate (TR) of these Arab sounds from LA_R and LA_N, we begin with data acquisition, which involves recording the phonemes then we extract the parameters characterizing the signal acquired by LPC coefficients Analysis of results is motivating because they yielded the way for TR THE reaches LA_N to the contrary which reaches 33,65% in the first pronunciation and by increasing the repetition that expresses the efficiency of our system.

.**Keys Words** : Arabic, Speech System (SFS), Linear Predictive Coding, Matlab

المؤلف المرسل: راي علي ، **الإيميل**: ms.ali.rai@gmail.com

1. مقدمة:

الكلام هو إحدى الوسائل المستخدمة عالمياً للاتصال بين أفراد الجنس البشري، وهو هبة الله سبحانه وتعالى للبشر الذين ميزهم به عن سائر المخلوقات الأخرى والكلام هو استحداث موجات صوتية بواسطة الحركة الإرادية للتركيب التشريحي في نظام توليد الكلام لدى الإنسان لتنقل المعلومات من المتكلم إلى السامع هذا التركيب يتألف من ربط قنوات عديدة لتشكل بمجموعها جهاز النطق وأعضاءه وشكلها مختلف من شخص إلى آخر ولهذا تتغير، (Vocal tract and speech organs) فالكلام هو سلسلة صوتية متصلة ببعضها البعض اتصالا وثيقا[1]، يمكن وصف الكلام صيغة إشارة تحمل معلومات معينة، أي تمثيلها بشكل موجة سمعية والإشكالية التي نطرحها هي كيف نقوم بتصميم نظام لتقطيع الكلمات العربية إلى أحرف وتمييزها.

ويتضمن النظام أربع مراحل رئيسية وهي كما يأتي:

المرحلة الأولى: يتم فيها تطبيق خوارزمية لتحديد بداية نطق الكلمة ونهايته.

المرحلة الثانية: تم اقتراح وتنفيذ خوارزمية تقطيع جديدة تعتمد على خصائص المنطلق الزمني وقواعد علم الفونولوجي في اللغة العربية.

المرحلة الثالثة: تتضمن استخلاص الصفات للحروف بالاعتماد على طريقة ترميز التنبؤ الخطي وبناء قاعدة البيانات للنظام والتي تشمل متجهات الصفات للأحرف المستقطعة من الكلمات LPC مع الأخذ بنظر الاعتبار ورود الحرف في مواضع مختلفة من الكلمة[2].

المرحلة الرابعة: تشتمل تمييز الحروف العربية حسب كيفية نطقها فقد تم استخدام طريقة التي تستخدم أساسيات (DTW [dynamic time warping]) انحراف الوقت الديناميكي البرمجة الديناميكية لإيجاد مسار التطابق الذي يمتلك اقل قيمة تراكمية للمسافة إذ أن الكلمات المستخدمة في التقطيع والتمييز تعود للأشخاص الأربعة المكونين لقاعدة البيانات[3]

1.2. أنظمة تمييز على الكلمات المعزولة:

أنظمة التمييز على الكلام المنفصل تتعرف على الكلمات المنطوقة بصورة منفصلة.

2.2. أنظمة التمييز على الكلام المتصل غير المستمر:

تعد كحالة وسطية بين التمييز على الكلمات المنفصلة والتمييز على الكلام المستمر ففي هذا النظام يتم دخول الكلمات إليه بنفس طريقة الكلمات المنفصلة[4]

3.2. أنظمة التمييز على الكلام المستمر (المتصل):

في هذا النظام يتم إدخال جمل كاملة بنطق طبيعي بدون أية شروط على طريقة النطق للمستخدم وهذا النظام من أصعب الأنظمة تطبيقا في التمييز على الكلام لقلة الدقة في نطق الكلمات عندما تكون في داخل الجملة.

3.مراحل عملية تمييز على الكلام:

1.3. مرحلة تسجيل وقراءة الإشارة الصوتية

وهي المرحلة المسؤولة عن تسجيل صوت الكلمة باستعمال تقنيات التسجيل المتوفرة، لتحضير الإشارة الصوتية لمجموعة من المعالجات التي سيتم عليها فيما بعد.

