



## Master 2 Sciences Cognitives – 2018-2019

### Étude et simulation informatique d'un modèle de perception et production de parole continue Aspects hiérarchiques

La question des relations entre mécanismes de perception et d'action s'est installée au cœur des réflexions et des travaux à la fois en sciences cognitives, avec la découverte des neurones miroir et les avancées en neuroimagerie, et dans le domaine de la robotique autour des enjeux d'apprentissage et d'adaptabilité.

Nous étudions cette question dans le cadre de la communication parlée, en développant des modèles computationnels d'agents capables à la fois de percevoir et de produire des stimuli de parole simples (voyelles, consonnes, syllabes) et en analysant sur ces modèles en quoi les capacités de production peuvent interagir avec les capacités perceptives, les contraindre, les enrichir ou les compléter.

Ceci a conduit à la mise au point **d'un modèle bayésien d'agent communicant**, que nous avons baptisé COSMO, pour *Communicating about Objects using SensoriMotor Operations*. Ce modèle probabiliste provient de la modélisation de la situation de communication, dans laquelle deux agents veulent communiquer à propos d'un objet de l'environnement. Le modèle COSMO est basé sur l'hypothèse fondamentale que cette situation de communication peut être internalisée et émulée dans le cerveau de chaque agent, qui est alors en mesure d'agir aussi bien en tant que locuteur qu'auditeur.

Dans ce contexte, nous avons développé plusieurs instances du modèle COSMO, traitant de questions comme l'évolution des systèmes phonologiques, de la perception de parole dans le bruit, des propriétés fonctionnelles complémentaires de la voie auditive et de la voie motrice, de l'acquisition de répertoires syllabiques et phonémiques, ou encore de l'apparition d'idiosyncrasies dans le développement.

Cependant, **dans nos travaux précédents, les unités de parole, c'est-à-dire les unités distinctives représentées par le modèle, sont définies a priori**. Ainsi, par exemple, dans le modèle COSMO SylPhon, nous comparons les apprentissages des répertoires syllabiques et phonémiques, en supposant préalablement leur existence. En réalité, le système cognitif est certainement structuré autour d'unités de représentation qui résultent des propriétés des interactions du système : à la fois des interactions avec l'environnement (par exemple, les régularités du langage à apprendre) mais aussi des interactions entre les sous-parties du système (par exemple, les stabilités statistiques des relations sensorimotrices).

#### *Objectifs*

Notre objectif principal est donc **d'étendre l'architecture du modèle COSMO pour y représenter les aspects hiérarchiques associés à la structuration en unités de parole**. Pour cela, nous adapterons des outils de modélisation développés dans d'autres contextes.

En effet, dans le modèle des cartes bayésiennes (de navigation de robots mobiles), des hiérarchies de cartes sont définies, dans lesquelles les unités d'un niveau de représentation traduisent des propriétés comportementales des niveaux sous-jacents. Par exemple, une carte abstraite « observe » d'autres cartes, et les considère chacune comme un lieu abstrait, connectés par leurs comportements produits. Cela permet une structuration des propriétés temporelles par les interactions sensorimotrices des modèles sous-jacents, qui peuvent être complexes. Dans le modèle BAP (de lecture et d'écriture de lettres manuscrites), nous avons défini un modèle où des séquences temporelles complexes pouvaient être représentées par des ensembles discrets de points de contrôles. Cela permettait une représentation compacte de trajectoires complexes, basée également sur des propriétés saillantes au niveau sensoriel et moteur.

L'objectif est donc une **version du modèle COSMO reposant sur les formalismes des cartes hiérarchiques sensorimotrices et des modèles de points de contrôle, afin d'étudier la structure des représentations de parole**. Nous étudierons notamment les interactions entre les niveaux de représentations,



et leur apprentissage (par exemple, la manière dont l'exploration d'un niveau est contraint par les comportements des niveaux sous-jacents).

### *Travail proposé*

Pour réaliser l'objectif, l'étudiant devra **implémenter un modèle probabiliste, reprenant et combinant des éléments existants du modèle COSMO, d'une part, et des modèles de cartes sensorimotrices et du modèle BAP, d'autre part**. On utilisera ensuite ce modèle dans **des simulations illustratives, permettant de caractériser les propriétés mathématiques du modèle, afin de le comparer aux connaissances expérimentales du domaine**. Une revue de la littérature des modèles computationnels de perception de parole et des unités de parole qu'ils représentent sera attendue.

### *Contexte pratique*

Le travail sera réalisé au sein d'une **équipe pluridisciplinaire alliant modélisateurs** – spécialistes de la programmation bayésienne – et **spécialistes de la parole et de la cognition**. Il sera encadré conjointement par Jean-Luc Schwartz, chercheur à GIPSA-Lab (parole) et Julien Diard, chercheur au LPNC (modélisation bayésienne).

**Une poursuite en thèse est envisagée**, dans l'objectif d'étudier des modèles COSMO plus complets, impliquant des **structures temporelles emboîtées** (syllabes, mots, phrases), **des représentations multimodales** (auditives, somatosensorielles, motrices, voire, visuelles), et permettant de simuler des tâches de perception et de production en ligne, mais aussi **l'apprentissage des répertoires syllabiques et lexicaux** sur la base d'exemples, dans un contexte d'interaction avec un agent maître. Cette thèse comportera des **enjeux théoriques** sur les processus d'interaction et d'apprentissage et leurs corrélats neuro-anatomiques, mais aussi des **enjeux technologiques** pour le développement d'agents communicants autonomes, dotés de capacités d'apprentissage et d'interaction vocale.

### *Compétences demandées*

Le candidat devra avoir soit une formation en sciences cognitives et un intérêt pour la modélisation, soit une formation en programmation et simulation informatique, et une affinité pour la modélisation mathématique en sciences cognitives.

### *Contacts*

Jean-Luc Schwartz (GIPSA-Lab)  
[jean-luc.schwartz@gipsa-lab.grenoble-inp.fr](mailto:jean-luc.schwartz@gipsa-lab.grenoble-inp.fr) 04 76 57 47 12

Julien Diard (LPNC)  
[julien.diard@univ-grenoble-alpes.fr](mailto:julien.diard@univ-grenoble-alpes.fr) 04 76 82 78 07  
<http://diard.wordpress.com>

### *Financement*

Indemnités de stage assurées par le projet ERC « Speech Unit(e)s »