

## **Psychologie cognitive expérimentale**

M. Stanislas DEHAENE, membre de l'Institut  
(Académie des sciences), professeur

### ENSEIGNEMENT

#### **Cours : Le bébé statisticien**

Dans la continuité de l'année précédente, le cours a abordé une hypothèse qui fait actuellement l'objet d'intenses explorations théoriques et expérimentales en sciences cognitives : l'idée que le cerveau humain et animal contient des mécanismes d'inférence statistique approchant les équations normatives de l'inférence bayésienne<sup>a</sup>.

Dans le cours précédent, nous avons vu que cette théorie mathématique simple, qui caractérise le raisonnement plausible en présence d'incertitudes, rend compte d'une grande variété d'observations psychologiques et physiologiques. Lorsque notre cerveau reçoit des entrées ambiguës, il semble en reconstruire l'interprétation la plus probable. Cette inférence est hiérarchique et donne accès à des connaissances abstraites. La prise de décision pourrait résulter d'une combinaison de ce calcul bayésien des probabilités avec une estimation des conséquences de nos choix. L'architecture du cortex pourrait avoir évolué pour réaliser, à grande vitesse et de façon massivement parallèle, des inférences bayésiennes. L'algorithme utilisé pourrait expliquer la manière dont notre cerveau anticipe sur le monde extérieur et dont il répond à la nouveauté.

L'objectif du cours de cette année était d'explorer l'hypothèse que tous ces éléments sont présents chez le très jeune enfant, dès la première année de vie et peut-être dès la naissance. En effet, le bébé humain semble doté de compétences pour le raisonnement probabiliste. Le cerveau de l'enfant émet des prédictions sur le monde extérieur et semble disposer d'un puissant algorithme d'apprentissage de régularités statistiques. L'apprentissage des visages, des objets, des mots ou des règles linguistiques pourrait-il s'expliquer par une inférence bayésienne ?

---

a. Les enregistrements des cours sont disponibles en audio et en vidéo sur le site Internet du Collège de France : <http://www.college-de-france.fr/site/stanislas-dehaene/course-2012-2013.htm> [Ndlr].

### *Des scientifiques au berceau ?*

Les sciences cognitives contemporaines ont réfuté la vision de William James (1890) selon laquelle le bébé « assailli par les yeux, les oreilles, le nez, la peau et les entrailles, ressent tout cela comme une vaste et bruyante confusion ». Cette réfutation s'est produite en deux temps.

Dans un premier temps, depuis les années 1970, des centaines d'expériences ont mis en évidence les multiples compétences du bébé. Celui-ci dispose, dès la première année de vie, d'un noyau de connaissances fondamentales (*core knowledge*), tels que les concepts d'objet (Spelke, 1983), de nombre (Mehler & Bever, 1967 ; Wynn, 1992), d'espace (Landau *et al.*, 1981), ou encore un éventail de phonèmes et de règles pertinentes qui facilitent l'apprentissage du langage (Chomsky, Gleitman, Mehler).

Sur cette base, dans un second temps, s'est développée la métaphore du bébé comme un scientifique ou un détective. Selon Gopnik, les observations que fait l'enfant seraient intégrées à des « théories intuitives du monde physique, biologique et psychologique. Ces théories, comme des théories scientifiques, sont des représentations complexes, cohérentes et abstraites de la structure causale du monde extérieur » (Gopnik & Schulz, 2004). Le bébé élaborerait des modèles mentaux du monde et évaluerait leur plausibilité vis-à-vis des observations qu'il fait. Il disposerait donc de compétences précoces pour la manipulation des probabilités, l'évaluation, en parallèle, de nombreux modèles, la sélection des variables pertinentes, l'élimination des variables de non-intérêt, et le repérage des ambiguïtés ou des interprétations multiples.

Ces hypothèses fortes ont suscité un nouveau mouvement expérimental visant à tester la présence d'une intuition des probabilités et de capacités d'inférence chez le très jeune enfant. Ces recherches ont été couronnées de succès. Par exemple, dès huit mois de vie, un bébé est capable d'anticiper le résultat approximatif d'un tirage aléatoire d'une urne. Plus surprenant encore, il est également capable d'inférer, en sens inverse, le contenu probable d'une urne après quelques tirages, ce qui constitue le fondement même de l'inférence bayésienne. À seize mois, les expériences de Laura Schulz montrent qu'il fait déjà toute une chaîne d'inférences hiérarchiques à partir d'un tout petit ensemble d'observations. Par exemple, s'il ne parvient pas à faire fonctionner un objet, il déduit des circonstances précises de son échec si c'est son action ou l'objet lui-même qui est en cause (Gweon & Schulz, 2011). Quand un tirage dévie de l'échantillonnage aléatoire, il attribue à la personne qui puise dans l'urne une intention précise (Gweon, Tenenbaum, & Schulz, 2010).

Dès la seconde année de vie, l'enfant semble donc tirer de ses observations des modèles mentaux, en tenant compte des intentions des personnes qui agissent autour de lui, et en utilisant des inférences bayésiennes pour remonter aux causes cachées de leurs actions. La métaphore d'un « détective » ou d'un « scientifique au berceau » ne paraît pas inappropriée. Cependant, cette sagacité n'a rien de mystérieux, car il est possible d'en modéliser mathématiquement le fonctionnement. Divers algorithmes permettent d'approcher les règles de l'inférence bayésienne et pourraient être utilisés par l'enfant (T. D. Ullman, Goodman, & Tenenbaum, 2012).

L'approche bayésienne réconcilie nativisme et constructivisme. En effet, elle nécessite, d'une part, de postuler un grand nombre de mécanismes fondamentaux innés (sens de probabilités, raisonnement probabiliste avec des distributions, vocabulaire de concepts de nombre, d'espace, etc.). D'autre part, elle démontre que, sur cette base, en s'appuyant sur un mécanisme générique d'inférence bayésienne,

des principes extrêmement abstraits peuvent être inférés à partir d'interactions avec le monde extérieur. Ainsi, même la notion de causalité ne doit pas être postulée d'emblée – le bébé pourrait découvrir que cette hypothèse rend compte, mieux que d'autres, de ce qu'il observe (Goodman, Ullman, & Tenenbaum, 2011).

