وزارة التعليم العالي و البحث العلمي

جامعة الجزائر2 أبو القاسم سعد الله

كلية اللغة العربية و آدابها قسم علوم اللسان

و اللغات الشرقية

طرق تحليل الإشارة الكلامية و تأثيرها في التعرف الآلي على الكلام العربي

أطروحة مقدمة لنيل شهادة الدكتوراه علوم

تخصص علوم اللغة

الإسم و اللقب	الدرجة العلمية	الجامعة الأصلية	الصفة
د. جمال موسى	أستاذ التعليم العالي	جامعة الجزائر2	رئيسا
د. عظامو آمال بهاء هدى	أستاذ التعليم العالي	جامعة الجلفة	مشرفا ومقررا
د.كمال فرات	أستاذ التعليم العالي	جامعة الجزائر2	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

إهـــــداء

إلى من قرن الله اسمي باسمهما، من فوق سبع طباق، و أوصى ببرهما من سابع سماء

إلى التي نلتمس الجنة تحت قدميها، إلى نبع العطاء، ونبض الحياة في وجودنا، إلى البلسم الشافي لكل ألم و جراح، إلى معلمتي الأولى، والدتي العزيزة.

إلى مَثل الأبوة الأعلى، و نبراسي الذي ينير دربي، إلى من رفعت رأسي عاليا افتخارا به ...والدي.

إلى التي ملأت حياتي عبقاً وحباً وجمالاً ، وأهدتني فرحة قلبي ونور عينيّ ... زوجتي الغالية

إلى من حبهم يجري في عروقي و يلهج بذكراهم فؤادي، إلى الوجوه النيرة...إخوتي

إلى كل الأهل و الأقارب الذين لم تسعهم ورقتي هاته

إلى كلّ هؤلاء أهدي لهم ثمرة جهدي المتواضع ...هذا العمل...عربون محبّة، وعرفاناً وردّاً للجميل.

كلمة شكر

لا بدَّ لنا و نحن نخطو هذه الخطوات في الحياة العلمية من وقفة نعود بها إلى أعوام قضيناها في رحاب الجامعة مع أساتذتنا الكرام، الذين قدموا لنا الكثير باذلين بذلك جهودا كبيرة لأجل بناء جيل الغد، لبعث الأمة من جديد و قبل أن نمضي نتقدم بأسمى آيات الشكر و الامتنان و التقدير و التحية إلى الذين حملوا أقدس رسالة في الحياة إلى الذين مهدوا لنا طريق العلم و المعرفة

إلى جميع أساتذتي الأفاضل من الابتدائية إلى مرحلة ما بعد التدرج

و أخص بالتقدير و الشكر الأستاذة: **عظامو أمال بهاء**

و التي شاء القدير أن تتفضل عليَّ مرتين...الأولى بتدريسي في أثناء مرحلتي الجامعية الأولى و الثانية بإشرافها على هذا العمل و سهرها على إتمامه على أكمل وجه

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، ناهيك عن العراقيل التى ذللها لنا بسعة و رحابة صدرٍ

و كذلك نشكر كل من ساعدنا على إنجاز هذا العمل و قدم لنا يد العون، و نخص منهم

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

أخلص التشكرات إلى **الدكتور داود مسعود** على كل ما قدمه من توجيهات

كما أتقدم بجزيل الشكر إلى الأساتذة الأفاضل أعضاء لجنة المناقشة على قبولهم تقييم عملنا هذا فلهم منا كل الشكر و التقدير.

مقدمة

مقدمة

مر أكثر من عشرين عاماً على ابتكار الحاسوب الشخصي، وعلى الرغم من أنه كان حصيلة قرون من التطور المعرفي والتقني الذي شاركت فيه الإنسانية جمعاء، إلا أن مبتكريه في الولايات المتحدة الأمريكية لم يأخذوا بالاعتبار في أثناء تصميمه من ناحية العتاد والبرمجيات، سوى حاجات أسواقهم وثقافتهم ولغتهم، ولم يكن لديهم رؤية إستراتيجية لتطوره وانتشاره على مستوى العالم. لكن ما إن حدث ذلك الانتشار، ودخل استخدام الحاسوب في كافة المجالات، حتى برزت الحاجة لمواءمته مع اللغات الأخرى، ومنها اللغة العربية، وظهرت تحديات إمكانية تنفيذ ذلك على أنظمة لم تصمم أصلاً للغات مختلفة، إلا أن بنية الحاسوب الشخصي المفتوحة، وتطوره السريع، سهل نسبياً إضافة برمجيات وتقنيات ملحقة، خاصةً بثقافات ولغات أخرى، والتي كانت اللغة العربية واحدة منها.

بدأنا نشهد في منتصف التسعينيات نتائج البحث والتطوير لتقنيات حاسوبية تهدف إلى معالجة اللغة العربية ذاتها. ومن هذه التقنيات: التدقيق الإملائي والنحوي، التحليل الصرفي، الترجمة الآلية، تقنيات البحث واسترجاع النصوص وتوليد الخطوط العربية آلياً، التعرف الآلي على الكلام العربي.

يُعتبر العلاج الآلي للكلام من أهم الفروع التي لاقت اهتماما بالغا من الباحثين نظرا لتطبيقاته المتنوعة، فهو يعتمد على عملية تحويل الإشارة الصوتية، الملتقطة إلى مجموعة معطيات؛ قد تكون هذه المعطيات هي النتيجة المرجوة النهائية، وقد تستخدم كمدخل إلى نظم معالجة لغوية لاحقة للحصول على فهم الكلام، و لإعطاء تفسير لهذا الكلام والتصرف على أساسه، و قد صار لدينا خوارزميات ونظم علاج آلي للكلام عالية الأداء تتألف من خمسة مجالات رئيسة هي : التعرف (التمييز) الآلي للكلام و الإدراك و الإنتاج ، والترميز وتركيب الكلام, و هذه النظم لا تتعلق بالهندسة فحسب بل هناك مجموعة من المعارف الأساسية حول مفهوم الإدراك والكلام واللسانيات وعلم الأصوات نحن بحاجة إلى معرفتها. و بعبارة أخرى، فإن دراسة الكلام تشمل تخصصات

عـدة في آن واحـد، مثـل الإلكترونيـات و الصـوتيات واللغويـات، وتقنيـة المعلومـات، علـم الـنفس، و الـذكاء الاصطناعي والرياضيات، وما إلى ذلك.

التعريف بالموضوع و أهميته:

إن التعرف الآلي على إشارة الكلام هو من المواضيع التي تعرف أعمال بحث عديدة، فهو عملية تقطيع الإشارة إلى مقاطع ذات معنى، يجب أن تعالج لاحقاً لأجل الوصول إلى شيء يمكن القول عنه بأنه معلومات مفيدة تمكننا من التعرف على كل مقطع منها. و تُعد غالباً هذه العملية من المعالجات الأولية في أغلب أنظمة معالجة الكلام، وذلك لأجل الوصول إلى الهدف المتمثل بالحصول على صفات مفيدة وتحمل الكثير من المعلومات.

إن الشروع في وضع خوارزمية للتعرف الآلي على الكلام لا بدَّ أن يمر بمرحلة تحضير قاعدة للبيانات تحتوي على مقاطع إشارة الكلام المراد التعرف عليها، والتي هي الحروف العربية، فهذه المرحلة تعد مرحلة هامة وتتطلب دقة عالية ومعرفة 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 من الجمل باللغة العربية تكون الحروف فيها في مختلف المواضع (في بداية، و وسط، و نهاية الكلمة) و بعدها نقوم بتسجيل هذه الجمل المختلفة بواسطة عدة أشخاص، ثم عملية تقطيع الإشارة المتحصل عليها إلى مقاطع تُعالَج و تُحلّل لاستخراج مجموعة من القيم تسمى معاملات الإشارة الكلامية، هذه المعاملات تختلف حسب التقنية المستخدمة لتحليل المقاطع الصوتية من أجل الحصول على قاعدة البيانات، التي تستعمل لاحقا في عمليتي التدريب والاختبار لخوارزمية التعرف الآلي على الكلام العربي و هما المرحلتين الأخيرتان في إنشاء النظام.

إن مرحلة تحليل إشارات الكلام لاستخراج المعاملات الخاصة بكل مقطع صوتي، تلعب دورا أساسيا في مدى نجاعة أي برنامج للتعرف، لهذا فإننا نلاحظ أن خوارزميات التعرف قد تختلف عن بعضها البعض في بعض السياقات والقواعد التي تتبعها في أثناء عملها على الرغم من أن النتيجة التي يجب أن تعطيها واحدة. فكيف يمكن أن نستغل التقنيات الحديثة التي تستعمل في التعرف على الأنماط كنماذج ماركوف المخفية (المصنفات) ، و التي هي عبارة عن خوارزمية تضم سلسلة من المعالجات تتم على شكل مراحل تنفذ بالتعاقب , لأجل هدف أساسي هو تمييز الكلام العربي بدقة عالية ،و ما هو تأثير نوع التقنية المستخدمة في استخلاص المعاملات أو الصفات من الإشارة المدخلة على دقة تعرف النظام الذي يتبع هذا النوع من المصنفات ؟

دوافع اختيار الموضوع

لقد كان وراء اختياري لهذا الموضوع عدة أسباب أهمها ما يأتي:

•قلة البحوث والدراسات المتعلقة بالتعرف الآلي على الكلام العربي و خاصة دراسات المقارنة من هذا النوع و التي من شأنها تسهيل عمل الباحثين في هذا الميدان و توفير قاعدة بيانات تساهم في تطوير البحوث و توفير الوقت على الباحثين.

• مقارنة أهم طرق تحليل إشارة الكلام العربي و أثرها في نجاعة برامج التعرف الآلي على الكلام العربي التي تعتمد الخوارزميات الإحصائية.

• إثراء المكتبات العربية بدراسات فيزيائية للأصوات العربية.

• رغبتي في المساهمة في كسر الحاجز بين اللغة العربية وعالم الحاسوب.

• مسيس الحاجة إلى التعرف الآلي على الكلام الذي يسمح للمستخدم بتطبيق المهام المختلفة على الأجهزة شفويا: النسخ، والسيطرة، والترجمة،... .

منهجية العمل: لتحقيق هذا الهدف، فقد نظمنا عملنا على النحو الآتي:

المقدمة: فيها بيان تعريف بالبحث وطرح الإشكال، ثم عناصر البحث وأهدافه و مناهجه.

الفصل الأول: نتناول فيه اللسانيات الحاسوبية العربية من ناحية المفهوم و المبادئ و الأسس التي يقوم عليها هذا العلم، كما نتطرق فيه إلى الإنجازات التي وصل إليها الباحثون العرب في هذا الميدان.

الفصل الثاني: يعطي لمحة عامة عن التقنيات المستعملة في تحليل الإشارة الصوتية واستخراج السمات الأساسية منها.

الفصل الثالث: شرح وتقديم لبعض الخوارزميات المستخدمة في نظم التعرف الآلي على الكلام.

الفصل الرابع و الأخير ، نناقش فيه بتفصيل أكبر منهجية العمل, باستخراج المعاملات من الإشارة الصوتية بعدة طرق , و استخدامها في كل مرة من أجل التعرف الآلي على الكلام و المقارنة بين النتائج المحصل عليها من حيث نسبة تمييز الكلام مع العلم أن تطبيق نظام التعرف الآلي على الكلام يكون على مدونة أصواتها باللغة العربية.

وفي خاتمة هذا البحث سجلت أهم النتائج المحققة في هذه الدراسة ، حيث قدمت مقارنة بين مختلف النتائج المحققة في التعرف الآلي على الكلام العربي و مدى تأثير نوع المعاملات الناتجة عن تحليل إشارة الكلام في مدى نجاعة نظام التعرف.

المصادر و المراجع التي تم اعتمادها لإنجاز البحث، فهي متعددة و متنوعة, منها القديم و الحديث, و غرض هذه المزاوجة هو الاستفادة من مجهودات القدماء و المحدثين على حد سواء .

أخيرا أطمح من خلال الجهد الذي بذلته، رفقة توجيهات المشرف الفاضل أن أكون قد أضفت لبنة في خدمة اللغة العربية عامة و علم اللسانيات الحاسوبية خاصة, و كلي أمل أن يتقبله الله تعالى, و أن يكون إفادة للغتي الحبيبة.

الفصل الأول:

اللسانيات الحاسوبية

1.مقدمة

استطاع الإنسان، عبر عصوره المختلفة، أن يجلي ما خفي عليه من ظواهر كونية، وبدأ في تطوير العلوم التي يمتلكها، وسعى إلى امتلاك معارف جديدة، في سبيل وصوله إلى قمة المعرفة .وكان مما وصل إليه الإنسان المعاصر، في ميدان الدرس اللغوي، ما عُرف حديثاً بعلم اللسانيات الحاسوبية، الذي يُعد الآن من أبرز العلوم اللغوية، التي ظهرت في العصر الحديث.

2. المفاهيم الرئيسة للسانيات الحاسوبية

اللسانيات الحاسوبية هي أحدث فروع اللسانيات، ولعلها تكون أهم هذه الفروع جميعاً في عصر تتعاظم فيه أهمية الآلة والتقنية والمعرفة.

و يبدو جلياً أن هذا العلم فرع بينيّ ينتسب نصفه إلى اللسانيات وموضوعها اللغة، ونصفه الآخر حاسوبي وموضوعه ترجمة اللغة إلى رموز رياضية يفهمها الحاسوب، أو تهيئة اللغة الطبيعية لتكون لغة تخاطب وتحاور مع الحاسوب، بما يفضي إلى أن يؤدي الحاسوب كثيراً من الأنشطة اللغوية التي يؤديها الإنسان، مع إقامة الفرق في الوقت والكلفة.

وتقوم اللسانيات الحاسوبية على تصور نظري يتخيل الحاسوب عقلاً بشرياً، محاولة استكناه العمليات العقلية والنفسية التي يقوم بها العقل البشري حين ينتج اللغة ويستقبلها، من ثَمَّ يفهمها ويدركها، ولكنها تستدرك على الحاسوب أنه جهاز أصمُّ لا يُسْتَعمل إلاّ وفق البرنامج الذي صممه الإنسان له، ولذلك ينبغي أن نوصِّف للحاسوب المواد اللغوية توصيفاً دقيقاً يستنفد الإشكالات اللغوية التي يدركها الإنسان بالحدس.[1]

و ينحو عبد الرحمن حاج صالح هذا المنحى في دلالة المصطلح "اللسانيات الحاسوبية" على الحقل الذي تمتزج فيه اللسانيات بالمعلوميات حيث يقول: ((إن الدراسات و البحوث العلمية في اللسانيات الرتابية

[1]وليد أحمد العناتي, "مقال اللسانيات الحاسوبية العربية رؤية ثقافية"، مجلة فكر و نقد، عدد82. أكتوبر 2006 المغرب.

(الحاسوبية) ازدهرت في الوطن العربي في هذه الآونة وتكاثر إلى حد ما الباحثون في هذا الميدان الذي تتلاقى فيه علوم الحاسوب وعلوم اللسان، وهو ميدان علمي وتطبيقي واسع جدا كما هو معروف، إذ يشمل التطبيقات الكثيرة كالترجمة الآلية ، و الإصلاح الآلي للأخطاء الطباعية ، وتعليم اللغات بالحاسوب..))[1]

إن التعاريف السابقة لا تختلف كثيرا عمَّا عرفت به اللسانيات الحاسوبية، في أول مؤتمر دولي يقام بشأنها، بأنها ((علم جديد تتقاطع فيه اللسانيات مع جهاز صوري تفرزه العلوم المنطقية الرياضية، ويخضع للقيود التي تفرضها الآلات المعقدة للمعالجة الآلية للمعلومة، ويؤدي البحث في هذا المجال إلى إنشاء نموذج خوارزمي)) [2] .لقد تم في هذا التعريف تحديد ملامح اللسانيات الحاسوبية التي تتمثل في تلاقي جهازين مفهومين حديثين نوعا ما هما: اللسانيات في تطورها المتزايد، و العلوم المنطقية الرياضية في رؤيتها الصورية ، بمجال تقني حديث النشأة ، كما أعلن هذا التعريف عن الهدف من ظهورها وهو صياغة نموذج خوارزمي، وهو خطاطة منطقية رياضية للغة قابلة لأن تشغل في الحاسوب [3].

يحيل مصطلح اللسانيات الحاسوبية في الأدبيات العربية إلى المجال الذي ترتبط فيه اللسانيات أو علوم اللغة بعلوم الحاسوب ،فمما لا شك فيه أن معالجة اللسان العربي كلسان من الألسن الطبيعية تدخل في علم مخصوص وليد التطورات التكنولوجية المتقدمة، ألا وهو اللسانيات الحاسوبية، مجالها البحثي دقيق و جديد يعرض لآخر النظريات و التطبيقات الحاسوبية المجربة على جميع الألسن الطبيعية. و هذا ما يؤكد على ارتباط هذا التخصص بالتكنولوجيا و الإعلام الآلي ،حيث يلتقي فيه الجانب النظري اللساني بكل خلفياته المعرفية و المنهجية و الجانب التقني المعلوماتي بكل تطوراته مما سمح بصياغة ما اصطلح عليه بـ"الهندسة اللسانية" أو "تكنولوجيا اللسان"، مما أعطى للسانيات الحاسوبية طابعا تقنيا شديد الارتباط بالآلة، ويؤكد على ذلك طريقة صياغة هذا المصطلح

[1] رضا بابا أحمد، اللسانيات الحاسوبية، مشكل المصطلح و الترجمة، ص2

[2] المرجع نفسه، ص3

[3] المرجع نفسه ص2.

(اللسانيات الحاسوبية) ؛فقد تم وصفه بـ" الحاسوبية " التي تشير إلى نسبته و تعلقه بالحاسوب ، وهي الآلة التي تتجلى فيها معالجة المعلومات بطريقة آلية[1].

وكغيره من فروع اللسانيات، ينتظم اللسانيات الحاسوبية مكوّنان: أحدهما تطبيقي والآخر نظري. ويميل إلى ذلك **نهاد الموسى** مع تفصيل أكثر، حيث يعتبر اللسانيات الحاسوبية نظاما بينيا بين اللسانيات وعلم الحاسوب، يعني بحوسبة جوانب الملكة اللغوية[2].

أما التطبيقي فأول عنايته بالناتج العملي لنمذجة الاستعمال الإنساني للغة، وهو يهدف إلى إنتاج برامج ذات معرفة باللغة الإنسانية. وهذه البرامج مما تشتد الحاجة إليه من أجْل تحسين التفاعل بين الإنسان والآلة؛ إذ إن العقبة الأساسية في طريق هذا التفاعل بين الإنسان والحاسوب إنما هي عقبة التواصل[3].

وأما النظري (أو اللسانيات الحاسوبية النظرية) فتتناول قضايا في اللسانيات النظرية، تتناول النظريات الصورية للمعرفة اللغوية التي يحتاج إليها الإنسان لتوليد اللغة وفهمها[4].

وأمّا منتهى الغاية التي تجتهد اللسانيات الحاسوبية أن تُحَصِّلها فهي أن نهيّئ للحاسوب كفاءة لغوية تشبه ما يكون للإنسان حين يستقبل اللغة ويدركها ويفهمها ثم يعيد إنتاجها على وفق المطلوب.

والكفاية هنا هي المؤدى الضمني لمفهوم **تشومسكي**. وهي تتألف على المستوى النظري من[5]:

أولاً:استدخال قواعد اللغة"العربية"، في نظامها الصوتي، وأنساقها الصرفية، وأنماط نَظْمها الجُمْليّ، وأنحاء أعاريبها، ودلالات ألفاظها، ووجوه استعمالها وأساليبها في البيان، وأحكام رسمها الإملائي.

---

[1] رضا بابا أحمد، المرجع السابق، ص2، 3

[2] د نهاد الموسى، حصاد القرن في اللسانيات، الأردن ، مؤتمر عبد الحميد شومان ،ط د ت،ج2،ص47

[3] نهاد الموسى، العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، ص: 53-54

[4] المرجع نفسه، ص54.

[5] نهاد الموسى، الأساليب في تعليم اللغة العربية، ص: 123-124

وبيان ذلك أن لكل لغة نُظمها الصوتية والصرفية والنحوية الخاصة التي تتأسس على قواعد يكتسبها الناطق باللغة على نحوٍ غير واعٍ، ويؤديها بتلقائية. ومن أمثلة ذلك من العربية:

– ففي النظام الصوتي يَصْدُرُ العربي في نطق لغته عن قوانين صوتية لا تتهيأ له بالمعرفة المباشرة، فهو يُماثِل في(ال) الشمسية، ويُقَصِّر العلّة قبل الساكن، ويراعي قوانين التركيب المقطعي، فلا يبدأ بصامتين ولا ينتهي بهما إلاّ في الوقف حصراً.

– وفي النظام الصرفي يعرف الأبنية الصرفية للاسم والفعل، وما هو مشترك بينهما، ويعرف أبنية اسم الفاعل واسم المفعول واسم المكان، فإذا التبست الأبنية الصرفية استعان بالسياق ليفض اللبس.

– وفي النظام النحوي يعرف قواعد إعراب الأسماء والأفعال والحروف، ويعرف قواعد تركيب الجملة الاسمية، وتركيب الجملة الفعلية، وما يعرض لكل واحدة منهما من امتدادات من اليسار أو اليمين

– وفي الدلالة يعرف معاني المفردات المعجمية، ويستعين بالسياق ليفضّ الالتباس بين الألفاظ المشتركة.

وكذا القول في النظام الأسلوبي والهجائي. وهذه القواعد اللغوية متناهية مهما بلغ عددها.

ثانياً: إنتاج ما لا يتناهى من الأداءات اللغوية الصحيحة، فبالرغم من تأسيس اللغة على قواعد محدودة، إلاّ أننا نستطيع أن نولِّد من هذه القواعد عدداً غير محدود ولا معدود من الأداءات اللغوية. وهذا ما عرفته العرب بـ"القياس والتمثيل" وعرفه تشومسكي بـ"اللاتناهي".

غير أن هذه التطبيقات الحاسوبية الكثيرة التي تعالج اللسان العربي ليس من اليسير أن تجمع في أصول واحدة و أسسها الابستمولوجية غير واضحة ، وبالتالي لم توضع لها المقدمات التعليمية التي تسهل على القارئ العربي المتعلم أو الباحث أن يستفيد منها رغم ذلك .ويستنتج مما تقدم أن اللسانيات الحاسوبية هي مجال تتداخل فيه التصورات اللسانية والحاسوبية و تتلاقح لتشكل نظريات تعمل على معالجة الوقائع اللغوية وفق منهج حاسوبي لتتمخض عن ذلك تطبيقات متعددة تشمل تلك الوقائع اللغوية لكن في إطارها الآلي .وبالتالي، وإن كانت

اللسانيات علمًا متجذراً في الفكر الإنساني غير أن ارتباطها بالحاسوب هو من ابتداع القرن العشرين، عصر ثورة المعلومات .

ويرجع السبب في ذلك الارتباط إلى كون الحواسيب تمثل أوج ما بلغه التقدم التكنولوجي ، وأهم ما تحتاج إليه الحياة المعاصرة، لأنها تساعد على حل كثير من مشكلاتها المعقدة ، ويتم ذلك بالتواصل مع الحواسيب ، ويخدم الأهداف المتعلقة به والتي تنحصر في حل المشكلات المعقدة التي تتصل بحوسبة اللغة [1].

3. أسس اللسانيات الحاسوبية :

اعتمدت اللسانيات الحاسوبية في بدايتها، على التحليل الإحصائي للمفردات اللغوية في كتاب معين أو لدى كاتب معين لإعداد فهارس أبجدية لتلك المفردات، وتحديد تواترها في مؤلفاته ، ثم خطت خطوة مهمة جدا في اختصار الزمن عندما أعدت معاجم إلكترونية أحادية اللغة أو ثنائيتها أو متعددة اللغات.

تلاها خلال الستينيات من القرن العشرين و حتى الآن إنتاج برامج وأنظمة للترجمة الآلية وغيرها من التطبيقات اللسانية الحاسوبية،كان بعضها تجاريا بحتا. تُغذى بها الحواسيب لتترجم جملا مكتوبة أو منطوقة و مصطلحات كاملة في مجالات متعددة أهمها السياحية و التجارية و المرتبطة بالخدمات كمصطلحات التحية و الاستفسار عن الأسعار و الأماكن و الزمن وغيرها ،لكن كغيرها من التخصصات الحديثة و الأجنبية المنشأ، تتعدد المصطلحات الدالة على اللسانيات الحاسوبية والتي عكف الباحثون على استخدامها سواء كان ذلك في المراجع الأجنبية أم العربية ، مما يسبب مشكلا في توظيف المصطلح و ترجمته إلى اللسان العربي[2].

وقد كان أول مخترع أخرج منهجا في الترجمة الآلية وهو"مختبر جورج تاون" و كان من أقدم الاختصاصيين في الحاسوب الذين شعروا بأهمية الترويج الفعلي بين علوم الحاسوب ، وعلوم اللسان هو الباحث الأمريكي **د.ج. مايس** ، ثم **ف.إيجنف**، ويصرح هذا الأخير وهو من زملاء **تشومسكي** أن الترجمة الآلية المناسبة هي التي تعتمد

[1] رضا بابا أحمد اللسانيات الحاسوبية، مشكل المصطلح و الترجمة ،ج3.ص3-4

[2] المرجع نفسه، ص 21.

على أوصاف بنوية مناسبة للألسن المترجم منها وإليها ، ومنذ ذلك الوقت اهتم الباحثون بالنظريات اللغوية كأساس للعلاج الآلي للغات ، ثم توالى بعد ذلك افتتاح المراكز الحاسوبية للغة في أوروبا والاتحاد السوفياتي[1].

و اللسانيات الحاسوبية أحدث فروع اللسانيات في عصر تتعاظم فيه أهمية الآلة و التقنية، والمعرفة التكنولوجية ، وقراءة في عنوان هذا المحور، نجد أن النصف الأول من العنوان ينسب إلى اللسانيات ، ونصفه الآخر حاسوبي وموضوعه ترجمة اللغة إلى رموز رياضية يفهمها الحاسوب [2].

هذا ما أدركه "مازن الوعر " من علماء اللسانيات ، وعلماء الهندسة ، و الحاسبات الالكترونية في مؤتمر " اللسانيات التطبيقية ومعالجة الإشارة و المعلومات"من أنه لا يمكن للسانيات الحاسوبية (المعلوماتية) أن تكون علما قائما برأسه وله هويته ، ومبادؤه ومناهجه ،وتطبيقاته التكنولوجية إلا من خلال التعاون ، و التنسيق بين العلماء اللسانيين وبين علماء الهندسة الالكترونية وبين علماء الحاسبات الإلكترونية[3].

إن جهود هؤلاء العلماء تقوم على تصميم برنامج يصف للحاسوب المواد اللغوية وصفا دقيقا يستنفد الإشكالات اللغوية التي يدركها الإنسان بالحدس في إطار تناسق فروع كثيرة ، تتآلف لتشكل مبادئ اللسانيات الحاسوبية (المعلوماتية) من خلال البحث في اللغة البشرية كأداة طبيعية لمعالجتها في الآلة و الحاسبات الإلكترونية و استفاء مبادئ هذا العلم من: اللسانيات العامة بجميع مستوياتها التحليلية و الصوتية و النحوية و الدلالية ، ومن علم الحاسبات الإلكترونية (الكمبيوتر)،ومن علم الذكاء الاصطناعي ، وعلم المنطق ، ثم علم الرياضيات ، وإذا كان تحليل الكلام وتركيبه قد طبق أولا على المسائل الرياضية ،فإن العلماء أدركوا فيما بعد أن اللغة الطبيعية

[1]- أ.عيسى مومني ، بيبليوغرافيا اللسانيات .ط د ت. ص108ـ109.

[2] - المرجع نفسه، ص107.

[3] - المرجع نفسه، ص 107

البشرية هي نظام رياضي اتصالي كأي نظام من الأنظمة ، وهذا التمثيل للمعرفة الإنسانية و الآلات التكنولوجية كالحاسبات الإلكترونية مرتبط ارتباطا وثيقا بتحليل اللغات الإنسانية و تركيبها[1].

1.3 اللسانيات العامة :

اللسانيَّات العامَّة هي المصطلح المقابل ل (Linguistique générale) في الفرنسيَّة أو (General linguistics) في الإنجليزية، ويُستعمل المصطلح لإثبات الفرق بين اللسانيات العامة واللسانيات الخاصة، واللسانيات الخاصة هي التي تهتم بلغةٍ خاصة، وذلك من نحو لسانيات العربية أو الألمانية أو اليابانية أو غيرها، في حين أنَّ اللسانيَّات العامَّة تهتم بكل اللغات، أو بالأحرى بمجموع الخصائص المشتركة للُّغات.

وموضوع اللسانيات العامة على وجه الحصر، وبطبيعة الحال , هو اللغة الطبيعية أو اللغة البشرية، وإن شئْنا الدقة هو الوقوف على السمات المشتركة بين مجمل اللغات، سواء تعلقت هذه السمات بالجانب المعجمي أو الدلالي، أو الجوانب الصوتية أو التركيبية أو غيرها .

وبلا شك إنَ لكل لغة سماتها أو خصائصها اللسانية؛ ولكن هذا لا ينفي وجود خصائص مشتركة بين اللغات، والمشترك بين اللغات هو الأداء اللغوي ذاته، باعتبار أن اللغة قائمة على الكلام وعلى التعبير وعلى التواصُل، وهي قائمة على بنية واضحة، أو على نظامٍ لسانيٍّ يشتمل على أنساق مضبوطة، وإن اختلفت هذه الأنساق من لغة إلى أخرى قليلًا أو كثيرًا.

و الكلام البشري مبدؤه واحدٌ، باعتباره أصواتا دالةً، وهذه الأصوات بحاجة إلى التأليف فيما بينها؛ لتنشأ منها كلماتٌ أو وحداتٌ دالة، ولتنشأ عن هذه الكلمات عبارات أو جمل أو نصوصٌ لها دلالاتها ولها مقاصدها، يُعبِّر بها كلُّ فرد عن مقاصده.

[1] – عيسى مومني، بيبلوغرافيا اللسانيات، ص107ـ108.

