

Université Sorbonne Paris Nord  
Sup Galilée  
Spécialité Informatique



## Conduite et gestion de projet

Spécialité : INFO2

Livrable : Sprint 1

---

# Interface de Génération Sonore par IA

---

### *Étudiants :*

Seyfeddine JOUINI  
Yosra SASSI  
Sirine TLILI  
Hasna ELGARANI  
Kaoutar BRAHIMI

### *Client :*

Vittascience

28 novembre 2024

# Table de Matières

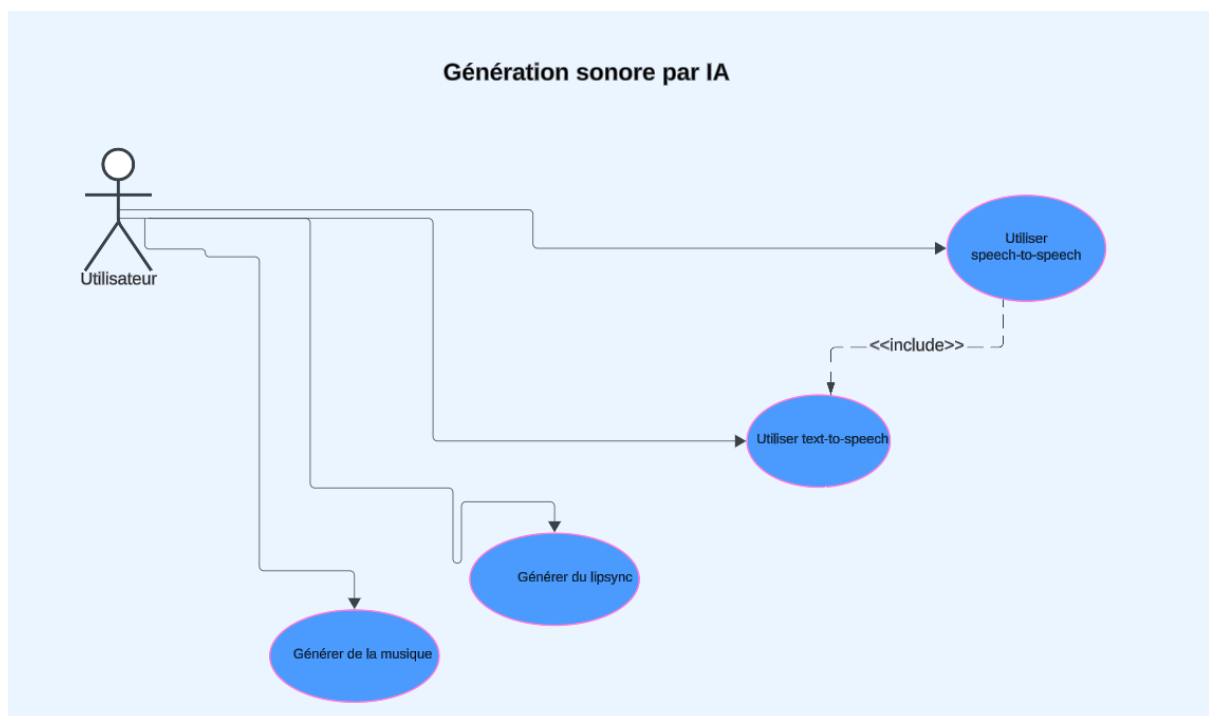
<b>Introduction</b>	<b>2</b>
<b>1. Diagramme de cas d'utilisation</b>	<b>2</b>
Hiérarchisation des cas d'utilisation	3
<b>2. Description des cas d'utilisation</b>	<b>3</b>
Cas d'utilisation 1 : Génération de speech à partir d'un prompt	3
Cas d'utilisation 2 : Speech-to-speech	4
Cas d'utilisation 3 : Lipsync à partir d'un audio et d'une image	4
Diagramme de séquence : Lipsync à partir d'un audio et d'une image	5
Cas d'utilisation 4 : Génération de musique	5
<b>Conclusion</b>	<b>6</b>

## Introduction

Ce document présente le livrable du **Sprint 1** pour le projet de développement d'un système d'intelligence artificielle dédié à la génération et transformation de contenus audiovisuels. Il inclut un **diagramme de cas d'utilisation** avec quatre fonctionnalités principales : génération de speech à partir d'un prompt, lipsync, génération de musique et speech-to-speech. Chaque cas d'utilisation est accompagné d'une hiérarchisation basée sur leur priorité. Enfin, des **diagrammes de séquence** illustrent les interactions internes du système pour ces fonctionnalités.