### *Le sens des probabilités chez l'enfant*

Le bébé dispose-t-il déjà d'un sens des probabilités qui lui permette de se représenter des distributions de probabilité, de les mettre à jour en appliquant la règle de Bayes, et de les utiliser pour générer des prédictions qu'il compare aux données reçues du monde extérieur ?

Pour Jean Piaget et Bärbel Inhelder (*La genèse de l'idée de hasard chez l'enfant*, 1951), la réponse ne peut être que négative. L'enfant ne peut comprendre l'irréversibilité du hasard avant le stade des « opérations réversibles ». Selon Piaget, ce ne serait que vers 7-11 ans qu'apparaîtrait l'idée du hasard, « mais les opérations élémentaires d'ordre (ordonner une suite ABC...) sont loin de suffire à la constitution d'un schème complet de permutations proprement dites ». Cette maîtrise apparaîtrait vers 12 ans.

Piaget et Inhelder s'appuient sur une série d'expériences dans lesquels les enfants de 4-5 ans répondent effectivement sans tenir compte des probabilités. Mais leur erreur consiste à s'appuyer exclusivement sur l'introspection verbale des enfants (ou leurs dessins), plutôt que sur les connaissances implicites que l'on peut déceler dans leur regard, leur surprise, ou leur choix spontané. À l'aide de ces techniques, les sciences cognitives contemporaines démontrent que les bébés sont compétents pour le calcul probabiliste. Un bébé de douze mois, qui voit pour la première fois une urne contenant trois objets bleus et un jaune, manifeste sa surprise lorsqu'un tirage aléatoire voit sortir l'objet jaune plutôt que l'un des bleus. Sa surprise est directement proportionnelle à l'improbabilité du résultat observé (Teglas *et al.*, 2011). Sa capacité d'anticipation est probabiliste, mais elle n'est pas fréquentielle. Selon Luca Bonatti, il n'a pas besoin d'observer de nombreux tirages pour en calculer les probabilités : sa déduction probabiliste se fonde exclusivement sur un modèle interne de la logique de la situation. Ce modèle interne combine des sources d'informations spatiales, temporelles et probabilistes. Le bébé réalise vraisemblablement une simulation mentale probabiliste, en s'appuyant sur les principes de solidité et de continuité spatio-temporelle des objets physiques. L'hypothèse d'une simulation interne rend compte de toute une série d'expériences sur le sens des probabilités, le concept d'objet, l'occlusion et le sens du nombre chez le bébé (Teglas *et al.*, 2011).

Selon cette perspective, le cerveau du bébé projette en permanence sur le monde extérieur des interprétations abstraites fondées sur un ou plusieurs modèles internes en compétition. Cet échantillonnage constant des modèles plausibles pourrait-il expliquer l'activité spontanée et stochastique du cerveau, visible dès la naissance sous forme d'un électro-encéphalogramme fluctuant ? Josef Fiser et ses collaborateurs testent cette hypothèse chez le furet nouveau-né (Berkes, Orban, Lengyel & Fiser, 2011). À l'aide de 16 électrodes placées dans le cortex visuel, ils enregistrent l'activité spontanée ou évoquée, soit par des scènes naturelles, soit par des stimuli artificiels (grilles) qui ne respectent pas les statistiques du monde naturel. L'analyse de la distribution de l'activité neuronale spontanée et évoquée montre qu'au cours du développement, l'activité spontanée se met progressivement à ressembler à l'activité évoquée par des stimuli naturels. Ainsi, le cerveau du furet internaliserait un modèle stochastique des transitions spatio-temporelles présentes

dans le monde naturel, et le réactiverait en l'absence d'entrées extérieures. À l'avenir, il sera important d'étendre ces expériences au bébé humain : l'examen de l'activité corticale spontanée, dans les circuits visuels ou linguistiques, pourrait fournir une nouvelle manière d'évaluer les modèles mentaux de l'enfant.

### *L'apprentissage de régularités statistiques*

L'imagerie cérébrale du très jeune enfant a été utilisée pour évaluer sa capacité à projeter des anticipations sur le monde extérieur et à émettre des signaux d'erreur lorsque ces prédictions sont violées. Lorsqu'on présente, à un bébé de quelques mois, une série de sons « ba ba ba ba » suivie d'un son identique (ba) ou déviant (ga), le cerveau de l'enfant, comme celui de l'adulte, réagit en émettant un potentiel évoqué spécifique des stimuli nouveaux (*mismatch response*). Cette réponse, qui traduit la détection d'une déviation d'une régularité auditive, est présente même chez le nouveau-né endormi, et chez le prématuré de six mois et demi de gestation. Elle traduit donc un calcul élémentaire et probablement inconscient des probabilités.

Les régularités auditives auxquelles les bébés sont sensibles ne sont pas triviales, mais apportent la preuve d'une hiérarchie d'inférences abstraites. Après avoir été exposé, de façon répétée, à la vision d'une bouche muette, mais dont le mouvement articule clairement le son « a », le cortex auditif du bébé de dix semaines répond à la nouveauté *auditive* du son « i » par rapport au son « a » (*Bristow et al.*, 2009). Des régions distinctes de son cerveau répondent au changement de voix et de phonème. Le bébé répond également, de manière distincte, aux changements locaux et globaux dans une séquence de sons – un test supposé mesurer l'accès des informations auditives à la conscience. Enfin, face à un visage en rotation, l'extraction de régularités statistiques permet à l'enfant d'apprendre que le profil et la vue de face appartiennent au même visage. Shimon Ullman montre qu'un algorithme simple, exposé à une heure de vidéos, extrait automatiquement la présence de mains et de visages (S. Ullman, Harari & Dorfman, 2012).

En conclusion, ces données suggèrent que même les très jeunes enfants disposent déjà de mécanismes prédictifs qui internalisent les relations statistiques spatiales et temporelles entre les entrées sensorielles qu'ils reçoivent, et les utilisent pour prédire les entrées futures et pour réagir avec surprise lorsque ces prédictions sont violées. L'extraction de telles régularités statistiques, au sein d'un réseau cortical hiérarchique, permet d'apprendre, sans aucune supervision, à détecter des catégories comme les visages et les mains, et à les reconnaître de façon invariante. Ainsi le très jeune enfant apprendrait-il à décomposer le monde visuel et auditif afin d'en découvrir les objets les plus saillants.

