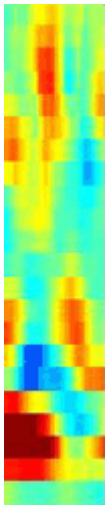
Francisco Varela's Home Page

"Le cerveau n'est pas un ordinateur

On ne peut comprendre la cognition si l'on s'abstrait de son incarnation"



La Recherche, No.308
Avril 1998, p.109-112
Entretien avec Francisco Varela par Herve Kempf

Francisco J. Varela, né en 1946, docteur en biologie de Harvard en 1970, est neurobiologiste. Après avoir longtemps travaillé aux Etats-Unis, il est entré au CNRS en 1988 ou il est actuellement Directeur de recherche. Il dirige l'équipe de Neurodynamique au Laboratoire des Neurosciences Cognitives et Imagerie Cérébrale (CNRS UPR 640), à Paris. Il a notamment publié L'inscription corporelle de l'esprit *(Seuil, 1993) et* L'arbre de la connaissance *(Addison-Wesley, 1994).*
Pour plus d'information consulter: http://www.ccr.jussieu.fr/varela/welcome.html

---

La Recherche - Steven Pinker vient de déclarer que "presque tout le monde dans les sciences cognitives" partage l'idée que le cerveau est une sorte d'ordinateur neuronal produit par l'évolution (*Wired, mars 1998)***. Ceux qui ne le pensent pas, dit-il "sont flamboyants, mais peu nombreux et peu représentatifs". Vous sentez-vous flamboyant et marginal ?**

Francisco Varela : La notion d'un ordinateur "neuronal" n'est pas évidente, parce qu'un ordinateur, stricto sensu, c'est un système symbolique. Tout le courant qu'on appelle aujourd'hui le cognitivisme, ou computationnalisme, travaille avec l'hypothèse que le niveau cognitif est autonome. Leur hypothèse est qu'il y a une différence entre les niveaux d'implémentation, entre le soft et le hard. Donc parler de l'ordinateur "neuronal", c'est déjà plonger dans la perplexité la moitié des chercheurs du domaine, pour lesquels ce que fait le cerveau n'est pas directement pertinent pour comprendre la cognition. Si vous parlez à un informaticien, il se moque de savoir sur quel ordinateur il travaille, ce qui l'intéresse, c'est la structure de son programme. De la même façon, ceux qui s'intéressent à la computation en tant que telle ne se préoccupent pas de son support. Ils ne dénient pas l'intérêt de la biologie,

ils disent que ce n'est pas leur domaine. Par ailleurs, il y a énormément de manières de tisser le passage entre l'hypothèse computationnaliste et le cerveau. La question revient à analyser ce qu'est un calcul fondé sur une machine non symbolique, comparable à un cerveau plutôt qu'à un ordinateur. C'est une question qui m'intéresse, comme elle intéresse beaucoup plus de gens que ne le croit Pinker. Et ce que nous disons, c'est que l'avenir des sciences cognitives repose sur une vision que l'on pourrait appeler "dynamique", par opposition à la vision computationnelle.

De quoi s'agit-il ?

Que plutôt que de travailler avec des symboles et des règles, il faut travailler avec des systèmes formés de variables réelles en utilisant des équations différentielles. Cette approche génère actuellement énormément de résultats et n'est pas du toute marginale. En revanche, cela forme bien une coupure entre les computationnalistes et les *"dynamicistes"*. **(**Cf. l'excellent collection des textes dans: R. Port and T. van Gelder (Eds.), *Mind as Motion*, MIT Press, 1995). Une des conséquences cruciales de l'approche dynamique, qu'on ne trouve pas dans la vision computationnelle, c'est que vous obtenez des propriétés émergentes, c'est-à-dire des états globaux de l'ensemble de vos variables, parce qu'il y a une interdépendance intrinsèque. Il n'y y a pas besoin de niveau hiérarchique ou chef d'orchestre pour coordonner la chose, c'est la dynamique même qui va la porter.

La cognition ne serait pas une propriété abstraite qui pourrait être reproduite dès lors qu'on aurait trouvé les règles d'agencement et de manipulation des symboles, mais proviendrait de l'interaction des composants biologiques du cerveau ?

