

Images de la pensée

Benoît Virole

1995-2022

Résumé

Ce texte constitue un chapitre de l'ouvrage *Sciences cognitives et psychanalyse* publié en 1995 aux Presses Universitaires de Nancy. Il a été écrit après une participation à un projet de recherche INSERM (U330) sur l'imagerie fonctionnelle dans la schizophrénie. Aujourd'hui, en 2022, beaucoup d'éléments présentés dans ce texte sont dépassés sur le plan technique, mais, à notre sens, la réflexion épistémologique sur le statut des images fonctionnelles en psychopathologie reste valide. De même, la présentation de la problématique de l'implémentation neuronale de la représentation mentale reste d'actualité.

Mots-clefs

Imagerie fonctionnelle Psychanalyse Sciences cognitives Épistémologie

Historique

Depuis une quinzaine d'années, le développement de l'imagerie médicale est venue nourrir les avancées des neurosciences. Ces techniques ont modifié profondément les pratiques exploratoires en substituant parfois à l'examen classique une *vision infraclinique* des troubles potentiels des fonctions supérieures. Sur un plan plus fondamental, l'imagerie permet de voir les activations des aires cérébrales en réponse à des tâches cognitives proposées aux sujets et ouvre le champ des liens entre le substrat neurophysiologique et les fonctions cognitives. Les succès de l'imagerie posent un certain nombre de questions d'ordre épistémologique dont nous voudrions ici essayer d'évoquer la portée générale. La première grande question est d'ordre historique, l'imagerie fonctionnelle s'inscrit-elle dans le *continuum* du regard clinique et n'en constitue-elle qu'un approfondissement ou bien fait-elle rupture sur le plan conceptuel ? Souvent présentée comme ouvrant une ère nouvelle dans la connaissance du substrat biologique de la pensée, l'imagerie fonctionnelle s'inscrit dans l'historicité de la connaissance scientifique. François Dagognet a rappelé que les fondements historiques de l'imagerie fonctionnelle sont d'abord à rechercher

dans l'histoire du regard clinique (Cf. aussi Canguilhem [8] et Foucault [13]).

1. Première en date de ces techniques, *l'électroencéphalographie* (EEG) vient de connaître une seconde jeunesse (1995) grâce à l'informatisation des données qui permet une visualisation multidimensionnelle des potentiels évoqués électriques. Les nouvelles techniques d'EEG permettent ainsi d'avoir une idée dynamique, en temps réel du fonctionnement de groupes de neurones relativement bien définis. La cartographie EEG cérébrale permet une exploration fonctionnelle plus fine et quantifiée de l'activité électrique du cerveau. Les études EEG connaissent aussi un regain d'intérêt par la découverte par Basar et coll. dans le bruit de fond de l'EEG de formes dynamiques semblables à des attracteurs étranges[3].
2. Les inconvénients de l'EEG pour recueillir des potentiels électriques dans certaines zones du cortex sont maintenant palliés par la technique de la *magnéto encéphalographie* qui mesure les champs magnétiques à l'extérieur de la boîte crânienne provenant des courants intracellulaires existant dans les dendrites des neurones activés (neurones pyramidaux). La mesure du champ magnétique évoqué par un événement permet de localiser sa source et d'en déterminer sa force proportionnellement au nombre de neurones concernés.

3. *Les potentiels évoqués et onde P 300* consistent à isoler de l'activité EEG globale celle qui est directement liée à une stimulation sensorielle itérative (auditive, visuelle ou somesthésique). La morphologie et la latence du signal électrique obtenu renseignent sur l'état fonctionnel des voies sensorielles et du cortex associatif. L'étude des potentiels évoqués auditifs tardifs permet une meilleure exploration des fonctions cognitives. En particulier, l'amplitude et la latence de l'onde P 300 sont liées au niveau global de vigilance modifié par l'âge et certaines maladies mentales (démences).
4. *La scintigraphie* permet l'obtention au travers de l'inhalation de marqueurs radioactifs à l'aide de capteurs de radio-activité placés de façon symétrique de part et d'autre du crâne, d'images colorées représentant l'état de la vascularisation des structures intracrâniennes.
5. *La résonance magnétique nucléaire (IRM)* est basée sur l'étude du comportement des protons (noyaux d'hydrogènes) après application d'une impulsion magnétique. Les protons peuvent en effet être considérés comme des aimants possédant un vecteur magnétique (spin). L'IRM consiste à étudier le retour de ces protons à un état d'équilibre après l'application du champ magnétique. Lors de la phase de relaxation aboutissant au retour à l'équilibre, les protons émettent un signal qui sera à la source de l'image. L'IRM présente l'intérêt d'une totale innocuité (absence d'irradiation) et de pouvoir réaliser des coupes dans tous les plans de l'espace.
6. *Le scanner* permet d'obtenir une image radiographique en trois dimensions au travers d'un traitement numérique. Sa résolution est excellente et la TDM (tomodensitométrie) ouvre la voie à une *métabimage* en allant jusqu'à dépasser les seuils oculaires humains. Selon le philosophe des sciences Dagonnet : « L'homme est donc exclu du circuit : il se borne à enregistrer les conclusions qu'il souhaite, puisque l'analyse du rayon, sa réfraction, son traitement ont été confiés à un centre de calcul. » [10, p.136]. L'imagerie permet l'antécédence de l'identification de la maladie avant les signes cliniques, ce qui ne va pas de soi sur le plan d'une épistémologie de la clinique, car selon Canguillem, « la santé, c'est l'état d'inconscience du sujet de son corps » [8].
7. L'étude de la vitesse d'incorporation d'acides aminés nécessaires aux *synthèses protéiques* permet chez l'animal d'apprécier l'activité des diverses structures cérébrales impliquées dans des tâches cognitives.
8. Enfin dernière en date, et la plus spectaculaire, *la tomographie par émission de positons (TEP)* est une technique d'imagerie non invasive permettant d'obtenir une représentation spatiale de la distribution de radio-éléments dans un organe en fonction du temps. On emploie des isotopes qui émettent des électrons chargés positivement, des *positons*, qui se propagent sur des distances de l'ordre de deux millimètres avant de rencontrer un électron négatif et de produire un choc matière/antimatière générant deux photons aux trajectoires diamétralement opposées. Ces deux photons sont détectés par une caméra et les données sont ensuite traitées pour localiser le point d'émission des deux photons. Le résultat est projeté sur un plan à deux dimensions, constituant ainsi une image. Cette méthode permet de quantifier le débit sanguin, le métabolisme du glucose et de l'oxygène (oxygène-15) et éventuellement la pharmacologie de récepteurs et des neuromédiateurs. L'augmentation du métabolisme des neurones activés lors d'une tâche cognitive se traduit, entre autres manifestations, par une augmentation de la consommation de glucose et du débit sanguin régional. Cependant les quantités mesurées en TEP sont des débits sanguins, donc dépendent étroitement de la répartition des apex vasculaires et ne donnent pas directement l'augmentation de métabolisme des corps cellulaires, mais de leurs arbres dendritiques. Lorsqu'une zone apparaît activée en TEP, on ne connaît donc pas précisément la microstructure de son câblage et de ses connexions.

