وزارة التعليم العالي و البحث العلمي
جامعة الجزائر–2– أبو القاسم سعد الله
كلية اللغة العربية و آدابها و اللغات الشرقية
قسم علوم اللسان

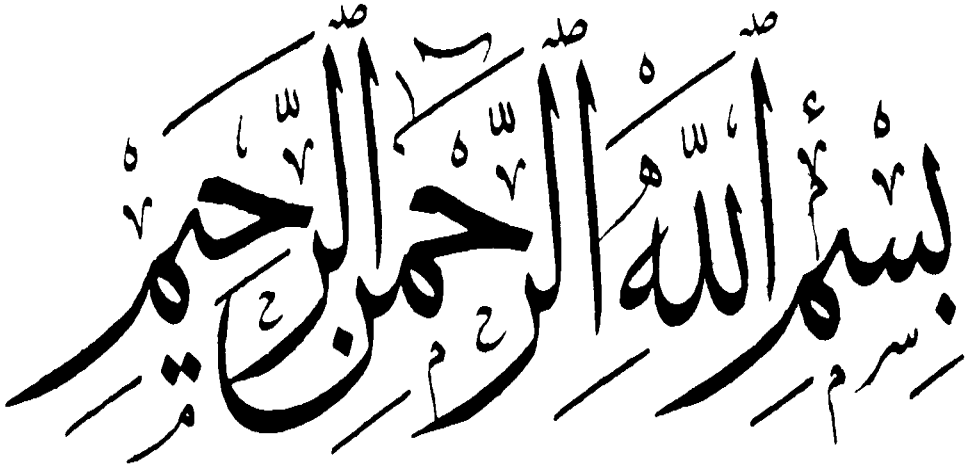
التمييز الآلي للكلام العربي بالإعتماد على الفونيمات الأحادية

أطروحة لنيل شهادة دكتوراه العلوم
تخصص : علوم اللسان و التبليغ اللغوي

إعدا الطالب :
راي علي
إشراف الدكتور(ة):
متيش أمال بهاء هدى

الإسم و اللقب	الدرجة العلمية	الجامعة الأصلية	الصفة
د. عبد المجيد سالمي	أستاذ التعليم العالي	جامعة الجزائر2	رئيسا
د. عظامو آمال بهاء هدى	أستاذ التعليم العالي	جامعة الجلفة	مشرفا ومقررا
د.متيش لحسن	أستاذ التعليم العالي	جامعة الجلفة	عضوا مناقشا
د. عيسيو محمد	أستاذ التعليم العالي	جامعة هواري بومدين	عضوا مناقشا
د. عليان حسينة	أستاذ التعليم العالي	CERIST	عضوا مناقشا

السنة الجامعية: 2020/2019


بسم الله الرحمن الرحيم

إهداء

قال الله تعالى: " ولئن شكرتم لأزيدنكم "

نرفع أيدينا قبالة السماء بحمد الله على ما من علينا من نعمة الهداية و الإرشاد و التوفيق

فله الحمد من قبل و من بعد.

أهدي ثمرة جهدي هذا إلى من قال في شأنهما

" و بالوالدين إحسانا"

إلى الذي تعب كثيراً من أجلنا وفرح كثيراً بنجاحنا ، و انتظر بفارغ الصبر أن يراني طالبا في الدراسات العليا ، إلى : أبي العزيز"حفظه الله ورعاه وأطال عمره

إلى التي غمرتني بالعطف والحب والحنان مذ كنت في المهد صبياً ـ ولازالت تغمرني ـ سِتّ الحبايب : " أمي الحنون " حفظها الله ورعاها و أطال عمرها ..

إلى التي ملأت حياتي عبقاً وحباً وجمالاً ، وأهدتني فرحة قلبي ونور عينيّ :إبنتي الغالية: " ياسمين "وكانت لي اسما على مسمى و نِعمَ الزوجة : زوجتي الغالية: " حميدة"

إلى إخواتي : محمد و لخضر و فتحي والهتاك : سمرة، خضرة، بركاهم ، أمينة ، قادة ، توفيق وكل أصدقائي وزملائي في الدراسة والعمل إلى هؤلاء جميعا أهدي ثمرة جهدي المتواضع هذا

راي علي

كلمة شكر

أحمد الله العلي القدير حمدا كثيرا يليق بجلال وجهه وعظيم سلطانه ،

لا تضاهي آلاءه ونعمه المصبغة ، وإن اجتهدت لذلك.

وأصلي وأسلم وأبارك على شفيعنا ونبينا وسيدنا محمد صلى الله عليه

وسلم وبعد:

نشكر االله تعالى أولا...

لا يفوتني أن أتقدم بجزيل الشكر و العرفان و الامتنان بشكر الخالص إلى الأستاذة الكريمة الدكتورة " عظامو آمال بهاء هدى " لها الفضل الأول في إرشادنا للبحث في هذا الموضوع حفظها الله ، وأفادتني الكثيرة بتوجيهاتها العلمية والمنهجية القيمة .

وأثني بالشكر إلى الأستاذ الدكتور "متيش لحسن" الذي لم يبخل علينا بسديد توجيهاته وصواب رأيه وغمرني بكرم نفسه ، و عظيم تواضعه ، وطيب حفاوته. كما أتوجه بشكري الخاص إلى الأستاذ الدكتور عبد المجيد سالمي رئيس قسم علوم اللسان بجامعة الجزائر2 على كل ما قدمه لنا من نصائح و إرشادات كما أقدم تشكراتي للأستاذ الدكتور دواد مسعود الذي وجهني ولم يردني خائبا كلما سألته أو طلبت منه النصح و التوجيه .

كما أتقدم بجزيل الشكر والعرفان إلى كل من مد لي يد العون والمساعدة ، من قريب أو من بعيد في إنجاز هذا العمل...نخص منهم

أعضاء مخبر "نمذجة و محاكاة و تحسين الأنظمة المركبة" بجامعة الجلفة

دون أن أنسى توجيه الثناء والشكر الأعضاء المناقشة هذه الرسالة لتحملهم عناء قراءتها،

فلهم منا كل الشكر على مجمل نصائحهم وتوجيهاتهم وانتقاداتهم التي تنير مسارنا العلمي.

راي علي

المقدمة العامة

بسم االله الرحمن الرحيم

الحمد االله رب العالمين والصلاة والسلام على أشرف الأنبياء وسيد المرسلين نبيينا محمد بن عبد االله ، و على آله وصحبه أجمعين . أما بعد :

ارتبطت اللغة بالإنسان أشد ارتباطا منذ غابر العصور فكانت متنفسة وكانت القناه التي يمرر فيها مقصده وعلى هذا استطاع الإنسان أن يعيش داخل بيت اللغة ويترعرع فيها ، بل كانت هي الوجود الرابط بين أمه وأخرى ، ولما تفطن لذلك وعرف أنه باللغة يستطيع خرق كل الحواجز اتخذها في أحيان كثيرة لتحقيق أهدافه ، ومهما يكن الأمر فإن اللغة و إشكالاتها تعد مطلبا من مطالب الباحث اللساني الذي يبين مكانتها و أسرارها .

ومن هنا بدأت الدراسات اللغوية في جميع الأبحاث منزلة متميزة جعلتها مادة للبحث ، و أخذت تتطور شيئا فشيئا مع نمو العقل البشري وتطور الوسائل التكنولوجيا للبحث العلمي ، حيث أصبح الإنسان يستعمل أدق الآلات التي من شأنها أن تساعد في حل كثير من الإشكاليات المستعصية ، وكشف أدق الأسرار التي تحير عبقريات الإنسان ومواهبه.

ويتناول هذا العمل في طياته موضوعا من المواضيع المهمة التي تثار بشأن اللغة كعلم يدخل في إطار الدراسات و الأبحاث العلمية المستحدثة ، وهو موضوع التمييز الآلي للكلام العربي و مدي مساهمته في النهوض بالعلوم و المعارف الإنسانية .

يعتبر العلاج الآلي للكلام من أوسع مجالات البحث حيث يعتمد عليها الإنسان في التواصل وهو الكلام من حيث موضوع التخاطبية الكلامية بين الإنسان والآلة منذ قرابة خمسة عقود ، اهتمام المهندسين وعلماء الكلام معاً ، ذلك لأن الكلام هو طريقة التواصل الأسهل والأكثر طبيعية بين البشر ، وتسمح مثل هذه الواجهات للسواد من الناس ، بالتواصل مع الشبكات الحاسوبية والحصول على المعلومات دون الحاجة إلى أن

يكونوا معلوماتيين ، وتتطلب هذه الواجهات تقنيات إنتاج الكلام من الحواسيب (تركيب الكلام) ، وفهم الآلة الكلام أو تعرف الكلام (Speech Recognition).

و قد توفرت في عصرنا الحالي كل الوسائل العلمية التي تساعد على دراسة الأصوات اللغوية بشكل يجعل الشرح و التحليل أكثر سلاسة و دقة و من ثم يسهل عمل الباحث في هذا الميدان ، ويسهل أيضا على المتعلم أن يأخذ ما شاء من التقنيات التي تمكنه من بلوغ مقصده في العلم.

إن الكلام الذي ننطق به يحتوي على تيار مستمر من الأصوات ، ولا تنفصل تلك الأصوات التي تكون الكلمات عن طريق التوقف أثناء الكلام ، ولكن يجب أن نستخلص تلك الأصوات و الكلمات من ذلك الكم المتصل من الهواء الخارج من الجهاز النطقي أثناء الكلام ، فالكلام يخرج في شكل كتل صوتية متتابعة في سلسلة صوتية أثناء الكلام المتصل ، حيث تتألف كل رسالة من سلسلة أصوات تتوافق مع سلسلة من الإشارات كل إشارة كلمة تملك وجهين هما الدال و المدلول ، ويتألف كل دال بدوره من سلسلة من الأصوات البدائية التي لا دلالة لها ، وهي الفونيمات.

إن الشروع في انجاز نظام للتعرف الآلي على الكلام لابد أن يبدأ بمرحلة تحضير قاعدة للبيانات تحتوي على تسجيلات إشارة الكلام المراد التعرف عليها و التي هي الحروف العربية فهذه المرحلة تعد مرحلة هامة و تتطلب دقة عالية و معرفة بمختلف الحروف و خصائصها حيث أن دقة النتائج المتحصل عليها تتعلق بجودة المعطيات التي توفرها قاعدة البيانات و لتحضير هذه الأخيرة نقوم بإنشاء مدونة تحتوي على عدد كاف من الجمل باللغة العربية تكون الحروف فيها في مختلف المواضع (في بداية و وسط و نهاية الكلمة) و بعدها نقوم بتسجيل هذه الجمل بواسطة عدة أشخاص ثم عملية تقطيع الإشارة المتحصل عليها إلى مقاطع تعالج لأجل الحصول على قاعدة بيانات تستعمل لاحقا في عمليتي التدريب والاختبار لنظام التعرف الآلي على الكلام المقترح ، وبالتالي

لابد من دراسة الأصوات دراسة علمية موضوعية بعيدة عن الذاتية وقد بدا واضحا ظهور مخابر صوتية حدث بعدها رقمنة هذه الدراسات الصوتية بظهور برامج حاسوبية تختصر العمل وتضمن النتائج ، ومنها البرنامج (**Speech Filing System**) **SFS** الذي اعتُمِد في هذا البحث.

إنّ هذا المشروع يهدف إلى استغلال الإمكانيات الكبيرة لتقنية نماذج ماركوف المخفي التي أظهرت درجة عالية من الدقة في تمييز الكلام ، وسيتم تحليل مناهج للتمييز على الكلام وهي : الطريقة الصوتية الإحصائية ، وهنا سيتم التركيز بصورة كبيرة على التقنية الإحصائية للتمييز على الكلام بسبب أنها الأكثر استخدامًا فى معظم أنظمة التعرف على (HMMs) الكلام الحديثة وخاصة نماذج ماركوف المخفي.

و بعد تحضير قاعدة البيانات تبدأ مرحلتا التدريب ثم الاختبار لنظام التمييز ، حيث يعتبر فيها التعرف علي مقاطع صوتية نابجة عن التداخل بين الأصوات عند الانتقال من حرف إلى آخر، و هنا تظهر أهمية التقنيات التي تستخدم في التمييز الآلي على الكلام مثل نموذج ماركوف المخفي HMM ، لما تمتاز به من قدرة على التمييز بين الأصوات المختلفة في الإشارة الصوتية من خلال معاملاتها.

هناك جملة من التحديات تظهر عند إنجاز نظام للتمييز الآلي للكلام ، من بينها أن الأصوات بطبيعتها تتداخل و تؤثر إحداها على الأخرى إضافة إلى وجود الضوضاء مما قد يؤثر على دقة التمييز.

كان وراء اختياري لهذا الموضوع عدة أسباب ألخصها في النقاط التالية:

- أهمية الوسائل العلمية الحديثة في تطوير البحوث اللغوية.
- إطّلاعي على المنهج المطبق في تعليم علم الأصوات.
- تطوير وحدة برمجية آلية للتعرف على الكلام يتضمن وحدتين : سمعية و نطقية ، من اكتشاف وتصحيح أخطاء نطق اللغة العربية بمستوى مقبول من الدقة .

- تطبيق طريقة نماذج ماركوف المخفية لكي نطور نظاماً يتعرف ويصحح نطق اللغة العربية .
- رغبتي لكسر هذا الحاجز المفتعل بين الدراسات النظرية والدراسات التطبيقية .
- توفير قاعدة بيانات تسهل على الباحثين القيام بالتجارب المختلفة نظرا لعدم توفرها.
- قلة البحوث والدراسات المتعلقة بالتمييز الآلي على الكلام العربي.
- حب الإطّلاع على مباحث ونظريات ومدى تطبيق نتائجها في مختلف الفروع العلمية الأخرى.
- الاعتماد على الجانب التقني والإستفادة منه للوصول إلى نتائج أقرب.
- إجراء عمليات التحليل الصوتي الآلي للفونيمات الأحادية.
- التعرف الواضح و التام على ظواهر الصوتية

كما أنني لم أجد أي دراسة تتعرض لهذا الموضوع بهذه الطريقة التي تجمع بين الوصف الآلي والفيزيائي لهذا الموضوع .

و قد انطلقت من إشكالية مدى إمكانية إخضاع اللغة العربية للدراسة العلمية ، المتعلقة بالتمييز الآلي على الكلام العربي و ما جدوى هذا النوع من الدراسات؟

وما هو تأثير معاملات تحليل الإشارة المدخلة على دقة التمييز وهل الاختلاف بين المتكلمين يؤثر على دقة تمييز الكلام؟ لهذا فإننا نلاحظ أن الخوارزميات التمييز قد تختلف بعضها على بعض في السياقات والقواعد التي تتبعها أثناء عملها على الرغم من أن النتيجة التي يجب أن تعطيها هي واحدة ، وألخص هذه الأهداف والأهمية في النقاط التالية :

- توفير معطيات عن أصوات اللغة العربية وخصائصها الأكوستيكية.
- إثراء البحث العلمي في مجال المعالجة الآلية للغة العربية.
- استعمال بعض التطبيقات على أنظمة التمييز الآلي على الكلام .

وقد اعتمدت في بحثي هذا منهجا وصفيا يعتمد على وصف وتحليل خصائص الصوتية والنظام اللغوي ولتحقيق هذا الهدف ، فقد نظمنا عملنا على النحو التالي:

تشتمل **المقدمة** إحاطة حول الموضوع وطرح الإشكال ، ثم عناصر البحث وأهدافه بالإضافة إلى الخطة المعتمدة.

في **الفصل الأول** تطرقت فيه إلى مفهوم أصوات اللغة العربية وخصائصها ، التعريف بالصوت الطبيعي واللغوي ، علم الأصوات النطقي ، أقسام علم الأصوات ،علم الأصوات الفيزيائي ، و مفهوم الفونيم .

وقد عرج **الفصل الثاني** أن أعطي لمحة عامة عن معالجة الإشارة الكلامية والتقنيات الرئيسية لتحليل الإشارة الصوتية واستخراج المعاملات لتمثيل الإشارة الصوتية.

و في **الفصل الثالث** شرح و تقديم الأساسيات الرياضية لنماذج المركوف المخفية HMM.

أما في **الفصل الأخير فتناولت فيه الدراسة العملية والنتائج** ، وسنناقش فيه بتفصيل أكبر منهجية العمل ، استخدام نماذج ماركوف المخفية (Hidden Markov Models) في عملية تمييز الآلي للكلام العربي بالإعتماد على الفونيمات الأحادية .

وأنهي البحث **بخاتمة** تتضمن جملة من أهم النتائج المتوصل إليها، إضافة إلى بعض الاقتراحات الرامية إلى تشجيع الباحثين في هذا المجال .

ولقد اعتمدت في دراستي هذه على مجموعة من المصادر والمراجع أهمها : دراسة الصوت اللغوي لأحمد مختار عمر ، استخدام نماذج ماركوف المخفية في التعرف على صور الوجه الاعتيادي والمحددة ، أحمد حسن ، رشا رعد هادي ، اللغة العربية معناها مبناها ، تمام حسان ، التعرف على الكلمات المعزولة ، ذو الفقار ابراهيم ، أصوات اللغة العربية ، غدير أحمد حامد ، عبد الغفار حامد هلال ، كمال بشر، علم الأصوات ، محمد حسن حسن جبل ، المختصر في أصوات اللغة العربية ، الصوتيات العربية منصور بن محمد الغامدي.

ولقد اعترضتنا في إنجاز هذا البحث بعض الصعوبات ، التي من شأنها أن تزيد الباحث الجاد إصرارا ، ومحاولة تجاوزها للوصول إلى نشوة الانتصار بتحقيق النتائج المرجوة ، ومن هذه الصعوبات صغر حجم المدونة ، وإن كان يبدو للبعض أن ذلك يسهل عملية البحث ، و كما أن كثرة الكتب و ضرورة الجمع بين ما هو قديم وحديث جعل الزمن ينفلت منا ، صعوبة اختيار المدونة في ظل عدم وجود قاعة مكيفة لإجراء مثل هذه التسجيلات ، صعوبة التعامل مع المصطلحات الصوتية ، ولكني أقرّ أنني قد واجهت البعض منها خاصة فيما يتعلق بالحصول على المراجع المتخصصة الأدبية التقنية .

لا يفوتني أن أتقدم بجزيل الشكر و العرفان و الامتنان بشكر الخالص إلى الأستاذة الكريمة الدكتورة " عظامو آمال بهاء هدى " لها الفضل الأول في إرشادنا للبحث في هذا الموضوع حفظها الله ، وأفادتني الكثير بتوجيهاتها العلمية والمنهجية القيمة ، وفي الختام أسأل الله أن أكون قد وُفقت إلى ما قصدت إليه في دراستي ، فإن وفقت لما أردت ووفيت ما قصدت ، فالحمد الله الذي هدانا لهذا وما كنا لنهتدي لولا أن هدانا الله ، و إن تكن الأخرى فالخير أردت ، و الجهد بذلت ، و منه التوفيق و العون و إن كان الجهد البشري لا يخلو من النقص ، فالكمال الله تعالى وحده.

الفصـل الأول:

أصوات اللغة العربية و خصائصها

1. مقدمة :

إن الاعتناء بأصوات اللغة العربية هو المنطق الأول في دراسة مستويات التحليل اللساني في الدراسات اللغوية الحديثة ، والذي يعرف بالمستوى الصوتي ، ويختص بدراسة أصغر وحدة في الكلام ، وهي الصوت اللغوي ليحظى بعد ذلك باهتمام الباحثين ، باعتبار مسائله دراسات وصفية تحليلية لها نتائج تطبيقية في ميادين متباينة ، في فروعه الثلاثة النطقي و الفيزيائي و السمعي .

كما لها دور رئيسي في حياة الإنسان لاستخدام الأصوات في الكلام الذي يعد أحد وسائل التعبير في التعاملات المختلفة في الحياة ، إن الصوتيات إحدى فروع علوم اللسانيات ويشكل المستوى الأدنى من مستويات الدراسات اللغوية ، حيث تكون المستويات الأعلى عقلية مجردة بينما الصوتيات علم ملموس ، فهو يتعلق بأصوات اللغة من حيث مخارجها وخصائصها الأكوستية وسماعها ، وله ارتباط مباشر بعلم لساني آخر أعلى منه وهو الفونولوجيا (phonology) الذي يشمل دراسة النظام الصوتي للغة وعلاقة الأصوات ببعضها ، وتأثير بعضها على بعض ، فالفونولوجيا يحدد الأصوات الأساسية لكل لغة أو ما يعرف بالفونيمات (phonemes) وتشكل فونيمات اللغة ما يعرف بالنظام الصوتي(sound system) ، ويتكون النظام الصوتي من صوامت (consonants) و صوائت (vowels) ، وفي المستوى الفونولوجي نطبق القوانين الفونولوجية فتتأثر بعض الأصوات نتيجة لوجودها في بيئة صوتية محددة ، وسوف ندرس في هذا الفصل خصائص أصوات اللغة العربية من حيث الخصائص النطقية والأكوستيكية ، حيث مهدت لذلك بتوضيح عن الجهاز النطقي ، دراسة تشريحية ونطقية ، والجهاز السمعي كذلك ، محاولا إعطاء فكرة عن كيفية حدوث الصوت اللغوي عند الإنسان ، وكيفية إدراكه من خلال الأذن ، كما تم توضيح المصطلحات الصوتية الأكوستيكية ، كالموجة الصوتية وترددها ، والمكونات الصوتية لها ، وعلى هذا الاعتبار فقد حظي الجانب الصوتي في اللغة بالكثير من الدراسة و الفحص و التحليل ، أدى إلى تفريع

الصوتيات إلى عدة أقسام ، كل منها يتناول جانبا معينا في دراسة المفاهيم الأساسية العامة المتعلقة بالصوتيات و فروعها و جوانبها .

2 . الأصوات اللغوية :

يرى علماء اللغة أن الأصوات اللغوية تتكون من وحدات مستقلة بالإمكان نطق صوت معين منعزلا عن غيره من الأصوات بغض النظر عن المعنى الذي يقع فيه ، وهذه الأصوات مختلفة ، أو الوحدات الصوتية المستقلة عن بعضها [1]، و التي يعبر عنها بالصوت الواحد ، وهي ما يطلق عليها علماء اللغة الغربيون المحدثون " فونيم " بأنه عائلة من الأصوات المترابطة فيما بينها في صفات لغة معينة [2]، والتي تستعمل بطريقة تمنع وقوع أحد الأعضاء في كلمة من الكلمات في نفس السياق الذي يقع فيه أي عضو أخر من العائلة ، أو بعبارة أخرى فهو أصغر وحدة صوتية خالية من أي معنى يمكن تحديده من معنى منطوق، وتحتوي كل لغة على عدد من الفونيمات، ومن هنا يجب التفريق بين الصوت المنطوق ، أو ما يشار إليه (allophène) آو (phone) ، وبين الفونيم أو ما يسمى بالصورة الذهنية للصوت أو الوحدة الصوتية ، ويمكن أن أجمل القول في أي صوت في أي لغة يستعمل في كلمة من الكلمات مرات متكررة ، وكيفية النطق قد تختلف من كلمة إلى أخرى ، ومن موضع إلى أخر ضمن الكلمة الواحدة اختلافا بسيطا ، فخشي العلماء أن تنحرف أصوات العربية ، فبداية القرن الثاني قاموا بوصف الأصوات معتمدين على التجربة باللسان و الأذن [3] ويقول د. أحمد مختار في هذا الإطار (وعلى أي حال فمن

---

[1] أحمد قريش ، اختلاف القدامى و المحدثين في تحديث المخارج و الصفات بعض الأصوات ، جامعة أبو بكر بلقايد تلمسان ، ص58.

[2] صالح القرمادي ، دروس في علم الأصوات اللغوبة ، نشريات مركز الدراسات ، تونس 1966 ص 135

[3] رمضان عبد التواب ، المدخل إلى علم اللغة و مناهج البحث اللغوي ، مكتبة الخانجي ، القاهرة ، ط2، 1405 هـ. 1985م، ص14.

المنطقي أن يكون البحث في الصوت اللغوي عند العرب قد بدأ في شكل جمع المادة اللغوية ، وقد تم هذا الجمع أولاً عن طريق المشافهة والحفظ دون منهج معين في ترتيب المادة المجموعة وتبويبها) [1].

لم يكن علماء اللغة وأصحاب المعاجم بعيدين عن هذا الأمر فقد تناولوا بعض المشكلات الصوتية ، إما في مقدمات معاجمهم ، أو في ثنايا المادة اللغوية كما فعل الخليل بن أحمد [2].

ويعتبر العرب من السابقين لغيرهم في علم الأصوات ، حيث ترجع نشأته إلى القرن الثاني الهجري ، ذلك العصر الذي يعج بنشأة الكثير من العلوم العربية وازدهارها ، وكان غاية هذا العلم عملية منذ البداية ، إذ جعلت مبادئه ومعطياته أساساً لعلوم العربية من نحو ، وصرف ، ومعجم ، و غير ذلك من معارف أدبية ونقدية وبلاغية. وتابعه في ذلك المبرد في كتابه (المقتضب) حيث عالج الإدغام في الجزء الأول من كتابه ، وقدم له بدراسة للأصوات ومخارجها[3].

أما المعجميون فقد تناولوا بعض القضايا الصوتية ، وذلك في المعاجم التي اهتم أصحابها بالأصوات وتقلباتها أمثال الخليل بن أحمد الفراهيدي ، الذي جعلها أساساً لتصنيف العين[4]، وأبو بكر بن دريد الذي أتبع نظام التقلبات في معجمه (جمهرة اللغة) [5].

إن الناطق بكلمة ، في كل لغة من اللغات ، يؤديها بصغة متواصلة وكان ما تلفظ به شيء واحد لا يقبل التجزئة ولكن داخل هذه الوحدة الصوتية يمكن إجراء تجزيئات وتحديد وحدات متتالية صغيرة غير قابلة للتجزئة هذه الوحدات تسمى الأصوات[6].

---

[1] احمد مختار عمر ، البحث اللغوي عند العرب ، القاهرة ، عالم الكتاب ، ط6، 1988 م ، ص 79-80.

[2] الخليل بن أحمد الفراهيدي ، كتاب العين ، (100 هـ 174 هـ)ج 11، ص 72.

[3] أبو القاسم محمد بن يزيد بن عبد الأكبر ، الكامل في اللغة والأدب ، إمام العربية في بغداد في زمنه ، ج 3 ، ص 144.

[4] نفسه ، الخليل بن أحمد الفراهيدي كتاب العين ، ج 2 ، ص 314.

[5] أبو بكر بن دريد ، هو أبو بكر محمد بن الحسن بن دريد، والمقصور والممدود والجمهرة في اللغة ، البصرة ، ج 6 ، ص80.

[6] مصطفى حركات ، اللسانيات العامة و قضايا العربية ، ط1، مكتبة العصرية ، بيروت ،1998م ص15.

وقد شارك أصحاب الموسوعات الأدبية في هذا المضمار وفي مقدمتهم الجاحظ في كتابه (البيان والتبيين) الذي تحدث فيه عن تنافر الحروف [1] .

وأول من تناول المباحث الصوتية في مؤلف منفرد هو أبو الفتح عثمان بن جني في كتابه (سر الصناعة) وتناول فيه عدد الحروف وترتيبها ووصف مخارجها وفروعها المستحسنة والمستقبحة.

ومن الذين اهتموا بأفراد الدراسة الصوتية ، ابن سينا فقد ألف رسالته (أسباب حدوث الحروف) التي تناول فيها طبيعة الأصوات من حيث قابلية الأذن لإدراكها وشدة الصوت ،كما تناول مخارج الحروف كسابقيه وانفرد بحديثه عن السواكن والعلل وتسمية الأول صوامت والثانية مصوتات وفصل كل هذا تفصيلاً دقيقا لم يكن عند سابقيه .

وكان جل ما قام به العرب في هذا المجال موجهاً إلى قيم الأصوات ووظائفها في اللغة وهو ما اصطلح عليه علماء اللغة المحدثون علم وظائف الأصوات (Phonologie) أو علم الأصوات التنظيمي ، وذلك لأنه يعالج الأصوات داخل النظام اللغوي المعين[2].

3. تعريف الصوت :

منذ وجد الإنسان و هو مشوق إلى معرفة ما حوله من الكائنات ، وقد سلك في سبيل وصوله إلى هذا الهدف مسالك مختلفة [3]، إلى أن توصل إلى أيسر السبل وأسهلها وهو الكلام الذي اعتبر من أجلّ النعم في حياة الإنسان ، إذ اللغة عبارة عن أصوات يعبر بها كل قوم عن أغراضهم[4].

---

[1] الجاحظ ، البيان والتبيين ، مكتبة الخالجي القاهرة ، (ط3) ، 1968 ، ص72.

[2] كمال بشر ، در اسات في علم اللغة ، القاهرة مصر ج 2، ص 67.

[3] إبراهيم محمد نجا ، التجويد والأصوات ، مطبعة السعادة ، القاهرة مصر ، ص 4.

[4] ابن جني ، الخصائص ، تحقيق محمد علي النجار ، دار الكتاب العربي ، بيروت ، ج 1 ، ص33.

1.3.الصوت لغة : يعرف الصوت لغة ‹ بـأنه الجرس ويقال الصات ، يصوت ، ويُصاتُ صوتًا ، وأَصَاتَ وصوت به ،كله ندى›[1]ويعرفه الفيومي فيقول: ‹ الصوت في العرف جرس الكلام ، و الجمع الأصوات ، هو مذكر، ورجل صائت إذا صاح وصيت قوي ، والصيت بالكسر الذكر الجميل في الناس ›.

2.3.الصوت اصطلاحا: ‹اضطراب طبيعي خارجي يعرض لجميع الأجسام وخاصة الهواء ، وهذا الاضطراب من جنس وصنف الظواهر الاهتزازية و التموجية وهو حركة جسم في اتجاهين، فهو تموج ينتشر في الهواء أو في غيره من المواد القابلة للاهتزاز، ومركز استقباله هو الأذن›[2].

فنجد من بين علماء اللغة العربية الذين وصفوا أصواتها وصفا دقيقيا لغويا ولهم دراسات مستفيضة منهم ابن الجني الذي درس الأصوات اللغوية [3]، وقدّم تعريفات عدة لمجموعة من المصطلحات ، وفي تعريفه لغويا يقول ‹أعلم أن الصوت عرض يخرج من النفس مستطيلا متصلا ، حتى يعرض له في الحلق و الفم ، والشفتين مقاطع ، تثنيه عن امتداده واستطالته ›.

وفي كلام الجاحظ [4] فقد عرفه في كتاب (البيان و التبيين) بقوله : ‹ الصوت هو آلة اللفظ و الجوهر الذي يقوم به تقطيع وبه يوجد التأليف ، ولا تكون حركات اللسان لفظا ولا كلاما موزونا ، ولا منثورا ، إلا بظهور الصوت ›.

كما يشير ابن فارس في تعريفه للصوت ، لقوله ‹ هو حبس لكل ما وقع في أذن السامع ›

أما الدكتور أحمد مختار عمر فيعرف الصوت اللغوي بأنه: ‹ الصوت الناتج عن اهتزاز الأوتار الصوتية ، حين مرور الهواء بها ، أي هو الصوت الذي يكون مصدره الجهاز النطقي عند الإنسان ›.

---

[1] ابن منظور ، لسان العرب ، دار صادر، بيروت ، ط1992،1 ، ص 57 .

[2] خولة طالب إبراهيم ، مبادئ في اللسانيات ، دار القصبة للنشر، الجزائر، 2006 م، ص 46.

[3] ابن جني ، سر صناعة العرب (تحقيق مصطفى الشقا و آخرون) ، مطبعة الباني القاهرة ، 1994 م ، ص6.

[4] المرجع السابق ، البيان و التبيين ، ص 79.

أما الدكتور تمام حسان في تعريفه للصوت [1] ‹ الأثر السمعي الذي تحدثه تموجات ناشئة من اهتزاز جسم ما ›.

ومن الواضح أن الصوت **الجرس** : هو عبارة عن ظاهرة طبيعية تدرك بحاسة السمع ، وقد أثبت علماء الصوت تجارب لا يتطرق إليها الشك ، أن كل صوت مسموع يستلزم وجود جسم يهتز ،كما أثبتوا أن هزات مصدر الصوت تنتقل في وسط غازي أو سائل أو صلب ، حتى تصل إلى الأذن الإنسانية[2].

نفهم من هذا القول أن الصوت هو الركيزة الأساسية التي استمد منها المصطلح الصوتي وجوده فهذا الأخير يقف عند جوانب مختلفة لها علاقة وطيدة بالصوتي وعلى هذا الأساس تفرع المصطلح الصوتيً إلى مصطلحات خاصة بجهاز النطق ومصطلحات خاصة بمخارج الأصوات ومصطلحات خاصة بصفاتها.

وقد اختلف العلماء في تحديدهم لمفهوم كلمة (صوت) وذلك بسبب اختلاف العلوم التي تخصصوا في دراستها ، فكانت النتيجة هي تعدد المفاهيم التي حددت بها الكلمة ، و اختلفت لكنها تكاملت .

من الناحية الفيزيائية حدد الصوت بعدد من المفاهيم عند غير عالم من علماء الطبيعة (Physics) وذلك على النحو التالية :

1- الصوت اهتزازات ميكانيكية في أي وسط عادي (غاز ، سائل ، صلب) [3].

2- الصوت سلسلة تتابعات من التضاغطات (Compressions) ،والتخلخلات (Pare factions) المتتالية في الهواء [4].

3- الصوت هو الطاقة التي تصل إلى الأذن من الخارج [5].

---

[1] تمام حسان ، اللغة العربية معناها و مبناها ، المكتبة الثقافية ، الدار البيضاء ، المغرب ، ص 66.

[2] إبراهيم أنيس ، الأصوات اللغوية ، مطبعة الأنجلو المصرية ، 1979م، ط5 ، ص6.

[3] محمد عبد الرازق كرجيه ، د. فيصل عبد الحليم ، ما نسمع وما لا نسمع ، بغداد ، مكتبة النمرود 1988م، ص5.

[4] المرجع نفسه ، ص846.

[5] الكسندر و آخرون ، الصوت ، ترجمة محمد عز الدين فؤاد ، القاهرة ، دار الكرنات ، 1962 م ، ص 13.

وأما من الناحية الفسيولوجية – النفسية ، فقد حدد الصوت بعدد آخر من المفاهيم عند غير عالم من علماء النفس الفسيولوجي (Psychophysiology) ، وذلك على النحو التالي :

1-الصوت خبرة حسية في الدماغ تنقل إليه عبر الأعصاب السمعية للأذن [1].

2- الصوت إحساس سمعي ناتج من دخول التتابعات السريعة من التضاغطات والتخلخلات المتتالية الحادثة في الهواء إلى الأذن البشرية.

الصوت هو الإحساس السمعي (Auditory sensation) أي السمع.

إذا نظرنا إلى المفاهيم التي ذكرناها يتبين لنا بعض الأمور هي :

أولاً : أن كلمة (صوت) لها معنيان هما:

أ) معنى موضوعي – فيزيائي .

ب) معنى فسيولوجي – نفسي .

ثانياً : أن زاوية النظر التي اعتمدت لتحديد مفهوم الصوت مختلفة من الناحية الفيزيائية عنها من الناحية الفسيولوجية النفسية .

ثالثاً : أنه لا يوجد مفهوم واحد من المفاهيم السابقة الفيزيائية منها والفسيولوجية النفسية يمكن أن يوصف بأنه جامع مانع ، إلا من حيث زاوية النظر التي اعتمدت .

وأما من حدد مفهوم الصوت بأنه إحساس سمعي ناتج من دخول التتابعات السريعة من التضاغطات والتخلخلات المتتالية الحادثة في الهواء إلى الأذن البشرية فقد ميز من ناحية الصوت بأنه إحساس سمعي من الضوء مثلا من حيث هو إحساس بصري كما أنه بين ، ومن ناحية أخرى أن هذا الإحساس السمعي أثر ناتج من تضاغطات وتخلخلات دخلت الأذن البشرية.