4. مرحلة التقسيم إلى إطارات واستخلاص المعاملات:

التقطيع إلى إطارات وتطبيق النوافذ Windowing في عملية حساب تشابه بين تابعين رياضيين حيث يكون الأول هو الإشارة أو الإطار المدخل والثاني هو التابع الرياضي المعبر عن النافذة أي أن تطبيق النافذة هو مجموع التفاف بين الإشارة والنافذة ويعبر عنها بالعلاقة[5]

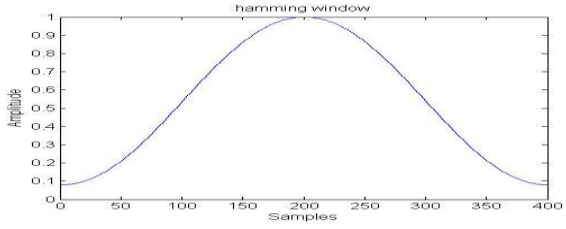
$$s(n) = \sum_{i=n-m+1}^{n} x(i) * w(n-i)$$

- **نافذة هامينغ**: Hamming Window :

هي أشهر أنواع النوافذ ولذا استخدامات كثير في لمجال معالجة الإشارة كما سنرى في هذا البحث حيث أن هذه النافذة تعطى بالمعادلة:

$$w(n) = 0.54 - 0.46 * \cos(2 * pi * \frac{n}{N-1})$$

حيث طول النافذة ، **الشكل. 1 نافذة هامينغ** هذه النافذة:



5- **الفونيمات:** تعد الفونيمات الوحدة الرئيسة التي تنقل المعنى اللغوي للكلام، فالفونيم هو اصغر وحدة صوتية محددة بعدد من الصفات الواضحة والمناسبة التي تميزه من غيره من الفونيمات (Phonetics) لتلك اللغة [6].

إن دراسة وتصنيف هذه الأصوات الكلامية نفسها تعرف بعلم الأصوات من 29 فونيم من الفونيمات (Standard Arabic language) تتألف اللغة العربية الفصحى وبذلك فان الفونيمات، (Vowels و6 فونيمات من فونيمات العلة (Consonants) الصامتة.

ويمكن (Consonants) والصامتة (Vowels) يمكن أن تصنف إلى صنفين العلة تصنيف الصوامت على النحو الأتي:

1- لأصوات الانفجارية (The plosive sounds)

عند النطق بهذه الأصوات يسمع انفجار خفيف ناتج عن اندفاع الهواء المنحبس خلف مخرج الصوت، وهناك مصطلح آخر للانفجارات هو إذ أن مصطلح الوقفات يشير إلى قفل مجرى الصوت قفلاً كاملاً Stops" .الوقفات الأصوات الانفجارية في اللغة العربية هي : ب، ض، د، ك، ت، ء، ق،[7].

2- الأصوات الاحتكاكية (The fricative sounds)

الأصوات الاحتكاكية عند نطقها لا يسمع معها مثل هذا الانفجار وذلك لان الهواء لا ينحبس عند مخرج الصوت بل ينساب في يسر وهوادة وعليه فان الأصوات الاحتكاكية في اللغة العربية هي :غ، ع، ج، ز، ظ، ذ، ه، ح، خ، ص، ش، س، ث، ف.

3- الأصوات الأنفية (The nasals sounds)

والقناة الأنفية في إنتاجها (Vocal tract) هي الأصوات التي تشترك القناة الصوتية ويوجد في اللغة العربية حرفان أنفيان هما: م، ن.

4- الأصوات الإنزلاقية (The glides sounds)

أهم ما يميز الأصوات الإنزلاقية إن الهواء المنحبس عند النطق بها يخرج من جانب من جوانب اللسان أومن كليهما واللغة العربية تمتلك ثلاثة أصوات إنزلاقية هي: ل المرققة، ل المضخمة، ر.

5-أصوات أشباه الحركات (Semi _ Vowels sounds)

.تمتلك اللغة العربية صوتين من أشباه الحركات هما: و، ي[8].

6- تسجيل الكلام :

المتكلمون: هم أربعة أشخاص اثنان من الذكور واثنان من الإناث تتراوح أعمارهم بين 26- 20 عاما، قام كل واحد منهم بتسجيل ما يقارب ال (35) كلمة مختلفة .عند إصدار

الكلمة، أمام لاقط الصوت تتغير الموجات الصوتية إلى موجات تناظرية ثم تقوم بطاقة الصوت) بتحوي لها إلى عينات في هذا البحث تم تسجيل الكلمة بمعدل عينات يساوي Sound card)، في 64 bit (Mono) واستخدام الأسلوب الأحادي11025 HZ وتمثيل كل تقطيع الكلمة العربية...

