

F I 3 E

UNIVERSITÉ DU QUÉBEC À RIMOUSKI

DÉVELOPPEMENT D'UN AVATAR INTERPRÈTE EN LANGUE FRANÇAISE PARLÉE COMPLÉTÉE (LfPC)

Cédric Fortin, sous la direction Mohamed Tarik Moutacalli

Département de mathématiques, informatique et génie

SOMMAIRE

Introduction	1
Problématique et objectifs	2
Méthodologie	3
Prototype	4
Conclusion	5
Références	6

INTRODUCTION

Ce projet porte sur :

- L'accès à l'information pour la communauté sourde et malentendante
- La Langue française Parlée Complétée (LfPC)

Questionnement initial

- À l'aide des technologies, comment répondre aux besoins de cette communauté



Crédit : Image de freepik

Pourquoi la LfPC?

- Sa simplicité, contrairement à une langue des signes. Elle consiste en des positions de mains autour du visage, en aditions aux mouvements de lèvres de l'interlocuteur
- Les interprètes LfPC sont plus rares au Québec
- Permet l'apprentissage de la langue orale

PROBLÉMATIQUE ET OBJECTIFS



Crédit : Image de freepik

TECHNOLOGIES

Unreal Engine

MetaHuman

C++ pour Unreal Engine

Intelligence artificielle de
conversion de paroles en
phonèmes

Comment concevoir **un système automatisé et efficace d'interprétation LfPC** reposant sur un système d'intelligence artificielle de traitement de la parole et de conversion phonétique pour l'interprétation en animation réaliste sur un **avatar 3D en temps réel**.

Objectifs

- Développer un avatar interprète 3D
- Solution en temps réel, conversion des paroles en animations
- Solution portative
- Évaluer les performances de l'interprète virtuel en comparaison à l'existant

01. REVUE DE LITTÉRATURE

- Étude de l'existant
 - Peu de recherches récentes
 - Très peu de mentions de la création complète de l'avatar

02. CRÉATION DE L'APPLICATION

- a. Création des animations
- b. Création d'un avatar MetaHuman
- c. Création d'un algorithme de conversion de phonèmes en unité LfPC
- d. Ajout d'une logique d'exécution des unités LfPC en animations sur l'avatar
- e. Intégration d'une intelligence artificielle de conversion de paroles en phonèmes

02. VALIDATION DE L'AVATAR

- Validation à l'aide d'un jeu de données LfPC. Évaluation du pourcentage de réussite et comparaison avec l'existant