[1] أمجد عبد الرازق كرجيه ، فيزياء الصوت والحركة الموجية ، الموصل منشورات جامعة الموصل، 1987م، ص8.

ويأتي من حدد مفهوم الصوت بأنه الإحساس السمعي ، أي السمع ، فيفهم من كلامه أمران هما:

الأول : أن مالا يدرك صوتاً مسموعاً لا يعد صوتا.

الثاني : أن الصوت ينعدم عندما يستبعد العضو الحساس للصوت (الأذن) من المكان[1]، وعليه فالصوت هو ما أثار الأذن لتؤدي إلى إحساس سمعي ، ولا تتأثر الأذن لتؤدي إلى إحساس إلا بتحقق شرطين اثنين هما :

أ. أن تكون ترددات الأصوات واقعة ضمن مدى الترددات السمعية التي تدركها الأذن و هو المدى الذي يترواح بين (20-2000 ذ/ت) [2].

ب. أن تكون الأذن واقعة في مجال الطاقة الصوتية المسموعة أو قوة إسماع الأصوات del (Sonority).

وكان اهتمامهم بالأصوات اللغوية شيئا طبيعيا بالنظر إلى أنها – أي الأصوات اللغوية – شيئا محسوسا وهي المادة الأولى التي تتألف منها اللغة ، فهي أول ما يقدم على دراسته عادة علماء اللغة قديما و حديثا بالفحص و التحليل و التصنيف و بيان خصائص الوحدات الصوتية و اختلاف النطق وصورة تعاملها في أثناء الكلام وما يترتب عن ذلك من تغيرات صوتية في بنية الكلمة المفردة أو المتصلة بغيرها .

ومنه نستنتج أن المحدثين و القدماء لم يختلفوا في مصطلح الصوت اللغوي ، أو في تحديد مفهومه وهذا ما يدل على تقارب وجهات النظر حول هذه الظاهرة اللغوية.

---

[1] ابن منظور ، لسان العرب ، دار صادر ، بيروت ط1، ، ج2 ، الصوت ، 1955-1992 م ، ص 13.

[2] المرجع السابق ، فيزياء الصوت والحركة الموجية ، ص 55.

4. أنواع الصوت وفروعه:

الصوت نوعان[1] : عام وخاص :

1.4. الصوت الطبيعي و هو الصوت العام :

يعرف الصوت في علم الطبيعة (الفيزياء بأنه الأثر السمعي الذي ينشأ من اتصال جسم بالأخر)[2] ، أو هو الحدث الذي يختص السمع بإدراكه وينشأ من التقاء جرمين أحدهما بالأخر.

والمراد بالأثر السمعي أو الحدث الذي يختص السمع بإدراكه تلك الظاهرة الطبيعية التي هي عبارة عن الذبذبات أو الاهتزازات الصادرة من الجسمين الملتقيين ، وتنتقل خلال الوسط الناقل للصوت كالهواء في شكل موجات متتابعة ، حتى تصل إلى أذن السامعين ، فإذا انعدم هذا الاهتزاز أو تلك الذبذبات انعدم السمع ،ويمكن إدراك ذلك بطرق شوكه رنانة فإنك تلاحظ اهتزازها ويتبع ذلك الأثر السمعي وإذا سكنت اختفي الصوت[3].

معنى ذلك أن هنالك أسساً تقوم عليها دراسة الصوت فأي صوت مسموع يستلزم ما يأتي [4]:

1- وجود جسم يهتز .

2- وسط تنتقل فيه الذبذبات الحاصلة من اهتزازات الجسم المتذبذب ويمكن أن يكون ذلك الوسط غازياً أو سائلا أو صلبا كما ذكرنا سابقا.

---

[1] عبد الغفار حامد هلال ، أصوات اللغة العربية ، ط3 ، مكتبة و هبة ، القاهرة ، 1416هـ،1996 م، ص24،23.

[2] المرجع السابق ، الأصوات اللغوية ، ص7.

[3] المرجع السابق ، فيزياء الصوت والحركة الموجية ، ص 6.

[4] المرجع نفسه ، الأصوات اللغوية ، ص 9.

2.4. الصوت اللغوي وهو صوت الخاص:

هو ككل الأصوات ينشأ من ذبذبات مصدرها في الغالب الحنجرة عند الإنسان ، فعند اندفاع النفس من الرئتين يمر بالحنجرة ، فتحدث تلك الاهتزازات التي بعد صدورها من الفم أو الأنف ، تنتقل خلال الهواء الخارج على شكل موجات حتى تصل إلى الأذن ، ولكن الصوت الإنساني معقد ، إذ يتركب من أنواع مختلفة في الشدة ومن درجات صوتية متباينة [1].

كما أن لكل إنسان صفة موجة صوتية (Sound wave) خاصة تميز صوته من صوت غيره من الناس ، فليس صوت الإنسان في أثناء حديثه ذات شدة أو درجة واحدة ، و الصوت اللغوي عملية حركية يقوم بها جهاز النطق ، وتصاحبها أثار سمعية معينة ، تأتي من تحريك الهواء بين مصدر إرسال الصوت ، وصول كلمة الشمولي يؤدي بصفة متواصلة ، وكأنه لا يقبل التجزؤ ، لكن داخل هذه الوحدة الصوتية Sound nnit يمكن إجراء تجزيئات ، وتحديد وحدات صوتية متتالية صغيرة غير قابلة للتجزؤ من هذه الوحدات ، يطلق عليها الأصوات [2]. فتحدث به تموجات تنبه أطراف الأعصاب ما تشعر به المراكز السمعية في المخ الذي بدوره يقوم بتفسيرها وفك لغزها وإعطائها قيمتها ومعناها ، استناداً إلى ما اختزنته سابقاً من علاقة بين الرمز الصوتي ومدلوله.

3.4. الفرق بين الحرف والصوت:

الحرف هو الوحدة التجريدية المرسومة تشمل صوتا أو أكثر ، وقد لا يكون صوتا حينما لا ينطق ، وقد يكون صورة مرسومة للصوت ، الحرف هو الرمز الدال على الصوت أما الصوت هو ما ينتج عن العملية الحركية ذات الأثر السمعي المنطوق ، والصوت هو حركة أعضاء النطق المسموعة ، وهما شيئان مختلفان ، أما الحرف فهو مجموعة من الأصوات يجمعها نسب معين ، فهو فكرة عقلية لا عملية عضلية ، وإذا كان الصوت مما يوجده

---

[1] المرجع السابق ، فيزياء الصوت والحركة الموجية ، ص16.

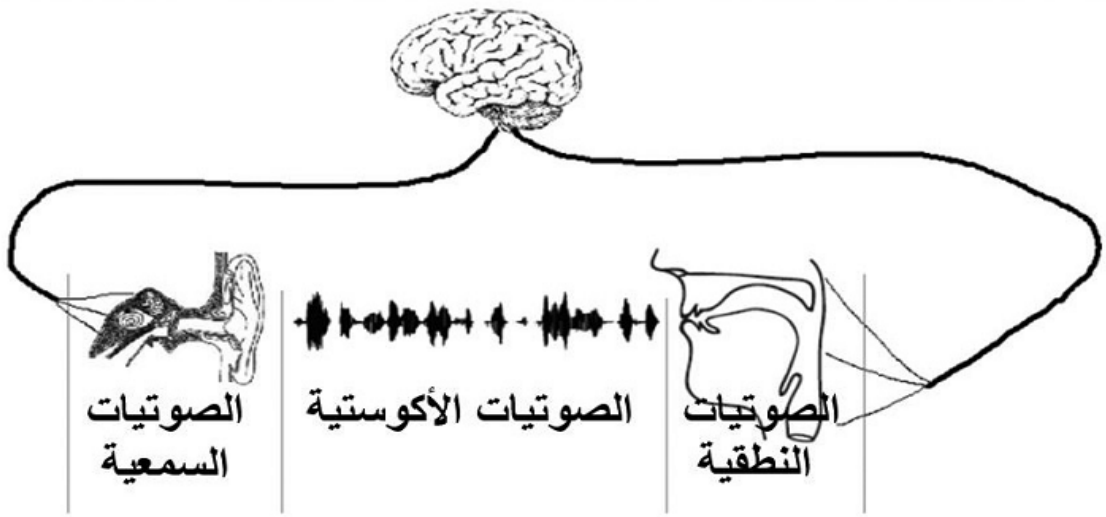
[2] مناف مهدي محمد الموسوي ، علم الأصوات اللغوية ، عالم الكتب ، بيروت ، لبنان ، 1998 م ،ص 18 .

المتكلم فإن الحرف مما يوجده الباحث [1]، أما الصوت فهو الأثر المسموع من التضييق على الهواء في مخرج من مخارج الأصوات ، فاللغات تختلف في الحرف لأن كل لغة لها رموز أحرف تختلف عن رموز أي لغة أخرى ، أما الصوت إذا اتحد في الصفات بين اللغات المختلفة فسيكون واحدا ولا يختلف اختلافا بينما يظهر للسامع العادي فالباء الصوت الشفوي الانفجاري غير المفخم [2].

5. فروع علم الصوتية:

علم الصوتية (phonétique) هو علم يدرس الأصوات البشرية بمعزل عن الوظائف اللغوية التي تؤديها هذه الأصوات ، فهو علم لا ينتمي انتماء صريحا إلى اللسانيات[3].

ينقسم علم الصوتية إلى ثلاثة فروع هي : الصوتيات النطقية و الصوتيات الأكوستوكية و الصوتيات السمعية (الشكل 1.1) .



الشكل 1.1: فروع علم الصوتية الثلاثة [4]

[1] ابراهيم السامرائي ، نزهة الألباء في طبقات الأدباء ، مكتبة المنار، الزرقاء الأردن، ط3 ،1405 هـ /1985م.

[2] Dictionary of language and linguistics. Applied Science Publiahers LTD , London ,p.128

[3] مصطفى حركات ، الصوتيات والفونولوجيا ، دار الأفاق الأبيار –الجزائر ، ص11.

[4] عبد الله بن يحي الفيفي ، مدخل إلى اللسانيات الحاسوبية ، ط1، 2018م ، الرياض، ص 16 .

1.5. علم الأصوات النطقي:(Articulatiry Phonetics)

علم الأصوات النطقي (Articulatiry Phonetics) أو علم الأصوات الفسيولوجي (Physiological Phonetics)[1] ، هو ذلك الفرع من علم الأصوات الذي يهتم بدراسة حركات أعضاء النطق لإنتاج أصوات الكلام ، وهو أقدم الفروع الثلاثة ظهورا ، و أوسعها انتشارا بين اللغويين ، ويختص هذا الفرع من علم الأصوات بثلاثة جوانب :

أولا : دراسة جهاز النطق الذي يصدر الصوت اللغوي ، ويتكون من ثلاثة أقسام رئيسية ، وهي : أعضاء التنفس (الرئتان و القصبة الهوائية) و الحنجرة بما تحويه من أعضاء وغضاريف و أوتار ، و التجاويف الأنفية و الفموية التي تحوي معظم أعضاء النطق ،كاللسان و اللهاة و الحنك اللين و الأسنان [2].

الثانيا: دراسة إنتاج الصوت اللغوي ، أي دراسة إنتاج الأصوات اللغوية المنطوقة ، وتصنيفها ، و التفريق بينها من حيث : مخارجها والكيفية التي تنطق بها .

ثالثا : وظيفة الصوت المنطوق ، ويشمل دراسة الأصوات الوظيفية ، أو وظائف الأصوات ، وتدخل هذه الدراسة ضمن علم الأصوات التنظيمي أو التشكيلي المعروف بالفونولوجيا (Phonology) أو علم الفونيمات (Phonemics) ، غير أن تعدد هذه الجوانب وتنوعها يقتضي تعدداً في مناهج علم الأصوات ، أو يستلزم تقديمه إلى فروع يقابل كل فرع منها جانباً من جوانب الصوت ، و يقوم بدراسته وتحليله وفقاً لطبيعته ومكوناته[3] وهو العلم الذي يعني بتعيين أعضاء النطق وتحديد وظائفها وتبيين دور كل منها في عمله النطقى ، و ما يعرض لها من

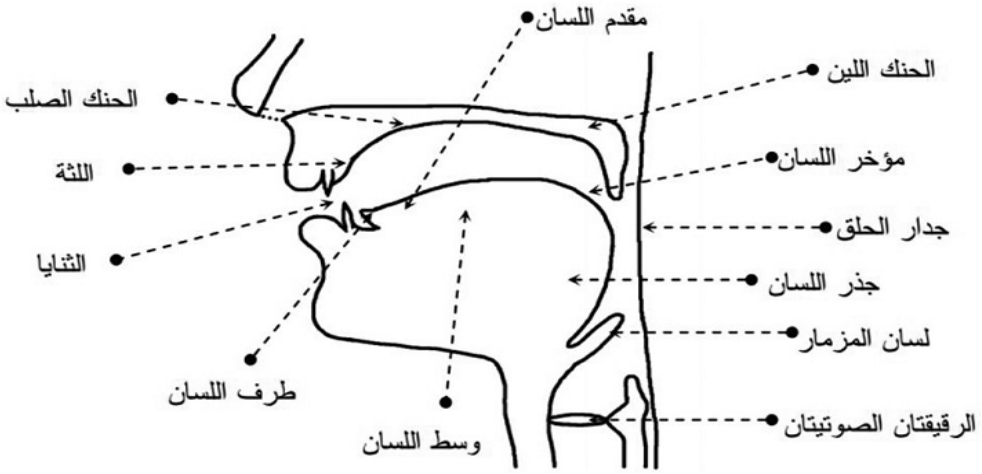
---

[1] أحمد مختار عمر ، دراسة الصوت اللغوي ، عالم الكتب ، مصر، 1997 م.ص77

[2] تمام حسان ، مناهج البحث في اللغة ، ص82.

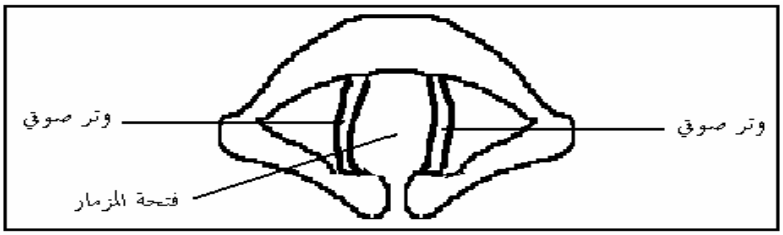
[3] نفسه ، ص23.

حركات ، يخلص إلى تحليل مكانيكية الجهاز النطقي عند إنتاج الأصوات اللغوية المختلفة بواسطة المتكلم هذا ، ويطلق عليه بعض العلماء إسم علم الأصوات الفسيولوجي[1].



الشكل 2.1: أعضاء النطق في الجهاز الصوتي [2]

وأهم عضو في جهاز النطق البشري هو الوتران الصوتيان ، و يلتقيان في الحنجرة تحت اللسان المزمار ، و هما على شكل شفة .



الشكل 3.1: الوتران الصوتيان [3]

[1] المرجع السابق ، الأصوات اللغوية ، ص23.

[2] ابراهيم بوداود ، فيزياء الحركات العربية بين تقديرات القدامى والقياسات المحدثين ، جامعة وهران سنة 2012م ، ص52.

[3] منصور بن محمد الغامدي ، الصوتيات العربية ، مكتبة التوبة ، السعودية ، 2001 م، ص18.

1.1.5. وصف جهاز النطق : و يتكون جهاز النطق الإنساني من الأعضاء التالية :

2.1.5. جهاز التنفس (L'appareil respiratoire) : الذي يقدم الهواء الجاري المطلوب لإنتاج معظم الأصوات اللغوية .

3.1.5 . الحنجرة (la gorge): التي تنتج معظم الطاقة الصوتية المستعملة في الكلام ، و تعد بمثابة صمام ينظم تدفق تيار الهواء .

4.1.5 . التجاويف فوق المزمارية (Cavités au-dessus de la glotte): التي تقوم بدور حجرات الرئتين، وفيها تتم معظم أنواع الضوضاء التي تستعمل في الكلام[1].

5.1.5 . أعضاء التنفس : وهي تشمل الرئتين و القصبة الهوائية :

6.1.5 . الرئتان (Les poumons) : و هما كتلتان مخروطيتان قابلتان للتمدد و الانكماش ، و يهمنا من عملهما أنهما تدفعان الهواء بعد دخوله إليهما في عملية التنفس فيخرج من تجويف جهاز النطق و تتكون الحروف في أثناء خروجه[2].

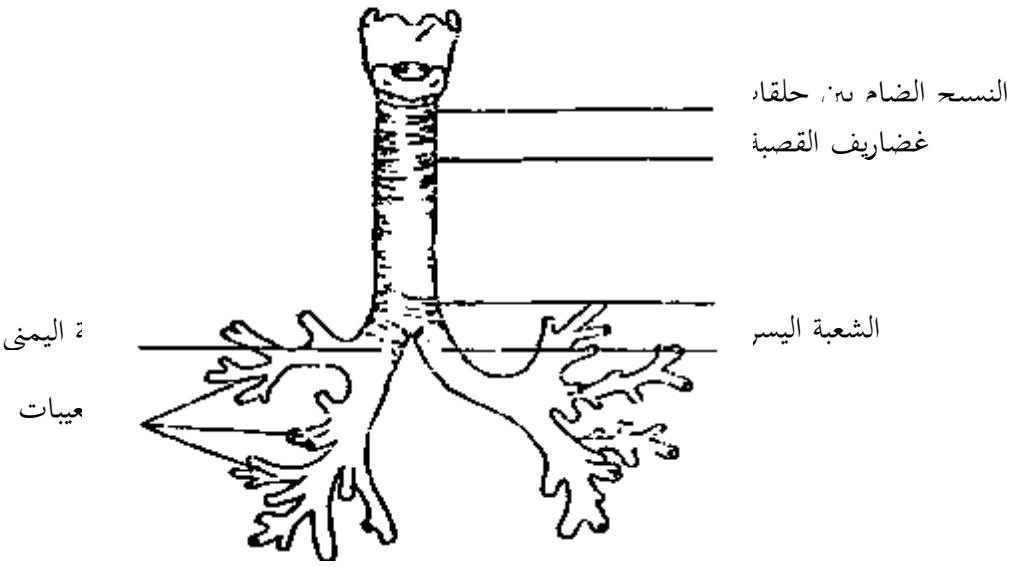
7.1.5. القصبة الهوائية (la trachee artere) : تقع القصبة الهوائية تحت الحنجرة ، و هي عبارة عن أنبوبة مرنة أعلاها قاعدة للحنجرة و أسفلها تتفرع إلى أنابيب أقل منها قطرا ، تصل إلى الرئتين و تتشعب فيهما لتوصل الهواء إليهما ، و دورها في إحداث الصوت أنها توصل الهواء الخارج من الرئة إلى الحنجرة و ما فوقها ، كما أنها تعد غرفة رنين للصوت حين يحدث في الأجزاء العليا منها ،لأنها دائما منفتحة لا تنطبق بحكم الحلقات الغضروفية المكونة لها[3].

---

(1) المرجع السابق ، دراسة الأصوات اللغوي ، ص99.

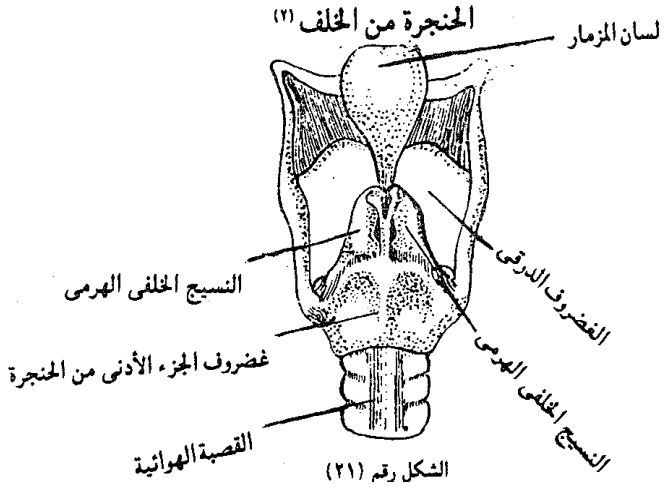
(2) خالد إسماعيل حسان ، في اللسانيات العربية المعاصرة ، مكتبة الآداب ، مصر، 2008 م، ص20.

(3) محمد حسن حسن جبل، المختصر في أصوات اللغة العربية ، مكتبة الآداب ، مصر ، 2012 م ، ص12.



الشكل4.1:القصبة الهوائية

8.1.5.الحنجرة (la gorge) : عبارة عن صندوق غضروفي يقع على قمة القصبة الهوائية و هي مفتوحة من الأعلى و من الأسفل و هذا يسمح بمرور الهواء من القصبة الهوائية إلى الحلق فالفم أو الأنف و العكس[1].

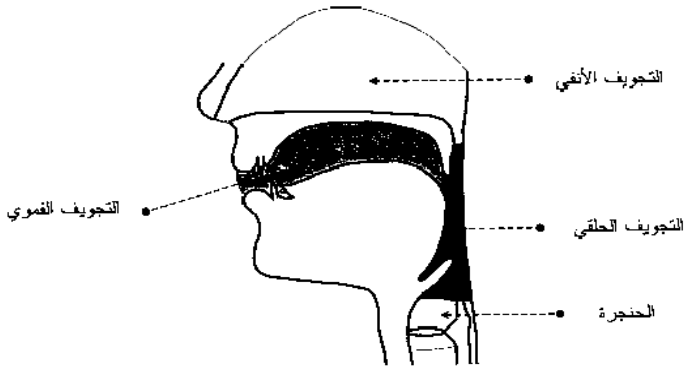


الشكل 5.1: الحنجرة من الخلف[1]

[1] المرجع السابق ، الصوتيات العربية ، ص36.

9.1.5. تجاريف ما فوق المزمار (Cavités au-dessus de la glotte):

اما تجاويف ما فوق المزمار فتشمل ما يأتي.



الشكل 6.1: تجاويف الجهاز الصوتي[1]

أ . تجويف الحلق (La cavité de la guttural):

يمتد تجويف الحلق من الحنجرة إلى بداية التجويف الفموي والتجويف الأنفي ، فهو عبارة عن أنبوب يصل بين الحنجرة من جهة و أنبوبي التجويف الأنفي و التجويف الفموي من جهة أخرى[2].

ب. تجويف الفموي (La cavité buccale):

هو أكثر التجاويف تعقيدا كما أن غالبية الأصوات اللغوية تخرج منه ، و يمتد من الشفتين إلى أعلى التجويف الحلقي و يحتوي على الشفتين ، و الوجنتين ، و الأسنان ، و اللسان ، و الحنك الصلب ، و الحنك اللين و الفكين الأعلى و الأسفل [3].

---

[1] المرجع السابق ، دراسة الصوت اللغوي ، ص33.

[2] داود مسعود ، التعرف الآلي على الكلام العربي المتصل ، جامعة الجزائر، 2016م ، ص11.

[3] المرجع نفسه ، دراسة الصوت اللغوي ص11.

ج. تجويف الأنف (La cavité de la nez):

يتكون التجويف الأنفي من تجويفين يفصل بينهما حاجز يمتد من بين فتحتي الأنف إلى مؤخرة التجويف الأنفي ، حيث يلتقيان في فتحة واحدة تشرف على التجويف الحلقي ، ويفصل الحنك الصلب و الحنك اللين بين التجويفين الأنفي و الحلقي ، و للتجويف الأنفي و ظيفة إحيائية مهمة ، إذا يقوم بتنقية و تلطيف الهواء داخل الرئتين[1] .

وعلم الأصوات النطقي هو من أقدم فروع علم الأصوات وأرسخها قدماً وأكثرها حظاً من الانتشار في البيئات اللغوية كلها ؛ ويرجع السر في ذلك إلى وظيفة هذا الذي سبقا ذكره.

2.5. علم الأصوات الأكوستيكي أو الفيزيائي:

علم الأصوات الفيزيائي Physical Phonetics أو الأكوستيكي Acoustic Phonetics ، فرع من علم الأصوات ، يهتم بدارسة الخصائص المادية أو الفيزيائية لأصوات الكلام منذ خروجها من فم المتكلم أو من أنفه أو منهما معا ، وانتقالها في الهواء أو في الوسط الناقل[2] حتى إلى أذن السامع ، و الدراسة في هذا الفرع من علم الأصوات تشمل جانبين :

أولا: دراسة الموجات و الذبذبات الصوتية التي يحدثها المتكلم باستعمال أجهزة الصوت و أعضائه المشارة في علم الأصوات النطقي .

ثانيا : دراسة الوسيط ، وهو الوسط الذي ينتقل عبره الكلام إلى أذن السامع [3].

ويطلق علماء الأصوات على كلا الجانبين مصطلح : أكوستيكية الصوت أو فيزيائية الصوت ، الذي يشمل فيزيائية الصوت بوجه عام وفيزيائية الصوت الكلامي بوجه خاص ، ويعالج عددا من المفاهيم الصوتية ،

---

[1] المرجع السابق ، مناف مهدي محمد الموسوي ، علم الأصوات اللغوية ، ص33.

[2] حسام البهنساوى ، علم الأصوات، مكتبة الثقافة الدينية، ط1 /2004 ، ص13.

[3] المرجع نفسه ، ص 48.

كحدوث الصوت ومصدر الصوت وحركته ، والموجات الصوتية ، وانتقال الصوت ، والذبذبات والتردد ، والحزم الصوتية .

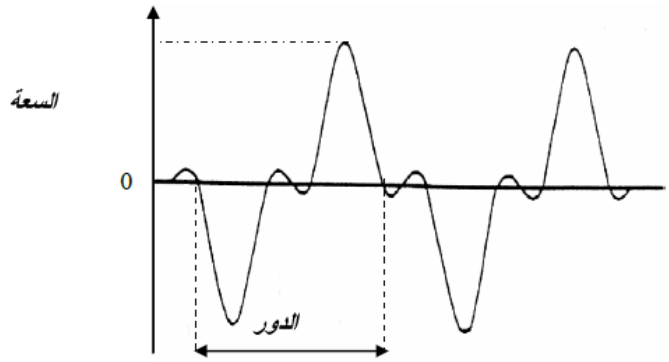
1.2.5. خصائص الصوت اللغوية الفيزيائية: (Les propriétés physiques des sons)

للصوت في الفيزياء خصائص كثيرة ، ولكنها لا تعنينا كلها في هذه الدراسة ، وإنما تميز في الأصوات وسنكتفي منها بالخصائص التي يمكن للأذن البشرية تمييزها [1] ، فهي مركبة إلى ثلاث صفات فسيولوجية هي : الشدة والارتفاع والطابع وسنعرف هذه الخصائص وعللها الفيزيائية واحدة تلو الأخرى .

1.2.5.أ. الشدة: (Intensité)

هي الصفة الفسيولوجية التي تميز فيها الأذن الصوت الشديد القوي من الصوت ، كأن يتحدث الإنسان بصوت مرتفع ، أو يهمس همسات خفيفة ، أو يستمع الشخص إلى حديث آخر مباشرة ، أو بمكبر صوت[2].

وعلتها الفيزيائية هي سعة اهتزاز طبقة الهواء بجوار الأذن ، التي ينتج عنها تغيرات محسوسة في الضغط وتسمى أيضا «علو أو حجم الصوت Loudness»



الشكل 7.1: الموجة الصوتية المركبة[3]

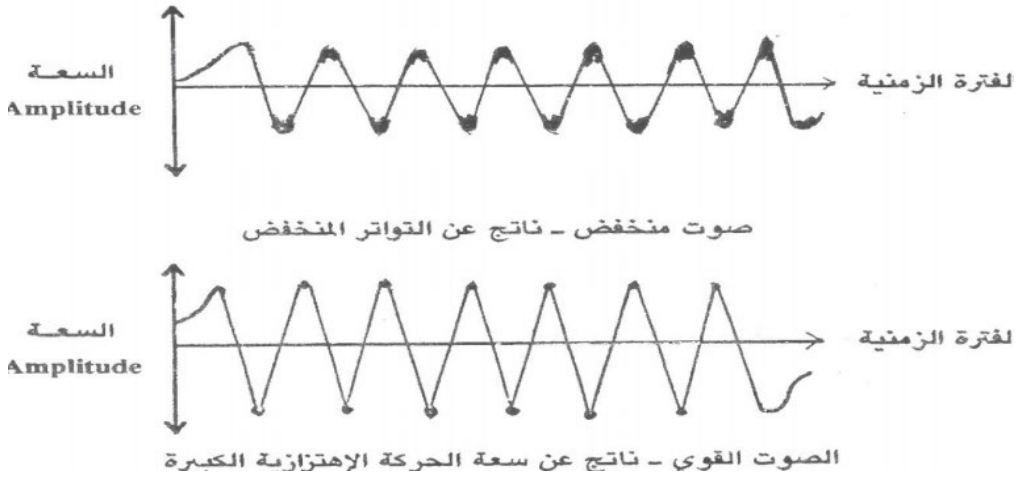
ويتحدد ذلك بالنظر في سعة الذبذبة التي تمثل البعد بين نقطة الاستراحة وأبعد نقطة يصل إليها الجسم المتحرك.

---

[1] هشام جبر، نظرية الاهتزازات و الأمواج الميكانيكية ، ديوان المطبوعات الجامعية ، 1999م ، الجزائر ، ص121.

[2] رضا الزلاقي ، الصوامت الشديد في العربية الفصحى ، جامعة الجزائر، 2006م ، ص 34.

[3] المرجع السابق ، فيزياء الحركات العربية بين تقديرات القدامى والقياسات المحدثين ، ص67.



الشكل 8.1: خصائص الموجة الصوتية [1]

1.2.5.ب. الارتفاع : (La hauteur)

هو الصفة الفسيولوجية التي تميز فيها الأذن الصوت الحاد الرفيع من الصوت " الغليظ الأجش ، كالاختلاف بين صوتي المرأة والرجل ، وبين زقزقة العصافير ونعيق الغربان " .

والعلة الفيزيائية لاختلاف الأصوات في الارتفاع هي الاختلافات في تواترها ، وتزداد الأصوات حدة بزيادة التواتر.

يعرف الارتفاع أيضا درجة الصوت أو حدته (pith) ويطلق عليه أيضا مصطلح التردد ويقاس تردد حركة الجسم أو تردد الذبذبات بعدد الدورات في الثانية ، والدورة عبارة عن تكرار كامل لنمط الموجة.

1.2.5.جـ. الطابع: (Le timbre)

وهو الصفة الفيزيولوجية التي تميز فيها الأذن بين صوتين متماثلين شدة وارتفاعا يصدرهما منبعان مختلفان «نوع الصوت Timbre» ، كأن نتعرف على صوت صديق دون أن نراه ، أو نتعرف على نوع آلة من سماع تسجيل صوتها.

---

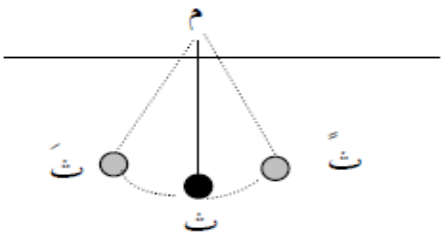
[1] ابن سينا ، تحقيق محمد حسان ، أسباب حدوث الحرف الطّيان ومراجعة شاكر الفحام وأحمد راتب النفاخي ، مطبوعات مجمع اللغة العربية ، دمشق 1983م، ص26.

علته الفيزيائية : إذا سجلنا صوتين يصدر أحدهما عن رنانة ، ويصدر الآخر عن كلارينت ، لهما ارتفاع واحد وشدة واحدة فنجد أن الأول منحنى جيبي والثاني دوري ، فنقول أن طابع الصوت له علاقة بشكل المنحنى الممثل للصوت الدوري ، ولما كان أي منحنى دوري ناتجا عن تركيب عدة منحنيات جيبية هي مدروجاته ، أمكننا أن نستنتج العلة ، دوري ناتجا عن تركيب عدة منحنيات جيبية ، أمكننا أن نستنتج العلة الفيزيائية لاختلاف الطابع في الأصوات .

تنقسم هذه الخصائص إلى نوعين : الخصائص أو الصفات الفيزيائية المحظة و الخصائص أو الصفات الفيزيائية النفسية ، الصفات الأولى صفات موضوعية راجعة إلى طبيعة الظاهرة الصوتية نفسها أما الثانية هي صفات ذاتية ناتجة عن إدراك الأذن للأصوات[1].

2.2.5. الصفات الفيزيائية المحظة :

3.2.5. الأصوات الدورية و الأصوات الغير دورية : إن الصوت من جنس الظواهر الاهتزازية ، حيث أن لهذه الحركة الاهتزازية مميزات تتوقف عليها طبيعة التموج و بالتالي طبيعة الصوت ، ومن هذه الخصائص الكيفية التي تهتز بها فإذا حدثت الحركة الاهتزازية في فترات متساوية كحركة النواس أو الرقّاص نقول أنها حركة دورية مثلما يوضح الشكل الأتي :



الشكل 9.1 : النواس [2]

[1] المرجع السابق ، الصوامت الشديدة في العربية الفصحى ، ص99 .

[2] المرجع السابق ، مبادئ في اللسانيات ، دار القصبة للنشر، الجزائر، 2006 م ،ص 38.

3.2.5. أ. التردد أو التواتر (Fréquence): [1]

هي حركة اهتزازية معينة أي عدد الدورات الكاملة في الثانية ، أو عدد الذبذبات في الثانية وكل جسم يتذبذب يكون له تردد خاص يتحكم فيه الوزن والطول ونسبة شد الأوتار الصوتية وكتلة التجاويف وشكلها (فمثلا كلما خف وزن الجسم المتذبذب يزداد تردده) ، ويشار إلى أن التردد يقاس بعدد الدورات في الثانية Cps (Cycles par second).

- الجسم ثقيل يتذبذب تذبذبا أبطأ من تذبذب الجسم الخفيف .
- الشوكة الرنانة ذات الذراعين الطويلين تتذبذب أبطأ من الشوكة ذات الذارعيين القصيرين.

3.2.5. ب . السعة (Capacité)

تمثل سعة الذبذبة (Amplitude) البعد بين نقطة الاستراحة ، وأبعد نقطة يصل إليها الجسم المتحرك ، وسعة الذبذبة هي المسئولة عن التوتر (Intensity) ، فكلما زاد الاتساع زاد التوتر ، والمصطلح العلو (Loudness) و الوصف الذي يطلق على التوتر المدرك (perceived intensity) [2].

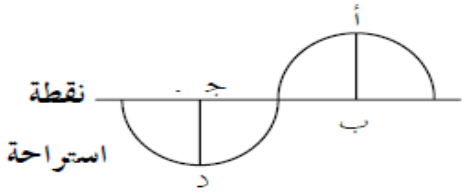
ويمكن توضيح سعة الذبذبة بالشكل 10.1 وهي المرموز إليها بالرمزين أ–ب– والرمزين جـ –د

---

[1] عصام نور الدين ، علم الأصوات اللغوية الفونيتيكا ، دار الفكر اللبناني ، لبنان ، 1992م، ص 20.

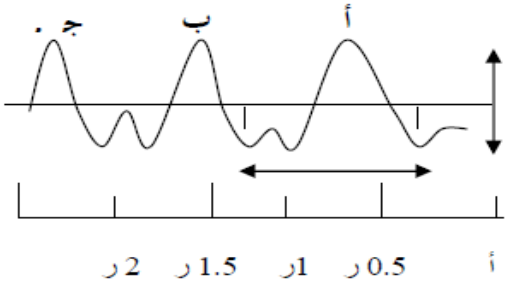
[2] المرجع السابق ، دراسة الصوت اللغوي ، ص25.

وسنمثل الآن الأصوات بالأمواج للتعرف على ماهية هذه الأمواج من ناحية التردد، والسعة ، ونقطة الاستراحة ، وغيرها.