L'approche dynamique travaille avec des variables biologiques, avec des activités neuronales plutôt qu'avec des symboles, avec des états globaux du cerveau appréhendés par l'imagerie fonctionnelle. Ce type de travail récuse la séparation entre la cognition et son incarnation. Cela représente au sein des sciences cognitives, outre l'opposition entre calcul et dynamique, un clivage entre cognition comme un niveau autonome et cognition incarnée. Beaucoup de chercheurs en sont venus à considérer qu'on ne pouvait pas comprendre la cognition si on l'abstrayait de l'organisme inséré dans une situation particulière avec une configuration particulière, c'est-à-dire dans des conditions écologiquement situées. On parle de *situated cognition*, en anglais, ou *embodied cognition*, cognition incarnée, ou encore d'*enaction* un néologisme que j'ai introduit il y une dizaine d'années (F.Varela, *Invitation aux Sciences Cognitives*, Seuil, Points Sciences, (1988, 2eme edition 1996); *The Embodied Mind: Cognitive science and human experience*, MIT Press, Cambridge, 1991 (trad. fr. *L'Inscription Corporelle de l'Esprit***,** Seuil, Paris, 1993). Le cerveau existe dans un corps, le corps existe dans le monde, et l'organisme agit, bouge, chasse, se reproduit, rêve, imagine. Et c'est de cette activité permanente qui émerge le sens de son monde et les choses.

Cela nous éloigne de l'intelligence artificielle ?

Non, parce que cette coupure cognition autonome/cognition incarnée traverse toutes les sciences cognitives, et pas seulement les neurosciences cognitives. Un des promoteurs de la robotique incarnée est Rodney Brooks, le directeur du Laboratoire d'intelligence artificielle du Massachusetts Institute of Technology (MIT). Il y a une dizaine d'années, il a lancé une espèce de déclaration de guerre en affirmant qu'on ne réussirait à faire des robots vraiment autonomes que s'ils étaient incarnés dans un contexte matériel, situés dans un monde sensible, sans une représentation du monde. Et il ne s'agissait pas d'un monde défini par une liste de

propriétés, comme on le fait habituellement dans les simulations informatiques. La vision incarnée séduite aussi beaucoup des chercheurs dans des autres domaines, par exemple ce qui travaillent sur le problème du développement de l'enfant.

Comment cela se traduit-il dans leur démarche ?

Dans l'analyse psychologique instaurée par Piaget*, l'enfant atteint vers l'âge de trois ans le stade conceptuel, c'est-à-dire qu'il acquiert la maîtrise de l'activité symbolique, il passe du singulier à l'universel. En fait, Esther Thelen a montré récemment (Thelen, E. and L.Smith (1993), *A Dynamical Systems Approach to the Development of Cognition and Action*, MIT Press, Cambridge) que l'acquisition de la capacité d'abstraction est inséparable de cycles de perception-action que l'enfant réalise sur certains objets, par exemple tous les objets qui contiennent de l'eau, les tasses, les verres, etc. L'enfant va les manipuler intensément pendant une certaine période. Et si vous l'empêchez de le faire, ou s'il est empêché de le faire à cause d'un handicap, cela gêne son développement cognitif. Thelen renverse la théorie antérieure en disant qu'à la base de ce développement, il y a l'incarnation sensori-motrice, le fait que toute perception entraîne une action, que toute action entraîne une perception, donc que c'est une boucle perception-action qui est la logique fondatrice du système neuronal. Brooks a adopté le même point de vue : si j'arrive à construire une machine qui ne sait rien à priori de son environnement, mais qui est dotée des boucles sensori-motrices efficaces, que fera-t-elle ? Elle se baladera partout, comme un bon cafard, et testera sa boucle de réaction/action, jusqu'à la rendre tellement robuste que, après multiples générations, elle se débrouillera dans n'importe quel environnement. L'hypothèse est que sur cette base vont pouvoir émerger des significations telles désirables/indésirable, des catégories universelles de type classes d'objets, voire le langage.

L'essentiel, c'est donc que l'organisme ait une interaction avec son environnement ?