وهي من الصيغ المستخدمة (Wav) التسجيل .خزنت هذه العينات الصوتية داخل ملفات بصيغة

(Windows).بشكل واسع في بيئة النوافذ.

7-الجمل الأصلية المراد تعلم عليها

1-مبادئ الحق والحرية والمساواة

2-إلى صديقي العزيز علي، كيف حالك؟ أرجوا أن تكون بخير

3-زيادة الحسنات ورفع الدرجات

4- ضربه بمضرب من حدي

5-الحضارة الإسلامية متأثرة بحضارة الأمم الأخرى

6-مبادئ الحق والحرية والمساواة

7-الخير حيث كان

8-والفارسية القديد

9-خير لي في ديني

10-كما حرم الخمر والميسرة والعادات السيئة

11 -ظهرت الحضارة الإسلامية متأثرة بحضارة الأمم الأخرى

12-ثمانية والخمسين

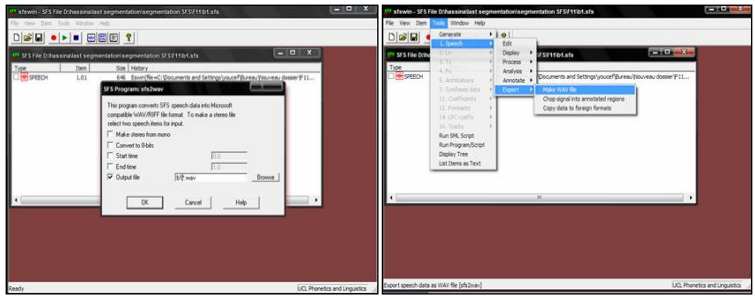
13-من كلامها أنها ضلت الطريق

14-والله ذو الفضل العظيم

15-بما أسلفتم في الأيام الخالية

16-توضع اللحم في القدر

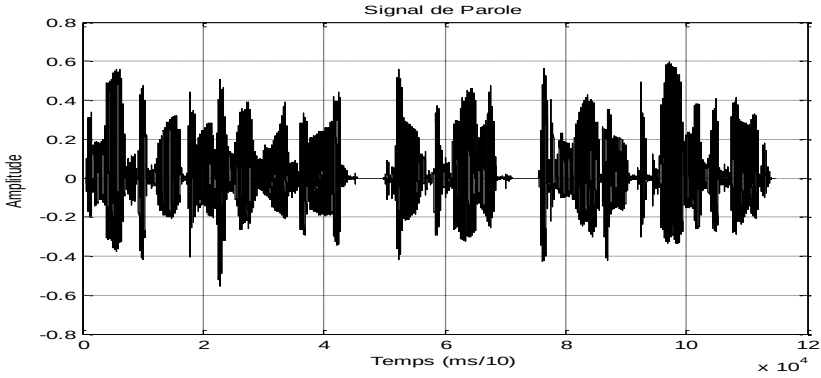
الشكل 1: الجمل الأصلية المراد تعلم عليها



الشكل 2: تسجيل ومعاينة الأصوات باستعمال برنامج SFS

8. تقطيع

المرحلة تجزئة تلعب دورا هاما جدا في النظم RAP، ويمثل كل جزء عبارةعن صوت. تم تنفيذ تجزئة



