دار المنظومة
DAR ALMANDUMAH
الرواد في قواعد المعلومات العربية

العنوان: دراسات لغوية : تعلم الكلام

المصدر: البحث العلمي

الناشر: جامعة محمد الخامس - المعهد الجامعي للبحث العلمي

المؤلف الرئيسي: بودرى، مارك

مؤلفين آخرين: الهاشمى، التهامى الراجد(مترجم)

المجلد/العدد: ع 36

محكمة: نعم

التاريخ الميلادي: 1986

الصفحات: 157 - 170

رقم MD: 390185

نوع المحتوى: بحوث ومقالات

قواعد المعلومات: HumanIndex

مواضيع: الالفاظ اللغوية ، الدراسات اللغوية ، الكلام ، الاصوات اللغوية ، اللسانيات ، الاتصال الشفوى

رابط: http://search.mandumah.com/Record/390185

© 2021 دار المنظومة. جميع الحقوق محفوظة.
هذه المادة متاحة بناء على الإتفاق الموقع مع أصحاب حقوق النشر، علما أن جميع حقوق النشر محفوظة. يمكنك تحميل أو طباعة هذه المادة للاستخدام الشخصي فقط، ويمنع النسخ أو التحويل أو النشر عبر أي وسيلة (مثل مواقع الانترنت أو البريد الالكتروني) دون تصريح خطي من أصحاب حقوق النشر أو دار المنظومة.

دراسات لغوية

تعلم الكلام

مارك بودري
تعريب د. النهامي الرابي الهاشمي

إن اللغة الاكثر ملاءمة للتحاور مع نظّامة إنسان آلي هي هذه التي نستعملها إلاّ أنّ أبسَط لفظ فيها يتكون من عدة عناصر صائتة، تَرَدُّدُها، جرْسها، عُلوُّها وإيقَاعُها كل هذا يختلف حسب الالة المصوتة لكل شخص وحسب نبره ايضا. ومع ذلك فان التعرف الاكوستيكي والصّوتي للألفاظ قد اكتسب جزئيا، وسيصبح من الممكن معالجة المستويات الرفيعة للتركيب والدلالة والتفكير.

سترتفع تكلفة التعرف على الكلام وتركيبه الآلي، حسب تنبؤات الـ **BIPE**، الى 784 مليون دولار في سنة 1985 مارّة سنة 1990 الى اربعة ملايير دولار ونصف مليار دولار. سيصل هذا السوق، في فرنسا، حسب نفس المصادر، وفي نفس الحقب، بالتّتابع الى 165 و 956 مليون فرنكاً.

فمن التحكم الصوتي في مساحة زجاج الى التحكم في الانسان الآلي وفي عتاد مطارف النظامات ذات الانتشار الكبير، مرورا بالمنفّذ الى بنوك المعطيات، تكون التطبيقات الكُمُونية لأنظَمَةِ تعرّفِ وتركيبِ الكلام هائلةً جداً.

إن الدور الاساسي الذي تحتله هذه التطبيقات في قضية تبسيط التفاعل بين الانسان والآلة ليوضّح لنا، بجلاء فائدتها الاقتصادية، انها لتكوّن، ولا شك، إحدى الشروط الاساسية الكبرى لإزالة عقبات كأداء تعوق غير المتخصصين في النّظامات وفي تقنيات اخرى موازية لها من النفاذ اليها.

العنوان الاصلي للمقال.

L'apprentissage de la parole Science et vie - Numero hors serie - Mars 1982

(1) مكتب الاعلام والتنبؤات الاقتصادية

Bureau d'Information et de Prévisions Economiques

ان تفهمنا النظامةُ عندما نكلّمها وأن يكون في مقدورها ان تجيبنا صوتيا، هما، من الناحية التقنية، مشكلتان اثنتان، صعوبة كلٍّ منهما من نفس الصنف. رغم انهما مختلفتان فلن يعالجا كلاًّ على حدة. يمرُّ فعلاً، حلول هاتين العمليتين من خلال تفهم دقيق لكل من آليات الحديث[2] وآليات الفكر عند الانسان. ان التقدم مرهون في جزء مهم بآليات الذكاء الاصطناعي.

