

# L'INTELLIGENCE ARTIFICIELLE

## L'ORDINATEUR EST A L'ÉCOUTE

(suite voir n° 1429)

### APRÈS L'EUPHORIE DES ANNÉES 50 L'ÉCLIPSE

LES études sur la reconnaissance de la parole ont connu, il y a une quinzaine d'années une grande vague d'intérêt de la part de nombreux centres de recherches ; mais le problème s'est rapidement révélé beaucoup plus complexe qu'on ne pouvait le supposer initialement, les méthodes d'études ainsi que les moyens de traitement se révélant insuffisants ou mal adaptés.

Après l'euphorie des années 1955-1965, succède un optimisme modéré (certains parlent de désenchantement). Depuis les toutes premières études, peu de progrès sensibles ont été enregistrés ; la reconnaissance de la parole, décomposée en mots, ou ensembles de mots dégagés de tout contexte, a atteint un optima, qui décroît faiblement avec l'étendue du vocabulaire, mais rapidement avec le nombre de locuteurs. Ainsi, pour un locuteur unique, un automate peut reconnaître, en temps réel, un vocabulaire de plusieurs dizaines de mots, avec une probabilité de bonne compréhension comprise entre 80 et 95 % ; lorsque le vocabulaire augmente, le temps de réponse s'accroît, les performances de l'automate se dégradent. Si l'on veut augmenter le nombre de locuteurs à reconnaître, il faut réduire le vocabulaire : quelques dizaines de mots pour quelques dizaines de locuteurs sont reconnues correctement avec une probabilité de 60 à 80 %.

On estime que la reconnaissance automatique, et en temps réel, pour un vocabulaire illimité, et pour un grand nombre de locuteurs, fera encore l'objet de longue recherches étalées sur une période de dix à vingt ans.

Après avoir connu une éclipse au cours de la dernière décennie, les études sur la reconnaissance de la parole ont repris un essor très net dans de nombreux pays, en particulier aux U.S.A., au Japon, et en U.R.S.S. Les premières conséquences apparaissent sur l'orientation nouvelle de la recherche, aux Etats-Unis : la Société I.B.M. reporterait actuellement tout son potentiel dans le domaine de la parole vers la reconnaissance ; un groupe de travail, crée dans le cadre de l'Advanced Research Project Agency, définit un plan de recherche commun à plusieurs laboratoires importants pour le développement de « systèmes de compréhension de la parole ».

On peut distinguer sept catégories d'options, d'appareils, sur lesquels on peut construire des machines obéissant à la parole : les transducteurs d'entrée, les unités sélectionnant les informations pertinentes, les intégrateurs, les convertisseurs analogues-numériques, la logique, les systèmes capables de délimiter les « formes acoustiques » et les transducteurs de sortie.

Le dialogue homme-machine, qui se fait encore par des langages codés utilisant des supports du type cartes ou rubans perforés, aurait tout intérêt à utiliser directement la parole. Il présenterait ainsi une diminution des risques d'erreur, dus à la grande

redondance de la parole, un gain de temps et une facilité de mise en œuvre importante.

Les applications de la reconnaissance, par les machines, de la voix parlée présentent des perspectives intéressantes en informatique ; elles ne se limitent cependant pas à la seule informatique. En matière de télécommunications, la numérotation parlée semble utopique, mais la

consultation purement vocale d'un centre de renseignements et l'automatisation complète d'un service du réveil peuvent être envisagées : les services offerts doivent rester simples pour être techniquement solubles, mais le grand nombre des usagers potentiels les rend intéressants. Dans le domaine postal, l'introduction du tri vocal simplifierait la manipulation des sacs et paquets :



Photo n° 24. — Pour l'introduction de données diverses dans un calculateur, la voix est, pour l'homme, le moyen le plus rapide, et certainement le moins fatigant. (Cliché : National Physical Laboratory, Teddington.)

de sérieuses études ont été entreprises sur ce thème aux Etats-Unis. Une application voisine est celle envisagée en Union Soviétique, au port de Mourmansk pour la répartition des cargaisons entre les navires.

