

Psychologie cognitive expérimentale

M. Stanislas DEHAENE, professeur

Enseignement

Leçon inaugurale

La leçon inaugurale, prononcée le 27 avril 2006, sera publiée aux éditions Fayard à l'automne 2006. Elle a porté sur les lois de la psychologie et les stratégies de recherche qui pourraient permettre de les établir.

« La psychologie est la science de la vie mentale. » Ainsi William James cernait-il, dès 1890, le domaine de ce qui allait devenir la psychologie cognitive. Celle-ci s'affirme comme une part intégrante des sciences de la vie, qui exploite toutes les méthodes de la biologie, depuis la génétique jusqu'à l'imagerie cérébrale ; mais une science de la vie *mentale*, qui tente d'énoncer des lois générales de la pensée, un domaine intime et subjectif que l'on aurait pu penser inaccessible à la méthode scientifique.

La diversité des cultures, des personnalités, et des compétences humaines semble rendre hasardeux le projet d'une science psychologique unifiée, capable d'énoncer des lois d'une portée générale. De fait, depuis les vingt dernières années, les laboratoires de psychologie expérimentale se sont spécialisés, chacun s'attachant à comprendre un aspect restreint de la cognition.

Par-delà les hasards de l'histoire évolutive et culturelle de l'espèce humaine, se pourrait-il pourtant que notre vie mentale soit régie par quelques principes généraux d'architecture cérébrale ? Renouant avec l'esprit du programme psychophysique de Fechner, Wundt, Ribot ou Piéron, la psychologie doit se donner un objectif ambitieux : pousser l'analyse des fonctions cognitives supérieures jusqu'à un niveau de formalisation comparable à celui de la physique, par la formulation de théories mathématiques et de modèles neuro-informatiques. Les lois psychologiques, même si elles peuvent être exprimées sous forme d'algorithmes formels, ne seront comprises en profondeur que lorsqu'elles auront été mises en relation avec les différents niveaux d'organisation du système nerveux. L'imagerie cérébrale

présente ainsi une opportunité exceptionnelle d'approfondissement du champ de la psychologie. Loin de ne constituer qu'une « néo-phrénologie », elle donne accès à l'architecture fonctionnelle et aux mécanismes des fonctions cognitives, plus directement que la traditionnelle étude du comportement.

Pour illustrer ces propos, la leçon inaugurale a présenté la dissection d'une fonction cognitive, l'arithmétique mentale, qui offre un prétexte à passer en revue quelques questions fondamentales d'architecture de la cognition animale et humaine. On pourrait penser que l'arithmétique n'est qu'une invention culturelle récente de l'humanité. Pourtant, un sens du nombre est présent chez le nourrisson et de nombreuses espèces animales. Il suit des lois psychophysiques simples qui régissent également la plupart des dimensions perceptives. Plus d'un siècle après leur formulation mathématique, ces lois reçoivent une validation directe avec la découverte, chez l'animal, de neurones codant pour les nombres, dans des régions homologues de celles observées dans le cerveau humain.

Les propriétés de ces populations de neurones rendent compte des variations de nos temps de réponse dans des opérations arithmétiques simples. Les règles de la chronométrie mentale peuvent se déduire de la physique statistique de réseaux neuronaux qui implémentent, en première approximation, un algorithme de prise de décision optimale.

Si l'on commence ainsi à comprendre les opérations mentales les plus automatiques, l'un des défis de la psychologie reste d'élucider les mécanismes de leur contrôle conscient. Loin de constituer le dernier refuge de la subjectivité, la cognition consciente suit également des lois psychologiques universelles. La compétition entre deux objets de perception, la collision de deux opérations mentales ou le détournement de l'attention créent des conditions reproductibles où l'on peut, à volonté, faire apparaître ou disparaître la perception consciente. Les recherches actuelles délimitent le contour des opérations subliminales et les contrastent aux états corticaux globaux et synchronisés observés lors de la prise de conscience. Elles laissent ainsi entrevoir une modélisation neuronale objective des contenus subjectifs de la conscience.

Cours 2005-2006

Le cours a porté sur le rôle de l'imagerie cérébrale en psychologie cognitive.

Historiquement, la psychologie cognitive s'est principalement appuyée sur les sciences du comportement. Par l'analyse des réponses des participants, de leurs erreurs et de leurs temps de réaction (chronométrie mentale), dans des tâches conçues pour séparer différents modèles de l'architecture d'une tâche cognitive, le psychologue cognitif tente d'inférer l'organisation des représentations mentales et des algorithmes de traitement de l'information chez l'adulte, l'enfant, et l'animal.

Un malentendu entoure fréquemment l'usage de l'imagerie cérébrale humaine. Certains prétendent qu'elle ne vise qu'à localiser, dans le cerveau, les bases

cérébrales de fonctions cognitives dont l'existence et l'organisation fonctionnelle auraient déjà été établies par des méthodes comportementales. Ainsi, la neuro-imagerie ne serait qu'une forme élaborée de « néo-phrénologie ».

À l'encontre de cette vision réductrice, le cours se proposait de montrer, à travers de nombreux exemples, comment les diverses méthodologies qui composent la neuro-imagerie fonctionnelle permettent de poser des questions d'organisation fonctionnelle qui vont bien au-delà de la simple localisation, et relèvent de plein droit de la psychologie. La neuro-imagerie représente une opportunité pour le psychologue dans la mesure où elle donne accès aux mécanismes cognitifs d'une façon plus directe. Les études comportementales, telles celles fondées sur le temps de réaction résument en une seule mesure une série d'étapes de traitement. L'examen de l'état d'activité d'une structure cérébrale fournit un ensemble plus différencié de mesures, chacune pouvant constituer un index relativement pur d'une étape de traitement. Ainsi, la neuro-imagerie apporte-t-elle des réponses nouvelles à certains des objectifs majeurs de la psychologie cognitive, notamment :

- la décomposition des tâches cognitives en sous-systèmes ou « modules » ;
- la mesure du décours temporel des opérations cognitives ;
- l'analyse de l'organisation interne des représentations mentales.

Principes et limites de la neuro-imagerie cognitive

Le psychologue ne peut se dispenser de comprendre les bases fondamentales des méthodes d'imagerie qu'il envisage d'utiliser. Après un bref rappel historique, soulignant les contributions conjointes de la physique fondamentale et des études du métabolisme cérébral, le cours s'est donc attaché à décrire les principes théoriques et pratiques de la méthode dominante actuellement, l'imagerie fonctionnelle par résonance magnétique (IRMf).

Sur le plan physique, la résonance magnétique nucléaire consiste en l'absorption par un noyau atomique pourvu d'un spin non-nul (fréquemment le proton), d'un rayonnement électromagnétique d'une fréquence particulière, en présence d'un champ magnétique. Le retour à l'équilibre des spins entraîne l'émission d'une onde électromagnétique mesurable à distance, dont les constantes de relaxation (T1, T2) caractérisent différents tissus. L'utilisation de gradients de champ magnétique permet de sélectionner une coupe, et de coder en fréquence et en phase l'origine spatiale des signaux à l'intérieur de cette coupe.

Ces principes sont à l'origine de l'IRM anatomique, exploitée depuis plusieurs décennies dans de nombreux hôpitaux. Mais comment rendre le signal de résonance magnétique sensible à l'activité cérébrale ? Ogawa et ses collaborateurs ont exploité le fait que la déoxy-hémoglobine est paramagnétique et perturbe le signal RMN (effet sur le T2 apparent ou T2*). Les travaux des groupes de Sherrington, Sokoloff, et Raichle avaient, depuis fort longtemps, montré que des modifications des paramètres hémodynamiques accompagnent l'entrée en activité

du cortex. L'activité cérébrale peut donc être mesurée indirectement par ce que l'on appelle l'« effet BOLD » (Ogawa, Lee, Kay, & Tank, 1990), une chaîne causale complexe dont les mécanismes biophysiques et moléculaires restent imparfaitement compris, dans laquelle l'activité cérébrale entraîne :

- une surconsommation d'oxygène, mais également une augmentation du débit sanguin ;
- une augmentation du ratio oxy-/déoxy-hémoglobine ;
- une diminution de la susceptibilité magnétique ;
- une augmentation du paramètre $T2^*$ et donc du signal IRM.

Bien qu'aucune injection de produit de contraste ne soit nécessaire, les mesures d'IRMf sont effectuées dans des conditions contraintes. Le participant volontaire est immobilisé en position allongée à l'intérieur de l'aimant dont le diamètre est réduit. L'environnement de l'IRM est bruyant (> 90 dB), et nécessite une protection auditive. Le champ magnétique lui-même ne présente aucun danger connu, à condition qu'aucun objet métallique ne soit introduit dans le champ. Sa présence limite l'utilisation d'appareils électroniques utiles en psychologie (caméra, souris d'ordinateur, etc.). Toutefois, par le biais de dispositifs de stimulation sensorielle et de réponse adaptés (miroirs, écouteurs piézo-électriques, etc.), il est possible de réaliser la plupart des expériences classiques de psychologie dans l'aimant d'une machine d'IRM.