المستويات المختلفة للاتصال الشفوي

إن الاتصال الشفوي يشغل، ابتداءً من مستويات التجريد الاكثر سُمواً في التفكير الى مستويات العتاد الطبيعي لاستقبال وارسال الكلام، مدارج كثيرة.

نجد مختلف المستويات المتسلسلة : واقعي،دلالي، تركيبي، مُعجمي، صوتي وأكُوستِيكي

ان المستويات الثلاثة الاولى هي الاكثر سُمُواً، ذلك ان المستوى الواقعي له مساس بالعالم المادي المحيط بنا في حين ان المستوى الدلالي له علاقة بمعنى الالفاظ، اما المستوى التركيبي فله اتصال بقواعد البناء النحوي. بيّنتْ الاعمال المتعلقة بالحديث الطبيعي كل الصعوبة التي تعترض السبيل حين تباشر هذه المستويات على الآلة. كما توجد طبعاً هذه الصعوبة حين التعرف على الكلام وتركيبه.

وفعلاً فان هذه المستويات ليست منفصلة بعضها عن بعض تماماً. وهكذا فانه لا يمكن للجملتين : «تبيعُ سيارتها لأن ليس لها حصّار» و «تبيع سيارتها لأن ليس لها نقود» أن تحلّلا، بشكل سليم، على المستوى التركيبي إلا بمساعدة المستويين الدّلالي والواقعي.اما بالنسبة للغة المنطوقة فإن هناك قواعد اخرى تظهر، زيادة على تلك التي تتحكم في اللغة المكتوبة : «لا انا مش آتٍ»، «انا لست آتيا».

وبالرغم من الصعوبات المتعلقة بالمستويات العليا فإنَّ عدداً من انْظمة التّعرُّف تحاول استخدامها. وفعلاً فانه من الممكن الآن خلق «احاديث» Langages ، مصطنعة، بنيت على معجميات Vocabulaires مصغّرة، بفضل تركيبات مصغّرة هي ايضا، ومحددة تحديدا بليغا، وبقواعد للنطق دقيقة للغاية.

(2) اترجم بالمصطلح «الحديث» اللفظ الفرنسيlangage» برّرت هذا في كتابي : «بعض مظاهر التطور اللغوي» ابتداء من صفحة 72، وكذا في كتابي « الثنائيات اللسانية صفحة 127.

إنَّ المعرفة الدلالية والواقعية التي كثيرا ما تكون مُدمجة ضِمنياً في المَعجمية بفتح (الميم)، وتركيبات أحاديثها لا تدخل الا على تطبيقات محدودة ودقيقة.

واحدة من هذه الانظمة التي تعرف بانظمة لفهم الكلام الطبيعي هو «Laser 1000» الذي وضع من طرفIBM. انه يحاول ان يتعرف على جمل بُنِيَّتْ انطلاقاً من 1000 لفظة متعلقة بالبراءات الخاصة بـ «Laser».

ظهر مشروع للتعرف على الجمل المتعددة المخاطب لمعجميات من 1000 لفظة في سنة 1971 بالولايات المتحدة، وُظِّف من اجله 15 مليون دولار على مدة خمس سنوات. ورغم ان الاهداف المحددة لم تتحقق، فانه مَكَّن البحث من التقدم بشكل ملحوظ.

واليوم فان احسن الانظمة المنبثقة من هذا المشروع هي :

- و Hearey II de Carnegie Mellon Haspy

(University HWIN de Bolt Beranek-Newman)

- و كذا نظام شركة System Development Corporation

وفي فرنسا فان الانظمة

«Myrtille»في CRIN(3)

و «Keal» في CNET(4)

و «Esope» في LIMSI.

هي الانظمة الاكثر تطوراً. هذا وإن بحوثاً مماثلة تجري في معظم البلدان الاخرى المصنعة.

ما زالت كل هذه الانظمة في طور المختبر نظرا، لأهمية العتاد التقني الذي لا بُدَّ من التوفُّر عليه لتشغيلها. يَسْهل، بالنسبة للمستويات الثلاثة الباقية، وهي المستوى المعجمي والصوتي والاكوستيكي، التحكم فيها آليا.

إن الانظمة المحتملة التي يمكن ان تؤدى، آجلاً او عاجلاً، الى تطبيقات عملية، صناعية بالنسبة للسواد الاعظم من الناس، ستعتمد،كلياً تقريبا، على هذه المستويات السُّفلى.