Ces techniques d'imagerie fonctionnelle sont en grande partie complémentaires. Chacune d'entre elles explorent certaines dimensions du cerveau et en laissent d'autres dans l'ombre. C'est pourquoi ces techniques sont employées de façon conjointe tant sur le plan clinique que sur le plan de la recherche. Les meilleurs résultats d'analyse des corrélations entre cognition et imagerie sont obtenus en superposant les images obtenues par TEP sur celles plus précises sur le plan anatomique de l'IRM. L'EEG et les techniques de MEG permettent également d'étudier finement l'organisation temporelle du fonctionnement cérébral alors que l'IRM et la TEP permettent une localisation spatiale des variations métaboliques corrélées à l'activité cérébrale. L'imagerie vise à mettre en évidence les liens entre les aires fonctionnelles et les opérations cognitives. Les données les plus propres à fournir un enseignement sur les fonctions cognitives sont issues

de l'imagerie par émission de positons (TEP). Elles peuvent être classées en suivant une hiérarchie cognitive, des processus dits de « bas niveau », sensoriels et perceptif, jusqu'aux processus de « haut niveau », tels le langage et le raisonnement.

1. Dans l'analyse des tâches cognitives visuelles, l'augmentation de la cadence de flash d'un spot provoque une activation des aires cérébrales visuelles d'autant plus importante que le clignotement est plus rapide. Plus la sensation est forte ou complexe, plus le niveau d'activité est important (travaux de l'équipe de Saint-Louis).
2. Devant des tâches testant la perception de la couleur, on peut mettre en évidence une augmentation localisée du débit sanguin dans la zone interne du lobe occipital correspondant donc au centre de la couleur chez l'homme (équipe du Hammersmith, Londres.)
3. On a mis en évidence des réseaux neuronaux différents impliqués dans les tâches visuelles complexes (localisation spatiale, reconnaissance des visages...) (groupe du National Institute, Bethesda)
4. La TEP a permis de mieux connaître la spécialisation interhémisphérique lors de tâches cognitives telle la reconnaissance des visages et des objets usuels et d'étudier les aires fonctionnelles impliquées dans l'attention (équipe de Montréal). Les trois types d'attention (dirigée, sélective, partagée) correspondent à des localisations différentes de l'activation (équipe de Saint Louis).
5. L'étude des relations entre apprentissage d'un mouvement et modifications de l'activité cérébrale a montré que plus le mouvement est connu par le sujet (automatisme), moins le nombre de régions activées est important (équipe du Karolinska à Stockholm).
6. D'autres chercheurs ont étudié les fonctions localisées sur le lobe frontal dans leurs relations avec les processus décisionnels (mouvements volontaires), indépendants des stimuli externes.
7. La TEP a permis l'étude des liens entre les types de performance linguistique et les aires fonctionnelles aboutissant à l'identification d'une zone (aire 47 de Broadman) impliquée dans le codage sémantique (équipe de Saint-Louis).
8. On a pu mettre en évidence l'activation d'une zone du cortex frontal dans la sensation de douleur (le cingulum) (équipe du Hammersmith, Londres.).
9. La TEP a été utilisée dans l'analyse des effets du stress sur l'activité cérébrale et a montré une

augmentation sur les deux pôles temporaux (cerveau limbique). Ce résultat, discuté, a été interprété comme la conjonction des éléments sensoriels avec le déclenchement de réponses comportementales à des situations dangereuses (équipe de Saint-Louis, USA) [25].

