الصوت وسيلة المستقبل للتعامل مع الحاسبات

د . مصطفى الشافعي

شهدت السنوات الأخيرة إهتماما متزايداً في إستخدام الصوت كوسيلة هامة للتعامل مع الحاسب.فبالإضافة إلى حاسة اللمس المتمثلة في لوحة المفاتيح وحاسة البصر المتمثلة في شاشة العرض، فإن إضافة حاسة السمع يضيف بعدا جديدا في التعامل والتفاعل البشري مع الحاسب، وبالتالي سيفتح آفاقا جديدة في تطبيقات الحاسب.

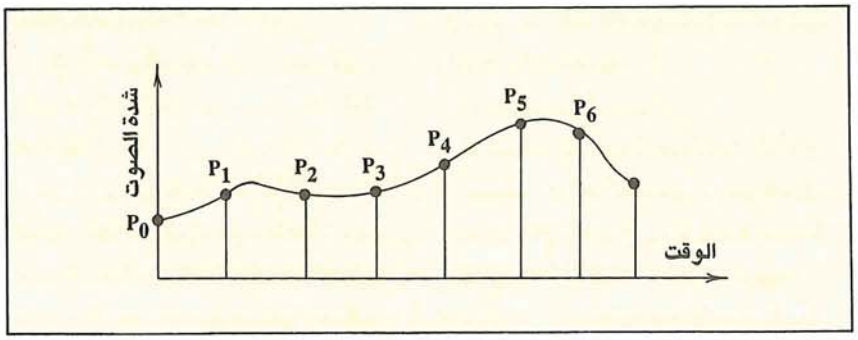
ومن أهم هذه التطبيقات إستخدام الحاسب في التعليم والتدريب وأجهزة مساعدة المعوقين و إرسال وتخزين الرسائل الصوتية والسكرتارية الآلية والحصول على المعلومات من بنوك المعلومات باستخدام الهاتف، وفي كثير من الخدمات الهاتفية الأخرى كالإستعلامات مثلا، وكذلك في تطبيقات كثيرة تتطلب إستعمال اليدين والبصر مع ضرورة إدخال البيانات أو طلب معلومات من الحاسب كما في عمليات فرز الطرود وفحص المنتجات الصناعية وقيادة الطائرات المقاتلة.

ويمكننا عموما أن نصنف التطبيقات الصوتية للحاسب إلى ما يلي:-

تشفير الكلام

تعد عملية تحويل الصوت إلى صورة رقمية اللبنة الأساس في نظم الهاتف الحديثة، وكذلك في النظم الحديثة لدمج المعلومات والأصوات. ولعل أكثر الطرق إنتشارا هي طريقة التضمين الرمزي للنبضة (Pulse Code Modulation) ، وتعتمد هذه الطريقة على قياس شدة الصوت على فترات متساوية بمعدل ثمانية آلاف مرة في الثانية ثم تحويل شدة الصوت إلى تتابع من الأرقام. وهذا يعني أن كل ثانية من الصوت تتحول إلى ثمانية آلاف رقم. وهذه الطريقة مع أنها الأكثر شيوعا لبساطتها، إلا أنها تعد مكلفة من حيث الذاكرة المطلوبة لتخزين المعلومات الصوتية أو نقلها.

ويبين شكل (١) أن شدة الصوت لا تتغير كثيرا من مادة إلى أخرى لأن هناك علاقة تربط هذه النبضات بعضها مع بعض. وقد بُنيت طريقة (ADPCM) على إستغلال هذه الخاصية لتقليل عدد الأرقام اللازمة وتمثيل ثانية من الصوت بأربعة آلاف رقم. وتوجد حاليا طرق أخرى كثيرة أكثر تعقيدا لتشفير الكلام، ولكنها تتفاوت في مقدرتها على تشفير الكلام وفي جودة الصوت الناتج. وبصورة عامة يمكن القول بأنه كلما تم تشفير الكلام كلما إزدادت العمليات الحسابية المطلوبة وقلت جودة الصوت الناتج. وهناك طرق حديثة مازالت تحت الإختبار لتحقيق تشفير يصل إلى ٦٠٠رقم / الثانية مثل طريقة (CELP) وهي طريقة تعتمد على بناء قاعدة من النماذج الرياضية تمثل تجويف الفم والحنجرة والأنف والقصبة الهوائية، وقاعدة أخرى لمصادر الصوت تمثل الترددات التي تصدر بالحبال الصوتية وغيرها. وعند إستقبال الصوت يقوم خوارزمي بالبحث في قاعدة النماذج الرياضية وقاعدة مصادر الأصوات عن أمثل نموذج رياضي وأمثل مصدر للصوت للحصول على أقل خطأ بين الصوت



● شكل (١) تحويل الصوت إلى تتابع من الأرقام تمثل شدة الصوت عند فترات متساوية .