(3) GRIN موجودة بجامعة Nancy I (توجد بالشمال الشرقي في فرنسا : المترجم).

(4) (Centre National dÉtudes et des Télécommunications à Lannion) CNET

المركز الوطني للدراسات الخاصة بالاتصال البعدي.

Lannion هي عاصمة شواطىء الشمال بفرنسا يقدر سكانها بحوالي 13797 نسمة (المترجم)

الاوّل من بينها هو المستوي المعجمي. يحتوي لروس Larousse الصغير على 20.000 لفظة، تضاف اليها الصيغ المشتقة منها والمتصرفة عنها وفعلاً حين يتعلق التطبيق بمحيط محددٍ جداً فان بعض المئات او العشرات من الالفاظ تكفي.

المستوى الذي ياتي اسفل من هذا مباشرة هو المستوى الصوتي. إن الالفاظ مُرمَّزَة(6) تحت شكل معياري بواسطة متتالية من الصوتات(7) يجوز ان تستعمل الصوتة(8) كوحدة للتعرف، ومع ذلك، يمكن لهذا الشكل المعياري، وهو لغة منطوقة، أن يُلحِقَهُ اضطراب تشكيلي(9) نجد ذلك في العربية(10) في التسهيل والبين بين في الهمزة مثل : «أإذا دخلنا» ينطق بها : «أهذا دخلنا» ، «يؤْمِن بالآخرة»ينطق بها «يومِنْ بالآخرةِ» ، و «التَّرخيم في النداء، مثل : «يا فاطمة» يُنطَقُ بها : «يا فاطم».

إن الالفاظ، زيادةً على ذلك، غير منفصلة في البلاغ المنطوق، ولكنها ملتصقة بلا انقطاع، ومن هنا فان التعرُّف عليها في مثل هذه السلسلة من الصعوبة بمكان. وبقطع النظر عن المعنى فان الجملة : «اسمعْ عليل الكلام» يمكن أن تؤوَّل : «اسمعوا علي الكلام»، «اسمعْ لي الكلام»(11).

المستوى الاخير هو المستوى الاكوستيكي. يقع نقل البلاغ الصوتي بواسطة ارتجاجات صائتة ناتجة عن تغيرات ضغط الهواء المحيط.

(5) مختبر تابع للمركز الوطني للبحث العلمي بـ Orsay في فرنسا.

(6) Code = رمز. أما «رمّز» بالتضعيف فاترجم بها اللفظ الفرنسي Coder,واما de'coder فاترجمها بـ «أرمز» بالهمزة التي تدل، في كثير من الافعال العربية على السلب، مثل عجم وأعجم، شكا وأشكا، الى آخره (المترجم).

(7) «الصُوتة» بضم الصاد هي القدر الادنى من الوحدات الصوتية، واما الوحدة منها فهي (صوتة) بفتح الصاد. (المترجم).

(8) Phoneme. يوجد في اللغة العربية حوالي 29 صوتة دون ادخال اصوات الامالة والاشمام والرّوم والتفخيم والترقيق والاختلاس (المترجم).

(9) Phonologique

(10) في المقال الاصلي المكتوب بالفرنسية امثلة بهذه اللغة، فضّلتُ ان اضرب الامثلة باللغة العربية لتعُمَّ الفائدة.

(11) في النص الاصلي امثلة بالفرنسية لا تفيد ترجمتها الامنية الى اللغة العربية في توضيح ما اراد تبليغه كاتب المقال للناس (المترجم).

لا يتكون البلاغ الصوتي، في المستوى الاكوستيكي من تجميع العناصر المعيارية مستقلة بعضها عن بعض، بل تتغير كل صوتة بتأثير من الصوتات السابقة واللاحقة. كما انها تتغير أيضا بمعلومات غير لسانية لها علاقة بالطبيعة العروضية[12]. إن التغيرات في علو الصوت لتهمُّ العروضية مثل ارتفاع الصوت في آخر الجملة الاستفهامية، كما يهمها الايقاع في صياغة العبارة وكذا سعة الصوتات.