10. Enfin, la TEP a été utilisée dans des études de psychiatrie biologique, en particulier au travers de l'étude des récepteurs dopaminergiques D2 dans les études sur la schizophrénie [20]. Les études sur les schizophrènes montrent une différence entre la distribution des débits sanguins entre les sujets normaux et les schizophrènes chroniques. En particulier, l'activation frontale est moins élevée chez les schizophrènes (Cf. revue dans J.L Martinot (1991) [20]).

Liens avec la neuropsychologie

Ces résultats de l'imagerie constituent-ils des faits nouveaux, originaux, appelant de nouvelles interprétations ou ne sont-ils que des données de confirmation des modèles cognitifs construits à partir d'autres méthodes de recueil de faits ? Ces résultats de l'imagerie fonctionnelle, parfois identifiée à une véritable *idéographie*[15], permettent d'établir des passerelles avec la neuropsychologie. Passerelles d'autant plus intéressantes que l'imagerie a détroné la neuropsychologie comme technique de diagnostic des troubles des fonctions supérieures, la libérant ainsi des questions anatomo-fonctionnelles en l'orientant vers les modèles cognitifs. Les relations entre les données des imageries par TEP et les données neuropsychologiques saluent ainsi les retrouvailles de ces deux voies d'approches.

En ce qui concerne les opérations perceptives et sensorimotrices, l'accord est quasi-total avec les données de la neuropsychologie. Les aires fonctionnelles activées correspondent bien aux zones dévolues cliniquement à ces fonctions et aux modules de traitement périphériques. Cependant, ces retrouvailles peuvent être ponctuellement l'objet de remises en cause. Ainsi, l'imagerie a permis de se rendre compte que certains modèles cognitifs n'étaient pas en accord avec les activations fonctionnelles sur le cortex.

Par exemple, pour la modalité visuelle, les principales aires d'activation au TEP sont le cortex strié et certaines aires jusqu'à la limite temporo-occipitale,

qui sont dévolues au codage des mots sous forme visuelle. Or, aucune activation des aires classiquement associées au codage phonologique n'est observée (Avan, 1992) [2]. Pour le traitement auditif des mots, le cortex auditif primaire (bilatéral) et certaines zones du cortex temporo pariétal, temporal antérosupérieur gauche (pour des sujets droitiers) sont activés alors que des logatomes n'activent pas les deux dernières régions. Les tâches d'associations activent une zone frontale inférieure gauche (à laquelle est attribuée un rôle dans les associations sémantiques) ainsi qu'une deuxième région (gyrus antérieur cingulaire.) Ainsi ces données sont en contradiction avec les modèles neuropsychologiques qui postulent qu'un mot lu a besoin de subir un codage phonologique pour accéder à une représentation sémantique.

D'autres auteurs ont montré qu'une évocation purement silencieuse de mots peut activer des aires motrices fournissant ainsi à nouveau des arguments à l'ancienne théorie motrice de la perception de la parole selon laquelle c'est l'écho proprioceptif des mots qui permet la catégorisation phonétique et finalement l'identification. Les modèles de la catégorisation phonétique utilisant exclusivement la notion de détecteurs de traits sans apports des informations motrices se verraient alors invalidés. Ainsi l'imagerie permet de remettre en chantier les débats entre les différents modèles de la catégorisation phonétique, en fournissant une nouvelle dimension argumentaire qui pourrait s'avérer décisive (Cf. la revue des théories de la perception catégorielle de la parole dans Petitot (1984) [23]).

Des phénomènes perceptifs aux phénomènes cognitifs de haut niveau comme le langage, l'imagerie fonctionnelle fournit donc des données de tout premier plan. Mais nous sommes toujours dans un espace cognitif référé aux paramètres perceptifs. En est-il de même si l'on descend à un niveau plus profond dans la hiérarchie des opérations cognitives et que l'on s'intéresse à la dimension sémantique par exemple? Sur le plan anatomo-fonctionnel, il est d'usage de situer les traitements corticaux des opérations sémantiques de façon beaucoup plus diffuse sur le cortex frontal et préfrontal. Or l'imagerie fonctionnelle a permis de localiser précisément une aire fonctionnelle correspondant au traitement

sémantique. Par exemple, Goldenberg et coll. ont montré que les opérations cognitives de traitement de phrases avec un contenu sémantique fortement associé à des percepts visuels s'accompagne d'une activité plus importante de la région dévolue au cortex visuel (région occipitale inférieure gauche) que celles de phrases au contenu purement conceptuel [11]. La plupart de ces résultats s'harmonisent de façon convaincante avec les conceptions de la modularité de l'esprit; la cognition est un ensemble de fonctions distinctes agencées de façon modulaire et séparant les fonctions périphériques des modules centraux.

On serait alors en droit d'attendre de l'imagerie fonctionnelle qu'elle localise aussi les processus cognitifs de très haut niveau, telles l'idéation mentale ou les opérations logico-mathématiques. La mise en parallèle des techniques d'imagerie et le paradigme fonctionnaliste des sciences cognitives montre cependant un paradoxe. L'imagerie a démontré la pertinence de la parcellisation des fonctions cognitives sur les aires fonctionnelles du cortex et s'accorde volontiers avec la notion d'une modularisation de l'esprit. Les deux approches devraient être ainsi étroitement complémentaires. Or le fonctionnalisme à la « Fodor » [12] suppose par définition l'indépendance des niveaux de description entre les fonctions cognitives et leur substrat matériel (qu'il soit biologique ou informatique).

