الحاسبات الالكترونية تتكلم

ترجمة سناء سعد الدين عثمان
جامعة الموصل

مترجمة عن مقالة
عن مجلة The COMPUTER SPEAKS SPECTRUM
الصادرة في شهر آب/١٩٧٩

الحصول على حديث غير محدود المفردات من نص

ما زال الحصول على الحديث من الحاسب عملية بالغة التعقيد رغم التقدم الالكتروني وحتى الآن يستخدم في بعض التطبيقات القليلة كتعليم فاقدي البصر أو تعليم الطلبة اللغات الاجنبية. ويعتقد الاخصائيون بان تطور البحوث والتقدم في أنابيب أشعة الكاثود وآلات الطباعة سيجعل هذا المجال ينتشر انتشاراً واسعاً في غضون السنوات العشر المقبلة، والتكنيك المتبع للحصول على الحديث من الحاسب هو تسجيل

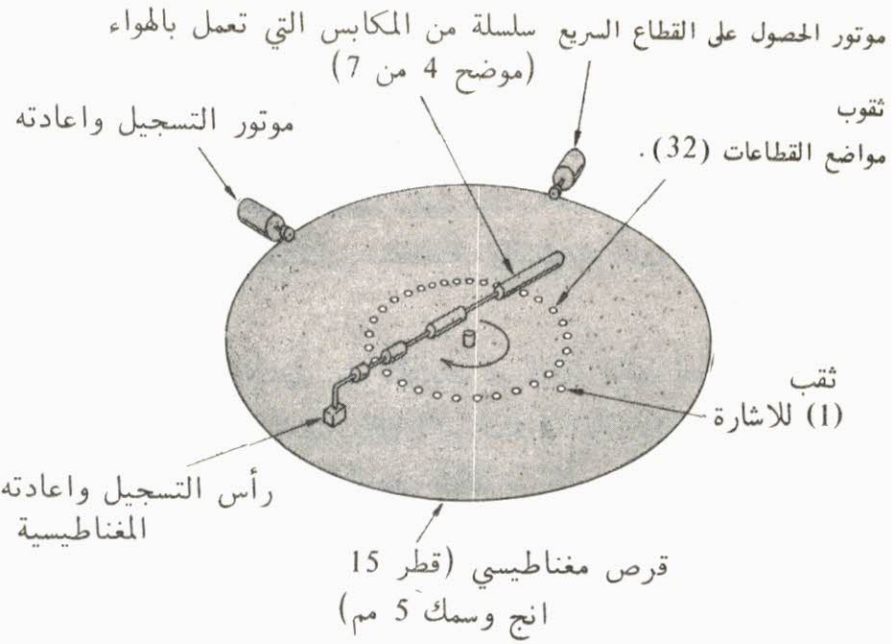
مناظر أو تسجيل رقمي مضغوط للحديث ، حيث نحصل على الحديث من نص يتم تسجيله من قبل اما على ما يشبه القرص السمعي أو قراءته من دوائر متكاملة في الذاكرة . ولعدم القدرة لمحاكاة كلام الانسان وقلة المفردات اللغوية التي يمكن الحصول عليها تكون كفاءة الحاسبات ذات الاغراض العامة التي تستخدم التكنيك السابق منخفضة ، ولذلك يتم الان الحصول على الحديث من نص .

ومن الصعب الحصول على الحديث ذو كفاءة عالية لمحاكاة كلام الانسان وذلك بسبب الخواص الديناميكية لآلية تكلم الانسان لكن الابحاث التي تجري الآن في المختبرات توصلت لرفع هذه الكفاءة ، الاّ ان كلام الحاسب ما زال غير مألوف لأذن الانسان .

عيوب التسجيل المناظر

الحصول على حديث ما هو الا عبارة عن تسجيل لحديث مناظر له تحت سيطرة الحاسب ، حيث يختار الحاسب الجمل المطلوب اخراجها ، اذ يوضع على شريط ممغنط تسجيل سماعي على مسار والعنوان الرقمي وزمن المعلومات على مسار آخر ، والعيب هنا الحصول على الجمل الملائمة بتتابع بطيء من الشريط وهذا لا يناسب أغلب تطبيقات الحاسب .