Que l'organisme soit suffisamment incarné dans un environnement pour pouvoir se débrouiller malgré le fait qu'il ne possède pas une représentation d'un monde préalable, son monde émerge avec ses actions, c'est un monde "enacté". Si je lâche ici un cafard, il va grimper l'escalier, faire le tour du jardin, revenir, éviter de tomber dans des trous ou de se coincer dans un coin. Il constitue une entité viable. En tant que système neuronal, le cerveau est fondé sur cette logique de la viabilité : dans son fonctionnement, tout revient à cette recherche de stabilité sensori-motrice. Durant l'évolution, où le système neuronal est-il apparu ? Pas chez les plantes, pas chez les champignons, pas chez les bactéries. Il est apparu chez les animaux. Pour se nourrir, les animaux ont trouvé la solution de manger des proies. Il leur faut donc se mouvoir - et la locomotion est la logique constitutive de l'animal. Et c'est là qu'apparaît le système neuronal, parce que pour chasser, se mouvoir, il faut une boucle perception-action. Sur le plan évolutif, le système neuronal est apparu comme ça, en reliant des senseurs à des muscles, et ces liaisons ont formé le cerveau. C'est sur cette base que des choses plus abstraites ont commencé à se greffer (Pour en savoir plus sur ce point voir H.Maturana et F.Varela (1994), *L'Arbre de la Connaissance*, Addision-Wesley, Paris).

Comme la manipulation de symboles ?

Oui, quoique très tardivement. L'évolution du système neuronal a pris en gros 1,5 milliard d'années. Les trois quarts de ce temps n'ont servi à réaliser qu'une seule chose, des animaux qui se débrouillent de façon sensori-motrice élémentaire. Pour le dire très vite, jusqu'au jurassique*, il n'y avait que des bêtes à la Rodney Brooks, qui ne faisaient rien d'autre que

bouger, chasser; voilà la long et fascinant histoire de l'incarnation dans le monde. C'est bien plus tard que s'est produit un saut quantitatif chez certains animaux, avec l'apparition du langage, d'éléments sociaux, et de capacités symboliques.

Comment se produit ce saut ? Pourquoi des propriétés abstraites symboliques émergeraient-elles chez le robot COG que développe Brooks ?

La réponse n'est pas encore claire. Mais l'activité symbolique n'apparaît pas toute seule. Elle fait toujours partie d'une situation, d'un contexte, qui produit une espèce de dérive vers certains types de récurrences sensori-motrices. Une fois acquise cette capacité d'agir-réagir sur le monde, des interactions entre individus ont pu se produire, formant un nouveau type de boucle de récurrence, qui est la rencontre avec l'autre. Et ça crée justement, dans cette voie de dérive, de nouveaux points de stabilité, des interactions d'orientation - chaque fois que l'autre se met comme ça, je me mets comme ça. Ce type de regard orientatif de l'un sur l'autre, c'est le début de la communication animale.

Mais alors pourquoi la société de cafards, par exemple, n'a-t-elle pas développé une capacité symbolique ?

Le système neuronal ne peut se développer que jusqu'à un certain point. Si votre corps est engoncé dans une armure, comme chez les cafards, ou si vous restez petit, le système neuronal plafonne. Chez les vertébrés, par contre, ce n'est pas vrai. Par exemple, les mammifères marins possèdent une capacité de communication que l'on connaît mal, mais qui n'est pas loin d'une espèce de partage assez profond. Il est toujours difficile de donner un poids à ce type de reconstruction historique. Mais ce qui compte c'est que l'apparition de l'esprit symbolique n'est pas un saut catastrophique, mais plutôt la continuité nécessaire de l'incarnation dans. On sait que l'appareillage symbolique de catégorisation et d'échange est très répandu. Ce qu'on trouve en plus chez les primates, évidemment, c'est le langage au sens strict du terme. Il y a là un phénomène de créativité évolutive sur lequel on sait très peu de choses. Mais le langage est apparu tardivement. Un million d'années par rapport à l'histoire de l'évolution, cela ne représente fait rien du tout.

Vous reliez les limitations de ces organismes aux "choix" évolutifs qui les ont contraints biologiquement. Cela signifie que l'abstraction est liée au cerveau. Comment le robot de Brooks qui, lui, n'est pas dans un substrat biologique va-t-il évoluer ?