Cependant, les deux voies se rapprochent sur le fait que leurs démarches respectives les entraînent vers une complication de plus en plus poussée de l'ensemble des aires fonctionnelles ou des modules de traitement. En outre, ces avancées font bon marché d'une aporie qui ne gêne pas l'application technique, réductionniste de l'imagerie mais qui est plus gênante sur le plan d'une compréhension plus holistique. Si l'imagerie fournit une vérification de l'existence d'aires fonctionnelles, et que la pensée peut ainsi être décomposée en des modules indépendants, au moins sur le plan anatomo-fonctionnel, comment se réalise l'unité subjective de la pensée? Car nous n'avons pas intuitivement conscience de la modularité supposée de notre pensée, et en cas de pathologie locale d'une fonction, le trouble peut être dénié. Cette distinction entre modularité fonctionnelle et unité subjective de la pensée touche les rivages de la philosophie par l'in-

roduction de la notion de sujet, mais elle est aussi directement impliquée dans les applications de l'imagerie en psychopathologie.

Images de la folie

L'imagerie fonctionnelle par TEP a été utilisée en psychopathologie en tentant de mettre en évidence des différences dans les niveaux d'activation corticaux entre des sujets exempts de troubles mentaux et des sujets malades mentaux. On conçoit l'importance de ce champ d'application de l'imagerie, en particulier par les relations qu'elle pourrait instaurer avec les données psychopharmacologiques. Il serait alors possible d'objectiver le dérèglement psychopathologique, non plus simplement dans ses conséquences cliniques, mais à la source du trouble. Arrêtons-nous cependant un instant sur le plan méthodologique.

Si un groupe de sujets présentant des troubles psychopathologiques apparentés sémiologiquement fournit de façon statistique une imagerie semblable à la même tâche cognitive et que, devant cette tâche un groupe de sujets ne se distinguant du premier groupe que par l'absence du trouble psychopathologique marquant le premier groupe présente une imagerie statistiquement différente, alors il existe une corrélation positive entre le facteur présence ou absence de trouble et le facteur localisation de l'activité cérébrale pour cette tâche.

Examinons d'abord les conditions de fiabilité des différents facteurs. Le repérage du trouble psychopathologique est apparemment le plus aisé dans la mesure où les cliniciens utilisent des critères standardisés permettant un accord de plusieurs cliniciens sur la présence ou non d'un trouble schizophrénique. Cependant l'exclusion du trouble schizophrénique chez les sujets du groupe témoin ne signifie pas que ces sujets ne présentent pas des troubles d'une autre nature (pensée obsessionnelle, rigidité caractérielle etc...). Le choix de la tâche cognitive constitue également un facteur d'imprécision. Prenons par exemple un test de fluence verbale. Ce test est dépendant certes des facteurs neuropsychologiques de fluence mais aussi des connaissances lexicales qui sont dépendantes des niveaux socio-culturels, et de l'entraînement des sujets à utiliser leur mémoire verbale au quotidien. Il

est bien sûr possible de chercher à harmoniser les deux groupes, mais comme les conditions pragmatiques d'utilisation du langage sont influencées par la schizophrénie, les variables présence ou absence de schizophrénie et celles de la performance psycholinguistique sont liées et covariant. Les schizophrènes réalisent donc moins bien la tâche cognitive, ce qui se traduit par une imagerie différente qui ne renseigne donc que sur la différence de réalisation des tâches cognitives due à un facteur x (qui peut être autant les conditions sociales de vie des schizophrènes, de leur chimiothérapie, etc..) entre les deux groupes et non sur la schizophrénie. On voit que si l'imagerie montre des distinctions entre les deux groupes de sujets, il n'est pas dit que cette distinction puisse être de façon univoque attribuée à la schizophrénie.

En termes logiques, on peut formaliser le problème de la relation entre imagerie et schizophrénie en supposant qu'il existe une implication directe entre la particularité de l'activation corticale et la maladie. Soit p une proposition (par exemple : absence ou présence d'une activation frontale objectivée par TEP à une tâche cognitive x et q une autre proposition (existence d'une schizophrénie), la relation peut être représentée ainsi :

$$p \supset q \quad (1)$$

et se lit p implique q . Or cette forme de l'implication causale peut en droit être remplacée par une relation symétrique :

$$q \supset p \quad (2)$$

où c'est la schizophrénie qui implique le dysfonctionnement frontal. Le choix entre ces deux options est très difficilement réalisable et on est ainsi confronté à la question classique de la cause première. La notion d'antécédence ne peut jouer, puisque les deux objets en relation sont cosubstantiels. Seule la mise en évidence de lésions structurelles du cortex frontal permettrait peut-être d'affirmer que la première proposition est vraie et que la seconde est fausse. Mais de toutes les façons, cette première proposition devra être complétée et remplacée par :

$$p_1 \wedge p_2 \wedge p_n \dots \supset q \quad (3)$$

où p_n représente la série des cofacteurs impliqués (existence de traitements, conditions de vie,...). Or ces cofacteurs ne sont pas tous neutralisés dans les études, et certains sont soit inconnus, soit non pris en compte, ce qui signifie que la proposition (1) peut être ramenée à :

$$p_1 \wedge c \supset q \quad (4)$$

avec c spécifiant l'ensemble des conditions annexes, y compris le hasard par description incomplète du phénomène. Ainsi les données de l'imagerie montrent bien qu'il existe une différence entre sujets schizophrènes et sujets sains. Mais il n'est pas possible d'affirmer sur le plan logique que cette différence est due à la schizophrénie ou à d'autres facteurs associés soit structurellement (traitements, conditions de vie...) soit du fait des contingences et du hasard. Cependant malgré ces difficultés méthodologiques, l'application de l'imagerie fonctionnelle en psychiatrie biologique semble prometteuse, bien qu'elle soit pour l'instant limitée à une appréciation localisatrice et/ou quantitative des différences des niveaux d'activation corticaux. Elle peut mettre en évidence des différences entre sujets pathologiques et sujets sains, mais il n'est pas certain qu'elle puisse seule aller plus loin et dégager la nature qualitative de ces différences.