ويمكن الحصول على صورة طبق الاصل من الحديث بواسطة اسطوانة دوارة صغيرة عليها جزء من فيلم يحتوي على مسارات للصوت وتستخرج الجمل الملائمة منه بواسطة رأس اعادة سماع ضوئية ، ولكن في مثل هذه الاجهزة يكون للحديث صدى فيها .



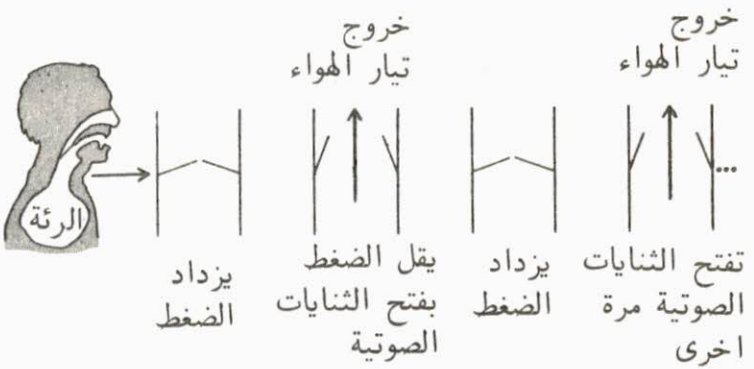
شكل ـ 1 ـ يوضح القرص السماعي ذو الحصول العشوائي لنظم التعليم والمعلومات مغطى بمادة الشريط السماعي. والتي يمكن الوصول الى اي قطاع ملائم فيها في كسر من الثانية بواسطة موتور اختيار القطاعات وفي نفس الوقت وبواسطـــــــة مكبس هوائي يتحرك الرأس فوق المسار العامودي المنتخب (يستغرق سماع القرص ٢٠ دقيقة).

ومن الاجهزة المشابهة والاكثر مرونة هو القرص السماعي للحصول العشوائي شكل (1) وفيه يكون الشريط السماعي على هيئة اقراص مسجل عليها الحديث او اي اصوات اخرى لمدة عشرين دقيقة مقسمة الى (128) مسار متحدة المركز ويستغرق المسار عشر ثوان ويقسم المسار الى (32) قطاع وبذلك نجد ان زمن الوصول الى اي موقع من (4096) موقع ـ [(128) مسار × (32) قطاع = 4096] ـ

هو كسر من الثانية وتكون عناوين المسارات والقطاعات على (12 BIT) ويحرك رأس اعادة التسجيل الى المسار الملائم مكبس يعمل بضغط الهواء وفي الوقت نفسه تدور الاسطوانة الى القطاع الملائم بواسطة عجلة صغيرة تدور بسرعة فائقة على حافة القرص وبمجرد الوصول الى المسار والقطاع المختارين تنخفض رأس اعادة التسجيل على القرص ثم تدور الاسطوانة تحت رأس اعادة التسجيل بالسرعة العادية بواسطة عجلة صغيرة بطيئة السرعة ، ويستخدم القرص السماعي في حاسب (Blato-Computer) لتعليم قواعد اللغات الاجنبية للمبتدئين ومن مميزات هذه الأجهزة انها تعطي اصوات ذات جودة عالية وحصول عشوائي حقيقي ومعدل بيانات منخفض ، فـ (19 BIT) تكفي لاختيار جملة تستغرق بضع ثوان ، اما عيوبها فتشمل الاحتياج الى الهواء المضغوط كما تلازمها مشاكل حركة الاجزاء الميكانيكية وتكون محدودة القدرة لتوليد جمل تامة وذلك لأن الجمل تتكون من كلمات مسجلة منفصلة ، فعند اعادة سماعها يكون لها تنغيم خاطىء حيث أن جمل الانسان ليست بينها فواصل بين الكلمات .

