

Annuaire du Collège de France

121^e année

2020
2021

Résumé des cours et travaux



COLLÈGE
DE FRANCE
— 1530 —



Annuaire du Collège de France

Cours et travaux du Collège de France

121 | 2024
2020-2021

Psychologie cognitive expérimentale

Stanislas Dehaene



Édition électronique

URL : <https://journals.openedition.org/annuaire-cdf/19350>

DOI : 10.4000/12ktu

ISBN : 978-2-7226-0778-1

ISSN : 2109-9227

Éditeur

Collège de France

Édition imprimée

Date de publication : 18 novembre 2024

Pagination : 175-195

ISBN : 978-2-7226-0777-4

ISSN : 0069-5580

Ce document vous est fourni par Collège de France



Référence électronique

Stanislas Dehaene, « Psychologie cognitive expérimentale », *L'annuaire du Collège de France* [En ligne], 121 | 2024, mis en ligne le 01 octobre 2024, consulté le 28 novembre 2024. URL : <http://journals.openedition.org/annuaire-cdf/19350> ; DOI : <https://doi.org/10.4000/12ktu>

Le texte et les autres éléments (illustrations, fichiers annexes importés), sont « Tous droits réservés », sauf mention contraire.

PSYCHOLOGIE COGNITIVE EXPÉRIMENTALE

Stanislas Dehaene

Membre de l'Institut (Académie des sciences),
professeur au Collège de France

La série de cours « L'influence du langage et des symboles sur la perception et la cognition » est disponible en audio et en vidéo, sur le site internet du Collège de France (<https://www.college-de-france.fr/agenda/cours/influence-du-langage-et-des-symboles-sur-la-perception-et-la-cognition>), ainsi que le colloque « The Representation of language in brains and machines » (<https://www.college-de-france.fr/agenda/colloque/the-representation-of-language-in-brains-and-machines>).

ENSEIGNEMENT

COURS - L'INFLUENCE DU LANGAGE ET DES SYMBOLES SUR LA PERCEPTION ET LA COGNITION

Introduction

Quels sont les liens entre langage et pensée ? Pourrions-nous avoir les mêmes pensées si nous n'avions pas la faculté de langage ? Les mots dont nous disposons nous incitent-ils à voir le monde sous un angle particulier ? Pensons-nous différemment en français, en russe ou en chinois ? Ou bien, ces langues ne sont-elles que des dispositifs de communication de nos représentations mentales, qui demeurent indépendantes du langage ?

Les sciences cognitives contemporaines apportent à ces questions classiques des réponses nouvelles. Le cours 2020-2021 a rassemblé des données issues de disciplines

telles que l'éthologie, la psychologie, l'anthropologie cognitive et l'imagerie cérébrale, afin d'analyser l'influence éventuelle qu'exerce la langue que nous parlons sur nos représentations mentales, au niveau tant perceptif que conceptuel.

Cours 1 - Introduction aux théories du relativisme linguistique et culturel

L'idée que la pensée humaine varie selon la langue que nous parlons est fréquemment appelée « thèse de Sapir-Whorf ». Edward Sapir (1884-1939), anthropologue et linguiste américain, affirmait en effet que « les êtres humains ne vivent pas seulement dans le monde objectif, ni seulement dans le monde des activités sociales telles qu'on les entend d'ordinaire, mais ils sont à la merci de la langue particulière qui est devenue le moyen d'expression dans leur société. » Son élève, le linguiste Benjamin Lee Whorf (1897-1941), prolongeait cette thèse :

le système linguistique [...] de chaque langue n'est pas seulement un instrument de reproduction servant à vocaliser des idées mais il est lui-même formateur d'idées [...].
Nous découpons la nature selon les tracés que notre langue a dessinés sur elle.

Cette thèse séduisante a fait son chemin dans la culture occidentale du xx^e siècle. Elle joue notamment un rôle central dans le roman *1984* de George Orwell (1949), dans lequel un dictateur façonne et impose une nouvelle langue, le « Novlangue », dans le but de restreindre la liberté de penser. En éliminant progressivement certains mots du vocabulaire, écrit Orwell, « le climat total de la pensée sera autre. En fait, il n'y aura pas de pensée telle que nous la comprenons maintenant. » La thèse whorfienne se diffuse également dans le grand public par le biais de légendes urbaines, telles que celle qui voudrait que la langue eskimo dispose de centaines de mots pour la neige, et que cela constituerait une preuve que le peuple esquimau catégorise le monde de façon idiosyncrasique. Cette idée est un mythe, comme nous l'avons expliqué dans le cours en nous appuyant sur l'article de Pullum (1989).

La communauté de sciences cognitives reste cependant considérablement divisée autour de la thèse de Sapir-Whorf. Pour certains chercheurs comme Lera Boroditsky ou Steve Levinson, le langage façonne ou restructure la pensée : les mots que nous utilisons ont un effet déterminant sur notre manière de conceptualiser et d'interagir avec notre environnement. Pour beaucoup d'autres, tels Leila Gleitman ou Steven Pinker, l'influence du langage sur les processus non linguistiques est minimale ou nulle. Les mots ne sont que des étiquettes sur des concepts non verbaux et universels que nous partageons tous, à d'infimes variations près. Leila Gleitman, par exemple, soutient que les structures linguistiques n'ont d'influence que quand le locuteur choisit d'adopter une stratégie verbale – et bien entendu, il n'est guère surprenant que la langue influence ainsi le traitement... linguistique. Dans *The Language Instinct*, Steven Pinker reprend la thèse de Jerry Fodor selon laquelle il existe un « langage de la pensée » bien différent du langage que nous utilisons pour communiquer. Selon cette hypothèse, on ne pense pas en français ou en russe, mais

dans un langage interne abstrait et universel, et c'est seulement au moment de communiquer que nous devons le figer dans une langue spécifique.

De fait, on peut considérer, schématiquement, qu'il existe trois versions de l'hypothèse de Sapir-Whorf. Selon la version forte, le langage interviendrait directement pour façonner nos représentations mentales. Selon une interprétation plus faible, le langage ne jouerait pas un rôle central dans la pensée, mais fournirait un outil supplémentaire, plus ou moins efficace, dans notre boîte à outils mentale. Enfin, selon une troisième interprétation, le langage ne nous biaiserait que dans la mesure où il fournit un moyen de transmission d'idées nouvelles, y compris des biais attentionnels ou culturels : par son acte de parole, le locuteur attirerait l'attention de l'auditeur sur telle ou telle distinction perceptive ou conceptuelle.

Le reste du premier cours a été consacré à discuter de cette troisième interprétation, dont la réalité ne fait aucun doute, mais qui n'est pas réellement « whorfienne ». Rappelons d'abord qu'il existe d'importantes variations d'exposition au vocabulaire d'une personne à l'autre et que, dès la petite enfance, ces différences se traduisent par des modifications de l'activité cérébrale : chez l'enfant entre 4 et 6 ans, l'activation de la région de Broca dépend directement du nombre de tours de conversation dans lesquels l'enfant a été engagé (Romeo *et al.*, 2018). Ensuite, de nombreuses expériences montrent que le choix des mots peut susciter, chez l'auditeur, des biais cognitifs et perceptifs importants. Ainsi, en psychologie sociale, la même activité cognitive (reproduire une figure de mémoire) peut induire ou non des biais de genre en défaveur des filles, selon qu'elle soit décrite comme une épreuve de « dessin » ou de « mathématiques » (Huguet *et al.*, 2001). Plusieurs travaux, notamment ceux de Sandy Waxman, montrent que la langue est un amplificateur de préjugés sexistes ou raciaux. Dans un autre domaine, les effets placebo et nocebo, ou encore l'hypnose, qui sont des phénomènes biologiques bien réels, illustrent à quel point les mots peuvent affecter, par exemple, la sensation subjective de douleur.

S'agit-il pour autant d'effets « whorfiens » ? Pas vraiment, en tout cas pas dans le sens où l'entendait Whorf lui-même. D'abord, ce sont des effets universels, sur un système perceptif universel – tous les locuteurs subissent l'effet placebo, quelle que soit leur langue. Ensuite, la transmission des idées et des connaissances est la fonction strictement normale du langage ; il n'y a donc rien d'étonnant à ce qu'elle modifie nos décisions. Dès la naissance, nous adhérons souvent étroitement à ce que disent les autres, soit parce que nous pensons qu'ils en savent plus que nous, soit parce que nous craignons de nous tromper, soit enfin pour préserver notre intégration au groupe social. L'imagerie cérébrale suggère que ces effets de transmission linguistique surviennent à un haut niveau de traitement cognitif, conscient, sensible à l'attention et à la distraction, tandis que les aires sensorielles de bas niveau ne sont pas affectées. La plupart des calculs sensoriels de bas niveau sont « cognitivement impénétrables », c'est-à-dire qu'ils ne changent pas en fonction des informations que nous recevons. Les illusions visuelles, par exemple, continuent de produire leur effet même lorsqu'on est informé de leur nature.