### L'INFORMATIQUE OFFRE LES DEBOUCHES LES PLUS IMMEDIATS

La reconnaissance automatique de la parole est un problème récent (les premières réalisations datent de 1951), lié au développement de l'automatisme et de l'informatique. Ainsi, l'informatique scientifique peut fournir des débouchés chaque fois qu'une interaction est nécessaire entre l'utilisateur et ses programmes en cours d'exécution : un système pourra alors être qualifié, à juste titre, de conversationnel.

En informatique de gestion, si le volume des informations provenant de la machine n'est pas trop important, la parole est le support adapté aux interrogations, créations, mises à jour de fichiers.

En automatique, on peut envisager la commande vocale des machines de production, directement ou par l'intermédiaire d'un système de conduite de processus, l'opérateur désignant, oralement, les modalités d'exécution d'une série de programmes déjà répertoriés. Les applications portentielles sont extrêmement nombreuses : de l'écriture des textes dictés (le « phonétographe » de Dreyfus-Graf) à la conduite des véhicules spatiaux (la voix pourrait devenir un instrument indispensable au pilotage). En 1969, un prototype de machine à écrire tapant directement sous la dictée fut présenté, au Salon de la Physique, qui se tenait à Londres : il fallait cependant, pour se faire comprendre de la machine, dénommée VOTEM (Voice-Operated Typewriter Employing Morsecode), dicter le texte en code morse, les traits étant prononcés « DAH » et les points « DI », les intervalles entre mots étant indiqués par « TOP ». Un second prototype, plus élaboré, répond directement à la dictée « en clair », mais il écrit en syntaxe phonétique.

Ces sept catégories forment un tout, et réagissent les unes sur les autres. Par exemple, la logique et les formes acoustiques seront différentes selon que le transducteur de la parole sera

un microphone dynamique passant 50 à 8 000 Hz, une ligne téléphonique admettant 300 à 3 400 Hz, ou un laryngophone transmettant 80 à 1 500 Hz.

L'étude du problème général de la reconnaissance de la parole semble être, actuellement, du domaine de la recherche fondamentale. Il faut entreprendre des travaux sur la perception, sur les règles linguistiques, effectuer des tests psycho-acoustiques. Les recherches, en France, se développent dans toutes les catégories d'options. Elles sont l'œuvre de nombreux laboratoires, à Paris et en province.

### LA RECONNAISSANCE VOCALE DEBUTE EN 1951

Voici vingt ans, lorsque commencèrent les travaux sur la reconnaissance vocale, les problèmes technologiques limitèrent l'ampleur des recherches : l'orientation faisait alors tout juste son apparition. En 1951, S.P. Smith présente un détecteur de phonèmes ; J.A. Dreyfus-Graf met au point son « phonétographe », appareillage analogique composé de 20 filtres passe-bande et de circuits identificateurs de phonèmes : le phonétographe utilise des « compresseurs sélectifs » qui augmentent l'émergence de certains phonèmes. Le résultat, obtenu en temps réel, est spectaculaire ; cependant l'appareil ne fonctionne qu'avec une seule voix, astreinte à une prononciation particulière qui accentue chaque phonème.

H.F. Olson et H. Belar, en 1962, envisagent la reconnaissance de « syllabes phonétiques », que le locuteur doit articuler séparément, avec un bref arrêt entre chaque syllabe ; il s'agit donc presque d'une reconnaissance par mots (et non plus par phonèmes). Il suffirait d'utiliser 2 000 telles syllabes pour couvrir quasiment toute la langue anglaise.

Quelques années auparavant, K.H. Davis, R. Biddulph et S. Baleshek avaient déjà abordé la reconnaissance de « manière globale », de mots courts : les dix chiffres (« zéro » à « nine », en anglais) sont reconnus analogiquement, avec un bon taux de réussite pour une seule voix. Des travaux similaires furent menés, en 1960, par P.B. Denes et M.V. Matthews.