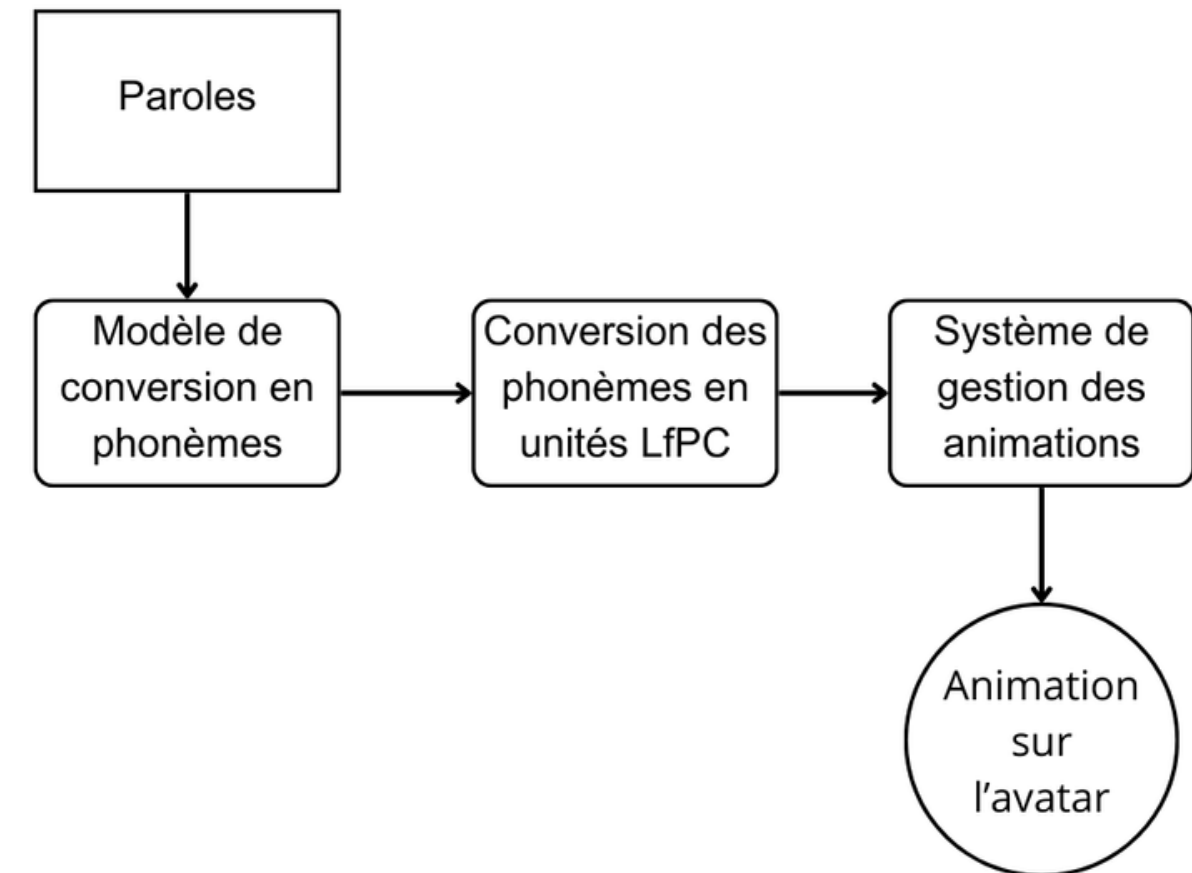


Schéma de conversion de paroles en animations

Qu'est-ce qu'une unité LfPC?

- Représente une syllabe
- Représente les mouvements labiaux à effectuer ainsi que la configuration manuelle et sa position relative au visage

PROTOTYPE

État

- Le prototype permet présentement d'effectuer la conversion de phonèmes en unité LfPC et d'en effectuer l'animation
- L'avatar est en temps réel

Enjeux courants

- Validation du fonctionnement sur mobile
- Validation auprès de personnes utilisant la LfPC



L'avatar effectuant l'unité LfPC «te»

CONCLUSION

POURSUITE DU PROJET

- Évaluation et amélioration du prototype
- Intégration de l'intelligence artificielle de conversion de paroles en phonèmes
- Validation des enjeux
- Évaluation de l'avatar



Crédit : Image de freepik

PARTENAIRES



SIVET

SERVICE D'INTERPRÉTATION
VISUELLE ET TACTILE

05

RÉFÉRENCES

- Attina, V., Beautemps, D., Cathiard, M.-A., & Odisio, M. (2004). A pilot study of temporal organization in Cued Speech production of French syllables : Rules for a Cued Speech synthesizer. *Speech Communication*, 44(1), 197-214. <https://doi.org/10.1016/j.specom.2004.10.013>
- Duchnowski, P., Braida, L. D., Lum, D., Sexton, M., Krause, J., & Banthia, S. (1998). Automatic Generation Of Cued Speech For The Deaf : Status And Outlook.
- Jovanova, B., Arsov, I., Preda, M., & Preteux, F. (2010). Online animation system for practicing cued speech. *International Journal of Image and Graphics*, 10(04), 497-512. <https://doi.org/10.1142/S0219467810003925>
- Lei, W., Liu, L., & Wang, J. (2024). Bridge to Non-Barrier Communication : Gloss-Prompted Fine-grained Cued Speech Gesture Generation with Diffusion Model (No. arXiv:2404.19277). arXiv. <https://doi.org/10.48550/arXiv.2404.19277>
- Malakar, M., & Keskar, R. B. (2021). Progress of machine learning based automatic phoneme recognition and its prospect. *Speech Communication*, 135, 37-53. <https://doi.org/10.1016/j.specom.2021.09.006>
- R. O. Cornett, “Cued speech,” *American annals of the deaf*, vol. 112, no. 1, pp. 3–13, 1967.
- Sankar, S. (2024). Automatic recognition and generation of French Cued Speech using deep learning.
- Sankar, S., Beautemps, D., Elisei, F., Perrotin, O., & Hueber, T. (2023). Investigating the dynamics of hand and lips in French Cued Speech using attention mechanisms and CTC-based decoding (No. arXiv:2306.08290). arXiv. <https://doi.org/10.48550/arXiv.2306.08290>
- Sankar, S., Lenglet, M., Bailly, G., Beautemps, D., & Hueber, T. (2025). Cued Speech Generation Leveraging a Pre-trained Audiovisual Text-to-Speech Model (No. arXiv:2501.04799). arXiv. <https://doi.org/10.48550/arXiv.2501.04799>
- Zhu, L., Liu, X., Liu, X., Qian, R., Liu, Z., & Yu, L. (2023). Taming Diffusion Models for Audio-Driven Co-Speech Gesture Generation (No. arXiv:2303.09119). arXiv. <https://doi.org/10.48550/arXiv.2303.09119>