تُمكِّننا العروضية من ادخال كثير من الظلال الرقيقة في البلاغ الصوتي يمكن ان يؤدي الى تغيير كامل في معنى الجملة. من اجل هذا وغيره كثير (كالضجيج المحيط، الصدى في الحجرة...) فان الصورة الاكوستيكية لنفس الصوتة (وبالاحرى لنفس اللفظ) لا يمكن ان تكون متشابهة ابداً.

ويجْدر أن نسجِّل ان مستمعاً انسانياً يستطيع، إن كانت الاخطاء أقل من 10٪ والنطق سليماً، والاستماع رائعاً، أن يُميز الصوتات في ألفاظ لا دلالة لها. لايدخل هذا المدرج في الحسبان المستويات العليا. وفي مقابل ذلك، فان البلاغ الصائت الذي يحتوى اكثر من 10٪ من الصوتات الفاسدة، يصير، في الحال، غير مفهوم. يَستعْمِلُ المستمع الانساني، لكي يُلَطِّفَ التَّغايريَّة المهمة في الاشارة الصائتة كل المستويات-معاً. إنه من جهة اخرى، وزيادة على ما ذكر، يستخدم اثناء الحوار الحيّ وسائل اخرى للادراك ليعيَّ أكثر البلاغ، إنه يستعمل، على الخصوص، النشاط الحركيّ للمخاطب حركات الشفاه والرأس واليدين الى آخره...

الالفاظ المنعزلة
صنف كامل من التطبيقات

نستطيع ان نستنتج مما سبق ان التعرف على الكلام وتحليله الآلي يجب ان يعتمد على معرفة دقيقة للمستويين الصوتيّ والاكوستيكي لكن هذه المعلومات ليست، في الوقت الحاضر، كافية لكي تضن لنا فهما دقيقا لحديث اصطناعي بسيط. وزيادة على ذلك، فان التقنيات المستعملة في المستويات الصوتية والمعجمية والتركيبية تتطلب عتادا ثقيلا وغالياً.

ولابد للوصول الى انظمة يمكن الاتجار فيها، اليومَ، من دفع قضية التبسيط الى أقصى حد ممكن.

(12) ما يقوله هنا كاتب المقال غريب، انه يعتقد ان العروضية ليست معلومة لسانية في حين انها من صميم الدرس اللساني قديماً وحديثاً (المترجم).

فإن اردنا تفادي الصعوبات المرتبطة بالمستويات العليا اقتصرنا في البدء على ما تحت المشكل، وهو التعرف على الكلمات المنعزلة.

إذا كانت ارادة التعبير عن جملة تامّة بواسطة متتالية من الفاظ، منعزلة بوقفة، غير ذي واقعية، فان هناك صنفا كاملا من التطبيقات حيث يمكن للالفاظ وحدها، او لجمل قصيرة جداً، ان تكون كافية لاعطاء أوامر أو معلومات.

ويُمكن ان نوفر صعوبة التعرف على الصّوتة. ذلك ان اللفظ يحدد حينئذ، إجمالياً، بواسطة مميزات اكوستيكية اكثر خشونة مستنبطة من الصّورة الاكوستيكية لِلإشارة اللّفظية.

فان نحن حققنا هذين التبسيطين ، بشكل لحظي، وصلنا الى المشكل الاكثر بساطة، الآ وهو التعرف الاجمالي او الاكوستيكي للالفاظ المنعزلة.

إن التطبيقات النموذجية هي تلك التي لا يستطيع المُسْتَعْمِلُ ان يتوصل فيها بسهولة مع النظامة، مستخدما الوسائل التقليدية من ملمس وشاشة لان يديه مشغولتان باعمال اخرى.

واننا لنجد هذه التطبيقات في ميادين كثيرة من الالعاب الالكترونية الى مساعدة قيادة الطائرة والتحكم في الانسان الآلي والقصور الذي تعين على انجازه النظامة، وتقديم المساعدة اثناء العمل في وسط صعب تحت الماء أو في الاجواء العليا. لنذكر بالنسبة لفرنسا نظام MOISE الذي ارسى قواعده الـ LIMSI (تكلفته 10.000 فرنكاً فرنسياً تقريباً) و نظام TOPAR الذي انجزته CEA (Saclay) ونظام مساعدة فرْز الرسائل الذي يدرس الآن في CRIN. اما بالنسبة لامريكا فإنه يوجد من بين الانظمة التي يمكن أن يقع الاتجار فيها Le Speech Lab. de Heuristics (تصل تكلفته الى 300 دولار) و (2000 دولار) Quick Talk de Thereshold (17.000 دولار).

