

”الحاسبات الالكترونية“

العدد ١٧ السنة ١١

١٤٠٧ هـ - ١٩٨٧ م

تحويل النص الى نطق في اللغة العربية

من بحوث مؤتمر بغداد العالمي الثاني لتقنيات وتطبيقات الحاسبات الالكترونية .

عبدالحق مرادي - المغرب

ترجمة: اكرم محمد عثمان - المركز القومي للحاسبات الالكترونية - بغداد

خلاصة :

يمكن تحويل النص المكتوب الى نص منطوق . اما باتباع طريقة التركيب القواعدي **Synthesis by Rules** أو بطريقة تركيب سلسلة العناصر الاساسية المحللة سلفاً . يعتمد نظام تركيب النص المنطوق في هذه الورقة على اعتاد الاسلوب الثاني . حيث يقوم النظام بتحويل النص العربي القياسي المدخل الى فيونيات **Phoneme** باستخدام مجاميع من الاصوات (**LPC Coded phone**) .

تتناول الورقة وصفاً موجزاً للغة العربية مع التأكيد على خصوصيتها بالمقارنة مع اللغة الانكليزية والفرنسية . كما تعرض الورقة طريقة تحويل النص العربي الى نطق باستخدام قاعدة تحويل الحرف الى نغمة مقابلة **Letter-phone** ، وطريقة بناء وتطور القاموس النغمي **diphone dictionary** . ان صلاحية وحدود الاصوات القياسية (**diphone**) موضحة بأفترض ان كل نغمة يجري تمثيلها بطور او صيغة واحدة في قاموس الاصوات القياسية .

وعلى هذا المبدأ تم تطوير (لوح الكتروني) يعمل بالاعتاد على المعالج المايكروبي **M6809** ويعتمد على الدائرة التكاملية الخاصة بالفهم اللغوي (**TMS S220**) حيث يسمح هذا اللوح بالتركيب اللغوي باستمرار (**Real time**) .

١ - المقدمة :

اللغة العربية يتكلم بها (١٥٠) مليون شخص ولم تجر لحد الان دراسات عميقة حول علاقة الانسان بالماكنة . وفي الوقت نفسه لم تجر أية بحوث بخصوص التكلم العربي من ناحية التركيب اللفظي أو من ناحية التمييز . قمنا في مختبراتنا بتعريف الصوت ، وخاصة بدراسة التركيب الصوتي بواسطة (LPC) او بواسطة المقاطع ، وقد شاركنا في اخراج اللعبة التثقيفية (Touch And Tell) من شركة تكساس بصيغتها العربية . ومن خلال خبرة الكاتب في التركيب الصوتي لمعجم محدد بأستخدام (LPC) أمكن توسيعها لشمول عدد غير محدد من الكلمات العربية . اللغة العربية فيها (٤٣) مقطعاً لفظياً . منظومة المقاطع ممكن تعريفها من تكوين (٦) مقاطع اساسية للصوت ولكن من المقاطع التالية (i) (u) (a) وللمقاطع البديلة الطويلة المقاطع (ii) (uu) (aa) وهذا التحليل برأينا لا يتطابق مع المقاطع الصوتية العربية . ويمكن البيان بأن نوعية الصوت اللين تعتمد كثيراً على محيط الانغام . إن خصائص اللغة العربية تحددها أمور اساسية من الانغام وعددها ٢٨ نغماً . بالمقارنة مع اللغة الفرنسية والانكليزية فان اللغة العربية منظومة لفظية يحددها ما يأتي :

- ١ - نعم حنجري ؟
 - ٢ - نعم بلعومي (ع) صوتي و (ك) غير صوتي .
 - ٣ - نعم لهات (q) ، (x) وغير صوتي و (ص) صوتي .
 - ٤ - نعم بلعومي (t) و (s) غير صوتي d و (v) صوتي .
- يبدو ان للانغام تأثير آلي على نوعية الصوت اللين . لكل الانغام العربية وحتى التوقف يمكن ان تضاعف . هذا الحديث سيتم الاشارة اليه في مجال توليد قاموس الدايفونات .

٢ - تحويل الاشكال الى مقاطع صوتية :

٢ - ١ تصميم برنامج لتحديد المقاطع العربية الصوتية :

من غير الممكن تحديد المقاطع الصوتية لقطعة مكتوبة باستخدام التركيب الكلامي في القواميس وذلك لكبر حجم الذاكرة المطلوبة . التقطيع يكون بعلاقة حرف الى صوت مع التقيد بقوانين التهجي او النطق . مقطع صوتي او عدة مقاطع صوتية تحدد بحرف واحد او مجموعة حروف وحسب طبيعة المطبوع . الارقام تترجم حسب التسلسل الكلامي لها . في اللغة العربية توجد احكام واضحة لطريقة النطق مقارنة للشواذ والكلمات الاجنبية الموجودة في اللغات الفرنسية والانكليزية .