La création des images nécessite une bonne connaissance de la réponse hémodynamique mesurée par IRMf en réponse à un stimulus ponctuel. Cette réponse, variable selon les personnes et les régions cérébrales, présente certaines caractéristiques stables. Après une brève période de descente du signal (« early dip »), la réponse monte fortement pour atteindre un pic 4 à 6 secondes après la stimulation, puis elle redescend souvent en dessous de la ligne de base avant de retrouver son niveau initial en une vingtaine de secondes. Ainsi existe-t-il un délai considérable entre l'activité neuronale et la mesure IRM, qui impose une limite à la résolution temporelle de la méthode. Toutefois, un point essentiel, validé empiriquement dans le système visuel, est l'*additivité* des réponses BOLD (Boynton, Engel, Glover, & Heeger, 1996) : lorsque la stimulation est répétée, à des intervalles de temps dépassant quelques centaines de millisecondes, l'activation évoquée est la somme des contributions des activations associées à chaque stimulus. Il en va de même lorsque la durée du stimulus est prolongée. La réponse globale peut donc être prédite par une opération linéaire, la convolution des périodes de stimulation avec une réponse hémodynamique canonique.

Cette propriété d'additivité est à la base de la méthode de régression multiple qui est employée pour analyser les données d'IRMf. Chaque élément de volume du cerveau, ou *voxel*, est soumis à une vaste régression linéaire à partir de prédicteurs créés par convolution de l'activité neuronale attendue et de la fonction hémodynamique canonique. Les contributions indépendantes de différents profils de réponse peuvent donc être testées statistiquement, et conduire à l'identification de régions distinctes du cerveau. Pour visualiser le résultat, on créera des « cartes

de paramètres statistiques » (*statistical parameter maps* ou SPM), représentant en fausses couleurs le niveau d'activité ou la valeur des paramètres statistiques de la régression en chaque point du cerveau où la statistique est significative au seuil choisi.

De nombreux travaux se sont attachés à examiner quels paramètres de l'activité cérébrale sont manifestés dans la réponse BOLD. Nikos Logothetis a réalisé un travail de pionnier dans ce domaine, en réalisant l'exploit technologique de mesurer simultanément l'activité neuronale locale (décharges neuronales et potentiels de champ local [LFP pour *local field potential*]) et le signal BOLD en IRM fonctionnelle (Logothetis, 2002 ; Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). En réponse à une stimulation visuelle de durée variable, il montre que les décharges neuronales sont brèves tandis que le potentiel de champ local montre un maintien durable. Le signal BOLD suit étroitement ce dernier décours temporel. Ainsi, le signal BOLD serait le reflet de l'intégration post-synaptique locale, au coût métabolique élevé, plutôt que des décharges neuronales elles-mêmes. Des exemples montrent que, dans des situations de forte inhibition, il peut même y avoir élévation du signal BOLD avec diminution des décharges neuronales. Ainsi, l'IRMf signalerait les entrées et les calculs locaux d'un réseau, mais pas nécessairement l'émission d'une réponse en direction des autres régions.

Ce message doit toutefois être nuancé car on s'accorde à penser que, dans la majorité des situations normales, le couplage entre les trois signaux (LFP, décharges neuronales et signaux BOLD) est très étroit. Cette corrélation est démontrée par Logothetis lui-même dans une étude de l'effet du contraste visuel. Une autre étude très récente, menée chez l'homme, a corrélé les enregistrements neuronaux unitaires intracrâniens réalisés chez deux patients épileptiques, avec les activations cérébrales observées en IRMf chez des volontaires sains, à qui on présentait le même stimulus complexe (un film). Une forte corrélation était observée entre les décharges neuronales, le potentiel de champ local entre 40 Hz et 130 Hz, et le signal BOLD (Mukamel *et al.*, 2005).

Dans le sens inverse, une étude a été menée pour explorer, chez le singe macaque, les décharges neuronales dans une région identifiée par l'IRMf pour sa réponse sélective aux visages par rapport aux autres catégories d'objets (Tsao, Freiwald, Tootell, & Livingstone, 2006). Dans cette région, qui pourrait être l'homologue de la région humaine fusiforme sensible aux visages (FFA pour *fusiform face area*), plus de 90 % des neurones répondent sélectivement aux visages. Ainsi observe-t-on à nouveau une excellente congruence entre les propriétés inférées par l'imagerie non-invasive, et celles identifiées par l'électro-physiologie.

En résumé, le psychologue cognitif qui utilise l'imagerie doit garder à l'esprit que le signal qu'il recueille est indirect et pourrait ne pas refléter fidèlement les propriétés de l'activité neuronale. Les méthodes d'imagerie sont toutefois en constante évolution, et le cours a brièvement abordé ce que pourrait nous réserver

l'avenir. La résolution spatiale est en amélioration constante. Certaines études chez l'animal, faisant notamment appel à la réponse IRM précoce ou « early dip », ou à des agents de contraste tels que le MION (micro-particules d'oxyde de fer), parviennent à identifier des structures colonnaires telles que les colonnes de dominance oculaire ou de sensibilité à l'orientation dans l'aire V1. Une nouvelle méthode est également apparue (Le Bihan, Urayama, Aso, Hanakawa, & Fukuyama, 2006). Elle mesure, par l'IRM de diffusion, le gonflement des cellules nerveuses qui accompagne leur entrée en activité. Le signal mesuré par cette méthode serait de nature différente du signal BOLD, mieux localisé et surtout plus rapide de 3 ou 4 secondes. L'avenir dira si cette nouvelle méthode prometteuse permet de suivre, avec un couplage plus étroit, l'évolution temporelle de l'activité cérébrale.

Méthode d'adaptation et organisation interne des représentations

Dans l'attente de méthodes non invasives qui permettraient de voir les colonnes corticales ou les neurones individuels chez l'homme, une stratégie indirecte joue un rôle important dans l'étude de l'organisation interne des représentations cérébrales. C'est la *méthode d'adaptation*, également appelée méthode d'amorçage (Grill-Spector & Malach, 2001 ; Naccache & Dehaene, 2001). Elle se fonde sur la découverte neurophysiologique que la plupart des neurones corticaux, quel que soit le niveau de codage où ils interviennent, montrent une diminution de leur réponse lorsque l'on répète plusieurs fois le même stimulus. En IRMf, un phénomène d'adaptation similaire peut être observé : le signal d'activation cérébrale décroît à mesure que l'on présente un même stimulus plusieurs fois, relativement à une situation de contrôle dans laquelle des stimuli différents sont présentés à chaque essai.

L'étude des propriétés fines de l'adaptation permet deux inférences utiles en psychologie cognitive (Naccache & Dehaene, 2001). Si le signal est moindre dans une région cérébrale donnée quand le stimulus A est suivi de A, que quand B est suivi de A, c'est d'abord que la région cérébrale en question discrimine les stimuli A et B, donc (vraisemblablement) qu'elle contient des populations partiellement différentes de neurones codant pour A et B. Cette inférence peut être valide, alors même que l'imagerie n'a pas la résolution suffisante pour voir, de façon directe, ces deux populations. Ainsi, la méthode d'adaptation offre une forme, certes indirecte, d'hyper-résolution.

En second lieu, l'adaptation permet d'étudier la métrique de similarité entre stimuli pour une région donnée, et donc le niveau de codage de l'information, en examinant quelles variations du stimulus comptent ou ne comptent pas comme une répétition. Dans le cortex occipito-temporal ventral, par exemple, on peut répéter la même image à des tailles différentes, et observer une adaptation inchangée (travaux de Kalanit Grill-Spector, Patrick Vuilleumier, etc.). Mieux, dans la région occipito-temporale gauche, la même adaptation se produit lorsque l'on

répète un mot à l'identique (maison-maison) ou dans une casse différente (maison-MAISON) (Dehaene *et al.*, 2001). On peut alors inférer que cette région code pour les images ou les mots d'une façon invariante pour les propriétés de surface telle que la taille ou l'allure exacte de l'objet.

Le cours a présenté de multiples illustrations du potentiel de cette méthode d'adaptation, empruntés notamment au domaine de la cognition numérique. À l'aide de cette méthode, il est en effet possible de montrer que le cortex intrapariétal code pour les nombres avec une métrique de proximité sémantique : à mesure que les nombres se rapprochent, l'adaptation augmente, indiquant un codage neuronal par des ensembles de neurones qui se recouvrent. L'utilisation de l'amorçage subliminal améliore également la méthode dans la mesure où elle permet d'éviter que des stratégies attentionnelles, non-locales à la région cérébrale étudiée, ne modifient globalement le niveau du signal d'IRMf.

La méthode d'adaptation est toutefois à utiliser avec précaution si l'on veut en tirer des inférences fortes sur le plan du codage neuronal. En effet, la méthode suppose qu'il existe une relation étroite entre la courbe d'accord des neurones et leur courbe d'adaptation en réponse à des stimuli plus ou moins distants de la répétition à l'identique. Or une étude récente (Sawamura, Orban, & Vogels, 2006) a testé directement cette hypothèse par l'électrophysiologie chez le singe éveillé, et a montré qu'il pouvait exister des conditions dans lesquelles elle n'était pas valide.