تسمح هذه الانظمة المختلفة بالتعرف على 16 لفظاً الى 100 لفظ مع معدل للخطإ يتراوح بين 5 وأقل من 1٪.

تظهر هذه الانظمة على شكل آلة مستقلة او خارطة محيطيّة للنظامة. إن كانت فوائدها تتمثل في انجازاتها وفي بساطتها فإن عيوبها، رغم ذلك مهمة للغاية، وذلك لاسباب كثيرة، منها : استعمال الفاظ معزولة معجم (بفتح الميم) ضعيف، ضرورة القيام بمرحلة التعليم (تَحتّم في بعض الاحيان تكرار اللفظ 10 مرات) وضرورة اعادة التدريب مع كل مخاطب.

فاذا ما اعتبر النظام كثيراً من المخاطبين دون القيام بتَعْليم كُلٍّ واحد منهم في كل مرة، سُمِّيَ هذا النظام : «متعدد المخاطب». ان تكلفة مثل هذا النظام في هذه الحالة مرتفع جدا : 80.000 دولار بالنسبة لـ : de dialog 1800 (موجود بالولايات المتحدة الامريكية، يسجل2٪ من الاخطاء بالنسبة لـ 85٪ من المستعملين) و 200.000 دولار بالنسبة لـ SRR 1101 de NEC (باليابان). وفِعلاً فإن هذه الانظمة تسمح بمعالجة عدة مخاطبين، بصفة لحظية وبغاية الربط مع الشبكة الهاتفية.

كيف يعمل هذا

فنأخذ، لكي نفهم كيف يعمل نظام للتعرف الشامل على الفاظ منعزلة المثال-TO PAR. هذه الآلة الحرة موجهة خصيصا للمعاقين حركياً.

إنها تسمح لمستعملها ان يتحكم بواسطة الصوت في جزء من محيطه (أمر بفتح واغلاق نور المصباح او النافذة او التلفاز الى آخره).

إن وحدة المعالجة، في مثل هذا النظام هي اللفظ (او جملة قصيرة جدا) متبوعة بسكوت لا يقل عن 0,2 ثانية وحين تكون بداية اللفظ أثبتت فان الثَّوابت تَحْسبُ فرج منظمة معطية صورة أكوستيكية مبسطة للفظ.

تمثِّل كل كلمة في مرحلة التَّعلم في الذاكرة بواسطة هذه الصورة الاكوستيكية وباسمها لكيْ يتكون معجم مرجعي.

أما في مرحلة التَّعرُّف، فإن اللفظ المجهول المُرْمَز حسب نفس المنهج يقارن بألفاظ المعجم. يختار اللفظ الاكثر شبها، لكن قد يقع الرَّفض اذا كان الشبه ليس قويا بما فيه الكفاية إن خَرْجَة الالة المناسبة للفظ المنطوق تنشط آنذاك مما يسمح بتحقيق الامر المطلوب.

تتطلب المقارنة بين لفظين بعض الاحتياطات، ذلك ان مدَّة لفظ لا تكون هي هي أبداً من تكرار الى آخر. وإننا نستطيع ان نعتبر، في تقريب أولي، وبالنسبة لناطق متعاون، انَّ هذا التغيير الموقت وزّع بشكل منتظم، على طول اللفظ. نستعمِل، لمعالجة الحالة التي تكون فيها التغييرات غير منتظمة، تقنيات للمقارنة ديناميكية لا ندخل في تفاصيلها.

رأينا أنَّنا اذا كنا نرجو التعرف على الكلام الطبيعي وبالتالي تفادي عيوب التعرف الاجمالي فلا بدَّ من المرور الى المستوى الصوتي. **يحتوي إذن المعجم على مراجع من المستوى الصوتي. ان مدرج المقارنة**

معقّد للغاية ولهذا لابد من محاولة الاخذ باعتبار جميع الاضطرابات المحتملة. يمكن لهذه الاضطرابات ان تجعل التحديد الدقيق للصوتيات امراً وعْراً.