C'est la question que se pose toute cette école. Il y a deux réponses possibles. L'une consiste à dire : nous allons faire le travail de l'évolution. Au lieu de construire des robots hyper-performants, à cinq million de dollars l'unité, on va construire des milliers de petites bestioles, avec des petites variantes, on va les laisser dans la nature, et on va sélectionner les plus résistants. Il s'agit de réaliser un contexte évolutif artificiel. Une autre option consiste à réaliser des bestioles viables, robustes, qui peuvent exister, et dans lesquelles on va introduire par copie des capacités, par exemple de mémoire - imitée du cerveau selon des règles de type dynamique neuronale - ou des capacités sociales beaucoup plus fines, des types d'interactions entre les agents. COG est fondé sur la deuxième option. Par exemple, il reconnaît un visage sur la base d'une sélection des traits qui est copié sur ce qui se passe dans notre cerveau. Par suivi du mouvement oculaire, on a repéré que quand des gens regardent les visages, ils forment une esquisse de certains traits du visage. Les chercheurs du MIT ont implémenté ce type de programme informatique. L'ingénieur est bricoleur plutôt qu'éleveur : la philosophie de Brooks est extrêmement pragmatique, il prend ce qui fonctionne, il bricole. Mais revenons

au terme *"évolutif"* employé par Pinker. Dans l'évolution qu'on a évoqué, il existe un tas de possibilités, et ce qui est retenu dans la suite des générations est finalement un effet de fluctuations statistiques. Mais il y a des gens qui continuent à penser que l'évolution ne fonctionne que par optimisation des forces sélectives. Pour chaque situation, il y aurait une espèce d'optimalité. Ce point de vue représente une troisième cassure dans le champ des sciences cognitives, entre les adaptationnistes et, le nom est encore difficile à trouver, je parle de *"dérive naturelle"*, Stuart Kauffman parle d'*"évolution douce"*. Stephen Gould est un éminent autre représentant de cette école. En revanche, Pinker adhère à l'option adaptationniste pure et dure, qui suppose en fait une correspondence exacte entre organisme et monde, une prédétermination.

La notion de dérive douce renvoie à l'hypothèse qu'on ne s'adapte pas à un monde extérieur qui est donné, mais qu'en faisant son travail de maintien de la viabilité, on transforme le monde.

Exactement. La pertinence des actions n'est pas donnée intrinsèquement, mais résulte de la construction de récurrences *via* la logique de l'incarnation sensori-motrice. Il va de soi que l'option qui considère le cerveau comme un système incarné n'est pas adaptationniste. La vision des couleurs en donne une bonne illustration (Thompson,E., A.Palacios and F.Varela (1992), Ways of coloring: Comparative color vision as a case study in cognitive science, *Beh.Brain Scien.* **15**:1-45). Si la vision adaptationniste a raison, les choses ont une couleur intrinsèque. Si la vision de l'évolution douce est pertinente, nous aurons une diversité des mondes chromatiques, non superposables les uns aux autres. Or quand on étudie comment voient les animaux, on constate qu'il y a une très grande diversité des mondes chromatiques, pentachromatique, tétrachromatique, trichromatique, dichromatique, qui ne sont pas superposables, et correspondent pourtant tous à des lignées animales tout à fait viables. Alors, qui voit la vraie couleur ? Nous, les pigeons qui voient en pentachromatique, ou les abeilles qui voient dans l'ultraviolet ? Quelle est la couleur du monde ?

Est-ce à dire que le monde n'existe pas en soi, mais seulement par les perceptions des êtres ?

Quand j'étais jeune, je pensais qu'il n'y avait pas un *"en soi"* du monde. Je n'en suis plus convaincu. Ce que l'on peut dire, c'est qu'on ne peut pas caractériser le monde par des attributs, mais seulement par des potentialités. Il y a certains contraintes qu'il faut respecter. Il y a une différence entre situation prescriptive et situation proscriptive. Dans la situation prescriptive, vous êtes forcés de faire ce que vous êtes en train de faire et rien d'autre. Dans la situation proscriptive, vous pouvez faire ce que vous voulez sauf certaines choses interdites. C'est intermédiaire entre l'anarchie et la règle absolue. Les adaptationnistes sont prescriptifs, notre vision est plutôt proscriptive. Les êtres vivants ne peuvent pas faire n'importe quoi. Du point de vue philosophique, c'est une vision qui renvoie à l'apparaître, a la phenomenalité plus qu'a une objectivité classique. Mais ceci n'équivaut pas à l'inexistence du monde, qui est la position dite "constructiviste".