Statut des images fonctionnelles

Cet exemple en psychopathologie montre bien que l'imagerie pose surtout la question centrale du statut épistémologique de la notion des corrélats et celle secondaire de la tendance à réduire la corrélation à la cause. Or cette réduction, inscrite dans la tendance naturelle à assigner des causes à la coexistence s'avère fautive. Il peut s'agir d'une dégradation dans l'illusion technologique de la notion de cause. La cause en médecine théorique, c'est ce sur quoi on peut *par principe* agir pour essayer de contrôler un phénomène, pour le provoquer si on en a besoin, pour l'éviter s'il est néfaste[6]. Mais la question centrale est

bien celle du statut des corrélats. Car ce que l'imagerie par TEP montre ce sont l'existence de corrélats visuels (zones d'activation) aux activités cognitives. Or cette notion de corrélats n'est pas univoque sur le plan logique. L'existence d'une covariation entre deux variables ne signifie pas que leur corrélation soit causale. La nuit n'est pas la cause du jour, et pourtant les deux phénomènes sont étroitement corrélés. En d'autres termes, le déplacement d'une marque visuelle sur le cortex en imagerie en miroir d'une modification des tâches cognitives présentées en entrée ne signifie pas forcément que les deux variables soient liées par une relation d'identité causale.

Ces points méthodologiques assez triviaux peuvent paraître superflus, mais ils sont peut-être plus nécessaires en psychopathologie qu'ailleurs dans la mesure où il convient de reconnaître la tendance forte à la recherche d'une causalité biologique unique (*le gène* ou *le déficit* en neuromédiateur). L'imagerie dans cette perspective, par sa puissance démonstrative, utilise les recours auto-didactiques des *icônes* mis en évidence par Peirce :

« ...une des grandes propriétés distinctives de l'icône est que par son observation directe peuvent être découvertes concernant son objet d'autres vérités que celles qui suffisent à déterminer sa construction. »[22, p.150]

Ce pouvoir de l'image pourrait bien favoriser ce mouvement à la fois simplificateur et réductionniste, s'apparentant par certains côtés à la pensée magique : *le faire voir génère le faire croire*. En d'autres termes, montrer une différence entre le cerveau d'un aliéné et celui d'un sujet normal fait irrésistiblement penser que cette différence est la cause de l'aliénation. Cela est peut-être acceptable pour des macro-structures, mais il n'est pas sûr que la visualisation des zones d'activation corticales puisse cependant prétendre construire le socle d'une rationalité neuropsychologique d'où la différence entre pensée normale ou pathologique pourrait être référée. Car si l'on descend plus en profondeur et que l'on s'avance vers les composantes des fonctions cognitives et donc vers le niveau des représentations neuronales des objets cognitifs, ce socle de rationalité s'avère des plus instables. La question de la représentation neuronale peut être

abordée par l'exposé de deux grandes conceptions antagonistes.

L'implémentation neuronale de la représentation

Selon la première inaugurée par la théorie localisationniste de Horace Barlow, quels que soient les niveaux d'abstraction et de complexité de l'objet, sa représentation neuronale est identique à l'activation d'un petit nombre de cellules corticales :

« Le système sensoriel est organisé de façon qu'une représentation du stimulus (...) soit réalisée avec un nombre minimal de neurones en activité. La perception correspond à l'activité d'un nombre réduit de neurones sélectionnés, chacun d'eux correspondant à un groupe d'événements extérieurs du même ordre de complexité qu'un ensemble d'événements symbolisés par un mot. » [4, p.371]

Dans ce modèle, la hiérarchie informationnelle est très forte entre les étages inférieurs (perception) vers les étages supérieurs contenant les cellules détectrices des formes élémentaires (détection de contours, de spectre...). Durant ce trajet *bottom-up*, le message subit une suite d'opérations de filtrage parvenant à extraire des caractéristiques de plus en plus abstraites (traits structuraux) par le biais de cellules de plus en plus sélectives. À un même niveau, (centres intégrateurs des voies ascendantes), les interactions entre les neurones sont de type d'inhibition latérale et s'apparentent aux réseaux neuromimétiques de type compétition comprenant une réduction progressive du nombre de neurones codant au fur et à mesure que l'on monte vers les centres corticaux. Malgré l'importance des arguments physiologiques en faveur de ce modèle, celui-ci souffre d'une grave aporie qui a été maintes fois soulignée. S'il existe autant de cellules cardinales que d'objets du monde à représenter sur le plan neuronal, alors il est nécessaire de poser que le cerveau contient autant de neurones cardinaux que d'objets du monde, ce qui correspond à un non sens logique, puisque les objets du monde sont dans infinitude virtuelle alors que le nombre de neurones est fixé.