الشكل 10.1: يرمز البعد بين (أ و ب) و (ج ود)

نقطة الاستراحة إلى سعة الذبذبة



الشكل 11.1: يعبر عن الدورات في الثانية (التردد)

3.5. علم الأصوات السمعي (Auditory Ponetics) :

علم الأصوات السمعي (Auditory Ponetics) فرع من علم الأصوات على الإطلاق ، يهتم بدراسة الإدراك السمعي لأصوات الكلام بعد وصولها إلى أذن السامع ، وهو ذو جانبين : جانب عضوي أو فسيولوجي، وجانب نفسي[1].

[1] المرجع السابق ، دراسة الصوت اللغوي ، ص 23.

أما الأول فوظيفته النظر في الذبذبات الصوتية التي تستقبلها أذن السامع ، وفي ميكانيكية الجهاز السمعي ووظائفه عند استقبال هذه الذبذبات وهي مرحلة تقع في مجال علم وظائف أعمال السمع (Physiology of Hearing) .

أي أنه يهتم بدراسة الذبذبات وموجات الصوت لحظة وصولها إلى الأذن وكيفية استقبالها وتحولها إلى رسائل مرمزة عبر الأعصاب إلى الدماغ ثم حل هذه الرموز في الدماغ ويذكر الجانب الثاني جهوده على البحث في تأثير هذه الذبذبات ووقعها على أعضاء السمع (الداخلية منها بوجه خاص) [1].

وهذان الجانبان متصلان غير منفصلين ، فهما وجهان لشيء واحد ، أو خطوتان متتاليتان لعملية استقبال الأصوات ، هذا وعلى الرغم من أن معظم الدارسين يجمع هذين الجانبين تحت إسم علم الأصوات السمعي إلا أن هناك من يجمعها تحت إسم علم الأصوات النفسي (Psychological Phonetics) مرجحين بذلك الجانب النفسي من العملية السمعية على جانبها الفسيولوجي ؛ على أساس أن الجانب النفسي هو ذو الأثر الواضح في سلوك السامع عند إدراكه للأصوات اللغوية المختلفة [2].

و الأذن هي أداة السمع ، أو الإلتقاط الذي يتلقى الإشارة الصوتية و يحولها إلى حركة تدب الأعصاب ، و تنتقل إلى الجهاز العصبى المركزي ،و تنقسم الأذن إجمالا إلى ثلاثة هي:

1- الأذن الخارجية the outer ear.

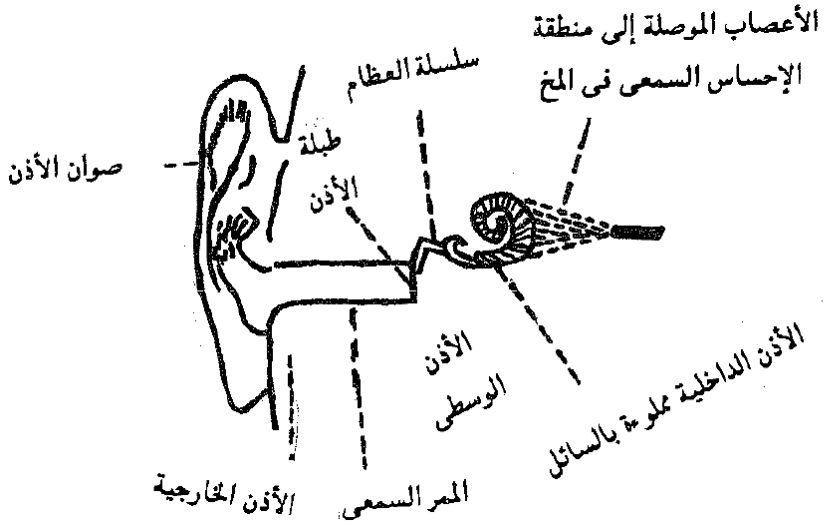
2- الأذن الوسطى the middle ear.

3- الأذن الداخلية the inner ear.

---

[1] عصام نور الدين ، علم الأصوات اللغوية ، الفونيتيكا ، دار الفكر اللبناني ،بيروت ، ط1 /1992 ، ص 153.

[2] المرجع السابق ، علم اللغة العام ، علم الاصوات ، ص 249.



الشكل12.1 : الأجزاء الرئيسية للأذن [1]

الرسم التوضيحي يبين مجال السمع في إطار حدود التردد و الشدة لكل النغمات المسموعة .

و الخط الأعلى في الرسم يمثل المستوى الذي تبدأ الأصوات عنده في تسبب الشعور بالألم في الأذن ، فإذا تجاوزت قوة الصوت 130 ديسيبل يوجد الشعور بعدم الراحة .

6. مفهوم المصوت أو الفونيم (le phonème)

إنطلاقا من مفهوم الترادف الصوتي ، يمكننا أن نعرف المصوت أو الفونيم بأنه مجموعة من الأصوات المترادفة سواء كان هذا الترادف ناتجا عن تنوع حر أو عن تنوع تركيبي [2] وهو أهم الوحدات الصوتية و أساس التحليل الصوتي ، غير أن اللغويين اختلفوا في حقيقته وتعريفه و تحديده ، يوصف الفونيم بأنه أصغر وحدة صوتية تعتمد على المخالفة التي تعني أن استبدال فونيم بأخر في كلمة ما يستلزم تغير معنى الكلمة أو وظيفتها [3]،فالنون والتاء في اللغة العربية فونيمان مختلفان لكل منهما وظيفة مميزة ، فكلمتا (ناب) و(تاب) يختلف معناهما بسبب

---

[1] المرجع السابق ، دراسة الصوت اللغوي ، ص25.

[2] المرجع السابق ، الصوتيات والفونولوجيا ، دار الأفاق الأبيار –الجزائر ، ص 26.

[3] محمد سعران ، علم اللغة ، مقدمة القارئ العربي ، ص 192.

وجود النون في الأولى و التاء في الثانية . وهذا التميز الفونيمي يكون في الصوائت أيضا كما في التقابل بين الفتحة و الضمة في كلمتي : (كَرَم) و(كَرُم) ، الفتحة فونيم و الضمة فونيم أخر.

وكذلك التقابل بين الحركات القصيرة و الطويلة ، فالفتحة القصيرة و الفتحة الطويلة في كلمتي : (مَطَر) و (مَطار) فونيمان مختلفان .

أ- الفونيم وحدة صوتية تميز كلمة من أخرى ؛ أي تقوم بالتفريق بين الكلمات من النواحي الصوتية والصرفية والنحوية والدلالية.

ب- الفونيم وسيلة مهمة في تسهيل عملية تعليم اللغات الأجنبية فالأصوات الفعلية المنطوقة في أية لغة كثيرة كثرة فائقة ، في حين أن فونيمات كل لغة تقل في عددها عن عدد هذه الأصوات المنطوقة بالفعل وبصورة ملحوظة.

ج- لفكرة الفونيم دور مهم في ابتكار الألفبائيات أو نظم الكتابة بصورة ميسرة [1]

مصطلح الفونيم (Phonème) أو علم الفونيم كما يسميه بعض اللغويين وهو يدرس في مظلة علم الفونولوجيا (Phonology) أي علم وظائف الأصوات الذي يبحث في الأصوات من حيث وظائفها في اللغة ومن حيث إخضاع المادة الصوتية للتقعيد ويمكن تقسيم الفونولوجيا حسب المركبات التي تكونها[2]إلى قسمين ، هما :

أ- **الفونولوجيا القطعية:** Segmental Phonology ويقوم هذا النوع على تحليل الكلام إلى قطع متميزة تسمى الفونيمات ومن أمثلة ذلك في اللغة العربية الصوامت والحركات.

[1] المرجع السابق ، دراسة الصوت اللغوي ، ص4.

[2] المرجع السابق ، مناهج البحث في اللغة ، ص111.

ب- الفونولوجيا فوق القطعية : (Suprasegmental Phonology) أو الفونولوجيا التطريزية (Prosodic Phonology)، وهي ظواهر صوتية تنبئ عن خواص الكلام وتحدد نوعياته ، وكيفيات أدائه ، وذلك مثل المفصل ، والتنغيم ، والنبر.

سجل العلماء تعريفات كثيرة جدا للفونيم وتختلف تلك التعريفات فيما بينهم حسب المدرسة الصوتية التي تناولته ومن أبرز التعريفات للفونيم ما ذهب إليه ماريو باي ، عندما قال عنه إنه : العلم الذي يعالج الخصائص الصوتية الوثيقة الصلة بلغة معينة من وجهة إحساس المتكلمين... وإذا كان من الممكن أن يشتمل الفونيم على صوت واحد: فون Phone (أو صوت موضوعي) فهو في الكثير الأعم يشتمل على مجموعة من الألفونات المتشابهة أو التنوعات الصوتية (Phonetic Variants) التي يتوقف استعمال كل منها أساسا على موقعه في الكلمة (أولا ، وسطا ، آخر) وعلى الأصوات المجاورة لها [1].. وهذا التعريف يضع لنا أسسة للفونيم فهو يتحدث عن خصائص الصوت البشري وتنوعاته المختلفة وما يطرأ على هذا الصوت من تغيرات.

1.6. اختبار فونيمة الصوت:

ليس كل صوت فونيما ، فكيف يمكننا تمييز صوت فونيما من غيره ؟ الفونيم كما سبق تعريفه قادر على تمييز معنى الكلمات أو تغييره إذن , فتمييز المعنى أو تغييره معيار أساسي يمكن الاعتماد عليه في تحديد فونيمية الصوت، وذلك يتم من خلال ما يسمى باختبار التبادل أو اختبار الصوتين من خلال الثنائية الصغرى (minimal pair,)أي اختبار الصوتين المراد اختبارهما بوضعهما في كلمتين متفقين في جميع أصواتها.

البحث عن كلمتين تتفقان وتتشابهان في جميع أصواتهما إلا في الصوتين المراد اختبارهما ، لإختبار الصوتين السين / س/ والصاد / ص/ مثال ، فلا بد من وضعها في كلمتين تتفقان في جميع أصواتهما وتختلفان في أن

[1] المرجع السابق ، دراسة الصوت اللغوي ، عالم الكتب ، مصر، 1997 م ، ص88.

إحداهما تشتمل على السين والأخرى على الصاد مثل (سار) و (صار). وفيما يلي مزيد من مثل هذه الثنائية التي يمكن استخدامها في اختبار فونيمية الأصوات[1]

الرقم	الثنائيات الصغرى	الأصوات المراد إختبارها
1	أليم – عليم	الهمزة والعين
2	تاب – طاب	التاء والطاء
3	جهر – ظهر	الجيم والظاء
4	كلب – قْلب	الكاف والقاف
5	جلس – جرس	اللام والراء

الجدول 1.1: اختبار فونيمية الأصوات

بالنظر إلى تبادل الصوتين في كل الثنائية: هل تبادلهما يؤدي إلى تغيير المعنى أو لا. إذا أدى إلى تغيير المعنى فهما فونيمان مختلفان كالهمزة / ء/ والعين /ع/ في الثنائية (أليم – عليم) , كذلك بقية الأصوات في جميع الثنائيات في الجدول السابق[2]

[1] نصر الدين إدريس جوهر، علم الأصوات لدراسي اللغة العربية , مكتبة لسان عربي, 2014 ، ص 154.

[2] المرجع السابق ، علم الأصوات الدراسي للغة العربية من إلندونيسيين, سدورجو ، ص 157.

2.6. مدارس الفونيم :

1.2.6. المدرسة النفسية : تعد هذه المدرسة " الفونيم صوت نموذجية يهدف المتكلم إلى نطقه ولكنه ينحرف عن هذا النموذج وقد يعود سبب ذلك لأنه من الصعب أن ينتج صوتين مكررين متطابقين أو لنفوذ الأصوات المجاورة أو أنه يعد صوتا مفردا وله تجريد ذهني أو صورة ذهنية ، يستحضرها المتكلم إلى عقله بالإرادة ويحاول بلاوعي أن ينطقها في الكلام فينجح في بعض الأحوال في تحقيق صورة الصوت بالنطق ولكنه في أحوال يخفق فيستحضر أقرب الأصوات إلى هذه الصورة وهذا شبيه بنظرية المثل عند أفلاطون[1].

2.2.6.المدرسة الفيزيائية : من رواد هذه المدرسة دانيل جونز ، الذي عرف الفونيم بأنه :" أسرة من الأصوات في لغة معينة متشابهة الخصائص ومستعملة بطريقة لا تسمح لأحد أعضائها أن يقع في كلمة في السياق الصوتي نفسه الذي يقع فيه الآخر".

ويفهم من هذا التعريف :

1. أن الفونيم لا بد أن يكون عنوان على مجموعة أصوات محكومة بالسياق.

2. أن أيا من أفراد هذه العائلة لا يمكن أن يرد في السياق الصوتي الذي يرد فيه الآخر.

3.2.6. المدرسة الوظيفية : يعد العالم تروبتسكوي ، رائد المدرسة اللغوية في "البراغ" منظر هذه المدرسة ؛ فهو يرى أن الفونيمات هي أصغر وحدات اللغة التي تستطيع " بطريق التبادل" أن تميز كلمة من كلمة أخرى.

ففي أعلى مستوى ، النص مركب من وحدات ذات دلالة (جمل ، شبه جملة ، كلمات ...) وأصغر هذه الوحدات تسمى مونيمات (monémes) . وفى المستوى الثاني فإن كل مونيم مركب من ناحية الشكل إلى وحدات خيالية من المعنى ، أصغر هذه الوحدات هي الفونيمات (phonémes) [2].

---

[1] محمد جواد النوري ، الفونيم وتجلياته في القرآن الكريم ، برواية حفص عن عاصم ، سورة البقرة ، 2014 ، ص32.

[2] المرجع السابق ، اللسانيات العامة و قضايا العربية ، ط1، مكتبة العصرية ، بيروت ،1998م ص21.

إن أراء هذه المدرسة ، جاءت مغايرة لآراء المدرسة الفيزيائية ، "لأنها تدخل التفرقة بين المعاني في تعريف الفونيم وما دام كل من (k) و (و) لا يفرقان بين المعاني في الإنجليزية فلا يعتبران فونيمين مختلفين ولكنهما يفرقان بين المعاني في العربية لذا يجب اعتبارهما فونيمين مختلفين في العربية.

هذه هي أبرز المدارس اللغوية التي تبنت نظرية الفونيم وعالجته وتوجد آراء أخرى عالجت نظرية الفونيم نذكر منها نظرية المدرسة التجريدية التي تعد الفونيمات مستقلة استقلالا كاملا عن الخصائص الصوتية المرتبطة بها كذلك توجد آراء حول نظرية الفونيم للعالم اللغوي دي سوسور الذي طالب الاعتماد على التأثير السمعي للتمييز بين الأصوات ، ويبدو ذلك جليا في قوله : إن الانطباع السمعي هو أساس أية نظرية صوتية فالانطباع السمعي له وجود لا شعوري عند المرء يسبق دراسة الوحدات الفونولوجية [1].

3.6.مكونات الفونيم : تنقسم الفونيمات إلى قسمين :

1- الفونيمات الأساسية

2- الفونيمات الثانوية

فالفونيمات الأساسية هي مجموعة من الصوامت والحركات التي تتكون منها اللغة ، والتي تشكل في طبيعتها أجزاء الكلام ، لذلك تسمى بالفونيمات التركيبية أو الجزئية ، و أما الفونيمات التي تتمثل ملامح صوتية تؤثر في الأصوات ولا تدخل في التركيب فهي الفونيمات الثانوبة ، وتسمى فوق التركيبة أو فوق القطعية : النبرة والتنغيم وغيرها[2].

الإتجاه السائد لدى العلماء الأصوات أن الفونيم يتكون من أسرة أو أنه وحدة صوتية " تجمع تحتها متعددات " ولكن العلماء اختلفوا في ماهية هذه المتعددات [3]، فأنقسموا إلى قسمين :

---

[1] المرجع السابق ، الفونيم وتجلياته في القرآن الكريم ، برواية حفص عن عاصم ، سورة البقرة ، 2014 ، ص34.

[2] باي ، ماريو ، أسس علم اللغة ، ترجمة أحمد مختار عمر ، القاهرة : عالم الكتب ، القاهر ، 1987م ، ط1، ص92

[3] المرجع السابق ، دراسة الصوت اللغوي ، ص 183.

القسم الأول : يرى أن الفونيم يتألف من مكونات تتمثل في تحققاته الصوتية التي يطلق عليها مصطلح ألوفونات (Allophnes) ولا يمكن تحديد ألوفون لفونيم ما إلا داخل السياق أو الموقع الصوتي الذي يرد فيه .

القسم الثاني : يرى أن الفونيم يتكون من الملامح الصوتية التمييزية (Distinctive Features) وتعنى خصائص صوتية يمكن أن تميز معنى منطوق من معنى منطوق أخر.

ومثال ذلك الكلمات الآتية : **ثلم** ، **وذلم** ، **وظلم** ، ففونيم الثاء يتألف من مجموعة من الملامح التمييزية الآتية وهي: **أسناني** ، **إحتكاكي** ، **مهموس**[1] .

أما فونيم الذال فهو يتألف من مجموعة من الملامح التمييزية الآتية وهي: **أسناني** ، **احتكاكي** ، **مجهور** .

أما الفونيم الظاء يتألف من مجموعة من الملامح التمييزية الآتية وهي: **أسناني** ، **احتكاكي** ، **مجهور** ، **مفخم** .

ففي الكلمات السابقة يوجد تقابل كبير في الملامح التمييزية ، ولكن الذي جعل كل فونيم منها (الثاء ، و الذال و الظاء) يختلف عن الآخر ، هو وجود ملمح واحد على الأقل مختلف كليا عن الآخر ، ففونيم الثاء يتمتع بملح الهمس ، في حين يتمتع فونيم الذال بملمح الجهر ، و فونيم الظاء يختلف عن فونيم الثاء بملمح الجهر والتفخيم ويختلف فونيم الظاء عن فونيم الذال بملمح التفخيم[2] .

7. الخاتمة :

تناولنا في هذا الفصل المفاهيم الأساسية التي ينبغي أن يعرفها كل دارس في مجال الأصوات اللغة العربية التي يصدرها الإنسان ويشكل منها كلماتها وتراكيبها ليدل بها على المعاني التي يريد توصيلها لغيره في مواقع استعمالها المختلفة بغية استخلاص القوانين الصوتية التي تكشف عن أحوالها في الكلمة أو الكلام ، التي تلعب دورا رئيسا في إكتمال النظام التواصلي بين أفراد المجتمع البشري ، والتي يمكن إجمالها بتناول المفاهيم العامة المتعلقة بالصوتيات وفروعها وأهميتها ، وفيه أيضا نلقي الضوء على خصائص الأصوات اللغة العربية و قد انصب اهتمام

---

[1] المرجع نفسه ، ص 35.

[2] المرجع السابق ، ص36.

الدراسات العربية القديمة على الجانب الفيزيولوجي للأصوات اللغوية ، فتحدثنا عن أعضاء النطق في معرض حديثهم عن إنتاج الأصوات و ما يعتري هذه الأعضاء من تغيرات من شأنها أن تعطي أثرا واضحا في الأصوات الناتجة ، و كانت وسيلتهم إلى ذلك الملاحظة الذاتية المعتمدة على الحس و التجربة و والممارسة الشخصية. فأثناء النطق المتكرر يتمكن الدارس من تحديد حركات أعضاء الجهاز الصوتي لدي و وباستعمال السمع يستطيع وصف الأصوات المنتجة ، وهو علم يبحث في مجال الأصوات اللغوية من حيث مخارجها وكيفية إخراجها وخواصها الأكوستيكية كموجات صوتية وكيف يتم سماعها وإدراكها ، و إلمام بكافة الفروع المعرفية الأخرى كفيزيائية الموجات الصوتية والتشريح ووظائف الأعضاء وعلم النفس للدخول إلى عالم الأصوات .

إن كل الدراسات و الأبحاث المتعلقة بالكلام لابد أن تمر بدراسة معمقة لمكوناته الأساسية و التي هي الأصوات اللغوية (الفونيمات) ، لما لها من خصائص وصفات تمكن من فهم الظواهر اللغوية المختلفة و الاستفادة منها، وقد تنوعت أساليب هذه الدراسة بين لغوية تدرس الأصوات من ناحية وظيفتها اللغوية و بين العلمية التي تدرسها كظاهرة طبيعية.

وقد قمنا بدراسة الأصوات اللغوية العربية من حيث مصدرها و خصائصها النطقية فتوصلنا إلى النتائج التالية:

- من خلال دراسة البنية التشريحية لأعضاء جهاز النطق و تحديد الأعضاء المسؤولة عن إحداث كل صوت على حدة ، كما لاحظنا توزيع الأصوات على طول جهاز النطق مما يكسبها صفات تساعد على التمييز بينها بسهولة.
- دراسة الأصوات تمكن نطقها بصورة صحيحة مما يساعد على الاستفادة من خصائصها.
- دراسة الفونيمات اللغة العربية وأنواعها التي تساعد في اختيار المعاملات المناسبة لتحسين أنظمة التمييز على الكلام.

الفصـل الثاني :

تقتيات المعالجة الإشارة الكلامية

1. مقدمة :

الكلام هو وسيلة الاتصال بين أفراد الجنس البشري ، وهو هبة الله للبشر ميزهم بها 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 ، فالكلام هو دينامكية الإشارة التي تحتوي على المعلومات الأساسية والتي تدعى بالموجة الصوتية ، هذه الموجة تنتج خلال عملية الضغط داخل فم المتكلم وهي نتيجة سلسلة حركات متناسقة ومتتالية لتركيب النظام اللفظي لدى الإنسان ، وقد تتحول كمدخل إلى نظم معالجة لغوية لاحقة للحصول على فهم الكلام لإعطاء تفسير لهذا الكلام والتصرف على أساسه.

تعد تقنيات التعرف على الكلام من أهم التقنيات الحديثة وقد تم تطوير العديد من الأنظمة المختلفة من حيث الطرق المستخدمة في استخراج السمات وطرق التصنيف ، تقنية التعرف على الكلام هي عملية استقبال الكلمات أو الجمل المنطوقة من خلال ميكرفون وتحويلها بدقة وكفاءة إلى مجموعة من الكلمات و يتحقق ذلك عن طريق اتباع خطوات معينة باستخدام نظام التعرف على الكلام .

وبما أن نظم تعرف الكلام تحتك بشكل كبير مع علوم أخرى فقد يكون من الصعب تحديد قواعدها الحقيقية التي من شأنها أن تؤدي إلى نتائج مرضية لذلك سنكتفي بما توصلت إليه البحوث في مجال معالجة الإشارة.

ظهر مصطلح تخاطب الإنسان مع الحاسوب (Human Computer Interface) ليصف كيفية التفاعل والتخاطب بين الإنسان و الحاسوب ، ثمة حاليا طيف ضخم من البرامج و الأجهزة التي تتطلب تفاعلا بين البشر والحاسوب ، قد يكون هذا التخاطب والتفاعل صوتيا أو كتابيا ، إن أشهر تقنيات لمعالجة الإشارة الكلامية التي تتيح تخاطبا نصيا وصوتيا بين الإنسان و الحاسوب هي نظم تحويل النصوص إلى كلام Text to Speech) System) ونظم تعرف الكلام (Speech Recognition Systems)

2. الكلام : (Seepch)

يوصف الكلام بأنه يمثل الصوت الناتج من جهاز التكلم الخاص بالإنسان (Speech System Production) لإنشاء تمثيل لغوي ينقل المعلومات من المستهل بالحديث إلى المستقبل للحديث ، أما الهدف منه فيتركز في نقل المعنى بين الأشخاص أو لزيادة الرابطة الاجتماعية بين الأشخاص أو الجماعات[1].

كما أن مفهوم الكلام يتضمن إلى جانب المحتويات الواقعية الحقيقية لإشارة الكلام معلومات عن شخصية المتكلم وهويته (من الذي يتكلم) ، الانفعال ، التنغيم ، فضلاً عن مضمون الرسالة المنقولة بالكلام ، كل هذه العوامل مهمة بالنسبة إلى مفهوم الكلام.

3. تحليل الكلام : (Speech Analysis)

إن الهدف الرئيسي لتحليل الكلام هو تمييز الإشارة الصوتية وكذلك التقليل من عرض الحزمة (bandwidth) وتمييز الإشارة الكلامية بعلامات قليلة ، وتتم هذه العملية في حيز من الزمن (Time domain) الذي يشمل إطارات مقاطع البيانات لاختبار مقاطع كلامية صغيرة (Frame) ، للبيانات التي يفترض أن تكون ثابتة خلال المدة الزمنية فضلا عن تحليل الكلام الذي أتاح للبشر التعامل مع الماكنة ، إذ استطاع الإنسان بناء دوائر إلكترونية دقيقة ذات مساحة صغيرة تمكنت من ترجمة أوامر البشر وفهمها [2] .

4. صعوبات تمييز الكلام:

تكمن صعوبة مسألة تمييز الكلام في تنوع شكل الإشارة الكلامية الموافقة للكلمة نفسها ، و يمكن تصنيف هذه التغيرات كما يأتي:

---

(1) Philips , " A multi-Band Spectral Subtraction Method for peech Enhancement". C.,(2001), http:// www. Utdallas.edu/~Liozou/suni-ms.pdf.

(2) Childers,D. G. ,(2000), Speech Processing and Synthesis Toolboxes ", John Wiley & Sons ,Inc.

1.4. تغيرات تتعلق بالمتحدث نفسه:

يختلف نطق المتحدث الواحد لنفس العبارة من وقت لآخر،[1] إذ يمكن أن تنشأ تغيرات صوتية بحسب تغيرات الحالة الفيزيولوجية والنفسية للشخص (هل هو مريض أو صحيح ، حزين أو فرح أو غاضب...؟).

2.4. تغيرات بين المتحدثين:

تتبع الخلفية الاجتماعية اللغوية لهم مثل ، اللهجة و أبعاد و شكل جهازهم الصوتي الفيزيولوجي.

3.4. اختلاف محيط المتحدث:

و هذا ناتج عن وجود إشارات صوتية غير مرغوب فيها (متحدثين آخرين ، ضجيج ، إغلاق باب...) ، أو عن تنوع الميكرفونات المستخدمة و مكان وضعها.

4.4. تغيير طريقة الكلام:

نطق الكلام نادرا ما يكون مضبوطا ، و يختلف نطق نفس العبارة من شخص إلى آخر، و بحسب معدل كلامه (سريع أو بطيء) ، وبحسب جودة صوته (يقصد الإفهام أو يتكلم على نحو عارض).

5.4. اختلاف نطق الصوت الواحد:

يمكن أن يختلف نطق الصوت الواحد تبعا لما إذا كان ينطق منفردا أو مع كلمات أخرى ، عندما لا يكون هناك حدود واضحة في الإشارة الصوتية بين الكلمات المتتالية ، و يمكن أن تكون هناك فترات صمت في منتصف الكلمة ، أو غياب أي توقف بين الكلمات ا لمتتالية.

[1] المرجع السابق ، التعرف الألي على الكلام العربي المتصل ، جامعة الجزائر ، 2016م ، ص39.

5. أنواع أنظمة التعرف الآلي على الكلام [1]:

توجد عدة أنواع لأنظمة التعرف على الكلام و قبل التطرق إليها ، لابد من ذكر المعايير التي على أساسها يتم تقسيم هذه الأنظمة، وبيئاتها التي سوف تعمل فيها اعتمادا على طريقة النطق حيث أنه يمكن أن يكون الكلام متصلاً أو منفصلاً إضافة إلى مدى قدرة النظام على تحديد بداية و نهاية النطق للكلام ، لذلك يمكن تقسيم نظم تمييز الكلام إلى ثلاثة أقسام على حسب الطريقة المستخدمة في التمييز و الاستخدام و هي:

1.5. أنظمة التعرف على الكلمات المعزولة:

أنظمة التعرف على الكلام المنفصل فهي تتعرف على الكلمات المنطوقة بصورة منفصلة ، و هنا يجب على المستخدم أن يتوقف بين كلمة وأخرى ، و بهذا يستقبل النظام كلمة واحدة و يحللها في وقت واحد ، أي ما يعني أن النظام يفرق بين الكلمات بواسطة سكوت المستخدم بين كل كلمة و كلمة، ولهذا النظام حالتين إما استماع أو صمت، ويكون كل نطق هادئ ونعني بهذا قلة الإشارة السمعية على كلا جانبي النافذة و تعني النافذة هي ناتج تقسيم الإشارة إلى كلمات [2] ، و يعتبر هذا النظام من أسهل الأنظمة في التمييز على الكلام لعدة أسباب:

- إناطة مهمة أي تفصيل الكلمات إلى المستخدم و ليس إلى النظام.
- عند نطق الكلمات كلمة كلمة تكون طريقة النطق أوضح و أدق للنظام.
- لا نواجه مشكلة (co-articulation) و هي إلتقاء الحرف في نهاية الكلمة الأولى مع الحرف في بداية الكلمة الثانية مما يسبب صعوبة في التعرف.

---

[1] ميس عبد القادر الكزبري ، تعرف الكلمات المعزولة باللغة العربية ، مذكرة لنيل شهادة الماجستير في المعلوماتية ، جامعة دمشق ، سوريا ، 2009 م.

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

2.5. أنظمة تمييز 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 فهي تفترض وجود وقفات زمنية واضحة تفصل بين الكلمات لكن هنا يتم تقليل فتره الصمت بين الكلمات بحيث تبدو كأنها متصلة أو كأنها جملة متقطعة بحيث تسمح للمستخدم بأن يتكلم بعدد من الكلمات في وقت واحد ، و في هذا النوع من التمييز تكون عملية التقطيع سهلة فهو يشبه النوع السابق لكنه أصعب في التمييز، و لكن يعاب على هذا النوع عدم مشابهة الواقع الإنساني الذي هو الغرض من هذه التقنية [1].

3.5. أنظمة تمييز الكلام المستمر (المتصل):

أنظمة التعرف على الكلام المستمر أكثر صعوبة من أنظمة التعرف على الكلام المنفصل [2] ، وفي هذا النظام يتم إدخال جمل كاملة بنطق طبيعي بدون أية شروط على طريقة النطق للمستخدم ، و هذا النظام من أصعب الأنظمة تطبيقا في تمييز الكلام لقلة الدقة في نطق الكلمات عندما تكون في داخل الجملة و ما يطرأ عليها من تشويه نتيجة لانسياب النطق و تدفقه ، و صعوبة تحديد حدود كل كلمة على حدة فهو يدرب على تمييز وحدات الكلام مثل الفونيمات و المقاطع و أن التمييز في الأنظمة المستمرة يتطلب تسلسلاً سمعيا ليتم تقطيعه إلى وحدات صوتية أصغر من الكلمة إلى مقاطع أو إلى أنصاف المقاطع أو إلى فونيمات و بعدها يتم تمييزها، [3] وهذه النوعية من التعرف تستخدم الفونيمات وحدة أساسية للتعرف، وتحتاج فضلاً عن الخصائص المميزة للفونيمات إلى معرفةٍ بإشارات اللغة و قواعدها ، و هذه المعرفة ضرورية لحل بعض المشاكل ، وفي هذا النوع من التعرف يكون حجم المصادر كبيراً وهذا ما يعتمد بصورة مباشرة على اللغة المستخدمة.

---

[1] المرجع السابق ، التعرف الآلي على الكلام العربي المتصل ، ص 41.

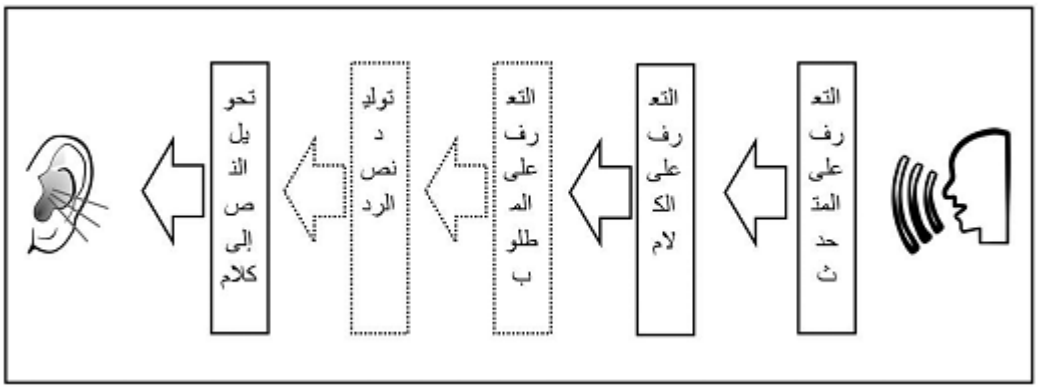
[2] فاتن بشير عبد الأحد ، تقطيع الكلمة العربية إلى أحرف وتمييزها ، مجلة الرافدين لعلوم الحاسبات والرياضيات المجلد 2005

[3] Coust, A.P., (2002), "Connected Word Recognition", http://ka.aidepot.com/paper/nodeq.htm1.

6. تطبيقات وتقنيات ذات علاقة بالأصوات اللغة العربية

هناك ثلاث تقنيات من أكثر التقنيات ذات العلاقة بالكلام التي تشكل تحديا كبيرا أمام الباحثين والمطورين وذات أهمية كبيرة للمستخدمين بجميع شرائحهم و هي : توليد الكلام آليا synthesis speech أو speech text –to ، التعرف الآلي على الكلام (automatic speech recognition) أو text-to-speech ، التعرف على المتحدث(speaker recognition) أو(speaker identification) هذه التقنيات الثلاثة مهمة في التعامل مع الآلة ، هذه التقنيات الثلاثة مترابطة إذ تشكل أساس التخاطب مع الآلة ، فالتعرف على المتحدث هو البصمة أو المفتاح الذي يوصل المتحدث إلى بياناته ويسمح له بتنفيذ أعماله مستخدما صوته ، والتعرف على الكلام وسيلة إيصال الأوامر والطلبات وإدخال المعلومات للآلة ، وتوليد الكلام ووسيلة استجابة الآلة للإنسان صوتيا. فالشكل يمثل تخاطب مع الآلة ؛ فهو مكون من خمسة نظم حاسوبية ، إلا أنه قد لا تتوفر جميع هذه النظم فيكتفى بواحد منها.

فعلى سبيل المثال فتح الباب بالبصمة الصوتية لا يتطلب إلا نظاماً واحدا للتعرف على المتحدث وكذلك استخدام الكفيف لنظام قارئ الكتاب الآلي لا يتطلب سوى نظام الناطق الآلي.



الشكل 1.2:أنموذج متكامل للتخاطب مع الآلة [1]

[1] المرجع السابق ، مدخل إلى اللسانيات الحاسوبية ، ط1، 2018م ، الرياض، ص 28 .

كما أن تحويل الموجة الصوتية إلى كهرباء كانت قفزة في تاريخ التقنيات فإن تحويل الموجات الصوتية إلى حروف وكلمات يعتبر قفزة في تقنية البرمجيات ، لأنه لا يمكن التعامل مع موجة أصوات الكلام من الناحية اللغوية ولكن يمكن التعامل مع النص بطرق عدة منها : فالترجمة الآلية لا تتم مباشرة من موجة صوتية بلغة ما ، وكذلك العمليات المعقدة الأخرى ذات العلاقة البالغة كفهم اللغة و التنقيب في النصوص وتلخيصها و فهرستها وتصنيفها و تقويمها و توليدها آليا ، كل هذه العمليات وغيرها الكثير منها تتطلب توفر نصوص لغوية ، لهذا فإن التعرف الآلي على الكلام في غاية الأهمية ليس فقط التطبيقات المباشرة ولكن لأنه يمثل بنية أساسية لعمليات أخرى أكثر تعقيدا .