Comment articulez-vous la question de la conscience avec l'évolution ?

Il faut commencer par distinguer conscience et expérience. Qu'est-ce qu'une expérience, pour une unité cognitive incarnée dans une situation ? Prenons un chien. Il a de la mémoire, il a des émotions, il a du tempérament. Tous ces modules cognitifs, qui se relient à des sous-circuits cérébraux, fonctionnent ensemble. Il y a une intégration à chaque instant, qui fait que ce que

le chien voit, comment il bouge la patte, l'émotion qui l'accompagne, tout ça n'est pas décousu. Il y a une espèce de suite d'émergences, de disparitions et de réémergences, des unités cognitives, modulaires mais intégrées. L'expérience, c'est un locus d'unité cognitive; c'est le thème centrale sur lequel je travaille actuellement. On peut s'accorder sur le fait qu'il y a une expérience de chien, que le chien est un sujet. Et les cafards ? On ne sait pas répondre à cette question. Mais la définition est opérationnelle, c'est-à-dire que si l'on peut établir une mesure de l'intégration de modules cognitifs, on pourra peut-être dire qu'il y a une micro-expérience chez le cafard, et un méso-expérience chez le chien.

Et la conscience ?

Donc l'expérience est quelque chose d'universel à partir d'un certain niveau d'intégration cognitive partagé, certainement, par tous les mammifères. Conclusion évidente : on peut facilement imaginer ce que c'est que d'être un singe. Il y a un article absolument magnifique de Nagel, "Qu'est-ce que ça fait d'être une chauve-souris" (Nagel T. (1974), ‘What is like to be a bat ?’ *Philosophical Review*, **83**: 435-450). Je ne sais pas ce que ça fait d'être une chauve-souris. Qu'il y ait un sujet cognitif, c'est certain. En revanche, il est beaucoup plus facile de se représenter ce que c'est que d'être un gorille, ou ce que c'est que d'être un bébé. Pourquoi poser la question ainsi ? Parce qu'on discerne l'inseparabilité entre l'expérience, telle que je viens de la définir, et les mécanismes d'émergence modulaires. Les sons, les odeurs, n'existent pas en tant que tels, mais seulement relativement au sujet cognitif, au chien par exemple. Ils sont pour le chien une manifestation phénoménale. Donc, au lieu de s'intéresser au mécanisme neurobiologie de l'odorat du chien, vous pouvez, comme le fait Nagel dans son article, regarder du côté phénoménal de cette activité cognitive. Une fois que l'unité de plusieurs modules cognitifs s'est produit, ça donne au sujet une perspective particulière sur le monde.

Et une capacité réflexive ?

L'apparition du langage va faire la différence entre avoir une expérience, reflétée par un comportement neuronal, et la capacité réflexive. Mais il ne faut pas confondre celle-ci et la conscience. La capacité réflexive, dans le vide, ça ne donne rien. Elle doit s'incarner dans un univers cognitif complexe. C'est à l'intérieur de l'expérience qu'il y a cette nouvelle capacité d'auto-description.

C'est l'expérience de soi ?

C'est l'expérience de se référer à soi, de se référer à sa propre expérience. La réflexivité est quelque chose d'absolument crucial, c'est la grande mutation se produit avec l'apparition du langage chez l'homme. Mais où j'ai des problèmes, c'est quand on essaye de coller la conscience à cette capacité réflexive sans faire état de l'énorme *background* que représente l'expérience. Certains chercheurs utilisent le terme de conscience primaire pour désigner la conscience non réflexive. C'est intéressant, parce que dans la vie quotidienne, 90 % de l'expérience est primaire, pas réflexive. On marche, on prend le métro, on peut même avoir des pensées sans qu'il y ait de réflexion.

Mais est-on ou non, avec la réflexivité, dans la conscience ? Est-il même utile de parler de conscience ?