Mesure du déroulement temporel des opérations cognitives

Une fois les étapes de codage et de traitement identifiées, l'une des questions essentielles que se pose la psychologie cognitive est celle de leur organisation temporelle. Une tâche cognitive peut-elle se décomposer en étapes élémentaires ? Combien de temps prend chaque étape ? Comment ce temps varie-t-il en fonction des paramètres de la tâche ? L'ordre d'exécution des étapes peut-il varier ? Leur exécution s'effectue-t-elle en parallèle ou en série ? La transmission des informations est-elle un processus continu ou discret ? Pourquoi le temps de traitement est-il variable d'essai en essai, ou selon les individus ?

Une imagerie cérébrale de haute résolution temporelle pourrait apporter des réponses à ces questions. Le cours a donc examiné dans quelle mesure l'IRMf, mais aussi d'autres méthodes comme l'électro- et la magnéto-encéphalographie, sont susceptibles de décomposer les activations cérébrales dans le temps.

Une idée répandue affirme que l'IRMf ne possède pas une résolution temporelle satisfaisante pour suivre le déroulement des processus cognitifs. C'est une idée fautive. Dans la mesure où le délai imposé par la fonction hémodynamique est relativement stable, il est possible de détecter des décalages de l'ordre de quelques centaines de millisecondes dans l'entrée en activité de différentes régions cérébrales, ou de la même région dans des conditions expérimentales

différentes. Ravi Menon et ses collègues montrent en particulier qu'il est possible de distinguer des régions dont la durée d'activation varie avec le temps de réaction, et d'autres dont l'entrée en activité indexe le moment de la programmation motrice (Menon & Kim, 1999 ; Menon, Luknowsky, & Gati, 1998).

Au laboratoire, nous avons exploité cette résolution temporelle par le biais de l'analyse spectrale. Dans un paradigme où un stimulus est présenté de façon périodique à intervalles longs (par exemple une phrase toutes les 15 secondes), l'activité évoquée dans les régions impliquées est également périodique. L'extraction de sa phase et de son amplitude donne de précieuses indications sur la latence et l'organisation des aires du langage. Nous avons pu montrer une organisation hiérarchique des régions temporales et frontale inférieure, avec une réponse aux phrases d'autant plus lente que l'on s'éloigne de l'aire auditive primaire, d'une part vers la partie postérieure du lobe temporal gauche, d'autre part vers le pôle temporal bilatéral et la région de Broca gauche (Dehaene-Lambertz *et coll.*, *Human Brain Mapping*, 2006). Cette hiérarchie temporelle, qui suggère des processus enchaînés d'intégration temporelle, est observée dès l'âge de 2-3 mois chez le nourrisson (G. Dehaene-Lambertz *et coll.*, *PNAS*, 2006).

La phase et l'amplitude des activations en IRMf peuvent être exploitées de façon plus fine encore. Supposons que l'expérience manipule un facteur F à plusieurs niveaux. Si F affecte le moment d'entrée en activité d'une région cérébrale, mais pas sa durée, on peut montrer mathématiquement que la phase doit varier linéairement avec F (avec une pente mesurant le ralentissement causé par F), tandis que l'amplitude du signal reste constante. Inversement, si F affecte la durée sans modifier le moment d'entrée en activité, alors la phase doit varier avec F (avec une pente égale à la moitié de la taille de l'effet de F sur la durée du processus mesuré), et l'amplitude doit également varier linéairement avec F. D'autres combinaisons d'effets sont possibles, et chacune est associée à une « signature » de variations de phase et de fréquence avec F. À l'aide de cette méthode, Mariano Sigman, post-doctorant au laboratoire, a montré qu'il était possible de décomposer une tâche cognitive complexe, qui durait un peu moins de deux secondes, en cinq étapes temporelles distinctes (Sigman *et coll.*, *NeuroImage*, 2006).

En dépit de ces avancées, l'IRMf reste limitée dans sa précision temporelle par le flou induit par la lenteur de la réponse hémodynamique. Un accès plus direct au déroulement temporel des opérations cognitives est fourni par l'électro- et la magnéto-encéphalographie (EEG et MEG). En effet, ces méthodes mesurent à l'échelle de la milliseconde les courants électriques et les champs magnétiques minuscules engendrés à l'extérieur de la boîte crânienne par l'entrée en activité synchrone de populations de neurones. Passer en revue l'ensemble des résultats de décomposition temporelle obtenus par EEG et MEG n'entraîne pas le cadre du cours. Nous nous sommes simplement intéressés à savoir si le gain de résolution temporelle s'accompagnait nécessairement d'une perte de résolution spatiale. Dans le cas de l'EEG, s'il est possible de suivre dans le temps une série d'étapes

de traitement, la diffusion et la superposition de leurs topographies à la surface du scalp rend difficile et ambiguë leur reconstruction spatio-temporelle (bien que des progrès significatifs aient été accomplis dans la reconstruction des sources corticales). La MEG, par contre, présente une sélectivité spatiale beaucoup plus grande. Les machines MEG ont connu récemment d'importants progrès. Certaines possèdent plus de 300 capteurs dont la sélectivité spatiale peut être augmentée par une configuration sous forme de « gradiomètre planaire ». Nous avons examiné un travail récent du groupe de Rita Hari, qui est parvenu à décomposer avec une grande finesse temporelle et spatiale cinq étapes successives d'imitation des actions faciales (Nishitani & Hari, 2002).

Autre avancée technologique, l'enregistrement de l'EEG peut maintenant être combiné avec celui des activations cérébrales en IRMf. L'enregistrement simultané EEG+IRMf requiert des amplificateurs adaptés, des électrodes spéciales compatibles avec le champ magnétique de l'imageur, mais surtout un traitement sophistiqué des signaux afin d'éliminer les artefacts issus des gradients d'IRM, du cardio-ballistogramme, et des mouvements induits notamment par la respiration du sujet. La pose des électrodes induit également une perte légère de signal IRMf. Cette méthode doit donc être réservée à des cas bien précis où les signaux d'intérêt ne pourraient pas être mesurés dans des expériences séparées (études de l'épilepsie, de l'activité spontanée endogène du cerveau, etc.).

L'enregistrement EEG+IRMf ouvre également la possibilité d'exploiter les fluctuations spontanées de l'EEG, d'essai en essai, pour proposer une nouvelle stratégie de localisation spatio-temporelle de l'activation cérébrale. L'idée est d'isoler une composante du potentiel évoqué qui occupe une fenêtre temporelle bien définie ; d'en mesurer les fluctuations d'essai en essai ; et d'utiliser ces fluctuations, après convolution avec la fonction hémodynamique canonique, pour prédire les signaux d'IRMf, afin de déterminer quelles sont les régions cérébrales impliquées. La méthode repose sur la relation solide qui existe entre les fluctuations du potentiel local et le signal BOLD mesuré en IRMf (voir plus haut). Il est à noter que c'est la seule méthode de localisation des sources en EEG qui ne nécessite pas de modélisation réaliste de la conductance des tissus. En principe, cette méthode peut fournir une étiquette temporelle aux activations d'IRMf, quelle que soit leur résolution spatiale intrinsèque. Nous avons examiné un exemple très récent d'exploitation de cette technique, qui a permis d'identifier dans le temps et dans l'espace une étape majeure du système de détection et de correction d'erreurs (Debener *et al.*, 2005). Il resterait à montrer, dans des travaux futurs, que plusieurs étiquettes temporelles, chacune associée à une fenêtre différente du potentiel évoqué, peuvent être ainsi apposées aux activations cérébrales afin d'en décomposer plus finement le décours.

L'imagerie cérébrale peut-elle décoder le contenu de la pensée ?

La dernière partie du cours a examiné dans quelle mesure les techniques d'imagerie cérébrale pourraient apporter un éclairage direct aux questions de

psychologie cognitive qui portent sur le contenu, à un instant donné, de l'état mental du sujet. Serait-il possible, à partir d'un examen de neuro-imagerie, d'inférer à quoi pense une personne, ou vers quel objet s'oriente son attention ? En théorie, on ne voit guère quel obstacle majeur pourrait s'opposer à des inférences de cette nature. L'imagerie cérébrale montre des corrélations systématiques entre l'état mental du sujet (supposé connu) et l'activité cérébrale. Le propre d'une corrélation étant d'être symétrique, il devrait être possible d'inverser le processus et d'inférer l'état mental à partir de l'activité cérébrale.

Une telle stratégie de décodage cérébral présenterait plusieurs intérêts. Tout d'abord, la capacité de décoder l'activation cérébrale fournirait une démonstration ultime de notre compréhension du codage cérébral des informations cognitives — si l'on parvient à inférer le contenu mental à partir de l'imagerie, c'est que l'on a compris, en partie au moins, dans quelles régions cérébrales et avec quel format cette information est représentée. On pourrait, ensuite, aborder de front la question de la généralisation de ce code neural à de nouveaux stimuli, et donc celle du degré d'abstraction des représentations cérébrales. En troisième lieu, le décodage cérébral serait susceptible d'apporter des informations invisibles dans le comportement, par exemple la mesure directe de l'orientation de l'attention spatiale. Un décodeur cérébral efficace permettrait peut-être de se passer de toute réponse comportementale. Enfin, le décodage pourrait présenter une utilité pratique dans le domaine des interfaces neuro-informatiques, du « bio-feedback », ou d'autres applications plus controversées comme la détection du mensonge.