تحويل الموجة الصوتية إلى كهرباء تتطلب مهندسا كهربائيا ليقوم بالمهمة ، أما تحويل الموجة الصوتية إلى نص فيتطلب متخصصين من علوم شتى (مبرمجين , مطورين) إضافة إلى قواعد بيانات و أدوات حاسوبية ، يضاف إلى ذلك التعقيدات المصاحبة لموجات أصوات الكلام ، فكل مرة ينطق شخص كلمة محددة تكون لها موجة صوتية مختلفة مهما تكرر نطقها .

1.6. التعرف الآلي على الكلام:

التعرف الآلي على الكلام هو تطوير نظام حاسوبي يتولى تحويل الموجة الصوتية إلى نص ، وليتعرف الحاسب على هذه الموجة عليه:

أولا : معالجة الإشارة الصوتية ليتمكن من التعامل معها .

ثانيا : تحديد الفواصل (الحدود بين كل صوت و آخر على الموجة الصوتية) .

ثالثا : استخلاص السمات الأكوستية للأصوات بين الفواصل .

رابعا : مقارنتها بالخصائص الصوتية المخزنة .

خامسا: اختيار المقارب لكل صوت حسب أعلى احتمالية ممكنة ليصل إلى نتيجة .

عندما تكون مهمة نظام التعرف الآلي على الكلام قد انتهت، إلا أن كل مرحلة من هذه المراحل تشكل تحديا كبيرا للمتخصصين من مبرمجين وباحثين ومطورين.

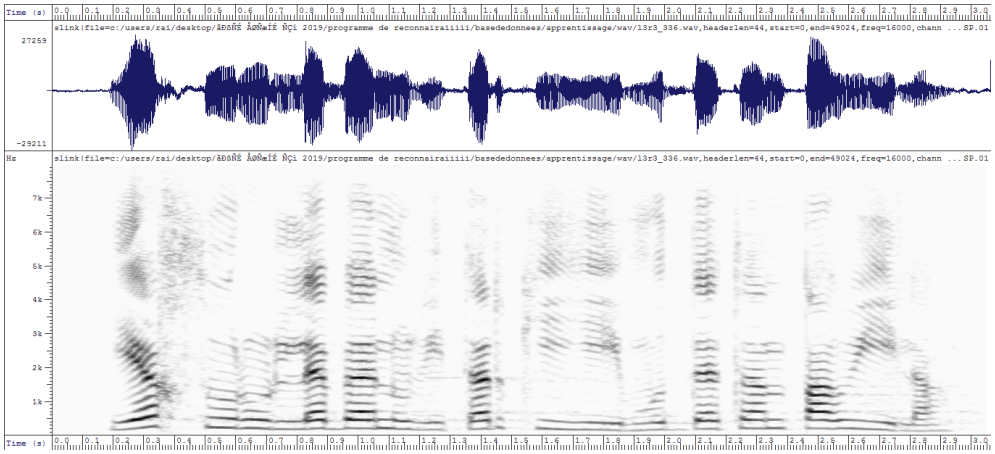
ولتنفيذ المراحل السابقة فإن نظام التعرف على الكلام المطلوب تطويره يحتاج إلى :

1- قاعدة بيانات صوتية لكلام عدد من المتحدثين تكون موجاته الصوتية مقطعة (segmented) و مرمزة transcribed ، وهناك عدد من قواعد البيانات للغة العربية منها LDC و (Westpoint arabic) وبنك الأصوات الهاتفية للمتحدثين [1] .

2- استخدام أدوات حاسوبية كأدوات أنموذج ماركوف المخفي (Hidden Markov Toolkits HTK) تقوم ببناء أنموذج أكوستي (Acoustic mode)من قاعدة البيانات بإستخلاص السمات الأكوستية لكل صوت aound features ، وتقوم لاحقا بمقارنة ما يصلها من كلام بما هو مخزن لديها في الأنموذج الأكوستي للوصول إلى النص المقابل للموجة الصوتية .

3- معالجة النص المستخلص ليكون نصا لغويا سليما .

[1] Alotaibi ,yousef A ,Sid –Ahmed Selouani M Alghhamdi , Ali H Meftah (2012) Arabic And English Speech Recognition Using Cross – Language Acoustic Models .11th International Conference On Information Science ,Science ,Signal Processing And their Application (ISSPA)



الشكل 2.2: موجة صوتية باستعمال برنامج SFS Speech Filing System [1]

يستخدم الباحثون والمطورون عددا من الآليات و الأدوات للرفع من كفاءة نظم التعرف التي تقاس باستخدام ملفات صوتية مخصصة لاختبار النظام ، فتحسب نسبة عدد هذه الفونيمات التي تعرف عليها النظام إلى إجمالي عدد الفونيمات في ملفات الاختبار.

يواجه مطورو أنظمة التعرف الآلي على الكلام عددا من التحديات من أهمها التباين في أصوات المتحدثين واختلاف صوت المتحدث الواحد من وقت لأخر[2] .

كما أن الخلفية الصوتية للمتحدثين تتسبب في إرباك النظام لعدم قدرته على فصل الموجات الصوتية الصادرة من المتحدث من الأصوات المنتشرة في الخلفية ،كما يصعب على هذه الأنظمة معرفة ما وراء الكلام كحالة المتحدث النفسية والعاطفية وما يريد.

[1] المرجع السابق ، التعرف الألي على الكلام العربي المتصل ، ص80.

[2] Nahar O. M Khalid, Mohammed., Abu G ShquierWasfi -Al KhatibHusni Al Muhtaseb Arabic (2016) Moustafa Elshafei - phonemes recognition using hybrid HMM/LVQ model for continuous . Speech recognition International of Journal Technology.

2.6. توليد الكلام آليا [1]:

تمهيد هذه الأنظمة إلى تحويل النص المكتوب إلى موجات كلام مسموعة ، وغالبا ما تقوم هذه الموجات بناء على: الموضح (intelligibility/comprehensibility) وهو مدى قدرة الإنسان على فهم كلام النظام بمعني تحويل الموجات الصوتية إلى النص الذي نطقه النظام في الأصل[2]، الطبيعة (naturantes) و هي مدى قربها من الكلام الطبيعى الذي يولده جهاز صوت الإنسان ، هذا التقويم يضعه مطورو هذه الأنظمة نصب أعينهم عند العمل على تطوير نظام لتوليد الكلام آليا ،كما أن المستخدمين يأخذون هذين المعيارين في الحسبان عند الشراء أو الرغبة في استخدام نظام من هذا النوع .

هناك عدد من الطرق لتوليد الكلام آليا للعمل ومن أكثرها شيوعا : توليد الموجات المجزأة concatenated speech synthesis, وتعتمد هذه الطريقة على استخلاص وحدة صوتية (speech unit) لكل صوت يمثل فونيما في اللغة من كلام طبيعي ، وعند توليد الكلام تستدعي الوحدات الصوتية التي تقابل النص المطلوب نطقه ، وتدمج مع بعضها لتمثل صوتا متصلا ، التوليد الآلي للكلام باستخدام النطق الرنيني Formant based synthesis , وهذه الطريقة لاتحتاج إلى قواعد بيانات صوتية ، وإنما تولد النطق الرنيني الذي يقابل كل فونيم ، وتدمج مع بعضها لتولد الموجة الصوتية .

التوليد الآلي للكلام بناءً على قواعد بيانات صوتية لأحد المتحدثين Data driven synthesis يستخدم المطورون لهذه الطريقة أدوات حاسوبية كأنموذج ماركوف المخفي[3]، والشبكات العصبية ، وقد شاع

---

[1] المرجع السابق ، مدخل إلى اللسانيات الحاسوبية ، ط1، 2018م ، الرياض، ص 30 .

[2] Chang y-y (2011) Evaluation of TTS Systems in intelligibilité and compréhension tasks .in proceedings of the 23rd conférence on computationnel linguistics and speech processing

[3] Wu , Zhizheng , Oliver Watts ,Simon King (2016)Nmerlin : An Open Source Neural Network Speech Synthesis System 9th ISCA Speech Synthesis Workshop

إنتشار الطريقة الأخيرة نظرا للنتائج التي تعطيها من حيث قربها من الصوت الطبيعي ، وتعتمد هذه الطريقة على قواعد بيانات صوتية سجلت ورمزت بعناية لأحد المتحدثين .

3.6.التعرف على المتحدث آليا:

تحمل الموجات الصوتية الصادرة عن الجهاز الصوتي خاصة بالأصوات اللغوية (الفونيمات) وتحمل أيضا معلومات أخرى غير لغوية منها حالة المتحدث النفسية [1]، المزاجية (سعيد ، حزين ، غضبان) ، فالمستمع قادر على التمييز بين المتحدثين ، فيتعرف على المتحدث المعتاد على سماع صوته إذا كان ضمن متحدثين آخرين .

7. مراحل عملية تمييز الكلام:

يمكن تبسيط مراحل مخطط تمييز الكلام في الإشارة الصوتية والتي يجب أن تمر بعدة مراحل وهي :

1.7. مرحلة تسجيل و قراءة الإشارة الصوتية:

وهي المرحلة المسؤولة عن تسجيل صوت الكلمة باستعمال تقنيات التسجيل المتوفرة ، لتحضير الإشارة الصوتية لمجموعة من المعالجات التي سيتم عليها فيما بعد[2]، حيث تتم عملية أخذ سلسلة من العينات تمثل مجال الإشارة الكلامية ضمن بطاقة تحصيل الصوت ، لذا علينا الانتباه إلى خصائص تسجيل هذه الإشارة و التي تتمثل في ثلاث محددات هي: تردد التقطيع ، عدد بتات الترميز، إضافة للتسجيل بأحد الخيارين Mono أو Stereo ، ويمكن إنجازها بعدة الطرق بحيث يمكن إيجاد نقاط بداية النطق بمقارنة المستويات الصوتية بالعينة المسجلة، أما نهاية النطق فهو أكثر صعوبة لأن المتكلم يميل إلى ترك ما ليس له معنى مثل إصطكاك الأسنان و التنفس والتنهد.

---

[1] الغامدي منصور بن محمد ، البصمة الصوتية ، المجلة العربية للدراسات الأمنية و التدريب ، ص 118.

[2] بركات ابراهيم أحمد ، التمييز الكلام العربي باستخدام الشبكات العصبونية الاصطناعية ، جامعة السودان للعلوم و التكنولوجيا ، 2006م، ص23

2.7. مرحلة التقسيم إلى إطارات و استخلاص المعاملات:

أي استخراج سمات الإشارة ، حيث تتضمن استخلاص الصفات للحروف بالاعتماد على طريقة ترميز التنبؤ الخطي وبناء قاعدة البيانات للنظام والتي تشمل متجهات الصفات للأحرف المستقطعة من الكلمات MFCC مع الأخذ بعين الاعتبار ورود الحرف في مواضع مختلفة من الكلمة [1].

3.7.تقنيات المقارنة و الربط للتعرف على الكلام :

تعتمد هذه الطريقة على اختيار الوحدة الصوتية المناسبة للتعرف (صوتيم ، أو ثنائية صوتية ، أو كلمة ، ، أو جملة) من قاعدة البيانات ، وهناك عدد ضخم من التقنيات المستخدمة وهذه المرحلة يتم فيها بناء النموذج الاحتمالي للتعرف على الفونيمات و تشتمل التمييز على الحروف العربية حسب كيفية نطقها [2]

4.7. التعرف على الكلام المعتمدة على المدونات :

وهي طريقة مقيدة بالمعطيات (data driven approach) ، تعتمد على وجود قواعد معطيات كلامية هائلة وجيدة ، تتضمن وحدات صوتية لغويا (phoneitic) [3] ، إن أي نظام لتمييز الكلام لابد و أن يمر بعدة مراحل ، وتعتبر مرحلة المقارنة والربط هي المرحلة الرئيسية التي يتحقق فيها معنى التعرف على النطق المراد لآلة التعرف عليه ، حيث يستخدم أي نظام لتمييز للكلام العربي بالاعتماد على تقنية ما ، و قد قمنا بتقسيم تلك التقنيات إلى فئات و هي:

---

[2] حليم ، علياء موفق عبد المجيد" تشفير إشارة الكلام بطريقة البعثرة بحث ماجستير ، علوم حاسبات ، كلية علوم الحاسبات والرياضيات ، جامعة الموصل 2003 م ، ص8.

[2] شهلة عبد الوهاب عبد القادر ،" تمييز أصوات الحروف العربية ماجستير ، علوم حاسبات ، كلية علوم الحاسبات والرياضيات ، جامعة الموصل ، 2002م ، ص 25.

[3] أميمة دكاك ، نحو النظام لتركيب الكلام باللغة العربية من نصوص ، المعهد العالي للعلوم التطبيقية و التكنولوجيا باستعمال الضم لأنصاف مقاطع صوتية وتنغيم الطبيعي ، جامعة سوريا ، 2018م ، ص 58.

1.4.7. تقنيات إحصائية:

حققت نماذج ماركوف المخفية رواجا في الفترة السابقة ، نتيجة للنجاح الكبير الذي حققته في التعرف على الكلام، ويرجع ذلك النجاح إلى الأساس الاحتمالي الذي بنيت عليه ، فالتعامل مع الكلام تعتريه مشاكل النقص و عدم المصداقية ، و تعمل هذه التقنية عن طريق دالة احتمالية لسلسلة ماركوف المخفية Hidden Markov HMM Model، التي تتكون من عدد محدد من الحالات و مجموعة من الدوال العشوائية المرتبطة بكل حالة ، تتغير الحالات نسبة لمصفوفة احتمالات انتقالية ، حتى الوصول للتوقع النهائي للنطق.

2.4.7. تقنيات المقارنة:

التقنية الأكثر شيوعا في هذا المجال هي تقنية تطويع الزمن الديناميكي (DTW) ، وتعمل هذه التقنية على مبدأ معالجة التغيرات في محور الزمن المسببة باختلاف معدل النطق ، لأن الكلمات التي تمثل بقيم صوتية متشابهة تقارن على أساس أنها متساوية الطول و أزمنتها متطابقة ، و تستعمل (DTW) الانحراف المتري لمقارنة متجهات الخصائص في كل من النطق و النطق المقارن و تحذف الفروق بين العينتين بتطويع محور الزمن بالنسبة للعينة التي تحقق أكبر توافق مع الأخرى [1].

3.4.7. تقنيات الربط:

أثناء السنوات الأخيرة ، ازدهر اتجاه حاسوبي جديد مستند على العلوم العصبية سموه الشبكات العصبية قياسا على العمل الداخلي للإنسان ، إن الفكرة الرئيسية وراء الشبكات العصبية هو محاكاة سلوك الربط العصبي في الدماغ ، و هي ممتازة في التعرف على الكلام لقدرتها على التكيف مع البيانات الناقصة أو المبهمة ، فالكلام قلّما يسلم من هذه العيوب ، كما تتميز بالسرعة بالمقارنة مع التقنيات الأخرى.

[1] المرجع السابق ، التعرف الآلي على الكلام العربي المتصل ، ص33.

8. تسجيل الأصوات:

الكلام ذو طبيعة معقدة للغاية ، لذا يتطلب تمثيلا و علاجا محددا من أجل تخفيف التعقيد من جهة ، وليتناسب مع الآلات من جهة أخرى ، لهذا نختار مع هذا الترميز ثلاث محددات:

1- تمتد حدود السمع ضمن المجال [20 – 20000Hz] ويتراوح أعلى مجال حساسية الأذن من –80] [8000Hz، فإذا ما قارنا مع طيف الإشارة الهاتفية نجد أن طيفها يمتد في المجال [300–3400Hz] و مع هذا تعطي صوتا مميزا للشخص المتكلم بكلمات مفهومة ، لذا يمكن اعتبار أعلى تردد للإشارة الكلامية 4000Hz.

و بتطبيق قانون شانون في تحصيل الإشارة الكلامية الذي يشترط أن يكون تردد التقطيع 2Fmax= Fs حيث Fmax التردد الأعظم للإشارة الكلامية ، يكفي أن نأخذ تردد التقطيع Fs=8000Hz ، ليعطي إشارة كلام مقبولة جدا و يجب الانتباه عند زيادة هذا التردد لترتفع نوعية الإشارة المسجلة و لكنها تؤدي لزيادة عدد العينات المأخوذة في الثانية مسببة زيادة في زمن المعالجة الرقمية.

2- عدد بتات الترميز ويقصد بها عدد البتات اللازمة لترميز قيمة مجال العينة أو عدد المستويات المأخوذة من الإشارة، و يمكن اختيار طول العينة بإحدى القيم الثلاث (8 bit,16 bit, 32 bit) ، مع العلم بأن زيادة هذه القيمة لا يؤثر على جودة الإشارة الكلامية بل سيعقد عمليات المعالجة مؤديا إلى زيادة زمنها.

3- التسجيل باستخدام قناة أو قناتين أي التسجيل بأحد الخيارين (Mono) و (Stereo) مما يؤثر على حجم ملف التسجيل ، كون الخيار (Stereo) يعطي ضعف حجم الملف الصوتي ، الذي تم تسجيله باستخدام الخيار (Mono).

9. استخلاص المعاملات:

تتمثل في تحليل الإشارة لاستخلاص المعاملات الصوتية بدقة حيث يتم معالجة الإشارة بتقسيمها إلى إطارات متساوية في الطول ثم يتم استخلاص المعاملات لكل إطار لتكون النتيجة النهائية متجهة من المعاملات الذي يمثل ذلك الإطار،[1] وهناك عدة طرق تستخدم لاستخلاص المعاملات لأية إشارة صوتية فيمكن استخدام الطاقة (Energy) و معدل قطع الصفر (Zero crossing) و معاملات (Mel – Frequency cepstrum coefficients (MFCC)) و معاملات التنبؤ الخطي ((Linear predictive coefficients (LPC)) و غيرها .

10. تحويل فورييه : Fourier Transform

إن تحويل فورييه يقدم أسلوبا لحساب التردّدات المكونة لإشارة معينة حيث تم إثبات أن أي إشارة أو تابع رياضي يمكن أن يكتب كمجموع توابع جيبية وكل من هذه التوابع له تكرار معين يسمى التّردد[2].

إن تحويل فورييه يقيس التشابه بين الإشارة أو التابع قيد الدراسة، وبين مجموعة التوابع الجيبية ذات التّرددات المختلفة كما توضح المعادلة التالية :

$$F(\omega) = \sum_{n=-\infty}^{+\infty} f(n) e^{-j\omega n} \dots\dots\dots\dots\dots\dots\dots\dots \quad (2.1)$$

حيث : F(ω) : هو تحويل فورييه للإشارة

f(n) : عينات الإشارة في اللحظات n

ω : تردد الإشارة الذي نرغب بحساب توافقها مع هذه الإشارة.

---

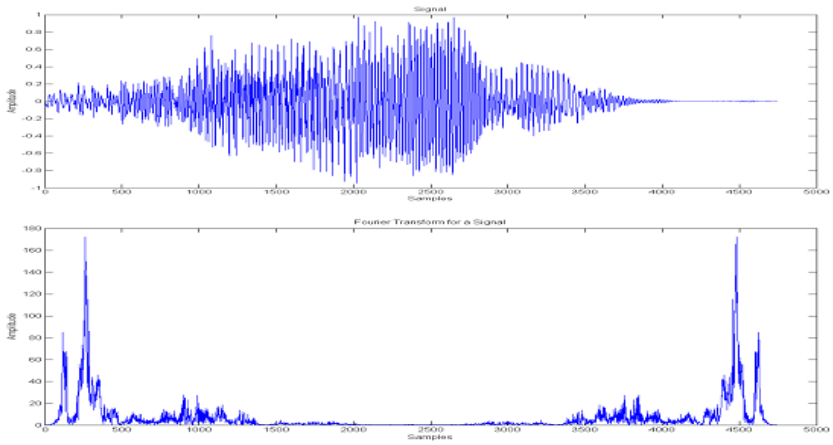
[1] يسرى فيصل الارحيم ، عمي انصاف جاسم ، تمييز الأرقام العربية المفردة النطق باستخدام الخوارزمية الجينية ، مجلة الرافدين لعلوم الحاسبات و الرياضيات المجلد (11) العدد (1) 2014 م ص 57

[2] ذو الفقار إبراهيم مرهم ، غدير أحمد حامد ، جامعة تشرين ، كلية الهندسة المعلوماتية قسم البرمجيات ونظم المعلومات ، التعرف على الكلمات المعزولة ، 2015م ، ص23

نلاحظ من المعادلة السابقة أن ناتج تحويل فورييه هو عبارة عن أعداد عقدية تمثل نسبة التوافق بين هذه الإشارة والتابع الجيبي الموافق للتّردد n ، حيث يمثل مجال هذه الأعداد العقدية مقدار التوافق بين الإشارتين ، وتسمى هذه المجالات طيف الإشارة التّرددي [1] .

نلاحظ أننا نقلنا الإشارة من مجال الزمن إلى مجال التّردد، حيث أن مجال الزمن هو تغير الإشارة بتقدم الزمن وكل لحظة زمنية تمثل مجال معين ، أما مجال التّردد فهو تغير استجابة الإشارة بزيادة التّردد ، حيث في الإشارات ذات التّردد المنخفض تكون القيم المرتفعة للتّرددات منخفضة المجالات في طيف الإشارة والعكس بالعكس.

إن أهم فوائد تحويل فورييه أنه يساعد في معرفة ترددات الإشارة حيث أنه تكون قيم مجالات الطيف التّرددي كبيرة عند التّرددات الأكثر فعالية في هذه الإشارة وضعيفة عند باقي التّرددات وبالتالي يمكن تحديد تردد الإشارة والتّرددات الغير مرغوبة والتي تعتبر ضجيج على الإشارة[2].



الشكل 3.2: يمثل هذا الشكل إشارة معينة و الطيف الترددي الخاص بها

إن تحويل فورييه متناظرة أي أن القيم بعد منتصف الطيف تماثل القيم قبل منتصف الطيف كما يبيّن الشكل السابق، لذلك يتم الاحتفاظ فقط بالنصف الأول من التحويل.

[1] المرجع نفسه ، ص44

[2] المرجع السابق ، ص 25

كما نلاحظ من الشكل أن أغلب الاستجابات تحدث في مجال التّرددات المنخفضة أي أن تردد الإشارة أكبر استجابة وهي بتردد 250 Hz تقريبا لذلك نجد أن لتحويل فورييه فائدة في تحديد التّرددات ذات الأثر الأكبر في الإشارة.

في تحويل فورييه نجد أن طيف الإشارة هو عبارة عن جداء طيف الإشارة الصادرة عن المنبع وطيف استجابة الوسط ، أي نستطيع القول أنه بتحويل الإشارة إلى مجال التّردد نحو مجموع الالتفاف إلى جداء أطياف الإشارات المشكلة لهذه الإشارة كما تبين المعادلتين التاليتين:

$$s(t) = \sum_{n=-\infty}^{+\infty} e(n)h(t-n) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.2)$$

تشكيل الإشارة في مجال الزمن حيث هي h استجابة الوسط ، e هي الإشارة الصادرة عن المنبع[1].

$$\boldsymbol{s(f) = E(f).H(f)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad \boldsymbol{(2.3)}$$

حيث هو طيف الإشارة المقابلة لاستجابة الوسط ، طيف الإشارة الصادرة من المنبع.

11. التقطيع إلى إطارات و النوافذ : Framing and Windowing

تقطيع الإشارة هو عملية تقسيم الإشارة إلى عدد من الأقسام يسمى كل منها إطار ويعالج كل إطار بشكل مستقل عن بقية الإطارات.

في أغلب الأحيان لا نستطيع التعامل مع الإشارة دفعة واحدة لأن ذلك قد يسبب نتائج غير مرضية ، حيث كما وجدنا في الفقرة السابقة أن تحويل فورييه لا يحدد اللحظات الزمنية الموافقة لتردد معين لذلك كان التقطيع هو الحل الأفضل لحساب التّرددات بشكل شبه محلي أي نكون معرفة جزئية عن اللحظات الزمنية الموافقة للتّرددات.

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

أما تطبيق النوافذ أو النوفذة 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:

$$s(n) = \sum_{i=n-m+1}^{n} x(i) * w(n-i) \dots\dots\dots\dots\dots\dots\dots \quad (2.4)$$

12 .أنواع النوافذ :

توجد أنواع عديدة جدا من النوافذ وكل هذه النوافذ (ما عدا المستطيلة) تميل لأن تكون منخفضة المجالات عند طرفي النافذة ومرتفعة في منتصفها وذلك لإلغاء تأثير عينات الإشارة من الطرفين وحصر تطبيق النافذة بمجموعة من العينات في وسط النافذة[1] .

سنستخدم في بحثنا هذا ثلاثة أنواع من النوافذ:

1.12. النافذة المستطيلة : (**Rectangular Window**) وتعتبر أبسط النوافذ ويعبر عنها بالمعادلة التالية:

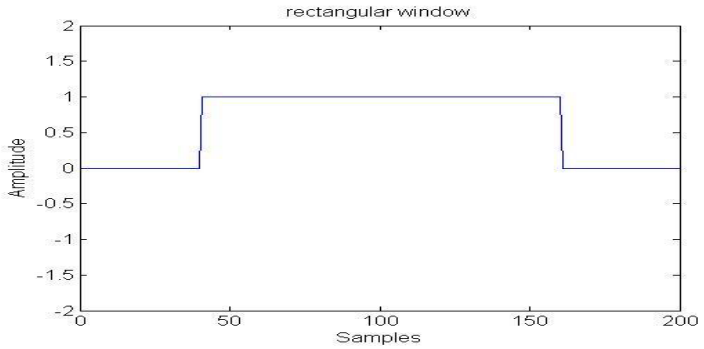
$$w(n) = \begin{cases} 1. & |n| \le \frac{N}{2} \\ 0. & n > N/2 \end{cases} \dots\dots\dots\dots\dots\dots\dots\dots \quad (2.5)$$

حيث أن N هو طول النافذة .

إن تطبيق هذه النافذة على إشارة معينة يكافئ تقطيع هذه الإشارة إلى إطارات كل منها بطول

وهذه النافذة تمثل بالشكل:

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



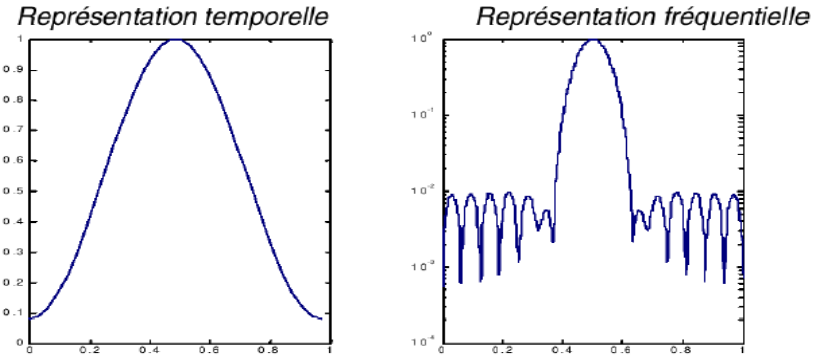
الشكل 4.2: يمثل النافذة المستطيلة

2.12. نافذة هامينغ (Hamming Window)

هي أشهر أنواع النوافذ ولها استخدامات كثيرة في مجال معالجة الإشارة ، حيث أن هذه النافذة تعطى بالمعادلة [1]:

$$w(n) = 0.54 - 0.46 * \cos(2 * pi * \frac{n}{N-1}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.6)$$

حيث n طول النافذة ، وشكل هذه النافذة [2]



الشكل 5.2: يمثل نافذة هامينغ[3]

[1] ميس عبد القادر الكزبري، تعرف الكلمات المعزولة باللغة العربية، ماجستير في المعلوماتية، جامعة دمشق، سوريا، 2009 م

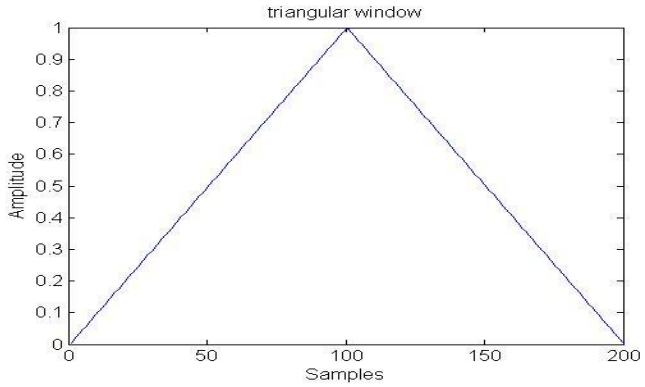
[2] المرجع نفسه ، ص23.

[3] المرجع السابق ، التعرف على الكلمات المعزولة ، ص29.

3.12. النافذة المثلثية :(**Triangular Window**) تعتبر هذه النافذة هي مجموع التفاف بين نافذتين مستطيلتين ويعبر عنها بالمعادلة:

$$W(n) = 1 - \left| \frac{n - \frac{N-1}{2}}{\frac{N}{2}} \right| \dots\dots\dots\dots\dots\dots\dots\dots \quad (2.7)$$

حيث N طول النافذة ، وشكل هذه النافذة:



الشكل 6.2: يمثل النافذة المثلثية

13. تسجيل ومعالجة الإشارة (Recording & Signal Processing)

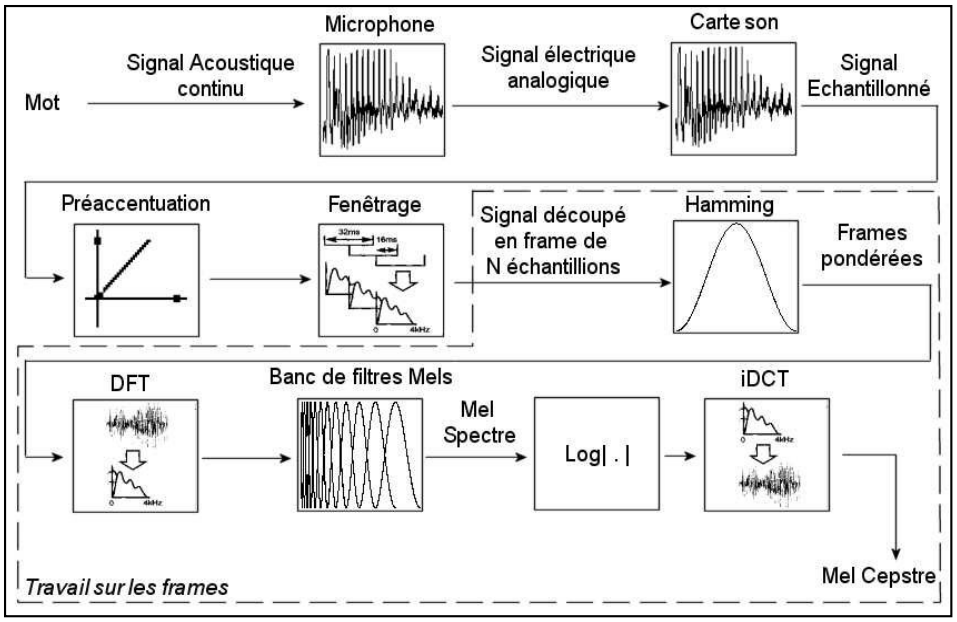
في البداية نقوم بتسجيل الإشارة الصوتية و التي ستمثل عينات اختبار حيث هذه الإشارة على التردد11025HZ ، حيث يتم في هذه المرحلة تطبيق مجموعة من التقنيات الرياضية على الإشارات لتحسين جودتها و التخفيف من أخطاء التسجيل و التي تنشأ لعدة عوامل مؤثرة من البيئة المحيطة .

14. تحليل طيف الإشارة (Spectrum Analysis)

إن تحليل طيف الإشارة هو تطبيق لتحويل فورييه (fourier transform) على قطاعات جزئية من الإشارة تسمى إطارات ، حيث يعتبر كل إطار مستقلة بشكل طيفي عن بقية الإطارات ، و الهدف هو معرفة الترددات الموجود في إطار زمني وتسمى هذه العملية بتحويل فورييه قصير الحد (STFT) Short-Time Fourier Transform ويكون الناتج هو رسم ثنائي يمثل إطارات زمنية STFT.

15. تقنية التحليل الإشارة الصوتية (MFCC) Mel Frequency Cepstral Cofficients

و يأتي شرح مفصل للمراحل التي يمر بها الصوت لاستخلاص المعاملات منه ، الشكل (الشكل 5.2) يوضح المخطط الكتلي لمعالجة (MFCC) و تشمل المراحل الآتية :



الشكل 7.2: مخطط عام يوضح مراحل حساب معاملات MFCC[1]

تعتبر هذه التقنية من أكثر التقنيات في استخراج السمات المتميزة للصوت في التمييز على الكلام وذلك بسبب دقة نتائجها والقدرة على التخلص الجزئي من ضجيج الإشارة، كما تعتمد بالأساس على ترشيح الترددات الخطية [2]، و إن محتويات الترددات لأصوات الإشارات لا تتبع مقياسا خطيا ، لهذا فكل نغمة في الترددات الطبيعية (f) تقاس بالهرتز (Hz)، أما طبقة الصوت فتقاس بمقياس Mel الذي يعطينا ترددات خطية بالمسافات التي يكون ترددها أقل من 1000Hz و يعطي ترددات لوغارتمية في المسافات الأعلى من 1000Hz لهذا يتم استخدام الصيغة التقريبية لحساب Mel لتردد معين إذا أعطينا f ب (Hz) :

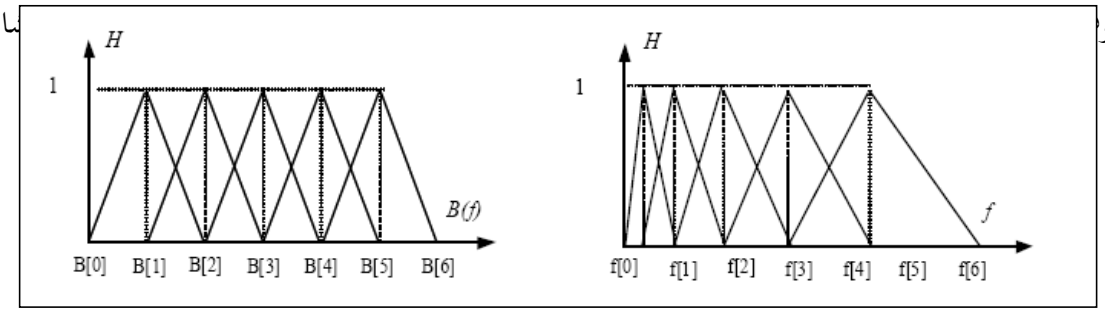
[1] المرجع السابق ، تمييز أصوات الأرقام العربية ، ص26.

[2] Brandstein M. ,2001. Speech Recognition of English Digits, Harvard University

يطلق على هذه المرشحات Mel Scale و تكون عبارة عن مجموعة من الإشارات المثلثية تمثل كل إشارة مرشحا لمجال ترددي مختلف ، ويمكن التعبير عن مراكز هذه المرشحات بالمعادلة :

$$\text{Mel}(f) = 2595 * \log10(1+f/700) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.8)$$

إن الهدف من مقياس الترددات Mel هو عمل بنك الترشيح (Filter Bank) الذي يحوي مرشحا واحدا لكل مكون من مكونات Mel المطلوبة و يتم بموجبه تجميع مرشحات تمرير الحزمة مثلثة الشكل (Triangular Bandpass Filter) يحدد عرض المرشح بوساطة ثابت التردد Mel التي في النهاية تمثل بنك الترشيح ، و يتم بواسطة هذه الطريقة تغطية كل الترددات بدون أي فقد للمعلومات بحيث أنه عند إدخال الإشارة الناتجة عن FFT إلى المرشح تظهر مواقع القوة لتلك الإشارة، و الحالة المثالية لعدد المرشحات في هذه الطريقة تساوي 20 والتي سو[illegible]ارة [1] [2].



الشكل 8.2: بنك الترشيح Mel

(1) يسرى فيصل الارحيم ، عمي انصاف جاسم ، تمييز الأرقام العربية المفردة النطق باستخدام الخوارزمية الجينية ، مجلة الرافدين لعلوم الحاسبات و الرياضيات المجلد (11) العدد (1) 2014 م.