وهناك أيضا منسق للنصوص يعمل على كل من جهازي الفارابي والرائد.

وبغض النظر عن اللغة المستخدمة تتطلب برامج النشر المكتبي العربية الإمكانات اللازمة مثل إمكانية الكتابة والرسم معا والتحكم في حجم الحروف ووضعها واستخدام الألوان واستخدام البرنامج لكتابة الوثائق المختلفة مثل الكتب والمقالات والإعلانات والنشرات والجرائد والمجلات وغيرها، وبالإضافة إلى هذه المتطلبات العامة فإن هناك متطلبات تختص بها اللغة العربية ومنها أنواع الخط المختلفة مثل الديواني والثلث والفارسي والكوفي والنسخ والرقعة وغيرها، ومن تلك المتطلبات أيضا معالجة الحروف وأشكالها والأرقام والكتابة من اليمين إلى اليسار.

٢ ـ معالجة اللغة العربية الطبعية

هناك عدد من البحوث في هذا المجال، وكما يظهر من تلك الجهود فإن الإنتاج مازال في طور البحث إلا في بعض المراحل الأولية في تحليل اللغة مثل التحليل الصرفي، وتتركز البحوث في هذا المجال في الوقت الحاضر على التحليل النحوي بما في ذلك الإعراب، أما بالنسبة لتحليل المعاني وإنشاء الجمل فما زالت الجهود قاصرة. وجدير بالذكر أن عناصر معالجة اللغة العربية يتفاعل بعضها مع بعض بحيث تصعب المعالجة بصورة متسلسلة خلال مراحل الصرف والتركيب والإعراب وتحديد المعنى، وهذا بدوره يؤدي إلى ضرورة تصور هذه المراحل كمجموعة من البرامج التي تعم في آن واحد وتتعاون فيما بينها لمعالجة النص.

٣ـ تحليل وإنتاج الكلام المنطوق

ما زالت معظم الجهود في هذا المجال في طور البحث بالرغم من وجود بعض الأنظمة البسيطة بشكل فعلي، ويتطلب هذا الجهد الكثير من البحث والإبتكار لطرق جديدة تختلف عن تلك التي تستخدم للغات أخرى لاختلاف الخصائص الصوتية للغة العربية عن غيرها، وتتحدد عناصر معالجة الكلام المنطوق في تحديد مجموعة من الأصوات التي تتركب منها الأحرف العربية ثم التعرف على هذه الأصوات وبالتالي الحروف، وتبرز هنا أيضا مشكلة تحديد الأصوات التي يبدأ وينتهي بها الحرف في الكلمة والكلام المتصل، وغالبا ما تستخدم النماذج الإحصائية في معالجة الكلام المنطوق، وهذا بدوره يؤدي في هذه الحالة إلى عدم إمكانية الحصول على نتائج صحيحة دائما، وهناك طريقة أخرى تعتمد على تعريف عناصر صوتية يتركب منها الحرف ويتم تمثيل ذلك بمجموعة من القواعد التي تضبط تركيب الأحرف، كما يجب أيضا تطوير معايير موحدة لتحديد الأصوات التي تتكون منها الأحرف ومجموعة محددة من الكلام المنطوق لاختبار البرامج المطورة حتى يمكن مقارنتها.

٤ ـ التعرف علي الكتابة العربية

نال التعرف على الكتابة العربية حظاً وافراً من جهود الباحثين. وعلى الرغم من ذلك لم يظهر منها على مستوى الإستخدام إلا القليل. وينطبق هنا نفس وضع الكلام المنطوق بالنسبة للطرق التي تستخدم في المعالجة ومنها الإحصائية والتركيبية والهندسية وما يجب تطويره من معايير.

٥ ـ نظم قواعد وجداول المعلومات

لم يتم في هذا المجال إلا ما هو موجود على جهاز آبل ماكنتوش ونظام الملفات على أجهزة صخر ونظم قواعد المعلومات على كل من الرائد والفارابي ، أما بالنسبة للأجهزة الأخرى فيتم استخدام طريقة تعريب البرامج، ولا يتطلب هذا الموضوع معالجة خاصة باللغة العربية إلا من حيث إظهار المعلومات بهذه اللغة وما يتعلق بتطوير الأوامر المستخدمة بها .