تكمن صعوبة المشكل في ان معدلات الخطأ في الصوتات لا تنزل إلا نادراً تحت 25٪ في سياق «وحيد الناطق» وتصعد بسرعة في سياق «عديد الناطق» (Multiloculeur) يُمرُّ تحسين النتائج على كل من اتقان التَّعَرُّف على المستوى الاكوستيكي وعلى اتقان معالجة الاضطرابات بواسطة المستويات العليا : المعجم والتركيبي والدّلالي. نستطيع ان نجتاز، للتغلب على بعض الصعوبات، الى مرحلة اولية، عوض البحث عن التحديد المباشر للصوتات.

يمكن أن تَعْتَبر الصوتة وكانها مجمع لمكونات أولية ذات مدة أقصر. إما ان تُحَلِّل الاشارة اللفظية الى مميزات اكوستيكية خاصة بمختلف الصوتات وإما ان توزع الى «نوافذ» تحليل ذات مدة ثابتة (0,05 الى 0,010 ثانية).

اما في منهجية معاكسة فإنّنا نرمي الي ابراز مجموعات من الصوتات : وحدات مقطعية أو ثنائية الصوتات وهي وحدات تبدأ من وسط الصوتة وتنتهي في وسط الصوتة الموالية. إن اللفظ «مدرسة» مثلا : مكون من م، مْ، مدْ، ر، سة ـ . يشير الخط الى السكتة، اما المراجع فتقدر بالمآت.

الدّراسة الاكوستيكية

سواء تعلق الامر بالتعرف الاجمالي او الصوتي او تعلق بتحليل الكلام، فلا بد من المرور بدراسة اكوستيكية للاشارة اللفظية التي تسمح باستخراج المميزات المثلى.

يمكن ان نميز انواعا ثلاثة كبرى من الصوتات. هناك، بالنسبة لجل الصوتات، اهتزاز للأوتار اللافظة (عدا إذا تعلق الامر طبعاً بوشوشة). هذه الاصوات المبحوحة هي التي تُقَدَّمُ بنية ذات مدة رتيبة تقريبا. تُلائم هذه المدة اهتزاز الاوتار اللافظة. إن التواتر الاساسي الملائم يمتد عادة من 90Hz الى 600Hz ابتداء من اصوات الرجال الى اصوات الاطفال.

ويمكن ان نعتبر كل مدة وكأنها ناتجة عن تراكم العديد من الاشارات الاولية[13]. إن الاشارتين الاقل ترددا (تتراوحان بين 200 و 300Hz) كافيتان

(13) تلائم هذه الاشارات الاولية مفهوم البانية (Fermant)

ترجمت، في كتابي : «بعض مظاهر التطور اللغوي» Formant بـ «بانية» المترجم.

لتحقيق مفهومية الكلام. إلا ان الاذن تستطيع ان تلتقط حوالي ست اشارات غيرها يكون دورها مهما في كل اللمسات التي يمكن ان تضيفها للبلاغ الملفوظ.

إن النوع الثاني هو نوع الصوامت مثل ط، ك السماة شديدة تتميز هذه الصوامت بالغلق التام لفترة وجيزة جداً لنقطة معينة في الممر المصوت(14). ثم انها تتميز ثانيا بتطور سريع نحو الصوتة الموالية.

اما النوع الثالث فهو نوع الصوامت س، ش أو ف السماة دعكية غير مبحوحة. ان الاوتار اللافظة لا تهتز وأن إشارة من ضجيج تظهر بدون بنية مدوية (نسبة الى المدة) ناتجة عن احتكاك الهواء عند اجتيازه بنقطة ضيقة جدا في الممر المصوت.

وطبعاً فان كل هذه المميزات الاكوستيكية مرتبطة بخصائص الجهاز الصوتي الانساني.

لقد سبق ان أكدنا على تتابع البلاغ الصائت. تمتاز الصوتات وعلى الخصوص الثوابت، بتطور إشارتها الاولية او البناة، انطلاقا من الصوتة السابقة الى الصوتة الموالية وزيادة. هذه التطورات التي لا تعرف، حتى الآن جيدا، هي الاصل الاساسي للصعوبات التي تواجه تحديد الصوتات(15).