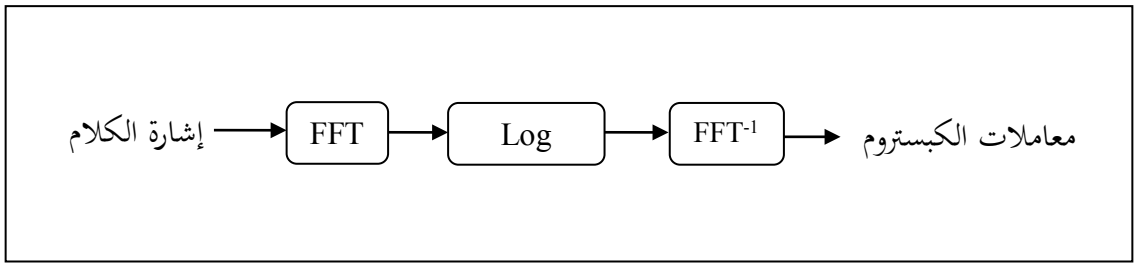
(2) T. Macro et T. Frank, «Implémentation d'une calculatrice vocale », EPITA, Université Pierre et Marie Curie, 2004 ، p32.

1.15. التحليل باستخدام الكبستروم (Cepstrum)

الاسم (Cepstrum) هو كلمة بدايتها (الحروف الأربعة) معكوس بداية كلمة الطيف (Spectrum) ، و هو يعرف رياضيا بأنه تحويل فورييه العكسي للوغاريتم مجال الطيف ، و يعبر عنه بالمعادلة:

$$\text{Cepstrum} = FFT^{-1}[\log|FFT[s(n)]|] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.9)$$

و باعتبار إشارة الكلام s(n) إشارة مقطعة فإننا نستخدم تحويل فورييه المتقطع ثم نحسب لوغاريتم المطال ثم نحسب تحويل فورييه العكسي كما يوضحه الشكل التالي:



الشكل 9.2: حساب معاملات الكبستروم[1].

2.15.معاملات MFCC:

في الخطوة الأخيرة يتم تحويل نتيجة بنك الترشيح إلى المنطلق الزمني و هي MFCC و التي تجهزنا بتمثيل جيد لمواصفات الطيف لجزء من إشارة الرقم الصوتي للإطار المحلل و تستخدم المعادلة الآتية لحساب MFCC

$$MFCCi = X_k \cos[\pi i (k - 0.5) / M] \qquad i = 1, 2, \ldots, p \quad \ldots\ldots\ldots (2.10)$$

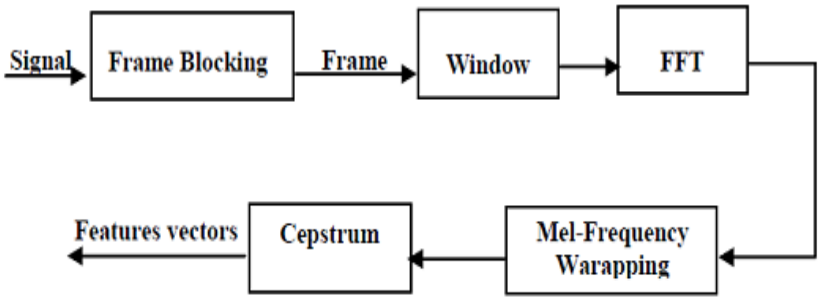
حيث أن:

M : عدد مرشحات في بنك الترشيح

X_k: هي إشارة الإطار الواحد Mel cepstrum .

P: عدد معاملات المتجه الواحد للإشارة المحللة.

(1) A. V. Aho, J. E. Hopcroft et J. D. Ullman, "The design and analysis of computer algorithms, chapitre 7 : The fast Fourier transform and its application", pp. 251-276, 1974.



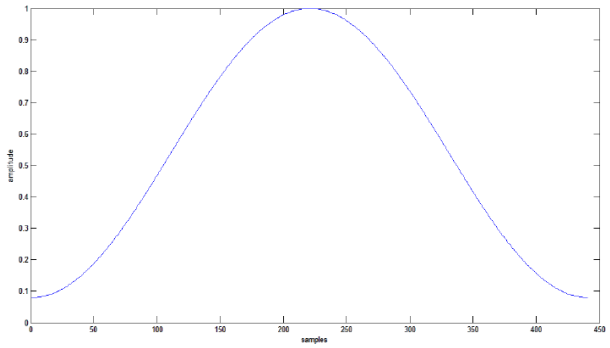
الشكل 10.2: مخطط يوضح مجال عمل تقنية MFCC لاستخراج السمات[1]

إن تقسيم الإشارة إلى إطارات (Frame Blocking) تعد أول مرحلة في استخراج السمات باستخدام MFCC وهي تقسيم الإشارة إلى إطارات ، يكون الداخل هو إشارة صوتية والخارج هو الإطارات الموافقة للإشارة ، والهدف الأساسي من عملية التقسيم هو حساب السمات الموافقة لتغير الإشارة بالنسبة للفاصل الزمني الذي يحدد طول الإطار ومن ثم تطبيق النافذة (windowing) هذه المرحلة هي الإطارات الناتجة عن التقسيم بعد تطبيق النافذة عليها ، حيث يتم تطبيق النافذة على كل إطار .

إن أفضل النوافذ و أكثرها استخداما من MFCC هي نافذة (Hemming) و التي تعطى بالمعادلة الأتية :

$$w(n) = \begin{cases} 0.54 - 0.46 * cos\left(\frac{2\pi n}{N-1}\right) & for\ n = 1,2 \dots N-1 \\ 0 & Otherwise \end{cases} \dots\dots\dots (2.11)$$

حيث أن N هي عدد عينات النافذة



الشكل 11.2: نافذة هامينغ من أجل طول إطار 25 ms

[1] Noelia Alcaraz mesegner , (2009) , ' Speech Analysis for Amtomatic speech recognition ' master of science in electronics , Norwegian university of science and technology .

تأتي مرحلة فورييه السريع Fast Fourier Transform (FFT) هذه المرحلة تتم فيها إخراج النوافذ لحساب الطيف الترددي لكل إطار والذي يعتبر داخل المرشحات الترددية .

وتطبيق المرشحات الترددية (Mel Frequency Warpng) تعتبر مرحلة البداية الفعلية لاستخراج السمات باستخدام MFCC ، كما أن هذه المرشحات يمكن بأي تابع رياضي ولكن الأفضل في MFCC أن يكون المرشح كما الشكل التالي :

$$F_{mel} = 2595.log_{10}\left(1 + \frac{f_{Hz}}{700}\right) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (2.12)$$

حيث من خلال هذا التردد يتم تحديد عرض المرشح أي مجال التّرددات التي يسمح بتمريرها و هي المرحلة التي يتم فيها فصل طاقة الإشارة عن استجابة المسار الصوتي .

ثم تأتي مرحلة تطبيق اللوغاريتم على خرج المرشحات الترددية (حساب الطاقة اللوغاريتمي) Log Energy Computation ، وذلك يؤدي إلى تسهيل الحصول على استجابة المسار الصوتي حيث أن هذه الاستجابة هي التي تمثل شكل المقطع الصوتي (phoneme) المنطوق والذي يعطى خرج هذه المرحلة بالمعادلة:

$$\left|log(s(f))\right| = \left|\log(\mathrm{E}(f).\mathrm{H}(f))\right| = \left|log(E(f))\right| + \left|\log(\mathrm{H}(f))\right| \dots \quad (2.13)$$

تطبيق التحويل الجيبي المتقطع Discrète Cosine Transform (DCT) هي مرحلة يتم فيها حساب المركبات التي تمثل كل إطار (cestral coefficennts) وهي السمات المميزة للإشارة وبتطبيق تحويل فورييه العكسي على لوغاريتم الإشارة فإن القسم الحقيقي من هذا التحويل يمثل المركبات الممثلة لهذه الإشارة وبالتالي هذا التحويل يؤدي إلى الحصول على المركبات التي تمثل الإشارة .

16. الخاتمة:

يعرض الفصل الثاني الدراسة المرجعية لتقنيات معالجة الإشارة الكلامية التي تلعب دروا هاما في التمييز الآلي للكلام العربي من حيث الطرق المستخدمة في استخراج السمات وطرق تصنيف الكلام الحالية و أشهر التقنيات المستعملة في التمييز للكلام العربي (المقارنة ، الربط ،إحصائية) بالإضافة إلى المقارنة بين هذه التقنيات وتبيان مساوئ و محاسن كل تقنية ، تتم فيها معالجة إشارة الكلام بشكل متعاقب للوصول إلى النتائج المرجوة .

حيث يتم تحويل إشارة الكلام إلى مجموعة من المعاملات تحسب في فترات زمنية منتظمة ، لتمثيل الإشارة في أفضل نموذج ، و من بين المعاملات الأكثر استخداما نجد معاملات MFCC حيث تعتبر هذه التقنية من أكثر التقنيات في استخراج السمات المتميزة للصوت في تمييز للكلام وذلك بسبب دقة نتائجها والقدرة على التخلص الجزئي من ضجيج الإشارة .

الفصل الثالث:

أساسيات الرياضية لنماذج الماركوف المخفية HMM

1. مقدمة :

نموذج ماركوف المخفي هو نموذج إحصائي قوي أثبت قدرته في مجال التنبؤ والتعريف بسبب قوة تراكيبه الرياضية ، وهو مستخدم بكثرة في تطبيقات نمذجة الإشارة ،كما وجدت لنماذج ماركوف العديد من التطبيقات في المجالات المهتمة بمعالجة الإشارة مثل معالجة الكلام .

كما أنها نجحت في تطبيقات معالجة اللغات الطبيعية في مستوى منخفض مثل تصنيف أقسام الكلام ، تقطيع العبارات واستخراج المعلومات.

آندري ماركوف (**A.Markov**) قدم اسمه للنظرية الرياضية بعمليات ماركوف (**Markov Proccess**) في أوائل القرن العشرين ، ولكن **باوم** (**Baum**) وزملائه من قاموا بتطوير نظرية نماذج ماركوف المخفية (**HMM**) في عام **1960**م.

2. سلاسل ماركوف المخفية :

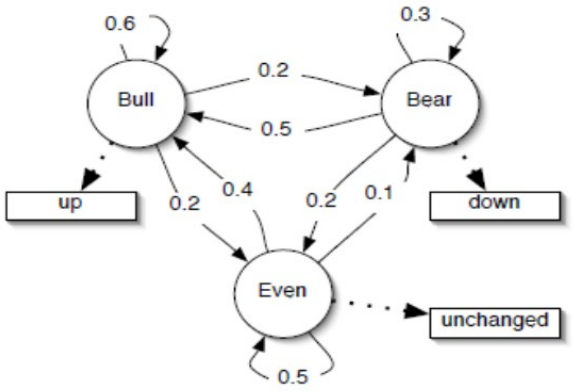
سلاسل ماركوف هي عملية عشوائية (عائلة من المتغيرات العشوائية) تحمل خاصية تسمى بالخاصية الماركوفية حيث في هذه العملية يمكن لتنبؤ المستقبل اعتمادا على الماضي المتمثل في الحاضر ، أي لا يوجد حاجة لمعرفة الماضي للتكهن بالمستقبل بل يكفي معرفة الحاضر .

أما بالنسبة لنماذج ماركوف المخفية فهي عملية عشوائية مزدوجة تتمثل في مستويين اثنين : المستوى الأول أو الطبقة الأولى هي الطبقة العليا المخفية وتتمثل بعملية ماركوف بحالات لا يمكن ملاحظتها (عملية ذات حالات مخفية) ومن هنا جاءت التسمية بالمخفية .

والمستوى الثاني هو المشاهدات وهو مخرج الحالات السابقة ، وفي الحقيقة الملاحظات (المشاهدات) عبارة عن تابع احتمالي بالنسبة للمستوى المخفي ، فحالات عملية ماركوف المختلفة ستنتج الملاحظات المخرج باحتمالات مختلفة .

1.2 عمليات ماركوف (Markov proeesses)

تعرف بأنها " الوسيلة التي يتم بها تحليل التغيرات الحالية لمتغير عشوائي معين من أجل التنبؤ بالمتغيرات المستقبلية لهذا المتغير ، ولقد اطلقت هذه التسمية نسبة للعالم الرياضي الروسي[1] **آندري ماركوف** (**A.Markov**) لنستعرض المثال التالي عن عملية ماركوف ، والنموذج المقدم يصف نموذج بسيط لمؤشر سوق الأسهم ، في النموذج ثلاث حالات (**bull، bear،even**) وثلاث ملاحظات (**unchanged down up**)



الشكل 1.3 : عملية ماركوف[2]

مثلاً إذا صادفنا سلسلة الملاحظات التالية [up-up-down] :فبالإمكان بسهولة معرفة الحالات الموافقة [bull-bull-bear] حيث كل ملاحظة لها مخرج ترتبط بحالة واحدة فقط .

2.2 نماذج ماركوف المخفية:

الشكل التالي يقدم مثالاً عن كيفية توسيع المثال السابق إلى (HMM) ، والنموذج الجديد يسمح لجميع الملاحظات أن تصدر من جميع الحالات باحتمالات المحدد [3].

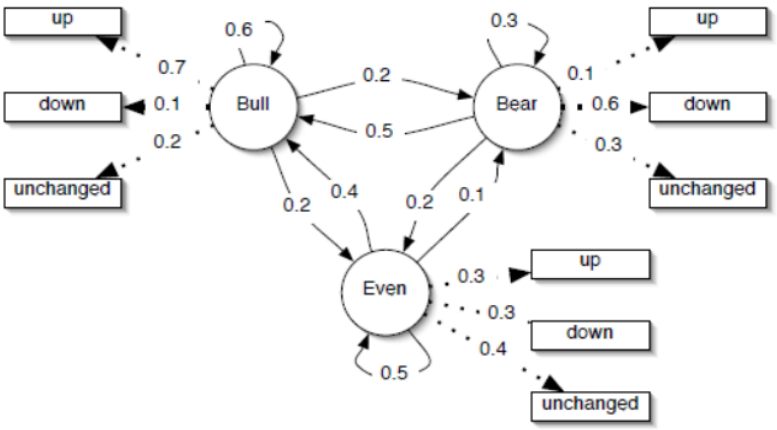
---

[1] أسعد عبد علي ، دراسة احصائية عن الهدر الزمني ، رسالة ماجستير ، جامعة بغداد ، كلية الاقتصاد ، 1988م ، ص20

[2] المرجع السابق ، التعرف على الكلمات المعزولة ، ص 62

[3] Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE.

[1].هذا التغيير يجعل النموذج الجديد معبراً أكثر، ويمكنه تمثيل الحدس بطريقة أفضل[



الشكل 2.3: نموذج الماركوف المخفي[2]

في هذه الحالة يمكن bull market أن يمر بأيام جيدة و أيام سيئة ، ولكن الأيام الجيدة أكثر غالباً. الاختلاف الرئيسي هنا أننا إذا علمنا سلسلة الملاحظات السابقة نفسها فإنه لا يمكن معرفة سلسلة الحالات التي ولّدت هذه الملاحظات ، و لهذا فإنّ سلسلة الحالات المخفية (Hidden) وعلى أي حالة يمكننا بالحسابات لمعرفة احتمالية إنتاج النموذج للسلسلة (الملاحظات) ، و أيضاً معرفة سلسلة الحالات التي على الأرجح ولّدت هذه السلسلة.

3.2 مثال توضيحي:

لنفرض لدينا N وعاء كبير ، وفي كل منها يوجد عدد من الكرات الملونة ، لنفرض أن M لون مختلف للكرات في هذه الأوعية [3].

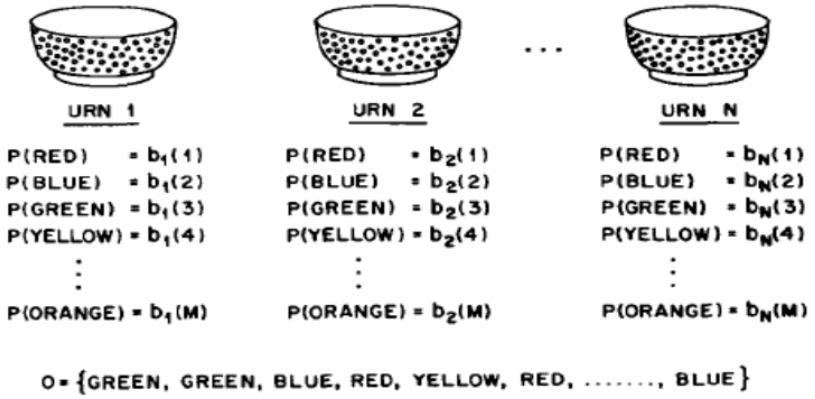
[1] المرجع السابق ، التعرف على الكلمات المعزولة ، ص 63.

[2] المرجع نفسه ، ص 64.

[3] Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE.

يقوم شخص ما بالعملية التالية للحصول على سلسلة من الملاحظات ، يختار الشخص وعاء ما عشوائيا ويسحب منه كرة و يسجل لونها ثم يعيدها ، ويكرر هذه العملية Tمرة ، إن اختيار الوعاء التالي يكون وفق عملية عشوائية اعتمادا على الوعاء الحالي (عملية ماركوف) .

في النهاية نحصل على سلسلة من ألوان الكرات المسحوبة (الملاحظات) يجب ملاحظة أنه في هذه العملية حالات النموذج هي أوعية الكرات (وعاءN) ومخرج كل حالة على الملاحظات المتمثلة بألوان الكرات (M لون محتمل) و تختلف احتمالية مخرج كل لون في كل حالة ، واختيار الوعاء التالي (الانتقال إلى الحالة التالية) يتم عن طريق مصفوفة احتمالات الإنتقال .



الشكل 3.3 : الأوعية و الكرات الملونة

3. رموز النموذج (عناصرHMM)

a. N ، عدد الحالات الكلي في النموذج أو النظام .على الرغم أن الحالات مخفية إلا أنه العديد من التطبيقات نجد دلالات فيزيائية مرتبطة بهذه الحالات (في المثال السابق نمثل كل حالة بوعاء)

بشكل عام فإن الحالات متصلة ببعضها بشكل يسمح الوصول من أي حالة إلى أخرى بشكل مباشر نرمز حالات المستقلة بـ : $s = \{s1, s2, \dots\dots\dots sn\}$

b. **M**, عدد رموز المشاهدات والملاحظات الممكنة في النموذج ، وتمثل المشاهدات لمخرج النظام الذي تجري نمذجته (في مثال الأوعية ألوان الكرات هي ما نلاحظه) [1] [2]

نرمز إلى الملاحظات المستقلة بـ : $s = \{v1, v2, \dots\dots\dots vm\}$

c. توزع الاحتمالية للانتقالات بين الحالات $A = \{a_{ij}\}$

$$aij = P[qt + 1 = sj | qt = si] \quad 1 \leq i,j \leq N \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (3.1)$$

d. التوزع الاحتمالي لرموز المخرج (المشاهدات):

In state j ,B=$\{b_j(k)\}$. Where :

$$\{bj(k)\} = P[vkat\ t | qt = j] \quad 1 \leq j \leq N\ ;\ 1 \leq k \leq M \quad \dots\dots\dots\dots\dots\dots (3.2)$$

E: توزيع الحالة البدائية : المعادلة التالية [3]:

$$\boldsymbol{\pi = \{\pi i\} \text{where :}}$$

$$\pi = P[q1 = Si], \quad 1 \leq i \geq N. \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (3.3)$$

$$\lambda = (\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\pi}).$$

4. أنواع نماذج ماركوف المخفية (Types of HMMs)

حتى الآن نفترض حالة خاصة من النماذج ، حيث كل حالة في النموذج يمكن الوصول إليها بخطوة واحدة من أي حالة أخرى من هذا النموذج [4].

---

[1] ونيسة بوختالة ، البنية الصوتية لقصار السورة القرآنية و أثرها في تعليم اللغة العربية 2007 ، ص44.

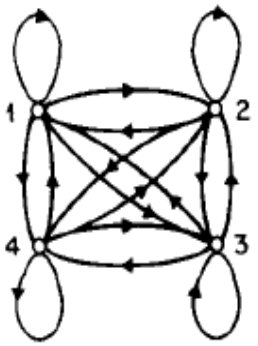
[2] المرجع السابق ، علم الأصوات اللغوي ، ص20.

[3] المرجع السابق ، التعرف على الكلمات المعزولة ، ص 89.

[4] Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE.

كما هو موضح بالمثال التالي : من أجل نموذج حيث N= 4 تكون مصفوفة الإنتقالات A و مخطط الإنتقال كما في الشكل التالي ، حيث a_{ij} هو عدد موجب بين 0 و 1

النموذج الثبوتي: و هو النموذج الذي تكون فيه كل الحالات انتقالية.



$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

الشكل 4.3: النموذج الثبوتي

توجد أنوع أخرى من النماذج لحساب الخصائص الملاحظة للإشارة التي تتم نمذجتها بشكل أفضل من النموذج السابق المتصل بشكل كامل .

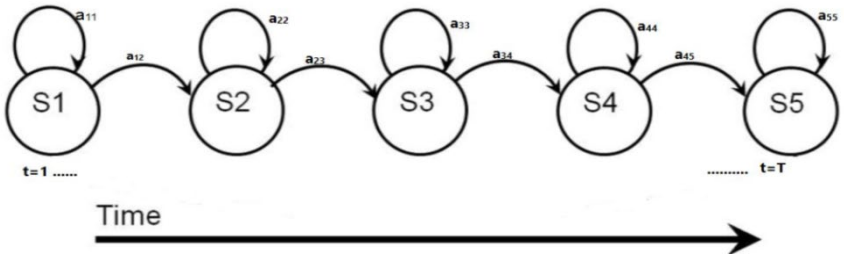
ومثل هذا النموذج يسمى (**Left–Right Model**) حيث أن السلسلة المرتبطة بهذا النموذج لها الخاصية التالية [1] مع تقدم الزمن (اللحظات الزمنية) فإن مؤشر الحالة يزداد أو يبقى نفسه ، أي أنه يتم التقدم من اليسار إلى اليمين .

لتوضيح هذا النموذج نأخذ المثال التالي : لنفرض لدينا نموذج حيث N=5 فإنه في اللحظة t=1 يكون النظام في الحالة s1 ومع تقدم مؤشر الزمن t إلى اللحظة t=2 فإن النظام إما أن ينتقل إلى الحالة s2 باحتمال الشكل العام لنموذج (**Left–Right**) كما في المثال التالي [2]:

[1] abiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE.

[2] المرجع السابق ، التعرف على الكلمات المعزولة ، ص 87

نموذج الأيسر– الأيمن : و هو الذي تكون فيه بعض الحالات انتقالية ، بحيث أن:



الشكل 5.3: نموذج Left-Right Model

إن هذا الشكل من النموذج له خاصية مرغوبة في سهولة نمذجة الإشارات التي تتغير خصائصها مع مرور الوقت (مثل الكلام) .

السمة الأساسية للنموذج السابق (بكافة أشكاله) أن احتمال الانتقال تتوقف على الشرط التالي :

$$Aij = 0 \text{ f or all } j<i \quad : \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.4)$$

وغير مسموح بالانتقال إلى حالة إذا كان مؤشرها أصغر من مؤشر الحالة الموجودين فيها (الحالية) و أيضا احتمال الحالة البدائية له الشرط التالي كما رأينا في المثال السابق :

$$\pi_i = \begin{cases} 0, & i \neq 1 \\ 1, & i \neq 1 \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.5)$$

وقد نضع بعض الشروط الإضافية على الانتقالات في النموذج بحيث نتجنب قفزات كبيرة بين الحالات ، حيث أنه كما رأينا فإنه يبدأ بالحالة الأولى دائما ويتحرك إلى الأمام فقط . مثلا إذا اعتبرنا الشرط التالي :

$$a_{ij=0, i>i+\Delta}$$

و في مثالنا السابق كانت قيمة دلتا تساوي 1 أي يمنع الانتقال إلى حالة غير التالية مباشرة ، إذا افترضنا من أجل دلتا تساوي 2 يمكن أن نمثل مخطط الانتقال ومصفوفة من أجل نموذج فيه N-4 كالتالي [1]:

[1] المرجع السابق ، التعرف على الكلمات المعزولة ، ص 89.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & 0 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}.$$

الشكل 6.3: نموذج Left–Right فيه دلتا تساوي 2

من الواضح أيضا في هذا النموذج أن الحالة النهائية الأخيرة S_n هي حالة ماصة أي أنها لها خاصية A_{NN}-1.

5. توليد ملاحظات من النموذج :

بإعطاء قيم مناسبة لكل من البارامترات السابقة استخدم النموذج الناتج لتوليد سلاسل ملاحظات :

$$O = \{ O_1 O_2 \ldots O_T \}$$

1– choose an initial state $q_1 = s_i$ according to initial state distribution π

2– set t =1

3– choose $Q_t = V_k$ according to symbol probability distribution in state $s_{i\ ;i,e,.\ ;b}(k)$.

4– Transit to a new state $q_{t+1} = s_j$ according to the state transition probability

Distribution for state $s_{i\ ,i,e,,.aij.}$

5– Set t= t+1 ; return to step 3 if t < t : otherwise terminate the procedurev .

6. المسائل الثلاثة الأساسية نموذج ماركوف المخفية :

إذا كان لدينا نموذج HMM كما درسنا فإننا نجد مسائل أساسية تهتم بإيجاد الحلول لها ليصبح ذو فائدة في تطبيقات حقيقية ، وهذه المسائل هي كالتالي :

حلول المسائل الأساسية لنماذج ماركوف المخفية:

لكي تكون نماذج ماركوف المخفية مفيدة في كثير من التطبيقات في العالم الحقيقي هناك ثلاث مسائل أساسية يجب أن يتم حلها و هذه المسائل هي[1] [2]:

المسألة (1):

تعمل مسألة التقييم على حساب احتمالية متسلسلة المشاهدات $p(O|\lambda)$ للنموذج عندما يكون المعطى هو النموذج $\lambda= (A, B, \pi)$ و متسلسلة المشاهدات $O =\{ O_1 O_2 \ldots O_T \}$، أي تحديد الإمكان لمتسلسلة المشاهدات (O) عندما يكون المعطى هو النموذج، و تحل عن طريق الخوارزمية الأمامية –الخلفية Forward-Backward Algorithm[3] [4] [5]

المسألة (2):

إذا كان المعطى هو النموذج $\lambda= (A, B, \pi)$ و متسلسلة المشاهدات

$O = \{O_1 O_2 \ldots O_T\}$، تتناول هذه المسألة كيفية تعديل معلمات النموذج $\lambda= (A, B, \lambda)$ لتعظيم $P(O|\lambda)$، و إيجاد متسلسلة الحالة المثلى لعملية ماركوف المخفية ، و التي تمثل أفضل توضيح للمشاهدات[6].

---

[1] ان احمد حسن ، رشا رعد هادي ، استخدام نماذج ماركوف المخفية في التعرف على صور الوجه الاعتيادي و المحددة حافاته ، المجلة العراقية للعلوم الإحصائية (21) 2012 م.

[2] L.R. Rabiner, " A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition " , Proceedings of IEEE , Vol. 77 , No. 2 , P. 257 – 286, (1989).

[3] T. Aarnio, "Speech recognition with hidden Markov models in visual communication", M.Sc. Thesis, College Computer Sciences, Turku-University, (1999).

[5] M. Stamp, " A Revealing Introduction to Hidden, Markov Models", IEEE , Vol. 51 , No. 7 , P. 347 – 356, (2003).

[6] المرجع السابق ، استخدام سلاسل ماركوف المخفية في تمييز حروف العلة في اللغة الانكليزية ، ص19.

المسألة (3): إذا كان المعطى هو متسلسلة المشاهدات $O = \{O_1 O_2 \ldots O_T\}$ و الأبعاد (N) و (M) ، تعمل مسألة التدريب على إعادة تقدير معلمات النموذج $\lambda=(A,B,\pi)$ الذي تكون فيه $P(O|\lambda)$ عظمى، وتحل هذه المسألة عن طريق Baum-Welch [1].

7. حلول المسائل الأساسية لنماذج ماركوف المخفية HMM:

1.7 حل المسألة (1) (مسألة التقييم):

إن الطريقة المباشرة لحسابها تكون من خلال سرد متسلسلة حالة ذات طول (T) (حيث أن (T) تمثل عدد المشاهدات) و لنتأمل المتسلسلة الثابتة للحالة:

$$Q = \{q_1 q_2 \ldots q_T\}$$

حيث أن (q_1) تمثل الحالة الابتدائية، أما الاحتمالية لمتسلسلة المشاهدات (O) بالنسبة لمتسلسلة الحالة في المعادلة (1) فهي:

$$P(O|Q,\lambda) = \prod_{t=1}^{T} P(O_t|q_t,\lambda) \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.1)$$

و ذلك بافتراض الإستقلال الإحصائي للمشاهدات، أي أن:

$$P(O|Q,\lambda) = P(O_1|q_1,\lambda) * P(O_2|q_2,\lambda) * \ldots * P(O_T|q_T,\lambda)$$

$$= \prod_{t=1}^{T} P(O_t|q_t,\lambda)$$

$$= \prod_{t=1}^{T} b_{q_t}(O_t) \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.2)$$

[1] M. Stamp, " A Revealing Introduction to Hidden, Markov Models", IEEE , Vol. 51 , No. 7 , P. 347 – 356, (2003).

و أن الاحتمالية لمتسلسلة الحالة (Q) هي:

$$P(Q \mid \lambda) = P(q_1, q_2, \ldots q_T \mid \lambda)$$

$$= P(q_1; \lambda) \prod_{t=2}^{T} P(q_t \mid q_{t+1}; \lambda)$$

$$= \pi_{q_1} \prod_{t=2}^{T} a_{q_{t-1} q_t} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.3)$$

و ذلك حسب الخاصية الماركوفية حيث أن (π) تشير إلى احتمالية الحالة الابتدائية للحالة (q1).

أما الاحتمالية المشتركة لـ (O) و (Q)، أي ورود بالوقت نفسه، فهي تمثل النتيجة للمعادلتين أي أن:

$$P(O, Q \mid \lambda) = P(O \mid Q, \lambda) * P(Q \mid \lambda) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.4)$$

و من هنا يمكن الحصول على الاحتمالية لـ (O) عندما يكون المعطى هو النموذج (λ) و ذلك عن طريق جمع هذه الاحتمالية المشتركة لكل المتسلسلات الحالة (q) ، حيث أن :

$$P(O \mid \lambda) = \sum_{\text{all } q} P(O \mid Q, \lambda) * P(Q \mid \lambda)$$

$$= \sum_{q1,q2,\ldots,qT} \pi_{q1} * b_{q1}(O_1) * a_{q1\,q2} * b_{q2}(O_2) * \ldots * a_{qT-1\,qT} * b_{qT}(O_T) \quad \ldots\ldots\ldots\ldots\ldots\ldots (3.5)$$

إن هذه الحسابات تكون غير عملية وإن كانت قيم (N) و (M) صغيرة، فمثلا إذا كانت (N = 5) (وهي تمثل عدد الحالات المخفية) و كانت (T = 100) (و هي تمثل عدد المشاهدات)، فيجب القيام بـ $2(100) * 5^{100} \approx 10^{72}$ عملية حسابية، و من ذلك يتضح بأن هذه الحسابات تتطلب جهدا كبيرا لحل المسالة (1) ،

ولذلك استخدم الباحثون الخوارزميات الأمامية-الخلفية لحل تلك المسألة، و فيما يأتي توضيح لهذه الخوارزميات[1]:

أولا الخوارزميات الأمامية [2] [3]:

بإمكاننا أن نعرف المتغير الأمامي $\alpha_t(i)$ بأنه احتمالية متتابعة المشاهدات الجزئية O1, O2, … , Ot والحالة S_i عند الزمن t عندما يكون المعطى هو النموذج λ، هذا و يمكن التعبير عنه رياضيا بالصيغة الآتية:

$$\alpha_t(i) = P(O_1, O_2, \dots, O_t, q_t = S_i | \lambda) \quad \dots\dots\dots (3.6)$$

و أن المتغيرات الأمامية يمكن حسابها من خلال الخطوات الآتية:

البداية:

$$\alpha_t(i) = \pi_i \, b_i(O_1), \; i = 1, 2, \dots, N \quad \dots\dots\dots (3.7)$$

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) \, a_{ij}\right] b_j(O_{t+1}), \quad t = 1, 2, \dots, T-1, \; j = 1, 2, \dots, N \quad \dots\dots\dots (3.8)$$

الإنهاء:

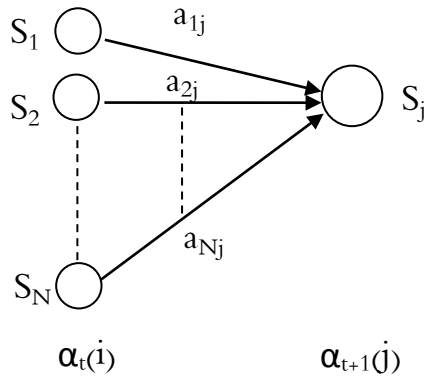
$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i) \quad \dots\dots\dots (3.9)$$

(1) المرجع السابق ، استخدام سلاسل ماركوف المخفية في تمييز حروف العلة في اللغة الانكليزية ، ص12.

(2) باسل يونس ، تمييز السلالم الموسيقية باستخدام نموذج ماركوف المخفي ، المجلة العراقية للعلوم الإحصائية (15) 2012 م.

(3) T.Pham, "Alignment-free sequence comparison with vector quantization and hidden Markov models", proceeding of IEEE of the Computational Systems Bioinformatics, (2003).

و يمكن توضيح الخوارزمية الأمامية من خلال الشكل (3) بالآتي:



الشكل 7.3: متتابعة العمليات المطلوبة لحساب المتغير

إن حسابات ال (α) تتطلب $N(N+1)(T-1)+N$ عملية ضرب و $N(N-1)(T-1)$ عملية جمع و هي أفضل من الحسابات المباشرة و التي تتطلب $(2T*N^T)$ عملية حسابية [1][2].

الخوارزمية الخلفية [3]:

بأسلوب الخوارزمية الأمامية نفسه يمكن أن نعرف رياضيا المتغير الخلفي $\beta_t(i)$ كما يأتي:

$$\beta_t(i) = P(O_{t+1}\, O_{t+2} \ldots O_T \mid q_t=S_i, \lambda) \quad \ldots\ldots (3.10)$$

أي أنه يمثل احتمالية متتابعة المشاهدات الجزئية من الزمن t+1 إلى T عندما يكون المعطى هو Si عند الزمن t و النموذج λ .

---

[1] S.Duran, "Keyword spotting using hidden markov models", M.Sc. Thesis, Bogazici University, (2001) P. 222

[2] L.R. Rabiner, " A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition " , Proceedings of IEEE , Vol. 77 , No. 2 , P. 257 – 286, (1989).