Dernier piège lié à l'hypothèse de Whorf : il n'est pas toujours facile de séparer les effets de la langue de ceux de la culture. Comme l'illustrent les travaux de Richard Nisbett, il semble exister de subtiles différences de style cognitif, par exemple entre les étudiants américains et chinois, mais ce n'est pas nécessairement la langue elle-même qui les cause, mais plutôt des habitudes culturelles.

Cours 2 - Pensée sans langage et hypothèse d'un langage de la pensée

La pensée peut-elle exister sans langage ? La thèse d'une identité entre langage et pensée a été défendue par certains philosophes. Dans le *Théétète*, Platon fait à dire à Socrate : j'appelle *pensée* « un discours que l'âme se tient tout au long à elle-même sur les objets qu'elle examine ». Merleau-Ponty va plus loin : « la pensée n'a rien d'intérieur, elle n'existe pas hors du monde et hors des mots » (*Phénoménologie de la perception*, 1945). Ray Jackendoff attribue cette idée à une sorte d'illusion métacognitive : « nous faisons très souvent l'expérience de nos pensées comme une sorte de monologue intérieur – nous entendons littéralement des mots, des syntagmes ou des phrases dans notre tête, et il est alors vraiment tentant de caractériser la pensée comme une sorte de langage intérieur. » Comme le souligne également François Recanati, cette introspection peut cependant s'expliquer par un simple parallélisme entre les deux domaines : les *contenus* de pensée seraient indépendants du langage, mais, dans l'espèce humaine, la parole accompagnerait toute *activité* de pensée un tant soit peu complexe.

Effectivement, l'analyse des représentations mentales animales, par les méthodes de l'éthologie et des neurosciences, donne tort à Platon et Merleau-Ponty : une pensée abstraite et complexe peut très bien exister en l'absence de tout langage. Le cerveau de la souris cartographie son environnement, celui du singe macaque représente le nombre d'objets... toutes ces espèces pensent sans mots, du moins si l'on définit la pensée comme la représentation mentale et la manipulation d'informations abstraites.

Il reste toutefois possible de soutenir que le langage vient à l'appui de la pensée, qu'il en constitue un outil essentiel. Quel en serait alors le rôle ? Andy Clark, à la suite de Frege, défend l'idée que l'expression de la pensée sous une forme linguistique permet sa manipulation explicite :

La formulation linguistique rend des pensées complexes accessibles aux processus d'attention mentale, ce qui permet de les soumettre à une variété d'autres opérations. Cela nous permet, par exemple, d'isoler différents éléments d'une pensée complexe afin de les examiner tour à tour ; de stabiliser des idées très abstraites en mémoire de travail ; et d'inspecter et de critiquer notre propre raisonnement.

Benjamin Whorf va encore plus loin. Pour lui, la formation d'idées abstraites dépend du langage : « le système linguistique (...) de chaque langue n'est pas seulement un instrument de reproduction servant à vocaliser des idées mais il est lui-

même formateur d'idées (...). » Le langage serait également responsable de la formation des catégories mentales : « nous découpons la nature selon les tracés que notre langue a dessinés ».

Les résultats expérimentaux de l'éthologie, mais aussi de la psychologie du développement de l'enfant, sont en complet désaccord avec ces thèses. Faute de temps, le cours s'est contenté de n'en citer que quelques exemples. Le cortex préfrontal du singe macaque catégorise un continuum d'images, sans mots bien sûr, simplement sur la base des récompenses qu'il reçoit (Freedman *et al.*, 2001). De plus, les principes qui sous-tendent l'apprentissage des catégories mentales semblent identiques chez l'Homme et chez le babouin (Chemla *et al.*, 2019). Enfin, *contra* Andy Clark, les singes macaques, bien que dépourvus de tout langage, sont capables de prêter attention à leurs représentations mentales et d'émettre des jugements de similarité de second ordre, ou encore un jugement introspectif sur leur niveau d'ignorance (cf. le cours 2010-2011 sur la métacognition et l'introspection).

Pour autant, n'est-ce pas l'apprentissage des symboles qui distingue notre espèce et qui permet aux représentations mentales de l'espèce humaine de dépasser celles des autres animaux ? L'examen des représentations animales fait apparaître que, même si de nombreuses espèces peuvent apprendre des dizaines de signes arbitraires, leurs représentations mentales ne sont pas identiques aux nôtres. Leurs signes s'écartent de nos symboles d'au moins deux manières. D'une part, les liens signifiant-signifié ne sont pas spontanément réversibles : lorsqu'un singe apprend que le symbole A prédit l'objet α , il ne généralise pas cette connaissance dans la direction opposée, c'est-à-dire de l'objet α vers le symbole A. Ainsi, les animaux semblent incapables de faire des allers-retours flexibles entre le signifiant et le signifié. D'autre part, et c'est le plus fondamental, les animaux ne parviennent pas à combiner plusieurs symboles selon une authentique syntaxe. Cela semble les empêcher de conceptualiser les régularités mathématiques ou scientifiques aussi profondément que nous. Le cours s'est attardé sur l'exemple des formes géométriques : à test rigoureusement identique, même un simple carré fait l'objet d'une représentation éminemment différente chez l'Homme et chez le babouin : l'espèce humaine les représente à l'aide d'une combinatoire de traits symboliques (parallélisme, angles droits, etc.), tandis que le babouin se cantonne à une représentation perceptive approximative, semblable à celle des réseaux de neurones artificiels à convolution (Sablé-Meyer *et al.*, 2021).

S'agit-il, pour autant, d'un effet linguistique ? Non, car l'effet de régularité géométrique s'observe chez tous les membres de l'espèce, y compris des enfants d'âge préscolaire qui n'ont pas encore appris le vocabulaire de la géométrie, ou des adultes Himba de Namibie, qui n'ont pas accès à une éducation formelle dans ce domaine. Il ne s'agit donc pas d'un effet whorfien, mais d'une caractéristique universelle de l'espèce humaine. Nous faisons l'hypothèse que seule l'espèce humaine possède :

- un système de liens symboliques, arbitraires, bidirectionnels, entre signifiants et signifiés ;

- une capacité combinatoire interne, un « langage de la pensée » qui permet de combiner ces symboles.

Selon cette hypothèse, nous possédons les mêmes noyaux de connaissances (*core knowledge*) que d'autres espèces de primates : objets, formes, personnes, nombres, positions, etc. Mais nous les recombinaisons en phrases dans un « langage de la pensée », ce qui nous permet de former une pyramide infinie de concepts enchâssés. Ce langage interne est universel : tous les humains possèdent cette capacité et peuvent donc accéder aux mêmes concepts. Cependant, l'espace conceptuel ouvert par cette capacité réursive est si vaste qu'aucun être humain ne peut l'explorer dans sa totalité, et que différentes cultures ne font pas les mêmes choix ou les mêmes découvertes. Le rôle de la communication linguistique, par le biais d'une langue spécifique, pourrait être d'orienter l'attention vers certaines combinaisons qu'une culture donnée, avec l'expérience des générations, juge plus importantes que d'autres. Il y aurait donc à la fois *universalité* (tous les membres de l'espèce humaine sont capables de penser les mêmes pensées) et *spécificité culturelle* (la disponibilité de certains mots, particulièrement au cours de l'apprentissage, biaise notre attention et affecte nos jugements perceptifs et cognitifs).

Cours 3 - La langue influence-t-elle la perception des émotions et des couleurs ?

Le troisième cours a examiné l'influence du langage sur la perception. Deux théories s'opposent. Selon Whorf et quelques autres chercheurs contemporains comme Lera Boroditsky, la perception peut être profondément modifiée par la disponibilité, ou non, d'un mot pour la décrire. Au contraire, selon Leila Gleitman et beaucoup d'autres chercheurs en sciences cognitives, les étapes initiales de la perception sont modulaires, indépendantes du langage, et d'éventuels effets linguistiques n'apparaissent que si les participants s'appuient sur le langage pour résoudre la tâche exigée. Un test clé permet de distinguer les deux théories : selon la seconde, l'éventuelle influence du langage doit disparaître si on empêche les sujets de faire appel à leurs processus linguistiques, par exemple en interférant avec ceux-ci.