ميكانيكية حديث الانسان

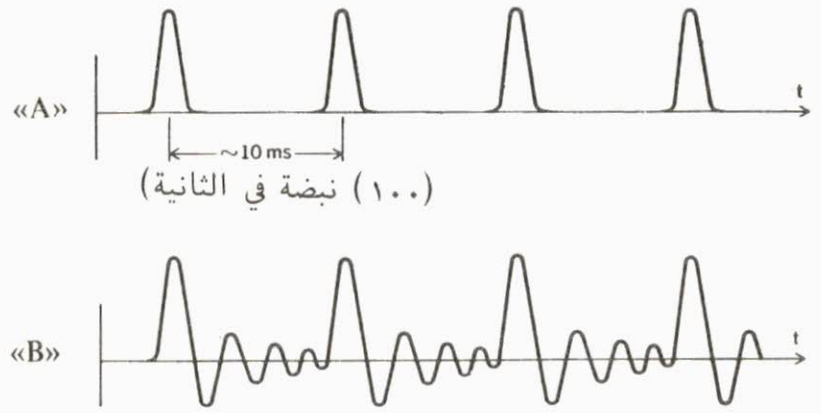
حيث ان اخراج الحديث الرقمي شكل من اشكال تخليق الحديث لذلك فمطلوب دراسة سريعة الى ميكانيكية حديث الانسان ـ شكل (2) ـ



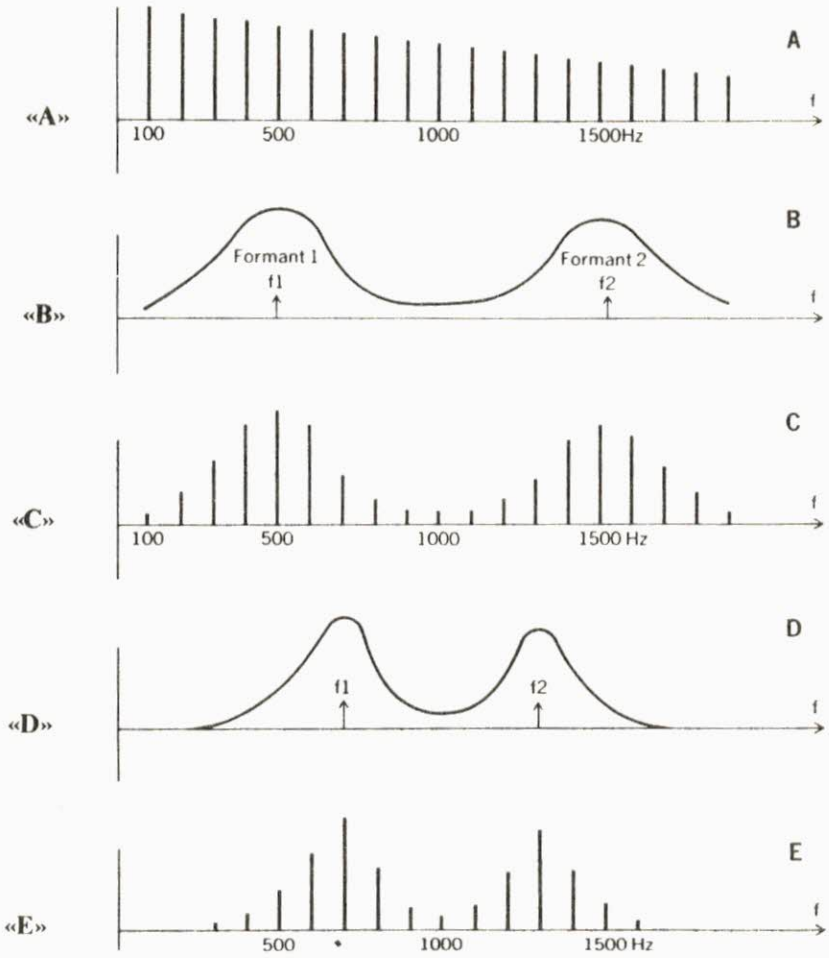
شكل ـ 2 ـ يوضح عمل الرئة والثنايات الصوتية كمذبذب مما يسمح بنبضات من الهوات داخل فتحة الفم وهذا هو منبع الاثارة للاصوات الرخيمة والاصوات الساكنة مثل V و Z الحركة الفعلية للثنايات الصوتية اكثر تعقيداً مما هو مبين بالرسم.