٦ ـ الرسم الآلي

تقتصر الجهود في هذا المجال على برنامج رسم مبسط على أجهزة صخر، ويبدو المجال فسيحا للإنتاج والإبداع في هذا المجال، وكما هو الحال بالنسبة لنظم المعلومات فيلزم هنا تطوير الأوامر باللغة العربية وإيجاد واجهة للتعامل باللغة العربية.

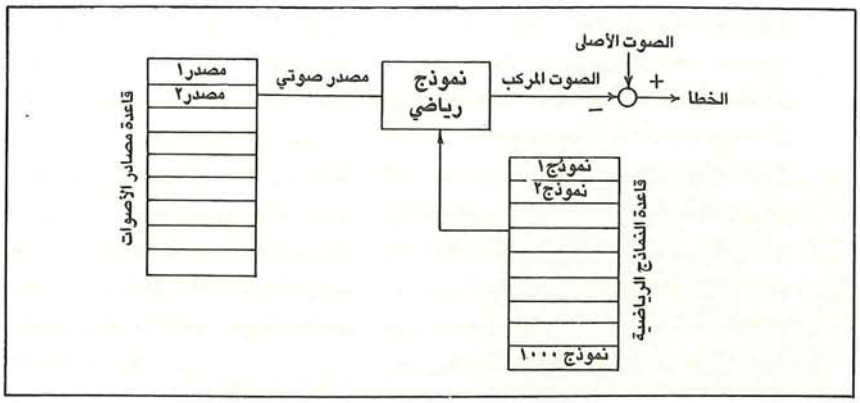
٧ ـ الدراسات الإسلامية

حظيت الدراسات الإسلامية بجهد وافر لتطوير بعض البرامج على الحاسب الآلي، فقد كان هناك عدد من البحوث التي اهتمت بالقرآن الكريم والسنة النبوية، ويبرز من هذه الجهود إلى طور الإستخدام الفعلي برامج القرآن الكريم والحديث المطورة على أجهزة صخر، ويجب توسيع مجال هذه الجهود، حتى تغطي النواحي الأخرى من الدراسات الإسلامية مثل الفقه وغيره، وتتمثل المعالجة في هذا المجال بصورة أساس في معالجة النصوص وتطوير نظم قواعد للمعلومات تتيح للمتخصص سهولة التعامل واستخراج المعلومات، وتطوير نظم تختص بالقواعد الفقهية والأحكام.

٨ ـ التعليم بمساندة الحاسب

بدأ الاهتمام باستخدام الحاسب الآلي في التعليم حديثا مما أدى بدوره إلى بعض الجهود في هذا المجال، وقد برز من الجهود بعض البرامج مثل برامج تعليم المواد الدراسية على أجهزة صخر وبرامج التعليم «الأستاذ». وما زالت هناك الحاجة للدراسات التي تقوِّم دور هذه البرامج وبالتالي تحسين وإبتكار طرق جديدة لمساندة التعليم بوساطة الحاسب الآلي، ويتطلب تطوير هذه البرامج مشاركة المدرسين في مراحل تطوير المادة وبالتالي ضرورة معرفتهم للتعامل مع الحاسب الآلي والبرامج.

يتبين مما ورد ذكره أن هناك كثير من الجهود في مجال تعريب الحاسب الآلي، وعليه يجب تنسيق الجهود وتكريسها حتى يمكن الحصول على النتائج بأقل جهد ممكن، ويجب على كل الأطراف العاملة في هذا المجال من أفراد وشركات خاصة وجهات القطاع العام التعاون فيما بينها ووضع خطة شاملة لمعالجة الموضوع ومن ثم توزيع المهام بينها وتدعيم الجهود بكل الإمكانات حتى يصبح استخدام الحاسب الآلي باللغة العربية متاحا على الصعيد العام وذلك لغرض تسخير هذه التقنية المتقدمة للإستفادة منها.



● شكل (٢) تشفير الصوت بطريقة CELP .

المكون والصوت الأصلي، شكل (٢). وتتميز هذه الطريقة في عدم الحاجة إلى تخزين أو إرسال الصوت الأصلي حيث يتم تخزين الرقم الخاص بالنموذج الرياضي والرقم الخاص بالمصدر الصوتي. إلا أن صعوبة هذه الطريقة تتمثل في الحجم الهائل للعمليات الحسابية المطلوبة والتي تقدر بنحو ١٤٠ـ٢٠٠ مليون عملية في الثانية.