و الكلام دفق صوتيٌّ يتحكَّم فيه المتكلِّم بالتقديم والتأخير، والحذف والتَّكرار، وتَصحبُه نبرةٌ أو نبراتٌ دالة، مثلما تصْحَبه تعبيراتُ الوجه، وحركةُ الجسم واليدين، وقد تختلف طبيعة الكلام باختلاف المتخاطبَيْنِ أو المتخاطِبِينَ، وباختلاف الموضوع والمقاصد الحاصلة من وراء الكلام، وقد يكون الكلام عاميًّا أو فصيحًا، وقد يكون حَسَنًا أو قبيحًا، وقد يكون حقيقيًّا أو مجازيًّا، وقد يأخذ أشكالًا شتى يُمليها المقامُ والمقال[1].

إن اللسانيات العامة تهتم بدراسة الأنظمة العامة للألسن المتمثلة في المستوى الصوتي والصرفي و التركيبي و الدلالي بالاعتماد على نظريات علمية قادرة على استنباط قواعد هذه المستويات وتفسيرها تفسيرا علميا ثم تطبيقها على لسان معين قصد التحليل و التفسير والإلمام بقواعده النظامية ، ولئن كانت هذه اللسانيات بفرعيها النظري والتطبيقي تحظى بمجالات واسعة من الدراسة سواء كان ذلك عند الغرب أو عند العرب ، فإن المشكل المطروح على البحث اللساني العربي يتمثل أساسا في المنهج وتطبيق النظرية. إذ يقع الكثير من الباحثين في عملية الإسقاط دون وعي بمشاكل النظرية نفسها، وهذا ما يجعل الدرس اللساني العربي غير قادر على استيعاب النظريات الغربية ، وهو ما يعوق تطويره وتحديثه ، ورغم اجتهاد الباحثين العرب الجدد في تفسير المستويات المعهودة للسانيات العامة أو التطبيقية مثل المستوى الصوتي و الصرفي و التركيبي و الدلالي ، فإن هذه المستويات مازال أمامها أشواط كبيرة للحاق بما تتوصل إليه اللسانيات الغربية من تسارع في البحوث، وخاصة المرتبطة منها بالتركيب و الدلالة في علاقتها بالمعالجة الآلية و الحاسوبية[2] .

2.3. علم الحاسبات

يعد الحاسوب ناتجا من نواتج التقدم العلمي والتكنولوجي ، ويعرف بأنه جهاز الكتروني يستطيع ترجمة أوامر مكتوبة بتسلسل منطقي، لتنفيذ محاولة إدخال البيانات و إخراجها ، و إصدار عمليات حسابية أو

[1] – عبدالحميد النوري عبد الواحد. https://www.alukah.net/literature_language/0/128422/ تاريخ الزيارة 2018/01/18.

[2]– خليفة ميساوي، المصطلح اللساني وتأسيس المفهوم، ط.1 ،2013 ، ص 30.

منطقية، ويقوم بالكتابة على أجهزة الإخراج أو التخزين، ويتم إدخال البيانات بواسطة مشغل الحاسوب عن طريق وحدات لوحة المفاتيح أو استرجاعها من خلال وحدة المعالجة المركزية التي تقوم بإجراء العمليات الحسابية والمنطقية وبعد معالجة البيانات يتم كتابتها على أجهزة الإخراج مثل: الطابعات أو وسائط التخزين المختلفة[1].

إن جهاز الحاسب يستخدم إشارات "نبضات" كهربائية ، وحيث أن الإشارة الكهربائية لها حالتان عادة: إما وجود الإشارة أو عدم وجودها ، أو إن أردت قل إشارة موجبة أو سالبة ، لذا فإن تمثيل البيانات داخل الحاسب يكون باستخدام هاتين الحالتين ، إلا أن بيانات الحاسب تعتبر بيانات رقمية ثنائية ، بمعنى أنه يستخدم أحد الرقمين (1 .صفر) لتمثيل البيانات بحيث رقم(1) يمثل وجود الإشارة أو أنها موجبة بينما الصفر يمثل عدم وجود الإشارة أي سالبة ، لذا يقوم الحاسب بالتعامل مع البيانات على أساس تمثيلها بالأرقام الثنائية (BYTE) و التي تتكون من الأرقام (1.صفر) والتي تسمى (BITS) فمثلا لتمثيل حرف الهجاء (أ) على الحاسب فإن الحرف يمثل برقم ثنائي يتكون من ثمان بتات بالشكل الآتي :(11000110)[2].

3.3. علم الذكاء الاصطناعي :

الذكاء الاصطناعي هو قدرة الآلة على محاكاة العقل البشري وطريقة عمله، مثل قدرته على التفكير، والاكتشاف والاستفادة من التجارب السابقة، فهو بذلك علم يبحث أولا في تعريف الذكاء الإنساني وتحديد أبعاده، ومن ثمة محاكاة بعض خواصه ، وهنا يجب توضيح أن هذا العلم لا يهدف إلى مقارنة العقل البشري الذي خلقه الله جلت قدرته وعظمته بالآلة التي هي من صنع المخلوق ، بل يهدف هذا العلم الجديد إلى فهم العمليات

[1] وليد العناتي، " الدليل نحو بناء قاعدة البيانات للسانيات الحاسوبية " ،مجلة علوم اللسان و التكنولوجيا، العدد 15، الجزائر 2009، ص 84-85.

[2] صالح بن إبراهيم السدراني ، صالح بن عبد العزيز ، تجميع وصيانة الحاسب الآلي ، المنهج الصحيح لمن أراد صيانة الكمبيوتر و التصليح ، إعداد المادة العلمية وصياغتها. الإصدار 10 ص05.

الذهنية المعقدة التي يقوم بها العقل البشري في أثناء ممارسته (التفكير) ومن ترجمة هذه العمليات الذهنية إلى ما يوازيها من عمليات محاسبية تزيد من قدرة الحاسب على حل المشاكل المعقدة .

ويمكن تعريف الذكاء الاصطناعي للحاسب الآلي بأنه القدرة على تمثيل نماذج محاسبة (computer models) لمجال من مجالات الحياة وتحديد العلاقات الأساسية بين عناصره ، ومن ثم استحداث ردود الفعل التي تتناسب مع أحداث ومواقف هذا المجال ، فالذكاء الاصطناعي بالتالي مرتبط أولا بتمثيل نموذج محاسبي لمجال من المجلات ، ومن ثم استرجاعه وتطويره ، ومرتبط ثانيا بمقارنته مع مواقف وأحداث مجال البحث للخروج باستنتاجات مفيدة ،يتضح أن الفرق بين تعريفي الذكاء الاصطناعي و الذكاء الإنساني المذكورين أعلاه هو القدرة على استحداث النموذج ، فالإنسان قادر على اختراع وابتكار هذا النموذج في حين أن النموذج المحاسبي هو تمثيل لنموذج سبق استحداثه في ذهن الإنسان، وثانيا في أنواع الاستنتاجات التي يمكن استخدامها في النموذج ، فالإنسان قادر على استعمال أنواع مختلفة من العمليات الذهنية مثل الابتكار و الاختراع و الاستنتاج و أنواعه في حين أن العمليات المحاسبية تقتصر على استنتاجات لبديهيات وقوانين متعارف عليها يتم برمجتها في البرامج نفسها .و يرتكز أصل علم الذكاء الاصطناعي في أبحاث بحتة، ونظرية تدرس أساليب تمثيل النماذج في ذاكرة الحاسب الآلي وطرق البحث والتطابق بين عناصرها واختزال أهداف بها وإجراء أنواع الاستنتاجات المختلفة مثل الاستنتاج عن طريق المنطق أو عن طريق المقارنة أو عن طريق الاستقراء[1].

يعتبر علم الذكاء الاصطناعي علما حديثا نسبيا حيث نشأ أواسط القرن العشرين،ويمكن القول –بالمعنى الواسع– أن هذا العلم يهتم بإمكانية قيام منتج اصطناعي بنفس أنواع العمليات التي تميز التفكير الإنساني[2].

[1] الذكاء الاصطناعي و تقنيات المعلومات،الأكاديمية العربية البريطانية للتعليم العالي .ط د ت، ص 4-5 .www.abahe.co.uk

[2] د.فريال حاج حسن ، الذكاء الاصطناعي و النظم الخبيرة . محرم 1423 هـ،ص 2 .

4.3. علم المنطق :

المنطق ((قانون التفكير الصحيح)) أي : أن قواعد المنطق وقوانينه بمنزلة المقياس و المعيار و الميزان ، وكلما أردنا التفكير و الاستدلال في بعض الموضوعات العلمية أو الفلسفية تعين علينا أن نزن ونعرض استدلالنا على هذه المقاييس و المعايير ، لئلا نقع في الخطأ في استنتاجاتنا، فالمنطق بالنسبة إلى العالم و الفيلسوف نظير المسطرة أو الشاقول عند المعماري في عملية البناء ، يعرف بواسطتها أن الحائط الذي أقامه عمودي أم لا ؟ وأن سطحه أفقي أم لا؟

هكذا قالوا في تعريف المنطق: ((إنه آلة قانونية تعصم الذهن من الخطأ في التفكير))[1]. وقد نقل علم المنطق من النصوص اليونانية ، وواضع هذا العلم ومدونه هو **أرسطو طاليس** اليوناني.[2]

5.3. علم الرياضيات :

الرياضيات هي علم مواضيعهُ مفاهيمُ مجردةٌ ، ويعرَّف هذا العلم بأنه دراسة القياس و الحساب و الهندسة ، بالإضافة إلى المفاهيم الحديثة نسبيا ،مثل البنية و الفضاء و الفراغ و التغيير و الأبعاد ، و قد يعرّفه البعض بشكل عام بأنه دراسة البنى المجردة اعتمادا على المنطق و البراهين الرياضية و التدوين الرياضي ، وقد يعرف علم الرياضيات أنه دراسة الأعداد و أنماطها ، ومن التعريفات التي أطلقها بعض العلماء على هذا العلم كذلك أنه علم القياس .

وتعبر الرياضيات لغة العلوم التي لا تكتمل إلا عندما يتم تحويل نتائجها إلى معادلات ، وتحويل ثوابتها إلى خطوط بيانية ، ثم إن الاصطلاحات الرياضية تدل على الكم ، والعدد يدل على كمية المعدود ، و المقدار قابل للزيادة أو النقصان .

[1] الأستاذ المرتضى المظهري، المنطق، دار الولاء للطابعة و النشر ، لبنان ،ط 2 ،2011، ص12.

[2] المرجع نفسه ص 11.

ويمكن وصف الرياضيات بأنها علم لحل المسائل وتطوير النظريات ، و في هذه الحالة يُنظر للرياضيات عادة باعتبارها لغة عالمية ذات رموز وقوانين مشتركة ، بغض النظر عن بلد المنشأ،حيث يستطيع علماء الرياضيات فهم بعضهم من خلال تلك اللغة ، و التي هي علم حي ، كما يرى الكثيرون أن للرياضيات قيمتها الخاصة، تلك القيمة التي تعمل من أجل الرياضيات، و الرياضيات من العلوم التي برع فيها العرب و المسلمون ، حيث أضافوا إلى هذا العلم إضافات كانت جملة أسباب تطوره في العصر الحديث، والذي تقدم بفضل العرب خلال القرنين 09 و10 للميلاد ، فبعد أن اطلع العرب على حساب الهنود أخذوا عنه نظام الترقيم بدلا عن نظام الترقيم على حساب الجمل .

نشأت الرياضيات مع قياس الإنسان ما يشاهد من ظواهر الطبيعة، إذ بدأ من قياس قسمة الطعام بين أفراد العائلة، ثم وصل إلى قياس الوقت و الفصول والمحاصيل الزراعية ، و تقسيم الأراضي و الغنائم و المحاسبة، وعلم الملاحة بالنجوم و الاستكشاف ، و القياسات اللازمة لتشييد الأبنية و المدن ، أما البُنى الرياضية التي يدرسها الرياضيون فغالبا ما يعود أصلها إلى العلوم الطبيعية، وخصوصا علم الطبيعة، ومن المعروف أن للرياضيات دورا " بارزا " في عدة علوم وهي: الفيزياء و الكيمياء ، و البيولوجيا، بالإضافة إلى العلوم الإنسانية[1]

4. تحديات اللسانيات الحاسوبية :

التحدي الأول : هو تحدي تنوع الألسنة البشرية بصورة كبيرة جدا ، وذلك التنوع هو في الأساس جزء من طبيعة اللغة البشرية ، فلا يمكن إنكاره أو تجاهله ، فعند دراسة اللغة البشرية نجد أن البشر لا يتكلمون لغة واحدة ، بل لهم ألسنة مختلفة باختلاف ألوانهم وأجناسهم ، فضلا لو أننا أردنا التركيز على لسان واحد دون غيره فسنجد أن ذلك اللسان له الكثير من اللهجات و أيضا عند الافتراض بدراسة لسان معين فسنجد أن لكل لهجة عدد من المستخدمين ، وكل شخص يختلف عن الآخر في استخدامه التواصلي للسان ، ولو افترضنا دراسة

[1] أ. أشرف مبارز زنقور ، د.داود سليمان مشاط، تاريخ الرياضيات و إسهامات العرب و المسلمين ، مركز النشر العلمي، جامعة الملك عبد العزيز ،2010.ص 1-2.

لهجة معينة في لهجة معينة في لسان مجموعة من الناس ، فسنجد أن هناك فونيمات و تنوعات صوتية لها ، و المطلوب هو دراسة هذه التنوعات ، واستخراج القواعد العامة و استكشاف خواص كل لسان حتى يسهل معالجته آليا.[1]

التحدي الثاني : يتمثل في المعرفة ، إذ ينبغي التعرف على ما يدور في أذهاننا و عقولنا حول اللغة ، وكل ما يتصل بها من عمليات معالجة ، ويتمثل هذا في ملكة اللغة أو لدى المتحدث الأصلي للغة ، وليس هذا فحسب بل يجب معرفة المعلومات غير اللغوية التي تساعدنا في فهم اللغة، وإذا تمكنا من ذلك فسنصوغ تلك المعرفة في صورة قواعد آلية وندخلها الآلة حتى تعمل بصورة أقرب ما تكون للعقل البشري ، وعندما نتغلب على التحدي الأول و نستطيع أن نستخرج القواعد العامة للغة، فيمكننا التغلب على التحدي الثاني ، ونستطيع إدراك المعرفة البشرية في اللغة، و إنتاج تطبيقات لغوية آلية على قدر عال من الكفاءة و الدقة العلمية ، تكون قادرة على تحليل اللغة البشرية و توليدها و فهمها [2]

5. اللسان العربي و الحاسوب :

في إطار المعلوماتية تزداد أهمية معالجة اللسان العربي من خلال الحاسوب، كي نواكب ركب الحضارة، و لما كان الحاسوب مولدا غير عربي اللسان – فهو إنجليزي المولد – فقد اتخذت تقنيات الحاسوب ونظم المعلومات اللسان الإنجليزي أساسا لها، وفرض هذا الأساس الإنجليزي قيودا تقنية على الألسن الأخرى ، وكلما اتسع التباين بين الإنجليزية وبين الألسن الأخرى ازدادت حدة هذه القيود .

ويمثل اللسان العربي واللسان الإنجليزي من وجهة نظر الحاسوب طرفي نقيض ، و من هنا كانت العقبات الصعبة أمام تعريب الحاسوب وأصبح حاجز اللغة من أشق الحواجز على المستخدم العربي , وقد نجح العرب في

[1] د. محمد بن سالم سعد الله، علم اللغة الحاسوبي وحوسبة القرآن الكريم كلية الآداب جامعة الموصل، ط د ت، ص07.
[2] المرجع نفسه، ص07ـ08

تعريب جزء من الحاسوب على مستوى اللغة المكتوبة ، لكن مازال أمر اللغة المنطوقة يحتاج لشوط طويل من التعريب .[1]

الأمر الثاني أن تغذية الحاسوب فيما عُرب فيه لم تتم بالشكل المطلوب في إطار المعجمية الحديثة والمستويات الدلالية والصرفية والتركيبية للسان العربي , وتحتاج لجهد مشترك من طرف اللسانيين والحاسوبيين .

و من المشكلات التي تعترض معالجة اللسان العربي حاسوبيا نذكر:

1.5.*في المستوى الصوتي :*

الثنائيات الصوتية في اللسان العربي ، مثل :

ت ⟵ ط ذ ⟵ ز

د ⟵ ض ث ⟵ س ... الخ

و مما لا شك فيه أن تحديد القيمة الصوتية بدقة ووضوح لكل وحدة صوتية أمر مهم للغاية، كي نتفادى الاختلاط الحادث عند كثير من المتكلمين بين هذه الأصوات المتشابهة والتي يوجد بينها فارق صوتي بالتفخيم أو الترقيق.

أيضا هناك مشكل الكلمات ذات النهايات الصوتية الواحدة في حين يختلف الحرف الأخير من كلمة لأخرى، مثل: عصا، منى، هدى، سعى، دعا ...الخ ، فينبغي تحديد الزمن الصوتي الذي يميز مبنى متشابها مع مبنى آخر [2].

2.5.في المستوى الصرفي :

على مستوى الصرف، ليس هنالك ترتيب معهود ينظم الأفعال والأسماء ، والمجرد والمزيد... إلخ في ترتيب وتنظيم يخرجنا من عشوائية عرض كلمات المادة الواحدة في اللسان العربي، اللهم في بعض المحاولات المعاصرة.

[1] د. محمد داود، العربية وعلم اللغة الحديث، دار غريب , 2001, ص 273 .

[2] المرجع نفسه، ص277.

إن ضبط هذا العنصر يعد خطوة في تيسير اللسان العربي للتعامل مع الحاسوب .

3.5.في المستوى التركيبي :

يشمل هذا المستوى التراكيب الصغرى في اللسان العربي التي لا تكون جملة ، مثل : المضاف ،والمضاف إليه، والصفة والموصوف ...إلخ

كما يشمل أيضا التراكيب الكبرى التي تكون جملة(فعلية أو اسمية) ، ومسألة الوجوه المختلفة للإعراب ينبغي تقنيتها بشكل محدد ، كذلك ضبط أنماط الجملة في اللسان العربي ضبطاً يقوم على اعتبار الواقع ، ويتسم بالدقة والوضوح كي ننجح في تيسير وتطويع اللسان العربي للحاسوب [1].

4.5.في المستوى الدلالي:

أما المستوى الدلالي فيعد من أعقد الأنظمة اللغوية ، وأشدها تعصيا على جهاز الحاسوب ؛وذلك عائد إلى أن الدلالة من أقل المستويات اللغوية فيما يخص التباين اللغوي –كما يقول الدكتور نبيل علي ، كما أنه يشيع فيها عدة ظواهر تخرجها من واقع الاستخدام اللغوي وحقيقته إلى المجاز ، كالاستعارة ، والكناية، والتشبيه ، وهذا أمر يتطلب تحديد تلك التعابير غير الحقيقة وتصنيفها دلاليا بما يساعد النظام الحاسوبي على تمثيلها ، ومن ثم معالجتها آليا [2].

ويمثل المعنى مشكلة كبرى بالنسبة للنظم الآلية، فتعدد المعنى للكلمة الواحدة، وحساسية السياق في تحديد دلالة الكلمة واختلاف الدلالة باختلاف الثقافات...، كل ذلك يجعل المعالجة الآلية للدلالة تنطوي على مفارقات يصعب بسببها تمثيل هذا المستوى حاسوبيا ، وبسبب هذا تجاوزت أول دراسة صادرة عن اللسانيات الحاسوبية العربية الحديث عن المعالجة الآلية لعنصر الدلالة في العربية .

[1] د. محمد داود، المرجع السابق، ص277 ،278.

[2] د .عبد الرحمن بن حسن العارف، مقال توظيف اللسانيات الحاسوبية في خدمة الدراسات اللغوية العربية ، جهود ونتائج ، جامعة أم القرى

علما أن هذا لا يعني أن المعالجة الآلية لجانب الدلالة في اللسان العربي قد أغفلت تماما ، بل كان لها حضورها ضمن المستويات اللغوية الأخرى، كالمستوى الصوتي ، والصرفي والتركيبي والمعجمي ،وضمن قضايا لغوية ذات صلة وثقى بالدلالة، كالترجمة الآلية ، وهذا ما نلمسه في الجهود التي بذلت لتغطية هذا الجانب من اللسانيات الحاسوبية ، سواء كان ذلك في صورة بحوث نظرية ، أو برامج تطبيقية .[1]

5.5.في المستوى المعجمي :

إن العمل المعجمي شاق و مرهق لاسيما في مجال اللسان العربي لأسباب كثيرة ، و لسنا الآن بصدد الحديث عن روعة أساليبه ودقة معانيه و تعدد ألفاظه ، ومن ثم كان من الصعب أن يقوم بالعمل المعجمي فرد واحد أو مجموعة من الأشخاص ، إن معجم عصر المعلوماتية لا يصفه اللغويون فحسب ، و إنما يصفه العلماء في اللغة و المتخصصون في علوم كثيرة ، لذلك يجب تكوين فرق عمل من تخصصات المختلفة "فريق لغوي" و فريق منطقي وفريق بلاغي و فريق هندسي و فريق فلسفي ... الخ متعاونين مع أفراد يتقن الواحد منهم ضربا من الفن الذي لا بد منه لإخراج معجم يرضي أساطين اللغة ،وغيرهم من العلماء ،ويفيد الطلاب و سائر القراء. إن هذا الفرق يجب أن تعمل من خلال هيئة حكومية متخصصة ترصد المبالغ اللازمة لإتمام ذلك المعجم فأهمية اللسان العربي ، واكتشاف مكنوناته و الاستفادة منها و الحفاظ عليها لا تقل أهمية عن أية ثروة مادية أخرى مثل الذهب و البترول ،هذه الثروات مآلها للنفاد يوما ما بعكس اللسان العربي الذي سيظل محتفظا بمكانته دائما و على الرغم من صعوبة هذا العمل و المشاكل التي تحول دون حوسبة المستوى المعجمي أو صناعة المعجم اللغوي الحاسوبي أو تعيقه كظاهرة الترادف و التضاد و الاشتراك أو المشترك اللفظي ، و دقته فلن يحول حائل بين إمكانية صناعة المعجم الحاسوبي أو حوسبة المستوى المعجمي إن تضافرت جهود فرق العمل من تخصصات مختلفة[2].

[1] نفس المقال السابق .

[2] د.سلوى حماده ، المعالجة الآلية للغة العربية ،المشاكل و الحلول ، دار غريب ، 2009، 14-19.

6. جهود الباحثين العرب لحوسبة اللسان العربي :

إن قضية تعريب الحاسوب استدعت منذ أمد طويل اهتمام العديد من الاختصاصيين في المعلوماتية و الالكترونيات و اللغويات ، ذلك أنه بات من المتأكد أن يتكلم الحاسوب باللسان العربي أي أنه يصبح قادرا على التعامل مع حروف اللسان العربي كمدخلات و مخرجات و معالجات ، إلا أن المتتبع لقضية تعريب الحاسوب منذ أوائل السبعينيات يشعر بأن قضية إظهار الحروف على شاشة الحاسوب ثم إظهارها على الطابعات استدعى جانبا كبيرا من المجهودات المبذولة وطنيا و إقليميا و دوليا [1] .

وقد كانت العلوم الشرعية من أسبق العلوم الإنسانية استخداما لتقنية الحاسبات الالكترونية و نظم المعلومات ، حيث بدأ العمل بها و الإفادة منها في السبعينيات من القرن الماضي في مجال البحث في ألفاظ القرآن الكريم ، و علم الحديث ، ثم عممت هذه العملية على المعاجم في شكل دراسات إحصائية، مثل دراسة الجذور الثلاثية و غير الثلاثية لمعجم الصحاح للجوهري 324 هـ ، حيث التقى إبراهيم أنيس مع أستاذ الفيزياء علي حلمي موسى سنة 1971م .[2]

و بهذا الشكل بدأ تطويع تقنيات الحاسوب لخدمة الدراسات الخاصة باللسان العربي: أصواتا ، وصرفا، و تركيبا، و معجما و دلالة .

[1] استخدام اللغة العربية في المعلوماتية ، المنظمة العربية للتربية و الثقافة و العلوم، تونس، 1996 م، ص 19.

[2] عيسى مومني ، بيبليوغرافيا اللسانيات، ص 109.

7. وصف المستوى الصوتي للسان العربي لأغراض البرمجة الحاسوبية :

- **أولا: وصف المستوى الصوتي (الفوناتيكي):**

يؤول اللسان العربي في نظامه الصوتي إلى أربعة وثلاثين فونيما ، ثمانية و عشرين صامتا تتمثل في الحروف الأبجدية له ، و ستة صوائت تمثلها الحركات الثلاثة القصيرة الفتحة و الكسرة و الضمة ، و الحركات الثلاثة الطويلة حروف المد ، و يعني في الجانب الصوتي وصف تلك الأصوات آحادا على المستوى الأكوستيكي الفيزيائي بحيث يكون لكل صوت صورة طيفية مرئية ذات ثلاثة أبعاد : بعد أفقي يمثل الوقت و بعد عمودي يمثل التردد ، و بعد ثالث يمثل درجة الشدة ، يظهر في شكل سواد على ورق خاص.

و البادي أن وصف الأصوات المنفردة على هذا النحو سيفضي إلى نتائج دقيقة لا تقبل اللبس ، و أن حدودا فاصلة قاطعة تمثلها أرقام رياضية حاسمة ستميز كل صوت عن بقية الأصوات ، و هو واقع الأمر الذي أكدته التجارب المعملية على المستوى الصوتي إلى حد كبير، غير أن التجربة أيضا تؤكد رصد حالات تباين نسبي في أداء النطاقين بتلك الأصوات تمثل –أحيانا – صعوبة حقيقية عند تمثيل الأصوات للحاسوب ،بل إن أداء الفرد الواحد قد يتباين تبعا لحالته الصحية و النفسية[1] .

و يزداد أثر التباين بين الأفراد لدى التعامل مع الصوائت للاتفاق الواقع بين كل صائت قصير و مثيله الطويل الذي لا يفترق عنه إلا في الوقت الذي يستغرقه ، فقد يكون في أداء العجلان من نطق الصوائت الطويلة ما يلتبس على الحاسوب بكونه صائتا قصيرا، و قد يوهم الأداء المتأني للصوائت القصيرة استيعاب الحاسوب لها على أنها طويلة .

[1] عيسى مومني ، المرجع السابق ، ص 10.

إن ابن اللغة في تمييزه لتلك الأصوات – استيعابا و إنتاجا- لا يصدر عن وصف فيزيائي ذي مخرجات رقمية رياضية ، و العقل البشري لا يستطيع أن يمايز بين تلك الأصوات على تباينها النسبي بين الأفراد ، و قد لا يستعين ابن اللغة في الأغلب الأعم بما قدمه علماء الصوتيات من صفات للأصوات و مخارجها و خصائصها[1].

و في المقابل فإن الحاسوب لا يمتلك تلك المقدرة الفطرية لدى ابن اللغة، و قد ينبهم لديه ما يتلقاه من أصوات لا يوجد لها مقابل مستدخل في ذاكرته، لذا يتحتم أن نقدم للحاسوب وصفا رقميا دقيقا للمستوى الصوتي للسان العربي ينبئ عن ماهية الأصوات التي يتلقاها ،ويختلف هذا الوصف بالضرورة عن الوصف الذي رسمه علماء الأصوات لابن اللغة و للناطقين بغيرها .

و على الرغم من أن الناطقين باللسان العربي يتباينون في نطق بعض الأصوات تبعا لانتماءاتهم اللهجية ،كنطق أهل الشام الجيم معطشة ، ونطق أهل الجزيرة لها قليلة التعطش ، ونطقها لدى المصريين على نحو ما عرف بالجيم القاهرية ،إلا أن ذلك لا يعد عائقا في سبيل الوصف ،و قد يغنينا عن الخوض في جدل أفصحها قبولها جميعا بوصفها تنوعات صوتية لفونيم واحد يقابلها جرافيم (رمز كتابي) واحد [2] .

- **ثانيا : وصف المستوى الصوتي الوظيفي (الفونولوجي):**

تخضع فونيمات اللسان العربي لقواعد فونولوجية تحكم تتابعها في سياق الكلمة أو الجملة ، وقد يطرأ على الفونيم تغيير في صفته تأثرا بما يسبقه أو ما يليه من الفونيمات ، ويعني ذلك أن الوصف الفوناتيكي لأصوات اللسان العربي لا يكفي لوحده ليمكن الحاسوب من تمييز الفونيمات ، و أنه يحتاج إلى وصف آخر لما يعتري فونيمات اللسان العربي – في سياقها الوظيفي – من تغييرات تحكمها القواعد الفونولوجية على مستوى التشكيل الصوتي ، فالحرف المضعف مثلا تزيد المدة التي تستغرق نطقه ، و الدال في (أردت) تنطق تاء بسبب التأثير الرجعي ، و الباء في (سبت) يصيبها من أثر التاء المهموسة ما يزيل عنها صفة الجهر ، ولفظ الجلالة (الله) بعد غير الكسر تكون فيه

[1] د.وجدان محمد صالح كنالي ، اللسانيات الحاسوبية العربية الإطار و المنهج، ص 10.

[2] المرجع نفسه ، ص 10-11.

اللام مفخمة من قبيل التنوع الصوتي ، وكلمتا (سوط – صوت) قد يتطابقان في سياق الأداء النطقي ،ولام (ال) لا تنطق إذا تلتها إحدى الحروف الشمسية [1].

