

## Traitement de la parole

### INTRODUCTION

Depuis de nombreuses années, les films de science-fiction présentent des ordinateurs, des robots, dotés de la parole.

Depuis une trentaine d'années, la science est en passe de "rattraper" la fiction puisqu'il existe des applications permettant :

- à l'ordinateur de parler (synthétiseurs de parole),
- à l'homme de s'adresser oralement à un ordinateur (reconnaisseurs de parole).

Nous allons donc tout naturellement nous intéresser au domaine de la synthèse de parole d'une part, et de la reconnaissance d'autre part.

### SYNTHESE :

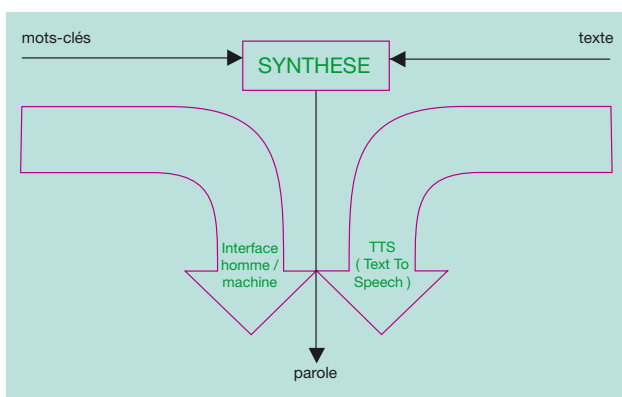
Texte (ou mots-clés) => signal de parole

Dans le cas de la synthèse, nous disposons d'une entrée texte (ou mots-clés) à partir de laquelle nous devons produire de la parole de synthèse (faire parler l'ordinateur).

Parmi les applications, nous pouvons citer :

- la lecture d'e-mails par téléphone,
- les serveurs vocaux,
- le contrôle (par exemple, dans les avions, remplacement des témoins lumineux par un message vocal clair à destination du pilote),
- les applications pour non-voyants telles celles qui permettent de faire lire des documents (journaux, factures) par l'ordinateur.

La figure ci-dessous représente deux types d'application de la synthèse selon que l'entrée est sous forme d'un texte continu ou sous forme de mots-clés.



### RECONNAISSANCE :

Signal de parole => information

Dans le cas de la reconnaissance, nous disposons d'un signal de parole dont nous devons déduire une information.

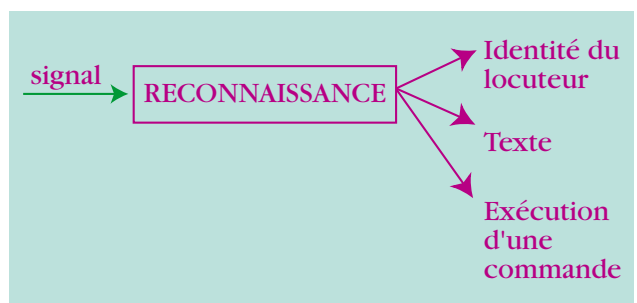
Parmi les systèmes de reconnaissance, nous pouvons citer :

- les systèmes de reconnaissance du locuteur,
- les systèmes de reconnaissance de la parole.

Les systèmes de reconnaissance du locuteur ont pour but de reconnaître le locuteur, c'est-à-dire la personne qui parle, et trouvent un grand champ d'applications dans le domaine des serveurs vocaux mais également du contrôle d'accès.

Les systèmes de reconnaissance de la parole ont pour but de reconnaître ce qui est dit (ordres à exécuter par exemple). Ces systèmes sont mis en œuvre dans les applications de dictée vocale, de commande vocale de machines,...

Ce domaine est appelé «reconnaissance du locuteur» ou «reconnaissance de la parole» selon le cas et est schématisé ci-dessous.



Ces deux aspects sont traités dans la suite mais notons que ces deux domaines peuvent cohabiter au sein d'une même application. Par exemple, dans une borne interactive (telles celles que l'on trouve dans les foires commerciales), nous avons à la fois de la reconnaissance vocale puisque la borne doit comprendre la demande du visiteur, et de la synthèse vocale puisque la borne doit répondre au visiteur.