التعرف على الكلام

يتطلب هذا النوع من التطبيقات تعرف الحاسب على الكلمات الصوتية مباشرة أو التعرف على المتحدث نفسه من تحليل الخصائص المميزة لصوته. ويعمل العلماء منذ أكثر من عقدين من أجل محاكاة الإنسان في مقدرته على تمييز الكلمات والتعرف على معاني الجمل ولكن يبدو أن الطريق سيكون طويلا، ويرجع ذلك تارة إلى قصور معرفة الكيفية التي تتم بها هذه العملية في العقل البشري، وتارة أخرى إلى عدم وجود الحاسب القادر على محاكاة العمليات الكثيرة المعقدة التي تتم في عقل الإنسان بنفس السرعة والكفاءة.

وهناك عدة عوامل تؤثر على قدرة نظم التعرف على الكلمات وكفاءتها تتضمن ما يلي:ـ

١ ـ إنفصال الكلمات

يُسهِّل نطق الكلمات منفصلة بينها سكتة خفيفة كثيرا من معرفتها مقارنة بنطقها متصلة، ويرجع ذلك إلى أن نطق الكلمات عادة ما يتأثر بتفاعل الكلمة بما قبلها أو بعدها .

٢ ـ عدد الكلمات

كلما زاد عدد الكلمات المطلوب التعرف عليها كلما كان هناك كلمات قريبة النطق من بعضها البعض حيث يصبح التمييز بينها أكثر صعوبة. هذا بالإضافة إلى إطالة الوقت اللازم لتحديد الكلمة المنطوقة.

٣ ـ عدد المتحدثين

إذا كان المتحدث واحداً يمكن تدريب نظام التعرف على الكلمات ليصبح على درجة عالية من الدقة، يرجع هذا إلى أن النماذج الرياضية المستخدمة لتخزين وتحليل الكلمات الصوتية عادة ما تكون حساسة لخصائص المتحدث لأنها تمثل نموذجاً رياضياً للمَخْرج الصوتي للمتكلم (كالفم والقصبة الهوائية ...إلخ) عند نطق كلمة أو صوت معين وهذا يختلف بالطبع من شخص إلى آخر.

٤ ـ تتابع الكلمات

يؤدي إستخدام عدد محدود من القواعد التي تحكم تتابع الكلمات وإتصالها في الجمل إلى يسر وسهولة التعرف على الجمل، ولذلك فقد أمكن الحصول على نسبة تمييز عالية تصل إلى ٩٩٪ في التطبيقات التي تلتزم بتركيبات محددة للجمل المستخدمة.

٥ ـ الوسط

يقصد بالوسط هنا ما إذا كان هناك ضوضاء حول المتحدث أم لا وكذلك إرتفاع الصوت وإنخفاضه، فالتحدث إلى الحاسب بصورة عادية وفي مستوى الضوضاء المعتاد في المصانع والمنازل والمكاتب مازال حلماً بعيد المنال.

نظام التعرف على الكلمات

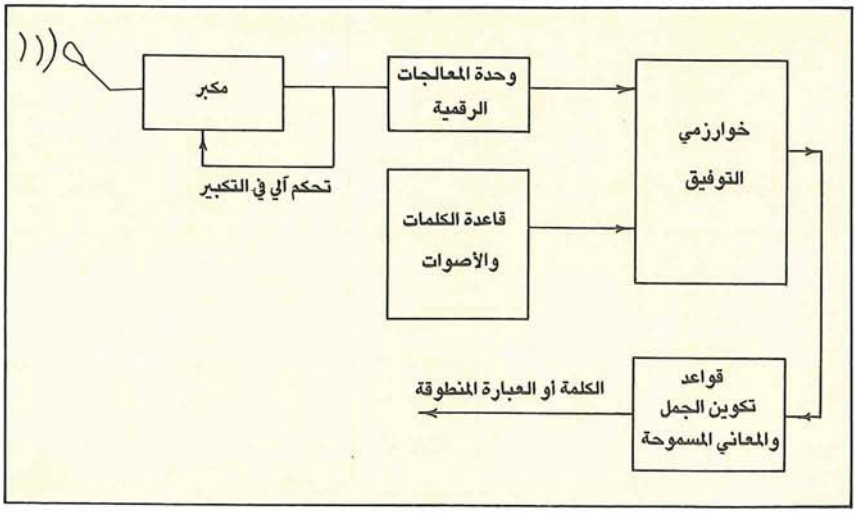
يتكون نظام التعرف على الكلمات الصوتية، شكل (٣)، من المراحل التالية:ـ

المرحلة الأولى، يتم إستقبال الإشارة الصوتية كما يتم ضبط مستواها بالتحكم الآلي في تكبير هذه المرحلة.