Nous avons examiné l'état de la littérature empirique dans deux domaines : la perception des émotions et celle des couleurs. En ce qui concerne les émotions, les travaux de Paul Ekman font référence. Prolongeant les idées de Charles Darwin dans *L'Expression des émotions chez l'homme et les animaux* (1872), Ekman montre que tous les groupes humains expriment et perçoivent de la même manière au moins six émotions de base : colère, dégoût, peur, joie, tristesse, surprise – une liste à laquelle Ekman ajoutera par la suite quelques autres émotions telles que le mépris, la honte, la fierté, la culpabilité... Leur expression sur le visage, et la capacité de les reconnaître seraient universelles. Différentes cultures peuvent cependant moduler l'expression de ces émotions, voire la réprimer dans certains contextes. Le cours a examiné en détail deux articles récents (Jack *et al.*, 2012 ; Jackson *et al.*, 2019) qui prétendent remettre en question les travaux d'Ekman, mais nous avons vu que leurs résultats ne

montrent que de faibles variations cross-culturelles et, surtout, ne prouvent absolument pas que ces différences soient déterminées par le lexique de la langue plutôt que par d'autres facteurs.

La situation est différente dans le domaine de la perception des couleurs, qui fait l'objet de très nombreuses expérimentations rigoureuses depuis les travaux princeps de Berlin et Kay (*Basic Color Terms: Their Universality and Evolution*, 1969). Berlin et Kay montrent que les termes basiques de couleur varient certes selon les langues, mais qu'ils forment une hiérarchie universelle. Les simulations mathématiques, notamment celles de Regier et collaborateurs (Regier *et al.*, 2007) indiquent que les termes de couleur, loin de varier arbitrairement d'une culture à l'autre, reflètent la subdivision quasi-optimale d'un espace perceptif irrégulier, mais partagé par tous les membres de l'espèce humaine. Selon le nombre de mots disponibles, le partage se fera différemment, mais il respectera la hiérarchie d'introduction des mots observée par Berlin et Kay : blanc et noir, puis rouge, vert ou jaune, bleu, marron, et enfin les mots moins fréquents : violet, rose, orange et gris. Une forte universalité se cache donc derrière la diversité apparente des langues.

Cependant, la question centrale de la thèse whorfienne demeure : ces variations linguistiques affectent-elles la perception des couleurs ? Les premières expériences de psychologie, dans ce domaine, montraient effectivement que, dans différentes tâches, le comportement des participants varie en fonction de la langue qu'ils parlent. Par exemple, si deux couleurs portent le même nom dans une langue donnée, alors les locuteurs de cette langue ont tendance à les juger plus similaires, toutes choses égales par ailleurs, que si ces couleurs se voient accorder des noms différents. Ils ont également tendance à les confondre lorsqu'on leur demande de les garder en mémoire à court terme. Ces différences apparaissent assez tôt dans l'enfance, au moment où les enfants acquièrent les termes de couleur. Cependant, Leila Gleitman et, à sa suite, Steven Pinker critiquent ces tâches subjectives ou de mémoire. En effet, rien n'indique que l'effet provienne bien de la perception, et non d'étapes ultérieures de recodage linguistique.

Prenons l'exemple du russe, qui fait une distinction obligatoire entre *goluboy* (bleu clair) et *siniy* (bleu foncé). Les travaux récents de Boroditsky et ses collaborateurs démontrent effectivement des différences entre locuteurs du russe et de l'anglais dans une tâche de jugement perceptif (choisir laquelle de deux tâches de couleur est identique à une troisième). Cependant, crucialement, ces différences disparaissent sous interférence verbale, lorsqu'on demande au sujet de réciter silencieusement des chiffres (Winawer *et al.*, 2007). De plus, de façon fascinante, cet effet n'existe que lorsque les couleurs sont présentées dans l'hémichamp droit, donc à destination de l'hémisphère gauche, dominant pour le langage (Regier et Kay, 2009). Au sein d'une même personne, la perception des couleurs serait donc « whorfienne » dans l'hémisphère gauche, linguistique, mais pas dans le droit !

Même si ces expériences montrent que le niveau linguistique intervient dans une décision perceptive, deux interprétations restent possibles : soit il existe une

activation supplémentaire, purement ascendante (*bottom-up*) des aires du langage, qui contribue à la prise de décision ; soit il existe un authentique retour descendant (*top-down*) de cette activation linguistique sur les étapes perceptives – et seule cette seconde interprétation serait réellement « whorfienne ». Les données disponibles sont faibles et ne permettent pas de conclure. Elles montrent toutefois que les effets perceptifs descendants, s'ils existent, sont minuscules : *grosso modo*, nous percevons tous le monde de la même manière.

La conclusion principale qui émerge de cette ensemble de recherches est celle d'une très grande universalité cross-culturelle de la perception des couleurs. Mais, par-delà le niveau perceptif, lorsque nous devons catégoriser ou retenir une information perceptive, nous nous aidons souvent d'un recodage linguistique qui aide ou qui interfère, surtout lorsque les stimuli sont très proches du seuil perceptif. Point fondamental, lorsque le code linguistique est rendu indisponible (comme c'est le cas chez un patient anomique, ou après interférence verbale chez le sujet normal), la perception continue d'être catégorielle, seules les frontières s'estompent légèrement. Ainsi, l'interprétation la plus plausible est celle de Gleitman et Pinker : le langage apparaît comme un outil cognitif supplémentaire qui, contrairement aux idées de Whorf, ne modifie pas radicalement notre perception.

Cours 4 - La langue influence-t-elle les concepts de personne, d'espace, de temps... ?

Le quatrième cours est d'abord revenu brièvement sur le rôle du langage dans la perception de la douleur (le langage et notamment les jurons peuvent atténuer la douleur en distrayant notre attention) et dans celle de la couleur (un patient aphasique avec anomie des couleurs continue de percevoir les frontières entre les catégories de couleurs, par exemple rouge/marron). Ces travaux confirment que la perception n'est pas directement altérée par le langage : ce sont uniquement les étapes tardives de recodage linguistique et d'accès à la conscience qui sont affectées.

Cependant, qu'en est-il des niveaux conceptuels ? À quoi pensent les patients aphasiques ? Leur capacité de penser se réduit-elle avec la perte du langage articulé ? Ou bien, reste-t-elle inchangée, comme le croit le critique d'art Tom Lubbock, lui-même aphasique :

Mon langage pour décrire les choses du monde est très petit, limité, [mais] mes pensées quand je regarde le monde sont vastes, sans limites, et normales – les mêmes que depuis toujours. Mon expérience du monde n'est pas réduite par l'absence de langage, mais est essentiellement inchangée.

L'introspection de Lubbock s'oppose totalement aux idées développées en 1973 par le philosophe Donald Davidson, qui énonce une version extrême de la thèse de Sapir-Whorf : « Parler une langue n'est pas une chose qu'un homme peut perdre tout en conservant la capacité de penser ». Qui a raison ?

Un premier domaine que nous avons examiné est celui la théorie de l'esprit (*Theory of mind* ou ToM). Selon le philosophe Andy Clark, « il se peut que ce soit le langage public qui soit responsable d'un ensemble de traits distinctifs de la pensée humaine : notre capacité de disposer d'une dynamique de second ordre ». Toute pensée réfléchie, de second ordre, serait alors une pensée linguistique, dépendante du langage. Selon cette thèse, les patients agrammatiques ne devraient plus être capables d'entretenir des pensées complexes, structurées selon de multiples niveaux d'enchâssement. Or, la théorie de l'esprit requiert la capacité de réfléchir aux pensées des autres et de les distinguer des siennes, notamment dans la tâche des fausses croyances. Cela implique l'utilisation de verbes mentaux qui prennent une phrase comme complément (« il pense que X »). De Villiers *et coll.* ont suggéré que le développement de la théorie de l'esprit dépend de celui du langage, et, empiriquement, les deux sont effectivement corrélés, notamment chez les enfants sourds avec retard de langage. Ainsi, selon la version forte de l'hypothèse de Sapir-Whorf, on devrait prédire que, si un patient ne peut plus formuler ni comprendre des phrases enchâssées (agrammatisme), il ne devrait plus être capable de réfléchir aux pensées d'autrui (perte de la *Theory of mind* = « aToMisme », selon l'heureux néologisme forgé par Naama Friedmann).