إن تصميم برنامج حاسوبي يستدخل الحدث الكلامي المنطوق ويحوله إلى نص مكتوب يتطلب الوقوف على تلك المواضيع التي تتغير فيها صفات الفونيمات تأثرا بما يجاورها على مستوى الكلمة، وإذا ما انتقلنا إلى مستوى الجملة لزمنا الانتباه إلى ما ينطق من الفونيمات ولا يكتب ،كالتنوين و المد في (هذا وأخوتها – الله–لكن) ، وكذلك ينبغي التحرز من الفونيمات التي تكتب ولا تنطق كهمزة الوصل و اللام الشمسية وألف (مائة) و التاء المربوطة عند الوقف و الألف الفارقة بعد واو الجماعة ، وسنفيد بالطبع من المخرجات ذاتها و إذا ما أردنا تصميم برنامج يتعرف المكتوب ويقرؤه ، ويلزمنا عندئذ شرط إضافي يقيد النص المستدخل ضبطا تاما بالشكل ؛ لكن افتراض أن يكون النص المكتوب خاليا من حركات المبني و علامات الإعراب غير مشكول–كما هو الحال في معظم النصوص العربية – يحتم علينا اللجوء إلى ثبت منطوق يودع في ذاكرة الحاسوب يجعل إزاء كل مادة معجمية مقابلا منطوقا للائتلاف مع ما يعتريه من اللواصق القبلية و البعدية ، ويكون اللفظ المحتمل لأكثر من وجه مقابلات منطوقة (بدائل) يمثل أكثرها ورودا للخيار الافتراضي، ثم يترك المحددات والأدلة السياقية الشكلية تعين إحدى تلكم البدائل ، و أما علامات الإعراب فيعوزها برنامج حاسوبي آخر تتكامل فيه النظم المورفولوجية و النسقية على نحو ما يسرد في وصف الجانب الإعرابي للسان العربي [2].

وحقيق بالذكر هنا الإشارة إلى ما يواجه الحاسوبيين في استدخال النصوص المطبوعة عبر الماسح الضوئي (scanner) ، و أنه ثمة حروفا يشكُلُ على الحاسوب معرفتها ، ويلجأ الحاسوبيون من أجل معالجة تلك المشكلات آليا إلى تزويد الحاسوب بثبت لما يتقارب من الحروف رسما ، وثبت آخر يرصد تتابع الحروف الأكثر شيوعا مرتبة حسب النسب المئوية لدرجة الشيوع.

[1] د. وجدان محمد صالح كنالي، اللسانيات الحاسوبية العربية،الإطار و المنهج ، ص 11.
[2] المرجع نفسه، ص 11.

أما إذا كان مطلبنا تمكين الحاسوب من تحديد الأخطاء الإملائية و الطباعية و تصويبها ، أو تقديم بدائل لها فسيسعفنا في ذلك ما تمخضت عنه الدراسات الإحصائية لما لا يتألف من الحروف العربية ، و يعيننا كذلك استقراء ما توصلت إليه الدراسات اللغوية من الأخطاء الإملائية الشائعة لدى من يكتبون بالعربية لتكون مرجعا هاديا في التصويب الآلي[1].

8. أثر العولمة على اللسانيات الحاسوبية:

ليس خافياً على أحد ما أصاب العالم- ومنه العالم العربي- من مدّ عَوْلمي حمل معه كثيراً من التغيرات السياسية والاقتصادية والاجتماعية والثقافية والتقنية. ولا شك أن العولمة قد مازت نفسها مِنَ الاستيطان القديم الإكراهي؛ إذ أطلقت العنان لأدواتها لتنشر أنماطها السياسية وقيمها الاجتماعية، وأطرها الثقافية، ونظمها الاقتصادية، هذه الأدوات هي: الفضائيات والإنترنت والهواتف النقالة... إلخ.

و لما كانت الولايات المتحدة الأمريكية هي قطب العولمة الأوحد والمسيطر على ذلك كله، فقد هيأت للغتها الإنجليزية فرصاً ذهبية لتنتشر على نحو غير مسبوق، أكان ذلك لا إرادياً أم بالإكراه والجبر الخفيّ. وبيان ذلك أن معظم المعرفة تُنْتَجُ باللغة الإنجليزية، وتُنْشَرُ وتُوَزَّعُ بها ورقياً أو إلكترونياً[2]، و لما كانت حركات الترجمة في العالم غير قادرة على ملاحقة الانفجار المعرفي العولمي، صار الأفراد و كذا الحكومات ينحون إلى التعامل بهذه اللغة. فصارت الإنجليزية لغة الرقي والمكانة الاجتماعية والمستوى الاقتصادي الرفيع.

و لما كانت العولمة تعتمد الحاسوب واستخداماته في نشر ثقافتها ولغتها، وصارت الإنجليزية هي المهيمنة على صفحات الإنترنت، وصارت اللغة الأولى في العالم[3]، وهددت اللغات المحلية والوطنية، فقد أدى ذلك إلى

---

[1] د. وجدان محمد صالح كنالي، المرجع السابق، ص 12.

[2] انظر بعض الإحصاءات المتعلقة باللغة الإنجليزية وهيمنتها في مجال الاتصالات والإعلام: نبيل علي، الثقافة العربية وعصر المعلومات، ص 273.

[3] انظر في تفصيل هذه المسألة : David Crystal, English as Global Language

استنهاض معظم دول العالم للنظر في كيفية مواجهة هذا الخطر اللغوي والثقافي، وصار الناس يتشبثون باللغة كونها مكوناً مهماً من مكونات الثقافة الوطنية التي ينبغي صيانتها والحفاظ عليها[1]، ومن بين هؤلاء كان العرب.

وقد ظهر لكثير من الناس ضرورة مواجهة الغزو العولمي بالسلاح نفسه، أي بوسائل الاتصال الحديثة ولا سيما الجانب المعلوماتي. وهكذا كانت اللسانيات الحاسوبية موئلاً للعرب لمواجهة العولمة واللغة الإنجليزية.

تشير الدراسات إلى أن بدايات اللسانيات الحاسوبية العربية كانت بأيادٍ أجنبية، ولأهداف تجارية خالصة، إذ كانت البلاد العربية سوقاً رائجة لتقنيات الحاسوب وبرامجه. ولكن الأمر بدأ يسير في الاتجاه الصحيح بعد ذلك. وبدأت برمجة الحواسيب بالعربية تتحرر شيئاً فشيئاً من الأيادي الأجنبية، و كان ذلك ماثلاً في:

- اشتغال الشركات العربية المؤسَّسة في الخارج (كندا وفرنسا) بحوسبة العربية.
- اعتماد هذه الشركات على عقول عربية، من اللسانيين والحاسوبيين.
- الاستثمار في قطاع حوسبة العربية في البلاد العربية.

ثم تسارعت مظاهر الاهتمام باللسانيات الحاسوبية وحوسبة العربية مع تسارع المد العولمي، وضرورة تيسير دورة الحياة العربية، وكانت هذه مرحلة مهمة من مراحل الحوسبة في العالم العربي، وأبرز مظاهرها:

- الاهتمام الحكومي المؤسسي باللسانيات الحاسوبية ومعالجة العربية آلياً بعقد الندوات والمؤتمرات المتخصصة وإنشاء مراكز البحوث التقنية المدعومة والممولة من الحكومات العربية.
- الاتجاه نحو التخصص الأكاديمي في هذا المجال، دراسة وتدريساً، لتوطينه في نفوس الدارسين أولاً، ثم في البيئة العربية ثانياً[2].

[1] مؤتمر اللغة العربية أمام تحديات العولمة، معهد الدعوة الجامعي للدراسات الإسلامية، بيروت، 2002.

[2] وليد أحمد العناتي, "مقال اللسانيات الحاسوبية العربية رؤية ثقافية" ص2.

وهكذا شهدت اللسانيات الحاسوبية العربية تطورات ملحوظة، وصارت خدمة اللغة العربية هدفاً أساسياً ومهماً من أهداف معالجة العربية آلياً. ولما وجد العرب –كغيرهم من الأقوام– في اللسانيات الحاسوبية وتطبيقاتها سلاحاً مهماً في مواجهة هيمنة الإنجليزية بدأوا في هذا على صعيدين:

أولهما: استخدام العربية في تصميم الحاسوب، بتعريب البرامج ولوحة المفاتيح، والطباعة العربية، بل تجاوزوا ذلك إلى ابتكار لغات برمجة عربية، وتصميم حواسيب خاصة تتعامل مع العربية على التعيين."فقد نجحت عديد الشركات العربية بالتعاون مع شركات أجنبية في صناعة الحواسيب، وتعريب لغات برمجة، وهكذا برزت لغة"نجلاء" وهي لغة "بايزك" عربية تشتغل على الحواسيب من نوع الفارابي من إنتاج الصناعات الحاسوبية السعودية، ولغة"الخوارزمي"وهي لغة مقتبسة عن"البايزك" أيضاً وتشتغل على الحواسيب من سلسلة "الرائد" و "بايزك" الحاسوب"صخر"من إنتاج الشركة العالمية[1].

ثانيهما: النشر الإلكتروني باللغة العربية. فكان أن انتشرت مواقع كثيرة تنشر صفحات هائلة بالعربية، على الصعيد الشخصي أو المؤسسي(الصحف، دور النشر، الجامعات، الدوريات، المؤسسات الحكومية).

ويؤمِّل هذا النشر تحقيق أهداف لغوية وسياسية وثقافية واجتماعية متعددة.

9. التعرف الآلي على الكلام العربي المنطوق :

اللغة منظومة اجتماعية و لكنها تتجسد في انتاجات فردية لولاها لما كانت اللغة حية, هذه الانتاجات قد تأخذ أشكالا مختلفة: خطاب , درس,قصيدة, شعر, رواية... هذه الانتاجات الفردية تسمى كلاما و لا يشترط في الكلام أن يكون منطوقا, فهو قد يأخذ الشكل المكتوب بأي طريقة من الطرائق: كتابة عادية, كتابة صوتية, كتابة بلغة المورس أو أي شكل من أخر كأن يعبر عن الكلمات بواسطة إشارات[2].

[1] محمد بن أحمد، اللغة العربية والنظم الحاسوبية والبرمجيات، ضمن كتاب: استخدام اللغة العربية في المعلوماتية، المنظمة العربية للتربية والثقافة والعلوم ، تونس ،1996 ، ص.133

[2] مصطفى حركات، اللسانيات العامة و قضايا العربية ،المكتبة العصرية صيدا بيروت ط1. 1998م

ما هو التعرف الآلي ؟ التعرف الآلي بأبسط أشكاله هو عبارة عن برنامج يستقبل المدخلات الصوتية عن طريق الميكروفون و يقوم بتحليلها و التعرف الصحيح عليها بحيث يمكن تحويل الصوت إلى أشكالٍ أخرى يمكن الاستفادة منها في مجالات متعددة، كما يمكن تعريف التعرف الآلي على أنه تقنية استقبال الكلام المنطوق من وسائط متعددة ثم تحليله و تحويله إلى أشكال أخرى يمكن الاستفادة منها في مجالات متعددة.و هذا التعريف أدق من التعريف الأول.

هنالك طريقتين أساسيتان في بناء القوالب للتعرف على الكلام : إما عن طريق تدريب النظام على الصوت الذي سوف يدخل إليه كلمة كلمة ، و بهذه الطريقة يتعرف النظام فقط على الكلمات التي تم إدخالها مسبقا, أو بطريقة الصوتيات بحيث يتم إدخال قائمة بأكبر قدر من الصوتيات, بحيث يعطينا تركيبها أكبر قدر من الكلمات المتعرف عليها من قبل النظام. وهذا يناسب اللغة العربية أكثر من اللغات الأخرى ، لأن الصوتيات العربية للنطق مخارجها أكثر من اللغات الثانية مما يسهل على النظام التفريق بين المدخلات.

عند بناء القوالب يتم التركيز على النقاط الآتية في اللغة العربية:

الحروف الساكنة وحروف العلة ، حيث أن حروف العلة من السهل التعرف عليها أكثر من الحروف الساكنة لأن طاقتها أقل.

أنماط المقاطع : في اللغة العربية هنالك خمس أنماط للمقاطع هي:

CV,CVV,CVC,CVVC, CVCC بحيث C يرمز للصوت الساكن و V للمتحرك و VV للصوت المتحرك الطويل. بحيث كل مقطع في اللغة العربية يكون لديه صوت ساكن والصوت الثاني في المقطع يكون متحركا ويكون أيضا أكثر من ساكنين في المقطع . لذلك فهذه القيود يمكن الاستفادة منها بصورة جيدة في إيجاد خوارزميات ممتازة لبناء أنماط المقاطع[1] .

[1] فهد العتيبي ، رسالة ماجستير(التعرف على الكلام العربي لمتحدث واحد باللغة العربية), سنة الطباعة 1422.

10. التحديات التي تواجه برامج التعرف الآلي على الكلام العربي:

تواجه برامج التعرف الآلي على الكلام العربي مشاكل عديدة بسبب انتشار اللغة العربية الواسع ، وبسبب تعدد اللهجات العربية، و كذلك بسبب صعوبة قواعد اللغة العربية. إن اللغة العربية تُعد سادس أوسع لغة في الانتشار ، فعدد متحدثي اللغة العربية 250 مليون متحدث, منهم 195 مليون شخص تعدّ العربية هي لغتهم الأم. وكذلك اللغة العربية هي اللغة الرسمية في أكثر من 22 دولة. وهذا العدد الكبير من الدول العربية يزيد من عدد اللهجات العربية. فعلى سبيل المثال قسّم بعض الباحثين اللغة العربية إلى عدة لهجات، فأولها العربية القديمة وهي اللغة الفصحى التي تتمثل في القرآن الكريم ، وثانيها وأكثرها انتشارا العربية الحديثة وهي اللغة العربية الفصحى المستخدمة في الجرائد والبرامج التلفزية والمجلات و المؤتمرات. وهناك أيضا العديد من اللهجات العامّية المنتشرة بين الناس، و التي يمكن تقسيمها إلى: اللهجة المصرية ، و اللهجة الخليجية, و لهجة بلاد الشام, و لهجة المغرب العربي[1]. وبالطبع فإن هذا التقسيم هو مجرد تقسيم عام للهجات العربية, ولكن في الحقيقة هناك أكثر من هذه اللهجات في اللغة العربية, فبالدولة الواحدة هناك أكثر من لهجة(غالبا). هذه الصعوبات كانت فقط بالنسبة للغة المنطوقة ، ولكن الأمر يزداد صعوبة في بعض التطبيقات فعلى ، سبيل المثال في الإملاء الصوتي تدخل صعوبات اللغة العربية المكتوبة إضافة إلى صعوبات اللغة المنطوقة. ومن هذه الصعوبات في اللغة المكتوبة, أن بعض الحروف في اللغة العربية تنطق و لا تكتب مثل الألف في الكلمات التالية (لكن, هذا, ذلك...وغيرها) وكذلك هناك بعض الحروف التي تكتب و لا تنطق مثل (همزة الوصل اللام الشمسية ، الواو في عمرو...وغيرها) وهناك حروف تنطق بخلاف ما تكتب مثل نون الإقلاب (منْ بعد تنطق ممبعد!). كذلك هناك تداخل في بعض الكلمات بين علامات الضبط بالشكل وحروف العلة, والتنوين والنون الساكنة. وإضافة إلى ذلك, الحروف العربية يتغير شكلها حسب موقعها من المقطع.

[1] Novel Speech Recognition Models for Arabic http://www.clsp.jhu.edu/ws2002/groups/arabic/

وزيادة على ذلك يجب علينا ألّا ننسى أن أهم العقبات التي تواجه برامج التعرف الآلي العربية هي قلة الدراسات والأبحاث باللغة العربية ، و كذلك عدم توفّر قاعدة بيانات كافية للصوتيات العربية, فأغلب قواعد البيانات هي مجرد اجتهادات شخصية[1]. وبالطبع لا يخفى علينا أن وجود قاعدة بيانات كاملة وكافية هي لبنة الأساس في التعرف على الكلام الآلي. وكل هذه الأسباب تقف عقبة أمام تطور التعرف على الكلام العربي آليا وتطوّر تطبيقاته, وتصعّب من مهمة الباحثين والعلماء في هذا المجال.

10. الفرق بين التعرف على الكلام والتعرف على الصوت:

إن الفرق بين التعرف على الصوت و التعرف على الكلام هو فرق جوهري، فالتعرّف على الكلام يهتم بالتعرف على اللغة نفسها و كلمات اللغة بغضّ النظر(غالبا) عن المتحدّث ، أمّا التعرف على الصوت فيهتمّ بالتعرف على المتحدّث نفسه و تمييزه من بصمة صوته عن غيره من المتحدثين. وأمّا بالنسبة لمجال التطبيقات والاستخدامات، فالتعرف على الكلام أوسع انتشارا و أسهل في التمثيل من التعرف على الصوت, فالتعرف على الصوت استخداماته تقريبا محصورة في مجال الأمن ، أما التعرف على الكلام فهو يدخل في معظم أمور الحياة . و هذا يمكن تأكيده عن طريق بعض التوقعات التي نشرت في مجلة **بي بي سي** العربية حيث توقّع بعض الباحثين أن تكون حصة السوق بالنسبة للتعرف على الكلام 96.2%, وأن تكون حصة التعرف على الصوت في السوق 0.7%, وباقي الحصة لبرامج تحويل النص إلى كلام.كما نرى في هذه الإحصائية أن برامج التعرف على الكلام ستنتشر أكثر بنسبة 137 مرة تقريبا من برامج التعرف على الصوت[2].

و بما أنّ التعرف على الصوت يكاد يكون محصورا في مجال الأمن, هذا يترتب عليه صعوبة التمثيل, لأنه يتطلب دقة أكبر في التطبيق, ونسبة أقل في الأخطاء. أما في مجال التعرف على الكلام, فقد لا يتطلب دقة عالية

[1] فهد العتيبي مرجع سبق ذكره.

[2] http://www.pcmag-arabic.com/news/news.php?id=EpFZlyEAllCMTvwyuW

كما يتطلبها التعرف على الصوت، وقد يسمح بنسبة خطأ أكبر من التعرف على الصوت. فهل سيتقدّم العلم إلى درجة يصل فيها التعرف على الصوت في الكمبيوتر الشخصي إلى مستوى من الدقة يستطيع من خلاله التفرقة بين أفراد الأسرة(مثلا) و استرجاع معلومات كل شخص على حدى فيحل الصوت مكان الكلمة السرية؟

12. الفرق بين التعرف الآلي على الكلام في اللغتين العربية و الإنجليزية:

لكل لغة في العالم خصائصها التي تميزها عن غيرها من اللغات، من حيث النطق والكتابة.و هذه الخصائص هي التي تشكل فرقا بين مختلف اللغات بالنسبة لأنظمة التعرف الآلي على الكلام, أمّا التقنيات المستخدمة لإنشاء هذه الأنظمة فهي نفسها في كل اللغات تقريبا. فمن الفروق بين اللغة العربية واللغة الإنجليزية أن اللغة العربية مخارج حروفها أوضح و أوسع من اللغة الإنجليزية, فاللغة العربية مخارج حروفها تبدأ من الجوف وتنتهي بخارج الشفاه، أما اللغة الإنجليزية فمخارج حروفها متقاربة. و هذه الميزة في اللغة العربية تجعل اللغة العربية أوضح وأسهل للتعرف الآلي على كلماتها. ومن الفروق أيضا أن اللغة العربية أبطأ في القراءة من اللغة الإنجليزية, وبرأيي أن هذا الفرق يجعل التعرف على الكلام العربي أسهل من الكلام الإنجليزي بحيث أن معظم برامج التعرف الآلي على الكلام تطلب من المستخدم أن يتكلّم ببطء حتى تستطيع فهمه.

و لكن مازالت برامج التعرف الآلي على الكلام الإنجليزي أفضل وأدق من برامج التعرف على الكلام العربي بالرغم من الميزات التي ذكرناها أعلاه، و ذلك لتأخر البحوث العربية في هذا المجال عن البحوث الإنجليزية ضف إلى اللهجات العربية أكثر بكثير من اللهجات الإنجليزية وهذا سبب آخر لتفوّق البرامج الإنجليزية على البرامج العربية.

13. بعض البرامج الموجودة للتعرّف الآلي على الكلام العربي:

من أشهر البرامج في التعرف الآلي على الكلام العربي: برنامج صخر للإملاء الصوتي (ASR), وبرنامج IBM Via Voice للإملاء الصوتي العربي. وفيما يلي شرح لبعض مزايا هذه البرامج:

- IBM Via Voice . :

هو من أقوى و أشهر البرامج في التعرف على الكلام العربي. و هو برنامج يقوم بتحويل النصوص المنطوقة إلى كلام مكتوب باللغة العربية.

ونُشر عنه بأنه يصل إلى نسبة دقة تتراوح بين 92-97%. وقد تم تطويره و إنشاؤه في فرع شركة IBM في القاهرة تحت أيدي خبراء مصريي الجنسية.

- **صخر 4.0** ASR :

هو برنامج للإملاء الصوتي باللغة العربية أيضا. و أحد أهم مزاياه أنّه يدعم عدة لهجات عربية مثل اللهجة المصرية واللهجة الخليجية. وقد نُشر عنه بأنه يصل إلى مستوى دقة تزيد على 95% .

وهناك بعض الجهود القائمة في مجال التعرف الآلي على الكلام العربي، و من هذه الجهود مشروع تقوم به مدينة الملك عبد العزيز بالتعاون مع شركة IBM باسم (مشروع التخاطب مع الحاسوب باللغة العربية عبر الهاتف) و أيضا مشروع تعاونت فيه شركة فيليبس مع شركة اتصالات الإمارات باسم (أورينتل) ، ويهدف هذا المشروع عند اكتماله إلى إيجاد نظام تعرّف صوتي فعّال و عملي باللغة العربية لاستخدامه في مختلف المجالات.

14.الخاتمة

علم اللسانيات الحاسوبية هو علم حديث و يُعد من أبرز العلوم اللغوية، التي ظهرت في العصر الحديث. يتكون هذا العلم، من عنصرين أساسيين، **أولهما اللسانيات،** وهو العلم الذي يدرس اللغات الطبيعية الإنسانية في ذاتها ولذاتها، سواء أكانت مكتوبة منطوقة أم منطوقة فقط. ويهدف هذا العلم أساساً إلى وصف أبنية هذه اللغات، وتفسيرها، واستخراج القواعد العامة المشتركة بينها، والقواعد الخاصة التي تضبط العلاقات بين العناصر المؤلفة لكل لغة على حدة .**وثانيهما الحاسوبية،** ويقصد بها، توظيف الحاسوب، بما يحتويه من إمكانات رياضية خارقة، وسعة تخزينية هائلة، في خدمة اللغة.

من خلال بحثنا في هذا الميدان يمكن ان نوجز ما توصلنا اليه في النقاط التالية:

- اللسانيات الحاسوبية دراسة علمية موضوعية للغة البشرية من خلال الألسنة الخاصة بكل قوم، وذلك باستخدام الحاسوب، أي دراسة ظواهر وحقائق وخصائص الألسنة آليا،أو معالجتها حاسوبيا.

- اللسانيات الحاسوبية فرع تطبيقي حديث، يستغل ما توفره التكنولوجيا المتطورة من أجل بلورة برامج وأنظمة لمعالجة اللغات الطبيعية معالجة آلية.

- اللسانيات الحاسوبية تتأسس على علوم شتى (اللسانيات العامة،علم الحاسبات الإلكترونية (الكمبيوتر)، علم الذكاء الاصطناعي،علم المنطق،وعلم الرياضيات)، مما جعل الباحثين يختلفون في تحديد منهج خاص واضح لها.

- اللسانيات الحاسوبية العربية نبتت على غرار اللسانيات الحاسوبية العامة استجابة لدواعي حضارية وإستراتجية ينشدها مستقبل اللسان العربي.

الفصل الثاني:

تحليل الإشارة الكلامية

1.المقدمة

أثار موضوع الواجهات 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 منذ عشرات القرون 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 التعرف الآلي على الكلام speech recognition.

و للتعرف الآلي على الكلام تطبيقات كثيرة مثل: إعطاء أوامر صوتية و تحكمية كطلب رقم هاتف صوتياً، و إدخال معطيات صوتية مثل رقم بطاقة الائتمان، و إعداد وثائق بنيوية كالوثائق الطبية...وغيرها الكثير[1].

و يُعد تعرّف الكلام إجرائية تُحوِّل الإشارة الصوتية، الملتَقَطة من هاتف أو ميكرفون إلى مجموعة كلمات؛ قد تكون هذه الكلمات هي النتيجة المرجوة النهائية، كما في التطبيقات المذكورة سابقاً، وقد تخدم كدخل إلى نظم معالجة لغوية لاحقة للحصول على فهم الكلام لإعطاء تفسير لهذا الكلام والتصرف على أساسه. وقد صار لدينا خوارزميات ونظم تعرف كلام عالية الأداء.

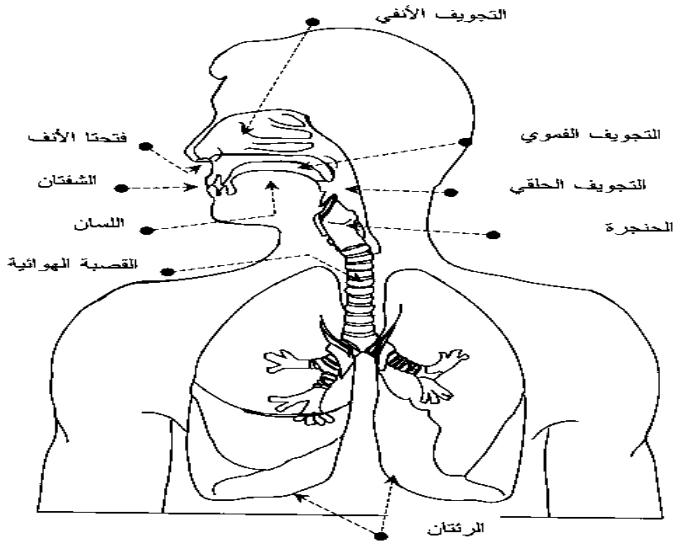
2.الجهاز الصوتي عند الانسان:

درس علم الأصوات الحديث جهاز النطق دراسة دقيقة مفصلة، وقد أفاد علماء الأصوات من علم وظائف الأعضاء، وعلم التشريح، وفيزياء الصوت، وساعدتهم الأجهزة الحديثة من آلات تصوير وتسجيل وغيرها من تحديد أعضاء النطق والأصوات التي تصدر عن طريق هذه الأعضاء والنقاط التي تحصر عندها أو تحتك بها تلك الأصوات.

[1] أميمة الدكاك, "نظام تعرف الكلمات المعزولة باستخدام نموذج ماركوف المخفي و استخدامه لإعطاء أوامر صوتية", المعهد العالي للعلوم التطبيقية و التكنولوجيا، دمشق 2005.

و يقصد بالجهاز النطقي لدى المحدثين من علماء اللغة مجموع أعضاء النطق المستقرة في الصدر والعنق والرأس ، ويسميه بعض اللغويين "بالجهاز الصوتي."

إن جهاز النطق عند الإنسان هو نفسه جهاز التنفس والذي له دور أساسي ، و هو دور بيولوجي يتمثل في عملية التنفس التي يضمن بها الإنسان الحفاظ على حياته ، بالإضافة إلى دور اللسان في التذوق و الأنف في الشم ، و لهذا تعتبر أهمية هذه الأعضاء في النطق ثانوية بالمقارنة بأهميتها البيولوجية و التي تتمثل في الحفاظ على حياة الإنسان[1].



[2]الشكل 1.2 : الجهاز الصوتي و الجهاز التنفسي

ويتكون جهاز النطق من الأعضاء الآتية: الرئتان، القصبة الهوائية، الحنجرة، الوتران (الحبلان) الصوتيان، الحلق، تجويف الفموي يحتوي على (اللسان ، اللهاة، الحنك الأعلى ، الأسنان، اللثة، الشفتان)، التجويف الأنفي (الخيشوم) .

[1] إبراهيم أنيس، الأصوات اللغوية، مكتبة الانجلو المصرية، مصر، 1999 م.

[2] مسعود بن محمد دادون.دروس في اللسانيات العامة – المستوى الصوتي

إن إحداث الأصوات اللغوية في الموقف الكلامي العادي يتطلب مجموعة من الآليات الفيزيولوجية. و قبل ذلك نلتفت إلى تعريف الصوت.

الصوت هو تردد آلي ، أو موجة قادرة على التحرك في الهواء في عدة أوساط مادية مثل الأجسام الصلبة و السوائل و الغازات، و لا تنتشر في الفراغ. و باستطاعة الكائن الحي تحسسه عن طريق عضو خاص يسمى الأذن.و من منظور علم الأحياء فالصوت هو إشارة تحتوي على نغمة أو عدة نغمات تصدر من الكائن الحي الذي يملك العضو الباعث للصوت ، تستعمل كوسيلة اتصال بينه و بين كائن آخر من جنسه أو من جنس آخر، يعبر من خلاله عما يريد قوله أو فعله بوعي أو بغير وعي مسبق ، و يسمى الإحساس الذي تسببه الذبذبات بحاسة السمع ، و الأصوات العادية التي يمكن للإنسان سماعها هي تلك التي تحدثها ذبذبات يتراوح ترددها بين 16 هرتز (عتبة السماع)، و16000 هرتز (عتبة الألم)[1].أما الأصوات الأدنى من عتبة السماع فهي ما تحت الصوت (infrasons)، و أما تلك التي هي أعلى من عتبة الألم فهي ما فوق الصوت (ultrasons).

هناك أعضاء معينة تتدخل في عملية الكلام يطلق عليها أعضاء النطق (وهي تسمية من باب التوسع و المجاز إذ تقوم هذه الأعضاء بالدرجة الأولى بوظائف عدة مثل التنفس للرئتين، والشم للأنف، والذوق للسان، والقضم للأسنان، أما الكلام فهو وظيفة ثانوية تقوم بها هذه الأعضاء)، وتتخذ هذه الأعضاء أوضاعا معينة محددة في أثناء النشاط الكلامي وتتطلب من المتكلم مجهودا[2].

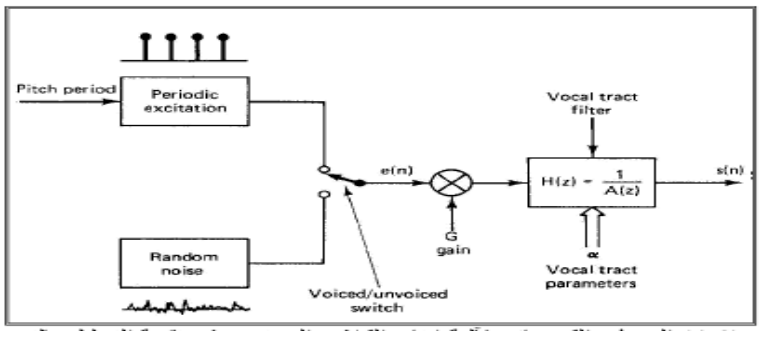
[1] مسعود بن محمد دادون, المرجع سابق.

