

# L'INTELLIGENCE ARTIFICIELLE

## COMMENT

### se faire entendre

### d'un ordinateur . . . .

**L**A reconnaissance automatique de la parole peut avoir plusieurs objectifs : l'apreuve des techniques générales de reconnaissance des formes et d'intelligence artificielle, la simulation du comportement humain, les entrées vocales dans un contexte de dialogue homme-machine.

Dès 1969, le Centre National d'Études des Télécommunications lançait l'étude de la machine de bureau vocale : le système devait pouvoir opérer en temps réel et être capable de reconnaître, après un rapide apprentissage, une vingtaine de mots prononcés par un locuteur à la fois; il était complété par une unité à réponse vocale, fournissant les résultats de calculs arithmétiques. Après la fin des mises au point, en septembre 1970, plusieurs locuteurs ont pu dicter des pages de calculs arithmétiques sur place ou à distance.

1970 a marqué le début d'une collaboration élargie, pluridisciplinaire entre spécialistes francophones, et marquée par des journées d'études annuelles, organisées au sein du Groupement des Acousticiens de Langue Française (G.A.L.F.), sur la perception, la synthèse, la reconnaissance de la parole; dans le cadre de l'AFCEP, un groupe de travail, animé par J.-Y. GRESSER, se penche, depuis plusieurs années, sur la reconnaissance automatique de la parole.

#### LA PARAMÉTRISATION DE LA PAROLE...

Dans une conversation courante, de nombreux phonèmes ne sont pas réellement prononcés : le débit de parole peut, ainsi, être plus rapide et la fatigue moins grande. Étant donnée la grande redondance au niveau « mots », d'une part, et au niveau « phrases » et « idées » d'autre part, l'information n'est pas perturbée. Par contre, un système de reconnaissance de la parole au niveau « phonèmes » sera, évidemment, perturbé. Dans l'avenir, pour des dispositifs capables de traiter une conversation, il sera toutefois nécessaire de reconnaître tous les phonèmes réellement prononcés, de façon à pouvoir reconstituer le message réellement émis, en fonction du vocabulaire des règles mises en mémoire. En attendant, le locuteur qui veut se faire comprendre par un ordinateur, doit s'imposer une prononciation lente et bien articulée.

Les ondes sonores émises par les organes vocaux peuvent être traduites en une tension électrique au moyen d'un microphone. Ce signal électrique peut-être considéré comme stationnaire de temps à autre, sur des durées de 100 millisecondes, environ; parfois aussi, il est purement transitoire. C'est

un tel signal que l'on cherche à décomposer en une suite de phonèmes.

La reconnaissance du signal vocal est conventionnellement décomposée en deux phases principales : le prétraitement et la reconnaissance.

La phase du prétraitement comprend toutes les opérations nécessaires à l'élaboration d'une série de paramètres à partir du signal temporel. Au cours de cette phase, le signal est segmenté en éléments minimaux; pour chaque segment minimal, on détermine les trois premiers formants, ainsi que les paramètres proprement dits, en utilisant une batterie de filtres passe-bandes à large bande passante.

L'élaboration des paramètres destinés à représenter le contenu du signal vocal est une étape essentielle. Les sens vocaux peuvent-être divisés en deux grandes catégories, suivant qu'il y a, ou non, vibration des cordes vocales, c'est la mélodie : un premier paramètre concerne l'existence de cette vibration, et, si elle existe, il faut en mesurer sa fréquence. Une troisième série de mesures définit la forme du spectre d'énergie du signal en fonction de la fréquence (fig. 16). A l'École Nationale Supérieure des Télécommunications a été étudiée une méthode de paramétrisation qui se différencie

des approches habituelles par la mise en œuvre de filtres à flancs doux et à large bande, donnant peu de retard et peu de déphasage, et l'utilisation conjointe des informations issues de deux filtres très voisins.

#### ... ET LA RECONNAISSANCE DE LA PAROLE

A l'issue de la phase de prétraitement, on dispose donc d'une décomposition des mots en segments possédant des caractères spécifiques. Ici se pose le problème de la reconnaissance, qui consiste à comparer les caractéristiques de chaque son incident à celles de phonèmes ou phonotomes\* mis en mémoire dans un ordinateur.

Dans un premier temps, un programme d'apprentissage permet de créer un dictionnaire réduit de phonèmes ou phonotomes normalisés.

\* La phonétique considère l'existence de sons élémentaires appelés « phonèmes » : on dénombre, en français, environ 36 phonèmes comprenant 16 voyelles et 20 consonnes. Un « phonotone » (ou « diphone », ou encore « diphonème ») est constitué par l'association de deux phonèmes.