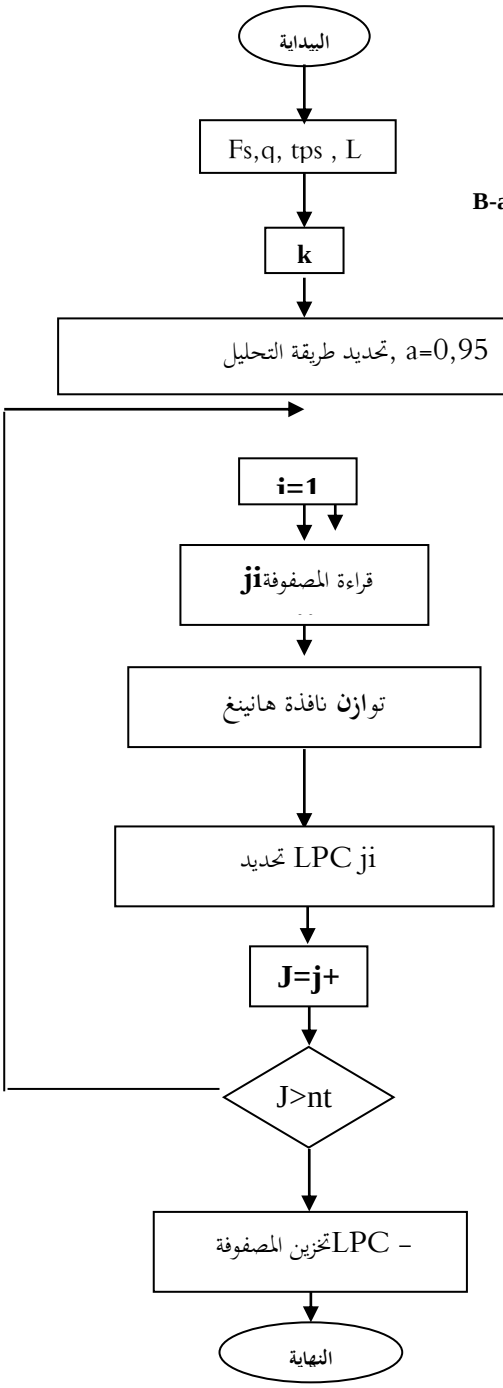
الشكل 3: إلى صديقي العزيز علي، كيف حالك؟ أرجوا أن تكون بخير

اشارة الجملة

Lieu d'articulation	Transcription Des arabisants	Phonèmes arabes	Types des phonèmes	Modes
emphatique	[d ?]	ض	v	occlusives
Uvulaire	[q]	ق	N- v	
pharyngale	[ħ]	ح	N- v	fricatives
Vélaire	[x]	خ		
emphatique	]ʕ[	ع	v	vibrante

الجدول 1: تصنيف فونيمات أصوات اللغة العربية

الحرف **k** ثابت



L=10ms ,
ntk=nombre de trame
tps = 20ms
jj = compteur de trame
k = le signal du Phonème
B-app = Nombre des occurrences de la basse d'apprentissage

الشكل 4 : المخطط التحليل الفيزيائي LPC

9- الإشارة وبناء قاعدة تحليل البيانات

تحتوي الإشارة الصوتية الرقمية على عدد كبير من المعلومات غير المهمة في عملية التعلم فضلاً عن أنها تحتاج إلى مساحة خزن كبيرة، فلتسهيل خطوات المعالجة وتقليلها نستخلص الصفات الأساسية من هذه الإشارة مما يؤدي إلى تقليص حجم البيانات.

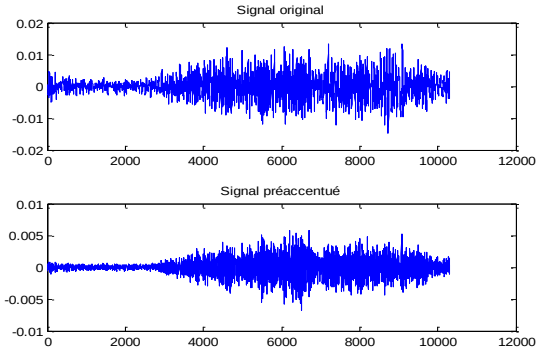
في هذا البحث يتم استخلاص الصفات للإشارة الصوتية للأحرف العربية المستقطعة كل على انفراد وبناء قاعدة البيانات من الأحرف المحللة التي (LPC) باستخدام طريقة التحليل تعد المرجع الرئيس لتمييز أي حرف، في هذا البند يتم أولا تحديد المتكلم وقراءة الملفات الخاصة بذلك المتكلم ثم يحلل كل ملف بصورة منفصلة لاستخلاص متجهات الصفات لغرض خزنها في قاعدة البيانات لذلك المتكلم. وتكرر هذه العملية مع جميع الأحرف لغرض استكمال البيانات، يتم استخلاص الصفات لكل حرف بتطبيق خطوات خوارزمية ترميز التنبؤ الخطي(P المذكورة أنفاً. علماً أن عدد معاملات المتجه الواحد لكل إطار من إطارات الإشارة يساوي (= 10). بالاعتماد على النتائج التي حصلنا عليها خلال التجربة والنتائج موضحة في الجدول)

10- مرحلة التعلم :

