

EN FRANÇAIS (6)

IN ENGLISH (6)

THURSDAYS 10:30 am – noon EST at: <https://ugam.zoom.us/j/96780028011>

(Video recordings will be accessible about a week after each seminar at: <http://dic.ugam.ca/en/events.html>)

### 17-Sep L'apprentissage des catégories visuelles par les humains et les réseaux neuronaux « deep learning »

[Christian THÉRIAULT](#) (UQÀM)

Les humains ont trois façons d'apprendre les nouvelles catégories visuelles : (1) l'apprentissage non supervisé (juste l'exposition passive aux membres et non membres de la catégorie), (2) l'apprentissage supervisé (ou renforcé) (essai-et-erreur actif, suivi de correction), et (3) l'apprentissage par l'explication verbale. Quand les humains apprennent des catégories difficiles, il leur arrive un changement de perception qui s'appelle la perception catégorielle : Les membres de différentes catégories deviennent plus différents (la séparation inter-catégorielle) et les membres de la même catégorie deviennent plus semblables (la compression intra-catégorielle). Pourquoi? Les réseaux neuronaux artificiels apprennent aussi les catégories visuelles (numérisées). Les changements dans leurs représentations internes lors de l'apprentissage supervisé donnent peut-être une explication

### 24-Sep Grounding word meanings in perceptual experience: A computer-vision approach

[Eritz GUENTHER](#) (U. Tübingen)

For about two decades, the fields of cognitive science and psychology have employed distributional semantic models such as LSA (Latent Semantic Analysis) as powerful computational models of semantic representation. These language-based models build meaning representations from the distributional patterns of words from large collections of natural text, and thus from approximations of the actual input experienced by humans. Until recently, this stood in contrast to computational models incorporating sensorimotor information, which was often approximated via participant ratings on sensorimotor features – the outcome of human experience. However, advancements in the field of computer vision allow us to model vision-based representations directly from visual input as approximated by large datasets of images. In this talk, I will present recent studies in which we employ these vision-based representations to investigate automatic (visual) sensorimotor activation (i) during word processing and (ii) in conceptual combination. In addition, I will present how a systematic mapping can be established between language-based and the vision-based representations, thus implementing a possible mechanism for the visual grounding of non-experienced concepts.

### 01-Oct Théorie de la simplicité et Cognition

[Jean-Louis DESSALLES](#) (Télécom Paris Tech)

Plusieurs auteurs ont noté que le fonctionnement cognitif humain est sensible à la complexité des structures (au sens de la taille de description minimale). La Théorie de la Simplicité est construite sur l'observation que de nombreux phénomènes cognitifs liés à la surprise, à l'intérêt narratif, à la pertinence, à l'intensité émotionnelle ou encore au jugement moral s'analysent sur la base d'une différence entre la complexité attendue et la complexité observée. Nous nous demanderons pourquoi cette capacité à repérer l'inattendu par la baisse de complexité est apparue au cours de l'évolution.

## 08-Oct Towards Understanding Understanding: Dialogue, Robots, and Meaning

Casey KENNINGTON (Boise State U)

Hallmarks of intelligence include the ability to acquire, represent, understand, and produce natural language. Although recent efforts in data-driven, machine learning, and deep learning methods have advanced natural language processing applications, important challenges remain. In my talk, I'll give an overview of general trends in understanding language on machines, what we can learn from children who acquire language seemingly with minimal effort, and what that means for future research. I will then explain my own research on grounding language into different physical modalities, what role emotion could play, and the potential importance of embodiment.

## 22-Oct Le Web sémantique en Ingénierie des environnements numériques d'apprentissage

Gilbert PAQUETTE (TÉLUQ)

La méthode d'ingénierie des systèmes d'apprentissage (MISA) a innové il y a une quinzaine d'année en intégrant la modélisation des connaissances au design pédagogique et aux principes du génie logiciel et cognitif. Elle est fondée sur quatre types de modèles : connaissances et compétences, scénario pédagogique, modèle médiatique et processus (logistique) de diffusion. Depuis, l'essor du Web sémantique a fourni un ensemble croissant de modèles sous forme d'ontologies pouvant être intégrés comme composantes d'une méthode d'ingénierie des environnements numériques d'apprentissage (ENA). Le but de ce séminaire est de présenter comment les typologies élaborées et intégrées à MISA peuvent être transformées en ontologies et utilisées pour la réalisation des ENA. La question des outils utilisant ces ontologies ne sera qu'effleurée, mais on en soulignera certains aspects de fonctionnement.

