

## **Atelier de prospective Langage et Cognition : Tendances et Perspectives pour la recherche en sciences cognitives**

Organisé par les laboratoires MoDyCo, CLLE-Erss, Irit et Ircam.

En association avec le GDR « Langues, langage oral, cognition : acquisition et dysfonctionnements – nouvelles approches » pour le thème 3.

Coordination : Christophe Parisse (MoDyCo)

Contact : Christophe Parisse – cparisse@u-paris10.fr – Tél : 01 40 97 74 35

L'atelier envisagé ici, focalisé sur le langage en tant qu'activité cognitive élémentaire, est à concevoir comme un groupe de réflexion dans le domaine des technologies cognitives. Il s'agit de réfléchir sur des thèmes fédérateurs dans les domaines de la linguistique computationnelle et cognitive, au sein desquels on peut poser des axes de recherches spécifiques à un domaine mais également transversaux.

L'atelier visera à couvrir les approches les plus novatrices en linguistique cognitive et computationnelle.

Un tel atelier part d'un constat simple : non seulement les capacités informatiques de calcul et de stockage actuelles ou envisageables dans un futur proche incitent à être ambitieux dans les thèmes de recherche abordés, mais elles permettent également de disposer de nouveaux types de données langagières (textuelles, orales et multimodales) qui permettent d'enrichir le terrain des sciences cognitives pour à terme toujours mieux comprendre, modéliser et simuler le comportement de l'esprit humain.

Les thèmes abordés seront plus précisément :

1. Outils pour l'interaction optimale avec le web ou via le web
2. Modélisation et simulation de fonctions langagières complexes
3. Nouvelles technologies pour l'apprentissage des langues, pour le dépistage et pour la remédiation des troubles langagiers

L'atelier serait organisé sur trois journées dans un format de table-ronde avec des présentations suivies de discussions approfondies. Le lieu sera choisi de façon à inciter les participants à assister à l'ensemble de l'atelier. L'accent sera mis dans les discussions de l'atelier sur les relations entretenues entre les sciences du langage, les disciplines connexes (informatique, psychologie, neuropsychologie, sociologie) et les applications à l'éducation, le développement d'applications industrielles en informatique et à la santé. Un tel atelier a pour but de mener à la production d'une synthèse écrite sur chacun des trois thèmes.

L'accent sera mis sur les relations entre la recherche fondamentale à l'interface des sciences du langage, de la psycholinguistique, des neurosciences et des sciences de l'ingénieur et la recherche appliquée au secteur de la communication électronique, de l'éducation et de la santé.

### **Présentation détaillée**

L'objectif de l'atelier Langage et Cognition est de constituer un groupe de travail multidisciplinaire focalisé sur les relations entre langage, cognition et nouvelles technologies de la communication, qui fournisse des éléments visibles (conférence ou colloque, ouvrage collectif, site web) dont puissent profiter les acteurs de toutes les disciplines des sciences de la cognition.

En effet, si le langage possède des caractéristiques communes avec d'autres modalités cognitives, il présente également sa spécificité et sa complexité propre, caractéristiques qui appellent des expertises linguistiques tout à la fois particulières et complémentaires

(expertises sur les aspects formels, fonctionnels et interactionnels, expertises sur les structures et les processus langagiers).

Dans ce cadre, l'objectif est de décrire et proposer de nouvelles directions de recherche et de développement scientifique dans les disciplines des sciences cognitives, mais aussi de fournir des indications thématiques et méthodologiques visant en particulier l'enrichissement des modèles langagiers et des outils qui peuvent en dériver. De ce point de vue, 3 angles d'attaque seront envisagés en parallèle : (1) en amont, une perspective théorique sur la façon de penser les nouveaux modèles de communication en relation avec l'évolution des technologies et les connaissances accumulées dans le domaine de la linguistique fonctionnelle et cognitive d'une part et des neurosciences d'autres part, (2) en aval, les ressources de plus en plus volumineuses en termes de données langagières et d'outils pour les traiter, (3) au bout de la chaîne : la constitution de nouveaux systèmes dans des domaines d'application variés. Les domaines sélectionnés dans cet ARP se décomposent en trois axes de réflexion principaux :

(1) outils pour l'interaction optimale avec le web ou via le web

La maîtrise, l'analyse et l'accès à des informations pertinentes adaptées aux besoins de l'utilisateur deviennent des enjeux stratégiques dans un environnement Internet. Au fur et à mesure que les bases de documents s'accroissent (tout type d'objets textuels et multimédias) et que les types de données se diversifient, que la taille des informations disponibles se multiplie, il faut révolutionner les moyens encore largement insuffisants dont on dispose pour interroger, parcourir et interagir avec toutes ces données, quels que soient leur volume.

-mettre en avant des avancées méthodologiques pour l'accès à l'information qui intègrent les contraintes cognitives des utilisateurs et qui exploitent au mieux les techniques de recherche d'information (PageRank, Structure des Réseaux de terrains, Indexation, Complexité, Visualisation/Navigation) et les sciences du langage (TAL, réseaux sémantiques, universaux, polysémie, désambiguïsation, cotexte).

