

RECONNAISSANCE ET SYNTHÈSE VOCALES AU SERVICE DE LA PÉDAGOGIE

Philippe GALIANA

Les derniers programmes de reconnaissance et de synthèse vocales ne sont plus à considérer comme un gadget. Du côté de leur efficacité tous les spécialistes s'accordent à le dire ; les programmes informatiques sont à présent suffisamment fiables pour reconnaître la plupart des locuteurs quel que soit leur accent et leur intonation. Les dernières avancées donnent des résultats très satisfaisants avec un taux de reconnaissance théorique de 95 %. Les progrès sont liés à l'utilisation de modèles statistiques, appelés Modèles de Markov Cachés qui en font un système plus intelligent. Plus on se sert du système, plus il s'améliore. D'autre part, ce qui pouvait sembler impossible dans le cadre d'une salle de classe l'est aujourd'hui : des micro-casques unidirectionnels avec filtrage sont fournis d'office avec les applications. Ils éliminent totalement les bruits ambiants qui nuisent à la compréhension de l'intervenant par le système. De fait, ces dernières avancées techniques font maintenant de la reconnaissance et de la synthèse vocales un outil à forte valeur pédagogique.

Les équipements

Côté budget, là non plus il n'y a plus d'excuse. Introduire la reconnaissance vocale et la synthèse vocale dans les pratiques pédagogiques courantes devient économiquement à la portée de tout établissement : que vaut aujourd'hui un ordinateur architecturé autour d'un Pentium 166 Mhz ? (puissance de calcul minimum requise) ; même les grandes surfaces n'en vendent plus ; pour ce qui est du logiciel, en 5 ans on est passé de programmes valant 25 000 francs à moins de 1 000 francs !

Une autre inquiétude des enseignants était celle du temps. Qu'ils se rassurent ! Ce que nous proposons ici n'est pas une discipline supplémentaire à essayer de loger dans les programmes scolaires. La reconnaissance et la synthèse vocale viennent s'ajouter de façon transpa-

rente, tel un plug-in, aux fonctionnalités de votre traitement de texte sous Windows. Et à raison de 140 voire 160 mots entrés par minute contre 50 en moyenne au clavier, l'exercice d'expression écrite prend bien moins de temps.

Motivation pédagogique

Bien sûr, tous les établissements sont aujourd'hui encore loin de pouvoir disposer d'un ordinateur fonctionnant au moins à 166 Mhz, mais il est déjà temps d'y penser voire de construire un projet visant à aider les enfants en difficulté ou handicapés par rapport à la lecture et à l'écriture car la reconnaissance et la synthèse vocales sont sur le point de bouleverser l'ordre des choses au plan de la communication homme/machine.

Certes, la reconnaissance vocale ne sait toujours pas traiter la parole spontanée, en prenant en compte les hésitations, les reprises et les phrases dont la structure est éloignée de celle de la langue écrite. On ne sait pas encore faire en sorte que les systèmes vocaux, au delà de la simple reconnaissance, comprennent aussi ce qui est dit, mais ces imperfections sont très utiles, elles doivent être exploitées. Elles nous apprennent à être vigilants et à rester critique sur le travail de la machine. En second lieu, la correction des erreurs de compréhension d'un mot permet de soulever les ambiguïtés de la langue et l'enseignant peut ainsi rappeler fort à propos certaines règles du français. Enfin, l'élève comprend rapidement qu'il doit faire des efforts dans le domaine de l'articulation et de la syntaxe.

Le rôle de l'enseignant est de se tourner vers le futur, et il est certain que ce qui est réservé aujourd'hui à tous les professionnels de la saisie, (médecins généralistes, avocats, écrivains, commerciaux, secrétaires) sera demain un banal outil de communication présent à tous les étages de la société. Dans les années à venir la reconnaissance et la synthèse vocales seront partie intégrante de notre vie quotidienne (la domotique et l'automobile vont en bénéficier largement). Le clavier de l'ordinateur est condamné à prendre de plus en plus de poussière, la souris à être bien moins asticotée...