— تحويل النص الى نطق .. —

الصعوبات هي كما يأتي : —

- أ — عدم وجود العلاقة ما بين الشكل والمقطع الصوقي .
 - اصوات مختلفة تمثل الحرف نفسه مثلاً الحرف (و) يأخذ صوتين مختلفين في الكلمتين (توب) و (توت) .
 - بعض الاحرف لا تعطي اي صوت أي أنها لا تلفظ مثل حرف (أ) في الكلمة ذهبوا .
 - مقاطع اشكال مختلفة يعطي الصوت نفسه مثل حرف (ة) وحرف (ت) في الكلمتين فتاة وبنات .
 - القوانين الخاصة في الكلمات غير القياسية مثل كلمة (لكنّ) .
 - قانون النطق للهمزة بتغيير طريقة لفظ الحرف (أ) إن كان في اول الكلمة يلفظ كهزمة او في داخل الكلمة فيلفظ (أ) .
- ب — المشاكل المتعلقة بدمج الكلمات .
- ج — طريقة تشكيل الارقام .

٢ - ٢ تحقيق وتطبيق برنامج للتحويل

بعض الحقائق تثبت كما يأتي :

اية كلمة تفصل عن كلمة اخرى بقاطع فراغ وكل المقاطع المتعلقة بها تذكر ضمن هذه الحدود .
تحصيل النص المطبوع يتحقق من الطبع على آلة طرفية نوع عربي لاتيني .
تستخدم مؤشرات خاصة لتحديد نوع مجموعة الاشكال الممثلة بأحكام خاصة ما بين الحرف والنطق (الارقام ، اشكال النطق....) .
تحليل النص يتحقق من خلال فتحة متغيرة الطول للتحليل والتي تتحرك على طول نص الاحرف .
التطابق ما بين النص المكتوب والنص الصوقي يتم عند تطابق طول فتحة التحليل والتي تؤدي الى تحديد احكام الحرف الى الصوت .
احكام الترجمة اعلاه تسمح بالعمل على اي شكل من النص .
الاحكام المكتوبة ببرامج (فورتران IV) على حاسبة PDP11/34 تسمح بمعاملة النص بسرعة فائقة .
استخدام تحليل آخر على حاسبة مايكروية نوع (MC 6800) ١٦ بت .

٢ - ٣ الاداء

كالعادة ، فإن برامج التحويل من التهجى الى الاشكال يجب ان تحقق (٤) مهام وهي :

- أ — أقل ما يمكن من الاخطاء .
- ب — امكانية عالية في تصحيح الخطأ .

... تحويل النص الى نطق ..

- ج - امكانية تحقيق ذلك على حاسبة مايكروية .
د - اظهار المعلومات وخواص العروض .
نعد أن النقطتين أ. ب اعلاه يمكن تحقيقها بحاسبة نوع PDP11/34 . الاعمال
الابتدائية التي تمت على حاسبة نوع (١٦) بت ، بينت امكانية تحقيق ما جاء في (ج)
اعلاه خواص العروض كانت محددة الى الاحرف الواضحة اللفظ .

٣ - توليد وتحقيق قاموس لفظي :

٣ - ١ طريقة التركيب اللفظي (دايوفون)

يمكن تعريف الدايوفون بأنه المجال من وسط مقطع لفظي الى وسط المقطع اللفظي الذي يليه .

وهي طريقة مناسبة للتوليد حيث ان تأثير الاندماج الكلامي لواحد من المقاطع اللفظية لا يتجاوز عادة منتصف الطريق للمقطع اللفظي الذي يليه . وبما ان التفرع الدايوفوني يحدث في حالة وضع الاستقرار المفصلي ، لذا فإن القليل من التنعيم كافٍ ما بين الدايوفونات المختلفة . وفي الوقت نفسه فان تحقيق المهمة الصعبة بأستنساخ المقاطع الصوتية باستخدام قواعد المقاطع الصوتية قد تجاوزناها ونقدر بأن حوالي (١٠٠٠ دايوفون) مطلوب لتحقيق خواص جيدة .