[3] المرجع السابق ، تمييز السلالم الموسيقية باستخدام نموذج ماركوف المخفي ، ص55

و يمكن حساب المتغيرات الخلفية من خلال الخطوات الآتية:

البداية:

$$\beta_t(i) = 1 \; ; i = 1, 2, \ldots , N \qquad (3.11)$$

التعاقب:

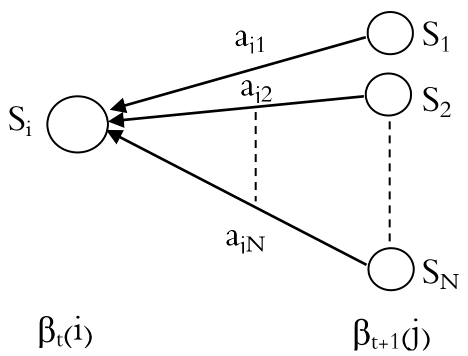
$$\beta_t(i) = \sum_{j=1}^{N} a_{ij}\, b_j\,(O_{t+1})\, \beta_{t+1}(j) \quad , t = T-1, T-2, \ldots , 2,1 \quad i = 1, 2, \ldots , N \qquad (3.12)$$

الإنهاء:

يمكن حساب الاحتمالية $P(Q|\lambda)$ باستخدام المتغيرات الأمامية و الخلفية كما يأتي:

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_t(i)\, \beta_t(i) \quad , \text{for any } t \qquad (3.13)$$

و الشكل8.3: يوضح العمليات المطلوبة لحساب الاحتمالية الخلفية:



الشكل 8.3: متتابعة العمليات المطلوبة لحساب المتغير الخلفي

2.7 . حل المسالة (2):

هناك بعض الطرائق الممكنة لحل المسألة (2)، أي لإيجاد متسلسلة الحالة المثلى المقترنة بمتسلسلة المشاهدات المعطاة منها :

خوارزمية (Viterbi) [1]

هي طريقة البرمجة الحركية المستخدمة لحساب مسار الحالة الانتقالية الأكثر احتمالا عندما يكون المعطى هو متسلسلة المشاهدات للرموز، و هي مشابهة للخوارزمية الأمامية ما عدا كون استخدام الـ "أكبر" "Max " بدلا من الـ "مجموع" " Σ ".

و لإيجاد أفضل متسلسلة حالة وحيدة $Q = \{q_1\ q_2 \dots q_T\}$ بالنسبة لمتسلسلة المشاهدات المعطاة $O=\{O_1\ O_2 \dots O_T\}$ يجب تعريف المقدار:

$$\delta_t(i) = \max_{q1, q2, \dots qt-1} P(q_1, q_2, \dots q_t = i, O_1\ O_2 \dots O_t | \lambda) \dots (3.14)$$

إذ أن $\delta_t(i)$ تمثل أعلى احتمالية على طول المسار الوحيد خلال الزمن (t) و المحسوبة بالنسبة لمشاهدات (t) الأولى و المنتهية في الحالة (S_i) ، فضلا عن استخدام $\psi_t(t)$ و التي تساعد في حفظ تتابع الأثر (Keep Track) للمسار الفعلي، أي أن $\psi_t(t)$ ستوضح أي حالة خلال الزمن (t−1) تقود لاحتمالية $\delta_t(i)$ الأعلى خلال الزمن (t).

[1] M. Stamp, " A Revealing Introduction to Hidden, Markov Models", IEEE , Vol. 51 , No. 7 , P. 347 – 356, (2003).

أما خطوات خوارزمية (Viterbi) فهي كالآتي:

البداية:

$$\delta_1(i) = \pi_i * b_i(O_1), \quad i = 1, \dots, N \qquad \dots\dots (3.15)$$

$$\psi_1(i) = 0 \qquad \dots\dots (3.16)$$

التعاقب:

$$\delta_t(i) = \max_{1 \le i \le N} [\delta_{t-1}(i) * a_{ij}] * b_j(O_t), \quad t = 2, \dots, T, \quad J = 1, \dots, N \qquad \dots\dots (3.17)$$

$$\psi_t(j) = \operatorname{argmax}_{1 \le i \le N} [\delta_{t-1}(i) * a_{ij}], \quad t = 2, \dots, T, \quad J = 1, \dots, N \qquad \dots\dots (3.18)$$

النهاية:

$$P^* = \max_{1 \le i \le N} [\delta_T(i)] \qquad \dots\dots (3.19)$$

$$q^*_T = \operatorname{argmax}_{1 \le i \le N} [\delta_T(i)] \qquad \dots\dots (3.20)$$

التتابع المعاكس

$$Q^* = \{ q^*_1 \dots q^*_T \} \qquad \dots\dots (3.21)$$

$$q_t^* = \psi_{t+1}(q^*_{t+1}), \quad t = T-1, T-3, \dots, 1 \qquad \dots\dots (3.22)$$

ثم تقرأ المتسلسلة الأفضل للحالات من متجهات الـ (ψ_t) [1] [2] [3].

---

[1] المرجع السابق ، استخدام سلاسل ماركوف المخفية في تمييز حروف العلة في اللغة الانكليزية ، ص33

[2] L.R. Rabiner, " A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition " , Proceedings of IEEE , Vol. 77 , No. 2 , P. 257 – 286, (1989).

[3] J.Parker, " Profile Hidden Markov Models " , Biophysical Journal , Vol. 17 , P. 1335 – 1348, (1999).

3.7 . حل المسألة (3):

هي المسألة الأكثر صعوبة لضبط بارامترات النموذج (A, B, π) و ذلك لتعظيم الاحتمالية لمتسلسلة المشاهدات لتحقيق أفضل احتمال لسلسلة الملاحظات لمعرفة النموذج.[1].

خوارزمية باوم ويلش (Baum–Welch) [2]:

هذه الخوارزمية تناقش خطوات التكرار بالاعتماد وبشكل ابتدائي على العمل الكلاسيكي للعالم (L.E.Baum) و زملائه و ذلك لاختيار معلمات النموذج ، و لوصف خطوات إعادة تقدير معلمات (HMM)، يتم تعريف المتغيرين الآتيين:

أما المتغير الأول هو:

$$\zeta t(i,j) = P(q_t = S_1 , q_{t+1} = S_j \,|O, \lambda)$$

$$= \frac{P(O , q_t = S_1 , q_{t+1} = S_j \,|\lambda)}{P(O|\lambda)}$$

$$= \frac{\alpha_t (i)\, a_{ij}\, b_j(O_{t+1})\, \beta_{t+1}(j)}{P(O|\lambda)}$$

$$= \frac{\alpha_t (i)\, a_{ij}\, b_j(O_{t+1})\, \beta_{t+1}(j)}{\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_t (i)\, a_{ij}\, b_j(O_{t+1})\, \beta_{t+1}(j)} \quad , t = 1, \dots , T-1 \qquad \dots\dots(3.23)$$

[1] مصون نبهان حمصي جربيني ، نظام تفاعلي ذكي من اجل التعليم على الشبكة العنكبوتية ، رسالة دكتراه ، جامعة حلب سوريا 2010 م.

[2] المرجع السابق ، تمييز السلالم الموسيقية باستخدام نموذج ماركوف المخفي، ص67

حيث أن البسط هو الاحتمال السابق ذاته ، والقسمة على$P(O|\lambda)$ يعطي تقييسِ احتمال المطلوب Si .

كما عرفنا سابقا في المسألة 2 ، الرمز الذي يدل على احتمال في كون النظام في الحالة Si عند اللحظة t

المتغير الثاني هو:

$$\gamma_{t(i)} = \frac{\alpha_t(i)\,\beta_t(i)}{\sum_{i=1}^{N} \alpha_t(i)\,\beta_t(i)} \quad \ldots\ldots\ldots\ldots (3.24)$$

كما يمكننا حساب المتغير $\gamma_t(i)$ باستخدام المتغير $\zeta_t(i,j)$)

$$\gamma_t(i) = \sum_{j=1}^{N} \zeta_t(i,j) \quad \ldots\ldots\ldots\ldots (3.25)$$

و باستخدام الصيغ المذكورة أنفا لإعادة تقدير نموذج جديد من النموذج القديم نقوم بالإجراءات الآتية:

العدد المتوقع للانتقالات في الحالة i عند الزمن t

$$\pi_i = \gamma_1(i) \quad (t=1) \quad \ldots\ldots\ldots\ldots (3.26)$$

$$a_{ij} = \frac{\text{العدد المتوقع للانتقالات من الحالة } S_i \text{ إلى الحالة } S_j}{\text{العدد المتوقع للانتقال من الحالة } S_i}$$

$$= \frac{\sum_{t=1}^{T-1} \zeta_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \quad \ldots\ldots\ldots\ldots (3.27)$$

$$b_j(k) = \frac{\text{العدد المتوقع للأزمنة في الحالة } j \text{ و رموز المشاهد } V_K}{\text{العدد المتوقع للأزمنة في الحالة } j}$$

$$= \frac{\sum_{t=1}^{T} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(t)} \quad \text{S, } O_t = v_k \quad ; j = 1, 2, \ldots, N \quad \ldots\ldots\ldots\ldots (3.28)$$

من المعادلات أعلاه نحصل على نموذج جديد يرمز له ب $\underline{\lambda}$ الذي يحقق: [1]

$$P(O|\underline{\lambda}) > P(O|\lambda) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.29)$$

و باستخدام الصيغ أعلاه يمكن إدراج الصيغ الآتية و ذلك لإعادة تقدير معلمات(HMM) (A,B,π) بالاعتماد على الخطوات السابقة و باستخدام (و بشكل متكرر) $\underline{\lambda}=(\underline{A}, \underline{B}, \underline{\pi})$ بدلا من $\lambda=(A,B,\pi)$ و تكرار حسابات إعادة التقدير فسوف تعدل الاحتمالية لـ (O) المشاهدة من النموذج إلى أن يتم التوصل إلى بعض النقاط المحددة، والنتيجة النهائية لخطوات إعادة التقدير هذه تعرف بـ (تقدير الإمكان الأعظم لـ (HMM) .[2]

كما يجب ملاحظة قيود النموذج $\underline{\lambda} = (\underline{A},\underline{B},\underline{\pi})$ أن تكون مقيدة بالقيود الآتية [3]:

$$\sum_{i=1}^{N} \pi_i = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.30)$$

$$\sum_{i=1}^{N} a_{ij} = 1 \quad ; i = 1, 2, \ldots, N \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.31)$$

$$\sum_{i=1}^{N} b_j(k) = 1 \; ; i = 1, 2, \ldots, N \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.32)$$

إن الجانب المهم من خطوات إعادة التقدير هو شرط (أو قيد) التصادفية لمعلمات (HMM)، و بالذات المتحققة في كل تكرار أي أن :

---

[1] المرجع السابق ، تمييز السلالم الموسيقية باستخدام نموذج ماركوف المخفي، ص68

[2] المرجع السابق ، استخدام سلاسل ماركوف المخفية في تمييز حروف العلة في اللغة الانكليزية ، ص57

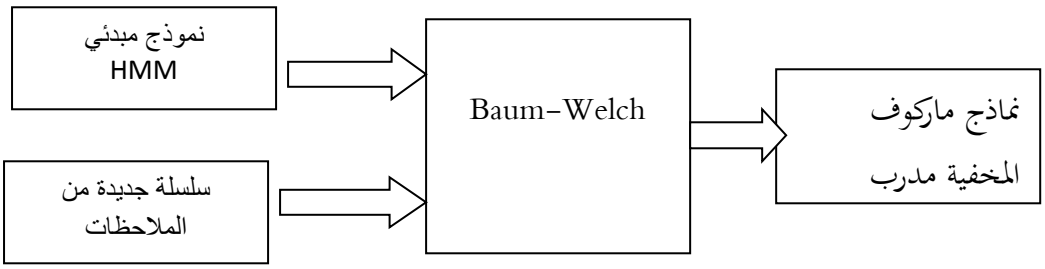
[3] L.R. Rabiner, " A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition " , Proceedings of IEEE , Vol. 77 , No. 2 , P. 257 – 286, (1989).

عملية إعادة التقدير التي تمثل عملية التكرار بدأت من اختيار أفضل قيم تخمينية (أو تقديرية) لـ $\lambda=(A, B, \pi)$ ، و إذا لم تتوافر هذه القيم فسوف يتم اختيار قيم عشوائية بحيث أن :

$\pi_i \approx 1/N$ $a_{ij} \approx 1/N$ $b_j(k) \approx 1/M$

مع التأكيد على أن تكون (A, B, π) دائما ذات صف تصادفي ، و أن لا تكون قيم عناصر المصفوفة متساوية [1].

مع ملاحظة أن تكون مقيدة بالقيود المذكورة سابقا [2].



الشكل 9.3: تدريب نموذج نماذج ماركوف مخفي باستخدام خوارزمية Baum-Welch [3]

[1] المرجع السابق ، ص259

[2] T. Aarnio, "Speech recognition with hidden Markov models in visual communication", M.Sc. Thesis, College Computer Sciences, Turku-University, (1999).p58

[3] Rabiner L.R. A Tutoril on Hadden Markov Models and Selected Application in Speech Recognition of the IEEE P275 –P285

يتم تدريب نماذج ماركوف المخفية باستخدام خوارزمية باوم ويلش و التي تعطى بالخطوات التالية:

1- تهيئة النموذج $\lambda= (A, B, \pi)$

2- حساب $\alpha_t (i)$

3- حساب $\beta_t(i)$

4- حساب $\zeta_t (i, j)$ و التي تمثل احتمال الوجود في الحالة i خلال الزمن t و الانتقال من الحالة j خلال الزمن t+1 لمتتالية الملاحظات $O = \{O_1,O_2,O_3, ..O_T\}$

و لنموذج $\lambda= (A, B, \pi)$

5- حساب $\gamma_t (i,j)$ و التي تمثل احتمال الوجود في الحالة i خلال الزمن t لمتتالية الملاحظات $O = \{O_1,O_2,O_3, ..O_T\}$ و لنموذج $\lambda= (A, B, \pi)$

6- إعادة تقدير النموذج $\lambda= (A, B, \pi)$

7- إذا ازدادت قيمة $P(O|\lambda)$ و $\lambda = \lambda$ عندئذ اذهب إلى الخطوة 9.

8- عندما يكون $P(O|\lambda) > P(O|\lambda)$ يدل على أنه تم تحسين النموذج و في الحالة المغايرة يتم تكرار الخطوات من 1 إلى 7.

9- في النهاية يجب التوقف إذا لم تكن $P(O|\lambda)$ تزايدية.

8. الخاتمة:

إن تقنية التعرف على الكلام المستخدمة تندرج ضمن الأسلوب الإحصائي للتعرف علي الأنماط وهي نماذج ماركوف المخفية لأنها حالياً أكثر تقنية في التعرف على الكلام و تضم مجموعة منتهية من الحالات التي يمكن أن تكون رمزا أو حرفا أو كلمة أو أي شيء آخر مما يجعلها قابلة للاستخدام في العديد من المجالات وهي عبارة عن نظام محطات الآلة المحدود كما أنها عبارة عن عملية تصادفية مضاعفة تستخدم للتعرف على الخواص الإحصائية ومن المميزات الخاصة التي تتمتع بها نماذج ماركوف المخفية هي الخاصية الماركوفية وهي أن الحالة التي تحدث في زمن ما تعتمد في حدوثها فقط على الحالة في الزمن الذي يسبقها مباشرة.

سلسلة ماركوف المخفية هي عبارة عن متتابعة من الحالات التي يمر بها نظام معين ، أو ظاهرة معينة خلال الانتقال عبر المحور الزمني أو أي دليل آخر، بحيث أن معرفة حالة النظام في زمن ما لا تعتمد على جميع تاريخ النظام .

الفصل الرابع:

الدراسة العملية و النتائج

1. مقدمة:

يهدف هذا العمل إلى إنجاز آلية تمكن من التعرف الآلي للكلام العربي بالإعتماد على قاعدة بيانات معينة والتي هي عبارة عن عينات من الجمل اللغة العربية وتعتمد هذه الآلية على أداة تقنية معروفة بـ(HTK) وهي تقنية مطورة جدا ومستعملة أساسا في البحوث المتعلقة بالتعرف الآلي للكلام العربي على المنطوق سيما ما يتعلق بتحليل الجمل.

كما تم استخدام نماذج ماركوف المخفية (HMM) في عملية تمييز الآلي للكلام العربي بالإعتماد على الفونيمات الأحادية ، بناء مكون معالجة الإشارة الرقمية و الذي يعتمد على سمات مقاطع الصوتية الخاصة بمجموعة من الجمل عن طريق تحليلها بتقنيات مختلفة ثم إدخالها الى نظام للتعرف الآلي للكلام المنطوق باللغة العربية برنامج يقوم بإنتاج القاموس الصوتي باستعمال قواعد نطق اللغة العربية ونشير هنا إلى ان هذا البرنامج ينتج القاموس مع الأخذ بعين الاعتبار عدد من الاختلافات التي قد تظهر في نطق الكلمة وهي عبارة عن سلسلة نهائية من الحالات والملاحظات والتي يتم توليدها عشوائياً ويتم الانتقال من حالة الى حالة اخرى عن طريق مصفوفة الاحتمالات و التي اقترحناها لكشف نسب تمييز لكل وحدة نطق وكشف الأخطاء اللغوية ، تتكون مراحل التعرف الكلام من ثلاثة مراحل :المرحلة الأولى : يتم استقبال الكلام و تحويله إلى صورة رقمية ثم تقسيمه إلى إطارات و استخلاص السمات المميزة له Mel-Frequency Cepstral Coefficients . و في المرحلة الثانية : نستخدم السمات المميزة للصوت في التمييز على الحروف الصوتية ، وينتج عن هذه عدة سلاسل محتملة من الحروف الصوتية و في المرحلة الثالثة : يتم تحويل السلاسل الفونيمية إلى كلمات وهذا بإستخدام الأدوات ماركوف المخفية (HTK) لأنها تتميز بأنها فعالة و مرنة من ناحية حرية اختيار و واسعة الاستخدام في بحوث التمييز للكلام وثرية بالأدوات وتحتوي علي كل الخيارات المطلوبة لتطوير أنظمة تمييز الكلام و قد قمنا بتحقيق هذا النظام بواسطة بعض البرامج المستعملة في انجاز نظام التمييز الآلي للكلام وهي:

1- مجموعة أدوات HTK

مجموعة أدوات HTK	
HLEd	Édition des fichiers d'étiquettes
HHEd	Édition des modèles
HCopy	Calcul des paramètres du signal
HInit	Initialisation d'un modèle
HRest	Ré-estimation d'un modèle
HERest	Ré-estimation des modèles enchaînés
HVite	Décodage en parole continue
HResults	Résultats du décodage
HList	Affichage des fichiers de données
HLStats	Calcul de statistiques sur les étiquettes
HSLab	Affichage du signal et des étiquettes

الجدول 1.4: مجموعة أدوات HTK

2- البرنامج (Speech Filing System) SFS

اختيار برنامج مستخدم حديثا في مجال تجهيز قواعد البيانات ، حيث تم الإعتماد عليه مؤخرا ، وذلك في خطة تطوير الأدوات المستخدمة في بناء قواعد البيانات الصوتية ، وبالنسبة لهذا البرنامج فقد أظهر كفاءة ملحوظة في مجال تحديد بدايات ونهايات المقاطع الصوتية تحديدا دقيقا ، وهو برنامج مصمم خصيصا لتحليل الإشارة الصوتية.

ويعد هذا البرنامج من أفضل البرنامج الصوتية ، لتعدد الخدمات الصوتية التي يقدمها ، ويعطي هذا البرنامج صورة دقيقة لمستويات الصوت من حيث القوة والشدة ، و الشكل الموجه والرسم الطيفي و البياني لها

وتم اختيار البرنامج للأسباب التالية :

1- الجمع بين الدراستين الإهتزازية والطيفية .

2- الدقة العلمية المطلوبة في مثل هذه الأبحاث العلمية .

3- هامش الإرتياب ضئيل في برنامج SFS ، مما يقلل من الأخطاء ويزيد من الدقة العلمية .

3- لغة البرمجة perl

2.الحزمة الفونية : هي قائمة الحروف الصوتية التي تكون نطق جميع كلمات اللغة ، يعرف الحرف الصوتي لكلمة معينة على أنه أصغر وحدة صوتية حيث يتم من خلاله الحصول على كلمة مختلفة عند استبداله بحرف أخر ، كما يعتبر الحرف الصوتي هو الوحدة الأساسية المستخدمة في التعرف الآلي على الكلام ، حيث يوضح الجدول (2.4).

الفونيم	الترميز المستخدم في النظام	الفونيم	الترميز المستخدم في النظام
الفتحة	AE	س(Seen)	S
الألف الممدودة	AE	ش(Sheen)	SH
الضمة	UH	ص(Sad)	SS
الضمة الممدودة	UW	ض(Dad)	DD
الكسرة(Kasra)	IH	ط(Tah)	TT
الكسرة الممدودة	IY	ظ(Thah)	DH2
همزة ء(Hamza)	E	ع(Ain)	AI
ب(Beh)	B	غ(Ghain)	GH
ت(The)	T	ف(Feh)	F
ث(Theh)	TH	ق(Qaf)	Q
ج(Jeem)	JH	ك(kaf)	K
ح(Hah)	HH	ل(lam)	L
خ(Khah)	KH	م(Meem)	M
د(Dal)	D	ن(Noon)	N
ذ(Thal)	DH	ه(Heh)	H
ر(Reh)	R	و(Waw)	W
ز(Zain)	Z	ي(Yeh)	Y

الجدول 2.4: الحروف الصوتية العربية الكاملة مع ترميزها المستخدم في النظام[1]

[1] مصطفى الشافعي ، مجلة التعرف الآلي على الكلام العربي ، التطورات و التحديات ، المؤتمر السادس للغة العربية ، ص123

3 . انجاز قاعدة البيانات:

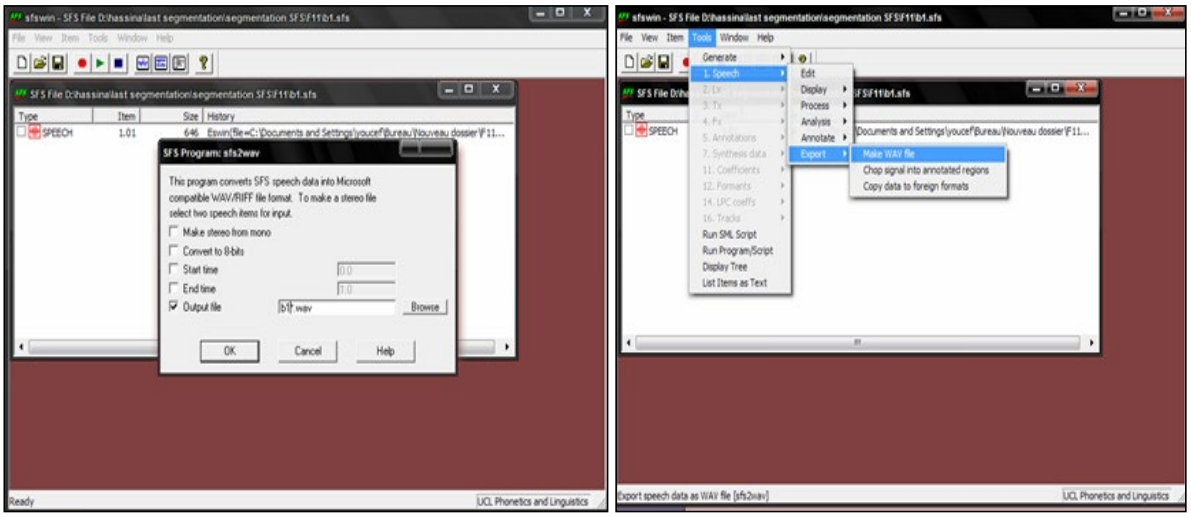
حيث نقوم بإنجاز مدونة تتكون من مجموعة من الجمل باللغة العربية تحتوي على الفونيمات في مواضع مختلفة (بداية و وسط و نهائية الكلمة) حتى يكون تمييز على الفونيمات في سياقات حقيقية مختلفة .

4. تهيئة قاعدة البيانات الأصوات باللغة العربية:

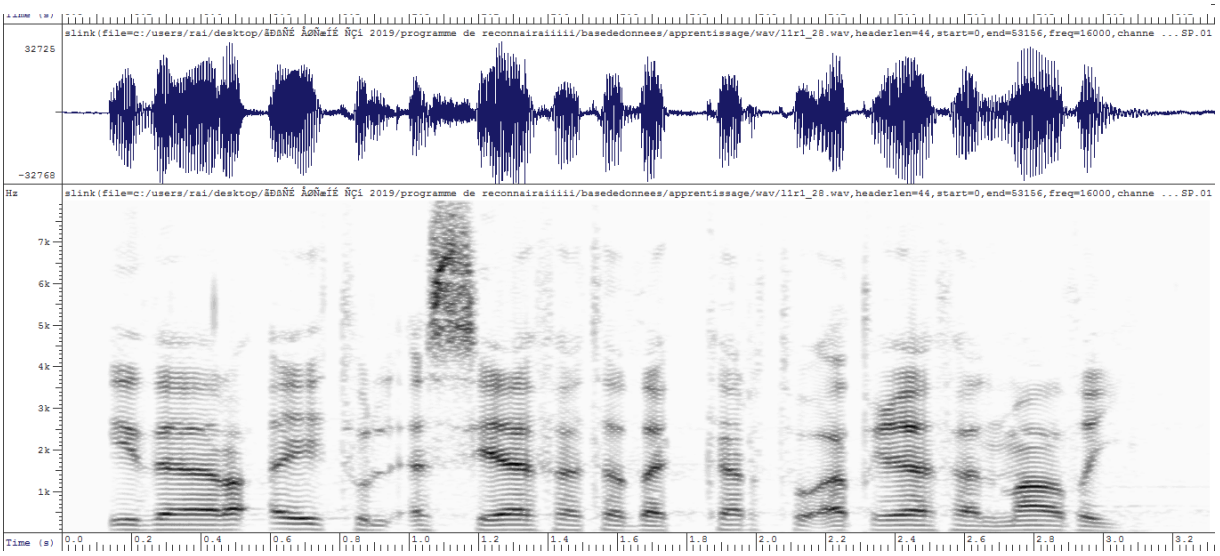
جرى التسجيل مع مراعاة شروط وسط تسجيل جيد وتحديد خصائص التسجيل التالية :

المتكلمون : هم ثلاثة أشخاص ذكرين وأنثى تتراوح أعمارهم بين (20-26) عاما ، قام كل واحد منهم بتسجيل ما يقارب مجموعة مؤلفة من 80 جملة تم تكرار تسجيلها 5 مرات ، أي لدينا 1200 تسجيل، استخدام نصفها للتدريب و نصفها للإحتبار.

عند إصدار الكلمة أمام لاقط الصوت تتغير الموجات الصوتية إلى موجات تناظرية ثم تقوم بطاقة الصوت card Soundبتحويلها إلى عينات رقمية ، في هذا البحث تم تسجيل الكلمة بمعدل عينات يساوي 11025HZ وتمثيل كل عينة بـ 16 bit واستخدام الأسلوب الأحادي Mono ، لذلك فانه عند تسجيل إشارة الكلام يتم تقطيع هذه الإشارة ضمن بطاقة تحصيل الصوت بتردد تقطيع مناسب، لذا علينا الانتباه إلى خصائص تسجيل هذه الإشارة و التي تتمثل في تحديد طريقة الترميز، لتسجيل خزنت هذه العينات الصوتية داخل ملفات بصيغة (Wav) وهي من الصيغ المستخدمة بشكل واسع في بيئة النوافذ (Windows) .



الشكل 1.4: تسجيل و معاينة الأصوات باستعمال برنامج (Speech Filing System) SFS

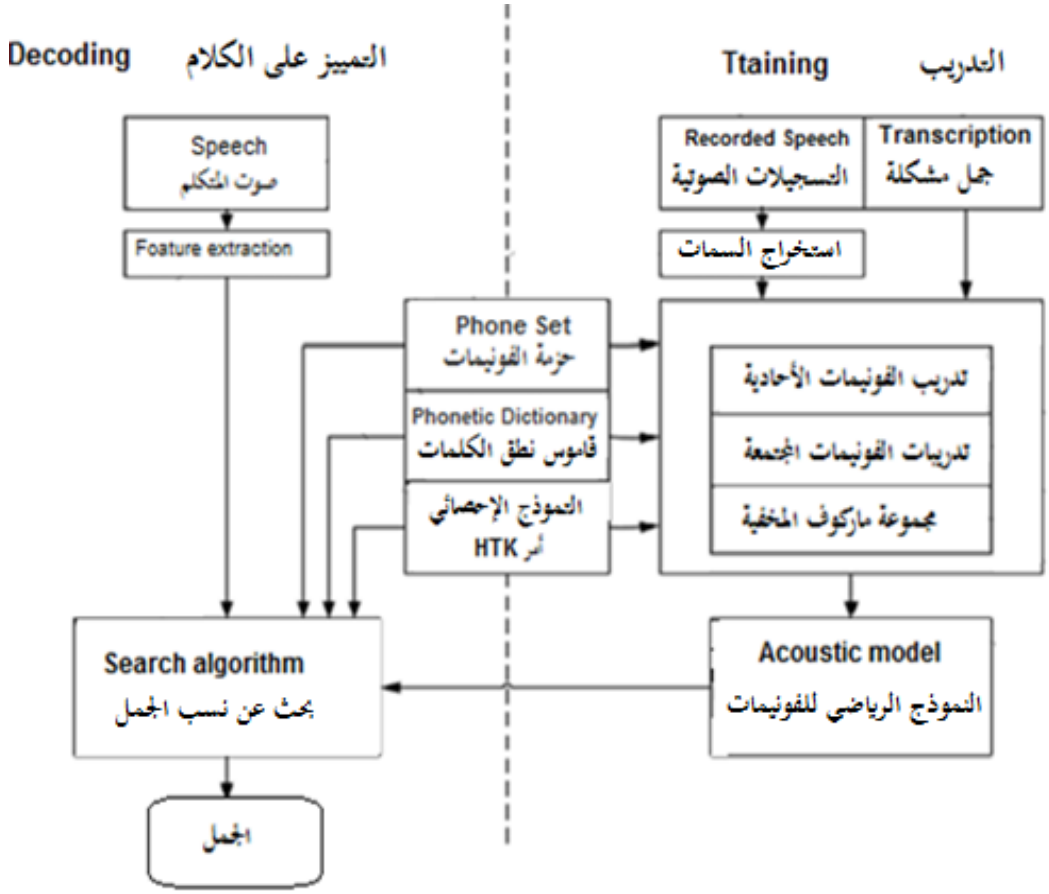


الشكل 2.4: الإشارة الصوتية مسجلة بصوت رجل عند تلفظ الجملة

" إِذَا رَأَيْتُ الْأَسْيَافَ تَذَكَّرتُ أَطْيَافَ الْمَاضِي"

الجمل الأصلية المراد تمييز عليها
أَرْذَالُ الّناسِ ذَوُو أَفْئِدَةٍ مَيّتةٍ
يَقْسِمُ الْوَهَّابُ أَرْزَاقَ الْعِبَادِ
إِذَا رَأَيْتُ ٱلأَسْيَافَ تَذَكَّرتُ أَطْيَافَ ٱلمَاضِي
الرَّسُولُ أَطْهَرُ الْبَشَرِ وَأَزَكَاهُمْ
ٱلأَعْشَى الشَّاعِرُ مِنْ أَدْهَى الشُّعَرَاءِ
سَأَلَهُ ٱلفقِيرُ فَمَا أَثْغَى شَيْئًا مِنْ بُخْلِهِ
قَسَمْتُ الْأُضْحِيَةَ إِلَى ثَلاَثةِ أَثْلاثٍ وَ أَكَلْتُ فِلْذَةً مِنْهَا
أَخَذْتُ فِي ٱلْبَحْثِ عَنْ أَقْوَالِ ٱلْعُلَمَاءِ
مِنْ أَضْأَلِ الطَّعَامِ رَغِيْفَ ٱلقمْحِ
إِلَزَامُ النَّاسِ مَالَا يَلْزَمُهُمْ إِجْحَافٌ بِحَقِّهِمْ
رَأَى النَّائِمُ أَضْغَاثَ أَحْلامٍ
الْبُطْءُ فِي اتِّخَاذِ الرَّأْيِ لَيسَ عَجْزًا
الْعَفَافُ عِتْقٌ لِلنَّفْسِ مِنَ الْعِشْقِ
بَكَى الْوَلَدُ خَوَفاً مِنْ لَمْخِ وَالِدِهِ
الْقرْآنُ مِنْ رَبِّكَ مَحْفُوظٌ عَنِ التَّحْرِيفِ
هَذِهِ الصُّورَةُ لَمْ يَمْسَحْهَا إِلَّا عَابِثٌ
الْقَوِيُّ يَطْغَى فَيَظْلِمُ الضَّعِيفَ
يُبْصِرُ الشَّاعِرُ الْجَمَالَ فَيُضْفِي عَلَيهِ مِنْ إِبْدَاعِهِ
الظُّفْرُ مِنَ الْبَشَرِ كَالظِّلْفِ مِنَ الْغَنَمِ

الشكل 3.4: الجمل الأصلية المراد التعرف عليها



الشكل 4.4: مكونات النظام التمييز على الكلام

5. مراحل تصميم برنامج:

1.5. تجهيز الملفات الضرورية التي تساعد على تعرف على الفونيمات الأحادية:

في البداية نقوم بإنشاء نظام تمييز يقوم بتحديد الموارد الضرورية التي نستعملها ، إذ نقوم بتعريف قاعدة البيانات في ملف (corpus) ، كما نقوم بتعريف liste Mots و هي قائمة الكلمات التي تشكل قاعدة البيانات المدونة corpus و نضيف إليها (SENT-END و SENT-START) و كلمات هذه القائمة مرتبة حسب الترتيب الأبجدي.

أَحَدُ
أَحْزَانٍ
أَحْسَسْتُ
أَحْضَرْتُ
أَحْفَظْ
أَحْقَرِ
أَحْلامٍ
أَحَذْتُ
أَخْضَرُ
أَخْمَاسًا
أَدًى
أَذْهَى
.
.
يُصِيبُ
يُفْصِحُ
SENT-END
SENT-START

الشكل 5.4: listeMots قائمة الكلمات التي تشكل قاعدة البيانات corpus

2.5.القاموس الفونيمي:

هو قاموس يحتوى على الكلمات التي تستخدم في التطبيق مع طريقة نطقها ، بالحروف الصوتية ، كما يعتبر القاموس الصوتي أحد أركان أنظمة التمييز الكلام العربي ، وذلك لأنه بمثل الحلقة الوصل بين الكلمات الموجودة في المدونة (البيانات التدريب) و بين الفونيمات الخاصة بكل كلمة .

حيث سيتم كتابة فونيمات كلمات listeMots كتابة صوتية و تحفظ في القاموس Dictionnaire.

```
أَحَدُ E AE HH AE D UH
أَحَطْتُ E AE HH AE TT T UH
أَحْزَانٍ E AE HH Z AE: N IH N
أَحْسَسْتُ E AE HH S AE S T UH
أَحْضَرْتُ E AE HH DD AH R T UH
أَحْفَظْ E AE HH F AE DH
أَحْقَرِ E AE HH Q AA R IH
أَحْلامٍ E AE HH L M IH N
أَحَذْتُ E AE KH AA DH T UH
أَحْضَرُ E AE KH DD AH R UX
أَخْمَاسًا E AE KH M AE: S AE N
أَدَّى E AE D AE
أَذْهَى E AE D H AE.
.
.
يُصِيبُ Y UH SS IY B UH
يُفْصِحُ Y UH F SS IH HH UH
SENT-END
SENT-START
```

الشكل 6.4: القاموس Dictionnaire

3.5. استخلاص الصفات و بناء قاعدة البيانات:

تعد تقنية استخلاص الصفات من إحدى التقنيات ذات التطبيق الواسع في مجالات عديدة حيث تمثل مرحلة تحليل الإشارة لاستخلاص الصفات الصوتية بدقة والتي تمثل المعاملات الصوتية ، حيث يتم معالجة الإشارة من البداية إلى النهاية وذلك بتقسيم الإشارة إلى إطارات متساوية في الطول ثم يتم استخلاص الصفات لكل إطار لتكون النتيجة النهائية متجه من الصفات ، حيث تم استخدام معاملات درجة النغم .