Afin de cerner les possibilités et les limites de la neuro-imagerie cognitive, nous avons examiné dans le cours de nombreux exemples publiés de décodage cérébral. Les plus anciens travaux concernent le décodage d'informations cognitives qui se traduisent par des changements macroscopiques de l'activation sur le cortex. En observant les cortex moteurs droit et gauche, il est ainsi possible d'inférer, avec un taux de succès supérieur à 95 %, si le sujet a répondu avec la main gauche ou droite (Dehaene *et al.*, 1998). De même, en examinant l'activité des régions fusiformes et parahippocampiques impliquées dans le codage visuel des visages et des lieux, est-il possible d'inférer la catégorie d'objets qu'une personne voit ou même se contente d'imaginer (O'Craven & Kanwisher, 2000). Soulignons ici que seules peuvent être reconstruites, par ce moyen, les grandes catégories de contenus mentaux qui sont associés à des régions macroscopiquement distinctes du cortex (mots, visages, maisons, etc.).

Plus récemment sont apparues des méthodes d'analyse multivariées qui tentent de restituer les informations codées par des propriétés neuronales « méso-scopiques » (colonnes corticales). L'idée en est simple. Lorsque l'on compare, dans chaque voxel, l'activation évoquée par deux conditions expérimentales minimalement différentes (par exemple regarder une chaise ou une théière), dans les cas les plus difficiles, il n'y a pas de grande différence d'activation. Il se peut, par contre, qu'il y ait de petites différences, distribuées plus ou moins aléatoirement dans les voxels en fonction de la densité plus ou moins grande de colonnes

corticales de l'un ou de l'autre type à la surface du cortex. Si ces différences sont indépendantes les unes des autres, leur configuration globale peut permettre de classifier les conditions expérimentales avec un degré de réussite qui dépasse le niveau du hasard. De nouvelles méthodes multivariées (machines à vecteurs de support) permettent d'extraire ces informations même lorsqu'elles sont enfouies dans un espace de grande dimensionnalité.

Nous avons examiné dans le cours trois articles récents qui ont exploité ces méthodes multi-variées pour extraire de l'IRM des informations fines sur la nature de la stimulation visuelle présentée aux sujets (Haynes & Rees, 2005 ; Kamitani & Tong, 2005, 2006). Leurs résultats montrent que (1) il est possible d'inférer, à partir d'un examen d'IRMf standard, l'orientation de barres ou la direction d'un mouvement présenté sur la rétine ; (2) cette inférence demande l'acquisition préalable, pour chaque sujet, de données permettant l'entraînement d'un classificateur multi-varié ; (3) une fois l'entraînement effectué, le classificateur peut être utilisé pour classifier de nouvelles images d'activité cérébrale, même acquises plusieurs semaines plus tard ; (4) lorsque deux stimuli sont présentés simultanément sur la rétine, le décodeur permet d'inférer celui auquel le participant fait actuellement attention ; (5) lorsque les stimuli sont masqués en sorte que le participant ne puisse pas rapporter consciemment leur identité, le décodeur continue de répondre avec un taux de succès supérieur au niveau du hasard. L'ensemble de ces résultats souligne la faisabilité du décodage d'un méso-code cortical des aires visuelles.

Au laboratoire, une expérience réalisée par Bertrand Thirion en 2005 a montré qu'il était possible de reconstituer, à partir des activations cérébrales, un stimulus plus complexe que la simple orientation ou la direction de mouvement. Dans cette expérience, les participants voyaient ou imaginaient des « dominos » constitués de cinq points occupant les cases d'une matrice 3x3. Sur la base d'une expérience préalable de rétinopathie, chaque voxel des aires visuelles était associé à une position préférée du champ visuel. Il était alors possible, étant donné une configuration d'activation des aires visuelles, de tenter de reconstituer le stimulus correspondant, chaque voxel fournissant une information statistique sur la présence ou non d'un stimulus au point correspondant du champ visuel. La reconstitution de stimuli visuels réels était excellente. Celle de stimuli imaginés était moins satisfaisante, toutefois le côté du champ visuel et, pour cinq hémisphères, l'identité de l'image mentale étaient reconstitués avec une probabilité supérieure au niveau du hasard.

L'avènement d'algorithmes de traitement des images d'IRM « en temps réel », qui rendent les statistiques disponibles moins d'une seconde après l'acquisition d'une coupe du cerveau, ouvre la voie à des applications pratiques du décodage cérébral. Il devient possible de lire l'activation d'une région cérébrale, et de l'utiliser pour commander la réponse d'un robot ou d'un logiciel. En montrant à la personne en train d'être scannée le niveau d'activation atteint par une région de son cerveau (« *bio-feedback* »), il est possible de lui apprendre à contrôler

cette activation. Un groupe de chercheurs a récemment employé cette technique afin d'entraîner les participants à contrôler la douleur en modulant le niveau d'activité du cortex cingulaire antérieur (deCharms *et al.*, 2005). Chez un groupe de patients souffrant de douleurs chroniques, cette méthode entraîne un changement durable de l'intensité subjective de la douleur. D'autres applications cliniques peuvent être envisagées, notamment pour faciliter la communication avec des patients paralysés ou atteints du syndrome d'enfermement (« *locked-in syndrome* »).

Une application plus controversée est la détection du mensonge. Parmi les travaux récents, Davatzikos *et coll.* (2005) sont parvenus à classifier correctement, à l'aide d'une machine à vecteur de support, près de 90 % des essais individuels dans une tâche où un volontaire nie posséder une carte à jouer qui lui a été assignée. Dans de nombreuses études, l'activation des régions préfrontales latérales et médianes contribue principalement à prédire avec succès les essais où le sujet ment, ce qui est compatible avec l'hypothèse que le mensonge requiert un effort pour aller à l'encontre de la réponse correcte, ainsi peut-être qu'un découplage du savoir d'autrui et du sien. Toutefois, aucune de ces études ne montre qu'un décodeur puisse généraliser à des mensonges de nature très variée. De fait, la sensibilité et la spécificité de ces méthodes sont douteuses. Le décodage des états mentaux sur la base de l'activation repose sur l'hypothèse qu'à chaque catégorie d'état mental correspond une catégorie restreinte d'états cérébraux. Dans le cas du mensonge, on imagine mal que l'immense variété des mensonges possibles se réduise, de façon simple, à un état d'activité cérébrale unique et aisément détectable.

Le cours s'est conclu en rappelant que, si l'imagerie cérébrale apporte des informations essentielles à la psychologie cognitive et contribue à une grande variété d'applications cliniques, les examens d'imagerie demandent une collaboration étroite du participant qui s'y prête. Il serait facile à celui-ci de perturber l'examen au point de le rendre inutilisable, par exemple en bougeant la tête ou en effectuant des activités mentales autres que celles demandées. C'est pourquoi, même s'il convient de rester extrêmement vigilant sur les aspects éthiques de l'utilisation de l'imagerie cérébrale, il paraît plus aisé d'en imaginer des applications utiles à la société que des usages mal intentionnés.

SÉMINAIRE

Le séminaire a porté sur la modularité et la décomposition des opérations mentales. En complément du cours, il s'agissait d'examiner comment les méthodologies autres que l'imagerie cérébrales peuvent contribuer à l'analyse de l'architecture des opérations cognitives. Quatre invités se sont prêtés aux débats sur ce sujet. Le professeur Ardi Roelofs (Institut Max-Planck de psycholinguistique, Nimègue, Pays-Bas) a décrit ses travaux sur la décomposition fonctionnelle de

la production verbale par le biais de méthodes chronométriques et d'interférence. Le professeur Saul Sternberg (Université de Pennsylvanie, Philadelphie, États-Unis) a présenté une méthodologie générale pour l'identification de « modules » de traitement de l'information. Emmanuel Dupoux, directeur d'études à l'École des Hautes Études en Sciences Sociales, a détaillé les niveaux d'organisation que l'on peut identifier au sein de la chaîne de perception du langage parlé. Enfin le professeur Alfonso Caramazza (département de psychologie, Université Harvard, États-Unis) s'est exprimé sur l'origine des catégories grammaticales de noms et de verbes, en s'appuyant sur un ensemble de données issues de la neuropsychologie, de la stimulation magnétique trans-crânienne, et de l'imagerie cérébrale.

ACTIVITÉS DE RECHERCHE DU LABORATOIRE

On ne décrira ici que quelques-uns des principaux axes de recherches du laboratoire, publiés entre 2005 et 2006.

Traitement subliminal des mots émotionnels

L'un des thèmes de recherche du laboratoire concerne la profondeur du traitement subliminal des mots. Est-il possible, par exemple, qu'un mot soit traité à un niveau sémantique sans pour autant être vu consciemment par le sujet ? Nos travaux antérieurs ont suggéré que tel était le cas dans le domaine limité des symboles numériques (Dehaene *et coll.*, 1998). En cernant les limites de la perception subliminale, ces expériences tentent de délimiter le contour et les fonctions spécifiques d'un « espace de travail global conscient » (Dehaene, Kerszberg et Changeux, 1998).