٣ - ٢ دايوفونات في محيط صوتي متعادل

إن صيغة الدايوفون تخزن في قاموس وتكون جاهزة للاستخدام في كل الحالات . لغرض تحقيق ذلك فإن اختيار المحيط الذي يستخلص منه الدايوفون عملية مهمة جداً . يتم استخلاص الدايوفونات من كلمات ذات ثلاثة مقاطع لا تعني شيئاً في جمل مشكلة لهذا الغرض . وإن المحيط الصوتي ومستلزمات التسجيل موضحة في (٦) .

٣ - ٣ توليد قاموس الدايوفونات

بعد التهذيب من النطق المحمل ، فإن الكلمات عديدة الفائدة تحلل بواسطة تكنولوجيا (LPC) . تحديد مواقع الدايوفونات تعتمد على ما يأتي : -
أ - عرض منحني الاشارة . فتستعمل هذه الطريقة لاختيار الجزء المناسب من منحني الاشارة للعرض وفي امكانية الانصات على الجزء المختار .

- تحويل النص الى نطق ..

- ب - الرسم البياني لمنحني الترددات الاصلية المضمنة وشدة المنحني المغلق .
ج - منحني الاستقرار الطيفي ويمكن تعريفه كالآتي :

$$S(n) = \frac{\sum_{k=0}^p |Y_k^{n+1} - Y_k^n - 1|}{C + \sum_{k=0}^p |Y_k^{n+1} + Y_k^n - 1|}$$

حيث ان (n) تشير الى هيكل الارقام . (p) مستوى الموديل و (c) عدد ثابت يسمح بتصميم المتغير S(n) حيث ان الطاقة قليلة جداً .

$$y(K) = \log(G) - \log(|x(K)|) \text{ و } K = 0,1 \dots p$$

و (G) هو الربحية في (LPC) و x(K) هو التحويل الفوريير (Fourier Tr) للـ (LPC) .
S(n) يحدد استقرارية الطيف للمقطع (n) نسبة الى دالة التحويل للمقطع (n+1) (n-1) وهي في اعلى نسبة في حالة تقاطع اللفظ وفي اوطأ نسبة في حالة الاستقرار .

د - التمثيل بالتخطيط الصوتي العددي

اول مقطع للدايوفونات قد تحقق . تم خزن الدايوفونات من اثني عشر معامل انعكاس . طبقة الصوت ، والربحية في (LPC) اخذت من طريقة التجنيل المشروحة في (٨) . الايجابيات التي يحققها ترميز الدايوفن بهذه الطريقة يمكن ان تغير بعض المتغيرات مثل طبقة الصوت ، الفترة الزمنية للمقطع والشدة . سلسلة من الاختبارات التركيبية في عدة محيطات صوتية مختلفة تساعد لتحقيق المقاطع الاولية . امكانية التغير والطبعة المخزونة في القاموس يمكن اعتمادها بشكل واسع في عدة محيطات .

٣ - ٤ أسلوب التركيب

الطريقة المعتمدة لتحقيق كلام اصطناعي باستخدام القاموس الدايوفوني المشروح لاحقاً تتحقق من تحويل مدخل من سلسلة من الاحرف الهجائية الى سلسلة من عناوين الدايوفونات وبعدها ارسال المقياس المعياري للشوايت لكل دايوفون .

٤ - ثبوت الدايوفويات كوحدة دنيا

منظومة التركيب المشروحة تستخدم مجموعة من (١٢٠٠) دايوفون . لكل دايوفون توجد نسخة فريدة في القاموس . الكلام المنتج بهذه المنظومة جيد ولكن يوجد بعض التحديدات للطريقة نفسها . المشكلة الاساسية يمكن تحديدها بما يأتي : -

٤ - ١ الانغام البلعومية : -

تحتوي اللغة العربية على بعض الاصوات الحاملة للنغم البلعومي .
t.ا و s.ا و d.ا و j.ا . هذه الانغام تؤثر على المقاطع التي تليها . تأثير الاندماج المفصلي
للانغام البلعومية يمكن ان يغطي كلمة بأكملها ، في حين ان طريقة الدايفونات تعتمد على
الحقيقة : ان نتيجة المنحنى للمقاطع الكلامية تعتمد فقط على المقاطع المجاورة لها .

٤ - ٢ المقاطع الكلامية التي تتأثر بقوة بما يجاورها : -

بعض الانغام الصوتية مثل (ء) تتأثر بشكل عال بحروف العلة المجاورة لها . ان مفهومية
الصوت الاصطناعي التي تستخدم مثل هذه الانغام غير مقبولة اذا كان يوجد مقطع
دايفوني واحد لها في القاموس .