فحينما يزداد الضغط داخل الرئة ترتفع الثنايات الصوتية للخارج ثم تهبط مرة أخرى بانخفاض الضغط الناتج من خروج تيار الهواء وتتكرر العملية بطريقة دورية وتستغرق الدورة (100) مايكرو ثانية للذكر البالغ مقابل تردد مقداره (100) هرتز بينما للانثى البالغة يبلغ (200) هرتز ، ونفخ الهواء في الحنجرة يثيرها مما ينتج عنه اشارات صوتية في الزمن نفسه وهذه الاشارة تكون في منتهى التعقيد ونجد من الافضل تحليل التردد السائد Frequency Domain في نفس الزمن ، وشكل (3-A) يوضح تحليل فورير للنبضات الدورية المثارة وشكل (4-A) يوضح تمثيل للتردد السائد والذي يتكون من التردد الاساسي ومضاعفاته ، ولان النبضات حادة فهناك توافق ذو معنى يصل الى عدد من الهرتزات وشكل (4-B) يوضح تجاوب التردد لفتحة الفم حينما يكون اللسان والشفاه والفك متحركين لعمل صوت

رخيم ومن المألوف ان يوصف تجاوب هذه الترددات بتقريبها مع عدد من ذبذبات الرنين التوافقية ويسمى في هذه الحالة التشكيل الموجي المميز. وهذا التشكيل الموجي المميز Formants (تناغم الاصوات ـ متوافق) مع أقل ترددين يمثل أهم رنين لتحديد ترخيم الصوت وفي الحقيقة يمكن ان يقدر تردد طبق الاصل على أساس أن متوسط طول القناة الصوتية (من الثنايات الصوتية الى الشفاه) للذكر هي 17 سم، كما يمكن تقريب هذه القناة الى انبوبة مستقيمة مغلقة عند ثنايات الصوت ومفتوحة عند الشفاه وبذلك يكون الطول الموجي للتردد FF1 هو (68) سم ـ (4 × (17) سم = (68) سم) ـ ويصبح التردد F1 هو سرعة الصوت (330) م/ث مقسوماً على (0.68) م فيساوي (500) هرتز تقريباً.



شكل ـ 3 ـ (A) تمثيل مبسط لنبضات الهواء المثارة في فتحة الفم لدورة نمطية طبق الاصل حوالي (10) مايكروثانية لانثى بالغة (B) شكل موجي معقد جداً والذي من الصعوبة جداً تحليله في الزمن نفسه.



شكل ـ 4 ـ يبين التحليل في التردد السائد أكثر سهولة فتحليل فورير لنبضات الهواء المولدة في شكل (A–3) ممثل في التردد السائد بواسطة خط طيف للتردد الاساسي وتوافقه.

