

# Interactions humain-machine, objets connectés, contenus numériques, données massives et connaissance

Aghilas SINI

29 août 2016

## **1 Projet SynPaFlex : Flexibilité pour la synthèse de la parole expressive**

De nos jours la synthèse de la parole à partir du texte permet d'atteindre de très bons niveaux de qualité. L'usage de grands corpus de parole a pour une grande partie contribué à ce succès. Malgré tout, la parole synthétique générée manque encore d'émotion, d'intention et de style. A l'heure actuelle, nous ne sommes pas capables de synthétiser une voix comportant l'expressivité nécessaire pour de la lecture de livres audio sans enregistrer un locuteur afin de créer un corpus de grande taille possédant ce style.

Certains travaux de la littérature s'intéressent à la prise en compte de phénomènes liés à l'expressivité et apportent des conclusions intéressantes permettant en partie de caractériser le fonctionnement et la matérialisation de ces phénomènes. Nous envisageons ici de traiter de manière conjointe l'émotion, l'intention et le style d'élocution, puisqu'en pratique ces notions sont très liées, et ce dans une perspective d'intégration à la synthèse de parole.

L'idée du projet SynPaFlex est de s'intéresser aux différentes caractéristiques de ce qui fait l'expressivité d'une voix afin de constituer un modèle de prosodie et un modèle de modification de la chaîne phonémique, adaptés à un locuteur. Ensuite, l'utilisation de ces modèles sera explorée afin d'intégrer de l'expressivité dans les systèmes de synthèse de la parole, notamment par concaténation. Enfin, un complément de ce travail portera sur les post-traitements, nécessaires pour pallier les défauts des unités sélectionnées par le moteur de synthèse. Une approche par conversion de prosodie des unités sélectionnées sera alors envisagée comme post-traitement de la synthèse.

L'ensemble de ces travaux portera sur l'étude du français et de l'anglais afin de conserver une certaine généricité.

Ces différentes étapes permettront d'apporter des connaissances sur la manière de modifier un système de synthèse, en terme de descripteurs des unités de parole, de fonction de coût de sélection des unités, de post-traitements.

Les enjeux majeurs du projet résident dans la faisabilité des applications de la synthèse de la parole expressive, applications qui pour l'instant restent peu répandues. Des débouchés sont notamment à attendre dans le domaine des jeux vidéo (diversification des voix de synthèse, création de voix expressives adaptées à la situation de jeu), de l'apprentissage des langues (dictée, style d'élocution) ou encore de l'assistance aux personnes.

## **2 Plus D'info sur le site officiel du projet**

SynPaFlex est un projet financé par l'ANR dont l'objectif est l'amélioration des moteurs de synthèse de la parole à partir du texte à travers deux axes de recherche principaux : La modélisation et la prédiction de variantes de prononciation La modélisation et la prédiction d'une prosodie adaptée au contexte d'élocution Les principales applications visées se trouvent dans les domaines du divertissement (livres audio, jeux vidéo), des jeux sérieux (environnements virtuels), de l'apprentissage des langues (dictée, style d'élocution) ou encore pour la construction d'aides vocales. Ce projet est porté par l'IRISA et implique des membres du LLF (Laboratoire de Linguistique Formelle) ainsi que de l'ATILF.