

LE PROJET CL² : Complexité, Langages et Langues

Projet piloté par François Pellegrino, Laboratoire Dynamique du Langage (CNRS, Université Lumière Lyon 2).

➤ Les partenaires

- Laboratoire Dynamique Du Langage (CNRS, Université Lumière Lyon 2) rattaché à l'IXXI.
- University of California, Berkeley (Etats-Unis).
- University of Texas at Austin, (Etats-Unis).
- CBU Medical Research Council, Cambridge, (Royaume-Uni).

➤ Durée et financement

Le projet est financé par l'ANR (Agence Nationale de Recherche), pour trois ans à hauteur de 185 000 euros. Il a débuté en décembre 2005.

Entretien avec François Pellegrino.

Vous dites que le langage est un système dynamique complexe à lui tout seul ?

Le langage est un objet scientifique à deux facettes : l'une étudie le langage comme capacité humaine à communiquer, capacité cognitive, et l'autre réside dans toute la diversité des milliers de langues parlées dans le monde. Certaines caractéristiques leur sont communes, et à l'inverse, on observe une très grande variation sur d'autres phénomènes, comme la manière dont les phrases sont construites ou les sons choisis et la manière de les combiner. Le langage est un objet complexe en soi, en particulier parce qu'on ne peut pas dire, je vais regarder le langage, je vais regarder le français. On peut étudier comment les gens l'utilisent, avoir une idée de comment ça fonctionne dans leur cerveau pour comprendre, pour produire une phrase en français, mais on ne peut pas dire, je regarde tout le langage. Le langage existe dans le cerveau des gens et dans l'utilisation qu'ils en font. Par ailleurs, le langage est quelque chose de structuré, où à partir d'un ensemble de sons, on les combine pour faire des mots et on combine ces mots pour faire des phrases. Ces différents niveaux interagissent d'une manière qui n'est pas du tout évidente, qui est même très mal comprise aujourd'hui. Toutes les langues ne vont pas utiliser les mêmes sons ni les combiner de la même manière. Enfin, ces interactions sont dynamiques parce que les langues évoluent en permanence. C'est donc un système dynamique complexe en lui-même.

Vous abordez également la complexité en étudiant l'acquisition du langage !

Oui, il y a un autre domaine très intéressant à la fois pour étudier la complexité et où la complexité peut nous permettre de mieux comprendre comment ça fonctionne, c'est l'acquisition du langage : comment un enfant acquiert sa langue maternelle ou un adulte une seconde langue. Il est probable qu'en mesurant la complexité de certains éléments du langage - ça peut être des sons plus complexes que d'autres ou des enchaînements de sons - on se rende compte que cette complexité intervienne dans les processus d'acquisition. Inversement, étudier l'acquisition du langage peut nous aider à identifier les éléments du langage qui sont plus ou moins complexes et être utile pour faire progresser les théories linguistiques. La complexité nous aide donc aussi dans la description des langues du monde, à mieux comprendre les liens entre la manière dont une langue, par exemple, construit ses mots à partir des sons et le choix des sons. Il y a donc des aspects liés à la cognition et à ce qu'on appelle la typologie linguistique, c'est-à-dire le travail de décrire des langues et de les classer, de regarder ce qui est commun, ce qui est différent d'une langue à l'autre.

Dans votre recherche, qu'avez-vous choisi de modéliser ?

*On peut s'intéresser à plusieurs choses en termes de complexité du langage, à l'interaction entre le choix des sons, la manière dont les sons sont combinés pour faire des mots et la manière dont ces mots sont combinés pour faire des phrases. On définit là, trois sous-disciplines de la linguistique que sont **la phonologie**, la science des sons ; **la morphologie**, comment on combine ces sons pour faire des mots et **la syntaxe**, comment on combine ces mots pour faire des phrases. **Dans ce projet**, on s'intéresse surtout à la syntaxe et la phonologie, et moi personnellement, je travaille surtout sur les sons. On a choisi d'étudier avec des modèles et des approches de systèmes dynamiques complexes, la manière dont les sons des langues sont choisis et interagissent. Dans la base de données que nous utilisons, il y a la description des systèmes de sons de centaines de langues, allant de 11 sons pour une langue de Papouasie jusqu'à 141 sons pour une langue d'Afrique australe. Cela amène au double constat que l'on peut efficacement communiquer dans les deux cas et qu'il y a beaucoup de degrés de liberté pour communiquer. Ce qui nous intéresse, c'est de comprendre comment ces sons vont être choisis et quelles sont les règles d'organisation des systèmes car on constate que les sons qu'utilise une langue ne sont pas pris au hasard. Par exemple, on distingue des voyelles orales (pour lesquelles l'air est expiré par la bouche) et des voyelles nasales (l'air est alors expiré à la fois par la bouche et le nez). En français, comme dans environ une langue du monde sur cinq, on trouve les deux types de voyelles par exemple dans 'sec' et 'cinq'.*

*En étudiant les systèmes où il y a beaucoup de voyelles, on se rend compte qu'on a par exemple une dizaine de voyelles orales, et autant de voyelles nasales qui correspondent aux voyelles orales, mais où en plus l'air sort par le nez. Il y a là une régularité qui n'est pas le fait du hasard et qu'on essaie de comprendre, avec des modèles qui modélisent cette structure et les interactions qu'il peut y avoir entre les sons. L'objectif est de voir si on arrive à identifier des principes d'organisation et d'évolution des systèmes de sons des langues du monde. Pour illustrer les interactions possibles au sein d'un système, j'ai un exemple tout simple en anglais. L'anglais utilise deux sons proches que nous, francophones, avons du mal à distinguer : le 'f', comme dans **fire** (feu) est le même qu'en français, mais le 'th' comme dans **to think** (penser) est différent. Ces sons sont assez difficiles à distinguer à l'oreille et peu de langues les utilisent ensemble car dès qu'il y a du bruit, on n'arrive pas facilement à les distinguer. Une des conséquences, c'est que ces sons ont tendance à ne pas coexister ensemble dans le système de sons d'une même langue. A terme, en anglais, il est probable qu'ils fusionnent et des mots comme par exemple, **fin** (la nageoire) et **thin** (mince) se prononceraient de la même manière.*

De quelle manière travaillez-vous dans le cadre de l'Institut ?