L'une de nos stratégies de recherche a consisté à enregistrer l'activité cérébrale associée au traitement de certains traits sémantiques. Pour des raisons strictement cliniques, certains patients se voient implanter des électrodes intracrânielles, fréquemment dans la région temporale interne et l'amygdale, une région impliquée dans la perception des stimuli évoquant des émotions négatives telle que la peur. Avec leur consentement, il est alors possible d'enregistrer les réponses intracrâniennes des patients à des stimuli sub- ou supra-liminaux. Lionel Naccache et Raphaël Gaillard, en collaboration avec l'équipe d'épileptologie de l'hôpital de la Salpêtrière (Professeurs Baulac et Adam) ont ainsi recueilli des enregistrements intracrâniens dans l'amygdale bilatérale chez trois patients au cours de la présentation subliminale ou consciente de mots émotionnels ou neutres (Naccache *et coll.*, *PNAS*, 2005). Les résultats ont montré une activation de l'amygdale différenciant les mots émotionnels des mots neutres. D'intéressantes différences entre activations consciente et non-consciente ont été observées. Comme prédit par notre modèle théorique (Dehaene, Kerszberg et Changeux, 1998), l'état non-conscient est associé à des activations brèves tandis que l'état conscient montre une amplifi-

cation et une plus grande durabilité de l'activation de l'amygdale pour les mots émotionnels.

Un second travail, réalisé principalement par Raphaël Gaillard et Antoine Delcul, a analysé le seuil d'accès à la conscience de mots émotionnels ou neutres pendant le masquage (Gaillard *et coll.*, *PNAS*, 2006). Notre théorie de l'espace de travail global conscient, comme d'autres théories (par exemple celle de la compétition biaisée proposée par John Duncan) prédit que le seuil doit être abaissé si le stimulus reçoit un traitement neuronal plus intense, par exemple parce qu'il possède une représentation neuronale plus étendue. Raphaël Gaillard et Antoine Delcul ont donc présenté comme cibles dans une expérience de masquage, des mots soit associés à une émotion négative (ex. : requin, viol...), soit neutres (ex. : maison, table). Indépendamment de cette valence, ils ont également fait varier l'ordre de passation des différents délais entre cible et masque. Le sujet devait à la fois rapporter s'il avait vu ou pas vu le mot et, dans tous les cas, tenter de le nommer. Les résultats ont montré que les deux variables — valence du mot, et ordre de présentation — ont un impact important sur le seuil de perception consciente. Le seuil non linéaire d'accès au rapport conscient est abaissé pour les mots émotionnels, mais également pour les mots qui ont déjà été perçus auparavant. Cette modification du seuil indique qu'avant même l'accès à la conscience, les informations de valence sémantique et de récence ont été extraites. L'ensemble de ces résultats fournit une méthodologie nouvelle pour étudier l'étendue de l'accès non conscient à la sémantique des mots.

Amorçage subliminal des mots écrits en Japonais

Dans la même ligne de recherches, un chercheur étranger en visite au laboratoire, Kimihiro Nakamura, et un chercheur post-doctorant, Sid Kouider, ont réalisé une expérience qui a étendu nos résultats d'amorçage subliminal. En 2001, nous avons montré que, chez des lecteurs adultes du français, la région occipito-temporale gauche impliquée dans l'extraction de la forme visuelle des mots montre un effet d'amorçage subliminal (Dehaene *et coll.*, *Nature Neuroscience*, 2001). Son activation, mesurée par IRM fonctionnelle, se réduit lorsque l'on présente un mot cible précédé d'une présentation subliminale du même mot comme amorce masquée. Cet effet d'amorçage reste inchangé si le mot est répété avec un changement de majuscules à minuscules (par exemple, amorce « piano », cible « PIANO »), ce qui indique qu'un niveau de représentation invariant pour la casse est associé à cette région et peut être activé en l'absence de toute conscience.

Notre nouvelle expérience a étendu ce résultat au cas de l'écriture japonaise. Dans cette langue co-existent plusieurs systèmes d'écriture, notamment les Kanji (vaste jeu de caractères représentant chacun un mot ou un morphème) et les Kana (jeu limité de caractères représentant les syllabes). La différence Kanji-Kana constitue une source majeure de variabilité des formes de surface de

l'écriture japonaise, qui dépasse très largement la différence entre majuscules et minuscules en français. Nous avons choisi un ensemble de mots qui s'écrivent indifféremment en Kanji et en Kana, et avons examiné si le cerveau des lecteurs des japonais peut reconnaître non consciemment qu'il s'agit des mêmes mots. L'expérience a montré que tel était le cas. Sur le plan comportemental, l'effet d'amorçage subliminal de répétition était significatif même lorsque l'amorce et la cible apparaissaient dans des notations différentes. En imagerie, cet effet d'amorçage inter-notations trouvait son origine, non pas dans la région occipito-temporale gauche (où seul apparaissait un effet d'amorçage Kanji-Kanji), mais dans une autre région latérale du cortex temporal gauche, associée par d'autres travaux (Devlin *et coll.*, 2004) aux associations sémantiques entre synonymes. Ainsi, un niveau de traitement abstrait, totalement indépendant des formes de surface de l'écriture, peut-il être atteint en l'absence de conscience. Ces résultats contribuent à un début d'analyse de la hiérarchie de régions corticales postérieures impliquées dans la lecture.

Traitement subliminal et accès conscient dans la tâche de « clignement attentionnel »

Afin de mieux comprendre la dynamique d'accès à la conscience, la thèse de Claire Sergent a été consacrée à l'analyse d'un autre paradigme expérimental qui permet de présenter des mots en dessous du seuil de conscience : le paradigme d'*attentional blink*, imparfaitement traduit en français par le terme de « clignement attentionnel » (Sergent *et coll.*, Nature Neuroscience, 2005). Dans cette situation expérimentale, un premier stimulus (T1), activement traité par le sujet, rend invisible, pendant une brève période de temps, la présentation d'un second stimulus (T2). Nos propres travaux ont d'abord consisté à trouver des conditions expérimentales dans lesquelles, de façon stochastique, le même stimulus T2 était parfois perçu parfaitement, et parfois jugé strictement invisible par les mêmes sujets. Ce paradigme fournissait donc des conditions optimales pour l'étude des différences entre traitements conscient et non conscient d'un mot écrit.

Nous avons ensuite recueilli des évaluations quantitatives de visibilité ainsi que des enregistrements des potentiels évoqués à 128 électrodes afin de mesurer l'évolution temporelle de l'activité cérébrale au cours de l'accès à la conscience. Les potentiels évoqués précoces (P1 et N1) étaient identiques, indépendamment de l'expérience du sujet, ce qui suggère que ces événements cérébraux ne sont pas les corrélats primaires de la perception consciente. Cependant, vers 270 ms, nous avons observé une soudaine bifurcation entre essais conscients et non conscients, avec une cascade d'événements cérébraux évoqués exclusivement par la condition consciente. L'analyse des sources corticales distribuées, réalisée avec l'aide de Sylvain Baillet et son logiciel BrainStorm, a montré que ces événements correspondaient à l'entrée en activité d'un réseau distribué impliquant notamment les régions temporales latérales et antérieures et de vastes secteurs du cortex pré-

frontal et cingulaire. Une réactivation pariétale et occipitale tardive était également visible.

Ces résultats ont été les premiers à révéler aussi précisément la séquence d'activité associée à l'accès à la conscience, par contraste avec une situation non consciente minimalement différente. Ils suggèrent que la transition vers l'accès conscient est associée au déclenchement optionnel, non linéaire et stochastique, d'une vague tardive d'activité cérébrale qui se déploie dans un réseau distribué et réverbérant dont le cortex préfrontal constitue l'un des nœuds essentiels.

Simulations d'un réseau thalamo-cortical global associé à l'accès conscient

Pour mieux comprendre l'origine de tels effets non linéaires d'amplification soudaine de l'activité corticale, Jean-Pierre Changeux et Stanislas Dehaene ont développé, depuis plusieurs années, un modèle neuro-informatique détaillé des réseaux thalamo-corticaux. Cette simulation, neuro-réaliste, comprend un vaste ensemble de neurones excitateurs et inhibiteurs, organisés en couches corticales et en aires cérébrales élémentaires, reliées à longue distance par des connections glutamatergiques ascendantes (NMDA) et descendantes (AMPA).

Dans un article récent (Dehaene et Changeux, *PLOS : Biology*, 2005), nous avons étudié les conditions dans lesquelles une activité neuronale spontanée émerge dans ce réseau, et l'interaction de cette activité spontanée avec une stimulation sensorielle externe. Les simulations ont caractérisé deux états principaux d'activité. En premier lieu, l'activité spontanée apparaît soudainement, avec des fréquences d'oscillation concentrées dans la bande gamma, lorsque l'on manipule l'intensité d'un signal ascendant de neuromodulation qui modifie l'« éveil » cortical. Cette activité spontanée modérée facilite l'entrée de stimulations sensorielles externes. En second lieu, au sein d'un tel réseau spontanément actif, nous observons parfois un « embrasement » soudain de neurones distribués dans plusieurs aires, interconnectés à longue distance, et constituant des états cohérents de représentation du réseau. Pendant cet embrasement, l'activité spontanée est élevée et peut bloquer le traitement d'entrées sensorielles concurrentes.