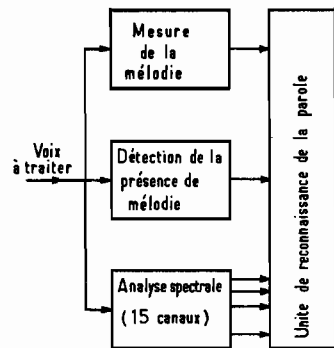


Fig. 16

Mais reconnaître une suite de diagrammes phonétiques n'est pas le résultat le plus intéressant en reconnaissance automatique de la parole. Il convient maintenant de créer un assemblage de ces diagrammes et de constituer des symboles phonétiques par association de plusieurs phonèmes ou phonotomes successifs. Puis, il devient nécessaire de regrouper les phonèmes en mots, c'est-à-dire de former des suites de symboles phonétiques, entre lesquels s'intercalent des blancs. Le principe de formation de mots peut-être très simple : le premier caractère du texte est comparé à tous les mots d'un lexique; s'il ne peut constituer un mot, on lui ajoute le caractère suivant, et l'on consulte, de nouveau le lexique... et ainsi de suite jusqu'à ce que l'on trouve le mot recherché. Après quoi, le caractère suivant est considéré isolément et l'on recommence une nouvelle investigation. Un programme (« LAIUS », réalisé par A. Braure) reposait initialement, sur ce principe.

Les essais de ce programme ont été effectués avec des phrases du type : « il règne une chaleur moite; on entend quelque part un léger sifflement, comme s'il y avait une fuite à la tuyauterie de chauffage »; dans le lexique figuraient tous les mots de ce texte, et quelques autres, comme « chat », « leur », « moi », « ouate », « an », « tant »... Le premier résultat fut curieux; qu'on en juge : « Il règne une chat leur mou a ton an tant quelque pas rein les jets sifflement comme si li avait une fut y tas la tuyauterie de chauffe a je »!

L'association de règles permet de diminuer considérablement les fautes. Néanmoins les résultats actuels ne sont, sans doute, pas suffisants pour permettre de répondre aux diverses questions auxquelles se heurtent les cher-

cheurs : faut-il chercher à reconnaître des phonèmes, des syllabes, des mots ou des phrases? en temps réel ou en temps différé? Faut-il prévoir des phases d'apprentissage par la machine seule ou par l'usager? pour l'ensemble des phonèmes et des prononciations normalisées? Quels sont les taux d'erreurs admissibles : 10 %, 1 % voire 0,01 %?

(à suivre)

MARC FERRETTI

### ON LIRA AVEC INTÉRÊT...

« *Reconnaissance de la parole et dialogue homme-machine* » par J.-Y. Gresser. L'ÉCHO DES RECHERCHES (CNET). Octobre 1973.

« *Reconnaissance automatique de la parole* » par M. Baudry, B. Dupeyrat, C. Frank. Communication présentée aux Journées du Groupe « Communication parlée » du G.A.L.F. - Lannion - Mai/Juin 1972.

« *Communication homme-machine sur support vocal* » par C. Guégen, A. Maïssis, L.-F. Pau. L'ÉCHO DES RECHERCHES (CNET). Octobre 1972.

« *Recherches actuelles sur l'extraction de caractéristiques et la reconnaissance de la voix parlée* » par M.-J. Vigneron, M.-Lamotte, J.-P. Haton, J. Brenont. AUTOMATISME. Décembre 1970.

« *Segmentation de la parole et reconnaissance des syllabes à l'intérieur des mots* » par G. Mercier. AUTOMATISME. Mars 1972.

« *Application des techniques statistiques à la reconnaissance de la parole* » par C. Berger-Vachon. AUTOMATISME. Mars 1972.

« *Idées générales sur la reconnaissance des formes appliquée à la parole* » par C. Roche. AUTOMATISME. Mars 1972.

« *Reconnaissance de phonème au moyen d'une cohérence artificielle* » par P. Alinat. AUTOMATISME. Mars 1972.

« *Reconnaissance de la parole en temps réel* » par J. Caelen, S. Castan, G. Perennou. AUTOMATISME. Mars 1972.

« *Analyse, synthèse et reconnaissance automatique de la parole* » par J.-S. Liénard. Thèse de doctorat soutenue en avril 1972 à l'Université de Paris VI.



partout  
des amis  
vous  
attendent!

devenez  
**radio-amateur**

pour occuper vos loisirs tout en vous instruisant.

Notre cours fera de vous un EMETTEUR RADIO passionné et qualifié.

Préparation à l'examen des P.T.T.

RAPY

**GRATUIT!**

DOCUMENTATION SANS ENGAGEMENT  
Remplissez et envoyez ce bon à

**INSTITUT TECHNIQUE ELECTRONIQUE**

ENSEIGNEMENT PRIVÉ A DISTANCE

35801 DINARD

NOM : \_\_\_\_\_

ADRESSE : \_\_\_\_\_

HPA 42