## 29-Oct Le robot cognitif et interactif: vers les robots assistants ou équipiers

Rachid ALAMI (LAAS CNRS, Toulouse)

Cet exposé aborde un ensemble de problèmes décisionnels essentiels auxquels est confronté le robot cognitif et interactif qui partage l'espace et la tâche avec un humain. Nous adoptons une approche constructive basée sur l'identification et la mise en œuvre de compétences collaboratives élémentaires inspirées des concepts de l'action conjointe. Le système est complet dans la mesure où il vise à mettre en œuvre un ensemble cohérent de capacités articulées de manière à ce que le contrôleur du robot puisse effectivement conduire, de manière flexible et acceptable par l'humain, une résolution de problèmes interactive et une réalisation de tâches collaboratives humain-robot. Ces capacités comprennent le raisonnement géométrique et l'évaluation de la situation reposant essentiellement sur la prise de perspective et l'exploitation

de la connaissance de chaque agent (humain et robot) dans un modèle cognitif distinct et une planification des tâches prenant en compte explicitement l'humain, l'estimation de son état mental, ses besoins et ses préférences.

Chatila, R., Renaudo, E., Andries, M., Chavez-Garcia, R. O., Luce-Vayrac, P., Gottstein, R., ... & Khamassi, M. (2018). Toward self-aware robots. *Frontiers in Robotics and AI*, 5, 88

## 05-Nov **Les auto-citations: caractéristique normale de l'accumulation des connaissances ou perversion de l'évaluation de la recherche?**

Vincent LARIVIÈRE (U. Montréal)

La science est une activité cumulative, où les connaissances passées servent de socle sur lequel s'appuient celles qui sont nouvellement produites. L'un des mécanismes par lequel cette filiation des connaissances se manifeste est l'acte de citer. Or, les citations sont également au cœur de l'évaluation de la recherche contemporaine, créant ainsi certains incitatifs pour les chercheurs à citer leurs propres travaux afin d'améliorer leur visibilité. Ainsi, l'autocitation est au cœur de la critique liée à l'usage des citations comme indicateurs de l'impact scientifique. Après avoir fait un bref survol des méthodes bibliométriques, nous examinerons, à partir d'un corpus de plusieurs dizaines de million d'articles, l'importance relative des auto-citations dans le paysage scientifique, leur relation avec l'âge et le genre des auteur.e.s, ainsi que leurs effets sur divers indicateurs d'évaluation de la recherche. Nous présenterons également les résultats d'une comparaison du contenu des textes cités à celui des textes citant, permettant ainsi de tester l'hypothèse selon laquelle les chercheur.e.s citent leurs propres travaux afin de gonfler leurs indicateurs d'impact. Référence générale sur la bibliométrie: <https://www.pum.umontreal.ca/catalogue/mesurer-la-science> (livre en libre accès)

## 12-Nov **Lifelong social learning in swarm robotics**

Nicolas BREDECHE (U. Sorbonne, Paris)

In swarm robotics, hundreds or more robots with limited computation and communication capabilities act together to accomplish a common task. Behavior programming is generally done by hand or optimized before deployment to address a specific task (e.g. shape formation) or to produce basic primitives (e.g. leader-following, flocking) that can then be combined by the user. However, it is not always possible to design beforehand the policies required to solve a task, as the environment may be unknown before deployment, or change through time. ---- In this talk, I will give an overview of our recent works in adaptive swarm robotics, where learning is performed in an online and distributed fashion. I will describe some recent evolutionary and social learning algorithms we have developed and implemented in both simulated and real robots. Originally loosely inspired by evolutionary biology and social learning, these algorithms face challenges that are usually not addressed in robotics, but familiar to biological collective systems. I will describe how the environment shapes learning in unexpected ways, as the swarm must address tasks in a self-sustainable fashion. I will also discuss how to achieve truly cooperative behaviours at the level of the collective, in particular concerning pursuing both individual benefits and social welfare

## 19-Nov **Grounded Language Learning in Virtual Environments**

Stephen CLARK (U. Cambridge)

Natural Language Processing is currently dominated by the application of text-based language models such as BERT and GPT. One feature of these models is that they rely entirely on the statistics of text, without making any connection to the world, which raises the interesting question of whether such models could ever properly “understand” the language. One way in which these models can be grounded is to connect them to images or videos, for example by conditioning the language models on visual input and using them for captioning. ---- In this talk I extend the grounding idea to a simulated virtual world: an environment which an agent can perceive, explore and interact with. More specifically, a neural-network-based agent is trained -- using distributed deep reinforcement learning -- to associate words and phrases with things that it learns to see and do in the virtual world. The world is 3D, built in Unity, and contains recognisable objects, including some from the ShapeNet repository of assets. ---- One of the difficulties in training such networks is that they have a tendency to overfit to their training data, so first we’ll demonstrate how the interactive, first-person perspective of an agent provides it with a particular inductive bias that helps it to generalize to out-of-distribution settings. Another difficulty is that training the agents typically requires a huge number of training examples, so we’ll show how meta-learning can be used to teach the agents to bind words to objects in a one-shot setting. Moreover, the agent is able to combine its knowledge of words obtained one-shot with its stable knowledge of word meanings learned over many episodes, providing a form of grounded language learning which is both “fast and slow”. [Joint work with Felix Hill.]