## 1. Diagramme de cas d'utilisation

La figure ci-dessous illustre le diagramme de cas d'utilisation de notre système :



## Hiérarchisation des cas d'utilisation

- **Priorité 1** : Génération de speech à partir d'un prompt (fonctionnalité clé initiale).
  - **Priorité 2** : Speech-to-speech (étendre les capacités vocales du système).
  - **Priorité 3** : Génération de lipsync (ajout de composantes visuelles).
  - **Priorité 4** : Génération de musique (extension créative).
- 

## 2. Description des cas d'utilisation

### Cas d'utilisation 1 : Génération de speech à partir d'un prompt

- **Description** : L'utilisateur entre un texte ("prompt") dans l'interface, et le système génère une voix synthétique correspondant au texte fourni.
- **Étapes clés** :
  1. L'utilisateur soumet un texte (prompt).
  2. Le système traite le texte pour produire des phonèmes.
  3. Une voix synthétique est générée et envoyée.

La figure suivante décrit le fonctionnement du Cas d'utilisation 1 :

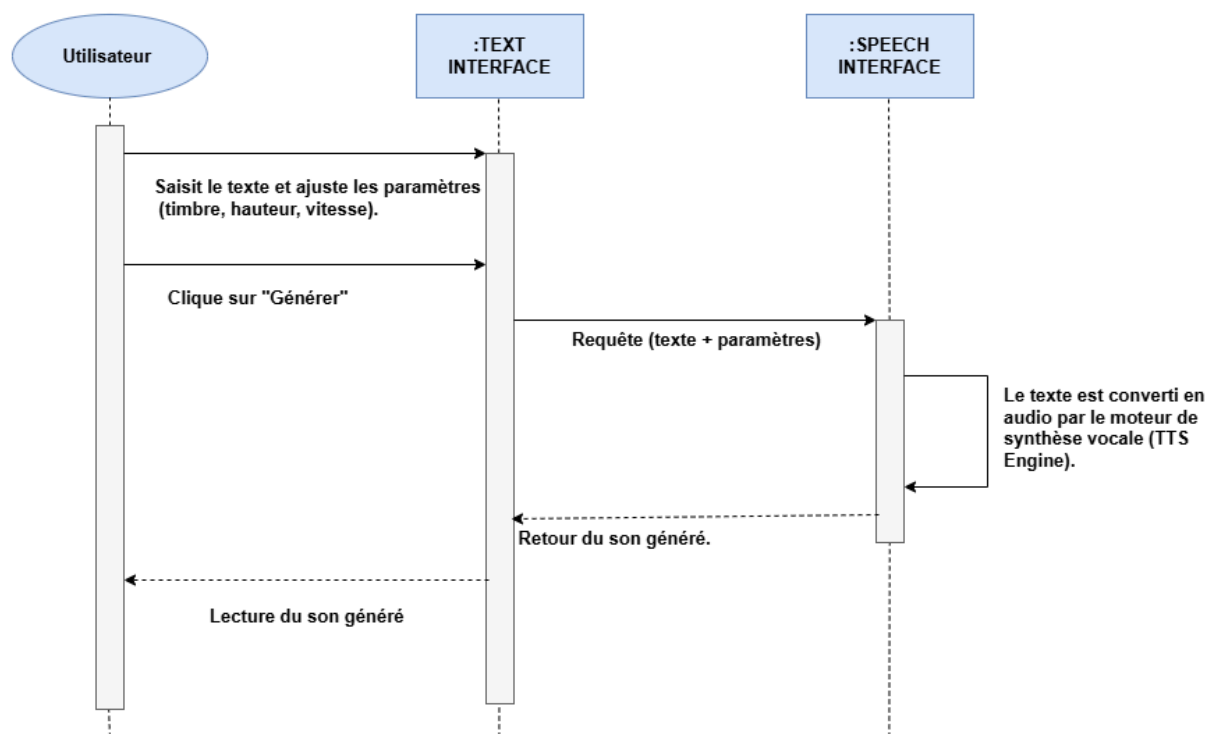


Diagramme de séquence : Génération de speech à partir d'un prompt

## Cas d'utilisation 2 : Speech-to-speech

- **Description** : Un utilisateur parle, et le système génère un discours transformé, en modifiant la voix.
- **Étapes clés** :
  1. L'utilisateur enregistre ou charge un fichier audio.
  2. Le système analyse l'audio pour en extraire des caractéristiques.
  3. Une voix transformée est renvoyée.

