

# HUMAVOX

Reconnaissance de la parole

HUman MAchine VOice eXchange

Synthèse vocale



## Introduction

Claviers et souris au placard ? Pas encore pour demain, mais l'amélioration soutenue des technologies de reconnaissance vocale laisse entrevoir le développement d'applications encore inconcevables il y a peu. Grâce à une avance certaine en termes de maturité, la synthèse vocale s'implante progressivement dans notre quotidien: horloges parlantes, services de renseignements, prévisions météorologiques en ligne, audioguides, rames de métros douées de la parole, etc.

## Objectifs

A partir de l'état de l'art en matière de traitement vocal, et avec le soutien de l'IDIAP <sup>(1)</sup>, les acteurs de ce projet se proposent de :

- réaliser un socle de développement utile à l'industrie
- promouvoir la reconnaissance avec vocabulaires limités
- faciliter la gestion des cas de non-reconnaissance
- développer l'aspect contextuel de la synthèse vocale
- caractériser l'interaction multimodale pour le canal audio
- exploiter la convergence des applications ASR <sup>(2)</sup> et TTS <sup>(3)</sup>

Des solutions concrètes sont en cours de développement pour permettre de démontrer les possibilités d'interaction de la reconnaissance vocale et de la synthèse de la parole dans un contexte de dialogue homme - machine.

Divers prototypes d'applicatifs industriels exploitent d'ores et déjà les résultats issus de la recherche fondamentale - ceux de l'IDIAP notamment.

## Applications

A l'heure actuelle, il existe trois démonstrateurs Humavox : commande vocale pour applications domotiques, publications électroniques en synthèse vocale, et dialogue vocal dédié à la gestion de dossiers de patients dans un cabinet médical. D'autres applicatifs sont en préparation : distributeurs de tickets, ballado-diffusion, etc.

## Motivations

La HES-SO VS (INFOTRONICS) possède un savoir-faire dans la conception de systèmes électroniques dédiés au traitement de l'information. Acquérir l'expertise nécessaire au développement de solutions de traitement de la parole sur des systèmes à ressources matérielles limitées constitue la motivation principale dans ce projet.

La HES-SO FR (ICT) développe de solides compétences dans le domaine des interactions multimodales. Son champ de recherche se situe à l'exacte interface de la reconnaissance de la parole, de la synthèse vocale, et donc des modes d'entrée-sortie utiles au dialogue homme - machine.

La HES-SO VD (IICT/SYSIN) est spécialisée dans les technologies de diffusion « polypublishing-crossmedia » : comment délivrer la bonne information, à la bonne cible, au bon moment, au bon format et via le canal approprié. La synthèse de la parole est une corde (vocale) de plus à son arc.

## Collaborations

Les entreprises souhaitant donner une valeur ajoutée certaine à leurs produits peuvent déjà s'inscrire dans le marché du futur toujours plus concurrentiel des « environnements intelligents ». Elles trouveront assurément, parmi les acteurs du projet, une oreille attentive pour mener des expériences pilotes communes dans ces domaines.

HES-SO VS - Frédéric Dessimoz, Michel Imhasly  
HES-SO FR - Omar Abou Khaled, Joe Farage, Elena Mugellini  
HES-SO VD - Steve Christin, Istvan Farkas, Daniel Rappo

<sup>(1)</sup> Institut Dalle Molle d'Intelligence Artificielle Perceptive

<sup>(2)</sup> Automatic Speech Recognition

<sup>(3)</sup> Text-to-Speech