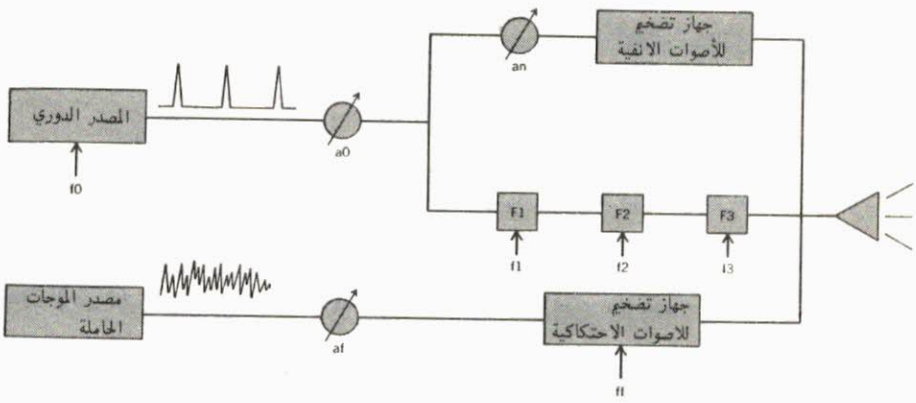
أما (B—4) فيمثل تجاوب التردد لفتحة فم ولسان وشفة وفك في اوضاع خاصة وهي مقربة لتردد رنين مزدوج (تشكيل موجي مميز) لمذبذب التوافق.

وبأمرار خط الطيف (A-4) خلال مرشحات مع تجاوب التردد لـ (B-4) ينتج لنا خط الطيف المرشح (C-4). المقابل لتحليل فورير لشكل الموجة المعقدة للوقت السائد (كما في شكل B-3) وهذا هو شكل التردد الذي نحتاج اليه لتوليد صوت محدد وتوليد الترقيم المختلف يتوقف على شكل فتحة الفم و (D-4) يوضح تغير تجاوب التردد ممثل بمقياس التردد، وبترشيح النبضات المولدة خلال تجاوب التردد المتعاقب ينتج تمثيل التردد السائد للترخيم الجديد (E).

وعندما تمر خطوط الطيف المثارة شكل (A-4) من خلال مرشح تجاوب التردد وشكل (B-4) فخط الطيف الناتج يكون معقد كما هو موضح في (C-4). أما شكل (B-4) فهو تحليل فورير للوقت السائد (TIME Domain) للاشارة الصوتية وهذه الاشارة تمثل الترخيم وللحصول على ترخيمات متعددة تتحرك الشفاه واللسان والفك باشكال مختلفة، وهكذا يتعاقب تردد الحنجرة مثل شكل (D—4) ومن الواضح بأن الدائرة التي تنتج اصوات مختلفة يمكن صنعها من مولد نبضات خلال مرشحات يمكن تغيير ترددا تها المركزية.

يمكن لدائرة الكترونية انتاج حديث معين من التردد السائد المتغير وشكل (5) يوضح مخطط بسيط لاصطناع المتغيرات، فالمتغيرات FO و aO تسيطرعلى التردد واتساع النبضات الدورية المثارة (الصوت) والمتغيرات FF و aF تسيطر على مدى التحكم واتساع الموجات الحاملة واشارة النبضات تشغل المرشحات الثلاثة F1 وF2 وF3 والذي ينتج عنه ترددات مركزية F1 وF2 وF3 المقابلة لتشكيل موجي مميز تستخدم للتقريب من تردد الفم كما ان هذه النبضات تمر خلال مذبذب والذي ينتج بالتالي خواص مميزة لكل حرف من الصوت، وكمثال لذلك هو الحاسب المتكلم (CT-1) وفيه تدخل لوحة اصطناع

الاشارات الى نظام الحاسب القياسي الصغير جداً والذي يعامل فيه كل متغير بعنوان منفصل ويكون اتساع الصوت واتساع الموجات الحاملة (8 BIT).



الحاسب المتكلم موديل CT-1

شكل (5)

مخطط مبسط للحاسب المتكلم (CT-1) لاصطناع المتغيرات ذات (8-BIT) ودرجة النغم (10) ، aO (اتساع الموجة الدورية المولدة) و aF اتساع الموجة الحاملة) و F1 F2 F3 الترددات الموجية المميزة و a_N فرع (الحروف الانفية) ويمكن السيطرة عليه بواسطة اجهزة معالجة غاية في الصغر Microprocessor