### *La découverte et l'apprentissage des mots*

L'apprentissage du langage relève-t-il d'une statistique similaire ? Sur le plan prosodique et phonologique, le bébé apprend extrêmement rapidement les propriétés de sa langue maternelle. Même les nouveau-nés de quelques jours préfèrent écouter des phrases dans leur langue maternelle – un apprentissage qui repose principalement sur les informations prosodiques lentes et de basse fréquence, et pourrait débiter *in utero*. Entre six et douze mois, les bébés détectent les catégories phonétiques de leur langue, et cessent de distinguer les contrastes qui ne sont pas pertinents pour eux (par exemple le contraste entre « r » et « l » pour des bébés japonais). Cet

apprentissage repose clairement sur une estimation statistique de la distribution des phonèmes entendus. En effet, une distribution bimodale des traits phonétiques incite l'enfant à stabiliser deux catégories distinctes de phonèmes, tandis qu'une distribution unimodale le conduit à stabiliser une seule catégorie. Pour les voyelles, l'apprentissage se ferait avant six mois, voire même *in utero*. Cependant, une expérience récente montre que cet apprentissage n'est pas accéléré chez les enfants prématurés, bien qu'ils bénéficient de plusieurs mois supplémentaires d'exposition au langage. Des contraintes biologiques fortes encadrent donc la maturation des réseaux cérébraux du langage.

En ce qui concerne le lexique mental, les premières preuves d'une reconnaissance des mots parlés sont observées très tôt au cours de la première année de vie, bien avant la production des premiers mots. À cinq mois, les enfants préfèrent écouter leur propre prénom plutôt que celui d'un autre enfant, même si celui-ci partage le même profil d'accentuation tonique. Exposés, pendant 10 jours, à 30 minutes d'histoires pour enfants, les enfants préfèrent écouter cette liste de mots déjà entendus plutôt qu'une liste de mots nouveaux.

L'extraction de régularités statistiques joue un rôle important dans ces apprentissages. L'expérience fondatrice de Jenny Saffran et de ses collaborateurs montre qu'un enfant de huit mois extrait, d'une séquence de syllabes « bakimoufadolu... » les probabilités de transition entre les syllabes et les utilise pour en déduire la présence de « mots » récurrents de trois syllabes (Saffran, Aslin & Newport, 1996). Ces mots deviennent des candidats pour une association avec des stimuli visuels, préfigurant ainsi, dès la première année de vie, l'apprentissage du lexique mental de la langue maternelle.

Enfin, sur le plan grammatical, une sensibilité à l'ordre des mots est perceptible dès douze mois de vie. Chez l'adulte, l'IRM fonctionnelle montre que l'apprentissage de régularités auditives modifie l'activation des régions temporales et frontale inférieure gauche. On peut donc supposer que, chez l'enfant, le substrat cérébral de l'apprentissage fasse également appel au même réseau, même si aucune expérience d'imagerie de l'enfant ne vient encore étayer cette supposition.

### *Vers une théorie bayésienne du lexique mental*

Dans un article important, Fei Xu et Joshua Tenenbaum proposent une théorie bayésienne de l'acquisition du sens des mots (Xu & Tenenbaum, 2007). Leur modèle suppose que l'enfant dispose d'un vaste espace d'hypothèses sur les référents possibles des mots. Chaque hypothèse consiste en un sous-ensemble d'objets auquel un mot peut renvoyer (par exemple : « tous les êtres vivants », « tous les chiens », « tous les dalmatiens », etc.). Chaque fois que l'enfant entend un mot dans un contexte donné, il met à jour la probabilité que chaque hypothèse soit vraie, en suivant les règles bayésiennes. Enfin, une hypothèse cruciale du modèle est que la vraisemblance varie en fonction inverse de la taille de l'hypothèse considérée.

À partir de ces axiomes, les auteurs montrent qu'il est possible de rendre compte d'une série d'observations empiriques importantes. Le sens d'un mot peut être appris à partir d'un seul exemple, ou de quelques-uns. Les exemples positifs suffisent : l'enfant n'a pas besoin de contre-exemples. Il peut acquérir un ensemble de mots pour des concepts qui se recouvrent. Les inférences sur le sens d'un mot sont graduelles, avec des degrés de confiance variables. Enfin, ces inférences

peuvent être influencées par le contexte d'apprentissage, particulièrement l'attention, les connaissances et les intentions du locuteur.

En particulier, le modèle bayésien rend compte, sans hypothèse supplémentaire, d'un principe linguistique classique, le principe d'exclusivité : chaque entité ne possède qu'un seul nom. Dès seize mois, en effet, lorsqu'ils entendent un nom nouveau, les enfants postulent que celui-ci renvoie à un objet dont ils ne connaissent pas déjà le nom. Cette propriété découle simplement d'un modèle bayésien hiérarchique, si l'on suppose l'enfant capable d'inférences conversationnelles telles que « si mon interlocuteur avait voulu nommer l'objet X, il aurait utilisé le mot X ». Ainsi, l'acquisition du lexique ne nécessite peut-être rien d'autre qu'un algorithme générique d'inférence statistique qui pourrait être présent chez d'autres espèces animales. De fait, l'apprentissage de plusieurs centaines de mots, en respectant le principe d'exclusivité, a été documenté chez un chien domestique.

Le psycholinguiste Dennis Norris a également montré que les principes génériques de l'inférence bayésienne rendaient compte d'une grande variété de phénomènes liés à l'accès au lexique chez l'adulte. Lorsque nous entendons un mot, nos entrées sensorielles sont nécessairement ambiguës. Dans ces conditions, la solution optimale consiste à calculer, pour chaque mot, la probabilité que ce mot ait été à l'origine des entrées sensorielles observées (écrites ou orales). Il faut accumuler, au fil du temps, l'évidence issue d'échantillons sensoriels successifs. Ce modèle très simple parvient à expliquer, par exemple, pourquoi le temps de réponse à un mot varie en fonction du logarithme de sa fréquence dans la langue ; ou bien pourquoi il dépend également de la proximité et de la fréquence relative des mots voisins qui n'en diffèrent que d'une lettre.