المرحلة الثانية، يتم تحويل الكلمة الصوتية إلى تتابع رقمي ثم تعالج رقميا لتحديد خصائصها المميزة طبقاً لنموذج رياضي معين.

المرحلة الثالثة، وهي مرحلة التوفيق (Matching)، وفيها يتم إستدعاء قاعدة الكلمات أو الأصوات المخزنة لمقارنة هذه الكلمات بالكلمة المنطوقة. ويستخدم في ذلك خوارزمي التوفيق، وهو خوارزمي يقوم بمحاولة إيجاد أقرب إثنين إلى خمسة من الكلمات المخزنة إلى الكلمة المنطوقة.

المرحلة الرابعة، ويتم فيها مقارنة الكلمات المرشحة مع تركيب الجملة وفقا لقواعد سبق تحديدها، وأيضا وفقا للمعاني المسموح بها، وعلى هذا يمكن رفض الكلمات



● شكل (٣) نظام التعرف على الكلمات .

المرشحة التي لا تتفق مع تركيب الجمل المسموح بها أو المعاني السابق تحديدها.

القراءة الصوتية للنصوص

يهتم هذا المجال بتحويل النصوص المخزنة في ذاكرة الحاسب إلى أصوات. وقد حظي هذا المجال بأهمية متزايدة في الآونة الأخيرة في عديد من التطبيقات مثل القراءة لدى فاقدي البصر وتلقي الرسائل الإلكترونية عن طريق الهاتف وفي مجالات التدريب والتعليم بمساندة الحاسب، وفي عدد من التطبيقات الصناعية وأجهزة الإنذار والأمان. وقد لقى هذا المجال نجاحاً كبيراً وهو يعد من المجالات القليلة التي إستطاع فيها الإنسان برمجة الحاسب لمحاكاة الوظائف العقلية للإنسان بدرجة عالية. ويعتمد الحصول على مزيد من التقدم في هذا المجال على مدى فهمنا لأصوات اللغة والتفاعل بينها وكذلك قواعد وتركيب اللغة ومعانيها.

نظام القراءة الصوتية للنصوص

يتطلب هذا النظام أولا معالجة النصوص غير المشكَّلة حتى تصبح مشكَّلة تماما ، وربما يتطلب قاموسا لترجمة الرموز والأرقام كتابة حرفية مثل ١٢٣=٢+٪....إلخ. أو الكلمات التي تنطق بطريقة تختلف عن كتابتها كأوائل السور في القرآن الكريم مثل: حم ، يس، أو بعض الكلمات الأخرى مثل: هذا، هذه ..إلخ.

ويتم بعد ذلك تقسيم النص إلى مقاطع كالتي تنطق في نفس واحد. فعادة يتم الوقوف عند نهاية الجمل القصيرة أو عند حروف الوصل. ومن الممكن في الجمل الطويلة الوقوف عند حروف الجر أو عند بداية جملة خبرية أو بعد عدد معين من الكلمات. بعد ذلك يتم تحويل النص إلى صورة حرفية أولية باستخدام شكل واحد من الحروف وتقسم الكلمات إلى مقاطعها ويكتب النص كما ينطق تماما،فمثلا :-