Nous relierons ces propriétés à de nombreuses observations expérimentales sur les bases neurales des états endogènes de conscience, et notamment le blocage de l'accès à la conscience qui survient dans les paradigmes de « clignement attentionnel » et de « cécité inattentionnelle » (*inattentional blindness*). Dans ces situations, un sujet normal, engagé dans une intense activité mentale, peut ne pas prendre conscience de stimuli sensoriels saillants mais dépourvus de pertinence. Le modèle reproduit les grands traits des résultats expérimentaux observés dans ces paradigmes, et notamment ceux obtenus dans notre laboratoire par Claire Sergent. Bien que nos simulations soient très simplifiées, leurs propriétés génériques clarifient ainsi quelques-uns des phénomènes fondamentaux associés à l'autonomie de la conscience d'accès.

Prise de décision numérique et interférence entre tâches cognitives

Mariano Sigman, post-doctorant au laboratoire, s'est intéressé à l'étude expérimentale d'une autre situation qui met en évidence l'interférence centrale entre tâches : le phénomène de « période psychologique réfractaire » (PRP pour *psychological refractory period*). Cette situation est très proche du clignement attentionnel, l'unique différence étant que les stimuli T1 et T2 ne sont pas masqués. Lorsque un intervalle court sépare T1 et T2, T1 est traité normalement, mais le temps de réponse à T2 est fortement ralenti.

Notre travail a mis à l'épreuve l'hypothèse, dérivée du modèle de l'espace de travail global, selon laquelle seule une étape centrale du traitement de l'information est responsable de la PRP (Sigman et Dehaene, *PLOS : Biology*, 2005). Notre modèle suppose que chaque tâche se décompose en trois étapes — perception, décision fondée sur une marche aléatoire, et réponse motrice — et que toutes ces étapes peuvent fonctionner en parallèle, sauf la décision elle-même ; cette dernière constituerait un goulot d'étranglement où seule une tâche peut s'exécuter à la fois. Nous avons testé ces idées dans une tâche de comparaison de nombre, bien connue du laboratoire, en explorant systématiquement l'espace des paramètres de notation numérique, distance numérique, complexité de la réponse, et délai de présentation de T1 et T2, relativement à une tâche concurrente de jugement de hauteur tonale. Les résultats ont révélé une décomposition de la tâche de comparaison où chaque facteur expérimental n'affectait qu'une seule étape de traitement. En particulier, la distance numérique affecte le processus de décision, qui est également la seule étape à ne pas s'exécuter en parallèle et à contribuer de façon prédominante à la variabilité des temps de réponse. Les autres étapes perceptives et motrices s'exécutent en parallèle avec la tâche auditive, et la modification de leur durée n'a pas d'effet sur la variance des temps de réaction.

Nous en avons déduit une architecture cognitive de la prise de décision qui relie deux branches précédemment distinctes de la psychologie cognitive : les modèles de marche aléatoire pour le temps de réaction, et les modèles à deux étapes de l'interférence centrale entre tâches. Mariano Sigman utilise à présent l'imagerie cérébrale en IRM fonctionnelle et en potentiels évoqués afin d'évaluer les bases cérébrales de cette architecture.

Rôle d'une région occipito-temporale gauche dans la reconnaissance visuelle des mots

Un travail expérimental sur l'organisation fine de la reconnaissance visuelle des mots, étape essentielle de la lecture, a été réalisé en collaboration avec Laurent Cohen et son équipe. Depuis plusieurs années, notre stratégie consiste à évaluer par l'imagerie cognitive la nature de la contribution d'une région cérébrale occipito-temporale gauche à la lecture (VWFA pour *visual word form*

area). Bien que la plupart de nos expériences soient menées avec l'IRM fonctionnelle chez le volontaire sain, des données remarquables ont été obtenues cette année lors de l'étude d'un cas de patient épileptique conjointement étudié par le comportement, l'IRMf, et les enregistrements intracrâniens (Gaillard *et al.*, *Neuron*, 2006).

Notre modèle de la lecture postule que la VWFA implémente une série d'étapes automatiques de reconnaissance invariante des lettres et des mots, indispensable à la lecture rapide. Cette théorie a cependant été contestée, notamment par Cathy Price, qui postule que cette région joue un rôle générique dans la reconnaissance et la dénomination des objets. Nous avons mis à l'épreuve ces possibilités chez un patient qui, pour des raisons d'exploration clinique, fit par deux fois l'objet d'une implantation d'électrodes intracrâniennes dans la région occipito-temporale ventrale, puis fut opéré afin d'ôter un foyer épileptique tout proche de la VWFA. Avant l'intervention, la lecture était normale, et l'IRMf montrait une mosaïque normale d'activations évoquées par les mots, les visages, les maisons et les outils. Les enregistrements intra-crâniens montraient que la VWFA répondait avec une latence rapide aux mots conscients ainsi qu'aux mots subliminaux. Après l'intervention, le patient présentait un déficit prononcé de la lecture des mots, tandis que la reconnaissance des autres catégories visuelles demeurait intacte. De plus, sur la carte du cortex visuel, seule l'activation spécifique aux mots avait disparu.

Globalement, ces résultats soutiennent l'hypothèse d'un rôle causal et spécifique de la région occipito-temporale gauche (VWFA) dans la reconnaissance visuelle des mots. Ils viennent confirmer, un siècle plus tard et avec des technologies avancées, les observations neuropsychologiques présentées par Déjerine dès 1892 sur le syndrome d'alexie pure et son lien avec le cortex occipito-temporal ventral.

Modélisation de l'architecture de la reconnaissance visuelle des mots

Comment le réseau occipito-temporal peut-il s'organiser afin de reconnaître les mots écrits ? Depuis plusieurs années, le laboratoire formule et met à l'épreuve une théorie du « recyclage neuronal » au cours de l'apprentissage de la lecture. Selon celle-ci, bien que rien au cours de l'évolution du cerveau des primates ne l'ait spécifiquement préparé à une activité culturelle aussi nouvelle que la lecture, l'apprentissage, sans jamais créer *de novo* des circuits cérébraux entièrement différents, peut minimalement modifier certains de nos réseaux afin de les réorienter vers des fonctions culturellement apprises. C'est ainsi que les régions occipito-temporales ventrales, impliquées chez le singe macaque dans la reconnaissance invariante des objets, des scènes et des visages, semblent pouvoir se reconverter, au moins partiellement, afin d'acquérir une sélectivité pour les lettres et les mots et de transmettre leur identité aux régions impliquées dans la phonologie et la sémantique du langage parlé.

Sur la base d'une analyse détaillée de l'organisation de la région inféro-temporale du singe, nous avons récemment proposé un modèle détaillé de son recyclage pour la reconnaissance des mots écrits (Dehaene *et coll.*, *Trends in Cognitive Science*, 2005). Le code neural des mots doit être abstrait, puisque nous parvenons à les reconnaître indépendamment de leur position, de leur taille, et de leur police. Il doit également être sensible à l'identité et à l'ordre des lettres. La plupart des codes qui ont été proposés en psychologie cognitive sont, soit insuffisamment invariants, soit incompatibles avec les contraintes de l'organisation du système visuel des primates. Nous avons proposé un modèle selon lequel une partie de la voie occipito-temporale ventrale se reconvertit pour former une hiérarchie de « détecteurs de combinaisons locales » sensibles à des fragments de mots de taille croissante : fragments de lettres dans l'aire V2, caractères dans l'aire V4, puis, dans les régions plus antérieures, identité abstraite des lettres, bigrammes, et graphèmes et morphèmes complexes.

Notre proposition présente l'avantage d'expliquer pourquoi la détection de « bigrammes ouverts » (paires ordonnées de lettres, éventuellement séparées par une ou deux autres lettres) constitue une étape importante de la reconnaissance visuelle des mots, comme l'ont notamment suggéré certains résultats empiriques de l'équipe de Jonathan Grainger. Une nouvelle série d'expériences d'imagerie et de neuropsychologie a été lancée en 2005 afin de mettre cette hypothèse à l'épreuve.

Architecture cérébrale du traitement du langage parlé

Les équipes dirigées par Christophe Pallier et Ghislaine Dehaene-Lambertz étudient l'architecture des aires du langage, respectivement chez l'adulte et le très jeune enfant. Dans le cadre d'un projet européen (réseau Neurocom), une série d'expériences ont été conduites afin de visualiser en IRM fonctionnelle les aires activées par les phrases et certaines paramètres de leur architecture fonctionnelle.

Dans un premier temps, l'expérimentation chez l'adulte a montré qu'il était aisé d'activer l'ensemble des aires périsylvienne du langage en présentant, de façon passive, de courtes phrases de 2 à 3 secondes présentées à des intervalles longs (> 14 secondes). Une méthode d'analyse spectrale, décrite dans le cours, a permis de démontrer une organisation hiérarchique des délais temporels dans le lobe temporal et la région frontale inférieure gauche (Dehaene-Lambertz *et coll.*, *Human Brain Mapping*, 2006). Cette architecture temporelle hiérarchique a ensuite été retrouvée par IRMf chez le nourrisson de quelques mois — une observation qui nous semble fondamentale dans la mesure où elle pointe peut-être vers un biais d'origine génétique qui structurerait l'acquisition ultérieure des structures hiérarchiques enchâssées du langage parlé.

