



**HAL**  
open science

## 24h pour écouter parler. Conversations sur la langue française : INNOVER

Joseph J Mariani

► **To cite this version:**

Joseph J Mariani. 24h pour écouter parler. Conversations sur la langue française : INNOVER. 2023. hal-04430140

**HAL Id: hal-04430140**

**<https://hal.science/hal-04430140>**

Submitted on 31 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

24h pour écouter parler. Conversations sur la langue française : INNOVER  
Joseph Mariani  
LISN, Université Paris-Saclay, CNRS-LIMSI

### Explorer

Débuter travaux en reconnaissance de la parole en 1975, début de l'informatique.  
Problème résolu en grande partie à présent  
Grande satisfaction pour un chercheur, pour une communauté de chercheurs, comme un explorateur qui découvre une terre inconnue  
IBM 1130, 16 KOctets de mémoire. Un million de fois plus de mémoire à présent  
Analyse du signal de parole meuble d'un mètre cube, à présent 3 lignes de code

### Evaluer

Evaluer les performances d'une technologie.  
Permet de choisir les bonnes approches. Jusqu'en 1980, débats théoriques sans support expérimental. Mise en place par la DARPA de campagnes d'évaluation dans le domaine de la reconnaissance de la parole, puis dans la reconnaissance d'images, la robotique, les véhicules autonomes. Cela a permis d'immenses progrès.  
Egalement déterminer si l'état d'une technologie permet de réaliser une application. Evite de financer des projets qui n'ont aucune chance d'aboutir.

### Arriver au bon moment

Mon sujet de thèse de Docteur-Ingénieur que j'ai soutenue en 1977 était le dialogue oral pilote-avion dans les avions de chasse. Sans doute les conditions les plus difficiles pour la reconnaissance vocale, à cause du bruit, de la prise de son dans un masque à oxygène et du stress du pilote qui dans certains cas est incapable de parler. Et surtout du fait qu'il n'existe pas de données pour effectuer l'apprentissage du système puisque ces données sont secret-défense. Cela fait que cela n'a pas été implanté sur le Rafale qui était en cours de design et, la préparation d'un avion de chasse se faisant sur des périodes de quelques décennies, n'est toujours pas implanté à ma connaissance sur un avion de chasse à l'heure actuelle.

Cela ne nous a pas empêché d'être les premiers en France à réaliser un système de reconnaissance vocale commercialisé, et à développer des applications comme la composition téléphonique vocale, une serrure vocale, mais c'était trop tôt par rapport aux besoins de la société.

### Persévérer

La plupart des sociétés françaises, mais aussi européennes, ont petit à petit abandonné les recherches dans ce domaine du traitement de la parole ou du texte, France Télécom devenu Orange, Bull, Alcatel, Thomson, Philips, Siemens, SEL, Telefunken, alors que les GAFAM sont arrivées en force aux Etats-Unis, mais aussi à présent les BATX, leurs équivalents chinois, qui ont compris l'enjeu économique, politique et culturel que représente la maîtrise du traitement automatique de la langue et de la connaissance. Très récemment, les publications chinoises sont plus nombreuses que les publications américaines et, fait notable, les européens sont quasi absents de celles qui touchent le secteur industriel.

### Imaginer

Il faut imaginer ce qu'une innovation pourrait amener dans la société. J'avais moi-même rédigé une pleine page dans Le Monde en 1980, il y a plus de 40 ans, sur l'utilisation des technologies

vocales, qui se généralisent à présent. Mais il faut aussi imaginer ce qu'on peut faire avec les technologies qui existent. Les modèles de langue ont été utilisés depuis des décennies pour la reconnaissance vocale, et des systèmes de dialogue de type chatbot ont aussi été réalisés, mais il fallait faire preuve d'imagination pour produire ChatGPT et susciter le formidable succès médiatique que cela connaît.

### Etre surpris

Et c'est également être surpris par des auxquelles on ne s'attendait pas. Je me souviens quand mon système de reconnaissance s'est mis à reconnaître la phrase qu'il venait lui-même de prononcer et à se répandre à lui-même. Ça a été un grand moment !

### Dénommer

Dénommer correctement les champs scientifiques est aussi important. J'avoue que je suis hérisé d'entendre « une » intelligence artificielle, alors que dans le milieu scientifique on a toujours parlé de « l'intelligence artificielle » de façon générique. L'intelligence artificielle peut signifier faire effectuer par une machine les fonctions qui sont habituellement le propre d'un être humain, entendre, lire, écrire, parler mais aussi conduire une voiture ou composer une chanson, ou cela peut concerner un type de méthodes pour traiter l'information.

Il est d'ailleurs amusant de se souvenir qu'initialement, on avait coutume de qualifier de méthodes d'intelligence artificielle celles qui utilisaient des connaissances humaines, et de reconnaissance des formes, celles qui consistaient à mettre en correspondance une forme à reconnaître, signal de parole ou scène visuelle, avec un ensemble de formes de références. Il y avait même en France une conférence RFIA qui rassemblait et donc dissemblait ces deux types d'approches. On constate à présent que ce sont ces méthodes de reconnaissance des formes qui se sont complexifiées grâce aux progrès de l'informatique et à la disponibilité de données de taille gigantesque à travers l'apprentissage automatique qui ont permis la réalisation des systèmes actuels que l'on qualifie d'intelligence artificielle.

### Diversifier

Répondre aux besoins de tous est une exigence. Certains se souviennent que les premiers claviers d'ordinateurs n'avaient pas de signes diacritiques, pas de « é ». Cela a été corrigé par les constructeurs pour répondre aux besoins des clients qui parlaient des langues qui en avaient besoin, éventuellement en créant des claviers virtuels et non plus mécaniques. Puis le problème s'est reposé avec les premiers logiciels de messagerie électronique, et a été résolu. Certaines sociétés informatiques avaient fait la liste des langues parlées dans le monde, et mettaient une limite à celles qui n'étaient pas intéressantes économiquement. Cela a été modifié avec l'arrivée des réseaux sociaux où les utilisateurs voulaient pouvoir s'exprimer dans leur langue, et la nécessité de devoir modérer des contenus dangereux, de l'intox. En 2019, il apparaissait que seulement 2% des langues parlées dans le monde bénéficiaient de technologies des langues. Nous avons donc organisé une conférence avec l'UNESCO, LT4All, technologies des langues pour tous, et on note depuis un élargissement important de la couverture linguistique de ces technologies.