En conclusion, au cours de la première année de vie, l'enfant compile probablement des statistiques lexicales qui lui fournissent une liste de candidats pour les mots de sa langue, ainsi qu'une hiérarchie d'hypothèses sur les objets et concepts auxquels ils renvoient. Au cours de la seconde année, pendant la phase d'explosion lexicale, ces deux ensembles vont être mis en relation. La théorie bayésienne semble susceptible d'expliquer la vitesse remarquable avec laquelle cet apprentissage se produit, et les contraintes que l'enfant applique pour en restreindre l'ambiguïté. Enfin, même à l'âge adulte, la théorie bayésienne continue de rendre compte de l'apprentissage de mots nouveaux et de la vitesse avec laquelle ces mots sont reconnus en temps réel.

### *L'apprentissage de règles linguistiques*

Comme l'a expliqué Noam Chomsky, la connaissance de la langue dépasse, de toute évidence, la simple évaluation des probabilités de transitions entre mots. Nous devons, au minimum, considérer des règles abstraites qui font intervenir des catégories de mots (noms, adjectifs, verbes). Même les probabilités de transition entre ces catégories ne suffisent pas : toutes les langues possèdent une structure arborescente en constituants enchâssés, elles sont régies par des règles syntaxiques récursives qui se traduisent par des dépendances à distance variable et arbitraire.

Selon Chomsky, la complexité de ces règles et la « pauvreté » des stimuli qu'entend l'enfant nécessitent de postuler l'existence d'une grammaire universelle, un ensemble de principes linguistiques préalables à tout apprentissage. Cependant, un article récent attaque ce point de vue en montrant que, sur la base de l'écoute de quelques dizaines de phrases, un algorithme d'apprentissage bayésien hiérarchique

parvient à sélectionner, parmi des millions de règles, celles de la grammaire universelle (Perfors, Tenenbaum & Regier, 2011). Il ne serait donc pas nécessaire de supposer que celles-ci soient innées.

La mise à l'épreuve empirique de cette idée reste presque entièrement explorée, car seules quelques recherches empiriques ont porté sur la capacité de très jeunes enfants à apprendre la grammaire. À dix-sept mois, les enfants repèrent l'alternance de mots de fonction (comme l'article « le ») et de noms communs (comme « chien »). À un an, ils extraient la structure grammaticale d'une séquence de syllabes, et la généralisent à des séquences nouvelles. Surtout, l'expérience princeps de Marcus et de ses collaborateurs montre que, dès sept mois, les bébés sont sensibles à des structures abstraites ou « algébriques » dans la répétition d'une série de syllabes (Marcus, Vijayan, Bandi Rao & Vishton, 1999). Prolongée chez l'adulte (Pena, Bonatti, Nespor & Mehler, 2002), cette recherche suggère que, dans le cerveau humain, deux mécanismes très différents seraient à l'œuvre au cours de l'apprentissage de séquences : (1) un mécanisme d'apprentissage statistique, sensible aux probabilités de transition, et (2) un mécanisme d'apprentissage de règles abstraites, tout-ou-rien, qui extrait des règles algébriques (ABB, AxC, etc.). Dès la seconde année de vie, le second mécanisme permettrait à l'enfant de découvrir des règles abstraites sur des items non-adjacents (Gomez & Maye, 2005).

Un article récent s'attaque au défi particulier que pose l'apprentissage des nombres et du comptage (Piantadosi, Tenenbaum & Goodman, 2012). En effet, si les très jeunes enfants apprennent facilement à réciter les nombres « un deux trois quatre... », ils n'en connaissent pas nécessairement le sens (Wynn, 1990). Pendant plusieurs mois, ils apprennent, un à un, le sens de chacun des nombres un, deux, trois, etc. Ce n'est que vers trois ans et demi, qu'ils font soudain un saut conceptuel remarquable : ils comprennent que chaque nom de nombre correspond à une quantité bien précise. Piantadosi et collaborateurs parviennent à rendre compte de ces observations par apprentissage bayésien au sein d'un espace de formules du lambda-calcul, un « langage de la pensée » doté de primitives et de lois de composition. L'algorithme proposé parvient à découvrir les règles récursives qui régissent un domaine comme celui des nombres.

### *Conclusion*

Même si l'hypothèse du bébé statisticien reste spéculative, elle s'avère extrêmement productive, sur le plan empirique autant que théorique. Une série d'expériences récentes bouleverse nos connaissances en démontrant l'étendue et la subtilité des inférences qu'un enfant de moins d'un an est capable de déployer. Il n'est pas exclu que le cerveau de l'enfant abrite, d'emblée, un mécanisme rapide d'inférence. Les modèles bayésiens de l'apprentissage ne sont qu'en partie innéistes. Certes, ils supposent l'existence d'une machinerie complexe d'inférence probabiliste et d'un très vaste espace d'hypothèses (qui, dans l'espèce humaine, à la différence peut-être de toutes les autres espèces animales, inclue les fonctions récursives). Mais ils reposent également sur une sélection hiérarchique qui élimine massivement les hypothèses inappropriées et converge rapidement vers des catégories ou des règles abstraites. En ce sens, ces modèles ne font que préciser la maxime proposée par Jean-Pierre Changeux dans *L'homme Neuronal* (1983) : « apprendre, c'est éliminer ».

### *Principaux travaux cités*

Berkes P., Orban G., Lengyel M. & Fiser J. (2011). Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science*, 331(6013), 83-87.

Bristow D., Dehaene-Lambertz G., Mattout J., Soares C., Gliga T., Baillet S. *et al.* (2009). Hearing faces : how the infant brain matches the face it sees with the speech it hears. *J Cogn Neurosci*, 21(5), 905-921.

Dehaene S., Bossini S. & Giraux P. (1993). The mental representation of parity and numerical magnitude. *Journal of Experimental Psychology : General*, 122, 371-396.

Gomez R. & Maye J. (2005). The Developmental Trajectory of Nonadjacent Dependency Learning. *Infancy*, 7(2), 183-206.

Goodman N.D., Ullman T.D. & Tenenbaum, J. B. (2011). Learning a theory of causality. *Psychol Rev*, 118(1), 110-119.

Gopnik A., & Schulz L. (2004). Mechanisms of theory formation in young children. *Trends Cogn Sci*, 8(8), 371-377.

Gweon H. & Schulz L. (2011). 16-month-olds rationally infer causes of failed actions. *Science*, 332(6037), 1524.

Gweon H., Tenenbaum J. B. & Schulz L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proc Natl Acad Sci USA*, 107(20), 9066-9071.