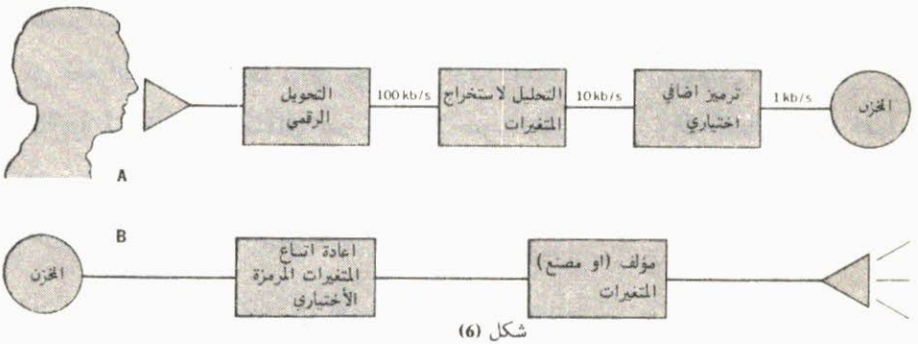
★ Nasal حرف انفي (الحروف التي تخرج من الانف مثل M, N)

★ Frication الحروف الاحتكاكية (الحروف التي يسمع الهواء عند نطقها مثل ل).

وحقيقة تكون المتغيرات بحالة مستقرة لفترة طويلة وهذا يعني ان جميع المتغيرات المصطنعة يجب ان تجدد باستمرار وبمعدل اكبر من الحديث المأخوذ من تغيرات شكل الفم ، أي يجب أن تجدد هذه المتغيرات بمعدل (100) مرة في الثانية لتعطي لنا انتقال تدريجي من صوت الى آخر . والتكنيك المتبع لعمل هذه المتغيرات هو ضغط التمثيل العددي للاشارة الصوتية مع متغيرات الترددات السائدة الملائمة للصوت .

الحديث الرقمي المضغوط لا يمثل كل الاصوات

يوضح شكل (6) طريقة لاصطناع المتغيرات لاخراج حديث رقمي مضغوط الى وسط مخزني فحديث الشخص برقم بحولي (100 kb/s) بواسطة محول ارقام مناظر للحديث وللتتبع الدقيق لاشارة الوقت السائد (Time Domain) للصوت والتمثيل الرقمي في مضمن ترميز النبضات في الخطوط التلفونية الطويلة حيث نجد بأنها تتكون من (8008 BIT) للحديث في الثانية أو (64 KB/S) .



شكل (6)

يحلل حديث الانسان في نظام الحديث المضغوط الى المتغيرات السائدة للتردد وتغير الزمن مع ترميزه اختيارياً وذلك لتقليل تكاليف التخزين ويمكن ان تستخرج البيانات من المخزن للحصول على الحديث.

وبسبب عدم التمثيل الكافي للوقت السائد فالمطلوب معدل بيانات عال للحصول حتى على صوت واحد في حالة مستقرة وشكل (A-6) يوضح تحليل فورير وتكنيك ربط لاستخراج متغيرات التردد المقابل للحديث الداخل. والتحويل من الوقت السائد الى التردد السائد يخفض معدل البيانات لحوالي (10 Kb/s)، وكمثال نجد أن في ماكينة الحاسب المتكلم المتغيرات تخزن في (8-BIT) ويجب ان نجدد بحوالي (100) مرة في الثانية وتنتج بيانات أولية بمعدل (7200 b/s) للحديث المضغوط وعلى هذا فزيادة ضغط الحديث يمكن ان يحسن متغيرات التردد وذلك بأخذ ميزة الحركة البطيئة للفم والترميز الذي يستخدم هذه المميزات يمكن ان يخفض معدل البيانات بحوالي (1000 b/s)، وهذه تعتبر النقطة الاقتصادية لخزن الحديث. وعملياً التحليل وبالتالي الترميز يمكن ان نحصل عليه من عملية تكاملية واحدة وبذلك يصبح الاساس لطريقة من طرق