L'intérêt de travailler avec l'Institut est clair : dans les derniers travaux qu'on a entrepris, on a mis en œuvre des méthodologies de modélisation pour les interactions entre les sons, qui ont été développées par les physiciens et qui ont été utilisées, par exemple, pour des problématiques en économie. Ces modèles de recherche demandent à être améliorés, adaptés aux besoins de chacun. Et c'est par ce dialogue entre des chercheurs ayant plutôt une culture de physique, ou de modélisation ou de science humaine pour ce qui est de l'objet qui nous intéresse, qu'on arrive aussi à développer et faire progresser ces modèles, pour qu'ils prennent en compte des phénomènes qui sont adaptés à la problématique du langage. Par exemple, un logiciel du commerce ne serait pas suffisant comme modèle et on mène donc des recherches en conception de modèles et le dialogue avec les physiciens et les modélisateurs est fondamental. Un autre point de contact avec l'Institut des systèmes complexes porte sur des aspects plus biologiques. Lorsque par exemple, on travaille sur la compréhension de la parole bruitée, on collabore avec des praticiens hospitaliers, des spécialistes de neurosciences qui s'intéressent au système d'audition, donc vraiment à l'oreille, à son fonctionnement neuronal. Là, on touche vraiment au fonctionnement du cerveau et du système sensoriel qui est également le cadre de processus de système dynamique, dans le sens où, si le traitement du son va de l'oreille vers le cerveau, en retour, le langage influence aussi la manière dont on perçoit le son. Cela donne lieu à des interactions généralement non linéaires, mal connues, et qui demandent à mettre ensemble, des gens qui travaillent en neurosciences, d'autres qui s'intéressent au langage, des modélisateurs et des biologistes qui comprennent le fonctionnement biochimique du nerf auditif, donc là aussi, dans une approche interdisciplinaire.

Quelles sont les répercussions concrètes de vos recherches, dans la vie au quotidien ?

Une retombée principale est de mieux maîtriser ce qui pose problème dans la compréhension de la parole dégradée, par du bruit ou pour des raisons de vieillissement, ce qui est très important parce que ces difficultés de communication ont un impact majeur puisque les gens qui en souffrent se retrouvent isolés. Mais on constate aussi que des gens qui n'ont pas de problèmes graves détectés ont malgré tout des

problèmes de compréhension dès lors qu'il y a du bruit autour. C'est donc essentiel de comprendre ces problèmes liés à la perception du langage, car aujourd'hui, si on est capable de diagnostiquer des problèmes d'audition de certaines fréquences ou certains niveaux d'intensité, on est beaucoup moins avancé pour ce qui est de la perception de la parole et de la compréhension du langage lui-même. Certains sons du langage seront-ils moins bien perçus que d'autres, comment la compréhension dans le bruit sera-t-elle affectée, etc. ? Il y a en fait beaucoup de choses à faire et la complexité donne un nouvel éclairage là-dessus !

Détecter plus tôt aussi les problèmes de l'acquisition du langage ?

Effectivement, le deuxième élément à plus court terme porte sur l'acquisition du langage. En regardant comment beaucoup d'enfants acquièrent le langage, on peut espérer détecter de manière précoce ceux qui vont avoir des problèmes dans cette acquisition. Aujourd'hui, on n'est pas réellement capable de le faire, parce qu'il y a énormément de variabilité entre les individus, pour l'acquisition du langage comme pour tout le reste. Et comme on ne sait pas bien ce qu'est la variabilité normale pour l'acquisition, on a du mal à décider si un enfant relève de cette variabilité normale ou pas. On sait que tous les enfants ne passent pas par les mêmes phases, au même moment ou dans le même ordre, au niveau du développement cognitif langagier. En fait, cette variabilité perdure chez les adultes ; on s'est rendu compte par exemple, en faisant des études sur des adultes sans problèmes d'audition, que si on dégrade des mots qu'on leur fait entendre, certains n'auront aucun problème tandis que d'autres de comprendront plus, alors qu'ils ont tous une audition dite normale ! Tant qu'on ne sait pas ce qu'est la variabilité normale, c'est très difficile de voir où est la frontière entre ce qui est normal et ce qui est problématique. L'acquisition du langage est quelque chose de très fin, de très complexe, et lorsqu'un enfant a de légers problèmes à un certain niveau, d'audition par exemple, il va compenser, il va peut-être plus regarder le visage des gens pour voir ce qu'ils prononcent, son système cognitif va peut-être travailler différemment et lui permettre d'acquérir le langage normalement. Par contre, si les difficultés s'accumulent et qu'il n'arrive pas à faire cette compensation, c'est sans doute là qu'on bascule dans une situation problématique. Dans ce contexte et si on est capable de graduer des éléments de l'acquisition du simple au complexe, cela peut nous aider à estimer si l'acquisition de l'enfant se passe bien ou pas.

SITE : www.ddl.ish-lyon.cnrs.fr

**CONTACT PROJET : François Pellegrino
francois.pellegrino@univ-lyon2.fr**