3. النموذج الكهربائي لآلية إنتاج الكلام

إن آلية إنتاج الكلام تبين أن الأوتار الصوتية تهتز عند إنتاج الأصوات المجهورة (اهتزاز شبه دوري) ، و لا تهتز عند الأصوات المهموسة, أما إنتاج الأصوات الاحتكاكية فيترافق دائما بوجود اضطراب في تدفق الهواء ، و هو ما يكافىء كهربائيا إشارة ضجيجية ، و تغيرات شكل الأنبوب الصوتي عند إنتاج الأصوات المختلفة يعزز بعض الترددات عن البعض الآخر. لذلك جرت عملية نمذجة عملية إنتاج الكلام كهربائيا بمرشح يمثل أثر الأنبوب الصوتي[1]، إشارة دخل هذا المرشح عبارة عن مزيج من قطار نبضات دورية (تمثل اثر اهتزاز الأوتار الصوتية) و ضجيج يمثل اضطراب تدفق الهواء لبعض الأصوات، و يوضح الشكل (2.2) النموذج الكهربائي لآلية إنتاج الكلام المستخدم في تقنية تحليل الإشارة بالتنبؤ الخطي(يتم شرحها لاحقا) حيث :

- المصدر هو قطار من النبضات بدور F/1+T, F يسمى تردد اهتزاز الأحبال الصوتية إضافة إلى إشارة الضجيج random noise

- المرشح : ينمذج أثر الأنبوب الصوتي ، و هو مرشح رقمي بأقطاب فقط من دون أصفار ، و يستخدم في تقنية التحليل بالتنبؤ الخطي.



الشكل 2.2 :النموذج الكهربائي لآلية انتاج الكلام.

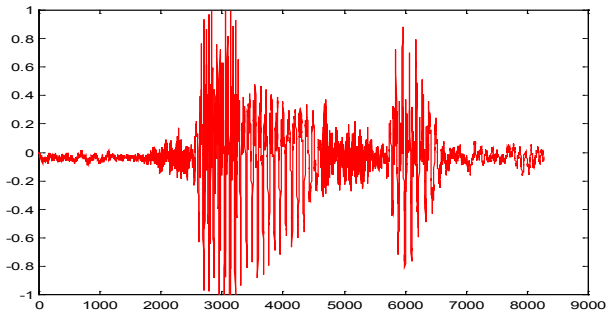
[1] ميس الكزبري "تعرف الكلمات المعزولة باستخدام سلاسل ماركوف المخفية".رسالة ماجستير جامعة دمشق كلية العلوم 2009 ص 12.

إن المصدر يميز الإنسان المتكلم و طريقة كلامه ، أما المرشح فيميز الكلام المنطوق ، لذلك فالتعرف على الكلام يتطلب معرفة شكل الأنبوب الصوتي أثناء إنتاجه ، و بتعبير أدق معرفة المرشحات الموافقة لنمذجته.

4.آلية تمثيل إشارة الكلام:

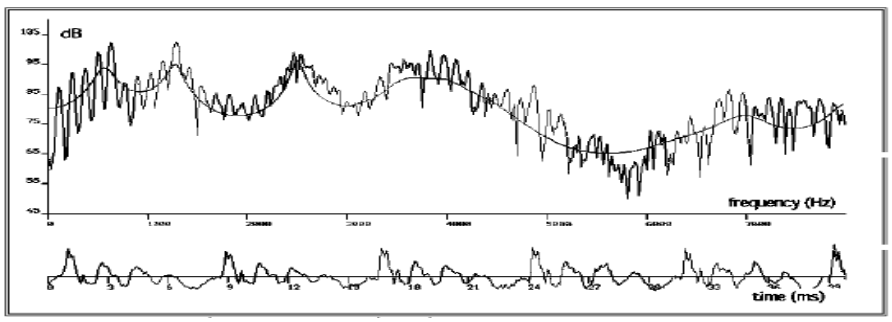
تمثيل الكلام أي الإشارة الناتجة يتم بثلاث طرق:

-**بشكل موجة** : أي تغيرات الإشارة بدلالة الزمن, و هذه التغيرات إما في مطال الإشارة و إما في طاقتها و نسميه عادة wave form.



الشكل 3.2: تمثيل زمني لإشارة كلمة "خمسة"

- **شكل طيف الإشارة:** حيث ندرس تغيرات مطال الإشارة أو طاقتها بحسب التردد[1].

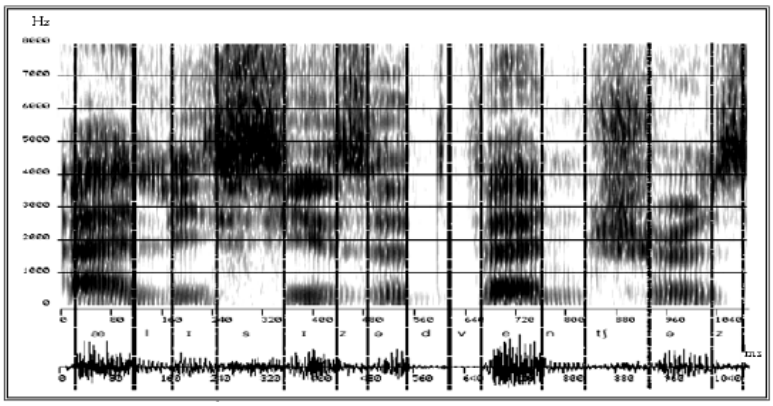


شكل4.2: نافذة كلام عند إنتاج صوت في الزمن(أسفل) يعلوها طيف الإشارة على هذه النافذ[1].

[1] D.jurafsky and J.Martin '' An Introduction to Natural Language Proccessing Computational Linguistices and Speech Recognition'' , prentice hall ,2000[1]

- **بشكل سوناغرام** spectrogram : و هي ثلاثية الأبعاد و ترتبط بالزمن و التردد و طاقة الحرف ، و هنا كأن نأخذ لقطات طيف الإشارة عند لحظات متتالية ، و لكن هذه الطريقة أكثر صعوبة في المعالجة.في كثير من الأحيان نرسم spectrogram كرسم ثنائي البعد و يتجلي البعد الثالث بدرجات لون الرمادي و الذي يمثل الطاقة.

تظهر الأشكال الآتية كيفية تمثيل الإشارة مع الزمن ،و تمثيل طيف هذه الإشارة و شكل السوناغرام الموافق[1]:



شكل5.2:نافذة كلام عند إنتاج جملة في الزمن(أسفل) يعلوها السوناغرام الموافق.

5.مراحل عملية التعرف الآلي على الكلام:

1.5. رقمنة الإشارة الكلامية و تحديد نهايات الكلمات

أهمية المعالجة الأولية تأتي من تهيئة الاشارة و إعدادها لتقليل الأخطاء و تسهيل تطبيق عمليات المعالجة اللاحقة عليها , و المعالجة الأولية تكمن في اختيار وسائط التحصيل.

1.1.5. اللاقط

نختار اللاقط بحيث تكون نسبة الإشارة إلى الضجيج عالية و خالية قدر الإمكان من التشوهات ,كما يفضل اختيار لاقط اتجاهي لحساسيته الدقيقة,بينما اللاقط غير الاتجاهي يسيء إلى الإشارة الكلامية بشكل كبير

[1] D.jurafsky and J.Martin '' An Introduction to Natural Language Proccessing Computational Linguistices and Speech Recognition'' , prentice hall ,2000

لتساوي حساسيته عن المحور مع اختلاف المنبع الصوتي مما يسمح لمركبات ضجيجية بالتأثير على الإشارة الكلامية.

2.1.5.بطاقة تحصيل الصوت

و هي ذات أهمية كبيرة حيث يتم فيها تحويل الإشارة الكلامية إلى إشارة رقمية,لذلك يجب اختيارها صغيرة الخطأ قدر الإمكان ، و أغلب بطاقات التحصيل تضيف مركبات مستمرة و ترددية يتم حذفها بمرشحات تمرير عال ، كما يوجد في بطاقات التحصيل الحديثة مرشح تمرير منخفض يقوم بحذف المركبات الترددية الأعلى من نصف تردد العينات ، لهذا ليس من الضروري تطبيق هذا المرشح برمجيا[1].

3.1.5.وسط التحصيل المحيط:

من المهم عند بناء نماذج الكلام أن يكون وسط التحصيل قليل الضجيج قدر الإمكان.لأنه كلما كان الوسط أقل ضجيجا كانت ، الصفات المستخلصة من الإشارات مطابقة للأصل.

2.5.مرحلة تسجيل الكلمات

عند تسجيل إشارة الكلام يتم تقطيع هذه الإشارة ضمن بطاقة تحصيل الصوت بتردد تقطيع مناسب ,أو بمعنى آخر أخذ سلسلة من العينات تمثل مطال الإشارة الكلامية, و باعتبار أن هذه العملية تتم ضمن بطاقة التحصيل ، لذا علينا الانتباه إلى خصائص تسجيل هذه الإشارة ، و التي تتمثل في تحديد طريقة الترميز المستخدم و من أهم طرق الترميز المستخدم ترميزPCM (pulse code modulation) و نختار مع هذا الترميز ثلاثة محددات:

[1] ميس الكزبري، المرجع السابق، ص 6.

أ- تردد أخذ العينات f_s: تمتد حدود السمع[1] ضمن مجال 20-20000hz و يتراوح أعلى مجال حساسية الأذن من 800-8000 hz فإذا ما قارناها 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 4000 hz .

و بتطبيق قانون شانون في تحصيل الإشارة الكلامية الذي يشترط أن يكون تردد التقطيع أكبر من أو يساوي التردد الأعلى $f_s \geq f_{max}$ حيث f_{max} التردد الأعلى أو الأعظمي للإشارة الكلامية ، يكفي أن نأخذ تردد التقطيع f_s= 8000hz ليعطي إشارة كلام مقبولة جدا. و يجب الإنتباه بأنه عند زيادة هذا التردد ترتفع نوعية الإشارة المسجلة و لكنها تؤدي لزيادة عدد العينات المأخوذة في الثانية مسببة زيادة في زمن المعالجة الرقمية.

ب- عدد بتات الترميز Q و يقصد بها عدد البتات اللازمة لترميز قيمة مطال 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 بالعكس سيعقد عمليات المعالجة مؤديا لزيادة زمنها.

إن اختيار عدد بتات الترميز 12 bit يحقق ترميزا جيدا ، إلا أن الخيار المتوفر في الحاسب هو 16 bit لهذا و تسهيلا للحساب فالمستحسن اختيار 16 bit [2].

ج- التسجيل باستخدام قناة أو قناتين أي التسجيل بأحد الخيارين mono أو stéréo مما يؤثر على حجم الملف الصوتي الذي تم تسجيله.

[1] Pilhofer ،Michael (2007). Music Theory for Dummies. p97.

[2] مقال أساسيات لابد لأي مبرمج معرفتها موقع انفورمتيك بواسطة وجدي عصام – مقالات عامة –سبتمبر 2013.

3.5.كشف حدود الكلمات

تأتي أهمية تحديد بداية و نهاية الكلمة المنطوقة في تقليل الحسابات اللازمة في عملية التمييز بشكل كبير ، و ذلك بتحديد مكان الكلمة و عزلها عن الصمت الموجود قبلها و بعدها ، و بالتالي تقليل حجم الكلمة المراد معالجتها ، يتم هذا بخوارزمية تعتمد على قيمتين هما طاقة الإشارة و معدل قطع الصفر[1].

4.5.حساب طاقة الإشارة

تكمن أهمية طاقة الإشارة بتمييز أماكن الصمت في الإشارة الكلامية, وتعرَف رياضيا بأنها مجموع القيم التربيعية لمطالات إشارة الكلام , بفرض أن إشارة الكلام $x(m)$ أي

$$E = \sum_{-\infty}^{+\infty} x^2(m) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.1)$$

إلا أن هذه المعادلة يمكن تطبيقها في حال الإشارة المستقرة و باعتبار أن الإشارة الكلامية تتغير بشكل بطيء نسبيا مع الزمن, فيمكن اعتبارها مستقرة ضمن مجالات زمنية قصيرة, لذا نقوم بتقطيع الإشارة إلى مجالات زمنية قصيرة تسمى بالإطارات (النوافذ) نعتبر الإشارة ضمنها مستقرة,و بالحصول على طاقة النوافذ الزمنية المتتالية نحصل توصيف جيد للإشارة, مع الانتباه إلى أخذ قيمة وسطية لطول النافذة يتراوح بين 100- 200 عينة عند تردد تقطيع يساوي 10khz أي ما يعادل زمنيا نافذة طولها 10-20 ms و بالتالي نحسب طاقة الإشارة بالمعادلة

$$E(n) = \sum_{n-N+1}^{n} \left[s(m).w(n-m) \right]^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.2)$$

حيث N عدد عينات الإشارة و s(m) عدد عينات إشارة الصوت المسجلة,و w(n) التابع المعبر عن النافذة المستطيلة و الذي يعطى بالعلاقة :

[1] ميس الكزبري، مرجع سبق ذكره، ص 14.

$$w(n) = \begin{cases} 1 & 0 \le n \le N-1 \\ 0 & otherwise \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.3)$$

و لتسهيل الحسابات نستخدم القيم المطلقة عوضا عن القيم التربيعية في المعادلة السابقة فتصبح:

$$E(n) = \sum_{n-N+1}^{n} |s(m)| - w(n-m) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.4)$$

عند حساب طاقة الإشارة لا يمكن أن نغفل معدل قطع الصفر z(n) أو ما يسمى معدل العبور بالصفر و هو عدد المرات التي تتغير فيها إشارة قيمة عينة الإشارة الكلامية بين السالب و الموجب, و بالعكس خلال المجال الزمني 10م.ثا و نهتم بحساب وسطي هذا المعدل خلال الضجيج IZC و الذي يعطى بالمعادلة :

$$\overline{IZC} = \frac{\sum_{k=1}^{k} z_k(n)}{K} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.5)$$

حيث K=10 و هي مجالات الحساب خلال 100 م.ثا الأولىمن التسجيل و خلال هذا المجال يجري حساب المقادير الآتية:

1- معدل قطع الصفر IZC.

2- الانحراف المعياري لمعدل قطع الصفر σ_{IZC} و يحسب بالمعادلة:

$$IZCT = \min\left(IF, \overline{IZC}, 2 * \sigma_{IZC}\right) \ldots\ldots (2.6)$$

3- طاقة الضجيج حيث حسبنا الطاقة على كامل المجال في المعادلة 4-2 و نحسب القيمة العظمى و الدنيا للطاقة و منها نحسب العينتين ITL و ITU بالمعادلات الآتية:

$$\begin{cases} I_1 = 0.03*(IMAX - IMIN) + IMIN \\ I_2 = 4*IMIN \\ ITL = \min(I_1, I_2) \\ ITU = 5*ITL \end{cases}$$

نعتبر IF=25 قطع للصفر خلال 10م.ثا عتبة ثابتة.

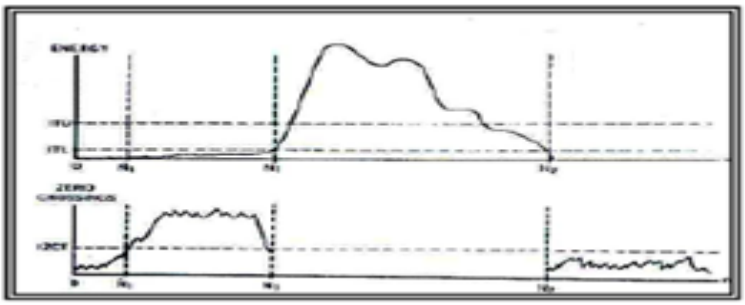
و مما سبق يمكننا حساب عتبة قطع الصفر IZCT بالمعادلة

$$IZCT = \min\left(IF, \overline{IZC}, 2*\sigma_{IZC}\right)$$

تبدأ الخوارزمية من بداية المجال حتى نتجاوز قيمة العتبة الدنيا ITL تعين هذه النقطة بشكل مبدئي بأنها بداية المقطع الصوتي بشرط ان تتجاوز الطاقة المعينة الطاقة الدنيا ITU بعد تجاوزها العتبة الدنيا ITL و نسمي هذه النقطة مبدئيا نقطة البداية .N

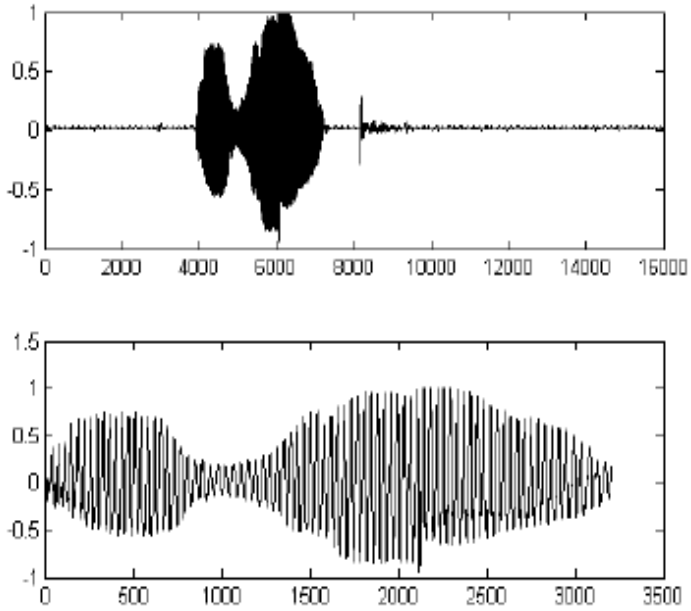
نتابع الخوارزمية حيث استخدمنا في التقدير المبدئي السابق طاقة الإشارة فقط ، و هذا التقدير فعال في حال كانت الطاقة ضعيفة و كانت قريبة من طاقة الضجيج ,لذلك نعود بالحساب لخارج المجال N1-N2 الذي حددناه في بداية و نهاية الكلمة ، و نستخدم معدل قطع الصفر باعتبار النقطة التي يتم تجاوز معدل قطع الصفر هي النقطة الفعلية لبداية و نهاية الكلمة ، بالتالي نعدل حدود الكلمة. وهذا موضح في الشكل (2.7) حيث حسبنا طاقة الإشارة و حددنا النقطتين N1 وN2 كنقطتي بداية و نهاية الكلمة ، و عند استخدام معدل قطع الصفر تم تعديل نقطة البداية إلى النقطة N1 و هي النقطة الأولى التي حدث فيها هذا التجاوز, بينما تبقى النقطة N2 هي نقطة النهاية لعدم تحقق الشرط السابق[1].

[1] L.R.Rabiner and M.R Sambur « An Algorithm for determining the endpoints of isolated utterances » The bell system Technical Journal vol 54 februry 1975.



الشكل 6.2: تصحيح نهايات الكلمة اعتمادا على معدل العبور بالصفر

الخوارزمية السابقة مهمة في الأوساط الضجيجية العادية أي عند استخدام برنامج التعرف، و لا تفقد أهميتها عند مراعاة شروط وسط التحصيل الجيد كونها متكيفة مع البيئة لاعتمادها على ضجيج البيئة المحيطة في حساب معدل الطاقة و معدل قطع الصفر.



الشكل 7.2 :مثال تحديد بداية و نهاية كلمة إغلاق

5.5.استنتاج أشعة سمات الإشارة الكلامية :

المعالجة السابقة هي لتمييز الإشارات عن بعضها البعض و حذف الإشارة غير المفيدة ، بالتالي لدينا الآن الإشارة تحتوي المعلومات التي تميز كل كلمة عن الأخرى ، و للحصول على هذه المعلومات نحتاج إلى معالجة الإشارة و تمثيلها بمجموعة من المعاملات ، و هذه المعالجة تمر بعدة مراحل :

1.5.5.ترشيح الإشارة بمرشح التعزيز الأولي

و نقصد به إلغاء أثر الشفاه ، حيث أن للشفاه أثرا كبيرا في عملية إنتاج الكلام ، و الذي يتجلى بإشعاع يخفف مطالات الترددات العالية في الإشارة, و يتم إزالة هذا الأثر بمرشح تمرير مرتفع من الدرجة الأولى يعطى تابع التحويل له ب :

$$H(z)=1-az^{-1}.....a\in[0.40,1] \quad .....(7.2)$$

و نختار قيمة a لتكون قريبة من الواحد و غالبا تعتمد في معالجة الاشارة الكلامية القيمة الآتية :

$$a=1-\frac{1}{16}=0.9375$$

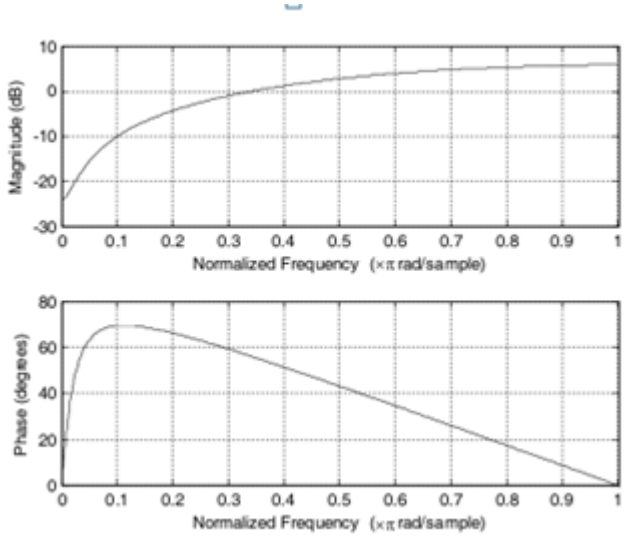
و بالتالي تعطى الإشارة s(n) الناتجة عن استخدام هذا المرشح على إشارة الكلام s(n)بالعلاقة التالية :

$$s'(n)=s(n)-0.9375*s(n-1) \quad .....(8.2)$$

و برمجيا نقوم بحساب العينة الحالية بجمع هذه العينة مع العينة السابقة بعد ضربها بالمعامل a أي

$$x(i)=x(i)+0.9375*x(i-1)$$

يعزز هذا المرشح الترددات العالية و يعوض الانحدار الموجود في طيف إشارة الكلام و هذا موضح في الشكل:



الشكل 8.2: خواص مرشح التعزير الأولي من حيث المطال و الطور

2.5.5.تقسيم الإشارة إلى نوافذ زمنية:

يستخدم التحليل الطيفي لجزء قليل من الإشارة لأن مواصفاتها ثابتة، حيث إنه يمكننا اعتبار إشارة الكلام إشارة مستقرة لا تتغير طبيعتها مع الزمن خلال نوافذ زمنية قصيرة نسبيا،و يتم تقسيم الإشارة إلى عدة نوافذ متداخلة بحيث تحوي كل نافذة N عينة، و نأخذ تداخلا بين النوافذ بإزاحة مقدارها M عينة للتأكد من عدم ضياع معلومات بين النوافذ ، فإذا اعتبرنا أن القيمة النموذجية لطول النافذة هي 30 ms و مقدار الإزاحة بين نافذة و أخرى هي 15 ms عندها من أجل تردد تقطيع Fs=8000Hz فإن N=240samples و M=120samples مع الانتباه إلى أن عدد النوافذ سيختلف من كلمة لأخرى ،كون الكلمات مختلفة الأطوال و بالتالي النافذة[1]:

$$x_i(m) = s(m+iM) \qquad m= 0, 1 , \dots ,N-1$$

[1] شهلة عبد الوهاب عبدالقادر، تمييز أصوات الأرقام العربية، قسم أنظمة الحاسبات، المعهد التقني الموصل، الموصل، العراق.

3.5.5.التمرير عبر النافذة:

يتم في هذه الخطوة تمرير كل إطار من خلال نافذة بفترة معينة عبارة عن دالة رياضية تشبه في وظيفتها عمل النافذة الاعتيادية بإدخال كمية الضوء ، و ذلك لتقليل عدم استمرارية الإشارة و إزالة التعرجات الحادة منها في بداية و نهاية كل إطار كما هو موضح في الشكل ، تقوم هذه النافذة بتنعيم الإشارة و إزالة التعرجات الحادة منها لكل قالب في الإشارة ، وتكون نتيجة النافذة ممثلة بإشارة المعادلة الآتية[1]:

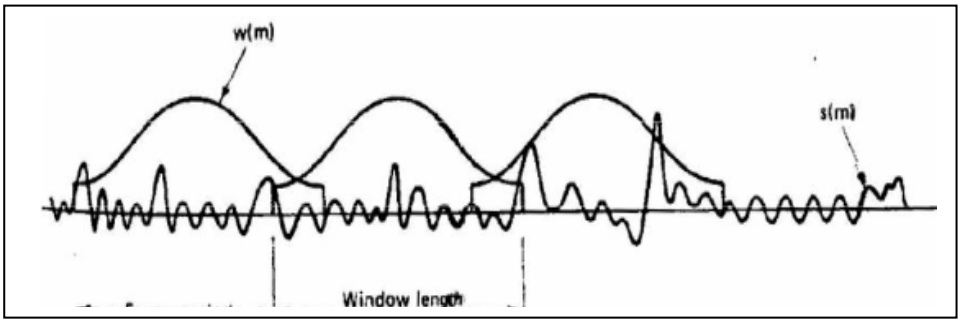
$$x_i(n) = x_i(n).w(n) \quad , 0 \leq n \leq N-1 \quad \ldots\ldots\ldots\ldots\ldots\ldots (2.9)$$

حيث:

x_i : عينات القالب الواحد .

N : حجم القالب أي عدد عينات القالب.

و هناك أنواع عديدة من النوافذ, و النافذة الأكثر شيوعا في أنظمة التعرف على الكلام هي نافذة **هامينغ** (Hamming window)، و السبب في استخدامها هو خواصها الطيفية الجيدة ، حيث يتم تثقيل العينة باتجاه مركز النافذة و نخفف من أثر التشوه عند الحواف، و بالتالي نتأكد من عدم تشوه الإشارة و الحصول على تعرف دقيق بنسبة عالية للكلمة، و يوضح الشكل الآتي هذه الخطوة.



الشكل 9.2: الضرب بنوافذ هامينغ

[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.

تعطى علاقة نافذة هانينغ بالمعادلة الآتية:

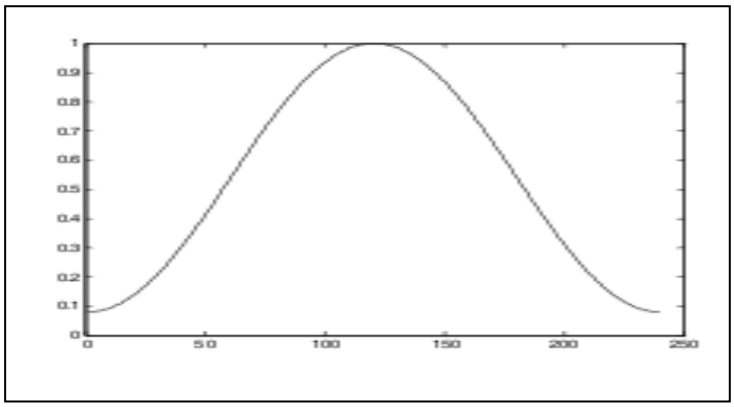
$$\text{Hanning}(n) = \begin{cases} \alpha - (1-\alpha)\cos\left(\frac{2\pi n}{N-1}\right) & 0 \le n \le N-1 \\ 0 & \text{otherwise} \end{cases} \qquad (2.10)$$

حيث N هو طول النافذة و $0 \le n \le N-1$ و تعتبر نافذة هامينغ حالة خاصة من نافذة هانينغ باعتبار $\alpha=0.54$:

و بالتالي تكون المعادلة لحساب النافذة

$$h(n) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) & 0 \le n \le N-1 \\ 0 & \text{otherwise} \end{cases} \qquad (2.11)$$

$$x_i(m) = x_i(m).h(m) \quad , 0 \le m \le N-1 \quad \ldots\ldots\ldots\ldots\ldots\ldots (2.12)$$



الشكل 10.2: شكل نافذة هامينغ[1] .

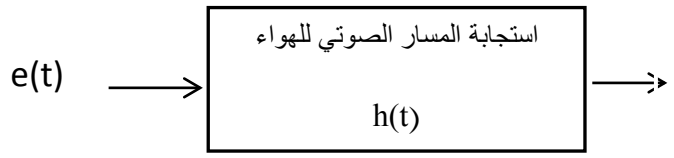
[1] ميس عبد القادر الكزبري، مرجع سبق ذكره.

6.طرق استنتاج أشعة سمات الإشارة الكلامية :

1.6. التحليل باستخدام الكبستروم (Cepstrum):

تحليل المركبات الطيفية: إن كلمة cepstral ليس لها معنى لغوي و إنما هي ترتيب معكوس للأحرف الأولى من كلمة spectral (طيفي) ، و ذلك للدلالة على أن هذه العملية تطبق على طيف الإشارات الترددي لاستخراج مجموعة المركبات التي تمثل هذه الإشارة, و كذلك الحال بالنسبة لكلمة cepstrum هي معكوس لكلمة spectrum .

