

Projet d'Ingénieur ou de Master

Effets de l'incongruence de l'animation d'une tête parlante audiovisuelle 3D sur la production de la parole

Cadre: Depuis longtemps, le Département Parole & Cognition de GIPSA-lab a développé une tête parlante virtuelle 3D qui peut produire de la *parole augmentée*, i.e. de la parole complétée par l'affichage animé des articulateurs visibles et non visibles (cf. figure à droite). Cette tête parlante peut être contrôlée à partir de données articulatoires mesurées sur le locuteur à l'aide d'un articulographe électromagnétique.



Tête parlante sur laquelle on voit les différents articulateurs (langue et lèvres en rouge, mâchoire en cyan, palais dur et voile du palais en jaune).

Problème: Dans la ligne de ces études, la question que nous considérons dans ce projet est de savoir si – et comment – l'animation incongruente d'une tête parlante audiovisuelle, i.e. l'affichage d'articulations ne correspondant pas au son joué, peut influencer la production du locuteur.

Projet: Le projet consistera donc à

- Mettre en œuvre un dispositif expérimental grâce auquel le locuteur peut observer des stimuli audiovisuels joués par la tête parlante, en même temps qu'il/elle est monitoré(e) à l'aide d'un microphone pour le son, d'une sonde ultrasonore pour la langue, et d'une caméra vidéo pour les lèvres (cf. figure à droite).
- Définir un corpus d'articulations approprié pour mettre en évidence les modifications articulatoires susceptibles d'induire des changements dans la production du locuteur. Comme nous sommes particulièrement intéressés par la vision de la langue, puisqu'elle n'est pas habituellement visible, nous proposons d'explorer deux traits présents dans le système vocalique français: (1) le trait haut / bas qui varie dans /ieea/ pour les voyelles avant, et dans /uoɔ(ɒ)/ pour les voyelles arrière; (2) le trait avant / arrière qui varie dans les trois séries de voyelles ayant approximativement les mêmes hauteurs: voyelles hautes /yu/, moyennes /øo/, et basses /æɔ/. Voir illustrations en annexe.
- Construire les stimuli audiovisuels incongruents par déformation des stimuli originaux de manière cohérente (par exemple afficher une articulation de /ɛ/ lorsqu'un son de /a/ est entendu, jouant sur la différence d'aperture mandibulaire). Faire varier également le rapport Signal / Bruit et le type d'affichage (avec ou sans langue ...)
- Faire passer les tests audiovisuels perceptifs aux sujets.
- Analyser la parole produite par les sujets aux niveaux acoustique (formants) et articulatoire (formes de langue et de lèvres), afin de déterminer si la vision de stimuli incongruents peut influencer les productions des sujets.

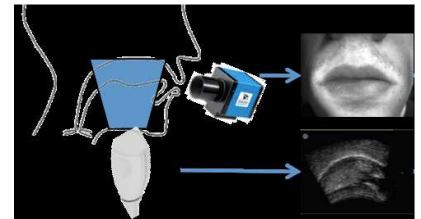


Illustration du dispositif de monitoring de la production de parole

Ressources disponibles: tête parlante virtuelle 3D, corpus de données d'animation, dispositif d'enregistrement audio/ultrason/vidéo et logiciel de contrôle, outils pour les tests de perception audiovisuelle

Compétences appréciées: Capacités d'intégration de logiciels sous linux. Connaissance de Matlab. Notions en production de parole et / ou en sciences cognitives / perception.

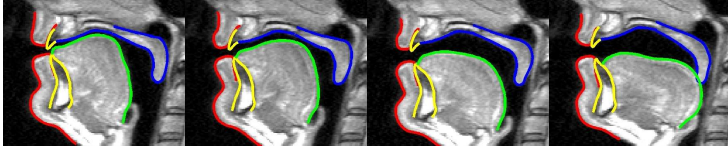
Mots-clé: tête parlante audiovisuelle, perception audiovisuelle de la parole perturbée, mesure des mouvements articulatoires, interfaces humain – machine.

Encadrement: Le projet se déroulera sur une durée de 4 ou 5 mois au DPC en collaboration avec Pierre Badin, Thomas Hueber, Gérard Bailly, et Frédéric Elisei.

Conditions: Indemnité mensuelle de 400€ garantie. Extension possible du stage avec rémunération supérieure.

Annexe: Illustration des traits vocaliques en français.

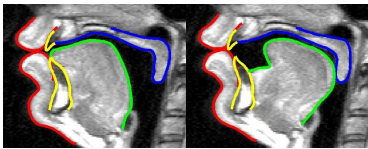
Variation du trait haut / bas pour les voyelles avant /ieɛa/



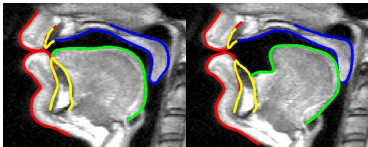
Variation du trait haut / bas pour les voyelles arrière /uoɔ/



Variation du trait avant / arrière pour les voyelles hautes /y/



Variation du trait avant / arrière pour les voyelles hautes /ø/



Variation du trait avant / arrière pour les voyelles basses /œ/