Un second point important est que la plupart de ces régions, le long du sillon temporal supérieur, sont sensibles à la répétition des phrases. Un effet d'adaptation a été observé : l'activation diminue à la seconde présentation d'une phrase.

Dans l'esprit de la méthode d'adaptation décrite dans le cours, nous avons exploité des variantes de cet effet dans plusieurs expériences d'imagerie qui visaient à comprendre la spécialisation fonctionnelle des aires du langage. Nous avons d'abord étudié l'invariance pour le locuteur, et montré que deux régions du sillon temporal supérieur gauche détectent la répétition d'une phrase de façon totalement indépendante de la personne qui parle (tandis qu'une région homologue droite s'intéresse au locuteur). Nous avons ensuite étudié la possibilité d'un effet d'amorçage syntaxique et sémantique (travaux de thèse d'Anne-Dominique Devauchelle). Les premiers résultats, en cours d'écriture, montrent que certaines régions, en particulier temporelles antérieures, s'intéressent au contenu sémantique des phrases dans la mesure où elles sont sensibles à la répétition du même contenu par des structures syntaxiques différentes. Inversement, nous n'avons pas pu, pour l'instant, observer de région cérébrale qui serait sensible à la seule structure de l'arbre syntaxique, indépendamment du contenu lexico-sémantique des phrases. Les travaux se poursuivent actuellement afin de comprendre quels niveaux d'organisation du langage parlé sont associés à chaque région activée.

Noyau universel de connaissances géométriques

Une équipe dirigée par Stanislas Dehaene et comprenant Véronique Izard, thésarde au laboratoire, Pierre Pica (CNRS), et Elizabeth Spelke (Harvard) s'intéresse depuis plusieurs années aux représentations proto-mathématiques qui fondent l'intuition des nombres, de l'espace et du temps. Grâce à Pierre Pica, une série d'expériences sur les intuitions géométriques a pu être menée au Brésil avec des volontaires Munduruku, un peuple d'Amazonie dont nous avons déjà décrit les compétences numériques (Pica *et coll.*, *Science*, 2004). Notre recherche pose la question classique de la nature de notre compétence géométrique (Dehaene *et coll.*, *Science*, 2006). La géométrie constitue-t-elle un noyau d'intuitions présentes chez tous les membres de l'espèce humaine, quel que soit leur langue ou leur niveau d'éducation ? Nous avons utilisé deux tests non-verbaux portant sur les primitives conceptuelles de la géométrie. Les résultats ont montré que les enfants et les adultes Mundurukus utilisent spontanément des concepts géométriques de base tels que la notion de point, de ligne, de parallélisme ou d'angle droit, afin de détecter des intrus dans un jeu d'images simples. De plus, ils utilisent les relations de distance, d'angle et d'orientation entre des points sur une carte géométrique schématisée pour localiser avec précision des objets dans le monde réel. Dans ces deux tests, les performances des enfants et des adultes Mundurukus étaient étroitement corrélées à celles d'enfants et d'adultes américains, même si ce dernier groupe montrait des taux de réussite globalement plus élevés.

Ainsi, nos résultats suggèrent qu'en dépit de la variabilité des cultures, nous partageons tous certaines intuitions géométriques profondes, même en l'absence de scolarisation, d'interactions avec des symboles graphiques et des cartes, et d'un riche lexique de termes géométriques. Dans les prochaines années, nous

espérons explorer plus avant les mécanismes, les origines, les limites et les bases cérébrales de ce noyau de compétences géométriques.

PUBLICATIONS DU LABORATOIRE

2005

BARTH, H., LA MONT K., LIPTON, J., DEHAENE, S., KANWISCHER, N. SPELKE, E. Non-symbolic arithmetic in adults and young children. *Cognition*, 98, 1-24, 2005.

DEHAENE, S., CHANGEUX, J.P. Ongoing Spontaneous Activity Controls Access to Consciousness : A Neuronal Model for Inattentive Blindness. *PLOS : Biology*, 3, 910-927, 2005.

DEHAENE, S., COHEN, L., SIGMAN, M., VINCKIER, F. The neural code for written words : a proposal. *Trends in Cognitive Neurosciences*, 9, 335-341, 2005.

DEHAENE-LAMBERTZ, G., PALLIER, C., SERNICLAES, W. SPRENGER-CHAROLLES, L., JOBERT, A., DEHAENE, S. Neural correlates of switching from auditory to speech mode. *NeuroImage*, 24, 21-33, 2005.

DODEL, S., GOLESTANI, N., PALLIER, C., ELKOUBY, V., LE BIHAN, D., POLINE, J.B. Condition-dependent functional connectivity : syntax networks in bilinguals. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 360, 921-935, 2005.

GAILLARD, R., NACCACHE, L. Amygdala recording of emotions evoked by subliminal words. *Medecine Science (Paris)*, 10, 804-805, 2005.

GLIGA, T. & DEHAENE-LAMBERTZ, G. Structural encoding of body and face in human infants and adults. *Journal of Cognitive Neuroscience* 17, 1328-1340, 2005.

HENRY, C., GAILLARD, R., VOLLE, E., CHIRAS, J., FERRIEUX, S., DEHAENE, S., COHEN, L. Brain activations during letter-by-letter reading : A follow-up study. *Neuropsychologia*, 43, 1983-1989, 2005.

HUBBARD, E.M., PIAZZA, M., PINEL, P., DEHAENE, S. Interactions between number and space in parietal cortex. *Nature Reviews in Neuroscience*, 6, 435-447, 2005.

MOLKO, N., WILSON, A. & DEHAENE, S. La dyscalculie développementale, un trouble primaire de la perception des nombres. *Médecine & Enfance* 25, 165-170, 2005.

NACCACHE, L., DEHAENE, S., COHEN, L., HABERT, M.-O., GUICHART-GOMEZ, E., GALANAUD, D., WILLER, J.-C. Effortless control : Executive attention and conscious feeling of mental effort are dissociable. *Neuropsychologia*, 43, 1318-1328, 2005.

NACCACHE, L., GAILLARD, R., ADAM, C., HASBOUN, D., CLEMENCEAU, S., BAULAC, M., DEHAENE, S., COHEN, L. A direct intracranial record of emotions evoked by subliminal words. *PNAS*, 102, 7713-7717, 2005.

NACCACHE, L., PUYBASSET, L., GAILLARD, R., SERVE, E., WILLER, J.C. Auditory mismatch negativity is a good predictor of awakening in comatose patients : a fast and reliable procedure. *Clin. Neurophysiol.*, 116, 988-989, 2005.

NACCACHE, L., HABERT, M.O., MALEK, Z., COHEN, L. & WILLER, J.C. Activation of secondary auditory cortex in a deaf patient during song hallucinosis. *J. Neurol.* 252, 738-739, 2005.

NAKAMURA, K., DEHAENE, S., JOBERT, A., LE BIHAN, D., & KOUIDER, S. Subliminal convergence of Kanji and Kana words : Further evidence for functional parcellation of the posterior temporal cortex in visual word perception. *Journal of Cognitive Neuroscience*, 17, 954-968, 2005.

SANGRIGOLI, S., PALLIER, C., ARGENTI, A.M., VENTUREYRA, V. & DE SCHONEN, S. Reversibility of the other-race affect in face recognition during childhood. *Psychological Science* 16, 440-444, 2005.

SERGENT, C., & DEHAENE, S. Neural processes underlying conscious perception : Experimental findings and a global neuronal workspace framework. *J. Physiol. Paris* 98, 374-384, 2005.

SERGENT, C., BAILLET, S., DEHAENE, S. Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, 8, 1391-1400, 2005.

SIGMAN, M. PAN, H., YANG, Y., STERN, E., SIBERSWEIG, D. GILBERT, C.D. Top-down reorganization of activity in the visual pathway after learning a shape identification task. *Neuron*, 46, 823-835, 2005.

SIGMAN, M., DEHAENE, S. Parsing a cognitive task : A characterization of the mind's bottleneck. *PLOS : Biology*, 3, e37, 2005.

VIARD, A., FLAMENT, M.F., ARTIGES, E., DEHAENE, S., NACCACHE, L., COHEN, D., MAZET, P. MOUREN, M.C., MARTINOT, J.L. Cognitive control in childhood-onset obsessive-compulsive disorder : a functional MRI study. *Psychol. Med.*, 35, 1007-1017, 2005.

COHEN, L., JOBERT, A., LE BIHAN, D., DEHAENE, S. Distinct unimodal and crossmodal regions for word processing in the left temporal cortex. *Neuroimage*, 23, 1256-1270, 2005.

Livre

DEHAENE, S., DUHAMEL, J.R., HAUSER, M., RIZZOLATTI, G. (Ed.) *From Monkey Brain to Human Brain*. (Fyssen Foundation workshop) MIT Press, 2005.