L'expérience montre qu'il n'en est rien. Naama Friedmann et ses collaborateurs démontrent l'existence d'une double dissociation : une lésion cérébrale peut rendre agrammatique sans affecter la théorie de l'esprit, et inversement, il se peut qu'un patient puisse produire des phrases complexes, mais ne comprenne plus les pensées d'autrui. Les travaux de Rosemary Varley montrent également d'autres dissociations, par exemple entre l'agrammatisme et la musique, ou encore la capacité de calcul arithmétique et même algébrique (Fedorenko et Varley, 2016). Ces domaines sont indépendants du traitement du langage.

Pour autant, la thèse de Whorf est-elle réfutée ? Pas dans sa version faible. D'abord, le langage pourrait jouer un rôle important au cours de l'acquisition des concepts (nous abordons ce sujet dans le cours 5). D'autre part, l'organisation d'une langue particulière pourrait moduler, subtilement, les représentations des concepts, même si celles-ci sont non verbales et reposent sur des réseaux cérébraux distincts. C'est pourquoi le cours a examiné des expériences récentes sur la perception du rythme, du genre, de l'espace et du temps. Dans tous ces domaines, les effets sont faibles et ne permettent pas de conclure à une influence directe de la langue – il semble plutôt s'agir de biais culturels. Nous avons notamment examiné les travaux de Steve Levinson et son groupe (Majid *et al.*, 2004). À première vue, ces travaux soutiennent fortement l'hypothèse de Sapir-Whorf, dans la mesure où ils mettent en évidence d'importantes variations de comportement, dans toute une série de tâches spatiales, selon la manière dont la langue des participants dénote les relations spatiales. En effet, alors que le français utilise principalement un cadre égocentrique, relatif à la position de l'observateur (« la fourchette est à gauche de la cuiller »), d'autres langues n'utilisent qu'un cadre absolu, fondé soit sur les points cardinaux (« la fourchette est au nord de

la cuiller » ; langues arrernte ou guugu yimithirr d’Australie), soit sur des points de repères géographique (« la fourchette est vers le haut de la colline par rapport à la cuiller » ; langue tzeltal du Mexique) ; tandis que d’autres langues encore n’utilisent qu’un cadre intrinsèque, centré sur l’un des objets (« la fourchette est au nez de la cuiller » ; langues mopan (Belize) ou totonac (Mexique).

Levinson et ses collaborateurs observent que ces manières de parler de l’espace influencent les performances – mais Leila Gleitman apporte plusieurs critiques très fortes. On peut, par exemple, biaiser les Américains à répondre comme les Mayas locuteurs du tzeltal en fournissant un point de repère en bout de table ; et à l’inverse, des Mayas peuvent très bien apprendre à répondre de façon égocentrique comme les anglophones (Li et Gleitman, 2002). Les résultats de Levinson sont donc dus au fait que les tâches qu’il utilise sont ambiguës. Dans ce cas, la pragmatique du langage induit les sujets à adopter une règle cohérente avec le vocabulaire de leur langue. Sans entrer ici dans le détail, il en va de même des travaux de Boroditsky sur la représentation du temps. Dans le domaine de l’espace comme du temps, les cadres de référence sont en petit nombre, sans doute innés, universels et accessibles à tous les êtres humains. Les langues et les cultures ne font que moduler et sélectionner parmi ce petit répertoire (pour une conclusion similaire, voir également les travaux de Hespos et Spelke, 2004).

Cours 5 - L’impact du langage sur l’apprentissage chez le jeune enfant

Si le langage n’influence guère la pensée à l’âge adulte, joue-t-il un rôle plus grand au cours du développement ? Le langage pourrait constituer une sorte d’échafaudage temporaire, nécessaire seulement à la construction de la pensée. D’une part, il pourrait faciliter l’acquisition des concepts élémentaires, en étiquetant directement les concepts à apprendre (c’est l’hypothèse de Lupyan) ou en attirant l’attention sur les propriétés intéressantes du monde extérieur (travaux de Sandy Waxman). D’autre part, il pourrait agir comme une colle (*glue*) qui permet d’assembler les primitives conceptuelles issues de nos noyaux innés de connaissance (*core knowledge*) afin de former des concepts composites que nous ne pourrions pas formuler autrement (hypothèse d’Elizabeth Spelke), ou bien que nous aurions beaucoup plus de mal à découvrir, étant donné l’immense espace de concepts combinatoires que notre cerveau est capable de représenter, mais qu’il n’a pas le temps d’explorer tous (c’est ma propre hypothèse).

Le cours a examiné plusieurs travaux empiriques qui mettent à l’épreuve ces possibilités. En premier lieu, l’expérience montre qu’étiqueter des objets par des mots facilite effectivement leur apprentissage, tant chez l’enfant que chez l’adulte. Cet effet est principalement attentionnel. Lorsqu’ils apprennent un nom d’objet, les enfants restreignent l’espace des hypothèses à la forme des objets, ce qui accélère leur catégorisation. Les étiquettes verbales aident également à individualiser les objets.

En second lieu, le langage permet d'exprimer des expressions combinatoires telles que « à gauche du mur vert », qui vont au-delà de ce que signifient les mots isolés. Elizabeth Spelke, avec Hermer-Vazquez, suggère que cette capacité pourrait expliquer la supériorité des enfants humains sur d'autres espèces animales dans une tâche de réorientation spatiale, où il s'agit de rechercher un objet caché à une position potentiellement ambiguë (Hermer-Vazquez *et al.*, 2001). Cependant, ces conclusions sont contestées, notamment parce que des patients sévèrement aphasiques restent capables de réaliser cette tâche. La conclusion la plus conservatrice semble être que le langage fournit, si nécessaire, un appui temporaire à cette tâche, mais que ce sont des représentations spatiales non linguistiques qui sont mises en jeu.

En conclusion, il n'est guère surprenant que les pensées des autres êtres humains, transmises par le biais du langage, affectent très précocement le développement cognitif de l'enfant, que ce soit dans l'apprentissage de catégories (noms) ou de propriétés (adjectifs), ou dans l'utilisation d'indices pour s'orienter dans l'espace. Ces résultats ne signifient pas que, sans langage, nous ne pourrions pas avoir les mêmes pensées ni les mêmes représentations mentales : elles seraient, tout simplement, plus difficiles à atteindre. Les données sont compatibles avec l'existence d'un immense répertoire de concepts non verbaux, fondés sur un « langage de la pensée » universel, qui engendre une très vaste combinatoire de représentations mentales potentielles parmi lesquelles la langue et la culture viennent sélectionner.

Cours 6 - L'influence du langage en mathématiques

Le cours s'est conclu par une analyse de la contribution du langage dans le domaine des mathématiques. Cela nous a permis de passer en revue les grandes conclusions du cours.

D'abord, dans le domaine du nombre et de l'arithmétique élémentaire, il ne fait aucun doute qu'il existe une vaste pensée sans langage, chez l'animal comme chez l'Homme. De nombreuses espèces animales représentent spontanément le nombre, ce qui leur permet d'estimer le nombre de congénères ou la quantité de nourriture. Les bébés de quelques mois, voire quelques jours, sont capables de représenter le nombre d'un ensemble d'objets, soit de façon exacte (*subitizing* pour les nombres 1, 2 ou 3), soit de façon approximative (au-delà de 3), indépendamment de la modalité (auditive ou visuelle). Ils peuvent également combiner ces nombres pour anticiper le résultat d'une opération arithmétique. La pensée arithmétique, au moins approximative, précède donc l'acquisition du langage. Et ce système préverbal de représentation des quantités continue d'être utilisé lorsque nous apprenons les chiffres arabes et que nous faisons des calculs, comme en témoigne l'effet de distance en comparaison numérique et en calcul approximatif.

