



Convergence phonologique et phonétique en conversation face-à-face

Contexte

Ce travail s'inscrit dans le cadre des travaux de l'équipe MPACIF sur la modélisation des boucles de perception-action multimodale en situation d'interaction face-à-face.

Des études récentes montrent que les boucles de perception-action peuvent amener les interlocuteurs non seulement à faire converger leurs productions linguistiques vers un répertoire de formes communes – comme le lexique ou la syntaxe – (Gutierrez-Clellen and Heinrichs-Ramos 1993) mais aussi à adapter mutuellement leurs productions vocales – formes phonologiques telles que le choix des variantes de prononciation, mais aussi caractéristiques phonétiques telles que la vitesse d'articulation, les formants vocaliques, la distribution de la fréquence fondamentale (cf. Figure 1), ou plus prosaïquement le niveau sonore (Ward and Nakagawa 2004; Komatsu and Morikawa 2005; Kousidis, Dorran et al. 2008). Ces convergences ont été observées aussi bien entre interlocuteurs humains que dans des dialogues homme-machine : les traces collectées lors de dialogues d'agents conversationnels avec des usagers montrent que ces derniers adaptent très rapidement leur lexique et la syntaxe de leurs interventions au niveau de langue utilisée par l'agent (Bell, Gustafson et al. 2003).

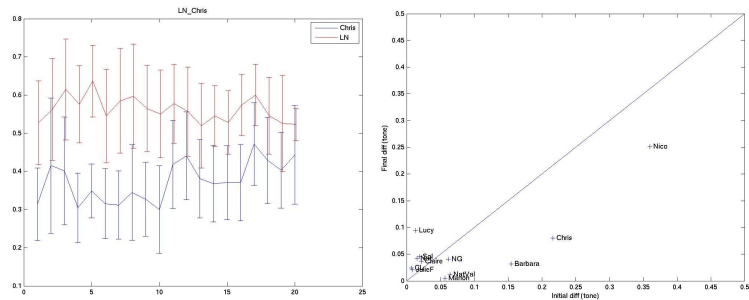
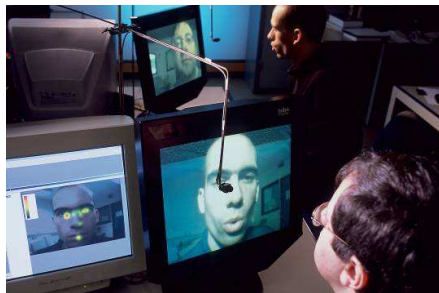


Figure 1 : Convergence des registres de fréquence fondamentale (F0) d'interlocuteurs en face-à-face. A gauche : dispositif d'enregistrement d'une conversation face-à-face médiatisée. Au centre, exemple d'évolution des registres (moyennes et écart-types des valeurs de F0 exprimées en tons) de deux interlocuteurs (Chris et Héléne). Chaque abscisse correspond à un tour de parole. A droite, les valeurs des écarts initiaux et finaux pour des faces-à-faces de la locutrice Héléne avec 11 différents interlocuteurs. Notez que la plupart des interactions produisent des rapprochements de registres comme pour Chris.

Sujet

L. Martin (M1, ICPS/UPMF) a effectué en 2009 une étude préliminaire sur la convergence phonétique dans des jeux de dominos verbaux (bateau, taudis, diffus, fusil, etc.) où chaque interlocuteur commence un mot avec la dernière syllabe du mot prononcé par son interlocuteur. L'étude proposée vise à poursuivre ce travail en se concentrant sur la caractérisation automatique de la convergence par des techniques de reconnaissance de parole déjà disponibles au laboratoire.

Le travail consistera à enregistrer des dialogues en situation d'interaction face-à-face et à caractériser l'évolution des caractéristiques phonétiques (fréquence fondamentale, débit, formants des voyelles, etc.) des productions vocales des interlocuteurs au cours du temps afin de caractériser les convergences éventuelles des représentations phonétiques en fonction des caractéristiques individuelles de chaque interlocuteur (sexe, origine géolinguistique, etc.).

Environnement de travail

Matlab, Praat

Collaborations

Ce travail s'effectue dans le cadre du PPF « Interactions Multimodales » et du projet ANR « Amorces » sur la communication face-à-face pour le travail collaboratif homme-robot. Le candidat s'insèrera dans une équipe de six personnes.

Responsables

Gérard BAILLY

GIPSA-lab

04 76 57 47 11

Gerard.Bailly@gipsa-lab.inpg.fr

Références

Raidt, S., G. Bailly, et al. (2007). Gaze patterns during face-to-face interaction. IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology - Workshop on Communication between Human and Artificial Agents (CHAA), Fremont, CA.

Bailly, G., F. Elisei, et al. (to appear). "Boucles de perception-action et interaction face-à-face." Revue Française de Linguistique Appliquée **XIII**.