La figure suivante décrit le fonctionnement du Cas d'utilisation 2 :

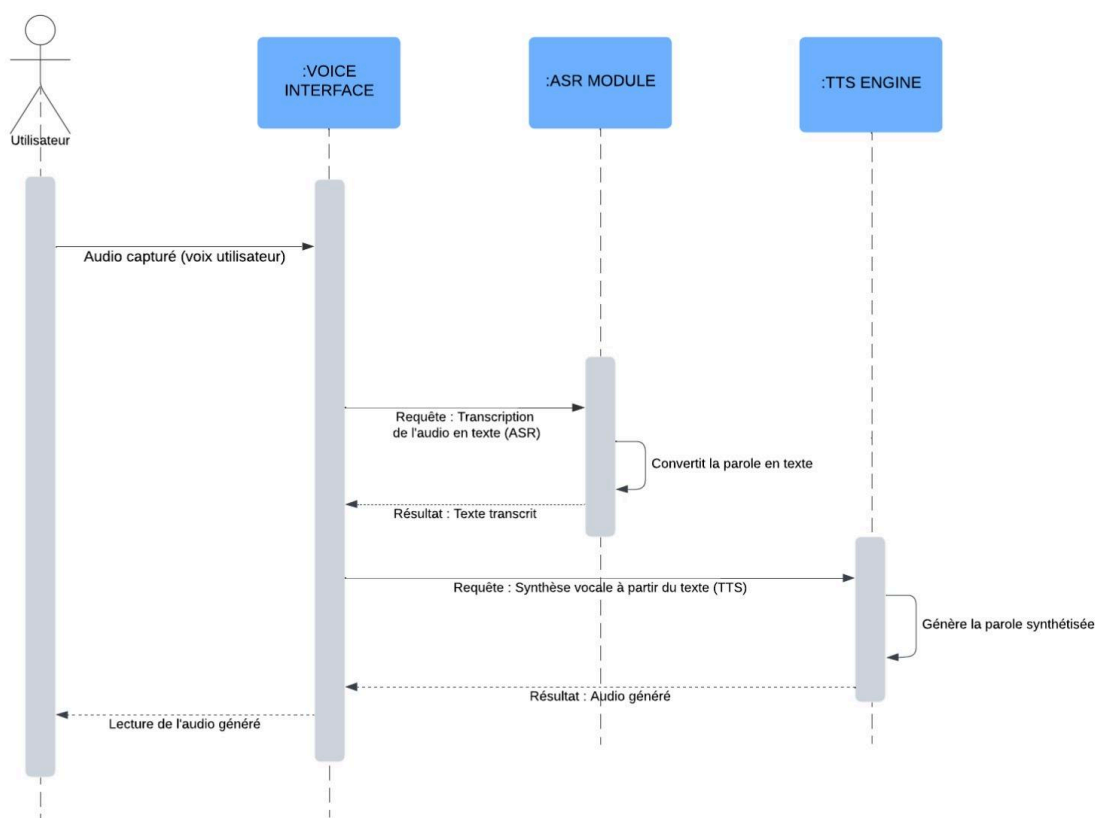
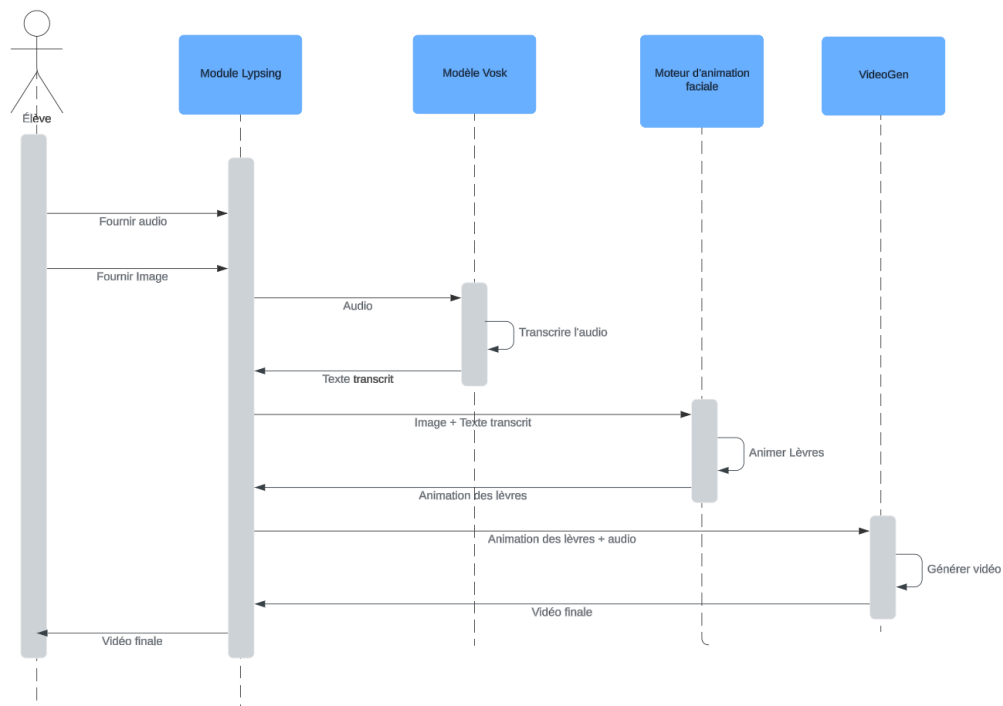