Quelques activités pédagogiques

Le champ des applications pédagogiques est vaste et encore mal défini mais on peut déjà signaler :

- la reconnaissance et la synthèse vocales ne sont pas encore du niveau à traiter la dictée de Pivot, mais, quel plaisir d'écrire quand on ne sait pas encore écrire ! Dès la grande section de maternelle, la production de textes pour le journal scolaire devient envisageable ;
- toujours en maternelle, le plaisir est tout aussi grand lorsqu'on peut « lire » le message d'un autre enfant alors qu'on ne sait pas encore lire (voir plus loin SpeakBack) ;
- pour les plus « vieux » du CE et du CM l'exercice d'expression écrite prend une tout autre dimension. Le nombre de mots entrés par minute est pratiquement le double de ce qui est produit au clavier. La motivation est plus grande. L'élève dispose d'un tout autre rapport production/temps ;
- effectuer une addition, une soustraction... pour vérifier ses calculs ;
- la mise en page est bien moins laborieuse. Il est inutile de fouiller dans les menus du traitement de texte (on peut même les faire disparaître). Quelques commandes orales suffisent pour agir sur les attributs des caractères, la mise en place des paragraphes devient d'une simplicité enfantine ;
- corriger un problème de ponctuation absente ou mal positionnée ;
- écrire sous la dictée de la machine (l'enseignant est plus libre pour circuler dans la classe) ;
- l'élève peut corriger la dictée vocale du maître ou d'un camarade ;
- au collège et en cours de langue en général, on observe que la timidité est bien moins grande face à la machine. L'élève s'exprime davantage ;
- l'ordinateur peut jouer le rôle d'un répétiteur pour aider l'élève à mémoriser un texte, une règle, un théorème, une poésie (synthèse vocale en boucle)

Quels logiciels ?

Tout d'abord, il faut bannir tous ceux qui ne disposent pas de synthèse vocale pour l'analyse des documents préexistants (lecture de

texte) et veiller à ce qu'ils soient intégrables aux applications Windows afin d'en faciliter l'utilisation par les plus jeunes.

La dictée doit pouvoir se faire en mode continu (c'est le cas de tous les logiciels récents). Choisir un dictionnaire si possible personnalisable et d'une capacité supérieure à 30 000 mots. Enfin, les macros vocales sont bien utiles aux tout petits qui n'ont qu'un seul mot à prononcer à la place de toute une expression.

Pour l'heure, seulement deux programmes on retenu l'attention du laboratoire TADT (Traitement et Analyse des Données Textuelles) de Paris VIII : ViaVoice Gold 4.3 (IBM) pour ses capacités à piloter le PC et Voice Xpress Plus (Lernout & Hauspie) pour sa facilité à éditer les textes.

Ce dernier logiciel, comprenant un dictionnaire plus léger, est fourni avec un éditeur de texte, ce qui permet de faire quelques économies substantielles sur l'achat d'un traitement de texte.

Le programme fait tout seul la différence entre une commande de mise ne page et la dictée sans qu'il soit nécessaire de passer d'un mode de fonctionnement à l'autre, ce qui simplifie l'utilisation du programme pour les plus jeunes utilisateurs qui ne sont pas encore trop familiarisés avec le PC.

À partir d'internet

Mis à part de ViaVoice Gold et Voice Xpress Plus dont on trouve respectivement une démonstration sur <http://ibm.com> et <http://www.lhs.com/speechtech/demos.asp>, on commence à rencontrer de petits logiciels de reconnaissance vocale spécialisés, en version d'évaluation ou même gratuits par simple téléchargement sur le réseau des réseaux. Cependant ils sont pour la plupart en anglais...

InteractiVoice Lite, système de commandes vocales pour Windows et ses applications et TextHelp ScreenReader shareware de synthèse vocale pour lire tous vos documents, sont téléchargeables à l'adresse : <http://www.download/pc>.

SpeakBack, programme sachant relire à voix haute toute information sélectionnée dans une application Windows, qu'il s'agisse de textes ou de tableaux de chiffres. Le programme sait épeler chaque terme et peut fonctionner en boucle tel un répétiteur. La vitesse de lecture, le volume de la voix et la tonalité sont paramétrables. SpeakBack est

adapté à la langue française, il respecte les liaisons et les intonations. Le tout à moins de 300 francs (<http://www.mysoft.fr>)

Conclusion

Face à de tels bénéfices potentiels que sont les quelques milliers de francs qui aideront les enfants à ne pas rester sur le bord de la route du progrès ?

La reconnaissance et la synthèse vocales fonctionnent - avec quelques imperfections, soit - mais tout comme une automobile dont il faut tenir le volant pour s'adapter en permanence aux difficultés du parcours. Alors pourquoi ne pas se laisser tenter ?

Philippe GALIANA

Traitement et analyse des données textuelles

Université de Paris VIII

galiana_philippe@compuserve.com

galiana@ipt.univ-paris8.fr

Article dicté en utilisant le programme Voice Xpress Plus.