En second lieu, au cours du développement de l'enfant, le langage sert d'échafaudage temporaire à l'acquisition des concepts. Il facilite l'acquisition du concept de nombre exact en permettant d'apprendre à compter un nombre précis d'objets, comme le

montre toute une série d'expériences chez les indiens Mundurucu et Pirahã d'Amazonie, ou chez les enfants sourds du Nicaragua. Elizabeth Spelke spéculé que cet apprentissage dépend de la maîtrise des règles de la grammaire générative, et que c'est la syntaxe de la langue naturelle qui permet de comprendre les grands nombres. Selon cette hypothèse, les locuteurs des langues différentes – par exemple avec des nombres différents (« vingt » comme en français, *versus* « deux dix » comme en chinois), une base différente, ou un ordre différent des mots (*drei und zwanzig*) – devraient posséder des représentations mentales distinctes. Elizabeth Spelke défend donc bien une théorie « whorfienne ». Cependant, les données de comportement et d'imagerie cérébrale dont nous disposons ne vont pas dans ce sens, au moins chez l'adulte. Elles suggèrent plutôt que les concepts mathématiques sont codés dans un « langage de la pensée » non verbal et universel. Selon mes propres conclusions, au cours du développement, un environnement linguistique riche et stimulant aiderait les enfants à acquérir les concepts de l'arithmétique, non pas parce que ces concepts sont codés verbalement, mais parce que les constructions linguistiques attirent l'attention des enfants sur les constructions mentales non verbales les plus appropriées ou les plus efficaces. Le langage agit donc comme catalyseur des concepts non verbaux : il les rapproche et suggère comment les combiner.

En troisième lieu, chez l'adulte, un recodage linguistique de l'information peut servir d'outil mental dans certaines tâches. La mémoire de chiffres, comme celle des tables de multiplication, bénéficie d'un stockage additionnel sous forme verbale. Ainsi, les patients aphasiques, s'ils conservent la plupart de leurs facultés arithmétiques, peuvent éprouver des difficultés avec les tables de multiplication – et les bilingues, avoir des difficultés à transférer leurs connaissances en calcul exact d'une langue à l'autre.

Enfin, et c'est le plus important, à l'encontre des idées de Whorf, les variations interlangues n'ont qu'un impact modeste sur les capacités perceptives et cognitives. Comme prédit par Leila Gleitman, c'est essentiellement lorsque la tâche fait directement ou indirectement appel au langage que le langage a un impact – et encore, celui-ci reste faible. Ainsi, les enfants apprennent plus rapidement à compter en chinois qu'en anglais ; cet effet semble lié à la plus grande simplicité du système de numération chinoise. La brièveté des noms de nombres en chinois a également un impact positif sur la mémoire. Enfin, l'inversion des dizaines et des unités dans les langues germaniques a un petit effet sur le calcul, mais uniquement lorsque le stimulus ou la réponse sont présentés verbalement : aucun effet n'est observé lorsque le problème est présenté en chiffres arabes et que la réponse est tapée au clavier, ce qui montre bien qu'il ne s'agit pas vraiment d'une influence linguistique au niveau conceptuel.

En conclusion, le langage n'est pas le support de la pensée, mais il lui sert de catalyseur. Les mathématiques s'appuient principalement sur des représentations non linguistiques, sur un « langage de la pensée » purement interne et non verbal. Cependant, leur apprentissage peut être facilité par la structure de langue. Les interactions linguistiques attirent l'attention sur les objets mentaux les plus

pertinents. La numération en base dix, par exemple, est rendue transparente dans la langue orale chinoise – et peut-être serait-il intéressant d’expérimenter l’utilisation d’une numération pseudo-chinoise auprès des élèves français de maternelle, en leur faisant nommer temporairement les nombres « dix cinq » au lieu de « quinze » ou « deux dix » au lieu de « vingt » ? Même si ces effets de catalyse linguistique ne sont pas strictement « whorfiens », ils pourraient être importants sur le plan pratique. Dès la maternelle, l’enseignement devrait combiner (1) la manipulation d’objets non verbaux concrets, proto-mathématiques (collections, constructions, jeux de plateaux, formes, tangrams...), mais aussi (2) la verbalisation explicite qui rend évident les processus de pensée et donne à l’élève des objectifs d’apprentissage clairs et précis.

Bibliographie très succincte

- Chemla E., Dautriche I., Buccola B. et Fagot J., « Constraints on the lexicons of human languages have cognitive roots present in baboons (*Papio papio*) », *Proceedings of the National Academy of Sciences*, vol. 116, n° 30, 2019, p. 14926-14930, <https://doi.org/10.1073/pnas.1907023116>.
- Fedorenko E. et Varley R., « Language and thought are not the same thing: Evidence from neuroimaging and neurological patients », *Annals of the New York Academy of Sciences*, vol. 1369, n° 1, 2016, p. 132-153, <https://doi.org/10.1111/nyas.13046>.
- Freedman D.J., Riesenhuber M., Poggio T. et Miller E.K., « Categorical representation of visual stimuli in the primate prefrontal cortex », *Science*, vol. 291, 2001, art. 5502, p. 312-316, <https://doi.org/10.1126/science.291.5502.312>.
- Hermer-Vazquez L., Moffet A. et Munkholm P., « Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks », *Cognition*, vol. 79, n° 3, 2001, p. 263-299, [https://doi.org/10.1016/S0010-0277\(00\)00120-7](https://doi.org/10.1016/S0010-0277(00)00120-7).
- Hespos S.J. et Spelke E.S., « Conceptual precursors to language », *Nature*, vol. 430, 2004, p. 453-456, <https://doi.org/10.1038/nature02634>.
- Huguet P., Brunot S. et Monteil J.M., « Geometry versus drawing: Changing the meaning of the task as a means to change performance », *Social Psychology of Education*, vol. 4, n° 4, 2001, p. 219-234, <https://doi.org/10.1023/A:1011374700020>.
- Jack R.E., Garrod O.G., Yu H., Caldara R. et Schyns P.G., « Facial expressions of emotion are not culturally universal », *Proceedings of the National Academy of Sciences*, vol. 109, n° 19, 2012, p. 7241-7244, <https://doi.org/10.1073/pnas.1200155109>.
- Jackson J.C., Watts J., Henry T.R., List J.-M., Forkel R., Mucha P.J., Greenhill S.J., Gray R.D. et Lindquist K.A., « Emotion semantics show both cultural variation and universal structure », *Science*, vol. 366, 2019, art. 6472, p. 1517-1522, <https://doi.org/10.1126/science.aaw8160>.
- Li P. et Gleitman L., « Turning the tables: Language and spatial reasoning », *Cognition*, vol. 83, n° 3, 2002, p. 265-294, [https://doi.org/10.1016/s0010-0277\(02\)00009-4](https://doi.org/10.1016/s0010-0277(02)00009-4).
- Majid A., Bowerman M., Kita S., Haun D.B. et Levinson S.C., « Can language restructure cognition? The case for space », *Trends in Cognitive Sciences*, vol. 8, n° 3, 2004, p. 108-114, <https://doi.org/10.1016/j.tics.2004.01.003>.
- Pullum G.K., « The great Eskimo vocabulary hoax », *Natural Language & Linguistic Theory*, vol. 7, n° 2, 1989, p. 275-281.
- Regier T. et Kay P., « Language, thought, and color: Whorf was half right », *Trends in Cognitive Sciences*, vol. 13, n° 10, 2009, p. 439-446, <https://doi.org/10.1016/j.tics.2009.07.001>.

Regier T., Kay P. et Khetarpal N., « Color naming reflects optimal partitions of color space », *Proceedings of the National Academy of Sciences*, vol. 104, n° 4, 2007, p. 1436-1441. <https://doi.org/10.1073/pnas.0610341104>.

Romeo R.R., Leonard J.A., Robinson S.T., West M.R., Mackey A.P., Rowe M.L. et Gabrieli J.D.E., « Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function », *Psychological Science*, vol. 29, n° 5, 2018, p. 700-710. <https://doi.org/10.1177/0956797617742725>.

Sablé-Meyer M., Fagot J., Caparos S., van Kerkoerle T., Amalric M. et Dehaene S., « Sensitivity to geometric shape regularity in humans and baboons: A putative signature of human singularity », *Proceedings of the National Academy of Sciences*, vol. 118, n° 16, 2021, <https://doi.org/10.1073/pnas.2023123118>.

Winawer J., Witthoft N., Frank M.C., Wu L., Wade A.R. et Boroditsky L., « Russian blues reveal effects of language on color discrimination », *Proceedings of the National Academy of Sciences*, vol. 104, n° 19, 2007, p. 7780-7785, <https://doi.org/10.1073/pnas.0701644104>.

COLLOQUE - THE REPRESENTATION OF LANGUAGE IN BRAINS AND MACHINES

Colloque coorganisé avec les professeurs Luigi Rizzi (chaire Linguistique générale) et Stéphane Mallat (chaire Science des données), les 25 et 26 juin 2021.