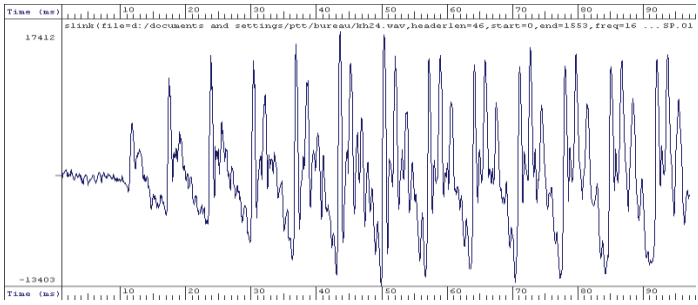
تتم عملية التعلم كما مبين في الشكل 10 الذي يوضح المخطط الانسيابي لعملية التعلم بعد أن يتم استخلاص الصفات لإشارات المصادر وبناء قاعدة البيانات للنظام يستلم النظام إشارة الاختبار التي سوف تمر بنفس المراحل السابقة إزالة السكوت من بداية ونهاية الإشارة، تقطيع إشارة الكلمة إلى أحرف، واخيرا استخلاص متجهات صفات كل حرف من حروف الكلمة.

يتم بعدها حساب اقل قيمة تراكمية بين متجهات صفات إشارة حرف الاختبار ومتجهات صفات التي ستوضح لاحقا، ثم تخزن النتيجة في (DTW) إشارة حرف المصدر باستخدام خوارزمية مصفوفة التطابق المحددة، وتكرر هذه العملية مع كل إشارات المصادر يستخرج في النهاية اقل قيمة من مصفوفة التطابق والحرف الذي يعود إلى تلك القيمة يعد ناتج التمييز (DTW) يتم تطبيق خوارزمية انحراف الوقت الديناميكي.

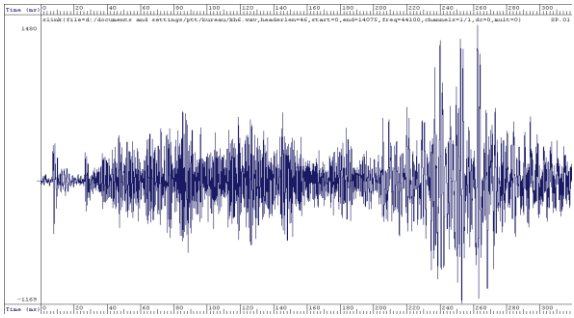
في البداية تعرف المتغيرات وفضاء (DTW). حيث يوضع المخطط الانسيابي بخوارزمية (LPC) البحث الذي سوف يستخدم لتطبيق الخوارزمية، فنستخدم سلسلة متجهات الصفات لإشارة المصدر وإشارة الاختبار الناتجة من عملية التحليل ونطابقها وان نتيجة التطابق تمثل اقل قيمة للكلفة وهي الكلفة التراكمية لمسار التطابق. نحتاج إلى مصفوفتين لخزن إشارة المصدر على الترتيب Test. و Ref وإشارة الاختبار اللتين يشار إليهم.



الشكل 5: اشارة الحرف 'خ'

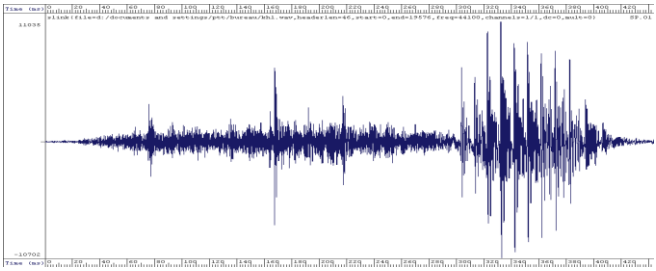


الشكل 6: تخزين اشارة الحرف 'خ ' ومقارنتة بالحرف السابقة و باستعمال نظامSFS متحدث انجليزي