تتابع جميع الفرق الفرنسية البحوث الاساسية على المستويين الصَّوتي والاكوستيكي. هناك، من بين الفرق التي لم تذكر بعد، الـ CERFIA(16) والـ ENSERG والـ IMAG(17) ومعهد البرمجة(18).

إن احد اجنحة معالجة الكلام المرتبطة بالتَّعرف هو ذلك الذي يهتم بالتحديد والتحقُّقِ من الناطق. ومعلوم ان تحديد الناطق هو مشكل تكميليٌّ

(14) اطلق النحاة العرب الاقدمون لا سيما الذين تطرقوا لما يعرف عندنا بـ «المشترك» عليه «الحيز».

(15) لا بد لكي يمكن للنظامة أن تعمل على اشارة صوتية من ان تكون هذه الاشارة مرمزة في الذاكرة يمكن ان يستعين هذا الترميز بتقنيات كثيرة مختلفة. هذا بالذات هو الذي يكون منبعاً لمشاكل عويصة إذ يمكن ان يكون عدد الذاكرات المحتاج اليها متعارضاً مع امكانيات النظامة الحالية. نذكر وبسرعة ان من بين هذه الامكانيات التنبؤ الحظي والنماذج الرياضية للمقود الصائت والممرات بالصفر... وأخيراً البناة وأطباق التردد المعطاة من طرف اسراب مرشحات او احصاء رياضي...

(16) يوجد هذا المركز في جامعة Toulouse Paul Sabatier ,

(17) المدرسة الوطنية العليا للكهرباء ومعهد الرياضيات بـ Grenoble

(18) جامعة Pierre et Marie curie بباريس 6.

لمشكل التعرف وان المطلوب هو العثور على الاختلافات الموجودة بين الناطقين لإزالتها. ما نبحث عنه هو الارتفاع والجرس والايقاع والنبر وان هذا المشكل يُقَدِّم لنا نفس الصعوبات التي يواجهنا بها التعرف على الكلام الطبيعي.

يمكن ان نُبَسِّطَ هذاالمشكل بالاقتصار على التحقق من هوية انسان مستعينينَ بخارطة مغنطسية مثلا وبنطق لفظ او اكثر. تقارن هذه الالفاظ مع نماذج سُجلت اثناء مرحلة التعليم، فان كان الشبه كافيا «قُبِلَ» الناطق.

نَسْتَعْمِل TEXAS INSTRUMENTS في الولايات المتحدة هذا النوع من التفتيش لمراقبة الدخول الى قاعات النظامات. (يلاحظ) ان الاخطاء تكون من صنفين، إما رفضُ مستعملٍ صادق واما قبول خدّاع. وفي كلتا الحالتين فان الخطر يقدر ب 1٪ وطبعاً فإنّ المقلدين الاحترافيين يقبَلون افضل. لكن النظام يبقى اكثر فطانة من المستمع الانساني.

توليف الكلام

آخر ما يمكن ان نتحدث عنه هو تأليف الكلام من بين الانظمة «الناطقة» المعروفة جداً، والتي يمكن الاتجار فيها الآن هي «الاملاء السحري»[19] Dictee magique.

وليس من بينها واحد يعتمد على تأليف الكلام بما في الكلمة من معنى إذ الامر يتعلق بأنظمة تعادل مسجلة إصْبَعية، إنّها، جميعها، تعتمد على اعادة ـ حسب الطلب ـ لبلاغات خزنت في الذاكرة بواسطة ترميز للكلام على المستوى الاكوستيكي. وزيادة على ان المفهومية ليست دائما على اتمها فان البلاغات محجَّرة، قليلة العدد، تحتاج لترميزها الى مدرج ثقيل ومنفر.

يمكن ان تستعمل ايضا : تقنية ترميز وارماز الكلام هذه (دون ان يكون هناك تعرف بالفعل) لنقل بلاغ صامت على خط تحويل. المهم هو استطاعة نقل عدد من البلاغات على خط واحد. ان الترميز الاكثر تركيزا هو الذي يُلائم نقل سلسلة صوتية (ولربما بعض المعلومات «الجمالية»[20].

(19) وهو النص الفرنسي لـ Speak and Spell الذي انجزته Texas Instument والذي يمكن طفلا صغيرا من تعلم الاملاء بضربه على ملامس لفظاً سمعه.