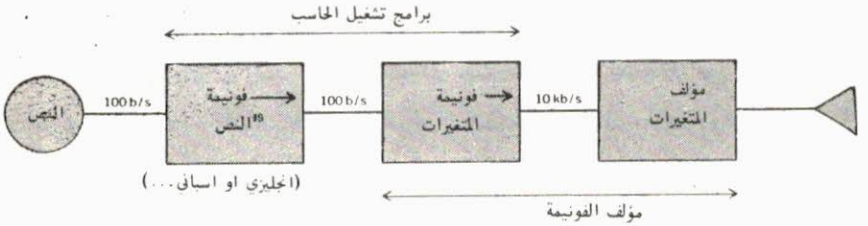
التسجيل الرقمي لمفردات الحديث.

فعند تكلم الانسان يمكن تسجيل هذا الحديث على قرص حاسب بشكل ارقام أو على شريط سماعي او تسجيل فوتوغرافي في حالة التشابه او التناظر ثم نعيد الحديث المسجل من وسط التخزين المخزن به.

وهنا نجد ان ميزة الحديث الرقمي المضغوط عن الشريط السماعي او القرص السماعي هو تجنب الاجهزة الميكانيكية والقدرة على تخزينه في الحاسب ونقله بمعدل بيانات اقل نسبياً، أما عيبه فهو أن طبيعة عمل الاشارة تجعل النظام مقتصر فقط على حديث الانسان وليس للاصوات الاخرى كالموسيقى مثلاً.

الحصول على حديث من نص

في الامكان الحصول على الحديث من سلسلة رموز مخزنة أو مولدة وشكل (7) يوضح نظام تحويل النص الى حديث، وفيه تحلل الرموز الى صوت وعادة تسمى فونيمة (Phonema) وهي وحدة كلام صغرى تساعد على تمييز نطق لفظ من آخر. وتكون الالفاظ في هذا النظام واضحة ثم يكبر ويكون لها درجة عالية للصوت لتأكيدها وفي المرحلة التالية تحول الى متغيرات ترددية ملائمة ونحصل على الحديث من متغير التأليف.



شكل (7)

الشكل يوضح نظام الحصول على الحديث من نص

والميزة الكبرى لمثل هذا النظام هو ان مفردات الكلمات والجمل غير محدودة فاي سلسلة رموز مستخرجة من مخزن أو مولدة يمكن ان تتكلم ، وبذلك يمكن الحصول على الحديث في هذا النظام على اي وسط شائع مثل الطابعة ، كما يلاحظ ان معدل التخزين والنقل المطلوبين قليلين جداً حوالي (100 B/S) مقارن بحوالي (1000 B/S) للحديث المضغوط وعيب هذا النظام انه غير متاح تجارياً لانه لا يفي بالغرض رغماً عن نجاحه في المختبرات ، وذلك لانه يعطي حديث عال الجودة كما ان مؤلف الفونيمة سهل في تشغيله عن مؤلف المتغيرات وأساساً تكون المدخلات لمؤلف الفونيمة هي سلسلة رموز بمعدل بيانات منخفض اما من الناحية الاخرى نجد ان المدخلات في مؤلف المتغيرات في منتهى التعقيد ومعدل البيانات عال . ويلاحظ ان هناك انواع من الفونيمة لكل منها مواصفات قياسية مختلفة حتى يمكن تحويل اللغات المختلفة كبيانات في البرامج .