Marcus G. F., Vijayan S., Bandi Rao S. & Vishton P.M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398), 77-80.

Pena M., Bonatti L.L., Nespor M. & Mehler J. (2002). Signal-driven computations in speech processing. *Science*, 298(5593), 604-607.

Perfors A., Tenenbaum J. B. & Regier T. (2011). The learnability of abstract syntactic principles. *Cognition*.

Piantadosi S.T., Tenenbaum J. B. & Goodman N. D. (2012). Bootstrapping in a language of thought : a formal model of numerical concept learning. *Cognition*, 123(2), 199-217.

Saffran J. R., Aslin R. N. & Newport E.L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928.

Teglas E., Vul E., Girotto V., Gonzalez M., Tenenbaum J.B. & Bonatti L.L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science*, 332(6033), 1054-1059.

Ullman S., Harari D. & Dorfman N. (2012). From simple innate biases to complex visual concepts. *Proc Natl Acad Sci USA*, 109(44), 18215-18220.

Ullman T. D., Goodman N. D. & Tenenbaum A. (2012). Theory learning as stochastic search in the language of thought. *Cognitive Development*, 27, 455-480.

Xu F. & Tenenbaum J.B. (2007). Word learning as Bayesian inference. *Psychol Rev*, 114(2), 245-272.

### **Séminaire : Interactions entre espace, temps et nombre : 20 ans de recherches**

Le séminaire s'est tenu sous forme de colloque le mardi 26 février 2013 dans l'amphithéâtre Marguerite de Navarre. Il a réuni quelques centaines de personnes issues de disciplines très diverses (sciences cognitives, psychologie développementale,



neuropsychologie, neurosciences, philosophie, anthropologie), dont une bonne moitié d'étudiants<sup>b</sup>.

L'objectif était de faire le point sur les interactions entre le nombre et l'espace, exactement vingt ans après la publication, par Stanislas Dehaene, Serge Bossini et Pascal Giraux (1993), d'un article mettant en évidence une association systématique et automatique : dès qu'une personne de culture occidentale réfléchit aux nombres, elle ne peut pas s'empêcher d'associer les petits nombres avec la partie gauche de l'espace, et les grands nombres avec la droite. Ce phénomène a été appelé l'*effet SNARC* (*spatial-numerical association of response codes*).

Pourquoi les concepts de nombre, d'espace, mais aussi de temps, sont-ils aussi étroitement liés ? Pas moins d'un millier de publications ont exploré les origines de cet effet. Même les bébés, et d'autres espèces animales, considèrent les dimensions de nombre, de taille, d'intensité et de durée comme des grandeurs avec une polarité commune, l'axe petit-grand. Une association systématique entre les petits nombres et la gauche est même retrouvée chez le poussin ! Il semble donc que les mécanismes de l'effet SNARC soient très anciens dans l'évolution. Chez l'homme, l'imagerie cérébrale suggère que le cortex pariétal, qui abrite de multiples codes neuronaux pour les dimensions numériques et spatiales, pourrait être le siège de cette interaction. Mais comment ces représentations sont-elles modifiées par la culture et l'éducation, particulièrement le sens de l'écriture, dont on sait à présent qu'il peut modifier et même renverser l'effet SNARC ? Vingt ans après, la « chasse au SNARC » continue.

Liste des intervenants :

- Stanislas Dehaene : The discovery of the SNARC effect and its brain correlates ;
- Martin Fischer and Samuel Shaki : Cultural and conceptual aspects of SNARC ;
- Wim Fias : Mechanisms underlying the SNARC effect ;
- Marco Zorzi : Neuropsychological impairments of number, space and time ;
- David Burr : A visual sense of number ;
- Giorgio Vallortigara : Space and number in animals ;
- Andreas Nieder : The neural coding of number and other dimensions ;
- Lola de Hevia : What babies know about space, time and number ;
- Manuela Piazza : How do young children expand their number sense ? ;
- Brian Butterworth : Space, time and number in dyscalculia.

Le colloque a été organisé avec les professeurs Martin Fischer (Potsdam) et Marco Zorzi (Padoue) et le Dr. Samuel Shaki (Ariel, Israël), avec le soutien de l'*European Society for Cognitive Psychology* (ESCOP), de l'*European Psychological Society* (EPS), de l'*École des neurosciences de Paris* (ENP), et de la Fondation Hugot du Collège de France. Nous remercions vivement tous ces organismes, ainsi que Giovanna Santoro (INSERM) et Sophie Benitta (Collège de France) pour leur aide efficace. Ces financements ont permis d'attribuer des bourses à cinq étudiants et un prix récompensant les deux meilleurs parmi les 38 posters présentés.

---

b. Les enregistrements vidéo de ce séminaire sont disponibles sur le site Internet du Collège de France : <http://www.college-de-france.fr/site/stanislas-dehaene/seminar-2012-2013.htm> [Ndlr].

## Autres enseignements

Dans le cadre des accords internationaux du Collège de France, quatre cours et séminaires ont été donnés :

À l'Institut d'études avancées de Pavie (IUSS) :

- Reading : How literacy changes our brain ;
- Arithmetic : The brain mechanisms of numeracy ;
- Language, music and mathematics : in search of brain mechanisms for syntax.

Au Center for Mind/Brain Sciences de l'université de Trente :

- Recent advances in understanding the mechanisms of consciousness.

## PUBLICATIONS 2012-2013

### Articles originaux

Faugeras F., Rohaut B., Weiss N., Bekinschtein T.A., Galanaud D., Puybasset L., Bolgert F., Sergent C., Cohen L., Dehaene S., Naccache L. Event related potentials elicited by violations of auditory regularities in patients with impaired consciousness. *Neuropsychologia*, February 2012, 50(3), 403-18 [doi: 10.1016/j.neuropsychologia.2011.12.015].

Pinel P., Fauchereau F., Moreno A., Barbot A., Lathrop M., Zelenika D., Le Bihan D., Poline J.P., Bourgeron T., Dehaene S. Genetic Variants of FOXP2 and KIAA0319/TTRAP/THEM2 Locus Are Associated with Altered Brain Activation in Distinct Language-Related Regions. *J Neurosci*, January 2012, 32(3), 817-25 [doi: 10.1523/JNEUROSCI.5996-10.2012].