(MFCC) (Mel -Frequency coefficient cestrum)

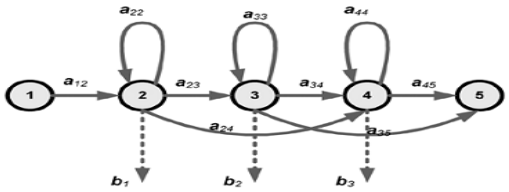
```
Example of an acoustical analysis configuration file #
SOURCEFORMAT = WAV
WINDOWSIZE = 250000.0 # = 25 ms = length of a time frame
TARGETRATE = 100000.0 # = 10 ms = frame periodicity
ENORMALISE=F
USEHAMMING = T # Use of Hamming function for windowing frames
ZMEANSOURCE=T
PREEMCOEF = 0.97 # Pre-emphasis coefficient
NUMCHANS = 26 # Number of filterbank channels
CEPLIFTER = 22 # Length of cepstral liftering
TARGETKIND = _D_A
NUMCEPS = 12
```

الشكل 7.4: ملف التكوين config

وتحسب هذه الضوابط اعتمادا على عملية توفرها وسيلة htk وهي عملية hcopy حيث تأخذ في المدخل ملفا صوتيا وتحسب معاملاتها وفق معطيات تثبيت معينة تخص حجم النوافذ وطبيعتها بالإضافة إلى ضوابط أخرى من وضع المستخدم [1].

ملف التكوين config يسمح بتحديد المعايير الضرورية من أجل مرحلة التحليل الصوتي.

نقوم بتدريب الفونيمات الأحادية : بناء نموذج رياضي لكل حرف صوتي أحادي نستعمل نموذج ماركوف المخفية HMM وهي سلسلة من الحالات ويتم الانتقال من حالة إلى الحالة أخرى عن طريق مصفوفة الاحتمالات ذو 3 حالات من الانتقالات المسموح بها من اليسار إلى اليمين يتم تهيئتها في المصفوفة الانتقالية



الشكل 8.4: نموذج ماركوف المخفي[2]

[1] S. Young, G.Evermann, M. Gales, T.Hain, D. Kershaw, P.Woodland, The HTK Book (for HTK Version 3.3), Cambridge University (2005).

[2] المرجع السابق ، مجلة التعرف الآلي على الكلام العربي ، التطورات و التحديات ، المؤتمر السادس للغة العربية ، ص125

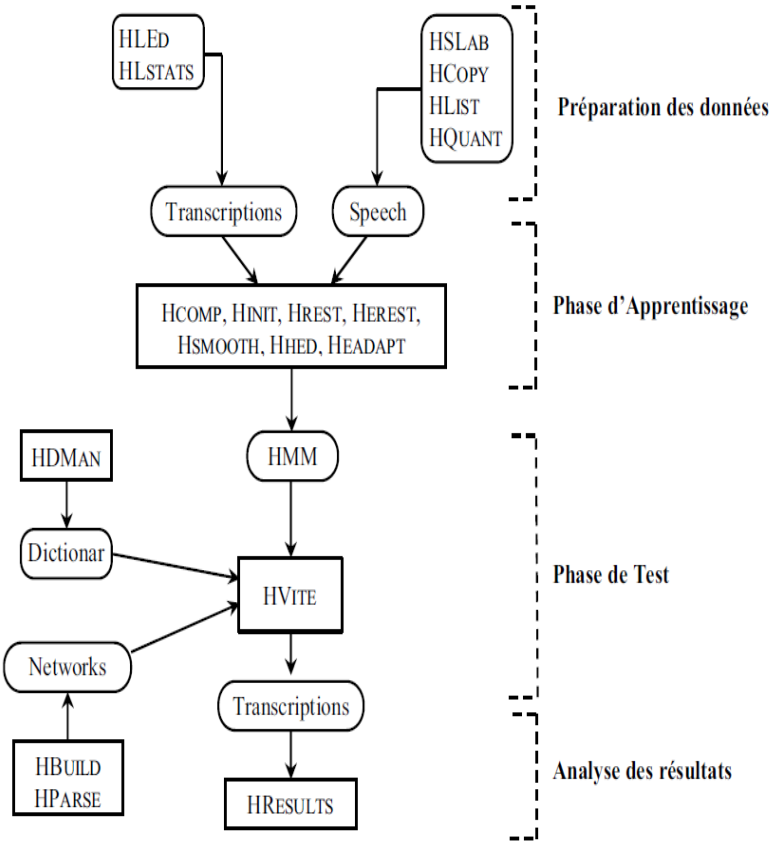
تسمح مرحلة التدريب بتشكيل النماذج المرجعية للنظام حيث أن نوعية هذه النماذج تؤثر بشكل كبير على نتائج نظام تمييز الكلام و يتم تنفيذ التدريب باستعمال أدوات HTK من خلال خطوتين رئيسيتين هما: التهيئة و إعادة التقييم ، و لهذا الغرض فان أكثر الأدوات المستعملة هما الأداتين: HCompV وHERest

إن الدور الذي يلعبه النموذج الأولي هو وصف شكل طوبولوجيا HMM والأرقام الحقيقة المستعملة في التعريف ليست ذات أهمية كبيرة لذلك يمكننا القول بأن حجم العامل VEC SIZE ونوعية الضابط MFCC وجب أن يخصص ويعرفا ، ووجب أيضا اختيار وضبط عدد الأحوال NUMSTATES ووجب ذكر عدد الإنتقالات المتاحة بين الأحوال وذلك بإعطاء قيم مختلفة عن الصفر بالنسبة للعوامل المرافقة للمصفوفة العبور TRAMSPوالأصفار من جهة أخرى .

```
~o <VecSize> 36 <MFCC_D_A> <StreamInfo> 1 36
~ h "proto"
>BeginHMM<
> NumStates> 5
> State> 2 <NumMixes> 1
> Stream> 1
> Mixture> 1 1.0000
>  Mean> 36
 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
>  Variance> 36
 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
> State> 3 <NumMixes> 1
> Stream> 1
> Mixture> 1 1.0000
>  Mean> 36
 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
>  Variance> 36
 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
> State> 4 <NumMixes> 1
> Stream> 1
> Mixture> 1 1.0000
>  Mean> 36
 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
>  Variance> 36
 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
> TransP> 5
0.000  e+0  1.000e+0  0.000e+0  0.000e+0  0.000e+0
0.000  e+0  6.000e-1  4.000e-1  0.000e+0  0.000e+0
0.000  e+0  0.000e+0  6.000e-1  4.000e-1  0.000e+0
0.000  e+0  0.000e+0  0.000e+0  6.000e-1  4.000e-1
0.000  e+0  0.000e+0  0.000e+0  0.000e+0  0.000e+0
>EndHMM<
```

الشكل 9.4: ملف التهيئة (prototype) HMM .

6. تدريب النظام :



الشكل 10.4: بنية نظام للتعرف بواسطة HTK [1] .

عمليا نحتاج إلى أربع خطوات لتدريب نماذج ماركوف المخفية : تحضير المعلومات وتكوين مجموعة ماركوف المخفية و الإختبار والتحليل.

1.6. تحضير المعلومات:

بعد تحضير المعلومات نقوم بتدريب النظام حيث لدينا قاموس يحتوى على الكلمات التي تستخدم في التطبيق مع طريقة نطقها لكل كلمة سيتم تقييمها في النظام مستخدمين في ذلك أمر HDMan ولأن الملفات

[1] S. Young, G.Evermann, M. Gales, T.Hain, D. Kershaw, G. Moore, J.Odell, D.Ollason, D. Povey, V.Valtchev, P.Woodland, The HTK Book (for HTK Version 3.3), Cambridge University (2005).

الصوتية للكلام لا تستخدم مباشرة فهي تحتاج إلى اجتياز عملية تحويل إلى خصائص مؤشرات صوتية مستخدمين فى ذلك أمر HCopy ، ونستخدم هنا MFCC لاستخلاص خصائص المؤشرات الصوتية ، بعدها سيتم تعريف قائمة الفونيمات (ListePhones) التي يتم استنتاجها بواسطة HDMan لتحديد معجم لقاعدة البيانات corpus بشكل قائمة و هي على النحو التالي[1]:

E ; AE:; L; M ; UH ; IH ; N ; Y ; T ; AE ; B ; IY ; TH ; JH ; H ; GH ; AA ; HH ; D ; Z ; F ; DH2 ; Q ; R ; KH ; DD ; AH ; UX ; S ; DH ; SH ; SS ; AH:; AA:; TT ; W ; K ; AI ; UW ; IX ; sil

بعدها نقوم ببناء برنامج باستعمال Perl يقوم بإنشاء ملف Transcription_mots يحتوي اسم الملف الصوتي المخصص للتدريب و الجملة التي تقابله بشكل قائمة من الكلمات.

عند بداية تدريب النظام يجب علينا أن نغذيه بمعلومات خاصة بالوحدات اللغوية الصوتية المتضمنة فى كل ملف صوتي للكلام ، لذلك ننشئ يدوياً ملفاً يحتوي على الكلمات التى تم نطقها فى ملفات الكلام ، وبعد ذلك يتم إنشاء ملف يحتوي على الوحدات الصوتية لهذه الكلمات مستخدمين فى ذلك أمر Hled وأخيراً يستخدم أمر Hparse لإنشاء ملف شبكة الكلمات الذى يحتوي على كل كلمة وكل انتقال من كلمة لأخرى أثناء التمييز المستمر على الكلام.

2.6. مجموعة أدوات HTK :

نقوم بإنشاء مجموعة أدوات HTK متماثلة أحادية الوحدات الصوتية وكل وسط أو تباين داخل هذه النماذج يكون متماثلاً ، ثم يضاف إليها نماذج التوقفات القصيرة و السكون لكشف السكون داخل إشارات الكلام ، ثم يعاد تدريب هذه النماذج وأخيرا يتم إنشاء نماذج لوحدات الصوتية مستخدمين في ذلك النماذج أحادية الوحدات الصوتية السابقة ، و لهذا نحتاج إلى خطوتين:

(1) S. Young, G.Evermann, M. Gales, T.Hain, D. Kershaw, G. Moore, J.Odell, D.Ollason, D. Povey, V.Valtchev, P.Woodland, The HTK Book (for HTK Version 3.3), Cambridge University (2005).

في الخطوة الأولى : مرحلة تهيئة نماذج HMM بواسطة الأداة HCompy

أما في الخطوة الثانية : تقييم نماذج HMM وأدوات التدريب المستخدمة هنا هي:

{ HERest , HCompV , HLed , HVite , HHEd }

3.6. الإختبار:

عند بداية الاختبار نستخدم أداة أمر HVite للتمييز على الكلام الموجود في الملفات الصوتية المدخلة إلى الحاسب مستخدمين فى ذلك نماذج ماركوف المخفية التى أنشأناها قبل ذلك وبعض الملفات التى أوجدناها خلال خطوة تجهيز البيانات ويستخدم أمر HVite لكل من التدريب ، خلال تدريب النظام يستخدم لتحديد مدى دقة وسائط نماذج ماركوف المخفية ، وهذا يساعد فى تحسين كودات و برمجيات التدريب وكشف الأخطاء فيها أثناء تقييم الوحدة الصوتية من أجل حساب نسب تقييمية للوحدة الصوت .

إن الهدف الرئيسي من وجود نموذج احصائي للغة باعتماد على المدونة (HVite) هو توقع حدوث كلمة معينة بعد كلمة أو كلمتين متتاليتين ، و بذلك امكانية تشكيل الجملة التي يراد التعرف عليها ، فعند تحديد المقطع الصوتي للتعرف عليه يبدأ النظام بالتعرف على الكلمة بالاستعانة بالنموذج الصوتي و القاموس الصوتي ، ثم تحديد الكلمة الأقرب في القاموس الصوتي و الحصول على كلمة جديدة مجاورة حيث يتم استخدام النموذج اللغوي لربط الكلمتين ، ثم ربط ثلاث كلمات ، وهكذا تستمر العملية حتى يتم تشكل جملة التي احتوى عليها المقطع الصوتي

إن نطق الكلام ليس عملية مضبوطة لأن نطق الكلمة نفسها قد يتم بطرق مختلفة من شخص إلى آخر أو بالنسبة للشخص نفسه ، مما يؤدي إلى تنوع شكل الإشارة الصوتية الموافقة لها ، و ينتج هذا الاختلاف في النطق بحسب سرعة كلام الشخص و جودة صوته إن كان يقصد الإفهام أو يتكلم على نحو عارض ، كما يمكن أن تكون التغيرات الفيزيولوجية و النفسية للمتكلم سببا في تغير النطق كالفرح و الحزن و الغضب و المرض...و غيرها ، و لجعل نظام تمييز على الكلام فعالا في التعامل مع هذا التنوع في النطق يجب النظر في جميع تغيرات

النطق الممكنة بتنظيم قاعدة التدريب و تحليل كل كلماتها المذكورة في القاموس(Dictionnaire) و اختيار الأقرب و الأنسب منها بالمقارنة مع الإشارة الصوتية المدخلة و يتم إنجاز هذا العمل بواسطة الأداة HVite:

يسمح تكرار خوارزمية باوم ولش المستخدمة بإعادة تقدير النماذج مع مراعاة التحسينات الأخيرة ، ثم يتم تخزين المعاملات الأخيرة الناتجة في مجلد جديد ، بهذا تختتم مرحلة التدريب لنماذج HMM.

4.6. التحليل:

نستخدم أداة أمر HResults لتحليل النتائج لقراءة الملف الناتج من أداة HVite وهذه المرحلة يتم البحث عن النماذج الفونيمات عن أحسن للحروف للصوت المنطوق وسيعمل القاموس لتحديد نسب التعرف لكل وحدة صوتية واردة في الكلمة ويدخل هنا نموذج الأحسن ،كما نعتمد على خوارزمية باوم ولش في إعادة تقدير النماذج من حيث عدد تكرارها ثلاث مرات و هذا يترجم قدرة و فعالية نماذج ماركوف المخفية ، الذي يكمن أساسا في أنه يقوم بتحسين نماذج HMM و يتوقف على إعادة تقدير معاملاتها (المتوسط و التباين). النماذج الناتجة سيتم حفظها في الملف hmm3 و هذا بإعادة تقدير نماذج HMM الواردة في الملف HMMn و حفظها في الملف ذو الترتيب الموالي أي HMMn+1.

7. تجارب و نتائج:

يمكن لنظامنا إجراء تصحيح للنطق إذا كان يحتوي على أي خطأ ، كما يمكنه الحصول على نسب تعرف على كل وحدة صوتية واردة في الكلمة.

لغرض توفير إمكانية تحقيق المهام الرئيسي لانجاز لنظام تمييز على الكلام العربي هناك العديد من المتغيرات التي يمكن أن تؤثر على نسبة تمييز و جودتها ، و للحصول على أحسن نتائج ممكنة بأقل قدر من العمليات سنقوم بعدة تجارب ثم نلاحظ من خلالها تغيرات نتائج تمييز و التي هي: **تغير حجم قاعدة البيانات الصوتية و تكرار تسجيل جمل المدونة و نوع المعاملات المستخدمة بالإضافة إلى تكرار التهيئة بـ HERest.**

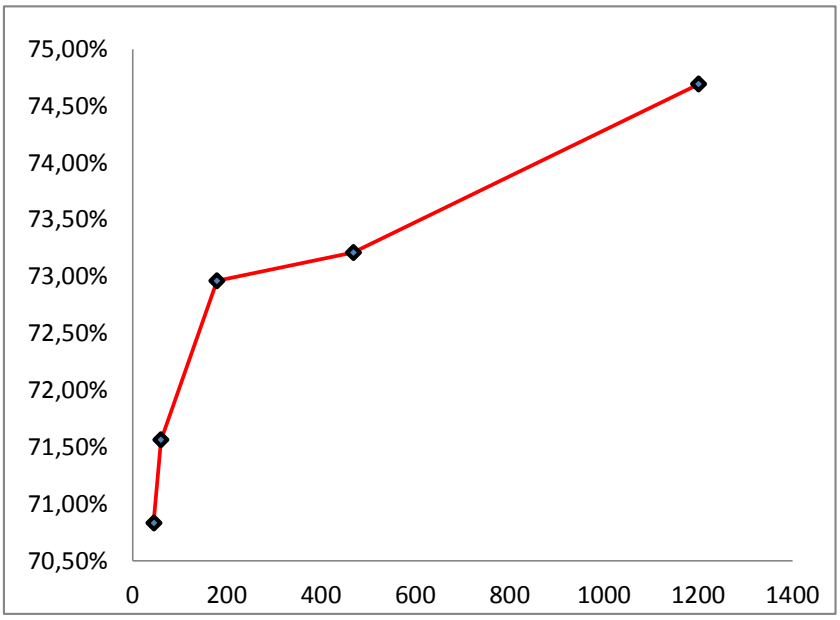
1.7. نسب التمييز المئوية بدلالة حجم قاعدة البيانات الصوتية :

يظهر لنا المنحنى التالي نسب التمييز المئوية بدلالة حجم قاعدة البيانات الصوتية ، حيث إن الاختبار تم على مجموعة مؤلفة من 80 جملة تم تكرار تسجيلها 5 مرات من طرف 3 اشخاص اثنان من الذكور و انثى أي لدينا 1200 تسجيل، استخدام نصفها للتدريب و نصفها للإحتبار.

ولقد قمنا بتحسين دقة البرنامج بزيادة حجم البيانات الصوتية ثم بتكيف البرنامج مع المتحدث كما هو مبين في الجدول لتالي :

المرحلة	نسب التمييز النطق
استعمال قاعدة بيانات صغيرة	70,83%
استعمال قاعدة بيانات كبيرة	73,96%
زيادة قاعدة البيانات الجديدة (1200) جملة	74,69%

الجدول 3.4: النسب تمييز المئوية بدلالة حجم قاعدة البيانات الصوتية



الشكل 11.4: منحنى تغير نسبة التمييز بدلالة حجم قاعدة البيانات الصوتية

نستنتج من المنحنى المبين أن حجم البيانات الأصوات المستعملة خلال عملية تدريب النظام له دور مهم

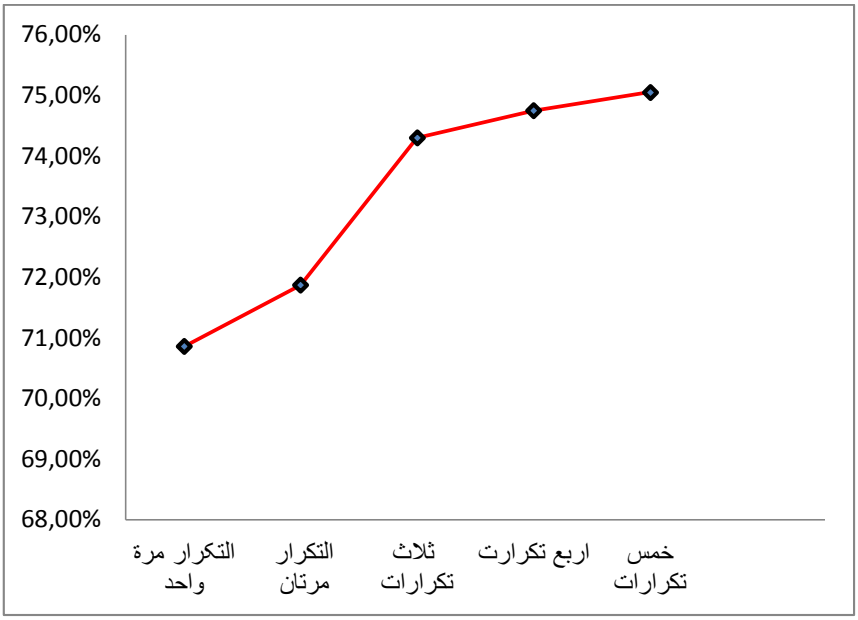
جدا في رفع نسب تمييز البرنامج على تقييم النطق و هذا يفرض علينا إعطاء أهمية كبيرة للمدونة من حيث الحجم و نوعية التسجيل ، وزيادة حجم قاعدة البيانات الصوتية الذي له تأثير فعال بزيادة دقة تمييز على الرغم من أن تأثيرها يكون واضحاً بزيادة الزمن المطلوب لإتمام عملية التمييز، وهذا ناتج إيجابي فمن المتوقع أن تكون هناك زيادة في نسبة تمييز مصاحبة للزيادة في حجم الملف مما يعطي دفعاً إيجابياً على النتائج العملية التي تم الحصول عليها عند تطبيق الخوارزمية .

كما أتضح من نتائج تطبيق الخوارزمية على عدد من الأمثلة المختلفة أن الخوارزمية لا تتأثر بصفة المتكلم أو نوع الكلام وصيغته المعتمدة في التطبيقات العملية للخوارزمية ، إذ انتخبت نماذج مختلفة من الأصوات لرجل وامرأة وطفل فضلا عن اختلاف أساليب إلقاء الكلام التي تباينت بين الأسلوب الأصوات واختلاف نبرات الصوت ومخارج نطق الحروف ، أو صيغته وإنما أعطى نتائج متقاربة من بعضها البعض ، لأنه بعد عملية التدريب سنتحصل على نماذج سهلة الاستخدام عند الاختبار حيث لا يظهر لحجم مدونة التدريب اثر في مدة تمييز.

2.7. نسب التمييز المئوية بدلالة تكرار التهيئة بـ HERest:

خوارزمية (Viterbi) هي طريقة البرمجة الحركية المستخدمة لحساب مسار الحالة الانتقالية الأكثر احتمالا [1]، و تستند هذه الخوارزمية على الاحتمال الذي يحسب كل مرة بين الكلمة المراد اختبارها و نموذجها مما يعطي أهمية كبيرة لحساب النماذج و إعادة تهيئتها و الذي يتم تنفيذه باستخدام الأمر HERest في مجموعة الأدوات HTK ، وتظهر نتائج تقدير تمييز الكلام المتحصل عليها بدلالة تغير تكرار التهيئة بـ HERest في المنحنى :

[1] Anil, K. J., Robert, P. W. & Jianchang, M. 2000, 'Statistical Pattern Recognition: A Review', IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1.



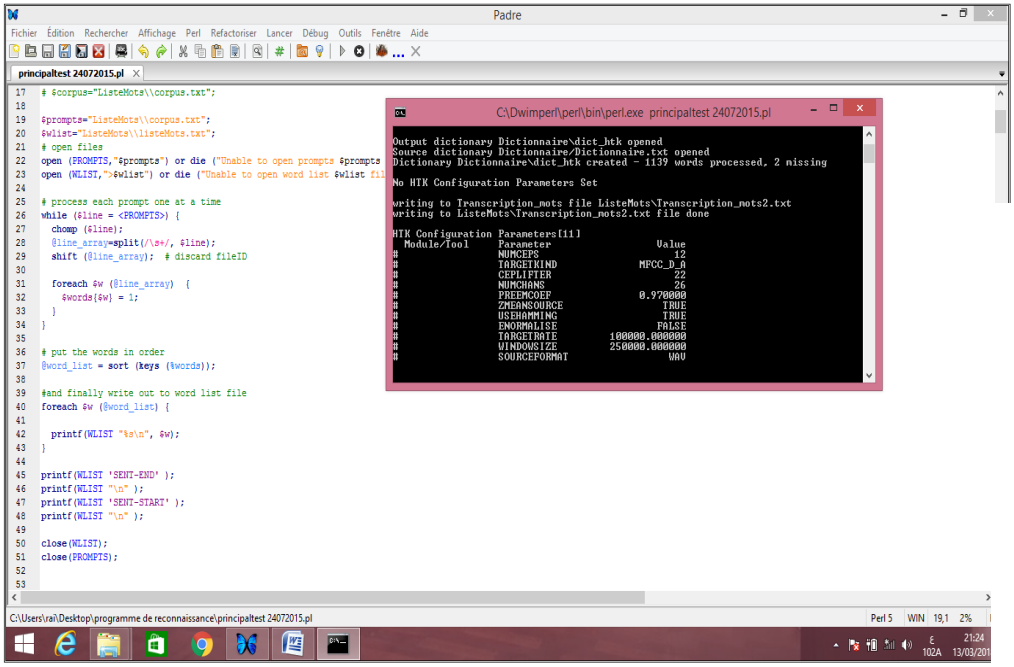
الشكل 12.4: منحنى تغير نسبة التمييز بدلالة تغير عدد تكرارت إعادة تقدير

و بما أن تكرار عملية إعادة تقدير النماذج تأخذ زمنا لا يستهان به في مرحلة التدريب، و نظرا لأن نسبة التمييز ابتداء من التكرار الثالث متقاربة ، إذا يمكن الاكتفاء بتكرار عملية إعادة تقدير النماذج 3 مرات فقط.

و تبين النتائج الموضحة أعلاه أن نظام التمييز المقترح الذي يعتمد على خوارزمية بوم ولش في إعادة تقدير النماذج ذو كفاءة عالية من حيث عدد الكلمات التي تم تقييمها بشكل صحيح و هذا يترجم قدرة و فعالية نماذج ماركوف المخفية ، الذي يكمن أساسا في أنه يقدم خوارزميات قوية للتدريب.

و بعض المشاكل التي واجهتنا هى تشكيل الكلمات لتى تلعب دورا أساسيا في عملية تمييز الكلام ، فإن التشكيل يساعد في تقليل الغموض أثناء عملية تمييز الكلمة ، التي يجب أن تأخذ بعين الاعتبار عند تحديد كيفية نطق الكلمات في القاموس الصوتي ، لأنها ألفاظ مختلفة يجب أن يفرق بينهما في الكتابة الصوتية للفونيميات في القاموس الصوتي ، وبالتالي فإن قضية التشكيل تعتبر أساسية في أنظمة تمييز الكلام العربي ، ومن المشاكل التي واجهتنا الأخطاء في التشكيل ، حروف بدون تشكيل ، حروف لها أكثر من تشكيل ،آداة التعريف قد تشكل كالأتي (ال،الْ ، أل ، أَل ،أُل) و التنوين .

و نلاحظ أن نسبة تمييز تسجيلات مجموعة التدريب كانت 74.69 % أما نسبة تمييز على مجموعة تسجيلات خارج مجموعة التدريب كانت 69.81 % حيث أظهرت النتائج أن نسبة تمييز الأكبر كانت لتسجيلات المتكلم الذي كان ضمن مدونة التدريب ، و هذا يعني أن النظام يعمل بشكل جيد لمتحدث ضمن مدونة التدريب كما أظهرت النتائج التالية :



الشكل 13.4: احدى واجهات التنفيذ بلغة Perl

Aligned transcription: l1r5_24.lab vs l1r5_24.rec

LAB يَقْسِمُ الْوَهَّابُ أَرْزَاقَ الْعِبَادِ

REC تَكْسِرُ الْوَهَّابُ أَرْزَاقَ الْعِبَادِ

Aligned transcription: l1r5_23.lab vs l1r5_23.rec

LAB أَرْذَالُ النّاسِ ذَوُو أَفْئِدَةٍ مَيّتَةٍ

REC الرِّجَالِ النَّاسِ ذَوُو أَفْئِدَةٍ عِزَّةَ

Aligned transcription: l1r5_28.lab vs l1r5_28.rec

LAB إِذَا رَأَيْتُ الْأَسْيَافَ تَذَكَّرْتُ أَطْيَافَ الْمَاضِي

REC إِلَّا رَأَيْتُ الْأَسْيَافَ تَذَكَّرْتُ أَطْيَافَ الْمَاضِي

Aligned transcription: l1r5_3.lab vs l1r5_3.rec

LAB ٱلْأَعْشَى الشَّاعِرُ مِنْ أَدْهَى الشُّعَرَاءِ

REC الْجَأْشِ الشَّاعِرُ مِنْ إِذَا الشِّعْرَ

Aligned transcription: l1r5_30.lab vs l1r5_30.rec

LAB سَأَلَهُ ٱلْفَقِيرُ فَمَا أَنْغَى شَيْئًا مِنْ بُخْلِهِ

REC السَّأَّالُ الْفَقِيرِ فَمَا أَنْغَى شَيْئًا مِنْ بُخْلِهِ

Aligned transcription: l1r5_32.lab vs l1r5_32.rec

LAB قَسَمْتُ الْأُضْحِيَةَ إِلَى ثَلاَثَةِ أَثْلَاثٍ وَ أَكَلْتُ فِلْذَةً مِنْهَا

REC سَمِيكٌ الْأُضْحِيَةَ إِلَّا ثَلاَثَةِ أَثْلَاثٍ وَ أَكَلْتُ فِلْذَةً مِنْهَا

Aligned transcription: l1r5_37.lab vs l1r5_37.rec

LAB أَخَذْتُ فِي ٱلْبَحْثِ عَنْ أَقْوَالِ ٱلْعُلَمَاءِ

REC أَكَلْتُ فِي ٱلْبَحْثِ عَنْ أَقْوَالِ ٱلْعُلَمَاءِ

Aligned transcription: l1r5_4.lab vs l1r5_4.rec

LAB مِنْ أَضْأَلِ الطَّعَامِ رَغِيفَ ٱلْقَمْحِ

REC إِنْ أَفْضَلِ الطَّعَامِ الرَّغِيفِ وَقَدْ نَحْذُو

Aligned transcription: l1r5_42.lab vs l1r5_42.rec

LAB إِلْزَامُ النَّاسِ مَالَا يَلْزَمُهُمْ إِجْحَافٌ بِحَقِّهِمْ

REC إِنْسَانِ النَّاسِ مَالَا يَلْزَمُهُمْ إِجْحَافٌ بِحَقِّهِمْ

Aligned transcription: l1r5_61.lab vs l1r5_61.rec

LAB الْبُطْءُ فِي اتِّخَاذِ الرَّأْيِ لَيسَ عَجْزًا

REC ٱلْبُلْبُلُ فِي اتِّخَاذِ الرَّأْيِ لَيسَ عَجْزًا لَا

Aligned transcription: l1r5_75.lab vs l1r5_75.rec

LAB الْعَفَافُ عِتْقٌ لِلنَّفْسِ مِنَ الْعِشْقِ

REC الْعَطْسِ حَتْوَةٌ لِلنَّفْسِ مِنَ الْعِشْقِ

Aligned transcription: l2r5_352.lab vs l2r5_352.rec

LAB بَكَى الْوَلَدُ خَوَفاً مِنْ لَمْخِ وَالِدِهِ

REC لَكَ الْوَلَدُ خَوَفاً مِنْ لَمْخِ وَالِدِهِ

Aligned transcription: l2r5_367.lab vs l2r5_367.rec

LAB الْقُرْآنُ مِنْ رَبِّكَ مَحْفُوظٌ عَنِ التَّحْرِيفِ

REC الْقَدَمُ رَبِّكَ مَحْفُوظٌ عَنِ التَّحْرِيفِ

Aligned transcription: l3r5_323.lab vs l3r5_323.rec

LAB هَذِهِ الصُّورَةُ لَمْ يَمْسَخْهَا إِلَّا عَابِثٌ

REC يَهْذِي الصُّورَةُ لَمْ يَمْسَخْهَا إِلَّا عَابِثٌ

Aligned transcription: l3r5_333.lab vs l3r5_333.rec

LAB الْقَوِيُّ يَطْغَى فَيَظْلِمُ الضَّعِيفَ

REC الْقَاضِي يَطْغَى فَيَظْلِمُ الضَّعِيفَ

Aligned transcription: l3r5_339.lab vs l3r5_339.rec

LAB يُبْصِرُ الشَّاعِرُ الْجَمَالَ فَيُضْفِي عَلَيهِ مِنْ إِبْدَاعِهِ

REC يُفْصِحُ الشَّاعِرُ الْجَمَالَ فَيُضْفِي عَلَيهِ مِنْ إِبْدَاعِهِ

Aligned transcription: l3r5_5.lab vs l4r5_5.rec

LAB رَأَى النَّائِمُ أَضْغَاثَ أَحْلامٍ

REC خَضْرَاءَ النَّائِمُ أَضْغَاثَ أَحْزَانٍ

==========HTK Results Analysis==========

Date: Tue Mar 13 19:32:19 2018

Ref : ListeMots\Transcription_mots.

----------Overall Results----------

PHONE: %Corr=74.69,

الشكل 14.4: نموذج حول النتيجة النهائية لنظام تمييز الآلي للكلام العربي بالاعتماد على الفونيمات الأحادية

8. الخاتمة:

لقد قدمنا خلال هذا البحث باستخدام نماذج ماركوف المخفية في بناء نظام للتمييز الآلي للكلام العربي بالإعتماد على الفونيمات الأحادية قائمة على مجموعة أدوات البنية البرمجية HTK ، وذلك باعتماد على سمات مميزة للإشارة تم الحصول عليها باستخدام تقنية MFCC ، و اقترحناها لتستخدم كبديل للتقييم البشري للنطق وهذا بمساعدة مجموعة مختلفة من المتكلمين ، فوجدنا أن النظام يعمل بكفاءة أكبر مع متكلم معتمد ضمن مدونة التدريب ، أكثر من متكلم خارج مدونة التدريب ، و قد كانت النتائج التي تم الحصول عليها مشجعة ، بالنظر إلى العقبات و المشاكل الكبيرة المتعلقة باستخدام أنظمة تمييز الآلي للكلام العربي .

ولقد قمنا بتحسين دقة البرنامج بزيادة حجم البيانات الصوتية ، وقد بدأنا العمل على توسيع قاعدة بيانات الأصوات لأنها تبين مما سبق أن حجم بيانات الأصوات المستعملة و نوع المعاملات الصوتية لها تأثير كبير على عملية تعرف 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 90 % بعد أن استعمل هذا الحجم الكبير من البيانات الصوتية خلال تدريب البرنامج.

الخاتـمـة العامة

الخاتمة

تعد تقنيات التعرف على الكلام من أهم التقنيات الحديثة ، وقد تم تطوير العديد من أنظمة المختلفة من حيث الطرائق المستخدمة في استخراج السمات وطرائق التصنيف ، لكن مهما كانت الخوارزمية المستخدمة أو طريقة التصنيف فإن في تكنولوجيا معالجة الكلام للتطبيقات الحقيقية يمكن للعديد من الحقائق أن تشوه أو تتلف الكلام ، مما يجعل الكلام لا يبدو بالطريقة التي تم تسجيله فيها.

في هذا العمل تم إنشاء نظام التمييز الآلي للكلام العربي بالاعتماد عمى السمات الصوتية المستخرجة من فيديوهات منطوقة لكلمات باللغة العربية ، وتحديد المرشح الأمثلي في مرحلة استخراج السمات لخوارزمية معاملات تردد ميل لزيادة نسبة التمييز. تم استخدام نماذج ماركوف المخفية في مرحلة التصنيف وتحديد عدد الحالات الأمثلي في المصنف.

من خلال البحث عن أفضل تطابق وحيد بين مجموعة من مفردات قاعدة البيانات الصوتية للغة العربية ، حيث وجد نتيجة للأبحاث المتعددة أن هذه النماذج تعطي نتائج جيدة ، و للحصول على أفضل النماذج التي تساعد في تحسين أداء النظام المقترح للتمييز الآلي للكلام العربي قمنا بإجراء عدة تجارب على مجموعة من المتغيرات تتمثل في تغير حجم المدونة و عدد مرات تكرار تهيئة النماذج.

ويمكننا من خلال العمل الحالي استنتاج ما يلي:

1. مجموعة أدوات HTK قادرة على التمييز بسرعة محسوسة و قدرة أداء عالية كما يبين من خلال متابعة أداء النظام في التمييز.
2. زيادة في نسبة التمييز ناتج عن زيادة حجم قاعدة بيانات الأصوات لأنها تبين مما سبق أن حجم بيانات الأصوات المستعملة خلال عملية تدريب النظام لها دور مهم جدا في رفع دقة البرنامج على تقييم النطق .

3. من أجل حل المشاكل و تسهيل خطوات المعالجة و تقليلها يتم استخلاص الصفات الأساسية للإشارة الصوتية و بناء قاعدة البيانات من معاملات (MFCC) مما يؤدي إلى تقليص حجم البيانات و إعطائها أفضل تمثيل ممكن لإشارة وهذت يقود إلى دقة أعلى في أداء نظام التمييز.

4. الاختلاف بين المتكلمين يؤثر على دقة تمييز الكلام ، حيث أن النظام يعمل بشكل جيد مع متكلم معتمد ضمن قاعدة بيانات التدريب ، أكثر من متكلم خارج مدونة التدريب.