La réalisation de P. Vicens, en 1969, est parmi les plus remar-

# AKAI®

## le N°1 DE L'AUDIO VISUEL

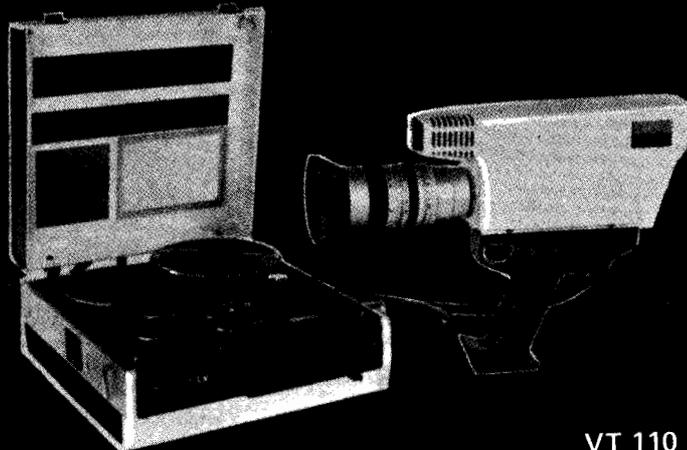


VT 100S

Enregistrement et reproduction instantanés SON et IMAGE avec contrôle permanent. Arrêt sur image pendant la reproduction. Alimentation par accus, secteur et batterie auto.

L'ensemble caméra, enregistreur et téléviseur de contrôle

6 400 F



VT 110

Haute-fidélité d'enregistrement et de reproduction grâce aux têtes Vidéo en Cristal de Ferrite.

MONITOR INCORPORE  
Poste synchro

L'ensemble caméra, enregistreur et téléviseur de contrôle.

9 800 F

EN DEMONSTRATION ET VENTE AU

stéréo hi-fi CLUB

# CIBOT

12, rue de Reuilly  
75012 PARIS  
345-65-10/343-66-90  
136, bd Diderot  
75012 PARIS  
346-63-76

quables de ces dernières années : la procédure d'identification des mots prononcés est extrêmement souple ; il est prévu une procédure d'apprentissage en cas de mauvaise identification. Le système donne d'excellents résultats (entre 90 et 100 %) pour un locuteur, en trois ou quatre cycles d'apprentissage, même pour un vocabulaire étendu (jusqu'à plus de 500 mots).

La première réalisation commerciale en matière de reconnaissance vocale date de 1971 : le « Voice Command Systems » de J.W. Glenn et M. H. Hitchcock reconnaît, de manière fiable, 24 mots isolés, moyennant 5 cycles d'apprentissage par le même locuteur.

Il semble qu'actuellement, la reconnaissance analytique, par phonèmes, ait cédé le pas à la reconnaissance globale de mots isolés, mémorisés par apprentissage : ce dernier mode de reconnaissance paraît devoir être plus rapidement opérationnel dans l'environnement des calculateurs.

Cependant, personne n'a, jusqu'à présent, résolu de manière satisfaisante le problème d'adaptation à un locuteur quelconque. Il faut d'ailleurs remarquer que, très généralement, les résultats de reconnaissance portent sur des mots prononcés de manière

Tableau 9. — Options et matériels

	Options	Matériels
1	Transducteurs de signaux vocaux	— Microphone 160 000 bits/s — Téléphone 32 à 64 bits/s — Laryngophone 16 000 bits/s
2	Présélecteurs d'informations pertinentes	— Compresseurs d'amplitudes — Analyseurs de spectre — Extracteurs de mélodie
3	Intégrateurs temporels	— Redresseurs et filtres passe-bas — « Fenêtre » de temps
4	Convertisseur analogique numérique	— Quantificateur de niveaux et de durées — Echantillonneur — Multiplexeur
5	Traitement logique des informations	— Mémoires — Calculateurs
6	Formes acoustiques à délimiter	— Phonèmes (30) — Syllabes (30 × 30 = 900) — Mots (900 × 900 = 810 000) — Phrases (30 <sup>16</sup> ≈ 10 <sup>22</sup> )
7	Transducteurs électro-optiques ou mécaniques	Ecrans de visualisation — Imprimantes — Phonacteurs, méglographes, phonétographes, vocographes...

comparable, en rythme, niveau, intonation ; il n'existe pas de système travaillant indifféremment sur la voix parlée, chuchotée, criée, ou chantée, même pour un seul locuteur.