البرمجة ليست بالعمل السهل

البرامج التي يعمل بها الحاسب لتحويل النص الى ترددات معقدة في كتابتها وذلك لتوقعنا ان كل صوت في الكلمة يمكن ان ينتج بشكل مشابه لطباعة خطاب كامل في كلمة وبذلك يجب أن يعطي البرنامج الترددات الملائمة لكل صوت كما هو مخزن وحيث ان شكل الفم يتغير تدريجياً من صوت الى آخر فمن الواضح ان الانتقال التدريجي يجب ان يتم بين كل مجموعة من المتغيرات الى الاخرى . كما ان التفصيلات الدقيقة لكيفية الانتقال السريع ووقت بدايته يختلف لكل صوت ، ولذلك يجب ان تحتوي البيانات على عدد كبير من الاوصاف الساكنة للصوت كما يجب ان يقوم البرنامج بعمل بعض العمليات المشابهة لشكل تغير في الفم من صوت الى آخر ومن طبقة الى أخرى وذلك لان طبيعة بعض الاصوات ديناميكية ، وكمثال لذلك نجد ان الحروف الساكنة

يتوقف فيها تيار الهواء لحظياً عند نطقها وكذلك تختلف الواحدة عن الاخرى ليس فقط في الخواص الاستاتيكية لكن بسرعة مميزة لكل واحدة عن الاخرى .

تحويل النص لحديث يختلف في مدى تعقده

تحويل النص الى الفونيمة ربما يكون سهل جداً أو معقد جداً ويعتمد هذا على اللغة ومدى الدقة المطلوبة ، وكمثال تحويل نص عادي من لغة الاسبرانتو لغة دولية مبتكرة بنيت على أساس مشترك من اللغات الاوروبية الى حديث بدرجة عالية من الجودة سهل جداً ، وهذا يرجع الى سهولة وعدم غموض الهجاء ونظام الصوت البسيط .

تنغيم حدة للصوت

يجب ان يؤخذ في الاعتبار عند تحويل النصوص الى أصوات عاملين مهمين الا وهما شدة وحدة الحديث الطبيعي وذلك لان عند هجاء بعض الجمل يجب ان تكون هناك بعض الالفاظ واضحة وطويلة بعض الشيء في نطقها كما أن طبقة الصوت ترتفع فيها كما تراعى الفواصل وعلامات الاستفهام في طبقة الصوت ، فنجد ان هناك بعض الاسئلة التي تحتاج للاجابة عليها بنعم او لا وتكون طبقة الصوت فيها منخفضة ومن حسن الحظ الى اغلب هذه الاسئلة تبدأ بكلمة تدل على السؤال وهذه الكلمات نعرفها وبذلك يكون ترميزها اقل صعوبة .

كما يمكن استخدام بعض الرموز الكبيرة لاعطاء المبرمج بعض السيطرة على ترتيب الجمل ويمكن لحرف هجاء واحد ان يعطي المبرمج سيطرة كاملة على الصوت وطبقاته وزمنه .

تكلم الحاسب الالكتروني مفيد في ادارة الاعمال

تجارياً تكون اغلب تطبيقات هذا النظام في ادارة الاعمال حيث تخزن المعلومات ويمكن للعميل ان يتصل تلفونياً بمركز الحاسب لتحديد مقدرة الانتاج او مقدرة المنتج ونجد ان الاجابة تأتي من الحاسب وهذه الاجابة مولدة من نص مكتوب وهو في الاصل عدد قليل من الكلمات وهذا التطبيق يمكن ان يكون بواسطة التسجيل الرقمي ولكن ميزة الانتاج في التسجيل المناظر هو تجنب الاجزاء المتحركة والميزة عن الحديث المضغوط هو ربما يكون السهولة لتصنيع وتطوير مؤلف المفردات بدون التوسع في طلب مجموعات جديدة من الكلمات وكتطبيق آخر لاخراج الحديث هو تزويدنا بتعليق مستمر للحركة على الشاشات المرئية كما يمكن كتابة هذه التعليقات على الشاشة ونجد هنا ان الحديث المضغوط له عيب هو ان معدل بيانات الحديث (1000 B/S) يمكن ان تستهلك اكثر من عرض الشريط في وحدة الحاسب الطرفية وعلى هذا يمنع الاستكمال اللحظي لصور الشاشة ونجد ان نقل السلسلة المرمزة الى مؤلف الفونيمة يستهلك عرض اقل من الشريط بحيث لا تتأثر الرسومات على الشاشة بدرجة كبيرة.