

Coordination des rythmes respiratoires dans une tâche collaborative Caractérisation et modélisation

Contexte

La gestion fluide des tours de parole entre interlocuteurs suppose que chacun soit prêt à prendre la parole lorsqu'il lui semble bon et acceptable de contribuer à la conversation. Il est donc essentiel de pouvoir inspirer l'air à un rythme permettant aux poumons de fournir l'énergie aéroacoustique nécessaire à la prise de parole.

Des recherches ont montré que les interlocuteurs accommodent mutuellement de nombreuses caractéristiques non seulement du contenu linguistique (notamment choix de mots et de la syntaxe) et des choix phonologiques (notamment de l'adoption d'un accent commun ou de rapprochement des variantes de prononciation utilisées) mais aussi au niveau des signaux échangés (débit de parole, volume sonore, etc). Capella *et al* (1979; 1980; 1981) ont notamment montré que la structure des pauses et des tours de parole d'un locuteur pouvait être relativement bien prédite par celle de l'interlocuteur et que ce système couplé était bien caractéristique de la dyade.

Cette accommodation mutuelle s'appuie à la fois sur des informations a priori (connaissance mutuelle de chaque système de croyance, des habitudes gestuelles de l'autre, etc) et du traitement en ligne des signaux. La plupart des modèles cognitifs considèrent que perception, action et cognition sont des systèmes couplés et que les processus mentaux co-évoluent avec leurs manifestations sensibles. Shockley et al (2007) ont ainsi montré que la prononciation simultanée de mots ayant la même structure accentuelle par deux locuteurs induisait des mouvements posturaux plus synchrones lorsque les locuteurs prononcent des mots ne partageant pas cette propriété et ceci d'autant plus s'ils pouvaient s'écouter.

Les signaux échangés participent donc à la régulation de nos activités cognitives, perceptives et motrices.



Sujet

Le but de ce travail est d'étudier et de modéliser la coordination conjointe des tours de parole et des rythmes respiratoires (estimés à partir des mouvements de l'abdomen et du thorax) à l'aide de modèles statistiques. Une première piste est l'analyse de récurrences (voir par ex. Webber Jr. and Zbilut 1994) qui permet de rechercher des patrons identiques ou liés dans des séries temporelles. Une deuxième piste (notamment explorée par Capella *et al*) est de construire des automates probabilistes type chaînes de Markov cachées (Hidden Markov Models) décrivant la succession des divers états d'une conversation (prise de parole, silence avec ou sans inhalation, etc.) et les comportements associés. Ces HMM ont la capacité de reconnaître des états à partir de comportements mais aussi de générer des comportements optimaux à partir d'une suite d'états (Zen *et al.* 2007): on peut donc à la fois observer les comportements des autres, les reconnaître et générer ses propres comportements à partir de l'observation d'interactions. Cet apprentissage par l'observation d'actions conduites par d'autres - permettant ainsi de se substituer à autrui - a des bases cognitives fortes (notamment de part l'existence de neurones-miroirs démontrée par Rizzolatti and Craighero 2004)¹.

Protocole expérimental

L'objectif est d'étudier comment un sujet négocie son rythme respiratoire et s'adapte aux stratégies de communication de son interlocuteur en essayant de minimiser la variabilité du contenu linguistique échangé lors de conversations libres ou de jeux collaboratifs tels que la map-task (Anderson *et al.* 1991). Nous avons pour ceci introduit un jeu verbal appelé dominos verbaux (Lelong and Bailly 2011) qui consiste à reprendre tout ou partie des éléments sonores prononcée par le partenaire en piochant dans un ensemble de mots proposés. L'avantage de ce scénario est d'avoir des tours de parole rapides (n'impliquant donc pas forcément une prise de souffle à chaque tour de parole) et de maîtriser diverses caractéristiques des mots échangés (vocabulaire facile vs. difficile, sons employés, etc).

Travail spécifique

Le travail consistera donc à définir et tester le protocole expérimental. Il sera alors appliqué à deux sujets cibles interagissant avec cinq interlocuteurs différents. Les voix et les mouvements respiratoires des deux interlocuteurs de chaque paire seront enregistrés de manière synchrone de manière à pouvoir rechercher des signatures caractérisant le degré de prédictibilité mutuelle des comportements des interlocuteurs.

Thématiques abordées dans le stage

- Capture de mouvement
- Modélisation statistique

Compétences requises

- Notions de statistique, maîtrise de Matlab

Contacts

Gérard Bailly

GIPSA-lab

04 76 57 47 11

Gerard.Bailly@gipsa-lab.grenoble-inp.fr

Indemnités de stage

Ce stage fait l'objet d'une indemnité fixée annuellement par le conseil de laboratoire, de l'ordre de 400€ mensuels.

¹ Voir http://www.scholarpedia.org/article/Mirror_neurons

Références

- Anderson, A., M. Bader, E. Bard, E. Boyle, G. M. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. S. Thompson and R. Weinert (1991). "The HCRC Map Task Corpus." Language and Speech 34: 351-366.
- Capella, J. (1979). "Talk and silence sequences in informal conversations. I." Human Communication Research 6: 3-17.
- Capella, J. (1980). "Talk and silence sequences in informal conversations. II." Human Communication Research 6: 130-145.
- Capella, J. and S. Planap (1981). "Talk and silence sequences in informal conversations. III: Interspeaker influence." Human Communication Research 7: 117-132.
- Lelong, A. and G. Bailly (2011). Study of the phenomenon of phonetic convergence thanks to speech dominoes Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issue. A. Esposito, A. Vinciarelli, K. Vicsi, C. Pelachaud and A. Nijholt. Berlin, Springer Verlag: 280-293.
- Rizzolatti, G. and L. Craighero (2004). "The mirror-neuron system." Annual Review of Neuroscience 27: 169-192.
- Shockley, K., A. Baker, M. Richardson and C. Fowler (2007). "Articulatory constraints on interpersonal postural coordination." Journal of Experimental Psychology: Human Perception and Performance 33: 201-208.
- Webber Jr., C. L. and J. P. Zbilut (1994). "Dynamical assessment of physiological systems and states using recurrence plot strategies." Journal of Applied Physiology 76(2): 965-973.
- Zen, H., T. Nose, J. Yamagishi, S. Sako, T. Masuko, A. Black and K. Tokuda (2007). The HMM-based speech synthesis system version 2.0. Speech Synthesis Workshop. Bonn, Germany, pp. 294-299.