إن أي إشارة هي عبارة عن استجابة وسط معين لكميات الطاقة التي يخضع لها, فالإشارة الصوتية هي استجابة المسار الصوتي للطاقة الناتجة عن ضغط الهواء الخارج من الرئتين. إن استجابة المسار الصوتي h(t) لهذه الطاقة هو ما يحدد شكل الإشارة الناتجة و بالتالي الصوت e(t) الناتج كما في الشكل:



الشكل 11.2: استجابة المسار الصوتي

يتم تمثيل هذه الإشارة رياضيا باستخدام الالتفاف (convolution) حيث أن الإشارة الناتجة هي مجموعة التعارف لإشارتين كما في المعادلة:

$$s(t) = \ (t) * h(t) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2.13)$$

إن الهدف من تحليل أي إشارة هو الحصول على استجابة الوسط h(t) ، حيث أن هذه الاستجابة هي التي تحدد ما هو الصوت المنطوق. لذلك نحتاج إلى فصل استجابة المسار الصوتي عن طاقة الإشارة, تعتبر هذه العملية

صعبة التطبيق في مجال الزمن بسبب مجموع الالتفاف (convolution) لذلك نحتاج إلى تحويل هذا المجموع إلى شكل يمكن استخلاص الاستجابة منه بحيث نستخدم تحويل فورييه لنقل الإشارة إلى مجال التردد ،و نطلق اسم الطيف لتردد الإشارة على خرج تحويل فورييه الذي يعطي بالمعادلة :

$$:s(f) = E(f) * H(f) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.14)$$

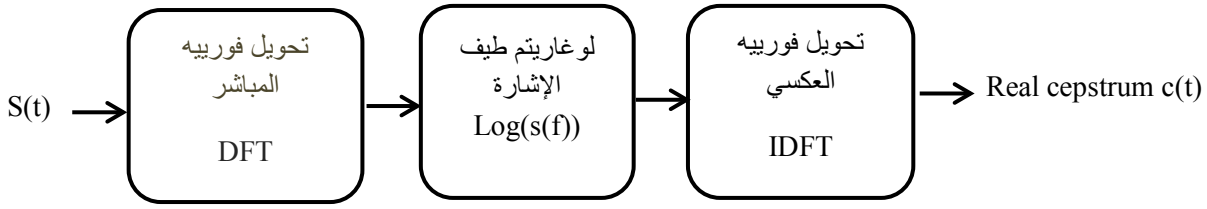
إن استجابة الوسط (المسار الصوتي) هي استجابة بطيئة التغير في الزمن وبالتالي فهي إشارة ذات تردد منخفض ، وطاقة المندفع في هذا الوسط هي إشارة سريعة التغير وبالتالي ذات تردد مرتفع فبالإمكان تحويل الضرب إلى جمع خطي لكل من الإشارتين و ذلك بتطبيق لوغارتم الطرفين كما في المعادلة :

$$\left|Log(E(f).H(f))\right| = \left|Log(E(f)\right| + \left|Log(H(f))\right| \ldots\ldots\ldots\ldots(2.15)$$

إن هذا الجمع الخطي للوغارتم الإشارتين يجعل بالإمكان الحصول على استجابة الوسط ، و الذي نحصل عليه بتطبيق تحويل فورية العكسي فنحصل على المركبات الطيفية الممثلة لإشارة الصوت المدخل ، أي انه بتطبيق تحويل فورييه العكسي على لوغاريتم الطيف الترددي للإشارة s(t) نحصل على المركبات الطيفيية الممثلة لهذه الإشارة والتي يطلق عليها real cepstrum, و يمكن تلخيص كل العمليات السابقة بالمعادلة:

$$C_s(n) = \frac{1}{N}\sum_{k=0}^{N-1} Log\left|S(k)\right|.e^{-j2\pi kn/N} \ldots\ldots\ldots n = 0.1\ldots\ldots N-1 \qquad (2.16)$$

أي أن التحليل الطيفي للمركبات cepstral analysis هو محاولة الحصول على المركبات التي تمثل استجابة الوسط للطاقة المقدمة من المنبع حيث تعتبر هذه المركبات هي القسم الحقيقي من تحويل فورييه العكسي المطبق على لوغاريتم الطيف الترددي للإشارة كما يوضح الشكل الآتي:



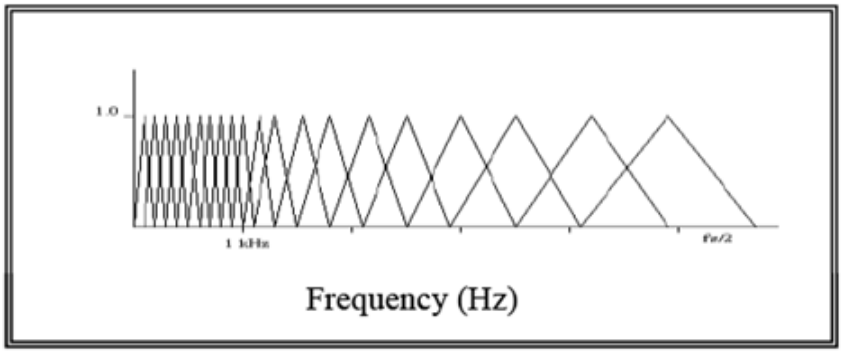
شكل12.2 :مراحل تحليل المركبات الممثلة

2.6 .تقنية درجة النغم Mel Frequency Cepstral Coefficients (MFCC)

تعتبر هذه التقنية من أكثر التقنيات شيوعا في استخراج السمات المميزة للصوت في نظم التعرف على الكلام ، و ذلك بسبب دقة نتائجها والقدرة على التخلص الجزئي من ضجيج الإشارة ، و كذلك سرعة تطبيقها وسهولته. كما أنه في هذه التقنية تتم مقاربة عملية السمع عند الإنسان أي محاولة استخراج سمات الإشارة بشكل يتوافق تقريبا مع آلية السمع عند الإنسان ، حيث أن الأذن البشرية حساسة للترددات المنخفضة تحت 1000 هرتز و ضعيفة الحساسية للترددات المرتفعة فوق 1000 هرتز. لذلك فإن هذا السلوك جعل بالإمكان تطبيق مجموعة معينة من المرشحات على الإشارة لكل مرشح مجال ترددي معين.

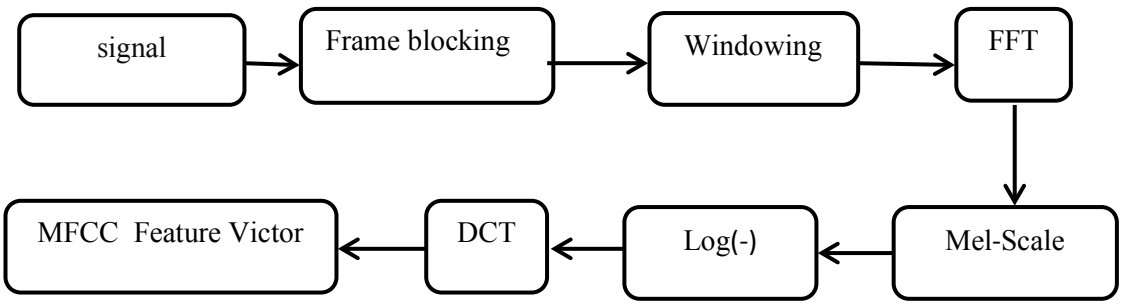
تسلك هذه المرشحات سلوكا خطيا من أجل الترددات المنخفضة وسلوكا لوغاريتميا من أجل الترددات المرتفعة (أي تتم زيادة عرض المجال الترددي للمرشح بشكل خطي من أجل الترددات المنخفضة تحت 1000 هرتز هو أن الأذن غير حساسة للترددات المرتفعة ، و بالتالي يكمن تقليل عدد المرشحات المميزة لهذه الترددات .يطلق على المرشحات Mel-Scale وتكون عبارة عن مجموعة من الإشارات المثلثية تمثل كل إشارة مرشحا لمجال ترددي مختلف ويمكن التعبير عن مراكز هذه المرشحات بالمعادلة :

$$F_{mel} = 2595.\log_{10}(1 + \frac{f_{Hz}}{700}) \quad \ldots\ldots\ldots\ldots(2.17)$$



الشكل 13.2: بنك الترشيح Mel

تعتمد تقنية MFCC على أسلوب التحليل الطيفي للمركبات بالإضافة إلى تطبيق مجموعة من المرشحات التي تحدد استجابة الإشارة من أجل مجالات ترددية مختلفة و بالتالي يمكن حصر مجال ترددات الإشارة و مركباتها في المرشحات التي تكون استجابة الإشارة فيها مرتفعة و تصبح السمات اللازمة لتمثيل هذه الإشارة أقل ، في تقنية MFCC تمر الإشارة الصوتية بعدد من المراحل و التي تكون نتيجتها مجموعة من السمات المميزة لهذه الإشارة ، هذه المراحل تتم بالترتيب وفق المخطط الآتي:



الشكل 14.2: مخطط عمل تقنية MFCC

يمكن اعتبار المراحل الثلاث الأولى مشتركة بين أغلب تقنيات تحليل الإشارة الصوتية لاستخراج السمات، و يمكن اعتبار البداية الفعلية لاستخراج السمات بتقنية MFCC بمرحلة تطبيق المرشحات الترددية Mel Frequency Warping .

1.2.6.تطبيق المرشحات الترددية

كما سبق و ذكرنا أن تقنية MFCC تحاول محاكاة آلية استجابة الأذن البشرية للترددات، لذلك و من أجل كل الطيف الترددي لكل إطار، نقوم بتطبيق مجموعة من المرشحات الترددية filter banc على هذا الطيف بغية حساب استجابة هذا الطيف لهذه المرشحات.

إن خرج هذه المرحلة هو الاستجابات الترددية للمرشحات التي تعتبر دخلا للمرحلة التالية، و هي المرحلة التي يتم فيها فصل طاقة الإشارة عن استجابة المسار الصوتي.

2.2.6.تطبيق اللوغاريتم على خرج المرشحات

إن MFCC تعتمد على التحليل الطيفي للمركبات لاستخراج السمات كما سبق و ذكرنا , حيث أنه يتم تطبيق اللوغاريتم على خرج المرشحات الترددية، و ذلك يؤدي إلى تسهيل الحصول على استجابة المسار الصوتي , حيث إن هذه الاستجابة هي التي تمثل المقطع الصوتي المنطوق phonème و بالتالي هذه الاستجابة هي ما يهمنا في الإشارة حيث يعطى خرج هذه المرحلة بالمعادلة:

$$\left|Log(E(f).H(f))\right| = \left|Log(E(f)\right| + \left|Log(H(f))\right| \ldots..(2.18)$$

3.2.6 تطبيق التحويل الجيبي المتقطع Discrete Cosine Transform

في هذه المرحلة يتم احتساب المركبات التي تمثل كل إطار حيث أن هذه المركبات هي السمات المميزة للإشارة (يتم تمثيلها باختيار 13 مركب) ، و هي MFCC والتي تجهزنا بتمثيل جيد لمواصفات الطيف لجزء من إشارة الرقم الصوتي للإطار المحلل وتستخدم المعادلة الآتية[1] لحساب معاملات MFCC:

$$MFCCi = X_k \cos[\pi i (k - 0.5) / M] \qquad i = 1, 2, \dots, p \quad \dots\dots\dots (2.19)$$

حيث أن: M : عدد مرشحات في بنك الترشيح

X_k: هي إشارة الإطار الواحد Mel-cepstrum .

P: عدد معاملات المتجه الواحد للإشارة المحللة.

3.6. تقنية المشتقات الأولى و الثانية لمعاملات الكبستروم:

و التي تسمى Delta and Delta-Delta Cepstrum و تعطى بالعلاقات الآتية:

$$\Delta c_i(n) = \sum_{k=-2}^{k=+2} k c_i(m+k) \qquad \dots\dots\dots\dots\dots\dots (2.20)$$

$$\Delta\Delta c_i(n) = c_i(m+1) - c_i(m-1) = \sum_{k=-1}^{k=+1} k \Delta c_i(m+k) \qquad \dots\dots\dots\dots (2.21)$$

إن كل قيمة من شعاع المشتقات $\Delta c_i(n)$ هي الفرق بين القيمة الموافقة في شعاع معاملات الكبستروم و القيمة ذات الترتيب نفسه في النوافذ المجاورة للنافذة الحالية، و لا نأخذ عددا كبيرا من النوافذ، لأن هذا يتنافى مع شرط استقرار إشارة الكلام الذي اعتبرناه في بداية هذه المرحلة من المعالجة، لهذا في حالة المشتقات من المرتبة

[1] يسرى فيصل الارحيم، عمي انصاف جاسم، مرجع سبق ذكره.

الأولى نأخذ أربع نوافذ مجاورة للنافذة الحالية (اثنتين قبلها و اثنتين بعدها)، أما في حالة المشتقات من المرتبة الثانية فإن قيم الشعاع $\Delta\Delta c_i(n)$ هي قيمة الفروق لمشتقات الدرجة الأولى $\Delta c_i(n)$ لنافذتين مجاورتين للنافذة الحالية.

إن اعتماد هذه المشتقات في نظم التعرف يؤدي إلى تحسين كبير في أدائها، حيث نلاحظ أن طول أشعة المشتقات المعرفة أعلاه تساوي طول شعاع معاملات الكبستروم، و يتطلب لحساب هذه الأشعة حساب أشعة معاملات الكبستروم لنوافذ الكلمة كلها، مما ينتج لدينا من الحسابات السابقة تسلسل من الأشعة $X=\{x_1,x_2,\ldots,x_n,\ldots,x_s\}$ بحيث كل شعاع x_n هو ناتج ضم شعاع الكبستروم $c_i(n)$ و شعاع مشتقاته من المرتبة الأولى $\Delta c_i(n)$ شعاع مشتقاته من المرتبة الثانية $\Delta\Delta c_i(n)$ للنافذة n أي:

$$x_n = (c_1(n), c_2(n), \ldots c_D(n),$$

$$\Delta c_1(n), \Delta c_2(n), \ldots \Delta c_D(n),$$

$$\Delta\Delta c_1(n), \Delta\Delta c_2(n), \ldots \Delta\Delta c_D(n)) \quad \ldots\ldots\ldots\ldots\ldots\ldots (2.22)$$

و إذا حددنا عدد معاملات الكبستروم $D = 12$ فإن معاملات كل شعاع x_n هو 36[1].

4.6. تحليل الإشارة باستخدام (LFCC) Linear Frequency Cepstrum Coefficients

تجهز هذه الطريقة بنموذج جيد و تقريبي لما يصدره جهاز النطق من أصوات و إزالة كل الترددات عديمة الأهمية , و تحسين جميع نواحي الإشارة و التي تساعد على دقة التعرف. مراحل عمل هذه التقنية تشبه التقنية السابقة عدا الإخراج التحليلي الطيفي فإنه لا يدخل الى مرحلة Mel Frequency wrapping بل يدخل مباشرة إلى مرحلة 'Cepstrum' حسب المعادلة الآتية :

$$LFCC_i = \sum_{k=0}^{NYQ-1} X_k \cos\frac{\pi / k}{NYQ} \ldots\ldots i = 1.2\ldots P \ldots\ldots\ldots\ldots (2.23)$$

حيث أن :

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

NYQ: نصف عينات القالب الواحد.

X_K: إشارة القالب الواحد Spectrum

P : عدد معاملات المتجه الواحد للإشارة المحللة.

5.6. التحليل بتقنية التنبؤ الخطي Linear Predictive Coding

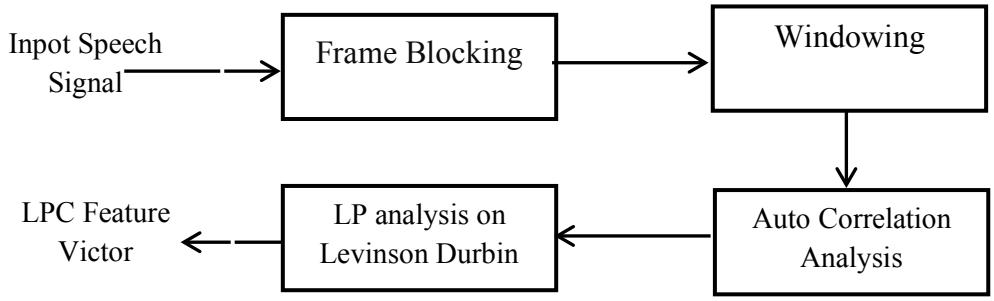
هي تقنية تحليل و تركيب (synthesizing) الكلام البشري ، وهي مستخدمة منذ زمن طويل ، ففي أحد معاهد تكساس عام 1987 أنشئت أول آلة النطق الكاتم (speech synthesizer) في دائرة متكاملة (integrated circuit) ، كانت هذه الإشارة الأولى لإمكانية صنع آلة تحاكى عمل الأوتار الصوتية البشرية باستخدام تقنية التشفير التنبؤي الخطي , فللتشفير التنبؤي الخطي إمكانية خفض تردد العينة تسمي (low bit rates) . و بالرغم من الصوت الصناعي الذي تنتجه، فقد شقت التقنية طريقها في التطبيقات العسكرية، حيث عدم الاهتمام بالدقة مقابل خفض حجم البيانات و سريتها.

و فكرتها الأساسية هي أن العينات الموزعة زمنياً في الإطار يمكن أن يتوقع بديلا خطيا لها، هذا البديل يمثل خصائص الصوت الأساسية في الإطار. فمثلاً إطار مكون من 255 عينة يمثل 20 جزء من الثانية، يمكن تمثيل هذه الفترة الزمنية في 13 عينة فقط تمثل خصائص الكلام البشري كما لو كانت 255 عينة.

تُستخدم شفرة التنبؤ الخطي عموما في ضغط الصوت في شركات الهاتف علي سبيل المثال GSM، وتستخدم في تأمين الإشارات اللاسلكية بحيث كل صوت يشفر ويرسل في قناة صوتية. و تأتي أهمية تحليل التنبؤ الخطي (Linear Predictive Coding) كونها تعطي تقديرا لأقطاب تابع تحويل الأنبوب الصوتي, إن خوارزمية التنبؤ الخطي من الدرجة p تتنبأ بقيمة أي عينة في الزمن اعتمادا على ال p عينة السابقة, أي رياضيا لدينا المعادلة الآتية:

$$s_n = \sum_{k=1}^{m} a_k . s_{n-k} + e_n \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2.24)$$

حيث نحسب sn بالاعتماد على العينات السابقة التي عددها p و نقوم بحساب معاملات التنبؤ الخطي، و لنميز هنا بين الإشارة الفعلية sn و تقدير الإشارة الذي نحسبه باستخدام معاملات التنبؤ الخطي SNN



الشكل 15.2:مخطط استخراج سمات الإشارة بتقنية التنبؤ الخطي

7. صعوبات نظم تمييز الكلام العربي

– غموض الإشارة الصوتية و تشابه الكلمات: تستطيع بعض نظم التمييز الآلي الكلام تحقيق أداء عال و نسبة تعرف مرتفعة إذا كانت كلمات قاموس التمييز واضحة نسبيا و لا تشبه بعضها البعض، مثل التعرف على الأرقام من صفر الى تسعة . و لكن في حال كانت كلمات القاموس متشابهة أو غامضة فهذه النتيجة لا نحصل عليها ذاتها. مثل التعرف على بعض الحروف الانكليزية كالحروف المتشابهة في طريقة اللفظ (DCEB) أو في حال التعرف على الكلمات العربية التالية المتشابهة في اللفظ :تحرير, تحديد, تجديد...

– الضجيج و تشوه الإشارة الصوتية: يؤثر الضجيج بشكل كبير على أداء نظام التعرف، و لا يمكن التخلص من ضجيج البيئة المحيطة في الغالب، و أبسط أنواع هذا الضجيج ظاهرة ارتداد الصوت عن جدران الغرفة و التشوه المرافق لعملية التسجيل ، و يمكن تخفيف أثر الضجيج باستخدام ميكروفون ذي نوعية جيدة.

– استخدام عدد من الميكروفونات: يؤثر استعمال عدد من الميكروفونات ذات المواصفات المختلفة على أداء نظام التعرف, لأن كل ميكروفون يتميز بخطية معينة و تشويه معين , و بالتالي عند نمذجة الميكروفون بمرشح فان كل ميكروفون له خصائص مختلفة عن الآخر.

– تغير طريقة الكلام: نطق الكلام نادرا ما يكون مضبوطا، و يختلف نطق العبارة نفسها من شخص إلى آخر، و بحسب معدل كلامه (سريع أو بطيء)، و كذا بحسب جودة صوته (يقصد الإفهام أو التكلم على نحو عارض).

– تحديد بداية و نهاية الكلام: و بمعنى أبسط تحديد بداية و نهاية جملة أو كلمة حسب نظام التمييز المبني, فمثلا في أنظمة تمييز الكلام المستمر قد يتوقف المتحدث للحظة، علما أن الجملة لم تنته بعد، أيضا في أنظمة الكلام المنفصلة حيث يكون لفظ بعض الحروف غير واضح مثل ف, هـ, ح في أول الكلمة.

8.الخاتمة

إن الهدف الأساسي من إنشاء أنظمة التعرف الآلي على الكلام ، هو استغلالها لانجاز تطبيقات تقنية تُسهل لمستخدم الحواسيب التعامل معها مثل: إعطاء الأوامر و التحكم صوتيا و غيرها، و إنجاز مثل هذه الأنظمة يمر بعدة مراحل تتم فيها معالجة إشارة الكلام بشكل متعاقب للوصول إلى النتائج المرجوة، إلا أن هذه المعالجة تعترضها مجموعة من المشاكل و الصعوبات يتطلب حلها استخراج أقصى قدر من المعلومات المفيدة من الإشارة الصوتية المدخلة لتسهيل التعرف عليها، حيث يتم معالجة الإشارة بتقسيمها إلى إطارات متساوية في الطول، ثم يتم استخلاص المعاملات لكل إطار لتكون النتيجة النهائية متجها من المعاملات يمثل ذلك الإطار، و هناك عدة طرق تستخدم لاستخلاص المعاملات لأية إشارة صوتية فيمكن استخدام الطاقة (Energy) و معدل قطع الصفر (Zero crossing) و معاملات درجة النغم (Mel Frequency cepstrum coefficient) ومعاملات التنبؤ الخطي (LPC) (Linear coefficients predictive) .

الفصل الثالث:

طرق التصنيف المستخدمة في

التعرف الآلي على الكلام

1.المقدمة

إن مصطلح التعرف على الأنماط (pattern recognition) يُعرَّف كأداة تارة و كعلم تارة أخرى ، ففي الحالة الأولى يعني أنه ماكنة ذكاء اصطناعي و في الحالة الثانية يمكن تعريفه على أنه علم وصف / تمييز /تصنيف .

و يمكن تصنيف هذا العلم الى ثلاثة أنواع كعلم تمييز الأنماط الإحصائي و التركيبي العصبي:

- التمييز الإحصائي و يعتمد على إيجاد احتمالات في البيانات بدل الاعتماد على تركيبة البيانات، و يتم تصنيف البيانات فيه بناء على عدد من الصفات، لتصنف البيانات ضمن فضاء يضم x من الأبعاد و الهدف منه هو الحصول على مجموعة من المناطق (regions) تطابق الفضاء المتعدد الأبعاد الناتج.
- أما التمييز التركيبي فيعتمد على تركيبة هرمية من التصنيفات التي تضم بعضها بعضا.
- في حين يحاول التمييز العصبي المعتمد على الشبكات العصبية الاصطناعية أن يحاكي آلية الإنسان في تمييز الأنماط.

2. تصنيف نظم التعرف الآلي على الكلام

يمكن تصنيف نظم التعرف الآلي على الكلام وفق عدة معايير:

1.2. التصنيف وفق طريقة الكلام:

- تمييز الكلمات المنفصلة isolated word recognition: حيث يتم تمييز كل كلمة لوحدها و يعزلها الصمت عما قبلها و ما بعدها.
- تمييز الكلمات المتصلة connected word recognition:حيث يتم تسجيل مجموعة من الكلمات التي تقرأ دفعة واحدة دون تعمد وجود صمت واضح بينها. وهذه الأنواع من النظم تشبه سابقتها إلا أن تنفيذها أصعب.

- تمييز الكلام المستمر continuous speech recognition: و هذه أصعب أنواع النظم حيث يتم تسجيل كلمات متصلة مع بعضها البعض دون فواصل و يمكن أن تختفي بعض الحروف أو تتحول إلى أخرى أو يضاف سعال أو ضحك.و يتطلب هذا معرفة متقدمة بتقسيم الكلام و معرفة الفونيمات المختلفة وصولا للكلمات و الجمل و المعنى باستعمال القواعد الصرفية و النحوية و الدلالية.

2.2. التصنيف و فق نمط الكلام

- تمييز كلام مقروء حيث يتم لفظ الكلمات بوضوح و عناية كما لو أننا نقرأ من الكتاب.
- تمييز كلام عفوي ، حيث يتم نطق الكلمات بعفوية كما لو أننا نتكلم بصورة عادية و هنا قد لا يكون اللفظ سليما كما قد لا تطبق قواعد اللغة في هذه الحالة.

3.2. التصنيف و فقا للمتحدث:

- تمييز كلام يعتمد على المتحدث speaker dependent: حيث يتم تسجيل الكلام من متحدث وحيد أو من عدة متحدثيين معينين و يميز النظام كلام المتحدثيين الذين قاموا بتدريب النظام من قبل,ولكن لا يعني هذا أنه لا يميز كلمات متحدثيين أخريين إنما تكون نسبة التمييز أقل.
- تمييز كلام مستقل عن المتحدث speaker independent:حيث يتم تمييز كلام أي متحدث مهما كان عمره أو جنسه و سواء شارك في تدريب النظام أم لا و لكن يحتاج هذا النظام الى مجموعة كبيرة من المتحدثيين أثناء مرحلة التدريب.

4.2. التصنيف و فقا لعدد الكلمات

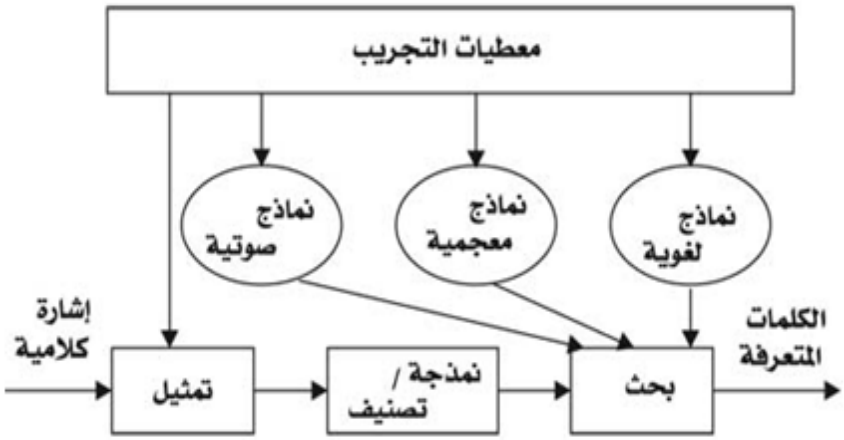
- تمييز عدد صغير من الكلمات حيث لا يتجاوز النظام عدد عشرين كلمة.
- عدد متوسط و يكون في حدود الألف كلمة.
- تمييز عدد كبير يفوق العشر الاف كلمة.

3. خطط نظم التعرف الآلي على الكلام:

يمكن تبسيط مخطط نظام التعرف على الكلام- أي نظام من التصنيفات السابقة- بالمطابقات التالية:

- مطابقة فونيمات بناء على نماذج صوتية.
- مطابقة كلمات بناء على قواميس.
- مطابقة جمل بناءا على نماذج نحوية و دلالية.

و المخطط التالي يبيين مراحل بناء نظام تعرف على كلام مستمر:



الشكل 1.3 : مكونات نظام نموذجي للتعرف الآلي على الكلام

و بصورة عامة يمر بناء أي نظام بمرحلتين أساسيتين :

- مرحلة التدريب أو مرحلة التعلم حيث يتم تسجيل الكلمات و بناء النظام بالاعتماد على حسابات تتم على هذه الكلمات.
- مرحلة التعرف: حيث يتم لفظ الكلمات و بمجموعة من الحسابات على هذه الكلمات يتعرف النظام على هذه الكلمات بالمقارنة مع الحسابات الموجودة لديه.

4.طرق التصنيف

هناك بعض الاختيارات الممكنة حول أي نوع من نماذج الإشارة هو المستخدم لتمييز خواص الإشارة ، وبشكل موسع استطاع العلماء تقسيم أنواع نماذج الإشارة إلى : نماذج التحديد والنماذج الإحصائية ، إذ إن نماذج التحديد (بشكل عام) تستخدم بعض الخواص المحددة المعروفة للإشارة ، مثلا تمثل موجة الجيب ، ففي هذه الحالة يكون التحديد لنموذج الإشارة (بشكل عام) واضحا ومباشرا ، وهذا مطلوب في تحديد (أو تقدير) قيم المعلمات لنموذج الإشارة (مثل : المدى ، التردد ، شكل (أو صورة) موجة الجيب) ، أما الصنف الثاني لنماذج الإشارة فهو مجموعة من النماذج الإحصائية التي تستخدم فقط لتمييز الخواص الإحصائية للإشارة ، ومن الأمثلة على هذه النماذج ، العمليات الكاوسية ،وعمليات بواسون وعمليات ماركوف وعمليات ماركوف المخفية ...الخ .

1.4.نماذج ماركوف المخفية

تعرف عملية ماركوف المخفية على أنها عملية تصادفية مخفية مركبة وذلك لأنها عملية ضمنية لا يمكن مشاهدتها ، ولكن يمكن مشاهدتها من خلال عملية تصادفية أخرى إذ إن الأخيرة تعطي متسلسلة من المشاهدات، فهي عملية مزدوجة ذات فضاء حالة منتهي مع عملية ماركوفية مخفية تتحكم باختيار الحالات.[1] درست (HMMs) ، التي هي من النماذج التصادفية ، في أواخر الستينات وبداية السبعينات من القرن العشرين وقد بدت المسألة صعبة جدا في بادئ الأمر ولكن بمرور الزمن ونتيجة للبحوث والدراسات التي جرت حول الموضوع فقد تم تسهيله نظريا وتطبيقيا في مجالات كثيرة من مجالات الحياة .

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

في عام (1989) قدّم الباحث (Rabiner) بحثا موسعا حول (HMMs) حاول من خلاله (وبشكل دقيق ومنظم) استعراض الجوانب النظرية لهذا النوع من الصياغة الإحصائية وإظهار كيفية تطبيقها في مسائل تمييز الكلام ، وقد عدّ هذا البحث المصدر الأساسي للعديد من البحوث الأخرى المختصة بنماذج ماركوف المخفية .

في عام (1999) ألقى (Parker) محاضرة بعنوان (ملف نماذج ماركوف المخفية) تضمنت المقدمة لـ(HMMs) وكيفية بنائها .[1] كذلك في عام (2001) قدّم (Ellis) بحثا بعنوان (تمييز المتسلسلة) أوضح فيه كيفية تمييز المتسلسلات الإحصائية فضلا عن نماذج ماركوف المخفية وحالاتها وطرائق حسابها[2] (Elllis ، 2001) .