Absolument. Quand je dis que je ne suis pas d'accord quand les gens parlent de conscience, c'est parce qu'ils ont tendance à coller ça uniquement au côté réflexif. Evidemment, je prends

la conscience très au sérieux. Mais si on veut avancer dans cette question, il faut au moins en poser les termes. Un point fondamental, c'est de comprendre la nature de l'expérience tout court. C'est une forme de conscience, mais c'est une conscience sans réflexion. La réflexion, ou la capacité réflexive, va donner à la conscience son statut humain. S'il n'y avait que de l'expérience, je serais plutôt gorille. Mais je sépare les deux questions. Le problème scientifique de l'étude de la conscience humaine n'est que un chapitre de l'étude de l'expérience, parce que cette capacité réflexive nous permet d'explorer la conscience comme si on avait des données phénoménologiques directes. Comment l'étudier ? Par exemple, on place un sujet dans une situation relativement contrôlée, et on lui demande de se rappeler la maison où il a été élevé dans son enfance. La plupart des gens y arrivent sans problème. Ensuite, vous leur demandez comment ils ont fait, c'est-à-dire quelle stratégie ils ont employé pour évoquer le souvenir. Le souvenir lui-même n'a pas d'importance, c'est l'acte de déclencher le souvenir qui représente la capacité consciente. Elle est assez unique, on ne peut pas demander à un gorille d'évoquer ses souvenirs. Donc la capacité réflexive nous ouvre une fenêtre phénoménologique, c'est-à-dire la possibilité d'explorer l'expérience humaine avec une grande rigueur de recueil des données phénoménales (Cf. Varela, F. (1996), Neurophenomenology: A Methodological remedy to the hard problem, *J.Consc.Studies* **3:**330-350).

Il s'agit des données mentales ?
Dans l'ancienne tradition psychologique, on parlait d'introspection. Dans la phénoménologie, on parle de description eidétique des contenus noétiques. Dans la psychologie cognitive on parle des rapports verbales. L'important c'est de se donner une méthode d'approche clair et discipliné **(**Cf. le numéro speciale de la revue *J.Consc.Studies*, "First-person Methodologies", Vol. 6(2-3), 1999).

Autrement dit, la conscience donnerait lieu à une boucle perception-action qui, au lieu de s'appliquer à la perception d'artefacts extérieurs, s'appliquerait à des contenus mentaux ?

Ce qui crée un dynamique tout à fait différente.

Mais est-ce une capacité propre aux humains ?

Je pense qu'elle est liée aux humains, tout simplement parce qu'elle est liée à la réflexion. Ce n'est qu'à travers le langage qu'on induit la capacité réflexive. La capacité réflexive semble intrinsèquement liée à l'apparition des capacités langagières et très vite, elle revient à ça.

Le fait que la capacité du langage semble, au moins à l'état d'ébauche, présente chez d'autres espèces signifie-t-il qu'il en va de même pour la conscience ?

Non, c'est là qu'est la différence : sans langage, pas de capacité réflexive. Vous avez bien une expérience, très complète, très riche, mais vous n'avez pas cette capacité réflexive. Un gorille ne peut pas faire ce geste de retournement, passer du contenu à l'acte de conscience qui donne lieu au contenu.

La conscience est donc proprement humaine ?

C'est un problème terminologique, pour moi ça n'est pas très important. Si vous appelez conscience seulement ce qui contient cette capacité réflexive, la conscience est purement humaine. Mais le point crucial c'est la capacité de dédoublement que fait apparaître le langage

: je peux décrire une description. C'est comme si on enlevait les yeux de la tête, même sans la capacité de voir, il resterait la capacité d'imaginer des objets. Bien que le rêve ou les images imaginaires soient liées à la vision, elles n'en sont pas pour autant inséparables. De même, la capacité réflexive a sans doute a été introduite par la capacité langagière provenant elle-même des interactions sociales. Or, l'exploitation de cette ressource manque dans les recherches concernant une science de la conscience actuelles (Voir par example S.Hameroff, A.Kazniak, and A.Scott (Eds.), *Towards the Science of Consciousness: The Second Tucson Discussions and Debates*, MIT Press, Cambridge, 1998). Pourtant on ne peut pas se passer d'exploiter ce domaine des actes de conscience vus de la première personne. En mettant à égalité de rigueur les données à la première personne et les données à la troisième personne, on doit pouvoir résoudre l'impasse dans laquelle se trouve le débat sur la conscience.