Ce colloque a porté sur la convergence et la divergence entre les approches computationnelles, neuroscientifiques et formelles du langage. La dernière décennie a été marquée par des avancées majeures dans les approches computationnelles du langage, conduisant à des applications pratiques. Dans le même temps, les neurosciences cognitives ont fait des progrès significatifs dans la compréhension des mécanismes de traitement du langage dans le cerveau, et la linguistique formelle a continué à faire des progrès constants dans la description structurelle du langage. Les trois approches ont suivi des voies largement indépendantes, ce qui n'est guère surprenant, étant donné les différences substantielles dans leurs méthodologies et leurs objectifs. Néanmoins, on aurait pu s'attendre à un niveau d'intégration plus élevé, puisqu'elles traitent toutes du même objet, le langage naturel. Cela ne s'est produit que dans une mesure limitée. Plus précisément, alors qu'il existe plusieurs exemples d'intégration des approches neuroscientifiques et linguistiques, les approches informatiques ont été largement élaborées de manière isolée des deux autres perspectives. S'agit-il d'une conséquence inévitable des différences d'objectifs et de méthodes, ou d'une situation temporaire, liée à l'histoire de ces domaines ? Peut-on envisager, dans les années à venir, une convergence plus grande vers un « modèle standard » du langage et des langues ? Nous avons abordé ces questions avec un groupe sélectionné de linguistes, de psychologues, de spécialistes du cerveau, et de spécialistes de l'informatique.

Programme :

- Introduction par Stanislas Dehaene, Stéphane Mallat et Luigi Rizzi ;

- Philippe Schlenker (ENS, Paris/NYU) : « *Grammatical inferences without words* » ;
- Emmanuel Chemla (ENS, Paris) : « *A linguistic learning bias found in humans, animals, and artificial learners* » ;
- Anne Christophe (ENS, Paris) : « *Bootstrapping the syntactic bootstrapper* » ;
- Naama Friedmann (Tel Aviv University) : « *Developmental syntactic disorders and what they say about the critical period* » ;
- Kyunghyun Cho (New York University) : « *Few-shot learning is still difficult with large-scale language models* » ;
- Tomas Mikolov (FAIR, Paris) : « *Neural language models: The successes, and the challenges* » ;
- Charles Yang (University of Pennsylvania) : « *Why children and machines learn differently* » ;
- Roger Levy (MIT) : « *Grammatical generalization and language processing in humans and machines* » ;
- Emmanuel Dupoux (EHESS/FAIR, Paris) : « *Simulating the emergence of linguistic categories in infants* » ;
- Marco Baroni (ICREA, Barcelone/FAIR, Paris) : « *On the proper role of linguistically-oriented deep net analysis in linguistic theorizing* » ;
- Andrea Moro et Cristiano Chesì (IUSS, Pavie) : « *The limits of competence and processing: The case of copular sentences* » ;
- Paola Merlo (University of Geneva) : « *Beyond the Benchmarks: Linguistically-informed notions of locality and similarity in distributed spaces* » ;
- Gennaro Chierchia (Harvard University) : « *Quantifiers in natural language: The interface between syntax and logic* » ;
- Benjamin Spector (ENS, Paris) : « *Logic, grammar and distribution: The case of polarity items* » ;
- Christophe Pallier (CNRS, Paris) : « *Probing syntax and semantics in the brain* » ;
- William Matchin (University of South Carolina) : « *The cortical organization of syntax: A minimalist perspective* » ;
- Stanislas Dehaene (Collège de France) : « *Why humans still trump machines: The language of thought hypothesis* » .

COURS À L'EXTÉRIEUR - *ADVANCES IN UNDERSTANDING
THE SINGULARITY OF THE HUMAN BRAIN*

Une série de trois cours a été donnée à l'université de Buenos Aires (UBA, Argentine), les 19, 20 et 21 novembre 2019 sous le titre « *Advances in understanding the singularity of the human brain* » :

- « *How we learn: The 4 pillars of efficient learning* » ;
- « *Reading acquisition and neuronal recycling* » ;
- « *From single words to sentences: Syntax in the brain* ».

RECHERCHE

L'équipe de recherche sur « les langages du cerveau », financée par l'European Research Council (programme NeuroSyntax) et comprenant notamment Fosca Al Roumi, Yair Lakretz, Samuel Planton et Mathias Sablé-Meyer, a poursuivi ses travaux aussi bien sur le plan comportemental qu'avec l'IRM fonctionnelle (IRMf) et la magnéto-encéphalographie (MEG). Rappelons que cette recherche teste l'hypothèse que le cerveau humain ne contient pas seulement des circuits dédiés au langage naturel, parlé ou écrit, mais également d'autres circuits qui encodent un langage de la pensée et sont capables de coder, de façon récursive, des régularités mathématiques, musicales, visuelles, etc. Cette faculté de représenter les informations sous forme d'expressions mentales enchâssées serait unique à l'espèce humaine.

Nous avons déjà montré, par les recherches comportementales de Marie Amalric et celles que Liping Wang a menées en IRMf, l'existence d'un langage de la géométrie qui code les régularités récursives des séquences spatiales. Fosca Al Roumi a prolongé ce travail en MEG, dont la résolution temporelle a permis de suivre le décours temporel de la représentation mentale d'une séquence visuo-spatiale, et de démontrer le codage cérébral des primitives géométriques et surtout des structures de groupe postulées par notre langage. Samuel Planton a ensuite montré que ce langage, inchangé, pouvait se généraliser à la modalité auditive et expliquer les régularités que les humains perçoivent dans des séquences binaires de deux sons. Fosca et Samuel sont actuellement en train d'étendre ce travail aux mécanismes cérébraux des représentations des séquences auditives, à l'aide de l'IRMf et de la MEG.

Mathias Sablé-Meyer, pour sa thèse, a étendu ces idées à la perception de formes géométriques statiques. Nous avons créé un test de recherche d'intrus dans lequel les participants devaient détecter une forme déviante parmi 5 répétitions de la même forme de base. Par exemple, la forme de base pouvait être un rectangle (avec des variations de taille et d'orientation) et la forme déviante, le même rectangle avec le coin inférieur droit déplacé. Nous avons testé 11 formes de base, toutes des quadrilatères, dont la régularité géométrique variait. Un énorme effet de régularité

géométrique est apparu : les intrus étaient plus faciles à détecter lorsque la forme de base possédait des régularités géométriques (côtés parallèles, côtés égaux, angles droits ou angles égaux). La complexité de la forme était prédite par le nombre total de telles propriétés symboliques. Ainsi, les carrés, rectangles, trapèzes ou parallélogrammes, qui possèdent des régularités compressibles, sont beaucoup plus faciles à représenter mentalement que des quadrilatères quelconques dépourvus de ces propriétés. Une série d'expériences a démontré que cet effet de régularité géométrique était hautement reproductible et qu'il se retrouvait, par exemple dans une tâche classique de recherche sérielle, ou même lorsque les emplacements des coins étaient présentés dans une séquence, ce qui fait le lien avec nos travaux précédents sur les séquences visuo-spatiales. Il est important de noter que l'effet est retrouvé chez des enfants d'âge préscolaire et chez des adultes sans éducation (les Himba de Namibie), ce qui suggère qu'il est largement indépendant de l'éducation. Les études IRMf en cours suggèrent que l'effet pourrait provenir du cortex intrapariétal antérieur droit.

Nos résultats montrent également que les réseaux de neurones artificiels qui dominent actuellement le domaine de l'intelligence artificielle sont incapables d'expliquer ces aspects mêmes les plus élémentaires de la représentation des connaissances géométriques ou musicales chez l'Homme, parce qu'ils ne disposent pas d'un mode de représentation symbolique et capable de composition récursive. Les travaux de Yair Lakretz posent la même question dans le domaine linguistique : les succès actuels des réseaux de neurones artificiels en traitement du langage (*transformers*, GPT2 et GPT3, etc.) sont-ils bien réels ? Sont-ils capables des mêmes performances qu'un être humain ? À travers plusieurs expériences, nous avons découvert que la réponse est négative. Par exemple, ces réseaux ne disposent pas de mécanisme satisfaisant capable d'accorder un verbe avec son sujet en présence d'autres noms qui interfèrent avec le sujet (par exemple : « Le patron qui récompense ses employés souhaite une augmentation. »). Dans le cadre de la thèse de Christos Zacharopoulos ainsi que dans des travaux en collaboration avec Marco Baroni, Yair Lakretz a étudié les mécanismes d'accord à longue distance chez l'Homme et dans les réseaux de neurones artificiels, et a montré des différences profondes : la représentation humaine du langage résiste aisément à la présence de multiples structures enchâssées, alors que ce n'est pas (ou peu) le cas des algorithmes actuels.