كذلك في عام (2003) قام (Zhai) بتقديم بحث حول الملاحظات المختصرة والمفيدة جدا في نماذج ماركوف المخفية (Zhai , 2003) . كذلك في عام (2003) قام (Stamp) دراسة بعنوان (مقدمة الكشف عن نماذج ماركوف المخفية) ، إذ إنه من خلال مثال بسيط عن درجات الحرارة ، أعطى فكرة موجزة حول مفهوم نماذج ماركوف المخفية وحالاتها المخفية ومسائلها الأساسية وطرائق حلها فضلاً عن طريقة البرمجة الحركية وكيفية تقييس (Scaling) نموذج ماركوف المخفي.

1.1.4.النظرية التصادفية : Stochastic Process

إن أية ظاهرة حقيقية تجري في حيز المعلمة المدروسة (كالزمن مثلا) هي عملية تصادفية إذا كانت حالات تلك الظاهرة في أي جزء من حيزها (في أي وقت مثلا) تمثل نتائج تجربة عشوائية تخضع لقوانين الاحتمالات ، هذا وتعرف العملية التصادفية أيضاً بالعملية العشوائية (Random Process) أو عملية الفرصة[3] (Chance Process).

[1] Rabiner,L.R. (1989). precedent.

[2] Ellis,D. (2001) " Sequence Recognition " , Computer, Speech, and Language, Vol. 1, no. 2, pp. 167-197

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

تعريف: العملية التصادفية $\{X(t) : t = T\}$ هي العملية المكونة من خلال قيم مختلفة تأخذها متغيرات عشوائية معينة ، معركة على تجربة ما ، عند قيم مختلفة من فضاء المعلمة المدروسة (عند أزمان مختلفة مثلا) حيث إن (T) تمثل مجموعة المعلمات وتشير إلى فضاء المعلمة (Parameter Space) .

2.1.4.فضاء الحالة (The State Space)

يعرف فضاء الحالة على أنه مجموعة من القيم المتعلقة بـ$X(t)$.

ويكون فضاء الحالة متقطعاً (Discrete) إذا احتوى على نقاط منتهية أو نقاط غير منتهية قابلة للعد ، أما في الحالات الأخرى فهو مستمر (Continuous)[1].

3.1.4.سلاسل ماركوف (Markov Chains) :

معظم النظم الشائعة الاستخدام في التطبيقات العملية تتمتع بخاصية كون أن حالة الظاهرة قيد التحليل في الزمن الحاضر و الزمن الماضي هي التي تحدد حالتها في المستقبل ، ولربما قد تدخل في ذلك عوامل أخرى وذلك حسب المسألة المدروسة. أما ما يخصنا فهو تلك النظم والظواهر التي تتميز بخاصية أنه إذا علمت حالة أي تأثير لحالات الظاهرة في الزمن الماضي على ما سيجري لها في المستقبل ، وتلك النظم التي تمتلك هذه الخاصية تُعرف بسلاسل ماركوف ، و أن الخاصية نفسها تُدعى بالخاصية الماركوفية (Markovian property)[2].

المصفوفة الاحتمالية الانتقالية : إن المصفوفة العشوائية (التصادفية) هي مصفوفة ذات عناصر عشوائية (غير سالبة) منتهية أو غير منتهية بحيث إن مجموع عناصر كل صف يساوي (1) ، فإذا كانت المصفوفتان $A= a_{ij}$ و$B= [b_{ij}]$ تصادفيتين فإن المصفوفة الناتجة عن حاصل ضربهما هي أيضاً مصفوفة تصادفية .

أما$p(i ;j)$ فهي تمثل احتمالية الانتقال من الحالة (i) في الخطوة (n) إلى الحالة (j) في الخطوة ($n+1$) حيث يتم

[1] Cox,D.R. and Miller, H.D. (1965). " The Theory of Stochastic Processes ", John Wiley & Sons , Inc. , New York , USA. P14.

[2] الربيعي و فاضل محسن, مقدمة في العمليات التصادفية, دار الكتب بغداد العراق .2000م.ص15

وضع العناصر p(i ;j) (لكل i,j ∈ E) على شكل مصفوفة مربعة مثلاً (P) وبافتراض أن (E) هي {... 2, 1, 0}، فإنه يمكن كتابة المصفوفة (P) بالشكل الأتي :

$$P = \begin{bmatrix} p(0,0) & p(0,1) & p(0,2) & \cdots & p(0,N) \\ p(1,0) & p(1,1) & p(1,2) & \cdots & p(1,N) \\ p(2,0) & p(2,1) & p(2,2) & \cdots & p(2,N) \\ \vdots & \vdots & \vdots & & \ddots \\ p(N,0) & p(N,1) & p(N,2) & & p(N,N) \end{bmatrix}$$

تعريف : يقال للمصفوفة المربعة (P) (المعرفة أعلاه) بأنها مصفوفة ماركوف (أو مصفوفة الانتقال) الممثلة للعملية {Xي n=N} ، على فضاء الحالة (E) ، إذا تحقق الشرطان الآتيان[1] :

- لكل قيم (i , j ∈ E) تكون $p(i , j) \geq 0$.
- لكل قيم (i ∈ E) يكون $\sum_{j \in E} p(i, j) = 1$.

4.1.4 نماذج ماركوف المخفية :

نماذج ماركوفة المخفية هي مجموعة منتهية من الحالات كل حالة تقترن بتوزيع احتمالي ، إما الانتقالات ما بين الحالات فتُحدد بوساطة مجموعة من الاحتمالات وتسمى الاحتمالات الانتقالية . بشكل عام ، تتولد الحالة الناتجة (أي المشاهدة) طبقاً لتوزيع الاحتمالية المقترنة ، حيث توجد احتمالية ناتجة فقط ولا توجد حالة ظاهرة يمكن أن تُشاهد ، ولهذا فإن الحالات تكون مخفية ؛ هذا هو معنى نماذج ماركوف المخفية بشكل عام (1989 و Rabiner) . وهناك تعاريف لمفهوم نماذج ماركوف المخفية منها : تعريف: وهي تمثل متسلسلات تصادفية (كما في سلاسل ماركوف) حيث إن الحالات لا تُشاهد بشكل مباشر ولكنها تقترن بدالة كثافة

[1] الربيعي و فاضل محسن، المرجع السابق، 2000 ،ص17.

احتمالية (pdf) ، وبشكل أوضح فإن متسلسلة الحالة {هو} = Q (أي متسلسلة الحالات المخفية) لا يمكن مشاهدتها.

تُعدّ نماذج ماركوف المخفية إحدى الوسائل المفيدة لدراسة النماذج الاحتمالية في السلاسل الزمنية، إذ إن معلومات (HMMs) حول الماضي تنقل على شكل متغير متقطع منفرد (وهي تمثل الحالة المخفية) .

إن نماذج ماركوف المخفية هي نماذج إحصائية تُستخدم لنمذجة البيانات حيث إنها تُستخدم وبنجاح في العديد من المجالات منها تمييز وفهم الكلام وسيطرة الرجل الآلي[1] .

في الآونة الأخيرة ، حازت (HMMs) على اهتمام كبير في مجتمع الـ (Bioinformatic) (أي مجتمع المعلوماتية الحياتية) ، إذ أنها تُستخدم في حل المسائل الخاصة بتنبؤ التركيب النانوي البروتيني Secondary Structure Prediction و إيجاد الجينة و تصنيف عائلة البروتين وتشكيل خرائط ارتباط الجينات لعلم الوراثة وخرائط فيزيائية [2].

5.1.4. عناصر نماذج ماركوف المخفية (Eléments of HMMS) :

ويمكن تلخيص عناصر (HMMs) بالأتي :

يتم التعبير عن نماذج ماركوف المخفية باستخدام العناصر التالية [3]:

- T = طول متسلسلة المشاهدات
- $Q=\{q_1,q_1,\ldots,q_N\}$ هي مجموعة الحالات لعملية ماركوف (أي الحالات المخفية)
- N: هي عدد الحالات المخفية في النموذج و أن $S=\{s_1,s_2,\ldots,s_N\}$ هي مجموعة من الحالات

[1] Zhai,C. (2003). " A Brief Not on the Hidden Markov Models ", IEEE Transactions on PAMI Vol. PAMI –18 , No.3 , P.479

[2] Precedent.Parker,J. (1999). " Profile Hidden Markov Models "

[3] بان 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.

- $V = \{v_1, v_2, v_3, \ldots, v_M\}$ مجموعة رموز المشاهدات.
- M : عدد رموز الملاحظات (المشاهدات) (V) و أن $V= \{v_1, v_2, v_3, \ldots, v_M\}$ هي مجموعة من المشاهدات.
- A : مصفوفة احتمالية انتقال الحالة $A=\{a_{ij}\}$ و هي تمثل احتمال الانتقال من الحالة i إلى الحالة j خلال وحدة زمنية واحدة (خطوة واحدة).

و تعطى بـ:

$$A=\{a_{ij}\} \ , \ a_{ij}=P[q_{t+1}=S_j \mid q_t=S_i] \ ; \ 1 \leq i,j \leq N$$

إذ أن $q_t=i$ تمثل الحالة i خلال الزمن t.

- B : مصفوفة التوزيع الاحتمالي لرمز المشاهدة في الحالة j $B=\{b_j(k)\}$، و هي تمثل احتمال مشاهدة الرمز v_K في الحالة j، مصفوفة احتمالية رابطة بين الحالات المخفية و المشاهدات إذ أن:

$$B=\{b_j(k)\} \ , \ b_j(k) = P(O_t = V_k \mid q_t = S_j) \ ; \ 1 \leq j \leq N$$

$$1 \leq k \leq M$$

إذ أن O_t تمثل ملاحظة الرمز i خلال الزمن t.

- π : مصفوفة توزيع الحالة الابتدائية، $\{\pi_i\}$ = π إذ أن:

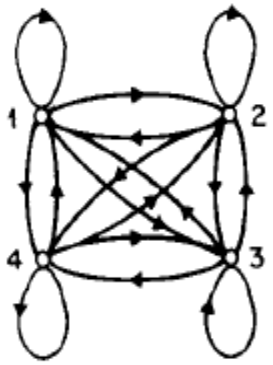
$$\pi = \{\pi_i\}, \ \pi_i = P[q_1 = S_i] \ ; \ 1 \leq i \leq N$$

- $O=\{O_1, O_1, \ldots, O_T\}$ متسلسلة المشاهدات.

6.1.4. أنواع نماذج ماركوف المخفية:

يمكن تصنيف نماذج ماركوف المخفية إلى نوعين و ذلك حسب الانتقالات بين حالاتها[1]:

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



الشكل 2.3: النموذج الثبوتي

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

$$a_{ij} = 0 \qquad , V \quad j < i$$

الشكل 3.3: نموذج الأيسر– الأيمن

و هذا النوع من النماذج يستخدم و بشكل واسع في نمذجة متسلسلات الإشارات.

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

2.4.الشبكات العصبية الاصطناعية

انتشرت تطبيقات الذكاء الصناعي في الفترة الأخيرة وخصوصا تطبيقات الشبكات العصبية حيث أنها تمثل أداة مناسبة لحل مشاكل عديدة من أبرزها تميز الأنماط وتصنيفها .

عرف عام 1943 بداية تطور الأنظمة العصبية الاصطناعية وقد حدد كل من **ميكلوج وبتس** (Meculloch and Pitt) الموديل الرسمي الأول للخلايا العصبية الحاسوبية من خلال موديل يتضمن كل العناصر الضرورية ومع ذلك فان انجاز وتنفيذ النموذج الالكتروني المدمج لهذا الموديل لم يكن عملي أو معقول من الناحية التكنولوجية خلال فترة الصمامات المفرغة علما أن هذا النموذج لم يطبق بشكل واسع لوصف المكونات المادية الحاسوبية بالنسبة للصمامات المفرغة وبالتالي لم يصبح قدم العالم **دونالد هيب** (Donald Hebb) في البداية مخططا تعليميا لتحديث اتصالات الخلايا العصبية التي يشار إليها بقانون هيب التعليمي Learning (Hebbian Rule) ، لقد ذكر هيب بأن المعلومات يمكن أن تخزن في روابط واتصالات ، وقد سلم بتقنية التعلم التي اثبت فوائدها في التطورات المستقبلية لهذا الحقل . و إن قانون هيب التعليمي يملك مساهمة أولية في نظرية الشبكة العصبية وقد تم بناء واختبار أول دراسة حاسوبية عصبية في الخمسينات حيث تم تطبيق الاتصالات تلقائيا ،و خلال هذه المرحلة كان مصطلح برسبترون (Preceptron) يطلق على الوحدة التي تمثلا لخلية العصبية ابتدع هذا المصطلح العالم فرنك برسبترون و أطلق على الوحدة التي تمثل الخلية العصبية. وكان هذا الاختراع عبارة عن ماكينة قابلة للتدريب عن تعلم وتصنيف نماذج معينة بواسطة تحوير الاتصالات إلى عناصر البداية. بهذه الطريقة فقد أصبح التماشي مع خيال المهندسين والعلماء وتم وضع خلفية أساسية للعمليات الحسابية لهذا النوع من الماكنات والتي مازالت تستخدم إلى يومنا هذا[1] .

[1] Kitano, H.."Designing neural networks using genetic algorithms with graph generation system". Complex system , 4 (4) ,461(476),(1990).

1.2.4.الشبكات العصبية

الشبكات العصبية تتألف من مجموعة من العقد تؤدي نوع خاص من الحساب بشكل جماعي، و إن كل عقدة هي وحدة قياسية حسابية صغيرة و هذه العقد يمكن أن تعمل بشكل متواز حيث أنها تعتمد على تفاعلات فيما بينها أو كيفية ارتباطها و عرفها بعض العلماء على أنها :

- نماذج رياضية تحاكي صفات الأنظمة البيولوجية التي تعالج المعلومات بطريقة متوازية مؤلفة من عناصر بسيطة نسبيا تدعى عصيبات .
- هي صنف أو كيان بسيط من الخوارزميات الرياضية التي تصاغ بشكل مخططات (Graphs) ، تصنف هذه المخططات إلى عدد كبير من الخوارزميات، و هذه الخوارزميات تقدم الحلول لعدد من المشاكل المعقدة.

إن أبرز نشاط تقوم به الشبكات العصبية هو عملية التصنيف و الترميز، و أبرز خواص الشبكات العصبية هي :

- مقاومتها للضوضاء .
- مرونتها في التعامل مع أنظمة الصور المشوهة .
- مقاومتها القصوى على تمييز الصور المتفككة أو المتحللة جزئيا .
- نستطيع القول بأنها عبارة عن تركيبات ذات عمليات متوازية مع وجود عدد كبير من الوحدات العاملة التي تحفز بواسطة ترابط العمليات الجارية وبالإضافة إلى خزين من المعلومات الموزعة بصورة متوازية .
- ذات عمليات غير خطية، أي مقدرتها على تقديم مخططات تتضمن علاقات غير خطية مضادة للضوضاء التي تجعل منها مصدرا جيدا لعمليات التصنيف والإسناد (classification predication).
- مقدرتها العالية على التكيف مع النظام اللوغاريتم التعليمي القوي وذات تنظيم داخلي يسمح باستخدامها للتكيف الداخلي الذي يعيش في محيط دائم التغيير.

2.2.4. أنواع الشبكات العصبية

بالإمكان التعرف على أبرز الأنواع الشائعة للشبكات العصبية متضمنة أنواع الإدخال والتعلم وبعض الاستخدامات الشائعة وكما في الجدول(1.3)[1]:

جدول 1.3 أبرز أنواع الشبكات العصبية واستخداماتها

نوع الشبكة العصبية	نوع الإدخال	أسلوب الإدخال	الاستخدامات الشائعة
شبكة Hopfield	Binary	Supervised	الذاكرة المرتبطة لتمييز حروف ASCII
شبكة Hamming	Binary	Supervised	الاتصال مع قناة متماثلة ثنائية
Carpenter/ Grassbery Classifier	Binary	Supervised	التجمع(نظرية الرنين التكيفية)
Perceptron	continuous	Supervised	تمييز الأشكال البسيطة وتصنيفها
Multi-layer perceptron	continuous	Supervised	تتميز الأشكال المعقدة وتصنيفها
Kohonen self organizing feature map	continuous	Supervised	تقييم المتجهات وتمييز الكلام,وتشبيه للشبكات العصبية البايولوجية

3.2.4 أساليب التعليم في الشبكات العصبية

بالإمكان استخدام طريقتين لتدريب الشبكات العصبية كوسيلة للتعلم او المحاكاة بين النظام والمستخدم وهذه الطرق هي :

- التعليم بوجود مشرف Supervised.
- التعليم مع عدم وجود مشرف Unsupervised.

الطريقة الأولى : يفترض وجود مشرف خلال عملية التدريب لكل نموذج ويستخدم للاختبار ويتم خلال (target output) وتتضمن نموذج المدجلات إضافة إلى الإخراج المرغوب لتحديد (real output) وان في

[1] Siddiqi, A. A.,&Lucas,S. M."A comparison of matrix rewriting versus direct encoding for evolving neural networks". In Proceeding of 1998IEEE,(1998).

هذه العملية يتم إجراء مقارنة بين الإخراج الحقيقي إلى ان يتم تنقيص أداء الشبكة بعد إعطاء مصفوفة الأوزان ويحدد متجه الخطأ (real vector) حيث انه يقارن بين (output vector) مع متجه المخرجات والمدخلات للشبكة ويقارن أيضا مع الإخراج المرغوب لتحديد الأخطاء من خلال المعادلة التالية :

Real output- target output = error

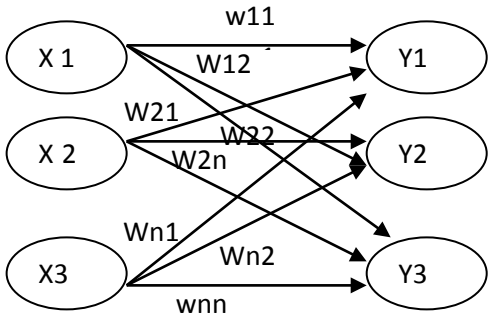
الطريقة الثانية: و في هذه الطريقة لايوجد مشرف لتقديم النماذج المطلوبة لذا يجب على النظام إن يتعلم بواسطة الاستكشاف والكيفية للخواص أو العوامل المهيكلة في نموذج الإدخال للمستخدم التجريبي ويجب إن يتم هذا التعلم من خلال تقوية الأوزان المنتخبة للعقد لكي تتطابق النماذج التجريبية ويميز التعلم بهذه الطريقة بصوره متكررة إلى إن يتم استقرار الأوزان .

4.2.4.معمارية الشبكات العصبية

أ- شبكات الطبقة المفردة (single layer network)

تمتلك هذه الشبكة طبقة واحدة من الأوزان, أما وحدات الإدخال فيمكن تعريفها بأنها الإشارات التي تستلمها من العالم الخارجي . وترتبط هذه الوحدات بوحدات الإخراج (التي تمثل الإجابة لوحدات الإدخال) وذات ارتباط كامل ومن أمثلتها شبكة (auto hetor Adeline preceptor)ذات طابع انسياب الخطأ خلفا ,ولا ترتبط وحدات الإخراج بأي وحدات أخرى باستثناء (شبكه Hopfield) التي فيها كل وحدات الإدخال هي نفسها وحدات الإخراج المخطط (3.4) يمثل نموذج لهذه الشبكات[1] .

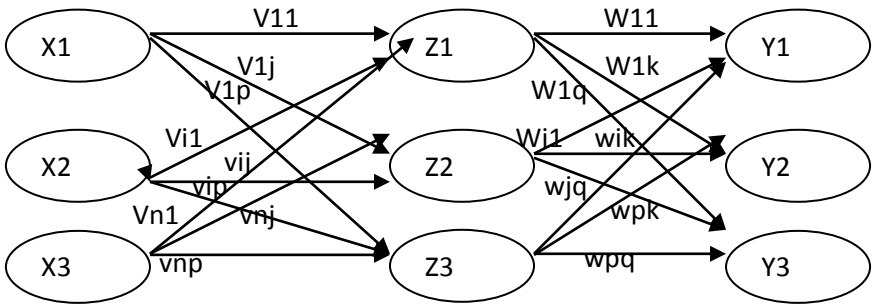
[1] Kitano, H.."Designing neural networks using genetic algorithms with graph generation system". Complex system , 4 (4) ,461(476),(1990).



الشكل 4.3: يمثل معمارية شبكات الطبقة المفردة

ب- الشبكات متعددة الطبقات (Multi layer network)

هي الشبكات التي تمتلك طبقة أو أكثر من العقد المخفية (Hidden nodes) يمكن تعريفها بأنها الوحدات التي لا تمثل وحدات إدخال ولا وحدات إخراج يمكن لهذه الشبكة التي تحل مشاكل أكثر تعقيدا من النوع الأول ولكن تدريبها يكون أصعب من أمثلتها (Medline ,Back propagation necogmatron miff).



الشكل 5.3: يمثل معمارية الشبكات متعددة الطبقات

ت- الشبكات العصبية التنافسية (Competitive Neural network)

تندرج ضمن معمارية قسم من الشبكات, تعتمد على طريقة تعلم (الفائز يأخذ الكل) حيث تتنافس الخلايا فيما بينها لكي تفوز واحده من هذه الشبكات في وقت معين لأن استجابتها أقوى من بقيه الشبكات ومن أمثلة هذا النوع من الشبكات:

– شبكة (Percetron)

هي من الشبكات ذات الطبقة الواحدة قد تمثلت هذه الشبكة في عام 1954 وقد سميت بالمدرك لأنها افترضت لتكون نموذج مبسط من التقنيات البالوجية لمعالجة المعلومات ذات العلاقة بالإحساس أي بمعنى إن المدرك في صيغتها المبسطة تحتوي على طبقتين منفصلتين من الخلايا تتمثل بطبقة الإدخال والإخراج وان هذه الخلايا تغذي خلايا طبقة الإخراج ولكن ليس العكس بالعكس وان الخلايا دخلت للطبقة لاتتصل يبعضها بالبعض الأخر . سير العمليات تتمثل بواسطة الأوزان وان تغذيه هذه الشبكة أمامية التغذية[1].

[1] Richard P .Lippmamn's "article An introduction to computing with neuron nets ",IEEE magazine, April 1987.

7. الخاتمة:

إن خوارزميات التصنيف هي شكل من أشكال تحليل البيانات والتي تستخلص نماذج تصف بشكل دقيق فئات وتصنيفات البيانات المهمة.

إن التحليل باستخدام خوارزميات التصنيف هو عبارة عن عملية مكونة من خطوتين، تسمى الأولى خطوة التعلم (Learn Step)، حيث يتم فيها بناء نموذج التصنيف، والخطوة الثانية هي خطوة التصنيف (Classification Step)، حيث يتم فيها استخدام النموذج من أجل التنبؤ بالفئات أو السمات لبيانات محددة.

ويتم في خطوة التعلم تحليل البيانات باستخدام خوارزمية التصنيف ومن ثم استخلاص قواعد التصنيف (Classification Rules)، والخطوة الثانية هي خطوة التصنيف ويتم فيها استخدام بيانات السجل المستكشف من أجل تقدير مدى تحقق قاعدة أو قواعد التصنيف، فإذا تحققت قاعدة التصنيف فإنه يمكن تطبيقها على السجل المستكشف.

إن تطور تقنيات التصنيف المختلفة كنماذج ماركوف المخفية جعلها أداة إحصائية فعالة لمعالجة العمليات التي تتغير مع تغير الزمن، فهي تساعد على ملاحظة مجموعة من الحالات و إيجاد احتمال وقوعها بعد تدريبها على حالات أخرى تمت ملاحظتها سابقا , و من المميزات الخاصة التي تتمتع بها نماذج ماركوف المخفية هي الخاصية الماركوفية و هي أن الحالة التي تحدث في زمن ما تعتمد في حدوثها فقط على الحالة في الزمن الذي يسبقها مباشرة، كما أن لنماذج ماركوف المخفية فضاء حالة يضم مجموعة منتهية من الحالات التي يمكن أن تكون رمزا أو حرفا أو كلمة أو أي شيء آخر، مما جعلها قابلة للاستخدام في العديد من المجالات كالتعرف على الكلام.

الفصل الرابع

الدراسة التطبيقية و تحليل النتائج

1. المقدمة

تعتمد أنظمة التعرف الآلي على الكلام بالأساس على أشكال الموجات الصوتية التي يتم من خلالها تشخيص الكلمة المستلمة من المتكلم ، فقبل 40 سنة من الدراسات و الأبحاث ، اقترحت خوارزميات عديدة للتعرف على الكلام تعتمد كل واحدة منها على تقنية معينة لاستخلاص مميزات الإشارات. تعد مطابقة الأنماط (pattern- recognition) واحدة من أهم هذه الدراسات، ففيها يخزن البرنامج النماذج التي يطلق عليها مصادر النظام (références) لكل صوت من مجموعة الأصوات لتشكل قاعدة البيانات, ثم تقارن إشارة الصوت (test) القادم بكل الأصوات لإيجاد أحسن تطابق ناتج للتعرف.

تم في هذا العمل استخلاص معاملات الإشارات الخاصة بمجموعة من الجمل عن طريق تحليلها بتقنيات مختلفة ثم إدخالها الى نظام للتعرف الآلي على الكلام المنطوق باللغة العربية ، حيث سنكرر عمل هذا النظام ثلاث مرات بحيث نستعمل في كل مرة نوعا واحدا من المعاملات المستخلصة من تحليل الإشارات الصوتية و قد اخترنا في عملنا هذا كلا من تقنية درجة النغم MFCC و تقنية المشتقات الأولى لدرجة النغم ΔMFCC, ثم المشتقات الثانية ΔΔMFCC ،كما دعمنا عملنا في الأخير بإضافة متغير طاقة الإشارة إلى التقنيات المختارة ومراقبة تأثير هذا الأخير على النتائج.

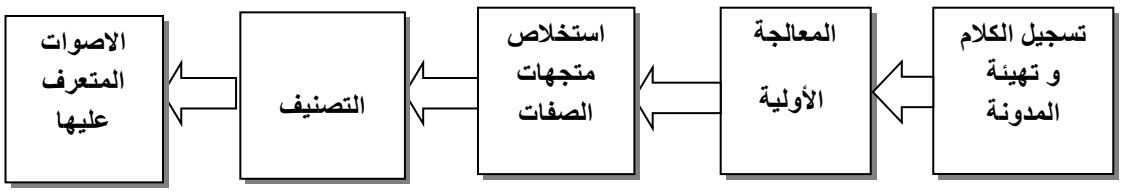
اعتمدنا في هذا البحث على مرحلتين أساسيتين تتألف منهما كل أنظمة التعرف و هي:

أ- استخلاص الصفات (feature extraction): تعد تقنية استخلاص الصفات من إحدى التقنيات ذات التطبيق الواسع في مجالات عديدة من أهمها تمييز الأنماط، وتعد إحدى الركائز المهمة التي يقوم عليها هذا المجال إذ تهدف إلى إيجاد سلسلة من متجهات الصفات (Feature Vector) التي تسمح بتمثيل محكم للإشارة الداخلة مما يسهل على النظام تمييز النمط الداخل مستخدماً تلك المعاملات الثابتة المستخلصة من إشارة الإدخال .يتم في مرحلة استخلاص الصفات اقتباس خصائص الشيء لغرض المقدرة على التعرف عليه فيما بعد من خلال تلك

الصفات التي تميزه عن غيره، و لم تقتصر عملية استخلاص الصفات على نوع محدد من البيانات فقد شملت الصورة و الكلام والفيديو. و مما ينبغي ذكره أن خصائص إشارة الكلام تتغير بصورة مستمرة مع الزمن نتيجة لعدد من العوامل، وهذا ما يجعل من المستحيل استخلاص صفات إشارة الكلام بصورة كاملة دون تقسيمها إلى عدد من المقاطع (Frames) و استخلاص صفات كل مقطع، وذلك يرجع إلى كون إشارة الكلام مستقرة نوعا ما في فترة قصيرة، وبذلك يكون الناتج عدد من متجهات الصفات بنفس عدد المقاطع المكونة للإشارة.

ب- مطابقة الصفات (matching feature): أو ما تسمى بعملية التمييز أو التعرف, تقارن فيها سلسلة متجهات الصفات لإشارة الاختبار مع كل الاحتمالات الموجودة في قاعدة البيانات باستخدام طريقة مطابقة الأنماط التي تعطي القرار من خلال جودة التطابق[1].

تتم مراحل عمل برامج التعرف الآلي على الكلام بصفة عامة وفق المخطط التالي:



الشكل1.4 : مراحل عمل نظم التعرف الآلي على الكلام

[1]Land J., 2000. Isolated Word Speech Recognition ofEnglish Digits, http://www.gerc.eng.ufl.edu/defult.htm

2.بناء نظام للتعرف الآلي للكلام

إن بناء نظام للتعرف الآلي للكلام المنطوق باللغة العربية سيمر بعدة مراحل حتى الوصول الى النتائج المرجوة.

1.2. مرحلة تهيئة البيانات.

تختلف مرحلة تهيئة البيانات من نظام الى آخر و ذلك تبعا للتقنيات المستخدمة في التحليل و التعرف, حيث تتطلب كل طريقة نظاما معينا لترتيب المعلومات قيد الدراسة.