الشكل7 اشارة الحرف 'خ ' ومقارنتة بالحرف السابقة وباستعمال نظام :SFS

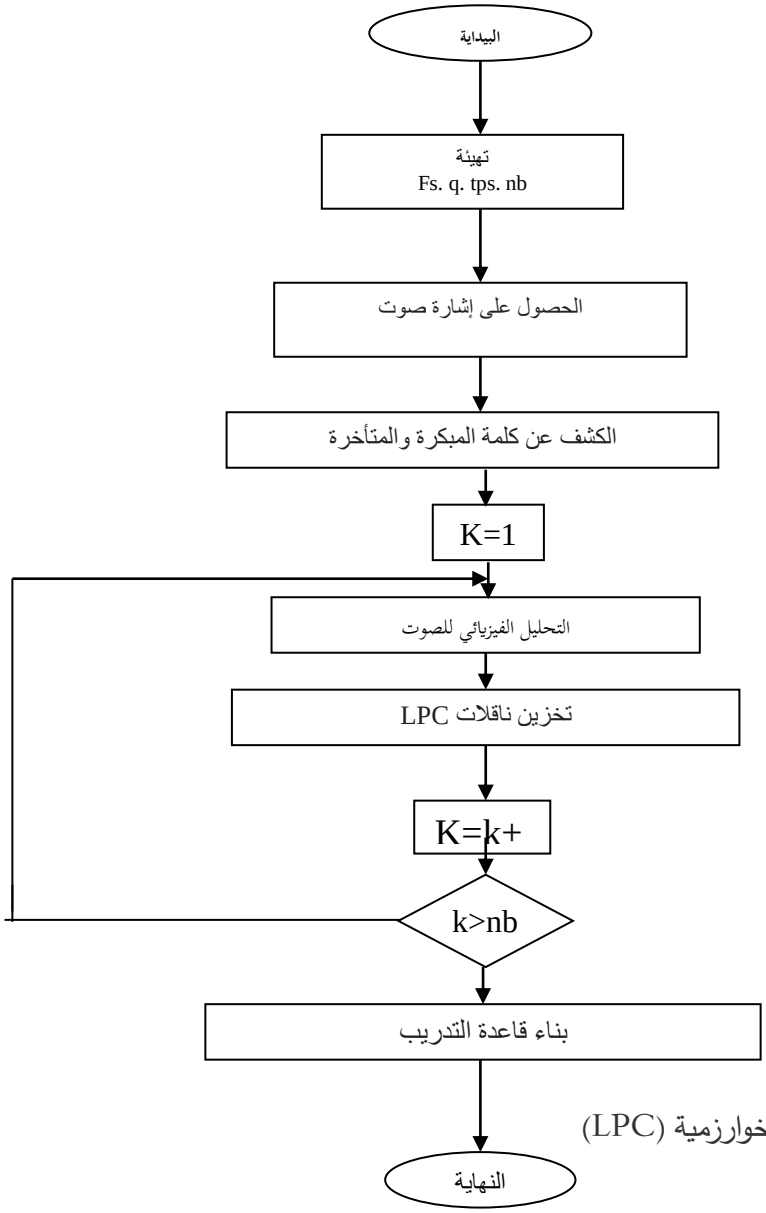
متحدث انجليزي



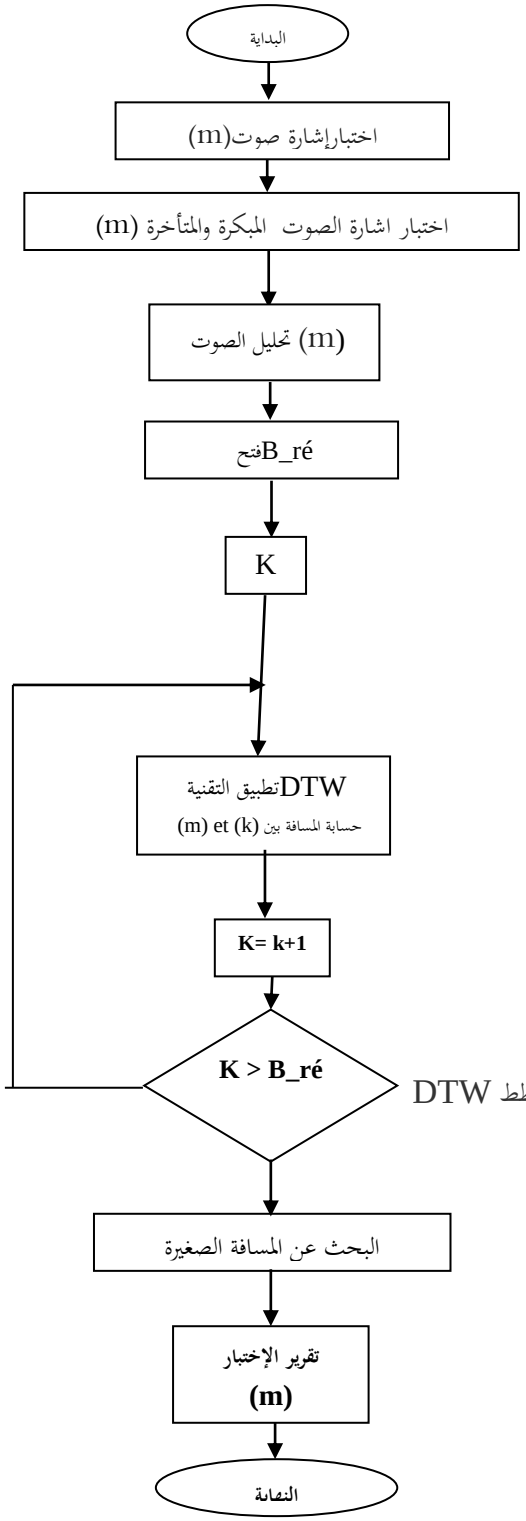
الشكل 8: اشارة الحرف 'س ' ومقارنتة بالحرف السابقة وباستعمال نظامSFS