Notre équipe s'est également attachée à tester l'hypothèse que ces représentations enchâssées, ces langages de la pensée, sont spécifiques à l'espèce humaine. Des tests de séquences visuo-spatiales ou auditives, et d'autres faisant appel à des formes géométriques, d'abord développés chez l'Homme, ont été ensuite proposés aux primates non humains, soit à NeuroSpin, en collaboration avec les équipes de Bechir Jarraya, Fanis Panagiotaropoulos et Timo van Kerkoerle, soit à l'extérieur, notamment avec Liping Wang à l'université de Shanghai (singes macaques), avec Joël Fagot à l'université d'Aix-Marseille (babouins) et avec Jared Tagliatella à Kennesaw State University (chimpanzés).

Les résultats actuels, encore fragmentaires, soulignent les limites de la cognition animale, qui semble confinée à la représentation de traits abstraits élémentaires. Les primates non humains ne semblent pas capables d'apprendre à combiner plusieurs informations pour former des représentations syntaxiquement complexes. Le test de l'intrus géométrique, par exemple, a été réalisé de manière identique chez les primates humains et non humains. En effet, un babouin peut facilement apprendre à utiliser un écran tactile et à toucher la forme, parmi six, qui diffère des autres. Avec Joël Fagot, nous avons entraîné 20 babouins à effectuer cette tâche avec une variété de formes non géométriques (fleurs, taches de couleur, etc.), puis nous avons examiné le transfert de leur apprentissage à des quadrilatères. Les babouins montreraient-ils, comme les humains, un effet de régularité géométrique ? La réponse est non : les babouins sont insensibles à la régularité géométrique, que ce soit immédiatement après le transfert ou après un entraînement intensif avec les formes géométriques. Contrairement aux humains, ils ne traitent pas les carrés et les rectangles différemment des autres quadrilatères irréguliers, et le modèle d'un langage symbolique ne rend absolument aucun compte de la variance de leurs réponses. Au contraire, leur comportement est bien prédit par les modèles classiques de réseaux convolutifs de la voie visuelle ventrale, qui, à l'inverse, sont largement inopérants chez les humains éduqués. Il est intéressant de noter qu'un mélange des deux modèles rend le mieux compte des performances des enfants d'âge préscolaire et des adultes non éduqués. Ainsi, deux stratégies sont disponibles pour résoudre la tâche de l'intrus géométrique : une stratégie perceptive, disponible chez tous les primates, dans laquelle les formes géométriques sont traitées dans le système visuel ventral comme n'importe quelle image ou n'importe quel visage ; et une stratégie symbolique, apparemment disponible uniquement chez les humains, dans laquelle les formes géométriques sont comprimées selon les « répétitions de répétitions » (symétries) discrètes et symboliques. Ce travail important a été publié dans *PNAS*.

Outre ces travaux de recherche fondamentale, en tant que président du Conseil scientifique de l'Éducation nationale (CSEN), j'ai également poursuivi mon travail de lancement de recherches appliquées à l'éducation, et d'écriture de synthèses de la littérature scientifique qui soient utiles aux enseignants. Une recherche a, par exemple, permis d'évaluer les capacités de lecture des élèves en début de sixième. Une seconde recherche, visant à évaluer leurs compétences dans le domaine des nombres décimaux et des fractions, est en cours de finalisation. Une partie importante de la thèse de Pauline Martinot est également consacrée à l'analyse des données longitudinales du programme EvalAide (évaluations nationales des élèves en début de CP, milieu de CP et début de CE1, en collaboration avec la DEPP) ; elle a fait l'objet d'une note en français à destination de l'Éducation nationale, et devrait faire l'objet d'une publication dans une grande revue. Un grand colloque international a été coorganisé avec Yann Algan sur le thème « Quels professeurs au XXI^e siècle ? ». Tous ces travaux et bien d'autres, dont les textes sont disponibles gratuitement sur le site du CSEN (<https://www.reseau-canope.fr/conseil-scientifique-de-leducation->

nationale.html), ont été en partie regroupés dans un premier livre intitulé *La Science au service de l'école*, paru en 2019 aux Éditions Odile Jacob, et dans un second livre intitulé *L'École éclairée par la science*, paru en 2021.

PUBLICATIONS

Al Roumi F., Marti S., Wang L., Amalric M. et Dehaene S., « Mental compression of spatial sequences in human working memory using numerical and geometrical primitives », *Neuron*, vol. 109, n° 16, 2021, p. 2627-2639.e4, <https://doi.org/10.1016/j.neuron.2021.06.009>.

Algan Y., Dehaene S., Huillery E., Pasquinelli E. et Ramus F., *Quels professeurs au XXI^e siècle ? Rapport de synthèse du Conseil scientifique de l'Éducation nationale*, Paris, Conseil scientifique de l'éducation nationale, 2021, https://www.reseau-canope.fr/fileadmin/user_upload/Projets/conseil_scientifique_education_nationale/Rapport_scientifique_Grenelle_de_l_education.pdf.

Berkovitch L., Charles L., Cul A.D., Hamdani N., Delavest M., Sarrazin S., Mangin J.-F., Guevara P., Ji E., d'Albis M.-A., Gaillard R., Bellivier F., Poupon C., Leboyer M., Tamouza R., Dehaene S. et Houenou J., « Disruption of conscious access in psychosis is associated with altered structural brain connectivity », *Journal of Neuroscience*, vol. 41, n° 3, 2021, p. 513-523, <https://doi.org/10.1523/JNEUROSCI.0945-20.2020>.

Chang C.H.C., Dehaene S., Wu D.H., Kuo W.-J. et Pallier C., « Cortical encoding of linguistic constituent with and without morphosyntactic cues », *Cortex*, vol. 129, 2020, p. 281-295. <https://doi.org/10.1016/j.cortex.2020.04.024>.

Ciccione L. et Dehaene S., « Grouping mechanisms in numerosity perception », *Open Mind*, vol. 4, 2020, p. 102-118, https://doi.org/10.1162/opmi_a_00037.

Ciccione L. et Dehaene S., « Can humans perform mental regression on a graph ? Accuracy and bias in the perception of scatterplots », *Cognitive Psychology*, vol. 128, 2021, art. 101406, <https://doi.org/10.1016/j.cogpsych.2021.101406>.

Cohen L., Salondy P., Pallier C. et Dehaene S., « How does inattention affect written and spoken language processing? », *Cortex*, vol. 138, 2021, p. 212-227, <https://doi.org/10.1016/j.cortex.2021.02.007>.

Dehaene S., *How We Learn: Why Brains Learn Better Than Any Machine... For Now*, New York, Penguin Books, 2021.

Dehaene S. (dir.), *L'École éclairée par la science*, Paris, Odile Jacob, 2021.

Dotan D. et Dehaene S., « Parallel and serial processes in number-to-quantity conversion », *Cognition*, vol. 204, 2020, art. 104387, <https://doi.org/10.1016/j.cognition.2020.104387>.

Feng X., Altarelli I., Monzalvo K., Ding G., Ramus F., Shu H., Dehaene S., Meng X. et Dehaene-Lambertz G., « A universal reading network and its modulation by writing system and reading ability in French and Chinese children », *ELife*, vol. 9, 2020, art. e54591, <https://doi.org/10.7554/eLife.54591>.

Dehaene S., Gui P., Jiang Y., Zang D., Qi Z., Tan J., Tanigawa H., Jiang J., Wen Y., Xu L., Zhao J., Mao Y., Poo M., Ding N., Wu X. et Wang L., « Assessing the depth of language processing in patients with disorders of consciousness », *Nature Neuroscience*, vol. 23, n° 6, 2020, p. 761-770, <https://doi.org/10.1038/s41593-020-0639-1>.

Heimler B., Behor T., Dehaene S., Izard V. et Amedi A., « Core knowledge of geometry can develop independently of visual experience », *Cognition*, vol. 212, 2021, art. 104716, <https://doi.org/10.1016/j.cognition.2021.104716>.

Lakretz Y., Dehaene S. et King J.-R., « What limits our capacity to process nested long-range dependencies in sentence comprehension ? », *Entropy*, vol. 22, n° 4, 2020, p. 446, <https://doi.org/10.3390/e22040446>.