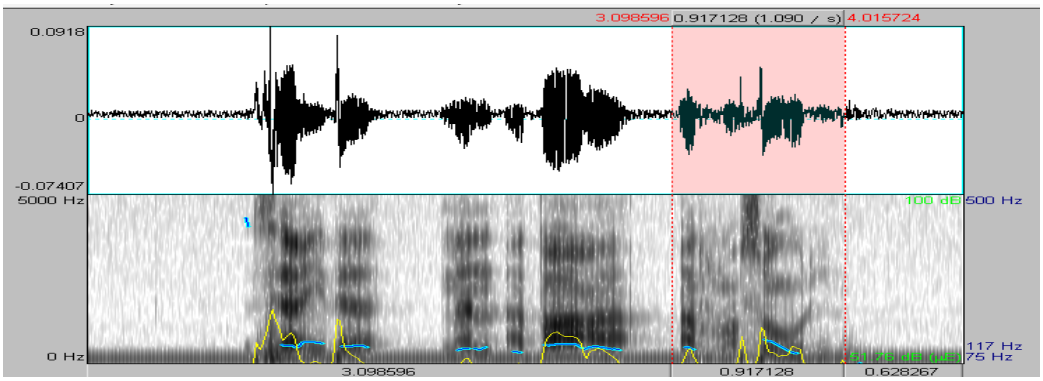
1.1.2 إنشاء المدونة

في هذه المرحلة يتم إنشاء المدونة التي ستحتوي على الأصوات التي ستتم دراستها كخطوة أولى, حيث تم تسجيل الأصوات من طرف أربعة أشخاص ، ذكران و أنثيين من البالغين و ذلك قصد تنويع المتكلمين، كل واحد منهم قام بتسجيل مجموعة من الجمل باللغة العربية عددها 110 جملة و إعادة نطقها خمس مرات و بذلك يصبح عدد ملفات المدونة هو 2200 تسجيل, هذه الجمل تم اختيارها بعناية بالغة حيث يراعى فيها استخدام أقصى عدد من حروف اللغة العربية في مختلف المواضع.

تمت مرحلة التسجيل باعتماد المعايير التالية:

- التسجيل بمعدل عينات 1025 هرتز
- تسجيل الأصوات بواسطة برنامج 'praat'.
- تمثيل العينات ب 16 بايت (bit)
- حفظ الأصوات بامتداد (mono wav) و هو الامتداد الشائع في نظام التشغيل ويندوز (windows)

.



الشكل4. 2: تسجيل و معاينة الأصوات بواسطة برنامج Praat

الجدول التالي يبين بعض الجمل المستخدمة في تكوين المدونة

جدول1.4 : بعض جمل المدونة

ٱلْفَرْعُ شِبْهُ ٱلأَصْلِ	بُثُوا الدَّعْوَةَ إِلَى بَذْلِ الْمَالِ
قَبَضَ عَلَى الْأَسْوِرَةِ بَقْبْضَتِهِ	انْكَسَرَ الضِّلْعُ بِسَبَبِ ٱلْعَدْو
أَثْقَلَ الدَّيْنُ كَاهِلَكَ	عَجِّلْ بِعَرْضِ مَا لَدَيكَ
أَثْخَنُوا فِي الْعَدُوِّ وَأَكْثَرُوا مِنَ السَّبْيِ فَجَرُّوا أَذْيَالَ الْهَزِيْمَةِ	فَصْمُ عُرَى الْأُخوَّةِ عَنْ عَمْدٍ لُؤْمٌ
يَبْدَأُ الطِّفْلُ بِالْحَبْوِ وَرَأْسُهُ أَرْغَبُ	صَدَرَ عَزْلُ خَالِدٍ لِعَمْرٍ
الْآيَاتُ إِلْهَامٌ لَمْ يَنْفُثْهَا الشَّيْطَانُ	كَانَ حَاتِمُ الطَّائِيُّ يُكْرِمُ الْأَضْيَافَ
دَخَلَ الْوَلَدُ فِي إِغْمَاءٍ بَعْدَ لَدْغِ الثُّعْبَان	عَسْفُ الرَّعِيَّةِ وَوكزُهَا يُوَلِّدُ الْكَرَاهِيَةَ
اِذْهَبْ إِلَى الْمَكْتَبَةِ لِلْإِتْيَانِ بِكُتُبِ ٱلعِلْمِ	يَجِبُ ٱلْحَمْدُ بَعْدَ الْعَطْسِ
خرَجَ اثْنَانِ مِنَ الطُّلَّابِ بِإِذْنٍ	كُنْيَةُ ٱلْجَرَادَةِ أُمُّ عَوفٍ
عُمْقُ الْإِخَاءِ سَبَبٌ لِدَوَامِ الصِّلَةِ	يَثْحَجُ الرَّاعِي ٱلعِجْلَ إِلَى ٱلْمَرْعَى
أَنْتَ تُشْعِرُ بِالْأُبُوَّةِ لِهَذَا الطِّفْلِ الْفَقِيْرِ	ضَرْبُ الْمُتَمَرِّدِ يَكْبَحُ ٱلعِصْيَانَ
حَانَ بَدْءُ مَوْسِمِ الْبَرْدِ	الْأَفْغَانُ مُقَاتِلُوْنَ أَفْذَاذ
رَأَى النَّائِمُ أَضْغَاثَ أَحْلامٍ	انْطَلَقَتْ أَفْوَاجُ الْحُجَّاجِ فِيْ جَوٍّ أَغْبَرَ

2.1.2 مرحلة تقسيم الجمل و ترميز الفونيمات

في الخطوة الموالية وبعد تسجيل الجمل نقوم بتقسيمها إلى كلمات في ملف نسميه قائمة الكلمات و يعرف في البرنامج بـ (liste des mots) ، ثم نقوم بترتيبها ترتيبا ألفبائيا بشكل عمودي و ذلك لتسهيل كتابتها كتابة صوتية اعتمادا على ترميز الفونيمات المكونة للكلمات بالطريقة التالية:

جدول 2.4: ترميز الفونيمات

الفونيم	الكتابة الصوتية	الفونيم	الكتابة الصوتية
الفتحة	AE	س	S
الألف الممدودة	AE:	ش	SH
الضمة	UH	ص	SS
الضمة الممدودة	UW	ض	DD
الكسرة	IH	ط	TT
الكسرة الممدودة	IY	ظ	DH2
همزة (ء)	E	ع	AI
ب	B	غ	GH
ت	T	ف	F
ث	TH	ق	Q
ج	JH	ك	K
ح	HH	ل	L
خ	KH	م	M
د	D	ن	N
ذ	DH	هـ	H
ر	R	و	W
ز	Z	ي	Y

تبعا لهذا الترميز سيتم حفظ جميع الكلمات في ملف نسميه القاموس و يعرف في البرنامج ب ' Dictionnaire', و نأخذ كمثال الجملة " الفرع شبه الأصل":

الفرع AE:LFARAIUH

شبه SHIHBHUH

الأصل AE:L AE:ESLI

في الخطوة الموالية، سيتم تعريف قائمة الفونيمات (liste phonemes) لتحديد معجم لقاعدة البيانات (corpus) بشكل قائمة و هي على النحو التالي:

E ;AE:;L;M ;UH ;IH ;N ;Y ;T ;AE ;B ;IY ;TH ;JH ;H ;GH ;AA ;HH ;D ;Z ;F ;D

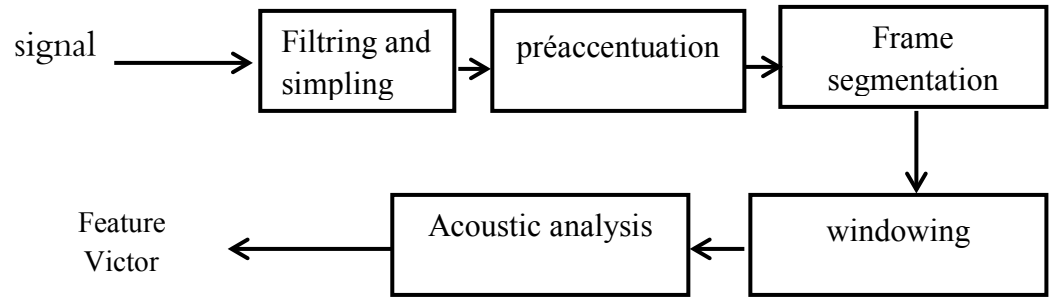
H2 ;Q ;R ;KH ;DD ;AH ;UX ;S ;DH ;SH ;SS ;AH:;AA:;TT ;W ;K ;AI ;UW ;IX

;sil

2.2.مرحلة تحليل الإشارات الصوتية و استخلاص الصفات:

و هي مرحلة أساسية في بناء البرنامج المستخدم للتعرف الآلي على الكلام حيث يتم على مستواها تغيير نوع التقنيات المستخدمة في تحليل الإشارات و مقارنة النتائج المتحصل عليها في الأخير.

عملية استخلاص الصفات تمت بتطبيق تقنيات درجة النغم (MFCC - ΔMFCC-ΔΔMFCC) لبناء قاعدة بيانات من المعاملات ، حيث يتم تحديد الجملة و قراءة الملف الخاص بها ثم يحلل هذا الملف بصورة منفصلة لاستخلاص متجهات الصفات لغرض خزنها في قاعدة البيانات, علماً أن عدد معاملات المتجه الواحد لكل إطار من إطارات الإشارة يساوي 12 معاملا من معاملات MFCC الأولى باستثناء معامل C0. و تكرر هذه العملية مع جميع الجمل لغرض استكمال قاعدة البيانات .



الشكل 3.4: مخطط استخلاص صفات الإشارة المدخلة

1.2.2 تهيئة ملف التحليل .configuration file:

بعد تحضير الملفات الصوتية نقوم بتهيئة ملف خاص بالتحليل يسمى " config" والذي يتضمن مجموعة من المعايير والشروط التي يتم اعطاءها للبرنامج ليقوم بتحليل الإشارات واستخلاص المعاملات وفقها.

بما أن التجربة تستخدم نفس مواصفات البرنامج, فإن مواصفات ملف التهيئة متشابهة. مثل تنسيق الملف (file format inpot)، ومعدل أخذ العينات (Rate Simpling) ، و حجم النافذة (size Window)، ونوع القناة (Type of channels)،كما يتم استخدام نفس نوع النافذة أيضا, بينما يظهر الفرق الرئيسي في نوع المعاملات المستهدفة (Target kind) وعددها.

يتم إعطاء القائمة الكاملة للمعاملات وقيمها في الجداول التالية:

جدول 3.4 : مواصفات ملف التحليل config الخاص بمعاملات MFCC

S n	*parameters*	*value*
1	Inpot format file	.wav
2	Simpling Rate (bits per sample)	16000
3	Type of channels	Mono
4	Window size	(25msec)
5	Frame periodicity	(10 msec)
6	used Window	Hamming
7	Number of Filter-bank channels	26
8	Target kind	MFCC-0
9	Number of MFCC coefficients	12
10	Pre-emphasis coefficients	0.97
11	Lenth of Cepstral Liftering	22

جدول4.4: مواصفات ملف التحليل config الخاص بمعاملات MFCC Δ

S n	*parameters*	*Value*
1	Inpot format file	.wav
2	Simpling Rate (bits per sample)	16000
3	Type of channels	Mono
4	Window size	(25msec)
5	Frame periodicity	(10 msec)
6	Windowused	Hamming
7	Number of Filter-bankchannels	26
8	Target kind	MFCC-0-D
9	Number of MFCC coefficients	12
10	Pre-emphasis coefficients	0.97
11	Lenth of Cepstral Liftering	22

جدول4. 5: مواصفات ملف التحليل config الخاص بمعاملات ΔΔMFCC

S n	*parameters*	*Value*
1	Inpot format file	.wav
2	Simpling Rate (bits per sample)	16000
3	Type of channels	Mono
4	Window size	250000.0(25msec)
5	Frame periodicity	10000.0(10 msec)
6	Window used	Hamming
7	Number of Filter-bank channels	26
8	Target kind	MFCC-0-D-A
9	Number of MFCC coefficients	12
10	Pre-emphasis coefficients	0.97
11	Lenth of Cepstral Liftering	22

3. مرحلة إعداد المصنف (classifier):

من أجل تقليص عملية البرمجة لمختلف أجزاء نظام التعرف الآلي على الكلام العربي، اخترنا أن نثري دراستنا هذه باستخدام وسيلة برمجة من بين مجموعة من أنظمة البرمجة المتعددة, الجدول رقم (4. 6) يوضح بعض أنظمة البرمجة.

جدول 6.4: بعض أنظمة البرمجة المستعملة في التعرف الآلي على الكلام[1]

	HTK	CSLU (ASR)	ISIP (ASR)	SPHINX
المؤسسة	Microsoft et Combridge University	Oregon graduate institut	Mississipi University	Camegie Mellon University
URL	htk.eng.cam.ac.uk	www.cslu.ogi.edu	www.isip.msstate.edu	fife.speeshcs.cmu.edu
لغة البرمجة	C	C, Tel/Tk	C++	C, perl, Java
بيئة العمل	Unix, Linux, windows	Windows	Unix	Unix, Linux, windows
تاريخ أول إصدار	1993	1992	1997	1987

وقد اخترنا نماذج ماركوف المخفية و التي تم تطويرها بجامعة كامبردج بواسطة **س.ج. يونغ** و فريقه. هاته الأدوات تم تطويرها لتسمح ببناء أنظمة التعرف الآلي على الكلام بالاعتماد على نماذج ماركوف المخفية.

تتمتع أدوات نماذج ماركوف المخفية بالمرونة خلال مراحل البرمجة لبناء نظام التعرف, حيث ان النماذج تستطيع تمثيل كلمات او أي قطع صوتية معتمدة على دالة كثافة احتمال مرتبطة بالحالات.

تعتمد مرحلة التعرف باستخدام نماذج ماركوف المخفية(HMM) على خطوتين رئيسيتين هما: التهيئة و إعادة التقييم. قبل بداية مرحلة التدريب يجب القيام بالخطوتين التاليتين[1]:

[1] A.K. Mohamed Cherif, Reconnaissance Automatique de la Parole par les HMM en Milieu Bruité, Doctorat, Université Annaba, Algerie, 2009.

تهيئة المصفوفة الانتقالية و ذلك بإنشاء نموذج HMM ذو ثلاث حالات من الانتقالات من اليسار الى اليمين لكل وحدة صوتية .

- تهيئة المتوسط بالقيمة صفر و التباين بالقيمة 1.(ملف التهيئة)

```
~o <VecSize> 36 <MFCC_D_A ><StreamInfo>
~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 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 1.0 1.0 1.0 1.0
<State> 3 <NumMixes> 1
<Stream> 1
<Mixture> 1 1.0000
<Mean> 36
```

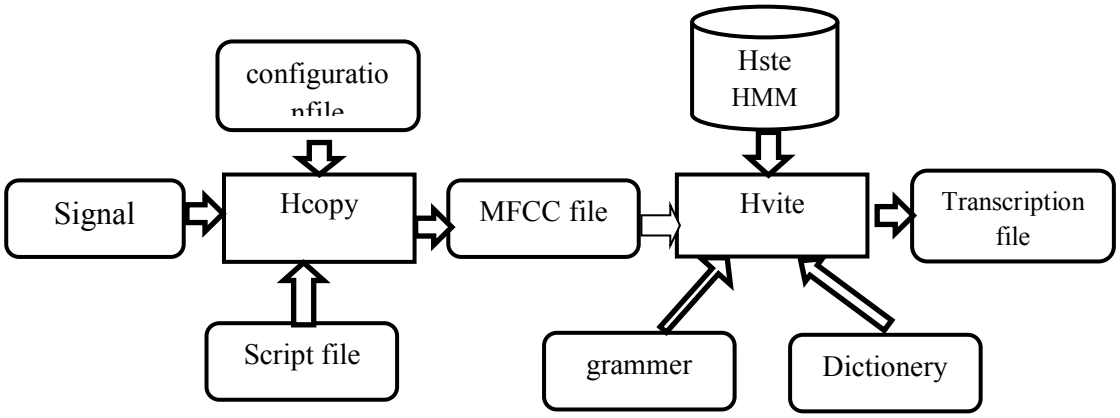
الشكل 4.4: نموذج ملف التهيئة

[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).

4.التعرف بواسطة الأداة HVite :

إن من صعوبات التعرف الآلي على الكلام، الإختلاف في طريقة لفظ الكلام بين المتكلمين و حتى لفظ الكلمة أو الجملة من طرف المتكلم نفسه, فنطقه للصوت في المرة الاولى لن يكون مطابقا تماما للنطق الثاني، و ينتج هذا الاختلاف في النطق بحسب سرعة كلام الشخص وجودة 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[1]، و لجعل نظام التعرف 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) واختيار الأقرب و الأنسب منها بالمقارنة مع الإشارة الصوتية المدخلة. و هو ما توفره مجموعة أدوات نماذج ماركوف المخفية بواسطة الأداة HVite.

تسمح الأداة HVite بترميز (Transcription) لحادثة متجهات الصفات الخاصة بكل كلمة أو فونيم, طريقة عمل الأداة في التعرف تتم وفق المخطط التالي:



الشكل 5.4 طريقة عمل الأداة HVite

[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).

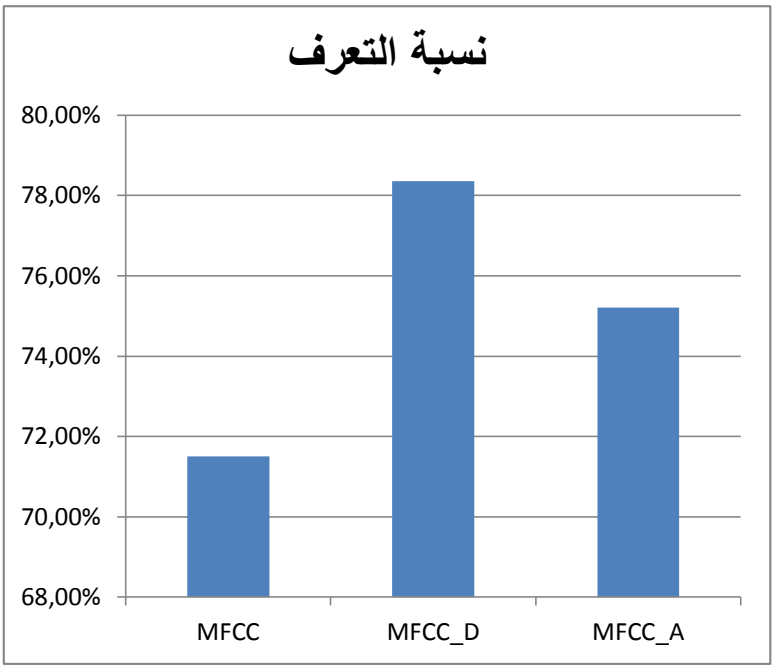
يسمح تكرار خوارزمية 'باوم وولش' المستخدمة بإعادة تقييم النماذج مع مراعاة التحسينات الأخيرة، ثم يتم تخزين المعاملات الأخيرة الناتجة في مجلد جديد. و هي المرحلة الأخيرة في تدريب نماذج HMM.

5. اختبار النظام بتقنيات التحليل المختارة و تحليل النتائج:

إن الاختبار تم على مجموعة مؤلفة من 110 جملة تم تكرار تسجيلها 5 مرات من طرف متكلمين متنوعين، أي لدينا 2200 تسجيل، تم استخدام نصفها للتدريب و نصفها للاختبار.

عند تنفيذ نظام التعرف بتطبيق معاملات MFCC ثم ΔMFCC ثم ΔΔMFCC تحصلنا على النتائج التالية:

1.5.متكلم موجود في مدونة التدريب

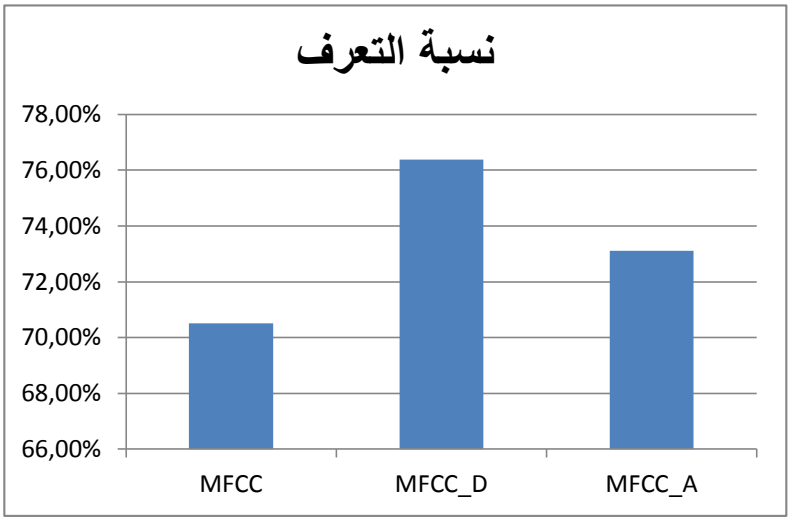


شكل 6.4. نسبة تعرف النظام لمتكلم من المدونة

يظهر لنا التمثيل البياني النسب المئوية للتعرف على الجمل بدلالة نوع المعاملات المستعملة حيث تظهر النسب المتحصل عليها أن النظام سجل دقة تعرف عالية بغض النظر عن نوعية المعاملات المستخلصة حيث سجلنا نسبة 73.48% بالنسبة لمعاملات MFCC و 79.48% بالنسبة لمعاملات ΔMFCC و بالنسبة لمعاملات ΔΔMFCC سجلنا نسبة 75.12%, إلا أن هذه النسب تظهر تفوق معاملاتΔMFCC و بالتالي هي الأنسب عند تطبيق النظام على متكلم من مدونة التدريب.

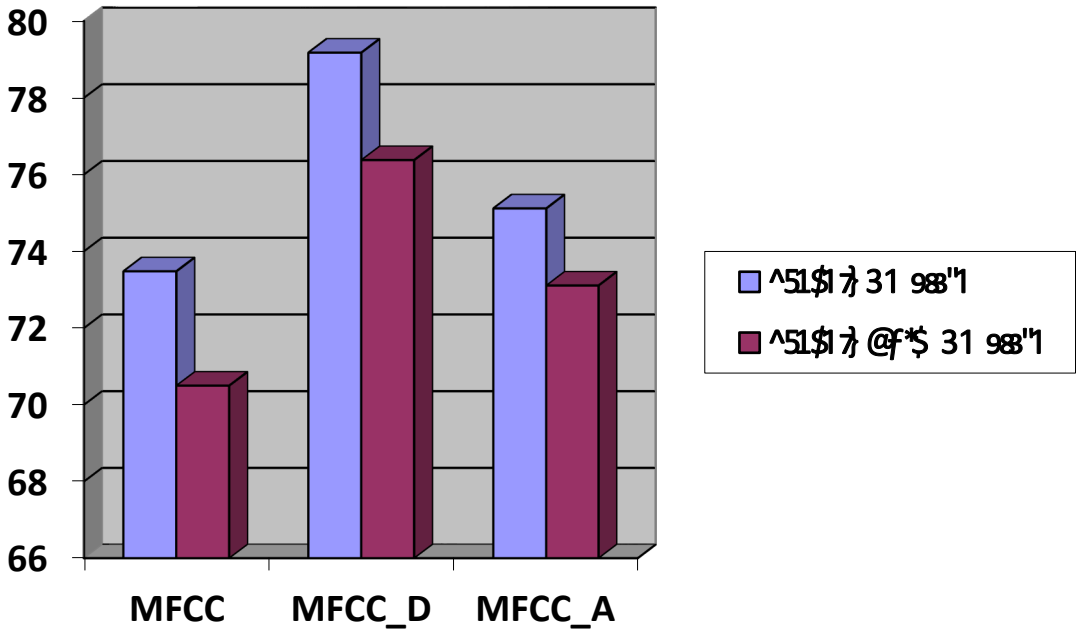
2.5.متكلم غير موجود في مدونة التدريب

عند استخدام متكلم غير موجود ضمن مدونة التدريب لاحظنا تناقصا في نسبة التعرف بالنسبة للأنواع الثلاثة من المعاملات, وقد كانت النتائج كالتالي: سجل النظام نسبة تعرف تقدر ب 70.5 %بالنسبة لمعاملات MFCC و نسبة 76.38 % عند استخدام معاملات ΔMFCC أما عند استخدام معاملات ΔΔMFCC فقد سجل النظام نسبة 73.11 %. الملاحظ في هذه التجربة أن أقل انخفاض في نسبة التعرف سجل عند استخدام معاملات ΔΔMFCC .



شكل 7.4. نسبة تعرف النظام لمتكلم من خارج المدونة

يمكن ان نوضح أهمية نوع المتكلم المستخدم في اختبار النظام و تأثيره البارز على نسبة التعرف من خلال التمثيل البياني التالي :



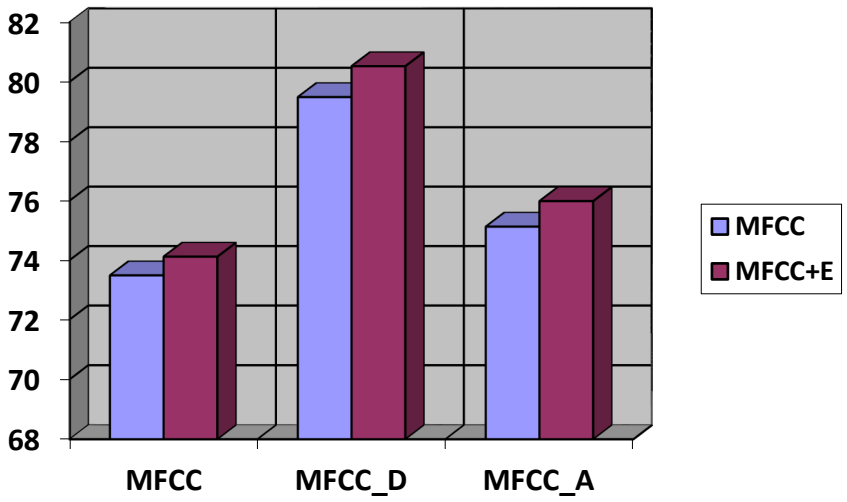
شكل 8.4.مقارنة نسب التعرف حسب نوع المتكلم.

نلاحظ من خلال الشكل السابق التفوق الظاهر لمعاملات المشتقات الأولى لتقنية درجة النغم حيث سجل النظام نسبة تعرف أعلى من النوعين الآخرين من المعاملات المستخدمة.

3.5.إضافة معامل طاقة الإشارة إلى المعاملات السابقة

بغرض دراسة تأثير طاقة الإشارة على نسبة التعرف قمنا بإجراء اختبار آخر على النظام, حيث أضفنا معامل طاقة الإشارة إلى كل نوع من المعاملات المجربة سابقا. النتائج المحصل عليها مبينة في الشكلين (9.4) و(10.4)

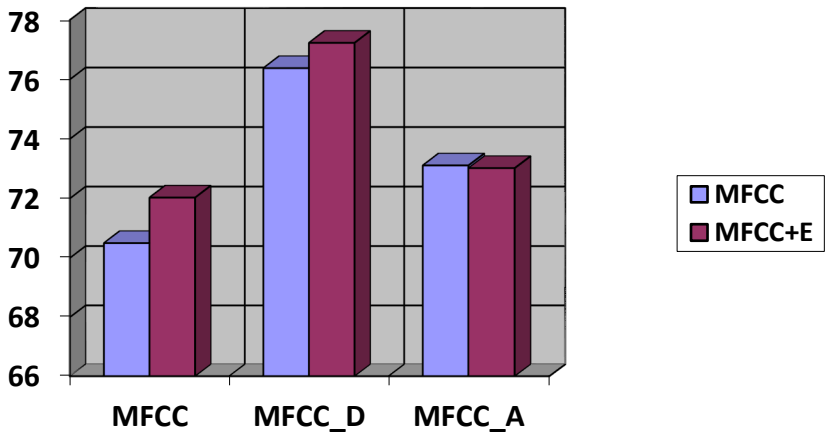
أ.مقارنة تغير نسبة التعرف بالنسبة لمتكلم من المدونة



شكل 9.4.مقارنة نسب التعرف حسب نوع المتكلم.

تأثير اضافة معامل طاقة الإشارة واضح في نتائج نسب التعرف المتحصل عليها لمختلف المعاملات , و هو أكثر وضوحا مع معاملات المشتقات الأولى لتقنية درجة النغم.

ب. مقارنة تغير نسبة التعرف بالنسبة لمتكلم خارج المدونة

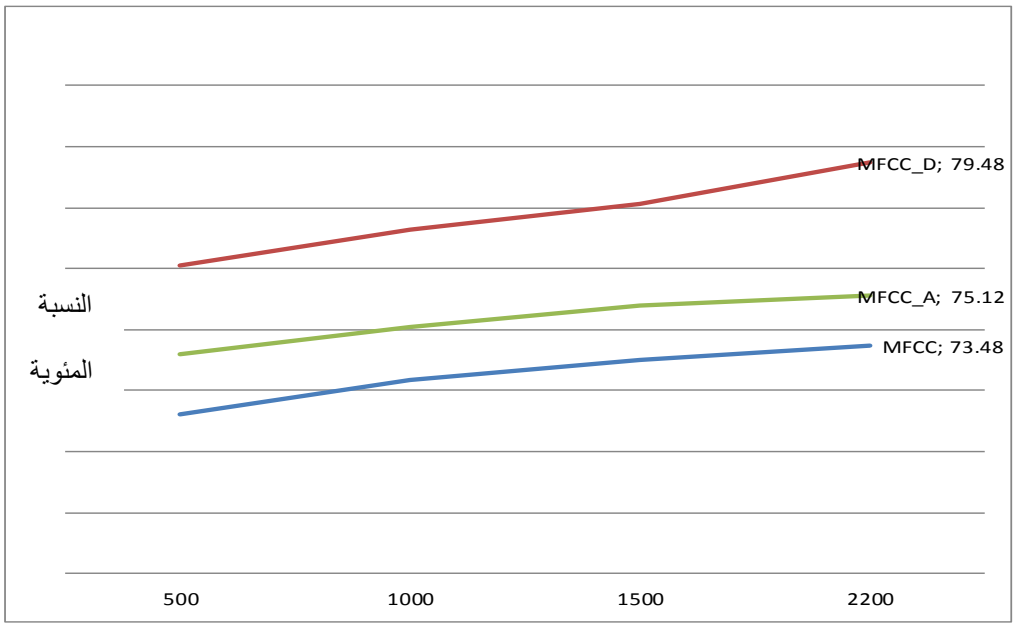


شكل 10.4.مقارنة نسب التعرف حسب نوع المتكلم.

نلاحظ من خلال الشكل أن اضافة معامل طاقة الإشارة بالنسبة لمتكلم من خارج المدونة لم يؤثر كثيرا في نسب التعرف و خاصة بالنسبة لمعاملات ΔΔMFCC.