11- بناء نظام قاعدة التدريب

يتم تطبيق خوارزمية انحراف الوقت الديناميكي(DTW) الموضحة في الشكل حيث يوضع المخطط الانسيابي بخوارزمية في البداية تعرف المتغيرات وفضاء البحث الذي سوف يستخدم لتطبيق الخوارزمية، فنستخدم سلسلة متجهات الصفات (LPC) لإشارة المصدر وإشارة الاختبار الناتجة من عملية التحليل ونطابقها وان نتيجة التطابق تمثل اقل قيمة للكلفة وهي الكلفة التراكمية لمسار التطابق. نحتاج إلى مصفوفتين لخزن إشارة المصدر وإشارة الاختبار اللتين يشار إليهما ب Test .وRef على الترتيب



الشكل9: لتطبيق الخوارزمية (LPC)



M = Phonème test
K= Phonème de référence
B_ré nombre de référence

الشكل10: الخوارزمية الانسيابي المخطط DTW

الحرف	Transcription API	عدد الحروف	عدد الحروف المراد اختبارها	النسبة المئوية (%)
خ	[x]	36	10	.77%72
ع	]ʔ[	36	10	.00%98
ض	[d ʔ]	38	10	76.25%
ق	[q]	38	10	74.25%
ح	[ħ]	36	10	2.50%8

الجدول 2: نتائج التقطيع لـDTW من معاملات LPC

الحرف	Transcription API	عدد الحرف	عدد الحروف المراد اختبارها	النسبة المئوية (%)
ح	[ħ]	36	5	25,25%
خ	[x]	36	5	30,25%
ض	[d ʔ]	38	7	%5039,
ق	[q]	38	5	27,75%
ع	]ʔ[	36	8	45,50%

الجدول 3: نتائج التقطيع لـDTW من معاملات LPC

12- الخاتمة:

لقد قدمنا خلال هذا البحث تصميم النظام تعلم أصوات اللغة العربية عن طريق متحدث انجليزي النطق باللغة العربية لذا قمنا بدراسة هذا المتكلم إنجليزي وإيجاد تقنية تعتمد على الحاسوب يتعرف على درجتها ومراقبة نسبة التحسن في النطق أثناء مراحل المعالجة، ونلاحظ أن معدل الاعتراف لتعليم أصوات اللغة الإنجليزية وصلت إلى 33.65% ويمكننا أن نقول أن المتعلمين لا ينطقون الفونيمات بشكل صحيح وعلى وجه التحديد الفونيمات [ح] و [ق]. هذا النظام غالبا ما يستعمل في مجال الاتصالات وإنجاز الأنظمة الذكية كأنظمة الأمن ونظام الإملاء على الحاسوب.

ولقد قمنا بتحسين دقة البرنامج بزيادة حجم البيانات الصوتية ثم بتكيف البرنامج مع المتحدث، وقد بدأنا العمل على توسيع قاعدة بيانات الأصوات لأنها تبين مما سبق أن حجم بيانات الأصوات المستعملة لها تأثير كبير على عملية تعلم من خلال عملية تدريب النظام مما نتج له دور مهم جدا في رفع دقة البرنامج على تقييم النطق، ومن المنتظر أن يحقق النظام دقة تفوق 60 % بعد أن استعمل هذا الحجم الكبير من البيانات الصوتية خلال تدريب البرنامج.