Lakretz Y., Hupkes D., Vergallito A., Marelli M., Baroni M. et Dehaene S., « Mechanisms for handling nested dependencies in neural-network language models and humans », *Cognition*, vol. 213, 2021, art. 104699, <https://doi.org/10.1016/j.cognition.2021.104699>.

López-Barroso D., Thiebaut de Schotten M., Morais J., Kolinsky R., Braga L.W., Guerreiro-Tauil A., Dehaene S. et Cohen L., « Impact of literacy on the functional connectivity of vision and language related networks », *NeuroImage*, vol. 213, 2020, art. 116722, <https://doi.org/10.1016/j.neuroimage.2020.116722>.

Malassis R., Dehaene S. et Fagot J., « Baboons (*Papio papio*) process a context-free but not a context-sensitive grammar », *Scientific Reports*, vol. 10, n° 1, 2020, art. 7381, <https://doi.org/10.1038/s41598-020-64244-5>.

Mashour G.A., Roelfsema P., Changeux J.-P. et Dehaene S., « Conscious processing and the global neuronal workspace hypothesis », *Neuron*, vol. 105, n° 5, 2020, p. 776-798, <https://doi.org/10.1016/j.neuron.2020.01.026>.

Milham M., Petkov C.I., Margulies D.S., Schroeder C.E., Basso M.A., Belin P., Fair D.A., Fox A., Kastner S., Mars R.B., Messinger A., Poirier C., Vanduffel W., Essen D.C.V., Alvand A., Becker Y., Hamed S.B., Benn A., Bodin C., Boretius S., Cagna B., Coulon O., El-Gohary S.H., Evrard H., Forkel S.J., Friedrich P., Froudust-Walsh S., Garza-Villarreal E.A., Gao Y., Gozzi A., Grigis A., Hartig R., Hayashi T., Heuer K., Howells H., Ardesch D.J., Jarraya B., Jarrett W., Jedema H.P., Kagan I., Kelly C., Kennedy H., Klink P.C., Kwok S.C., Leech R., Liu X., Madan C., Madushanka W., Majka P., Mallon A.-M., Marche K., Meguerditchian A., Menon R.S., Merchant H., Mitchell A., Nenning K.-H., Nikolaidis A., Ortiz-Rios M., Pagani M., Pareek V., Prescott M., Procyk E., Rajimehr R., Rautu I.-S., Raz A., Roe A.W., Rossi-Pool R., Roumazeilles L., Sakai T., Sallet J., García-Saldivar P., Sato C., Sawiak S., Schiffer M., Schwiedrzik C.M., Seidlitz J., Sein J., Shen Z., Shmuel A., Silva A.C., Simone L., Sirmmpilatze N., Sliwa J., Smallwood J., Tasserie J., Schotten M.T. de, Toro R., Trapeau R., Uhrig L., Vezoli J., Wang Z., Wells S., Williams B., Xu T., Xu A.G., Yacoub E., Zhan M., Ai L., Amiez C., Balezau F., Baxter M.G., Blezer E.L.A., Brochier T., Chen A., Crosson P.L., Damatac C.G., Dehaene S., Everling S., Fleysher L., Freiwald W., Griffiths T.D., Guedj C., Hadj-Bouziane F., Harel N., Hiba B., Jung B., Koo B., Laland K.N., Leopold D.A., Lindenfors P., Meunier M., Mok K., Morrison J.H., Nacef J., Nagy J., Pinsk M., Reader S.M., Roelfsema P.R., Rudko D.A., Rushworth M.F.S., Russ B.E., Schmid M.C., Sullivan E.L., Thiele A., Todorov O.S., Tsao D., Ungerleider L., Wilson C.R.E., Ye F.Q., Zarco W. et Zhou Y., « Accelerating the evolution of nonhuman primate neuroimaging », *Neuron*, vol. 105, n° 4, 2020, p. 600-603, <https://doi.org/10.1016/j.neuron.2019.12.023>.

Panagiotaropoulos T.I., Wang L. et Dehaene S., « Hierarchical architecture of conscious processing and subjective experience », *Cognitive Neuropsychology*, vol. 37, n° 3-4, 2020, p. 180-183, <https://doi.org/10.1080/02643294.2020.1760811>.

Pinho A.L., Amadon A., Fabre M., Dohmatob E., Denghien I., Torre J.J., Ginisty C., Becuwe-Desmidt S., Roger S., Laurier L., Joly-Testault V., Médiouni-Cloarec G., Doublé C., Martins B., Pinel P., Eger E., Varoquaux G., Pallier C., Dehaene S., Hertz-Pannier L. et Thirion B., « Subject-specific segregation of functional territories based on deep phenotyping », *Human Brain Mapping*, vol. 42, n° 4, 2021, p. 841-870, <https://doi.org/10.1002/hbm.25189>.

Pinho A.L., Amadon A., Gauthier B., Clairis N., Knops A., Genon S., Dohmatob E., Torre J.J., Ginisty C., Becuwe-Desmidt S., Roger S., Lecomte Y., Berland V., Laurier L., Joly-Testault V., Médiouni-Cloarec G., Doublé C., Martins B., Salmon E., Piazza M., Melcher D., Pessiglione M., van Wassenhove V., Eger E., Varoquaux G., Dehaene S., Hertz-Pannier L.

et Thirion B., « Individual brain charting dataset extension, second release of high-resolution fMRI data for cognitive mapping », *Scientific Data*, vol. 7, n° 1, 2020, art. 353, <https://doi.org/10.1038/s41597-020-00670-4>.

Planton S. et Dehaene S., « Cerebral representation of sequence patterns across multiple presentation formats », *Cortex*, vol. 145, 2021, p. 13-36, <https://doi.org/10.1016/j.cortex.2021.09.003>.

Planton S., van Kerkoerle T., Abbih L., Maheu M., Meyniel F., Sigman M., Wang L., Figueira S., Romano S. et Dehaene S., « A theory of memory for binary sequences : Evidence for a mental compression algorithm in humans », *PLOS Computational Biology*, vol. 17, n° 1, 2021, art. e1008598, [10.1371/journal.pcbi.1008598](https://doi.org/10.1371/journal.pcbi.1008598).

Rajalingham R., Kar K., Sanghavi S., Dehaene S. et DiCarlo J.J., « The inferior temporal cortex is a potential cortical precursor of orthographic processing in untrained monkeys », *Nature Communications*, vol. 11, n° 1, 2020, art. 3886, <https://doi.org/10.1038/s41467-020-17714-3>.

Sablé-Meyer M., Ellis K., Tenenbaum J. et Dehaene S., « A language of thought for the mental representation of geometric shapes », *PsyArXiv*, 2021, <https://doi.org/10.31234/osf.io/28mg4>.

Sablé-Meyer M., Fagot J., Caparos S., van Kerkoerle T., Amalric M. et Dehaene S., « Sensitivity to geometric shape regularity in humans and baboons: A putative signature of human singularity », *Proceedings of the National Academy of Sciences*, vol. 118, n° 16, 2021, <https://doi.org/10.1073/pnas.2023123118>.

Potier Watkins C., Caporal J., Merville C., Kouider S. et Dehaene S., « Accelerating reading acquisition and boosting comprehension with a cognitive science-based tablet training », *Journal of Computers in Education*, vol. 7, 2020, p. 183-212, <https://doi.org/10.1007/s40692-019-00152-6>.

Watkins C.P. et Dehaene S., « Can a game application that boosts phonics knowledge in kindergarten advance 1st grade reading? », *PsyArXiv*, 2021, <https://doi.org/10.31234/osf.io/pwumg>.

Woolnough O., Donos C., Rollo P.S., Forseth K.J., Lakretz Y., Crone N.E., Fischer-Baum S., Dehaene S. et Tandon N., « Spatiotemporal dynamics of orthographic and lexical processing in the ventral visual pathway », *Nature Human Behaviour*, vol. 5, n° 3, 2021, p. 389-398, <https://doi.org/10.1038/s41562-020-00982-w>.

Zhang H., Zhen Y., Yu S., Long T., Zhang B., Jiang X., Li J., Fang W., Sigman M., Dehaene S. et Wang L., « Working memory for spatial sequences: Developmental and evolutionary factors in encoding ordinal and relational structures », *Journal of Neuroscience*, 2021, <https://doi.org/10.1523/JNEUROSCI.0603-21.2021>.