6.دراسة تغير نسب التعرف الخاصة بكل نوع من المعاملات بدلالة تغير حجم المدونة:

لدراسة تأثر نوعية المعاملات المستخدمة بحجم المدونة , قمنا بتطبيق النظام عدة مرات مغيرين في كل مرة عدد التسجيلات المستعملة في مدونة النظام و مراقبة تغيرات نسبة التعرف و ذلك بالنسبة لكل نوع من معاملات التحليل المستعملة في البحث، حيث إن الاختبار تم على مجموعة مؤلفة في المجموع من 2200 تسجيل صوتي ,نقوم في البداية بتجريب النظام باستخدام 500 تسجيل صوتي فقط مختارة عشوائيا من التسجيلات المجهزة, ثم نرفع العدد الى 1000 تسجيل ثم الى 1500 تسجيل و نختم التجربة بالعدد الكلي المتوفر 2200 تسجيل . يظهر المنحنى التالي تغير نسبة التعرف عند تغيير حجم المدونة:



الشكل 11.4: منحنى تغير نسبة التعرف بدلالة حجم المدونة

نلاحظ أنه كلما زاد حجم المدونة في النظام , زادت نسبة التعرف بالنسبة للأنواع الثلاثة من المعاملات المستخدمة, مما يعني أن للمدونة و حجمها أثر ظاهر على النتائج المتحصل عليها من النظام, الا أن هذا التغير يتفاوت من تقنية تحليل لأخرى, حيث نلاحظ أن مجال التغير عند استخدام معاملات MFCC و عند استخدام معاملات ΔMFCC كان أكبر من التغير المسجل عند استخدام معاملات ΔΔMFCC أين سجلنا أقل تغير في نسبة التعرف و هذا ما يدل على أن استخدام معاملات ΔΔMFCC اقل عرضة للتأثر بحجم المدونة و بالتالي على نسبة تعرف النظام.أما استخدام تقنيتي MFCC و ΔMFCC يفرض علينا إعطاء أهمية كبيرة للمدونة من حيث الحجم و نوعية التسجيل.

و رغم أن زيادة حجم المدونة يؤدي إلى زيادة في نسبة التعرف إلا انه يؤدي أيضا إلى زيادة في العمليات الحسابية، و بالتالي زيادة في وقت المعالجة خاصة مرحلة التدريب، إلا أن هذا ضروري لزيادة فاعلية نظام التعرف بصورة عامة، لأنه بعد عملية التدريب سنتحصل على نماذج سهلة الاستخدام عند الاختبار حيث لا يظهر لحجم مدونة التدريب اثر في مدة التعرف، لهذا يتوجب علينا إنشاء مكتبة من التسجيلات باللغة العربية يتكفل بها مجموعة من الباحثين لتسهيل و تسريع وتيرة العمل بالنسبة للباحثين في هذا الميدان.

7. الخاتمة:

لقد قمنا في عملنا هذا بإنجاز نظام للتعرف الآلي على الكلام العربي و استخدام أنواع مختلفة من المعاملات الناتجة عن تحليل إشارات طيف الكلام لأجل الوصول إلى أعلى دقة ممكنة و استنتاج نوع المعاملات الأنجع من بين التقنيات المختارة و ذلك بالاعتماد على تقنية إحصائية في نظام التعرف.

لتقييم أداء النظام أجرينا العديد من التجارب و الاختبارات لتحديد كل مرة التقنية التي تسمح بتقييم أكبر عدد من الكلمات بشكل صحيح، فكانت النتيجة أن استعمال معاملات ΔMFCC يسمح لنماذج ماركوف المخفية من التعرف على الكلام العربي بشكل جيد و تحسين أدائه بالمقارنة مع النتائج المحصل عليها

عند استخدام معاملات MFCC و ΔΔMFCC, كما عززنا عملنا هذا باضافة معامل آخر عند التحليل و هو استخدام معامل طاقة الاشارة و ملاحظة تأثير هذا الأخير على نتائج التعرف أين لاحظنا زيادة في نسبة التعرف بالنسبة لكل نوع من المعاملات.

تبين لنا من خلال متابعة أداء النظام أن لحجم المدونة تأثير كبير على عملية التعرف، فزيادة حجم المدونة أدى إلى زيادة كفاءة النظام بصورة ملحوظة و هو ما يستدعى إعطاء أهمية بالغة لحجم المدونة عند انجاز هذا النوع من الأنظمة.

بصورة عامة بالنسبة لمختلف تقنيات التحليل و من أهم النتائج التي تحصلنا عليها في هذا الاختبار هو تفوق تقنية ΔMFCC بغض النظر عن حجم المدونة صغير, و هي تعطي نتائج أفضل كلما رفعنا حجم المدونة.

و قد قمنا بتطبيق هذا النظام على 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 متقاربة في نسبة التعرف حيث نستنتج منها أن النظام يعطي أفضل نسب تعرف عند استخدام تسجيلات لمتكلم من مدونة التدريب بغض النظر عن نوع معاملات التحليل المستخدمة.

خاتمة

خاتمة

إن ميدان التعرف الآلي على الكلام قد حضي بالكثير من الاهتمام ليكون التعامل بين الإنسان و الآلة أكثر مرونة و فاعلية، حيث نلاحظ أن هناك العديد من طرق تحليل الإشارات و استخراج المعاملات المميزة للإشارات لغرض استخدامها في انجاز الكثير من أنظمة التعرف على الكلام و لكل طريقة منها محاسن و مساوئ، و قد حضيت تقنية التحليل باستخراج معاملات درجة النغم MFCC بالكثير من الاهتمام نظرا لنتائج التعرف العالية التى توفرها بغض النظر عن طريقة التصنيف المستخدمة، حيث وجد نتيجة للأبحاث المتعددة أن هذه التقنية تعطي نتائج جيدة، و لذا فإننا في عملنا هذا لانجاز نظام للتعرف الآلي على الكلام العربي يعتمد على المعاملات المختلفة التي توفرها هذه التقنية و المقارنة فيما بينها بعدة طرق، متفرقة و مجتمعة وفق مراحل متعددة تمكننا من التعرف بشكل ناجح و فعال.

و للحصول على أفضل النتائج في تحسين أداء النظام المقترح للتعرف الآلي على الكلام قمنا بإجراء عدة تجارب على مجموعة من المتغيرات تتمثل في: تغير نوع المعاملات المستخدمة و حجم المدونة بالإضافة إلى إضافة معامل إشارة الطاقة للحالات السابقة.

يمكننا من خلال العمل الحالي استنتاج ما يلي:

1. معاملات المشتقات الأولى لتقنية درجة النغم ΔMFCC, تساعد على التعرف بسرعة محسوسة و أداء عاليين كما يتبين من خلال متابعة أداء النظام في التعرف.
2. إضافة معامل طاقة الإشارة سمح بزيادة كفاءة النظام في التعرف.
3. لحجم المدونة تأثير كبير على عملية التعرف، فزيادة حجم المدونة له تأثير فعال بزيادة دقة التعرف على الرغم من أن تأثيرها يكون واضحاً بزيادة الزمن المطلوب لإتمام عملية التعرف.

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

و قد أظهرت نتائج التعرف المختلفة التي تم التوصل إليها أن النظام كان قادراً على التعرف على الكلمات باستخدام معاملات المشتقات الأولى لدرجة النغم و ذالك بنسبة 79.48% في حالة المتكلمين من مدونة التدريب.

تعتبر هذه النتائج التي تم الحصول عليها مشجعة و مع ذلك فإن تنفيذ نظام للتعرف الآلي على الكلام العربي في ظل ظروف واقعية يتأثر بالعديد من المشاكل مثل الضوضاء، و نوعية التسجيل، وصعوبة المفردات و حالة المتكلم، حيث يبقى مهمة صعبة لا تزال تحتاج إلى العديد من البحوث للتعامل مع جميع جوانب التخاطب بين الإنسان و الآلة.

التوصيات:

- إجراء مقارنة بين طرق تحليل أخرى للإشارات الكلامية قصد تحسين أداء الأنظمة و النتائج المتحصل عليها.
- توسيع المدونة باستخدام عدد أكبر من المتكلمين و التسجيلات الخاصة باللغة العربية و تسجيلها بواسطة متكلمين مختلفين للمساعدة على تحسين أداء أنظمة التعرف الآلي على الكلام و مساعدة الباحثين في القيام بدراساتهم و أبحاثهم.
- استخدام طرق تصنيف أخرى كالشبكات العصبية الاصطناعية باعتماد معاملات درجة النغم في التعرف الآلي على الكلام.

قائمة المراجع و المصادر

1. أ. **رضا بابا أحمد**، اللسانيات الحاسوبية، مشكل المصطلح و الترجمة ، ط د ت

2. أ.**عيسى مومني** ، بيبليوغرافيا اللسانيات .ط د ت.

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

4. **إبراهيم أنيس**، الأصوات اللغوية، مكتبة اللأنجلو المصرية، مصر، 1999 م.

5. **ألان بونيه**، ترجمة: علي **صبري فرغلي**، الذكاء الاصطناعي واقعه و مستقبله، عالم المعرفة، الكويت، 1993.

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

7. **د مسعود بن محمد دادون**..دروس في اللسانيات العامة – المستوى الصوتي

8. **د نهاد الموسى** . حصاد القرن في اللسانيات، الأردن ، مؤتمر عبد الحميد شومان ،ط د ت

9. **د.أميمة الدكاك و د. يحي حربا**. نظام تعرف الكلمات المعزولة باستخدام نموذج ماركوف المخفي و استخدامه لإعطاء أوامر صوتية للتحكم في ذراع متحركة, المعهد العالي للعلوم التطبيقية و التكنولوجيا .2005.

10. **د.خليفة ميساوي**، المصطلح اللساني وتأسيس المفهوم ، ،ط.1 ،2013 .

11. **د.نهاد الموسى**. العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية .

12. **الربيعي و فاضل محسن**, مقدمة في العمليات التصادفية, دار الكتب بغداد العراق .2000م.

13. **رشيد عبد الرحمن العبيدي**، معجم الصوتيات ، مكتبة .د مروان العطية، العراق، 2007 م.

14. **سعد عبد العزيز مصلوح**، دراسة السمع و الكلام، عالم الكتب، مصر، 2005 م.

15. **سمير شريف أستيتية**, الأصوات اللغوية رؤية عضوية و نطقية و فيزيائية، دار وائل للنشر، الأردن 2002 م.

16. **مصطفى حركات**، اللسانيات العامة و قضايا العربية ،المكتبة العصرية صيدا بيروت ط1. 1998م

17. **صالح بن إبراهيم السدراني ، صالح بن عبد العزيز** ، تجميع وصيانة الحاسب الآلي ، المنهج الصحيح لمن أراد صيانة الكمبيوتر و التصليح ، إعداد المادة العلمية وصياغتها. الإصدار 10

18. **علم الأصوات اللغوية الفونيتيكا**، د.عصام نور الدين، دار الفكر اللبناني، لبنان، 1992م.

19. **مسعود بو دوخة**، محاضرات في الصوتيات، بيت الحكمة، الجزائر، 2013 م.

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

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

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

23. **د.خليفة ميساوي**، المصطلح اللساني وتأسيس المفهوم. ص 30، ط.1 ،2013

قائمة الرسائل و المجلات

الرسائل

1. **بركات ابراهيم، مصطفى الزين، طلال بهاء الدين**، تمييز الكلام العربي باستخدام الشبكات العصبونية الاصطناعية، بكالوريوس الشرف في علوم الحاسوب جامعة السودان للعلوم و التكنولوجيا، 2006 م.

2. **د.مسعود داود** رسالة دكتوراه التعرف الآلي على الكلام العربي المتصل جامعة الجزائر 2 /2016

3. **شهلة عبد الوهاب عبدالقادر**، تمييز أصوات الأرقام العربية، قسم أنظمة الحاسبات،المعهد التقني الموصل، الموصل،العراق.

4. **فهد العتيبي**, رسالة ماجستير(التعرف على الكلام العربي لمتحدث واحد باللغة العربية), سنة الطباعة 1422

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

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

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

المجلات

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

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

3. **بان 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

4. **باسل يونس ذنون الخياط**، المتتابعات الماركوفية مع التطبيق في المجال اللغوي، كلية علوم الحاسبات و الرياضيات، جامعة الموصل، المؤتمر العلمي الثاني للرياضيات–الإحصاء و المعلوماتية 2009 م.

5. **عمر صابر قاسم،** تهجين أنموذج ماركوف المخفي باستخدام شبكة يلمان العصبية الاصطناعية مع التطبيق، مجلة الرافدين لعلوم الحاسبات و الرياضيات المجلد(11)العدد (1) 2014 م.

6. **بان أحمد حسن متراس،** رشا رعد هادي المولى، استخدام نماذج ماركوف المخفية في التعرف على صور الوجه المشوه والمركب من صورتين مشوهتين، المجلة العراقية للعلوم الإحصائية (20) 2006 م.

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

8. **باسل يونس ذنون الخياط, شعاع محمود عزيز, زينة فالح صالح العجوز،** خوارزمية حاسوبية لتقدير مرتبة سلسلة ماركوف مع التطبيق، المجلة العراقية للعلوم الإحصائية (19) 2011 م.

9. **باسل يونس ذنون الخياط** هنادي داود سليم و مازن محمد غانم العناز، خوارزمية مقترحة للتحقق من مراوحة مشاهدات تنمذج وفقا للنموذج الماركوفي مع التطبيق، المجلة العراقية للعلوم الإحصائية (18) 2010.

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

11. **خليل إبراهيم السيف،سجى جاسم محمد،** كبس ملفات الكلام باعتماد خواص الإشارة، مجلة الرافدين لعلوم الحاسبات و الرياضيات المجلد(4)العدد (1) 2007 م.

12. **غيداء عبد العزيز** ، التعرف على النص العربي المطبوع باستخدام نموذج ماركوف المخفي، مجلة الرافدين لعلوم الحاسبات والرياضيات المجلد(7) العدد(2) 2010م.

13. **فاتن بشير عبد الأحد، إنعام غانم سعيد،** تقطيع الكلمة العربية إلى أحرف و تمييزها، مجلة الرافدين لعلوم الحاسبات و الرياضيات المجلد(2)العدد (2) 2005 م.

14. **وجدي عصام,** أساسيات لابد لأي مبرمج معرفتها مقال موقع انفورمتيك 2013/09/21

15. **وليد العناتي**، " الدليل نحو بناء قاعدة البيانات للسانيات الحاسوبية " ،مجلة علوم اللسان و التكنولوجيا، العدد 15 2009.

16. **صالح بن إبراهيم السدراني ، صالح بن عبد العزيز** ، تجميع وصيانة الحاسب الآلي ، المنهج الصحيح لمن أراد صيانة الكمبيوتر و التصليح ، إعداد المادة العلمية وصياغتها. الإصدار 10

17. **أ، أشرف مبارز زنقور ، د.داود سليمان مشاط**، تاريخ الرياضيات و إسهامات العرب و المسلمين ، مركز النشر العلمي، جامعة الملك عبد العزيز ،2010.

18. **وليد أحمد العناتي** ، "مقال اللسانيات الحاسوبية العربية رؤية ثقافية" .مجلة فكر و نقد عدد82. أكتوبر 2006.

قائمة المراجع الأجنبية

1. A.K. Mohamed Cherif, Reconnaissance Automatique de la Parole par les HMM en Milieu Bruité, Doctorat, Université Annaba, Algerie, 2009.
2. DOUIB Oualid, reconnaissance automatique de la parole arabe par cmu sphinx 4, magister, université FERHAT Abbas, Algerie, 2013.
3. J.Parker, " Profile Hidden Markov Models" ,Biophysical Journal , Vol. 17 , P. 1335 – 1348,(1999).
4. L.R. Rabiner, " A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition " ,Proceedings of IEEE , Vol. 77 ,(1989).
5. L.R.Rabiner, and B.Juang, "An introduction to hidden Markov models", IEEE ASSP MAGAZINE, (1986).
6. M. Stamp, " A Revealing Introduction to Hidden, Markov Models", IEEE , Vol. 51,(2003).
7. M.Hayashi, "Hidden Markov models to identify pilot instrument scanning and attention patterns", proceeding of IEEE International Conference on Systems, Man and Cybernetics, 2889-2896, Washington D.C.,Oct.5-8, (2003).
8. Pilhofer ،Michael (2007). Music Theory for Dummies ..Ellis,D.(2001) " Sequence Recognition " , Computer, Speech,and Language, Vol.1

9. S.Duran, "Keyword spottingusinghiddenmarkovmodels", M.Sc. Thesis, BogaziciUniversity, (2001).

10. T.Aarnio, "Speech recognition withhidden Markov models in visual communication", M.Sc. Thesis, College Computer Sciences, Turku-University, (1999

11. T.Pham, "Alignment-free sequencecomparisonwithvectorquantization and hidden Markov models", proceeding of IEEE of the ComputationalSystemsBioinformatics, (2003).

12. Zhai,C. (2003). " A Brief Not on the Hidden Markov Models ", IEEE Transactions on PAMI , Vol. PAMI –18 .

13. Humphreys, R, & Zumach,W. Automated star/galaxy discrimination with neural networks. The Astromical Journal,103 (1) ,318(331),(1992)

14. Kitano, H.."Designing neural networks using genetic algorithms with graph generation system". Complex system , 4 (4) ,461(476),(1990).

15. Miller ,G.F., Todd,P. M.,& Hegde,S.,U. "Designing neural networks using genetic algorithms ".In Schaker, J. D.(Ed.),proceeding of the Third International Conference on Genetic Algorithms(pp.379(384),(1989).

16. Richard P .Lippmamn's "article An introduction to computing with neuron nets ",IEEE magazine, April 1987.

17. Siddiqi, A. A.,&Lucas,S. M."A comparison of matrix rewriting versus direct encoding for evolving neural networks". In Proceeding of 1998 IEEE,(1998).

18. Yao, X. "Evolving artificial neural networks", Proceedings of the IEEE,87(9),1423(1447),(1999).

19. Cox ,D.R. and Miller, H.D. (1965). " The Theory of Stochastic Processes ", John Wiley & Sons , Inc. , New York , USA.

20. D.Jurafsky and J.Martin 'An intriduction to natural language proccessing computational linguistices and speech recognition'. Printice hall.2000

قائمـــــة الجداول

الجدول 1.3 ابرز أنواع الشبكات العصبية واستخداماتها 78

الجدول 1.4.بعض جمل المدونة 87

جدول 2.4: ترميز الفونيمات 88

جدول 3.4 : مواصفات ملف التحليل config الخاص بمعاملات MFCC 91

جدول4.4: مواصفات ملف التحليل config الخاص بمعاملات Δ MFCC 91

جدول4. 5: مواصفات ملف التحليل config الخاص بمعاملات ΔΔMFCC 92

جدول 6.4. بعض أنظمة البرمجة المستعملة في التعرف الآلي على الكلام 93

قائمة الأشكال

الشكل 1.2 : الجهاز الصوتي و الجهاز التنفسي 39

الشكل 2.2 :النموذج الكهربائي لآلية انتاج الكلام 41

الشكل 3.2: تمثيل زمني لإشارة كلمة "خمسة" 42

الشكل4.2: نافذة كلام عند إنتاج صوت في الزمن(أسفل) يعلوها طيف الإشارة على هذه النافذ 42

الشكل5.2:نافذة كلام عند إنتاج جملة في الزمن(أسفل) يعلوها السوناغرام الموافق. 43

الشكل 6.2: تصحيح نهايات الكلمة اعتمادا على معدل العبور بالصفر 49

الشكل 7.2 :مثال تحديد بداية و نهاية كلمة إغلاق 49

الشكل 8.2: خواص مرشح التعزير الأولي من حيث المطال و الطور 51

الشكل 9.2: الضرب بنوافذ هامينغ 52

الشكل 10.2: شكل نافذة هامينغ 53

الشكل 11.2: استجابة المسار الصوتي 54

الشكل12.2 :مراحل تحليل المركبات الممثلة 56

الشكل 13.2: بنك الترشيح Mel 57

الشكل 14.2: مخطط عمل تقنية MFCC 57

الشكل 15.2:مخطط استخراج سمات الإشارة بتقنية التنبؤ الخطي 62

الشكل 1.3 : مكونات نظام نموذجي للتعرف الآلي على الكلام 68

الشكل 2.3: النموذج الثبوتي 75

الشكل 3.3: نموذج الأيسر– الأيمن 75

الشكل 4.3: يمثل معمارية شبكات الطبقة المفردة 80

الشكل 5.3: يمثل معمارية الشبكات متعددة الطبقات

الشكل1.4 : مراحل عمل نظم التعرف الآلي على الكلام 83

الشكل4. 2: تسجيل و معاينة الأصوات بواسطة برنامج Praat 87

الشكل 3.4: مخطط استخلاص صفات الإشارة المدخلة 90

الشكل 4.4: ملف التهيئة 94

الشكل 5.4 طريقة عمل الأداة HVite 95

الشكل 6.4. نسبة تعرف النظام لمتكلم من المدونة 95

الشكل 7.4. نسبة تعرف النظام لمتكلم من خارج المدونة 97

الشكل 8.4.مقارنة نسب التعرف حسب نوع المتكلم. 98

شكل 9.4.مقارنة نسب التعرف حسب نوع المتكلم. 99

الشكل 10.4.مقارنة نسب التعرف حسب نوع المتكلم. 99

الشكل 10.4.مقارنة نسب التعرف حسب نوع المتكلم. 100

الفهـــــرس

الإهداء

التشكرات

المقدمة العامة أ-ب-

ج-د-هـ

الفصـــــــل الأول: اللسانيات الحاسوبية

1.المقدمة 8

2. المفاهيم الرئيسة للسانيات الحاسوبية 8

3. أسس اللسانيات الحاسوبية 12

3. 1. اللسانيات العامة 14

2.3. علم الحاسبات 15

3.3. علم الذكاء الاصطناعي 16

4.3. علم المنطق 18

5.3. علم الرياضيات 18

4.تحديات اللسانيات الحاسوبية 19

5. اللسان العربي والحاسوب 20

1.5 في المستوى الصوتي 21

2.5.في المستوى الصرفي 21

3.5.في المستوى التركيبي 22

4.5.في المستوى الدلالي 22

5.5.في المستوى المعجمي 23

6.جهود الباحثين العرب لحوسبة اللسان العربي 24

7.وصف المستوى الصوتي للسان العربي لأغراض البرمجة الحاسوبية 25

8. أثر العولمة على اللسانيات الحاسوبية 28

9.التعرف الآلي على الكلام العربي المنطوق 30

10.التحديات التي تواجه برامج التعرف الآلي على الكلام العربي 32

11. الفرق بين التعرف على الكلام والتعرف على الصوت 33

12. الفرق بين التعرف الآلي على الكلام في اللغتين العربية والإنجليزية 34

13. بعض البرامج الموجودة للتعرّف الآلي على الكلام العربي 35

14.الخاتمة 35

الفصـــــل الثاني : تحليل الإشارة الكلامية

1. مقدمة 38

2.الجهاز الصوتي عند الانسان 38

3.النموذج الكهربائي لآلية إنتاج الكلام 41

4.آلية تمثيل إشارة الكلام 42

5.مراحل عملية التعرف الآلي على الكلام 43

1.5. رقمنة الإشارة الكلامية و تحديد نهايات الكلمات 43

1.1.5. اللاقط Microphone 43

2.1.5. بطاقة تحصيل الصوت voice card 44

3.1.5 .وسط التحصيل المحيط 44

2.5. مرحلة تسجيل الكلمات 44

3.5. كشف حدود الكلمات Endpoint detection problem 46

4.5. حساب طاقة الإشارة 46

5.5. استنتاج أشعة سمات الإشارة الكلامية 50

1.5.5. ترشيح الإشارة بمرشح التعزيز الأولي 50

2.5.5. تقسيم الإشارة إلى نوافذ زمنية 51

3.5.5. التمرير عبر النافذة 52

6.طرق استنتاج أشعة سمات الإشارة الكلامية 54

1.6.التحليل باستخدام الكبستروم (Cepstrum) 54

2.6. تقنية درجة النغم Mel Frequency Cepstral Cooeffcients (MFCC) 56

3.6.تطبيق المرشحات الترددية 58

4.6.تطبيق اللوغاريتم على خرج المرشحات 58

5.6. تطبيق التحويل الجيبي المتقطع Discrete Cosine Transform 59

6.6. تقنية المشتقات الأولى و الثانية لمعاملات الكبستروم 59

7.6.تحليل الاشارة باستخدام(LFCC) Linear Frequency Cepstrum Coefficients 60

8.6. التحليل بتقنية التنبؤ الخطي Linear Predictive Coding 61

7 .صعوبات نظم تمييز الكلام العربي 62

8 .الخاتمة 64

الفصـــــل الثالث: طرق التصنيف المستخدمة في التعرف الآلي على الإشارات الكلامية

1. مقدمة 66

2. تصنيف نظم التعرف الآلي على الكلام 66

1.2.التصنيف وفق طريقة الكلام 66

2.2. التصنيف وفق نمط الكلام 67

3.2. التصنيف وفق عدد الكلمات 67

3. خطط نظم التعرف الآلي على الكلام 68

4. طرق التصنيف 69

1.4. نماذج ماركوف المخفية HMM 69

1.1.4.النظرية التصادفية 70

2.1.4.فضاء الحالة 70

3.1.4. سلاسل ماركوف المخفية 71

4.1.4.نماذج ماركوف المخفية 72

5.1.4.عناصر نماذج ماركوف المخفية 73

6.1.4.أنواع نماذج ماركوف المخفية 75

2.4.الشبكات العصبية الإصطناعية 76

1.2.4.الشبكات العصبية 77

2.2.4.أنواع الشبكات العصبية 78

3.2.4.أساليب التعليم في الشبكات العصبية 78

4.2.4.معمارية الشبكات العصبية 79

5. الخاتمة 82

الفصـــــل الرابع : الدراسة التطبيقية وتحليل النتائـــج

1. المقدمة 84

2.بناء نظام للتعرف الآلي للكلام 86

1.2. مرحلة تهيئة البيانات 86

1.1.2 انشاء المدونة 86

2.1.2 مرحلة تقسيم الجمل و ترميز الفونيمات 88

2.2.مرحلة تحليل الإشارات الصوتية و استخلاص الصفات 89

1.2.2 تهيئة ملف التحليل .config 90

3. مرحلة إعداد المصنف classifier 93

4.التعرف بواسطة الأداة HVite 95

5. اختبار النظام بتقنيات التحليل المختارة و تحليل النتائج 96

1.5.متكلم من مدونة التدريب 96

2.5.متكلم من خارج مدونة التدريب 97

3.5.إضافة معامل طاقة الاشارة 98

6.دراسة تغير نسب التعرف الخاصة بكل نوع من المعاملات بدلالة تغير حجم المدونة 100

8. الخاتمة 101

الخاتمة العامة 104

قائمة المصادر و المراجع

قائمة الجداول

قائمة الأشكال

فهرس الموضوعات

الملخص

ملخص

قمنا في هذا العمل بإنجاز دراسة مقارنة بين تقنيات استخلاص صفات الإشارة الكلامية و قد اخترنا معاملات درجة النغم MFCC (Mel Frequency Cepstral Cooeffcients) و المشتقات الأولى لدرجة النغم ΔMFCC, ثم المشتقات الثانية لدرجة النغم ΔΔMFCC, بالإضافة الى معامل طاقة الإشارة (Energy) , عن طريق استخدامها في بناء نظام للتعرف الآلي على الكلام العربي المتصل، باعتماد أدوات نماذج ماركوف المخفية كأداة للتصنيف, و مقارنة دقة التعرف المتحصل عليها من النظام تبعا للمعاملات المستخدمة.

لأجل ذلك قمنا بمجموعة من الاختبارات و التجارب التي توصلنا من خلالها إلى نتائج جيدة و مشجعة باعتبار المشاكل و الصعوبات التي تعترض أنظمة التعرف عموما كالتداخل بين الأصوات و الضوضاء و غيرها، و قد وصلنا في الأخير الى أن معاملات ΔMFCC, هي التي أعطت أعلى دقة تعرف خاصة عند اضافة معامل طاقة الإشارة , حيث تحصلنا على نسبة تعرف وصلت 79.48% بينما بلغت نسبة التعرف عند استخدام معاملات MFCC : 73.48% و نسبة 75.12%, بالنسبة لمعاملاتΔΔMFCC.

كلمات المفاتيح: العلاج الآلي للكلام, التعرف الآلي على الكلام, التحليل الصوتي, معاملات MFCC، معاملات درجة النغم.

Résume

Dans ce travail, nous avons construire un système automatique de reconnaissance de la parole arabe en utilistant les outils de Markov cachés comme méthode de classification. après en a fait une étude comparative entre les techniques d'extraction des paramètres du signal parole : MFCC et ΔMFCC et ΔΔ MFCC et l'énergie, et en comparant la précision de la reconnaissance obtenue à partir du système en fonction de paramètres d'analyse utilisée.

Pour cette raison, nous avons effectué une série de tests et d'expériences dans lesquels nous avons obtenu des résultats positifs et encourageants en compte tenu des problèmes et des difficultés rencontrés dans les systèmes de reconnaissance générale tels que le bruit.

Nous avons enfin constaté que les paramètres Δ MFCC donnaient la plus grande précision en particulier, lorsque le paramétre d'énergie du signal est ajouté, nous avons obtenu un taux de reconnaissance de 79,48%, et pour les paramètres MFCC étaient 73,48% et de 75,12% pour les paramètres ΔΔMFCC.

Mots clés: Traitement automatique de la parole, Reconnaissance vocale, Analyse vocale, parametres MFCC, Outils HTK.

Abstract

In this work, we completed a comparative study between the techniques of extracting the signal properties. We chose MFCC and Δ MFCC, ΔΔMFCC, and Energy, Building an automated system for automatic Arabic speech recognition by adopting HTK tools as a way of classification, and comparing the accuracy of the recognition obtained from the system according to parameter of analysis used.

For this reason, we have carried out a series of tests and experiments in which we have obtained good and encouraging results considering the problems and difficulties encountered in general recognition systems such as noise, and other interferences. Finally, we found that the Δ MFCC parameters gave the highest accuracy Particularly when the parameter of energy is added, we obtained a recognition rate of 79.48%, while the MFCC parameters were 73.48% and 75.12% for ΔΔMFCC

Key words: Automatic Speech processing, Automatic Speech Recognition, Voice Analysis, MFCC Transactions, HTK Tools.