5. تشكيل الكلمات يلعب دورا أساسيا في عملية تمييز الكلام ، فإن التشكيل يساعد في تقليل الغموض أثناء عملية التمييز على الكلمة ، التي يجب أن تأخد بعين الإعتبار عند تحديد كيفية نطق الكلمات في القاموس الصوتي ، وبالتالي فإن قضية التشكيل تعتبر أساسية في أنظمة تمييز الكلام العربي .

و قد أظهرت نتائج التمييز المختلفة التي تم التوصل إليها أن النظام كان قادراً على تمييز الكلمات و بنسبة 74.69%.

وتعتبر هذه النتائج التي تم الحصول عليها مشجعة و مع ذلك فإن تنفيذ نظام للتمييز الآلي للكلام العربي بالاعتماد على الفونيمات الأحادية في ظل ظروف واقعية يتأثر بالعديد من المشاكل مثل الضوضاء ، و نوعية التسجيل ، وصعوبة المفردات وحالة المتكلم ، حيث تبقى مهمة صعبة لا تزال تحتاج إلى العديد من البحوث للتعامل مع جميع جوانب التخاطب بين الإنسان و الآلة.

التوصيات:

- الدمج بين نموذج ماركوف المخفي و الطرق الأخرى للتخلص من المساوئ ، و تحسين الأداء و النتائج.
- إنجاز مدونة (قاعدة بيانات) تحتوي مجموعة كلمات أحد معاجم اللغة العربية وتسجيلها بواسطة متكلمين مختلفين للمساعدة على تحسين أداء أنظمة التعرف الآلي على الكلام و مساعدة الباحثين في القيام بدراساتهم و أبحاثهم.
- إن ارتباط النماذج الإحصائية بالحاسوب هو ما يمنحها خصوبتها وفي نفس الوقت طابعها المؤقت ، فظهور نوع جديد من الحواسيب قد يلغي كل النماذج الموجودة حاليا ، وهنا بالضبط تكمن أهمية النماذج ، وغناها ، وتطويرها للأبحاث ، لأنها ترتبط دائما بالتطور التكنولوجي لكل مرحلة .
- تؤول العربية في نظامها الصوتي إلى أربعة وثلاثين فونيما ثمانية وعشرون صامتا تتمثل في حروب الأبجدية العربية وستة صوائت تمثلها الحركات الثلاث القصيرة الفتحة والكسرة والضمة والحركات الثلاث الطويلة حروف المد ويعنى في الجانب الفوناتيكي بتمييز تلك الأصوات على المستوى الأكوستيكي الفيزيائي...والبادي أن تمييز الأصوات منفردة على هذا النحو سيفضي إلى نتائج دقيقة لا تقبل اللبس وأن حدوداً فاصلة قاطعة ...ستميز كل صوت عن بقية الأصوات.
- ومما لا شك فيه أن التأخر العلمي العربي قد انعكس سلبا على جودة التطبيقات اللغوية العربية التي توظف ثورة تكنولوجيا المعلومات لخدمة الإنسان ، فمن المعلوم أن الحاجة الملحة لإنتاج تطبيقات صوتية علمية على حد سواء قد أثرى علم الصوتيات في الدول المتقدمة بشكل كبير ، الأمر الذي لم يتحقق في البلاد العربية فندرة التطبيقات العلمية وقلة التطبيقات التجارية التي تعتمد على تقنية التعرف على الكلام أدت إلى ضعف التطور في هذا المجال ، وبذلك فإنه يمكن القول إن البحث العلمي في مجال الصوتيات العربية لا زال محدودا مقارنة مع اللغات الأخرى.

قائمة المصطلحات

قائمة المصطلحات (عربية /إنجليزية)

الصوتيات الأكوستيكية (الفيزيائية)	Acoustic
اتساع الموجة / اتساع الذبذبة	Amplitude
علم الأصوات الأكوستيكي	Acoustic phonetics
الأكوستيكي	Acoustic
اكتساب لغة الإشارة	Acquisition of sign language
ألوفون	Allophone
قياس	Analogy
علم اللغة	Anthropologicol linguistics
المعالجة الرياضية	Arithmetic processing
علم الأصوات النطقي	Articulatory phonetics
جهاز النطق	Articulatory système /Vocal Apparatus
التعرف الآلي على الكلام	Automatic speech recognition
آندري ماركوف	A.Markov
مجال السمع	Auditory range
الاحساس السمعي	Auditory sensation
عرض الحزمة	Band width
توليف الموجات المجزاة	Concatenated speech synthesis
ملف التكوين	config
تجويف الحلق	Cavity Pharyngeal
صوت صامت	Consonant

Drill	تدريب /تمرين
Decibel	ديسيبل (وحدة قياس شدة الصوت)
Duration	مدة (مدة تردد الموجة الصوتية)
External Ear	الأذن الخارجية
Formants	حزمنة صوتية
Fourier Transform	تحويل فورييه
Framing and Windowing	التقطيع إلى إطارات و النوافذ
Framing	التأطير
Filter pass high	مرشح تردد عالي
Filter Bank	المرشحات
Fréquency	التردد أو التواتر
Fonem Segmental	فونيمات رئيسية تركيبية أو القطعية
Fourier transform	لتحويل فورييه
Functional phonetics	علم الأصوات الوظيفي
Hamming Window	نافذة هامينغ
intensité	الشدة
Lungs	الرئتان
larynx	الحنجرة
Mel frequency filtering	ترشيح الإشارة وفقاً لتردد ميل
Mel Frequency Cepstral Cofficients (MFCC)	تقنية التحليل الإشارة الصوتية
Mel frequency scale	لمقياس ميل الترددي

Markov proeesses	عمليات ماركوف
Non Uniform Filter Bank	المرشحات غير المنتظمة
Linguistic symbol	رمز لغوي
Oral cavity	تجويف الفموي
Physiological phonetics	علم الأصوات الفسيولوجي
Pare factions	التخلخلات
pre-emphasis	تحسين الإشارة
phonemes	الصوتيات
Physical	الفيزيائي
phonème	الفونيم/صوتيم /وحدة صوتية
phonème	الصوت
Pitch	طبقة الصوت
Psychophysiology	النفس الفسيولوجي
Phonetics	الأصوات
Rectangular Window	النافذة المستطيلة
Recording & Signal Processing	تسجيل ومعالجة الإشارة
Reception	استقبال
Sound intensity	شدة الصوت
Spectrum power time short	ترشيح طيف الإشارة الصوتية
Short time power spectrum	طيف طاقة الزمن القصير
Sonority	قوة إسماع الأصوات
Signal	إشارة
Short-Time	قصير الحد

Speech Recognition	تعرف الكلام
Speaker recognition	التعرف على المتحدث
Speech Analysis	تحليل الكلام
Seepch	الكلام
Spectrum Analysis	تحليل طيف الإشارة
Sound Réception	استقبال الصوت
Sound source	مصدر الصوت
Sound nnit	وحدة صوتية
Sound wave	موجة صوتية
Speech therapy	علاج النطق /الكلام
Triangular Bandpass Filter	تمرير الحزمة المثلثية
Trachea	القصبة الهوائية
Triangular Window	النافذة المثلثية
Uniform Filter Bank	المرشحات المنتظمة
Vocal tract	المسلك الصوتي
Windowing	النافذة
Word	كلمة
Windowing	النوفذة
Zero crossing	قطع الصفر

قائمة المصادر و المراجع

فهرس المصادر و المراجع:

1. **ابراهيم السامرائي** : نزهة الألباء في طبقات الأدباء ، مكتبة المنار، الزرقاء الأردن، ط3 ،1405 هـ /1985م.

2. **إبراهيم أنيس**: الأصوات اللغوية ، مكتبة الانجلو المصرية ، مصر، 1999 م.

3. **إبراهيم محمد نجا** : التجويد والأصوات ، مطبعة السعادة ، مصر.

4. **ابن المنظور** : لسان العرب ، دار صادر ،بيروت ط1955،1م-1992م مج2 ، الصوت .

5. **ابن جني** : الخصائص ، الكامل في اللغة والأدب ، ، ج 3.

6. **أبو بكر بن دريد** : الاشتقاق ، ، الإعلام ج 6.

7. **أبو على الحسٌين بن عبد الله بن سينا تحقيق محمد حسان** : أسباب حدوث الحرف الطٌّيان : مطبوعات مجمع اللغة العربية ، دمشق 1983م.

8. **احمد عوض** : المجلس الوطني للثقافة و الفنون و الأداب ، الكويت .

9. **أحمد عوض** : موجز تاريخ علم اللغة في الغرب.

10. **احمد مختار عمر**: دراسة الصوت اللغوي ، عالم الكتب، مصر، 1997 م.

11. **أمجد عبد الرازق** : كرجيه وآخرون ، ما نسمع وما لا نسمع .

12. **أمجد عبد الرازق كرجيه** : فيزياء الصوت ، والحركة الموجيه ، الموصل منشورات جامعة الموصل ، 1987م.

13.**أحمد قريش** : اختلاف القدامى و المحدثين في تحديث المخارج والصفات بعض الأصوات ، جامعة أبو بكر بلقايد تلمسان .

14.**أميمة دكاك** : نحو النظام لتركيب الكلام باللغة العربية من نصوص ، المعهد العالي للعلوم التطبيقية و التكنولوجيا باستعمال الضم لأنصاف مقاطع صوتية وتنغيم الطبيعي ، الجمهورية العربية السوريا ، 2018م.

15. **بركات ابراهيم أحمد** : التمييز الكلام العربي باستخدام الشبكات العصبونية الاصطناعية ، جامعة السودان للعلوم و التكنولوجيا ، 2006م

16. **باي ، ماريو** : أسس علم اللغة ، ترجمة أحمد مختار عمر ، القاهرة : عالم الكتب ، القاهر ، 1987م ، ط1.

17.**تمام حساني** : اللغة العربية معناها مبناها مطبعة النجاح الجدّيدة الدار البيضاء ط 1، 1979م.

18.**تواتي** : دراسات وأبحاث لغوية مفاهٌيم في علم اللسان مؤسسة الحياة الجزائر ط1 ، 2006 م.

19.**الجاحظ** : **هو أبو عثمان عمرو بن بحر بن محبوب الكناني** ، البيان والتبيين ، ج5.

20.**حسام البهنساوي** : الدراسات الصوتية عند العلماء العرب والدرس الصوتي الحديث ، مكتبة زهراء الشرق، ، القاهرة مصر،ط 2005 م.

21.**حنفي بن عيسى** : محاضرات في علم النفس ، ديوان المطبوعات في الجزائر .

22.**خالد إسماعيل حسان** : في اللسانيات العربية المعاصرة، مكتبة الآداب، مصر، 2008 م.

23.**جعفر خير** ، دراسة ومقارنة فعالية خوارزميات استخلاص السمات الصوتية في تمييز نمط الشخصية المنفتحة للمتحدث ، مجلة جامعة تشرين للبحوث والدراسات العلمية ، 2018 م

24.**الخليل بن أحمد الفراهيدي** : العين ، ج 2 .

25.**خولة طالب الإبراهيمي**: مبادئ في اللسانيات ، دار القصبة للنشر، الجزائر، 2006 م.

26. **الدكتور احمد عوض** : المجلس الوطني للثقافة و الفنون و الأداب ، الكويت .

27. **رضا زلاقي** : الصوامت الشديد في العربية الفصحة جامعة الجزائر .

28. **وفاء محمد كامل** : المجلس الوطني للثقافة و الفنون و الأداب ، الكويت .

29. **الدليمي ، شهلة عبد الوهاب** : تمييز أصوات الحروف العربية ماجستير ، علوم حاسبات ، كلية علوم الحاسبات والرياضيات ، جامعة الموصل، 2002م.

30. **ذو الفقار ابراهيم مرهج ، غدير أحمد حامد**: التعرف على الكلمات المعزولة ،كلية الهندسة المعلوماتية قسم البرمجيات ونظم المعلومات جامعة تشرين ، 2015م.

31. **رمضان عبد التواب** : المدخل إلى علم اللغة ومناهج البحث اللغوي ، مكتبة الخانجي ، القاهرة ، ط2، 1405 هـ. 1985م.

32. **سمير شريف استيتية اللسانيات** : المجال والوظيفة والمنهج ، عالم الكتاب الحديث ، 2005 م.

33. **شهلة عبد الوهاب عبد القادر**: تمييز أصوات الأرقام العربية ، قسم أنظمة الحاسبات ، المعهد التقني الموصل، الموصل، العراق 2002م.

34. **عبد الحميد محمد أبو** : دراسات في التجويد والأصوات اللغوية ، مطبعة الأمانة، القاهرة، 1983 م

35. **عبد الرحمن أيوب** : أصوات اللغة ، ط1 ، القاهرة ، مطبعة الكيلاني ، 1968 م .

36. **علي فوزي خان و أخرون** ، مشروعات رفع كفاءة معلمي اللغة العربية ومتعلميها ، جامعة ملايا ، ماليزيا

37. **عبد الغفار حامد هلال** : أصوات اللغة العربية ، الطبعة الثالثة ، مكتبة و هبة ، القاهرة ، 1416 هـ /1996م.

38. **عبد القادر قنيني** : مبادئ علم وظائف الأصوات (الفنولوجيا) ، الدار البيضاء .

39. **عصام نور الدين** : علم الأصوات اللغوية ، الفونيتيكا ، دار الفكر اللبناني ،بيروت ، ط1 /1992م

40. **علي النجار** : دار الكتاب العربي ، بيروت ، ج 1 .

41. **عبد الله ابن يحي الفيفي** : مدخل في اللسانيات الحاسوبية ، الرياض ، ط1 ، 2017م.

42. **عمار العايب** : الصوت اللغوي ، دراسة فيزيولوجية ، الصوامت ، الحلقية ، و الحنجرية 2013م

43. **الفيومى** : المصباح المنير ، المكتبة العصرية بيروت ط1 1996م

44. **كمال بشر**: علم الأصوات ، دار غريب ، مصر، 2000 م.

45. **الكسندر و آخرون** : الصوت ، ترجمة محمد عز الدين فؤاد ، القاهرة ، دار الكرنات ، 1962 م .

46. **محمد حسن حسن جبل** : المختصر في أصوات اللغة العربية ، مكتبة الآداب ، مصر، 2012 م.

47. **محمد عبد الرازق كرجيه و د. فيصل عبد الحليم** : ما نسمع وما لا نسمع ، بغداد ، مكتبة النمرود 1988م.

48. **محمد علي الخولي**: معجم علم الأصوات ، مطابع الفرزدق التجارية ، السعودية ، 1982 م.

49. **محمود السعران** : في علم اللغة ، دار المعارف ، مصر ، 1962 م .

50. **مناف مهدي محمد الموسوي** : علم الأصوات اللغوية ، عالم الكتب ، بيروت ، لبنان ، 1998 م ،

51. **منصور بن محمد الغامدي** : الصوتيات العربية ، مكتبة التوبة، السعودية، 2001 م.

52. **مصطفى حركات** : الصوتيات والفونولوجيا ،دار الأفاق الأبيار –الجزائر.

53. **نصر الدين إدريس جوهر**: علم ألأصوات لدراسي اللغة العربية من إلندونيسيين, سدورجو: مكتبة لسان عربي, 2014م .

54. **الخليل بن أحمد الفراهيدي** : لغوي وأديب ، يرجع الفضل له في وضع علم العروض ، وهو صاحب كتاب العين ، (0 0 1 هـ 174 هـ) ج 11.

55. **ونيسة بوختالة** : البنية الصوتية لقصار السورة القرآنية و أثرها في التعليم اللغة العربية 2007م.

56. **ابراهيم بوداود** : فيزياء الحركات العربية بين تقديرات القدامى والقياسات المحدثين ، جامعة وهران سنة 2012م.

57. **محمد سعران** : علم اللغة ، مقدمة لقارئ العربي .

58. **محمد جواد النوري** : الفونيم وتجلياته في القرآن الكريم ، برواية حفص عن عاصم ، سورة البقرة ، 2014م

59. **هشام جبر**: نظرية الاهتزازات و الأمواج الميكانيكية ، ديوان المطبوعات الجامعية ، 1999م ، الجزائر.

الرسائل و المجالات :

1. **يسرى فيصل الارحيم ، عمي انصاف جاسم** : تمييز الأرقام العربية المفردة النطق باستخدام الخوارزمية الجينية مجلة الرافدين لعلوم الحاسبات و الرياضيات المجلد (11) العدد (1) 2014 م.

2. **احمد حسن، رشا رعد هادي** : استخدام نماذج ماركوف المخفية في التعرف على صور الوجه الاعتيادي و المحددة حافاته ، المجلة العراقية للعلوم الإحصائية (21) 2012 م

3. **حسن محمد الياس ، رنا بشار حسين** : استخدام سلاسل ماركوف المخفية في تمييز حروف العلة في اللغة الانكليزية ، المجلة العراقية للعلوم الإحصائية (9) 2006 م.

4. **باسل يونس ذنون الخياط ، زينة فالح صالح العجوز**: تمييز السلالم الموسيقية باستخدام نموذج ماركوف المخفي، المجلة العراقية للعلوم الإحصائية (15) 2012 م.

5. **فاتن بشير عبد الأحد** : تقطيع الكلمة العربية إلى أحرف و تمييزها ، مجلة الرافدين لعلوم الحاسبات والرياضيات المجلد 2005

6. **أحمد قريش** : اختلاف القدامى و المحدثين في تحديث المخارج والصفات بعض الأصوات ، جامعة أبو بكر بلقايد تلمسان ، مجلة اللغة و الأداب ، مجلة قاصدي مرباح –ورقلة ، العدد التاسع ماي 2010م .

7. **مصطفى الشافعي** ، مجلة التعرف الألي على الكلام العربي ، التطورات و التحديات ، المؤتمر السادس للغة العربية ، ص123

8. **مصون نبهان حمصي جربيني** ، نظام تفاعلي ذكي من اجل التعليم على الشبكة العنكبوتية ، رسالة دكتراه ، جامعة حلب سوريا 2010

9. **ميس عبد القادر الكزبري** : تعرف الكلمات المعزولة باللغة العربية، ماجستير في المعلوماتية، جامعة دمشق، سوريا، 2009 م.

10.داود مسعود : رسالة دكتوراه التعرف الآلي على الكلام العربي المتصل ، جامعة الجزائر ، 2016م.

11.**حليم ، علياء موفق عبد المجيد** ، تشفير إشارة الكلام بطريقة البعثرة رسالة ماجستير ، علوم حاسبات ، كلية علوم الحاسبات والرياضيات ، جامعة الموصل 2003م

12.أسعد عبد علي ، دراسة احصائية عن الهدر الزمني ، رسالة ماجستير ، جامعة بغداد ، كلية الاقتصاد ، 1988م .

قائمة المراجع الأجنبية :

1. AV. Aho, J. E. Hopcroft et J. D. Ullman, "The design and analysis of computer algorithms, chapitre 7 : The fast Fourier transform and its application", , 1974.

2. Alotaibi ,yousef A ,Sid –Ahmed Selouani M Alghhamdi , Ali H Meftah (2012) Arabic And English Speech Recognition Using Cross – Language Acoustic Models .11th International Conference On Information Science ,Science ,Signal Processing And their Application (ISSPA)

3. Bisrat Girma , (2012),' Non-uniform sampling based feature extraction for automatic speech recognition ' , master of science in electrical engineering , addis ababa university .

4. Brandstein M. ,2001. Speech Recognition of English Digits, Harvard University

5. Childers,D. G. ,(2000), Speech Processing and Synthesis Toolboxes ", John Wiley & Sons ,Inc.

6. Coust, A.P., (2002), "Connected Word Recognition", http:// a.aidepot.com/paper/nodeq.htm1

7. J.Parker, " Profile Hidden Markov Models " , Biophysical Journal , Vol. 17 ,

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

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

10. L.R. Rabiner, " A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition " , Proceedings of IEEE , Vol. 77 , No. 2 , , (1989).

11. M. Stamp, " A Revealing Introduction to Hidden, Markov Models", IEEE , Vol. 51 , No. 7 , , (2003).

12. Mohammad, B.(December 2014). Production and classification of speech sound, (seminar). Faculty of Informatics engineering, Tishreen University

13. Noelia Alcaraz mesegner , (2009) , ' Speech Analysis for Amtomatic speech recognition ' master of science in electronics , Norwegian university of science and technology .

14. Nahar O. M Khalid, Mohammed., Abu G ShquierWasfi -Al KhatibHusni Al Muhtaseb Arabic (2016) Moustafa Elshafei - phonemes recognition using hybrid HMM/LVQ model for continuous . Speech recognition International of Journal Technology.

15. Philips ,C.,(2001), " A multi-Band Spectral Subtraction Method for Speech Enhancement". http:// www. Utdallas.edu/~Liozou/suni-ms.pdf.

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

17. Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE.

18. S.Duran, "Keyword spotting using hidden markov models", M.Sc. Thesis, Bogazici University, (2001).

19. S.Young, G.Evermann, M. Gales, T.Hain, D. Kershaw, G. Moore, J.Odell, D.Ollason, D. Povey, V.Valtchev, P.Woodland, The HTK Book (for HTK Version 3.3), Cambridge University (2005)

20. Wu , Zhizheng , Oliver Watts ,Simon King (2016)Nmerlin : An Open Source Neural Network Speech Synthesis System 9th ISCA Speech Synthesis Workshop

18. T. Aarnio, "Speech recognition with hidden Markov models in visual communication", M.Sc. Thesis, College Computer Sciences, Turku-University, (1999).

19. T.Pham, "Alignment-free sequence comparison with vector quantization and hidden Markov models", proceeding of IEEE of the Computational Systems Bioinformatics, (2003).

20. T. Macro et T. Frank, «Implémentation d'une calculatrice vocale », EPITA, Université Pierre et Marie Curie, 2004

فهـرس الموضوعات

الفهـــــرس الموضوعات

الإهداء

التشكرات

المقدمة العامة أ-ب-ج-د-هـ

الفصـــــــل الأول: أصوات اللغة العربية وخصائصها

1. مقدمة 10

2. الأصوات اللغوية 11

3.تعريف الصوت 13

4. أنواع الصوت وفروعه 18

1.4. الصوت الطبيعي و هو الصوت العام 18

2.4 . الصوت اللغوي وهو صوت الخاص 19

3.4 . الفرق بين الحرف والصوت 19

5. فروع علم الأصوات 20

1.5 علم الأصوات النطقي 21

1.1.5 وصف جهاز النطق 23

2.1.5 جهاز التنفس 23

3.1.5 الحنجرة 23

4.1.5 التجاويف فوق المزمارية 23

5.1.5 أعضاء التنفس 23

6.1.5 الرئتانLes poumons 23

7.1.5 القصبة الهوائية la trachee artere 23

8.1.5 الحنجرة la gorge 24

9.1.5 تجاريف ما فوق المزمار(Cavités au-dessus de la glotte 25

أ . تجويف الحلق La cavité de la guttural 25
ب. تجويف الفموي La cavité buccale 25
ج. تجويف الأنف La cavité de la nez 26
2.5. علم الأصوات الأكوستيكي أو الفيزيائي 26
1.2.5.خصائص الصوت اللغوية الفيزيائية Les propriétés physiques des sons 27
1.2.5.أ. الشدة Intensité 27
1.2.5.ب. الارتفاع La hauteur 28
1.2.5.جـ. الطابع Le timbre 28
2.2.5. الصفات الفيزيائية المحظة 29
3.2.5. الأصوات الدورية و الأصوات الغير دورية 29
3.2.5. أ. التردد أو التواتر Fréquence 30
3.2.5. ب . السعة Capacité 30
3.5. علم الأصوات السمعي 31
6. مفهوم المصوت أو الفونيم le phonème 33
1.6. اختبار فونيمة الصوت 35
2.6 . مدارس الفونيم 37
1.2.6. المدرسة النفسية 37
2.2.6.المدرسة الفيزيائية 37
3.2.6. المدرسة الوظيفية 37
3.6.مكونات الفونيم 38
7. خاتمة 39

الفصـــــل الثاني : تقنيات المعالجة الإشارة الكلامية

1. مقدمة 42
2. الكلام Seepch 43

3. تحليل الكلام (Speech Analysis) 43
4. صعوبات تمييز الكلام 43
1.4. تغيرات تتعلق بالمتحدث نفسه 44
2.4. تغيرات بين المتحدثين 44
3.4. اختلاف محيط المتحدث
4.4. تغيير طريقة الكلام 44
5.4. اختلاف نطق الصوت الواحد 44
5. أنواع أنظمة تمييز الكلام 45
1.5. أنظمة تمييز على الكلمات المعزولة 45
2.5. أنظمة التمييز على الكلام المتصل غير المستمر 46
3.5. أنظمة التمييز على الكلام المستمر (المتصل) 46
6. تطبيقات وتقنيات ذات علاقة بالأصوات اللغة العربية 47
1.6. التعرف الآلي على الكلام 48
2.6. توليد الكلام آليا 51
3.6.التعرف على المتحدث آليا 52
7. مراحل عملية تمييز على الكلام 52
1.7. مرحلة تسجيل و قراءة الإشارة الصوتية 52
2.7. مرحلة التقسيم إلى إطارات و استخلاص المعاملات 53
3.7. تقنيات المقارنة و الربط للتعرف على الكلام 53
4.7. التعرف على الكلام المعتمدة على المدونات 53
1.4.7. تقنيات إحصائية 54
2.4.7. تقنيات المقارنة 54
3.4.7. تقنيات الربط 54
8. تسجيل الأصوات 55

9. استخلاص المعاملات 56
10. تحويل فورييه: Fourier Transform 56
11. التقطيع إلى إطارات و النوافذ Framing and Windowing 58
12 .أنواع النوافذ 59
1.12. النافذة المستطيلة Rectangular Window 59
2.12. نافذة هامينغ Hamming Window 60
3.12. النافذة المثلثية Triangular Window 61
13. تسجيل ومعالجة الإشارة Recording & Signal Processing 61
14.تحليل طيف الإشارة Spectrum Analysis 61
15. تقنية التحليل الإشارة الصوتية Mel Frequency Cepstral Cofficients (MFCC) 62
1.15. التحليل باستخدام الكبستروم Cepstrum 64
2.15. معاملات MFCC 64
16. خاتمة 67

الفصـــــل الثالث: الأساسيات الرياضية لنماذج ماركوف المخفية HMM

1. مقدمة 69
2. سلاسل ماركوف الملاحظة المخفية 69
1.2. عمليات ماركوف Markov proeesses 70
2.2. نماذج ماركوف المخفية 70
3. رموز النموذج (عناصرHMM) 72
4. انواع نماذج ماركوف المخفية Types of HMMs 73
5. توليد ملاحظات من النموذج 76
6. المسائل الثلاثة الأساسية نموذج ماركوف المخفية 76
7. حلول المسائل الأساسية لنماذج ماركوف المخفية HMM 78
1.7.حل المسألة (1) مسألة التقييم 78

2.7 . حل المسالة (2) 83
3.7 . حل المسألة (3) 85
8. خاتمة 90

الفصـــــل الرابع : الدراسة العملية و نتائـــج

1. مقدمة 92
2.الحزمة الفونية 94
3. انجاز قاعدة البيانات 95
4. تهيئة قاعدة البيانات الأصوات باللغة العربية 95
5. مراحل تصميم برنامج 98
1.5. تجهيز الملفات الضرورية التي تساعد على تمييز على الفونيمات الأحادية 98
2.5. القاموس الفونيمي 99
3.5. استخلاص الصفات و بناء قاعدة البيانات 100
6. تدريب النظام 103
1.6. تحضير المعلومات 103
2.6. مجموعة أدوات HTK 104
3.6. الإختبار 104
4.6. التحليل 106
8. تجارب و نتائج 106
1.8. نسب التمييز المئوية بدلالة حجم قاعدة البيانات الصوتية 107
2.8. نسب التمييز المئوية بدلالة تكرار التهيئة بـ HERest 108
9. خاتمة 113
الخاتمة العامة 115
قائمة المصطلحات 119

قائمة المصادر و المراجع 124
فهرس الموضوعات 133
قائمة الجداول 140
قائمة الأشكال 142
الملخص 145

فهرس الجداول

قائمـــــة الجـداول:

الرقم		الصفحة
1	الجدول 1.1: اختبار فونيمية الأصوات	36
5	الجدول 1.4: مجموعة أدوات HTK	93
6	الجدول 2.4: الحروف الصوتية العربية الكاملة مع ترميزها المستخدم في النظام .	94
7	الجدول 3.4: نسب التمييز المئوية بدلالة حجم قاعدة البيانات الصوتية	107

فهرس الأشكـال

قائمة الأشكال:

الشكل 1.1: فروع علم الأصوات الثلاثة 20

الشكل 2.1: أعضاء النطق في الجهاز الصوتي 22

الشكل 3.1: الوترين الصوتيين 22

الشكل 4.1: القصبة الهوائية و الحنجرة 24

الشكل 5.1: الحنجرة من الخلف 24

الشكل 6.1: تجاويف الجهاز الصوتي 25

الشكل 7.1: الموجة الصوتية المركبة 27

الشكل 8.1: خصائص الموجة الصوتية 28

الشكل 9.1: النواس 29

الشكل 10.1: يرمز البعد بين (أ و ب) و (ج ود) نقطة الستراحة إلى سعة الذبذبة 31

الشكل11.1: يعبر عن الدورات في الثانية (التردد) 31

الشكل 12.1: الأجزاء الرئيسية للأذن 33

الشكل 1.2:أنموذج متكامل للتخاطب مع الآلة 47

الشكل 2.2: موجة صوتية باستعمال برنامج SFS Speech Filing System 50

الشكل 3.2: يمثل هذا الشكل إشارة معينة و الطيف الترددي الخاص بها 57

الشكل 4.2: يمثل النافذة المستطيلة 60

الشكل 5.2: يمثل نافذة هامينغ 60

الشكل 6.2: يمثل النافذة المثلثية 61

الشكل 7.2: مخطط عام يوضح مراحل حساب معاملات MFCC 62

الشكل 8.2: بنك الترشيح Me 63

الشكل 9.2: حساب معاملات الكبستروم. 64

الشكل 10.2: مخطط يوضح مجال عمل تقنية MFCC لاستخراج السمات 65

الشكل 11.2: نافذة هامينغ من أجل طول إطار 25ms 65

الشكل 1.3: عملية ماركوف	70
الشكل 2.3: نموذج الماركوف المخفي	71
الشكل 3.3: الأوعية و الكرات الملونة	72
الشكل 4.3: النموذج الثبوتي	74
الشكل 5.3: نموذج Left-Right Model	75
الشكل 6.3: نموذج Left-Right فيه دلتا تساوي 2	76
الشكل 7.3: متتابعة العمليات المطلوبة لحساب المتغير	81
الشكل 8.3: متتابعة العمليات المطلوبة لحساب المتغير الخلفي $\beta_t(i)$	82
الشكل 9.3: تدريب نموذج نماذج ماركوف مخفي باستخدام خوارزمية Baum-Welch	88
الشكل 1.4: تسجيل و معاينة الأصوات باستعمال برنامج SFS	96
الشكل 2.4: الإشارة الصوتية مسجلة بصوت رجل عند تلفظ الجملة	96
الشكل 3.4: الجمل الأصلية المراد التعرف عليها	97
الشكل 4.4: مكونات النظام تمييز على الكلام	98
الشكل 5.4: listeMots قائمة الكلمات التي تشكل قاعدة البيانات corpus	99
الشكل 6.4: القاموس Dictionnaire	100
الشكل 7.4: ملف التكوين config	101
الشكل 8.4:نموذج ماركوف المخفي	101
الشكل 9.4: ملف التهيئة (prototype) HMM .	102
الشكل 10.4: بنية نظام للتعرف بواسطة HTK	103
الشكل 11.4: منحنى تغير نسبة التميز بدلالة حجم قاعدة البيانات الصوتية	107
الشكل 12.4: منحنى تغير نسبة التمييز بدلالة تغير عدد تكرارت إعادة تقدير	109
الشكل 13.4: احدى واجهات التنفيذ بلغةPerl	110
الشكل 14.4: نموذج حول النتيجة النهائية لنظام تمييز الآلي للكلام العربي بالاعتماد على الفونيمات الأحادية	112

الملخص

ملخص

إن العمل المنجز في هذا البحث يدخل في الإطار العام لنظم التمييز الآلي للكلام العربي **بالإعتماد على الفونيمات الأحادية** عن طريق الحاسوب ، التي تعتمد على الطرق الإحصائية في بناء نماذج للوحدات الصوتية

تم تطبيق خوارزمية MFCC لإستخراج السمات التي تستخدم على نطاق واسع كمدخل لغرض التمييز في أغلب نظم التمييز الآلي للكلام.

إن التجربة المنجزة في هذا البحث كانت قائمة حول نظام التمييز المركوفي المعلمي المبني على نماذج ماركوف المخفية Hidden Markov Models ، للتخفيف من كثرة البرمجة ، قمنا باستعمال البنية البرمجية HTK (علبة الوسائل لنماذج مركوف المخفي) و لقد تم اختيار هذه البنية البرمجية لمرونتها ولقابليتها للتغير بكافة أجزائها و هذا على مختلف مراحل إنجاز النظام .

لأجل تقييم قدرة النظام المنجز على التمييز قمنا بمجموعة من الاختبارات والتجارب التي توصلنا من خلالها إلى نتائج جيدة ومشجعة وتم تدريب نموذج احتمالي لتمييز كل كلمة على حدى حتى تحصلنا على نسبة تمييز بـ:74.69 %

كلمات المفاتيح: العلاج الآلي للكلام ، التعرف الآلي على الكلام ، التحليل الصوتي ، معاملات MFCC ، نماذج ماركوف المخفية HMM.

Résumé

Le travail réalisé lors de cette thèse s'inscrit dans le cadre général de la reconnaissance Automatique de la Parole (RAP) robuste , dans la construction de modules pour les unités vocales et linguistiques, l'algorithme MFCC a été appliqué pour extraire des caractéristiques largement utilisées en tant qu'intrants à des fins de discrimination, La plupart des systèmes RAP fonctionnent correctement dans un environnement aux caractéristiques acoustiques sont proches de l'environnement dans lequel s'est fait l'entraînement.

Le travail réalisé lors de cette thèse a été évalué par le système de reconnaissance Markovien de référence fondé sur les HMM. A fin de réduire au minimum la tache de programmation, nous avons utilisé la plate-forme logicielle HTK (Hidden Markov Model Toolkit) , nous avons choisi cette plate forme pour sa convivialité, sa souplesse et sa grande liberté de choix ont laisse tout au long de la construction des différentes parties du système de reconnaissance.

Nous avons utilisé le système de reconnaissance qui nous avons présenté pour estimer la prononciation et indiquer le taux de réussite de la prononciation des phonèmes, qui nous donne le taux de reconnaissance **74.69%** .

Mots clé: Le Traitement Automatique de la Parole, la reconnaissance automatique de la Parole, MFCC l'analyse de la parole. HMM Hidden Markov Model

Abstract

The work done in this thesis is part of the general framework of robust Automatic Speech Recognition (RAP), In the construction of modules for both vocal and linguistic units, the MFCC algorithm has been applied to extract features widely. Are Used as inputs for discriminatory purposes, most RAP systems work properly in an environment with acoustic characteristics that are close to the environment in which the training was conducted.

The work done in this thesis was evaluated by the Markovian reference system based on HMM. In order to minimize the programming task, we used the software platform HTK (Hidden Markov Model Toolkit), we chose this platform for its iflexibility and its suseepti bilinty of change of the different parts of the reconnaissance system.

In ordre to avalante the done system , we presented some exams and experiments wgich may give us good and motivated results , and we presented it to discriminate the pronunciation of each word till we got the rate **74.69%**

Key word : Speech Automatic Recognition, acoustic analysis, Hidden Markov Model (HMM).