(20) يسمى العتاد الذي يمكننا من هذا الترميز وهذا الارماز بـ VOCODER ،مستخرج من Voice Coder.

يُحَتَّمُ الارماز، في هذه الحالة الاخيرة، تكوين البلاغ الصوتي انطلاقا من سلسلة الصوتات المكونة له ، وهذا، بالضبط هو ما يسمى بـ «التأليف الحق».

على التأليف ان ينطلق اذاً من المستوى الصوتي لتركيب البلاغ الملفوظ، كما ان التعرف على الكلام الطبيعي يتطلب تحليلا صوتياً.

يَسْتَعْمِلُ الترميز على المستوى الاكوستيكي، كيف ما كان نوع التأليف المعتبر احدى التقنيات التي رأيناها في تحليل الكلام.

على المستويات العليا حينئذ ان تمكننا من تكوين تغايرية وتتابعية البلاغ الصائت، من الصوتة الى التي تليها، من لفظ الى لفظ بعده، وعلى الخصوص من خلال المعلومات العروضية. هذه المشاكل لا تحل في جميع عموميتها. إن الاذن لرقيقة الحساسية الى عوامل غير عادية في البلاغ الصائت. إن انظمة التأليف الحاضرة تعطي كلاماً مفهوماً غير ان صفاته الجمالية لا تَكْفي لتقديم تطبيقات خاصة بالجمهور (عدا الحالات المتعلقة بالالعاب).

نجد من بين الانظمة التي نعثر عليها في الولايات المتحدة الـ : Votrax (ترميز أكوستيكي للالفاظ). الذي يباع بالالاف المؤلفة من النسخ الصالحة للتطبيقات الصناعية، كما نجد النظام المعروف بـ Handivoice, للبكم، يقدم هذا النظام 47 صوتة و 17 جملة (تكلفته 2500 دولار). اما في فرنسا فنجد النظام Icolog المنبثق من المحلل Icophone في ملكية LIMSI, وقد دخل ميدان المتاجرة. اما المختبرات الرئيسية الفرنسية التي تهتم بتاليف الكلام فهي CEA (نظام SARA) و CNET و ENSERG الذي هو في ملك معهد البحوث الاكوستيكية والموسيقية في مركز Beaubourg بباريس، و LIMSI.

ولنشر الى تركيب اصوات مغناة ذات خاصيَّة عالية والمنجزة في IR-CAM والذي استعمل في حفلات موسيقية.

وعوض ان ننطلق مباشرة من المستوى الصوتي، يمكن ان نباشر تاليفا انطلاقا من نص مكتوب. يتطلب هذا ترجمة متوالية من الحروف الى متوالية من الصوتات مع ذكر بيانات من المستوى العروضي، إن أمكن. يمكن ان يقع هذا النقل في اكثر الحالات دون قليل تركيبي ولا دلالي. يوجد بعض الاضطراب في مثل هذا المثال «رأى جمال جمال الارض»، «وَلَدَ الرَّجُلُ»، «وَلَدُ الرَّجُلِ»، فلا بد من تكوين للاستثنائيات اذن.

ان احدى التطبيقات المهمة هي التركيب الآلي للمكفوفين الذي يجب ان يُقرَنَ بنظام للقراءة البصرية للحروف.

ان الاتجاه الآن يسير نحو تطوير الدورات المدمجة المتخصصة في نوع من التركيب الخاص. نَجدُ اذن في هذا السوق كبار صانعي المُكوِّنات، وهي : Texas Instruments و ITT و General instruments و National semi conductor و Sharp و Mutsushita و Hitachi

ولنختم لابدَّ من أن نؤكد على حيوية البحث في فرنسا. وتأتي فرنسا، في المعالجة الآلية للكلام بعد الولايات المتحدة واليابان. اغلب الفرق هي من القطاع العمومي الجامعي او من المركز الوطني للبحث العلمي، لكن بعض المختبرات الحرة تساهم في هذا المجهود مثل :

IBM France, Texas Instruments France, (CIT Alcate) و
(Tomson CSF) و

كما تساهم الادارات بنشاط كبير في تمويل البحوث. ولنشر اخيراً الى الانشاء الحديث لمجموعة من البحوث تعرف بـ GRECO, ستكون مهمتها في فرنسا، تنسيق النشاط البحثي على الكلام.