٤ - ٣ مضاعفات الانغام الانفجارية : -

لكل نغم عربي يمكن ان يضاعف . مضاعفات الانغام الانفجارية لا يمكن التعامل معها
بطريقة الانغام الصوتية الاخرى . يجب اجراء معاملات خاصة لهذه الانغام تعتمد على
احترام تناغم قاموس الدايفونات .

٤ - ٤ الاصوات اللينة القصيرة مقارنة بالاصوات اللينة الطويلة

ان اوقات مقاطع الصوت اللين وثيق الصلة في اللغة العربية . ان توليد المقاطع الطويلة
في المقاطع القصيرة او بالعكس يمكن ان يصغر حجم القاموس . هذا يحقق من خلال دراسة
نوعية الصوت الاصطناعي .

٥ - علم العروض

يتم تحقيق العروض من خلال استخدام الكلمة بكاملها . الكلمة تقطع الى عدة مقاطع
وتستخدم النبرة الاولى في المقطع المناسب . هنالك دراسة في طريقها وهي ((تشبيه العروض
لغة العربية)) .

٦ - التحقيق بالاجهزة

تم صنع لوح الكتروني لتركيب الصوت من خلال استخدام الحاسبة الصغيرة (M6809)
ووحدة مركب الصوت (TMS 5220) . تم إيصال اللوح بالحاسبة نوع PDP11/34 وهذا
أمكن تحقيق التركيب اللفظي في الوقت الحقيقي . منظومة مستقلة تعتمد على الهيكلية
نفسها مع قاموس دايفوني مخزون في ROM قيد التنفيذ .

- (1) R. EL MELANI
"Contribution a la reconnaissance de la parole arabe:
reconnaissance globale de mots isolés"
These do 3 Cycle, Marsh 1985, Rabat, Morocco.
- (2) A. RAJOUANI
"Contribution a la synthese de la parole arabe par ordinateur"
These de 3 Cycle Feb. 1983, Rabat, Morocco.
- (3) D. CHLADMI
"Contribution a la synthese par fomants de l'Arabe: Etude de la
pharyngalisation"
To be published.
- (4) G.E PETERSON, W. WANG, E. SIVERTSEN "Segmentation
techcniques in speech synthesis"
JASA, vol-30, n 8, August 1985.
- (5) R. SCHWARTZ and all "Diphone synthesis for phonetic vocoding"
ICASSP 79, Washington, pp. 891-894.
- (6) G. CHOLLET And all "On the generation and use of a segment
dictionary for speech coding, synthesis and recognition"
ICASSP 83, Boston, pp. 1328-1331.
- (7) M.STELLA
"Fabrication semi-automatique diction-naire de diphones"
Recherche/ Acoustiquo CNET, vol-VII, 1982/83.
pp-53-63
- (8) A.MOURADI
"Synthese de la parole arabe a partir du texte par la methode des
diphones"
Ph.D Thesis, July 1985, Rabat, Morocco.

(1) R. E. MILLER, "Contribution à la reconnaissance de la parole: l'analyse globale de mots isolés", *Thèse de l'École Supérieure d'Études Supérieures de Montréal*, 1985.

(2) A. R. LIOU, "Contribution à la synthèse de la parole: analyse par séquences", *Thèse de l'École Supérieure d'Études Supérieures de Montréal*, 1985.

(3) D. CHEN, "Contribution à la synthèse par échantillons de l'anglais: l'analyse par séquences", *Thèse de l'École Supérieure d'Études Supérieures de Montréal*, 1985.

(4) G. L. PETERSON, "SPEECH SYNTHESIS: A REVIEW", *IEEE Transactions on Audio and Electroacoustics*, vol. AU-18, August 1970.

(5) R. SCHWARZ, "A digital speech synthesis system for French", *IEEE Transactions on Audio and Electroacoustics*, pp. 881-894, 1976.

(6) G. CHOLLET, "On the generation and use of a segmental dictionary for speech coding, synthesis and recognition", *IEEE Transactions on Audio and Electroacoustics*, pp. 1128-1131, 1973.

(7) M. STETTIN, "Fabrication semi-automatique d'échantillons de diploons", *Recherches Acoustiques CNRS*, vol. VII, 1981, pp. 84-87.

(8) J. MOUTON, "Synthèse de la parole: analyse par séquences de diploons", *Recherches Acoustiques CNRS*, vol. VII, 1981, pp. 84-87.