Contre l'impasse des théories localisationnistes, Hebb a proposé en 1949 la notion d'« assemblées de cellules » dont l'activation globale, et non indivi-

duelle, constitue la brique élémentaire de la perception. Le concept d'*objet mental* proposé par Changeux reprend cette notion qui regroupe percept, image mentale et concept :

« L'objet mental est identique à l'état physique créé par l'entrée en activité (électrique et chimique), corrélée et transitoire, d'une large population ou assemblées de neurones distribuées au niveau de plusieurs aires corticales définies ». [9]

À l'opposé des théories strictement localisationnistes, la théorie de Karl Lashley [18] a proposé une organisation distribuée du système nerveux, où les engrammes neuronaux de la représentation sont délocalisés. Cette délocalisation a été en partie validée par le retour des modèles connexionnistes qui ont montré que dans un réseau neuronal, la représentation était délocalisée dans l'ensemble des poids synaptiques. À l'opposé du paradigme fonctionnaliste, le connexionnisme stipule que les opérations cognitives sont corrélées à l'existence de réseaux nerveux de grande dimension interconnectant des unités modulaires de structure et de fonction identique. Les unités fonctionnelles du cerveau peuvent être modélisées comme des assemblées de neurones transmettant une information.

Cependant toutes ces différentes théories s'apparentent toujours à un *atomisme logique* dont la règle centrale est que le *tout est identique à la somme des parties*. Ces théories prévoient une représentation particulière (neurones ou réseaux) pour chaque objet et pour chaque présentation de l'objet. Or cette conception entraîne une véritable explosion combinatoire mise en évidence par les modélisations en Intelligence Artificielle. La reconnaissance de formes assez simples nécessite des réseaux de dimension très grande rendant peu crédible leur réalité biologique. Pour un seul percept, le nombre de présentations physiques est très grand et nécessite un nombre virtuellement infini de neurones et de connexions.

Les modèles connexionnistes rencontrent également de nombreuses difficultés pour rendre compte des opérations cognitives plus complexes. Prenons le cas du traitement des règles de syntaxe par exemple,

1. Deux expressions syntaxiques différentes par un renversement propositionnel ne peuvent que générer la

même représentation neuronale du fait que l'on ne peut concevoir les réseaux neuronaux biologiques que comme des graphes non orientés. Sinon il faudrait postuler l'existence d'autant de réseaux différents que de flexions portées sur une même expression syntaxique.

2. Donc pour coder l'ensemble des renversements syntaxiques, il devient nécessaire de postuler l'existence de nouveaux réseaux. Cependant, l'introduction de nouveaux réseaux est peu réaliste voire impossible sur le plan physiologique puisqu'un réseau de neurones biologiques doit être entretenu par l'activité pour continuer à exister. Or, si un réseau codant pour une forme linguistique peu usitée est peu activé, alors il dégénère (du fait de la nécessité de l'entretien métabolique des neurones). Pourtant la productivité du langage montre que des formes linguistiques peu fréquentes peuvent être à tout moment performées par le locuteur.

Ainsi, le connexionnisme est soumis à un paradoxe de nature volumétrique : il existe à la fois trop de neurones aboutissant à des solutions multiples, et pas assez pour rendre compte de la diversité de la production linguistique. Une des solutions potentielles consiste à poser que tout objet cognitif complexe doit être appréhendé directement comme un ensemble de relations et non comme un juxtaposition de sous-objets. Il faut donc définir en plus des variables classiques définissant les activités des neurones individuels codant pour les propriétés des objets élémentaires des variables codant les combinaisons, associations entre les sous-objets. La théorie des corrélations de Malsburg postule que ce sont les corrélations temporelles rapides entre les activités de neurones qui codent les relations structurelles entre les objets. Ces corrélations sont de l'ordre de la milliseconde, ce qui constitue une échelle de temps très courte par rapport à la durée de vie des objets mentaux qui est de l'ordre de la seconde. Cette introduction des corrélations temporelles permet l'introduction d'une nouvelle variable, le *poids synaptique à variation rapide* qui est distinct des dynamiques synaptiques lentes correspondant au développement ontogénétique de la maturation des apprentissages. Le poids synaptique à variation rapide vient alors moduler la cinétique lente aboutissant à la constitution d'un système dynamique.

Mentionnons rapidement qu'un système dynamique est un ensemble de facteurs interagissants et dont les valeurs changent continuellement. Dans un tel système, les variables changent de façon continue alors que les paramètres du système restent constants. L'espace des états est l'ensemble de tous les états pris par le système. Un cas particulier des systèmes est celui des systèmes dissipatifs qui perdent ou gagnent de l'énergie. Les points représentant les différents états sur cet espace peuvent être reliés entre eux dessinant ainsi une trajectoire. L'espace des états rempli avec des trajectoires d'un modèle donné est appelé l'espace des phases du système. Les trajectoires tendent alors vers un ensemble limite que l'on nomme attracteur du système. La théorie des systèmes dynamiques isolés affirme qu'il n'existe et ne peut exister que quatre types d'attracteurs :

1. Le système s'installe dans un état stationnaire (attracteur statique)
2. Le système s'installe dans un régime oscillatoire périodique (attracteur cyclique).
3. Le système s'installe dans un état complexe composé de deux fréquences différentes et simultanées.
4. Le système s'installe dans un régime chaotique dit « attracteur étrange ».