Chapitres de livres

DEHAENE, S. Consciousness and Cautiousness. In *Progress in Brain research.*, S. Laureys (Ed.), 2005, pp. ix-x.

DEHAENE, S. Evolution of human cortical circuits for reading and arithmetic : The « neuronal recycling » hypothesis. In Stanislas Dehaene, Jean-René Duha-

mel, Marc D. Hauser & G. Rizzolatti (Eds.), *From Monkey Brain to Human Brain*. A Fyssen Foundation Symposium, chapter 8, pages 133-157. MIT Press, 2005.

DEHAENE, S. How a primate brain comes to know some mathematical truths. In *Neurology of Human Values*, J.P. Changeux, A.R. Damasio, W. Singer & Y. Christen (Eds.). Springer, 2005, pp. 143-155.

DEHAENE, S. Imaging conscious and subliminal word processing. In U. Mayr, E. Awh & S. Keele (Eds.), *Developing individuality in the human brain : A tribute to Michael Posner*. Washington : American Psychological Association, 2005, pp. 65-86.

DEHAENE, S., MAQUET, P., REES, G., FLYTCHÉ, D., LAMME, V., LOPES DA SILVA, F. Consciousness and the brain. In Fundação Bial (Eds.), *5º Simposio da Fundação Bial, Aquém e Além do Cérebro, Behind and beyond the brain*, 2005, pp. 159-165.

DEHAENE-LAMBERTZ, G. Neuropsychologie de l'enfant et troubles du développement. Hommet, I., Jambaqué, C., Billard, C. & Gillet, P. (Éds.), pp. 59-73 (Solal, Marseille, 2005).

DEHAENE-LAMBERTZ, G. Graines de Sciences. Jamous, M. & Saltiel, E. (Éds.), pp. 27-54 Paris, 2005.

NACCACHE. L. Visual phenomenal consciousness : a neurological guided tour. In *Progress in Brain Research.*, S. Laureys (Ed.), 2005, pp. 185-195.

2006

ALARIO, F.X., CHAINAY, H., LEHERICY, S. & COHEN, L. The role of the supplementary motor area (SMA) in word production. *Cognitive Brain Research*, 2006, 107, 129-143.

ANDOH, J., ARTIGES, E., PALLIER, C., RIVIÈRE, D., MANGIN, J.F., CACHIA, A., PLAZE, M., PAILLÈRE-MARTINOT, M.L., MARTINOT, J.L. Modulation of language areas with functional MR image-guided magnetic stimulation. *Neuroimage*, 29, 619-627, 2006.

ARTIGES, E., MARTELLI, C., NACCACHE, L., BARTREZ-FAZ, D., LEPROVOST, J.B., VIARD, A., PAILLÈRE-MARTINOT, M.L., DEHAENE, S., MARTINOT, J.L. Paracingulate sulcus morphology and fMRI activation detection in schizophrenia patients. *Schizophrenia Research*, 82, 143-151, 2006.

DEHAENE, S., CHANGEUX, J.P., NACCACHE, L., SACKUR, J., SERGENT, C. Conscious, preconscious, and subliminal processing : a testable taxonomy. *Trends Cogn. Sci.*, 10, 204-211, 2006.

DEHAENE, S., IZARD, V., PICA, P., SPELKE, E. Core knowledge of geometry in an Amazonian indigene group. *Science*, 311, 381-384, 2006.

DEHAENE-LAMBERTZ, G., DEHAENE, S., ANTON, J.L., CAMPAGNE, A., CIUCIU, P., DEHAENE, G.P., DENGHIEN, I., JOBERT, A., LEBIHAN, D., SIGMAN, M., PALLIER, C., POLINE, J.B. Functional segregation of cortical language areas by sentence repetition. *Human Brain Mapping*, 27, 360-371, 2006.

DUBOIS, J., HERTZ-PANIER, L., DEHAENE-LAMBERTZ, G., COINTEPAS, Y., LEBIHAN, D. Assessment of the early organization and maturation of infants' cerebral white matter fiber bundles : A feasibility study using quantitative diffusion tensor imaging and tractography. *NeuroImage*, 30, 1121-1132, 2006.

GAILLARD, R. NACCACHE, L. PINEL, P., CLEMENCEAU, S., VOLLE, E., HASBOUN, D., DUPONT, S., BAULAC, M., DEHAENE, S., ADAM, C., COHEN, L. Direct intracranial, fMRI and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron*, 50, 191-204, 2006.

GAILLARD, R., DEL CUL, A., NACCACHE, L., VINCKIER, F., COHEN, L., DEHAENE, S. Nonconscious semantic processing of emotional words modulates conscious access. *Proc. Natl. Acad. Sci. USA*, 103, 7524-7529, 2006.

GOLESTANI, N., ALARIO, F.A., MERIAUX, S., LE BIHAN, D., DEHAENE, S., PALLIER, C. Syntax production in bilinguals. *Neuropsychologia*, 44, 1029-1040, 2006.

GOLESTANI, N., MOLKO, N., DEHAENE, S., LEBIHAN, D., PALLIER, C. Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex*, sous presse.

GUTKIN, B.S., DEHAENE, S., CHANGEUX, J.P. A neurocomputational hypothesis for nicotine addiction. *Proc. Natl. Acad. Sci. USA*, 103, 1106-1111, 2006.

MERIAUX, S., ROCHE, A., DEHAENE-LAMBERTZ, G., THIRION, B., POLINE, J.B. Combined permutation test and mixed-effect model for group average analysis in fMRI. *Human Brain Mapping*, 27, 402-410, 2006.

NEW, B., FERRAND, L., PALLIER, C., BRYBAERT, M. Reexamining the word length effect in visual word cognition : new evidence from the English lexicon project. *Psychon. Bull. Rev.*, 13, 45-52, 2006.

POLINE, J.B., STROTHER, S.C., DEHAENE-LAMBERTZ, G., EGAN, G.F., LANCASTER, J.L. motivation and synthesis of the FIAC experiment : reproducibility of fMRI result across expert analyses. *Human Brain Mapping*, 27, 351-359, 2006.

SIGMAN, M., JOBERT, A., LEBIHAN, D., DEHAENE, S. Parsing a sequence of brain activations at psychological times using fMRI. *NeuroImage* (sous presse).

SIGMAN, M. DEHAENE, S. Dynamics of the central bottleneck : Dual-task and task uncertainty. *PLOS : Biology* (sous presse).

GLIGA, T. DEHAENE-LAMBERTZ, G. Development of a view-invariant representation of faces in humans. *Cognition* (sous presse).

DEL CUL, A. DEHAENE, S., LEBOYER, M. Preserved subliminal processing and impaired conscious access in Schizophrenia. *Archives of General Psychiatry* (sous presse).

DEHAENE-LAMBERTZ, G., HERTZ-PANIER, L., DUBOIS, J. Nature and nurture in language acquisition : Anatomical and functional brain imaging studies in infants. *Trends in Neurosciences*, (sous presse).

LAUFS, H., HOLT, J.L., ELFOND, R., KRAMS, M., PAUL, J.S., KRAKOW, K., KLEINSCHMIDT, A. Where the BOLD signal goes when alpha EEG leaves. *Neurobiology* (sous presse).

KLEINSCHMIDT, A. Cognitive control signals in visual cortex : Flashes meet spotlights. *Neuron* (sous presse).

KLEINSCHMIDT, A. COHEN, L., The neural bases of prosopagnosia and pure alexia : recent insights from functional neuroimaging. *Current Opinion in Neurobiology* (sous presse).

VINCKIER, F., NACCACHE, L., PAPEIX, C., FORGET, J., HAHN-BARMA, V., DEHAENE, S., COHEN, L. « What » and « Where » in word reading : Ventral coding of written words revealed by parietal atrophy. *Journal of Cognitive Neuroscience* (sous presse).

PIAZZA, M., MECHELLI, A., PRICE, C., BUTTERWORTH, B. Exact and approximate judgements of visual and auditory numerosity ; an fMRI study. *Brain Research* (sous presse).

PIAZZA, M., PINEL, P., DEHAENE, S., Objective correlates of an unusual subjective experience : a single case study of number-form synaesthesia. *Cognitive Neuropsychology* (sous presse).

Chapitres de livres

1. COHEN, L., VINCKIER, F., DEHAENE, S., Anatomical and functional correlates of acquired peripheral dyslexias. In : Cornelissen P.L., Hansen P.C., Kringelbach M.L., and Pugh K., editors. *The Neural basis of reading*. Oxford : oxford University press (sous presse).

2. COHEN, L., WILSON, A.J., IZARD, V., DEHAENE, S. Alcalculia and Gerstmann's syndrome. In : Godefroy O. and Bogousslavsky, J., editors. *Cognitive and Behavioral Neurology of Stroke*. Cambridge : Cambridge University Press (sous presse).

3. DEHAENE, S., COHEN, L. Neural coding of written words in the visual word form area. In : Cornelissen P.L., Hansen P.C., Kringelbach M.L., and Pugh K., editors. *The Neural basis of reading*. Oxford : Oxford University Press (sous presse).

Distinctions

Stanislas Dehaene a été élu membre de l'Académie des Sciences en décembre 2005.