## 26-Nov **Grounding dictionary meaning**

[Lopes](#) **MARCOS** (U. Sao Paolo, Brazil)

Language acquisition is a long-term dynamic process which comprises not only learning language sounds, words and phrase structures, but also grounding those linguistic entities in perception, feeling, and reasoning. We know for a fact that throughout this process babies are incapable of passively repeating ready-made expressions heard from adults' mouths, as they are constantly adapting their communication skills according to the available knowledge and linguistic capabilities. But how exactly is this done? Focusing on lexical meaning acquisition, we assume that the whole process is based on two key operations: (1) grounding words of a basic vocabulary set and (2) enhancing this vocabulary later on combinatorially by learning new words through verbal definitions composed of already grounded words. This correlates with the way dictionaries are made: a small set of words can recursively generate the whole of the largest, most exhaustive dictionary. From this point of view, dictionaries of any language, of any complexity can be described as acyclic oriented graphs. We aim to show that word acquisition by children reveals precisely those same structures.

## 03-Dec **Constraining networks neurally to explain grounding**

[Friedemann](#) **PULVERMUELLER** (Freie U. Berlin)

Meaningful use of symbols requires grounding in action and perception through learning. The mechanisms of this sensorimotor grounding, however, are rarely specified in mechanistic terms; and mathematically precise formal models of the relevant learning processes are scarce. As the brain is the device that is critical for mechanistically supporting and indeed implementing grounding, modelling needs to take into account realistic neuronal processes in the human brain. This makes it desirable to use not just ‘neural’ networks that are vaguely similar to some aspects of real networks of neurons, but models implementing constraints imposed by neuronal structure and function, that is, biologically realistic learning and brain structure along with local and global structural connectivity and functional interaction.

After discussing brain constraints for cognitive modelling, the talk will focus on the biological implementation of grounding, in order to address the following questions: Why do the brains of humans -- but not those of their closest relative species -- allow for verbal working memory and learning of huge vocabularies of symbols? Why do different word and concept types seem to depend on different parts of the brain ('category-specific' semantic mechanisms)? Why are there 'semantic and conceptual hubs' in the brain where general semantic knowledge is stored -- and why would these brain areas be different from those areas where grounding information is present (i.e., the sensory and motor cortices)? And why should sensory deprivation shift language and conceptual processing toward 'grounding areas' -- for example, in blind people, toward the visual cortex? I will argue that brain-constrained modelling is necessary to answer (some of) these questions and, more generally, to explain the mechanisms of grounding.

Pulvermüller, F., Garagnani, M., & Wennekers, T. (2014). Thinking in circuits: Towards neurobiological explanation in cognitive neuroscience. *Biological Cybernetics*, 108(5), 573-593. doi: 10.1007/s00422-014-0603-9

Pulvermüller, F. (2018). Neural reuse of action perception circuits for language, concepts and communication. *Progress in Neurobiology*, 160, 1-44. doi: 10.1016/j.pneurobio.2017.07.001

Schomers, M. R., Garagnani, M., & Pulvermüller, F. (2017). Neurocomputational consequences of evolutionary connectivity changes in perisylvian language cortex. *Journal of Neuroscience*, 37(11), 3045-3055.

Tomasello, R., Wennekers, T., Garagnani, M., & Pulvermüller, F. (2019). Visual cortex recruitment during language processing in blind individuals is explained by Hebbian learning. *Scientific Reports*, 9(1), 1-16.

## 10-Dec **Les humanités du numérique : vers une épistémologie des données et du calcul**

[Bruno BACHIMONT](#) (U. Sorbonne, Paris)

Alors que l'on évoque souvent ce que les humanités peuvent retirer de la mobilisation des outils numériques, renouvelant leurs objets et leurs méthodes, on néglige ce qu'en retour la tradition intellectuelle des humanités peut apporter à notre compréhension du fait numérique. Notre intervention portera sur les humanités du numérique, comme réciproque des humanités numériques et réponses qu'elles apportent aux enjeux contemporains, en explicitant les chantiers épistémologiques qui en découlent. Il sera alors possible de situer le rôle du calcul dans la construction du sens et les conséquences culturelles qu'on peut discerner.