عَلَىْ - - - عَ - لَ

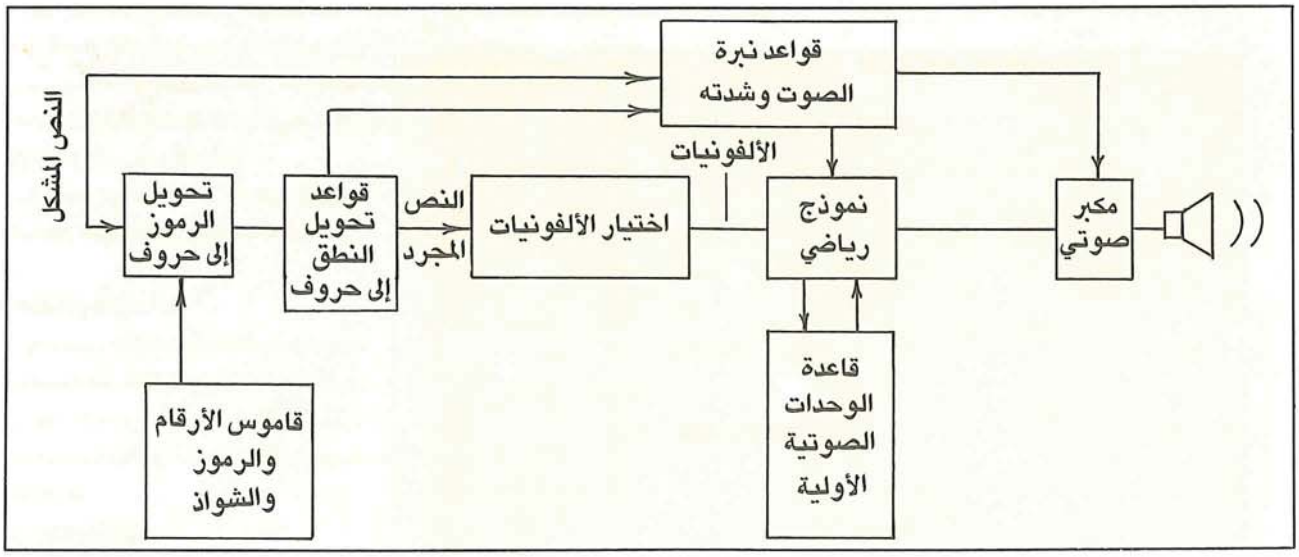
مَنْ رَأَى العَنْبَرْ - - - مَ رْ - رَ - ءَ ل - عَ مْ - بَ رْ

واللغة العربية غنية بقواعد مختلفة لإدغام الحروف أو إقلابها أو إخفائها أو إظهارها، والنص الناتج من هذه المرحلة يسمى النص المجرّد، شكل(٤).

يُحلّل بعد ذلك النص المجرّد للحصول على تقنين أدق للأصوات. فمن المعروف مثلا أن حركة الفتحة إذا أتت بعد حرف مرقّق كالباء في «بَاب» تنطق بطريقة مختلفة عما إذا أتت بعد حرف مفخّم كالضاد في «ضَار». وهذه الأصوات المختلفة لنفس الوحدة الصوتية تسمى الوحدات الصوتية الأولية في اللغة (الألفونيات/ Allophones) .

يتم بانتهاء هذه المرحلة تحويل النص المجرد سابق الذكر إلى سلسلة من الألفونيات بإستخدام قواعد إختيار الألفونيات، ولإعادة تكوين أصوات هذه الوحدات الأولية، فإنها تُستدعى من قاعدة الوحدات الصوتية الأولية للغة حيث تكون مخزنة في صورة جداول تحدد النموذج الرياضي المطلوب لتكوين كل وحدة صوتية أولية في اللغة. وإذا إستمعنا بعد ذلك إلى الصوت الناتج من هذا النظام نجده مازال آليا خالياً من الإنفعالات والإرتفاع والإنخفاض، لذلك فهناك معالجة أخرى تتم عادة بالتوازي مع ماسبق وهي التي تغير من نبرة الصوت وشدته تبعا للسياق حتى يبدو طبعيا ، شكل (٤) . فمثلا عندما يكون هناك حرف نداء فإن المقطع الأول ينطق بشدّة أعلى من غيره ثم تتناقص شدة الصوت تدريجيا حتى نهاية الجملة. وبالمثل هناك منحنى خاص لشدة الصوت يحدد العُلو والإنخفاض في الصوت للجمل التقريرية والجمل الإستفهامية وللتعجب... إلخ ، فإذا تمت إضافة ذلك سيبدو الصوت شبيها إلى حد كبير بالصوت الصادر من الإنسان.

وكما نرى فإن الحاسب لا يزال يفرض علينا لغاته وطرق التعامل معه، ولكن سيتمكن الإنسان إن شاء الله في وقت قريب أن يفرض عليه لغته وطريقته الطبعية في التعامل والإتصال. وجدير بالذكر أن التقدم سواء أكان في قراءة النصوص أو التعرف على الكلام يعتمد إلى حد كبير على اللغة من حيث الأصوات وقواعد تركيب الجمل والكلمات وكذلك معاني اللغة. ولهذا جدير بنا أن نهتم بالبحوث في هذا المجال، حتى نفرض لغتنا الحية على الحاسب ونستفيد من التطبيقات الكثيرة التي يتيحها هذا المجال.



● شكل (٤) نظام القراءة الآلية للنصوص العربية .