Si l'on se place dans une logique d'incarnation des actes mentaux, ils doivent être perceptibles ?

Oui, la conscience est étudiable. Le tout, c'est de poser la question au niveau juste. Et le niveau juste, c'est le côté émergent de l'expérience, et ensuite les capacités réflexives qui accompagnent la préparation de ces données phénoménales. Donc il y a deux niveaux enchevêtrés qu'il faut étudier séparément.

Ce côté émergent de l'expérience renvoie-t-il à la thèse que la conscience aurait émergé au cours de l'évolution du fait de l'avantage qu'elle apporte (Derek Denton, *L'émergence de la conscience*, Flammarion, coll. "Champs", 1995) **?**

Non. Je parle d'émergence dans le sens dynamique, pas dans le sens évolutif : les capacités cognitives se présentent sous des formes diverses et doivent être *"cousues"* ensemble en permanence, c'est ça que j'appelle le phénomène d'émergence,. En revanche, cela peut se manifester de diverses façons dans l'évolution. Une chauve-souris, une souris, un poisson, un insecte, chacun a son type d'expérience. Et puis la spécificité humaine a joué dans sa situation particulière, unique, qui a ouvert la possibilité réflexive. Là il y a effectivement une émergence sur le plan évolutif.

Qui aurait apporté un avantage évolutif ?

C'est une idée assez adaptationniste. Je m'en méfie, parce que cela suppose qu'il y a eu un paramètre optimum. Qu'est-ce qu'on a amélioré pour que ça soit sélectionné ? Mais les chemins de l'hérédité ne se guident pas par des adaptations. Il y a des interdépendances tellement complexes que l'on ne peut pas parler d'un pic adaptatif.

Alors, pourquoi aurait-elle émergé ?

Parce qu'il y avait, parmi toutes ces possibilités, la possibilité d'émerger. C'est un effet de situation. Cela aurait pu se passer ou ne pas se passer. Il y a un côté très aléatoire, dans le monde, lié à la notion d'évolution douce ou dérive évoqué plus haut. C'est comme si l'ontologie du monde était très féminine, une ontologie de la permissivité, de la possibilité. Tant que c'est possible, c'est possible. Je n'ai pas besoin de chercher à justifier par une optimalité idéal. Au milieu de tout ça, la vie tire les possibilités, elle bricole.

L'idée de machines émergeant comme êtres vivants devrait vous paraître concevable ?

Oui, je n'ai aucun problème avec l'idée qu'on crée des organismes dotés d'une identité somatique, d'une identité de type sensori-motrice, ou d'une conscience artificielle.

Cela fait peur aux gens, ils pensent à Frankenstein.

On change de registre. Cela renvoie à la dimension déontologique et éthique de la science, sa dimension sociale. On n'est plus dans l'ordre du possible, mais dans l'ordre du souhaitable, ce n'est pas la même chose. Je dis que c'est possible. Qu'on décide collectivement, démocratiquement, de le faire ou pas, c'est une autre question.

On en prend le chemin ?

Oui. Un robot comme COG, quand même, c'est un grand pas vers une machine du niveau d'un chien. Regardez ce qu'on peut faire avec des petits robots, c'est assez impressionnant.

Et pourquoi le fait-on ?

Pour les mêmes raisons que dans les autres champs scientifiques : en partie, parce que c'est passionnant à faire, que c'est fascinant, et en partie parce qu'il y a de gros contrats derrière, par exemple pour expédier de telles machines sur Mars •

Propos recueillis par Hervé Kempf

* Jean Piaget (1896-1980), psychologue suisse, a notamment décrit les stades du développement intellectuel de l'enfant.

* Le jurassique est la période de l'ère secondaire qui s'est écoulée de - 205 à - 135 millions d'années.

Francisco J. VARELA
LENA-Neurosciences Cognitives et Imagerie Cérébrale
CNRS UPR 640
Hôpital de la Salpêtrière
47, Blvd de l'Hôpital
75651 Paris Cedex
tel.(331).42.16.11.71 fax.(331).42.16.11.72