Wyart V., Dehaene S., Tallon-Baudry C. Early Dissociation between Neural Signatures of Endogenous Spatial Attention and Perceptual Awareness during Visual Masking. *Frontiers in Human Neuroscience*, 2012, 6 [doi: 10.3389/fnhum.2012.00016].

Wacongne C., Changeux J.P., Dehaene S. A neuronal model of predictive coding accounting for the mismatch negativity. *J Neurosci*, 2012, 32(11), 3665-3678 [doi: 10.1523/JNEUROSCI.5003-11.2012].

Monzalvo K., Fluss J., Billard C., Dehaene S., Dehaene-Lambertz G. Cortical networks for vision and language in dyslexic and normal children of variable socio-economic status. *Neuroimage*, 2012, 61(1), 258-274 [doi : 10.1016/j.neuroimage.2012.02.035].

Szwed M., Ventura, P., Querido L., Cohen L., Dehaene S. Reading acquisition enhances an early visual process of contour integration. *Developmental Science*, 2012, 15(1), 139-49 [doi: 10.1111/j.1467-7687.2011.01102.x].

Hesselmann G., Naccache L., Cohen L., Dehaene S. Splitting of the P3 component during dual-task processing in a patient with posterior callosal section. *Cortex*, March 2012 [doi: 10.1016/j.cortex.2012.03.014].

Maruyama M., Pallier C., Jobert A., Sigman M., Dehaene S. The cortical representation of simple mathematical expressions. *Neuroimage*, July 2012, 61(4),1444-60 [doi:10.1016/j.neuroimage.2012.04.020].

Dehaene S., Sigman M. From a single decision to a multi-step algorithm. *Curr Opin Neurobiol.*, 2012, 22, 1-9 [doi: 10.1016/j.conb.2012.05.006].

Vagharchakian L., Dehaene-Lambert Z.G., Pallier C., Dehaene S. A Temporal Bottleneck in the Language Comprehension Network. *The Journal of Neurosciences*, 2012, 32(26), 9089-9102 [doi: 10.1523/JNEUROSCI.5685-11.2012].

Schurger A., Sitt J., Dehaene S. An accumulator model for spontaneous neural activity prior to self-initiated movement. *PNAS*, August 6, 2012 [doi: 10.1073/pnas.1210467109].

Le Floch E., Guillemot V., Frouin V., Pinel P., Lalanne C., Trinchera L., Tenenhaus A., Moreno A., Zilbovicious M., Bourgeron T., Dehaene S., Thirion B., Poline J.B., Duchesnay E. Significant correlation between a set of genetic polymorphisms and a functional brain network revealed by feature selection and sparse Partial Least Squares. *Neuroimage*, 2012, 63, 11-24 [doi: 10.1016/j.neuroimage.2012.06.061].

Szwed M., Vinckier F., Cohen L., Dehaene S. Towards a universal neurobiological architecture for learning to read. *Behav Brain Sci*, 2012, 35(5), 308-309 [doi: 10.1017/S0140525X12000283].

Schneider E., Maruyama M., Dehaene S., Sigman M. Eye gaze reveals a fast, parallel extraction of the syntax of arithmetic formulas. *Cognition*, 2012, 125(3), 475-490 [doi: 10.1016/j.cognition.2012.06.015].

Nakamura K., Kuo W.J., Pegado F., Cohen L., Tzeng O.J.L., Dehaene S. Universal brain systems for recognizing word shapes and handwriting gestures during reading. *PNAS*, 2012 109(50), 20762-20767 [doi:10.1073/pnas.1217749109].

Thiebaut de Schotten M., Cohen L., Amemiya E., Braga L.W., Dehaene S. Learning to Read Improves the Structure of the Arcuate Fasciculus. *Cerebral Cortex*, Dec 2012 [doi:10.1093/cercor/bhs383].

McCrink K., Spelke E.S., Dehaene S., Pica P. Non-symbolic halving in an Amazonian indigene group. *Developmental Science*, 2012, 16(3) 451-462 [doi: 10.1111/desc.12037].

Kouider S, Stahlhut C., Gelskov S.V., Barbosa L.S., Dutat M., de Gardelle V., Christophe A., Dehaene S., Dehaene-Lambertz G. A Neural Marker of Perceptual Consciousness in Infants. *Science*, 2013, 340, 376-380 [doi: 10.1126/science.1232509].

Charles L., Van Opstal F., Marti S., Dehaene S. Distinct brain mechanisms for conscious versus subliminal error detection. *Neuroimage*, 2013, 73, 80-94 [doi: 10.1016/j.neuroimage.2013.01.054].

Weibel S., Giersch A., Dehaene S., Huron C. Unconscious task set priming with phonological and semantic tasks. *Conscious Cogn.* 2013, 22(2), 517-527 [doi: 10.1016/j.concog.2013.02.010].

Pegado F., Nakamur A.K., Braga L.W., Ventura P., Nunes Filho G., Pallier C., Jobert A., Morais J., Cohen L., Kolinsky R., Dehaene S. Literacy breaks mirror invariance for visual stimuli : a behavioral study with adult illiterates. *JEP :General*, 2013, sous presse.

## Chapitres de livres

Dehaene S. Cuando el reciclaje neuronal prolonga la hominización. In *La pizarra de Babel*, Sebastian Lapina, Mariano Sigman (Eds.), Libros del Zorzal, 2012, chap. 4, 91-105.

## Reuves, commentaires

King J.R., Bekinschtein T., Dehaene S. Comment on “Preserved feedforward but impaired top-down processes in the vegetative state”. *Science*, December 2011, 334(6060), 1203 [doi: 10.1126/science.1210012].

Dehaene S., Pegado F. O impacto da aprendizagem da leitura sobre o cérebro. *Grupo A, Revista Pátio*, Feb-Apr 2012.

Dehaene S. The Eternal Silence of Neuronal Spaces. *Science* 2012, 336, 1507-8 [doi: 10.1126/science.1222480].

## Film documentaire

Gibrat J.P. & Dehaene S. *Les chemins de la lecture*. 52 minutes, coproduit par ARTE, diffusé sur ARTE le 21 mars 2013.