11- التوصيات:

من خلال هذا البحث التوصل إلى الاستنتاجات الآتية:

- ان معالجة الإشارة باستخدام طريقة التحليل نتائجها تزودنا بنموذج جيد لإشارة الكلام وذلك بإعطائها افضل تمثيل ممكن للإشارة وهذا يقود إلى دقة أعلى في أداء التمييز .
- إن انحراف الوقت الديناميكي دقة (DTW) هي طريقة رياضية مرنة وتعطي نتائج ذات عالية في التمييز
- انجاز مدونة (قاعدة بيانات) تحتوي مجموعة كلمات احد معاجم اللغة العربية وتسجيلها بواسطة متكلمين مختلفين للمساعدة على تحسين أداء أنظمة التعرف الآلي على الكلام و مساعدة الباحثين في القيام بدراساتهم و أبحاثهم.
- استخدام طرق أخرى في عمليات التمييز مثل استخدام طرق تحليل أخرى الإشارة الصوتية مثل LFCC أو HMM لغرض استخالص الصفات بدال منLPC

قائمة المراجع والمصادر:

1- **عبد الرحمن أيوب،** أصوات اللغة، ط1، القاهرة، مطبعة الكيلاني، 1968 م،

2- **أمجد عبد الرازق كرجيه،** فيزياء الصوت، والحركة الموجيه، الموصل منشورات جامعة الموصل، 1987.

3- **حسن بصري،** طرق تعليم المفردات. مجلة الدراسات العربية، العدد الرابع، السنة الرابعة، أغسطس 1994م، يصدرها قسم اللغة العربية بمركز اللغات بجامعة ملايا.

.4- **بركات ابراهيم أحمد،** التمييز الكلام العربي باستخدام الشبكات العصبونية الاصطناعية، جامعة السودان للعلوم والتكنولوجيا، 2006

5- **حليم، علياء موفق عبد المجيد،** 2003" تشفير إشارة الكلام بطريقة البعثرة بحث ماجستير، علوم حاسبات، كلية علوم الحاسبات والرياضيات، جامعة الموصل.

6- **الدليمي، شهلة عبد الوهاب،** 2002،" تمييز أصوات الحروف العربية"ماجستير، علوم حاسبات، كلية علوم الحاسبات والرياضيات، جامعة الموصل

مراجع باللغة الأجنبية :

1- Kuacan R.S.(MARCH 2006) isolated word recognition from in-ear microphone data using hidden markov models (HMM) naval postgraduate school monterey calififfornia

2- Rabiner, L. & Juang, B. (1993). Fundamentals of Speech Recognition. Prentice-Hall International, Inc.

12. هوامش:

[1]– عبد الرحمن أيوب، أصوات اللغة، ط1، القاهرة ، مطبعة الكيلاني، 1968 م، ص 96

[2]– أمجد عبد الرازق كرجيه، فيزياء الصوت، والحركة الموجيه، الموصل منشورات جامعة الموصل، 1987. ص846.

[3]– حسن بصري،. طرق تعليم المفردات. مجلة الدراسات العربية، العدد الرابع، السنة الرابعة، أغسطس 1994م، يصدرها قسم اللغة العربية بمركز اللغات بجامعة ملايا

[4] - Kuacan R.S.(MARCH 2006) isolated word recognition from in-ear microphone data using hidden markov models (HMM) naval postgraduate school monterey calififormia

[5]– بركات ابراهيم أحمد، التمييز الكلام العربي باستخدام الشبكات العصبونية الاصطناعية، جامعة السودان للعلوم والتكنولوجيا، 2006م

[6]– حليم ،علياء موفق عبد المجيد، 2003 م"تشفير إشارة الكلام بطريقة البعثرة بحث ماجستير، علوم حاسبات، كلية علوم الحاسبات والرياضيات، جامعة الموصل.

[7]– الدليمي، شهلة عبد الوهاب، 2002،" تمييز أصوات الحروف العربية"ماجستير، علوم حاسبات، كلية علوم الحاسبات والرياضيات، جامعة الموصل

[8] - Rabiner, L. & Juang, B. (1993). Fundamentals of Speech Recognition. Prentice-Hall International, Inc.