Diagramme de séquence : Speech-to-speech

## Cas d'utilisation 3 : Lipsync à partir d'un audio et d'une image

- **Description** : L'utilisateur fournit un audio et une image. Le système génère une animation faciale synchronisée avec l'audio.
- **Étapes clés** :
  1. L'utilisateur charge un audio et une image.
  2. Le système synchronise les phonèmes de l'audio avec l'image.
  3. Une animation est produite et envoyée à l'utilisateur.

La figure suivante décrit le fonctionnement du Cas d'utilisation 3 :



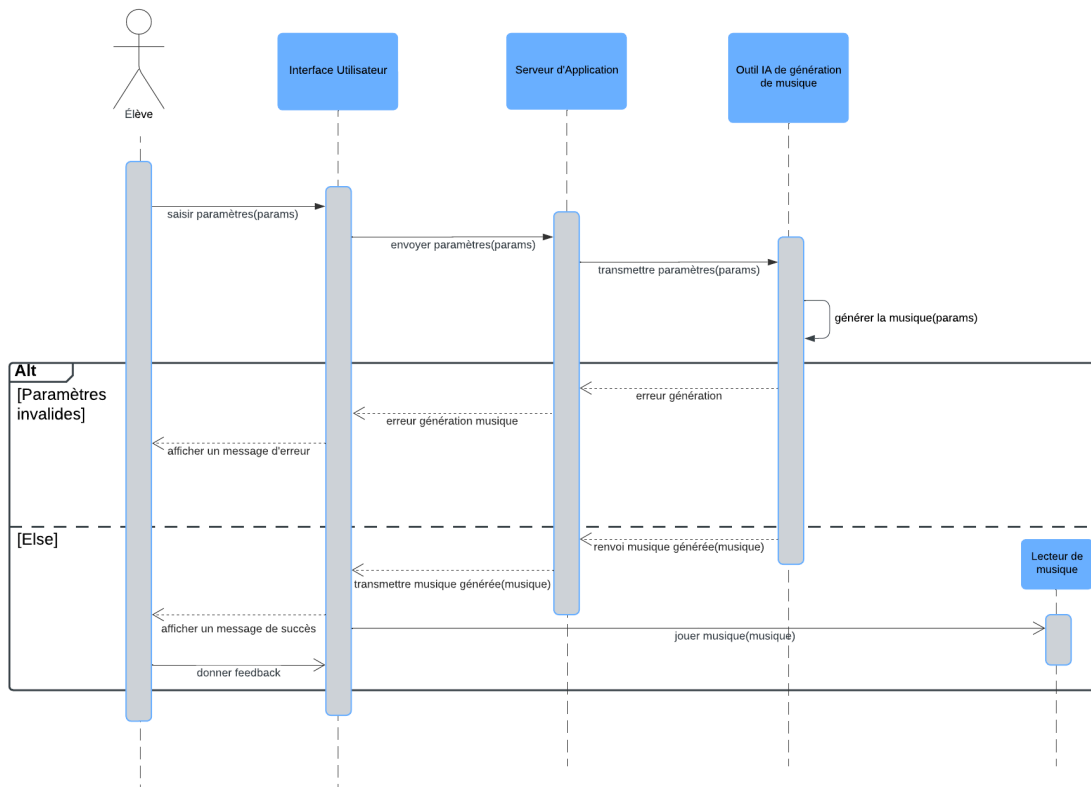
**Diagramme de séquence : Lipsync à partir d'un audio et d'une image**

#### Cas d'utilisation 4 : Génération de musique

- **Description** : Le système génère une composition musicale à partir de paramètres ou d'un contexte donné (style, ton, durée).
- **Étapes clés** :
  1. L'utilisateur définit les paramètres de style musical.
  2. Le système utilise un modèle de composition pour générer un fichier musical.
  3. L'utilisateur écoute la musique et donne son feedback.

La figure suivante décrit le fonctionnement du Cas d'utilisation 4 :

## Livrable : Sprint1



**Diagramme de séquence : Génération de musique**

## Conclusion

Ce livrable pose les bases fonctionnelles du système en définissant ses cas d'utilisation et leurs interactions internes. Il servira de référence pour orienter les développements futurs et garantir la cohérence des fonctionnalités.