## AUTRES ACTIVITÉS

### Principales conférences invitées

- « How the literate brain differs from the illiterate brain ». International meeting on brain, development and education, Neurospin, Saclay, 20 janvier 2012.
- « Reading as neuronal recycling ». 2<sup>nd</sup> Latin-American School for Education, Cognitive and Neural Sciences. El Calafate, Argentine, 7 mars 2012.
- « Signatures of conscious processing in the human brain ». 1<sup>st</sup> Seattle Brain Salon "Molecules and Mind », on the occasion of the inauguration of the Allen Institute for Brain Science, Seattle, 19 avril 2012.
- « Education matters : literacy, numeracy and the developing brain ». Roberts Research Institute, London, Ontario, 7 mai 2012.
- « The brain circuits for reading : How literacy alters our cortical representations and their connections ». Annual Lecture in honor of Norman Geschwind, Harvard Institutes of Medicine, Boston, 9 mai 2012.
- « How learning to read changes the brain ». Stanford University, Palo Alto, Californie, 11 mai 2012.
- « In search of the neural encoding of constituent structure ». Lorentz center, Leiden, 29 mai 2012.
- « Brain mechanisms of constituent structure in language and mathematics ». Conference on Neuroscience of Language, Jülich, Allemagne, 29 juin 2012.
- Trois conférences invitées sur « Cognition, Cerveau et Education », Rio, Florianopolis et Brasilia, Brésil, 7-17 juillet 2012.
- « How learning to read changes the brain ». Kavli Prize week, Trondheim, Norvège, 6 septembre 2012.
- « Electrophysiological signatures of conscious processing ». *Invited address*, International Society for Psychophysiological Research. New Orleans, 21 septembre 2012.
- « Psychological and physiological signatures of consciousness ». Keynote lecture, Belgian Brain Congress, Liège, Belgique, 27 octobre 2012.
- « Searching for Brain Mechanisms of Conscious Access and Introspection ». Colloque « Neurosciences and the human person », Académie pontificale des sciences, Vatican, 8-11 novembre 2012.
- « How the brain learns to read, and how we can improve reading education ». WISE Summit (World Innovation Summit for Education), Doha, Qatar, 14 novembre 2012.
- « Signatures of consciousness ». Italian Society for Psychophysiology. Venise, 24 novembre 2012.
- « Signatures of Conscious Processing in the Human Brain ». *Keynote lecture*, NIPS (Neural Information Processing), Lake Tahoe, États-Unis, 5 décembre 2012.
- « Advances in understanding reading acquisition ». *Keynote lecture*, 2012 Budapest CEU Conference on Cognitive Development. Budapest, 12 janvier 2013.
- « Signatures of consciousness in the human brain ». *Keynote lecture*, Annual Meeting of the Swiss Society for Neuroscience (SSN), Genève, 2 février 2013.
- « Advances in understanding reading acquisition ». 3<sup>rd</sup> Latin-American School for Education, Cognitive and Neural Sciences. Ilha de Comandatuba, Brésil, 8 mars 2013.
- « How the ventral pathway is changed by literacy ». Oxford Dyslexia Meeting, University of Oxford, 12 avril 2013.

- Trois cours sur « Cognitive neuroscience » à l'Institut d'études avancées de l'université de Pavie, Italie, 19-25 mai 2013.
- « Using MEG to track conscious access and its non-conscious consequences ». Association for the Scientific Study of Consciousness, San Diego, 15 juillet 2013.
- « From illiterate to literate brain ». East China Normal University, Shanghai, 26 août 2013.

### **Direction de thèses**

Lucie Charles (19/09/2013 ; directeur Stanislas Dehaene) : « Mécanismes conscients et non-conscients de la décision et de la méta-décision ».

### **Participation aux programmes nationaux et internationaux**

Responsabilité du pilier « Architectures cognitives » du *Human Brain Project*, projet « Flagship » de la Communauté européenne, financé pour dix ans, et visant à développer des simulations du cerveau humain.

Participation au comité de sélection du Prix international Kavli en neurosciences.

Participation au conseil scientifique de la Direction générale de l'enseignement scolaire (DGESCO) dirigée par Jean-Paul Delahaye.

## ACTIVITÉS DE RECHERCHE DU LABORATOIRE

### **Mécanismes cérébraux de la détection d'erreurs**

Dans le cadre de la thèse de Lucie Charles, un travail empirique et théorique a porté sur les mécanismes cérébraux de la détection d'erreurs, et a permis de préciser les limites du traitement non-conscient (Charles et Dehaene, *NeuroImage* 2013). La capacité d'évaluer notre propre comportement et de juger si nous avons répondu de façon correcte ou incorrecte fait partie des opérations dites « métacognitives ». Jusqu'à présent, beaucoup d'auteurs supposaient que de telles opérations renvoyaient nécessairement à des processus conscients. En effet, lorsque nous nous penchons en nous-mêmes et évaluons nos propres processus cognitifs, il semble très intuitif que nous en soyons conscients (pratiquement par définition). Cependant, il est également possible que certaines opérations cognitives automatiques et inconscientes soient spécialisées dans l'évaluation de la performance des autres systèmes cérébraux. Les paradigmes de présentation subliminale nous ont permis de mettre cette alternative à l'épreuve des faits.

Nous avons présenté des chiffres masqués à des degrés variables, de sorte qu'ils étaient parfois visibles et parfois invisibles. Tout en enregistrant l'activité de leur cerveau grâce à la magnéto- et à l'électro-encéphalographie, nous demandions aux sujets, à chaque essai, de (1) juger si les chiffres étaient plus grands ou plus petits que 5 ; (2) dire s'ils avaient vu le chiffre ou pas, ce qui permettait de trier les essais selon que les chiffres étaient conscients ou non-conscients ; (3) dire s'ils pensaient avoir répondu correctement à la tâche de comparaison.

Les résultats ont démontré l'existence de deux processus métacognitifs de détection d'erreurs, l'un conscient, l'autre non. D'une part, la réponse cérébrale appelée « négativité associée à l'erreur » (*error-related negativity*) survenait de

façon tout-ou-rien, uniquement lors des essais conscients. D'autre part, les réponses comportementales montraient que nos sujets parvenaient à juger si leurs réponses étaient correctes ou incorrectes, bien mieux que ne le prédisait le niveau du hasard, même lors des essais non-conscients.