D'après Malsburg, les réseaux neuronaux peuvent être décrits comme des systèmes dynamiques en prenant en compte les intercorrélations des neurones. Ce système peut être décrit comme suit :

1. Les niveaux d'activation élevés entre neurones ont tendance à renforcer la corrélation entre les deux neurones α et β (ce qui la reprise de règle de Hebb.)
2. La cinétique du niveau d'activation d'une synapse entre α et β est contrôlée par la corrélation entre les activités de α et de β . Le poids synaptique croît si la corrélation est élevée (feed-back positif). Selon Hopfield, la variable rapide codant les relations interneuronales est identifiable à un spin en électromagnétisme.
3. Les corrélations se propagent (si α est corrélé à γ et γ à β alors α et β sont corrélés).
4. Les chemins synaptiques indirects entre α et β (du type $\alpha - \gamma - \beta$ ou $\alpha - \gamma - \delta - \beta$) interviennent également.
5. Il apparaît un phénomène de coopération au niveau de neurones entre lesquels existent des chemins courts (peu de synapses). Les corrélations se propagent préférentiellement le long des chemins courts.

6. Feed-back et coopération définissent une organisation qui doit être complétée par une autre règle dite de « compétition » interdisant aux corrélations de grandir tous ensemble. Biologiquement cette contrainte est assumée par les réseaux d'interneurons inhibiteurs.

Le point central de ces théories assimilant le fonctionnement des réseaux neuronaux à des systèmes dynamiques est que les processus d'auto-organisation donnent naissance à des *topologies*. Par exemple, dans cette approche dynamique, l'objet de la perception n'est plus codé par une succession de neurones sélectifs vers une cellule cardinale, mais par une configuration de corrélations entre cinétiques rapides. Ainsi, la dynamique d'activation du système possède une propriété auto-organisatrice qui favorise de façon sélective les graphes organisés topologiquement et par voie de conséquence les graphes réalisant des homéomorphismes entre topologies déjà présentes dans le système.

Selon Bienenstock, il existe dans le système nerveux central un immense répertoire de topologies représentant toutes des objets différents. Ces graphes sont stockés sous la forme de poids synaptiques lents et mobilisent tous le même ensemble de neurones sélectifs élémentaires. Ces graphes sont superposés les uns sur les autres et se trouvent en compétition mutuelle. La dynamique d'activation, agissant sur une telle superposition de topologies se traduit par une coopération entre éléments d'une même topologie et par la compétition entre topologies différentes. Lors de la perception, l'information provenant des organes des sens active sélectivement le graphe représentant l'objet. Un homéomorphisme peut ainsi s'établir entre le graphe activé et le message sensoriel au niveau de la rétine et du cortex visuel primaire. Selon Bienenstock, la cognition et l'apprentissage perceptif s'articulent autour de la notion d'homéomorphisme : activation d'homéomorphismes entre objets extérieurs et objets mentaux lors de la perception, mais aussi impression dans le réseau synaptique lent, d'homéomorphismes nouvellement activés entre topologies anciennes lors de l'apprentissage perceptif. La construction de la représentation du monde est ainsi selon Bienenstock le résultat d'un échange permanent *entre une dynamique auto-organisatrice et une organisation extérieure*[5].

Cette vision de la cognition, très proche des thèses de Varela [28] sur la clôture informationnelle du SNC est encore trop schématique selon Bienenstock, il a proposé de l'affiner en tenant compte du contexte sémantique de l'objet que l'on peut décrire comme les autres graphes qui lui sont partiellement homéomorphes et de l'aspect séquentiel des opérations de perception : l'objet mental est un enchaînement dynamique de topologies plutôt que l'activation d'une topologie unique. Les zones d'activation mise en évidence par des épreuves cognitives pourraient correspondre à *un sommet de graphe topologique et non au lieu du phénomène cognitif*.

Conclusion

L'imagerie fonctionnelle nous fournit une preuve de la matérialité de la pensée, en tant qu'elle identifie la matière à la spatialité. Mais elle ne permet cependant pas de franchir le pas qui sépare le repérage d'un lieu où est situé un objet de la connaissance de la structure de cet objet. C'est donc une matérialité de surface, un reflet, certes instructif, mais qui possède des limitations internes. En ce sens l'imagerie ne fait pas rupture dans un processus de connaissance déjà éprouvé. L'écoute du battement cardiaque est devenue une exploration fonctionnelle efficace à partir du moment où l'on a disposé d'un bon modèle physiologique du fonctionnement du cœur. En ce qui concerne la pensée et son implémentation neuronale, les sciences cognitives peuvent incontestablement construire en partie ce modèle extérieur permettant d'organiser et d'interpréter les données de l'imagerie fonctionnelle. Il n'est pas sûr qu'elles puissent aisément fournir un modèle permettant de comprendre comment les représentations neuronales sont encodées et quelles sont leurs fonctions. L'avancée des techniques de l'imagerie fonctionnelle fait reculer les limites de notre connaissance du substrat neuronal de la pensée. Dans le même temps les sciences cognitives ouvrent la vision vertigineuse d'une pensée délocalisée. En ce sens, la découverte récente des dimensions fractales, qui se sont vues d'ailleurs appliquées rapidement sur la morphologie neuronale [27], ouvre une perspective nouvelle, un *nouvel imaginaire scientifique* pour la cognition. En quelque sorte, au même moment où l'imagerie fonctionnelle vien-

drait inscrire la cognition dans l'espace-temps usuel accessible à notre compréhension intuitive, d'autres avancées viennent dissoudre ce socle temporaire de rationalité. L'avenir permettra-t-il de posséder de nouveaux concepts, de nouveaux instruments pour aller plus loin dans l'objectivation de la pensée? Il est difficile d'en préjuger, mais il est possible d'imaginer que l'on puisse arriver en neurosciences à la situation des sciences physiques, à savoir que l'instrument d'observation modifie l'objet observé aboutissant à une relation d'indétermination. Ce n'est certes pas le cas avec la tomographie par émission de positons, mais rien ne dit que pour visualiser et essayer de connaître les topologies dynamiques des systèmes neuronaux, il ne faille un jour les perturber de telle façon que l'on puisse observer leur réaction. La neuropsychologie serait alors dans cette même curieuse situation épistémique qu'a vécu la physique, celle de l'inter-dépendance entre l'objet et la méthode.