(à suivre)

Marc Ferretti.

Tableau 10. — Des laboratoires travaillant à la reconnaissance de la parole

— Ecole nationale supérieure des télécommunications (Paris).
— Société lannionnaise d'électronique (Lannion).
— Laboratoire d'électricité et d'automatique (Nancy).
— Centre national d'études des télécommunications (Lannion).
— Laboratoire de physique électronique (Lyon).
— Laboratoire de reconnaissance des formes (Paris).
— Thomson-CSF (Cagnes-sur-Mer).
— Laboratoire de cybernétique (Toulouse).
— Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur (Orsay).
— Laboratoire d'acoustique et d'électronique (Paris).
— Centre d'études nucléaires (Saclay).
— Ecole nationale supérieure d'électronique (Grenoble).

## 50 ANS DE RADIODIFFUSION EN ALLEMAGNE

C'EST le 29 octobre 1923 qu'eut lieu, à 20 heures, la première émission allemande de radiophonie. Emises du Vox-Haus de Berlin les ondes de 400 mètres de longueur étaient d'une puissance assez faible : 1 kW.

Le nombre des récepteurs de radiodiffusion n'était que de 1 508 au 1<sup>er</sup> janvier 1924. Mais la population allemande fut rapidement attirée par les intérêts de la radio en sorte que, deux ans plus tard, plus d'un million d'appartements y contenaient des radio-récepteurs. En 1943, leur nombre dépassait 16 millions.

De nos jours il y a, en Allemagne Fédérale, 19 millions de récepteurs de radiodiffusion et 18 millions de téléviseurs.

Le premier émetteur vraiment puissant de radio fut réalisé et installé à Langenberg par la plus ancienne et la plus importante

entreprise allemande de radio : « Telefunken » ; sur fréquence moyenne ses émissions avaient une puissance de 15 kW. C'est encore « Telefunken » qui créa, en 1929, le premier radioémetteur à ondes courtes d'une puissance de 8 kW.

Depuis très longtemps « Telefunken » est uni avec « AEG » (Allgemeine Elektrizitäts-Gesellschaft = Société générale d'électricité) qui fut fondée il y a quatre-vingt-dix ans.

De nos jours, cette si grande entreprise a des succursales dans 39 pays ; elle emploie 148 000 personnes en Allemagne et 21 600 à l'étranger. Quant à son chiffre d'affaires, il s'élevait, au cours des six premiers mois de l'année courante, à 5,9 milliards de marks, soit plus de 10 000 milliards de nouveaux francs. Par rapport à 1972, cela représente une augmentation de 14 %.

Tous les deux ans, « AEG-Telefunken » organise des réunions de journalistes techniques les faisant bénéficier de plusieurs conférences relatant les récents progrès réalisés dans cette grande entreprise.

Une telle réunion eut lieu en octobre dans un immense hôtel récemment construit à Hambourg. Y furent invités 40 journalistes techniques allemands et 27 étrangers émanant de neuf pays : Autriche, Belgique, Danemark, Etats-Unis, Finlande, France, Hollande, Suède et Suisse. Le seul journaliste qui fut invité de notre pays est Eugène Aisberg, président de l'U.I.P.R.E. (Union internationale de la presse radiotechnique et électronique).

Ce grand colloque de presse était très bien organisé par Friedrich Bender, chef du service de presse et dirigé par le

Dr Hans Groebe, président de la grande entreprise.

En deux jours ont été prononcées onze références, chacune suivie de très vivantes discussions. On y a notamment exposé les problèmes d'appareils électroniques utilisés dans les satellites européens, de techniques électroniques d'éclairage, de procédés récents de transmissions à haute puissance et de méthodes assurant la sécurité des vols.

En outre, l'ensemble des invités fut conduit dans une des récentes usines de « AEG-Telefunken » située dans les proches environs de la ville.

Tous ceux qui ont ainsi bénéficié de cette invitation ont constaté que les progrès de la technique électronique vont de plus en plus vite et s'étendent dans tous les domaines des activités humaines.

E. GLACIMONTO.