-aller au-delà de la simple navigation hypertextuelle codée de manière figée lors la conception des données. On peut envisager des systèmes où les connaissances de navigation ne sont donc pas explicitement modélisées mais plutôt dispersées dans le texte. Dans cette approche, on considère qu'un texte peut être visualisé de différentes manières, appelées vues du texte, et que chaque vue peut donner lieu à différentes manières de naviguer dans le texte. Les connaissances navigationnelles sont alors modélisées indépendamment de l'objet texte et gérées dans une base de connaissances appelée « module de navigation ».

-gérer les interactions entre les utilisateurs et les systèmes (quel que soit le nombre d'acteurs humains ou de machines impliqués). En particulier, on peut envisager des systèmes plus évolués et plus « humains » grâce aux possibilités de synthèse de la parole : simulation et reconnaissance d'émotions, modèles génératifs de l'expressivité dans la parole, de la prosodie, transformations de voix et d'identité.

(2) modélisation et simulation de fonction langagières complexes : interactions et dialogues, apprentissage du langage, apprentissage de langues particulières, évolution des langues

L'augmentation de données de corpus de langage (en particulier grâce à la multiplication de corpus incluant son et vidéo et permettant des descriptions comportementales), l'amélioration de la puissance des systèmes et surtout des capacités mémoire, permettent de nouvelles approches en matière de modélisation de l'apprentissage et de l'évolution du langage, que ce soit en utilisant des paradigmes classiques (approche symbolique), des réseaux de type neuromimétique, des approches distribuées ou de type multi-agents.

L'élargissement des bases empiriques manifestant les dynamiques cognitives dans leur unité comme dans leurs variations suscitent un renouvellement du questionnement sur la nature du fait cognitif et sa projection sur l'humain envisagé entre l'individu et le(s) groupe(s) aux différentes échelles d'historicité (cognitivisme vs. autopoïèse et enaction, internalisme vs. externalisme). Le croisement des descriptions et des catégories linguistiques classiques avec leur mise en œuvre dans

l'expérience intersubjective (prosodie, kinésie, gestique, contact visuel) appréhendée dans sa multimodalité par les nouveaux corpus (voir aussi axe 1) met à portée l'implémentation matérielle des théories cognitives récentes.

(3) nouvelles technologies pour l'apprentissage des langues, pour le dépistage et pour la remédiation dans le cadre d'application clinique et scolaire

Les améliorations des interfaces et des capacités de simulation des traitements du langage offrent de nouvelles applications pour l'apprentissage des langues (langues maternelles et langues secondes) et pour le diagnostic et la rééducation des troubles de langage. Également, bien que l'Éducation Nationale et l'ensemble des acteurs du terrain scolaire portent une attention de plus en plus marquée au repérage ainsi qu'à la prise en compte des difficultés des enfants en proie à des troubles du langage, il manque encore une pleine prise en charge du dépistage, de l'accompagnement et de la remédiation précoces des troubles sur le terrain familial ou scolaire. Le contexte de l'ARP est un moyen de proposer des outils et des synergies dans ce sens.

La simulation du développement langagier, en parallèle avec les travaux similaires sur le développement cérébral, permettent d'envisager de reconstituer les trajectoires développementales d'un retard ou d'un déficit. La modélisation de mécanismes cognitifs et langagiers dans un cerveau en développement perturbé (surdité, syndromes génétiques, troubles neurodéveloppementaux comme l'autisme, troubles spécifiques du langage oral ou du langage écrit) peut avoir un impact sur la conception d'algorithmes du développement cognitif, sur la plasticité fonctionnelle corticale, et sur la façon de lier l'évaluation et la remédiation. On peut envisager d'anticiper les conséquences d'une rééducation grâce à la simulation informatique et ainsi pouvoir choisir les méthodes les plus appropriées selon les spécificités de chaque profil médical.

Cette division en trois axes peut sembler artificielle dans la mesure où des thématiques communes se retrouvent d'un axe à un autre. En outre, certains outils théoriques et pratiques abordés dans les axes 1 et 2 ont des retombées pour les applications de l'axe 3. En conséquence, on attend une interaction étroite entre les différents axes.

L'objectif de cet ARP est de donner aux acteurs scientifiques les moyens de réflexion et d'auto-évaluation qui sont nécessaires pour ne pas reproduire les élans, souvent généreux mais pas toujours productifs, du passé des recherches en intelligence artificielle. Cette visée prospective s'inscrit dans un cadre cumulatif, en d'autres termes, lancer les chercheurs sur des voies d'investigation nouvelles, originales mais également réalistes. Elle implique une connaissance exhaustive du terrain et une capacité à établir des passerelles entre des milieux d'expertise complémentaires et à faire émerger des synergies efficaces et nécessaires pour que de nouveaux programmes scientifiques soient lancés mais également aboutissent.