Références

- [1] Abraham F. D., *A Visual Introduction to Dynamical Systems Theory for Psychology* Santa Cruz, Aerial Press, Inc, P.O Box, 1990.
- [2] Avan P., Loth D., Menguy C., « Imagerie fonctionnelle cérébrale : Pet, Meg, deux techniques d'avenir en audition? » in *Les cahiers de l'audition* Vol 5., 6, septembre 1992, pp. 26-28.
- [3] Basar E., *Chaos in Brain Function* New-York, Springer-Verlag, 1990.
- [4] Barlow H.B., « Single Units and Sensation : a Neuron Doctrin for Perceptual Psychology » *Perception* vol. 1, 1972, pp. 371-394.
- [5] Bienenstock E., « Une approche topologique de l'objet mental » in *Les théories de la complexité*, Paris, Seuil, 1991.
- [6] Bonsack F., « Aspects épistémologiques des notions de corrélation et de causalité in *Elaboration et justification des modèles* présentés par P. Delattre et M. Thellier, Tome 2, Paris, Maloine, 1979.
- [7] Bouton C. P., *Discours physique du langage, Génèse et histoire de la neurolinguistique* Paris, Klincksiek, 1984.
- [8] Canguilhem G., *Le normal et le pathologique* Paris, Puf, 1966.
- [9] Changeux J.P., *L'homme neuronal* Paris, Fayard, 1983.
- [10] Dagonnet F., *Philosophie de l'image* Paris, Vrin, 1986.
- [11] Goldenberg G., « Patterns of Regional Cerebral Blood Flow Related to Memorizing of High and Low Imaregy Words - An Emission Computer Tomography Study », *Neuropsychology*, 25, 1987, pp. 473-485.
- [12] Fodor J.A., *La modularité de l'esprit*, 1983, Paris, Les éditions de Minuit, 1986.
- [13] Foucault M., *Naissance de la clinique* Paris, Puf, 1963.
- [14] Freud S., *Contribution à la conception des aphasies*, 1891, Bibliothèque de Psychanalyse, Paris, Puf, 1983.
- [15] Ingvar D. « L'idéogramme cérébral » *Encéphale*, 3, 1977. pp. 5-33.
- [16] Lacan J., *Les quatres concepts fondamentaux de la psychanalyse*, Le séminaire, livre XI, Paris, Seuil, 1973.
- [17] Lantéri-Laura G., *Histoire de la phrénologie, l'homme et son cerveau selon J.G. Gall*. Paris, Puf, 1970.
- [18] Lashley K. S., « In Search of the Engram » in *Physiological Mechanisms in Animal Behavior* Society of Experimental Biology Symposium, 4, Cambridge, pp. 454-482.
- [19] Leybaert J., Alegria J., « Difficultés de lecture des sourds : une approche psycholinguistique », in *vivre sourd aujourd'hui et demain*, Bruxelles, Edirsa, 1986.
- [20] Martinot J.L., « Physiologie cérébrale et psychopathologie, modèles de recherche en imagerie fonctionnelle », *L'information psychiatrique*, 7, septembre 1991, pp. 669-679.
- [21] Mazoyer B., « La cartographie des fonctions cognitives » *Sciences cognitives, Le Courrier du CNRS*, 79, octobre 1992.
- [22] Peirce C., *Écrits sur le signe* rassemblés, traduits et commentés par Gérard Deledalle, Paris, Seuil, 1978.
- [23] Petitot J., *Les catastrophes de la parole*, Paris, Maloine, 1984.
- [24] Raichle M.E., « Developing a Functional Anatomy of the Human Visual System with Positron Emission Tomography », in : *Vision and the Brain*, New-York, Cohen B. Ed, Raven Press, 1990, pp. 257-270.
- [25] Reiman E.M, Fusselman M.J., Fox P.T, Raichle M.E., « Neuroanatomical Correlates of Anticipatory Anxiety », *Science*, 243, 1989, pp. 1071-1074.

- [26] Reuchlin M., *Cognition : l'individuel et l'universel*, ouvrage collectif sous la direction de M. Reuchlin, Paris, Puf, 1990.
- [27] Reichenbach A., « A Comparative Fractal Analysis of Various Mammalian Astroglial Cell Types », *Neuroimage*, 1, 1992, pp. 69-77, 1992.
- [28] Varela F., *Autonomie et Connaissance : essai sur le vivant*, Paris, Seuil, 1989.

Pour citer cet article : *Sciences cognitives et psychanalyse*, Presses Universitaires de Nancy, 1995, pp. 93-106. & <https://virole.pagesperso-orange.fr/image.pdf> (2021).