Ces résultats suggèrent qu'une partie seulement des processus métacognitifs peut se dérouler sans conscience, et conduisent à distinguer deux types de processus de détection d'erreurs. Toute décision à choix forcé, même non-consciente, résulte d'une accumulation d'évidence probabiliste et, à ce titre, s'accompagne d'une information statistique approximative sur la précision de la réponse. Cependant, la capacité de détecter nos erreurs fait appel à un second système, vraisemblablement localisé dans le cortex cingulaire, et qui compare l'intention avec l'action en cours afin de détecter tout décalage entre nos plans et leur réalisation. Ce second niveau nécessite une représentation consciente de l'intention motrice et, à ce titre, dépend de la prise de conscience du stimulus présenté.

En conclusion, l'importance de ce travail réside dans le fait qu'il étend la limite des opérations non-conscientes en montrant que certains processus métacognitifs peuvent se dérouler non-consciemment – tout en montrant qu'il existe des limites très nettes à la profondeur de ce traitement non-conscient.

### **Accès à la conscience chez le très jeune enfant**

Un second travail notable, en collaboration avec Sid Kouider (École normale supérieure) et Ghislaine Dehaene-Lambertz, et publié dans *Science* en 2013, a permis de poser, pour la première fois, la question de la détection d'un traitement conscient chez le très jeune enfant.

La plupart des parents n'ont aucun doute sur le fait que leur bébé ait conscience du monde qui l'entoure. Cependant, cette assertion n'est pas facile à tester empiriquement dans la mesure où le bébé ne dispose pas d'un langage qui lui permette de rapporter ses états mentaux. Bon nombre de ses comportements pourraient s'expliquer par des processus automatiques et non conscients, analogues à la « vision aveugle » présente chez certains patients atteints de lésions cérébrales.

Pour pallier ces difficultés, nous avons proposé une idée simple. Puisque nous disposons, aujourd'hui, de solides signatures cérébrales du traitement conscient chez l'adulte, pourrions-nous les détecter également chez le bébé de quelques mois ? Nous avons donc adapté à l'enfant le test de masquage utilisé dans nos recherches antérieures chez l'adulte. À chaque essai, on présente un visage pendant un temps très bref mais variable, précédé et suivi d'une image qui sert de masque. Chez l'adulte, il existe un seuil de perception en deçà duquel le visage reste invisible. Au-delà de ce seuil, le stimulus, lorsqu'il devient visible, déclenche une réponse cérébrale tardive et non linéaire caractéristique de ce que nous avons appelé l'« embrasement conscient » (*conscious ignition*). Or, chez le bébé, le comportement et les réponses cérébrales présentent une organisation similaire. Sur le plan comportemental, l'enfant ne s'oriente que si la durée de présentation du visage excède un certain seuil. Sur le plan cérébral, les réponses précoces aux visages varient de façon strictement linéaire avec la durée de présentation, mais seule une étape tardive suit un profil non linéaire, avec le même seuil de déclenchement que le comportement d'orientation vers le visage. L'analogie de ces réponses non linéaires tardives chez l'enfant et l'adulte suggère que l'ensemble de l'architecture

perceptive, qui permet l'accumulation non-consciente d'évidence et le franchissement d'un seuil d'accès à la conscience, est présente dès cinq mois de vie.

La mesure de l'intensité et de la latence des réponses cérébrales vient cependant nuancer ce propos. Chez le bébé, l'ignition consciente est trois à quatre fois plus lente que chez l'adulte, et le seuil de conscience est nettement plus élevé. Il est probable que l'architecture fonctionnelle soit présente, mais repose sur des faisceaux de connexions corticales à longue distance immatures, ralentis par l'absence de myélinisation, et que le flux de conscience soit donc nettement ralenti chez le très jeune enfant.

### **Apprentissage de la lecture**

Le laboratoire a poursuivi l'analyse des mécanismes cérébraux qui sous-tendent l'apprentissage de la lecture. D'une part, une analyse des données d'IRM de diffusion, acquises chez la population d'illettrés, ex-illettrés et lettrés brésiliens dont les résultats d'IRM fonctionnelle avaient été publiés en 2010, a permis de montrer que la partie postérieure du faisceau arqué, dans l'hémisphère gauche, devient significativement plus anisotrope avec l'apprentissage de la lecture (Thiebaut de Schotten et coll., *Cerebral Cortex*, 2012). Ainsi, ce faisceau serait modifié, dans sa structure anatomique même, par l'apprentissage. Il jouerait un rôle essentiel dans l'échange de signaux entre les aires de la voie visuelle ventrale codant pour les chaînes de lettres, et les régions temporales supérieures et pariétales inférieures codant pour les phonèmes.

D'autre part, un travail en collaboration avec Amir Amedi, à l'université hébraïque de Jérusalem, a montré que l'aire de la forme visuelle des mots, la région occipito-temporale ventrale spécialisée dans la lecture, est présente même chez l'aveugle de naissance, qu'il lise en braille ou même par substitution sensorielle auditive (Striem-Amit et coll., *Neuron*, 2013). Ce résultat étonnant montre que cette région n'est pas uniquement visuelle, mais « méta-modale » et capable de reconnaître la forme des lettres quelle que soit la modalité de présentation.

### **Sens du nombre et de l'espace**

En ce qui concerne la compréhension des nombres, enfin, un programme de recherche a été initié afin de préciser le décours temporel des opérations qui permettent au cerveau de convertir en une quantité un nombre écrit en chiffres arabes. Pour ce faire, dans le cadre de sa thèse, Dror Dotan a développé un logiciel innovant sur tablette tactile qui permet de digitaliser toute la trajectoire du doigt lorsqu'un participant pointe vers la position d'un nombre sur une ligne numérique orientée de gauche à droite (Dotan et Dehaene, *Cognition*, 2013). Dans un premier travail, nous avons étudié dans quel ordre, et avec quelle latence, se fait le calcul de la position spatiale qui correspond à un nombre sur une échelle. Les résultats ont montré que le cerveau adulte traite presque immédiatement les deux chiffres en parallèle, avec des pondérations linéaires appropriées pour les dizaines et les unités. Cependant, pendant une étape intermédiaire, c'est le logarithme du nombre qui détermine la direction du doigt. Or, divers travaux antérieurs ont montré que chez l'animal, l'enfant préscolaire et l'adulte non-éduqué, le sens du nombre obéit à une organisation logarithmique. Ainsi, il semble que même chez l'adulte éduqué et entraîné, cette représentation primitive et approximative continue d'être présente et d'influencer, de façon non-consciente, nos jugements numériques.