

Levinas et les robots

Author : Julien De Sanctis

Categories : [Science & Techno](#)

Date : 8 février 2019

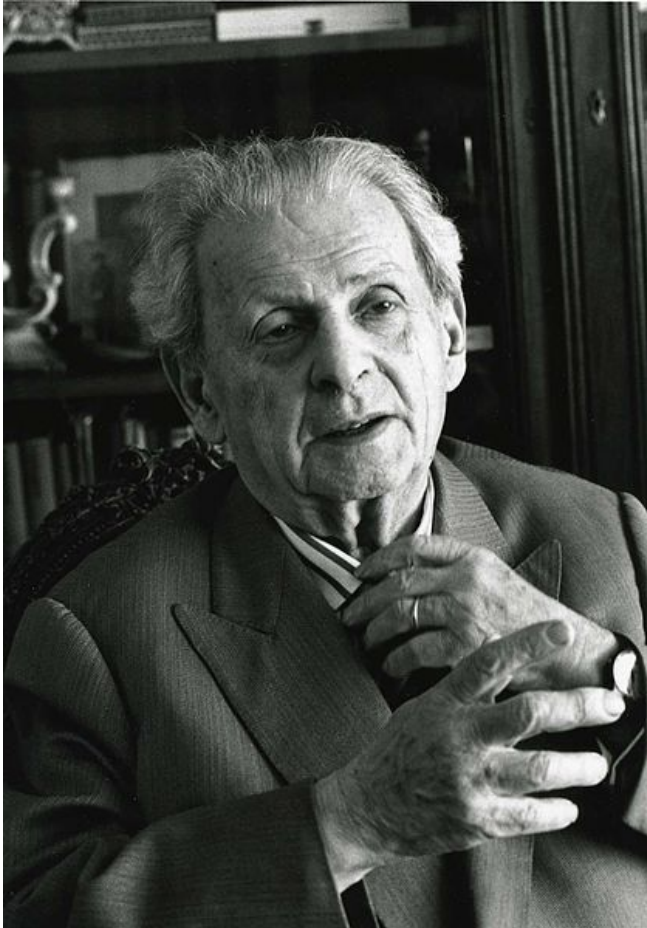
ANALYSE : L'auteur de *Totalité et infini* aurait certainement considéré que les robots, bien que développés de plus en plus à l'image de l'homme, n'ont pas de «visage», remarque Julien de Sanctis, philosophe des techniques. Mais est-ce aussi simple ?

Diplômé de l'ESSEC et de l'Université Paris Panthéon-Sorbonne, [Julien De Sanctis](#) est doctorant à l'UTC en Philosophie et Ethique appliquée à la robotique interactive. Il effectue sa thèse en CIFRE avec Spoon, une jeune start-up créée fin 2015. Chroniqueur pour [iPhilo](#) et [La Pause Philo](#), il travaille avec l'agence [Thaé](#) qui promeut la philosophie pratique auprès d'acteurs privés comme les entreprises.

La robotique sociale, également appelée robotique interactive, est un jeune champ de recherche (et de développement) dont l'objectif général est de produire des artefacts capables d'interagir socialement avec l'être humain. Ces interactions peuvent servir une fonction spécifique telle l'accès à des services numériques divers, ou être utilisées pour leurs vertus potentiellement socialisantes dans différents dispositifs thérapeutiques (accompagnement de l'autisme ou du grand âge). Du point de vue scientifique, on vante également les mérites «réflexifs» de la sociobotique en ce qu'elle représente l'occasion d'étudier, par contraste et confrontation expérimentale, la socialité humaine.

Au-delà, force est de constater qu'il n'existe pas, à cette heure, d'étude poussée du concept quelque peu fuyant de «robot social». En quoi un robot peut-il être qualifié de social ? Quelles représentations du social, conscientes ou non, orientent la conception des sociobots ? Ce sont là des questions que je prends en charge dans ma recherche doctorale, tant elles paraissent nécessaires à la compréhension de ce qu'est la robotique sociale. Mais l'enquête ne saurait s'arrêter à la dimension sociotechnique du domaine. En effet, les roboticiens produisent déjà des entités dont nous pouvons faire l'expérience et dont nous pouvons d'ores-et-déjà dire quelque-chose.

Lire aussi : [Sommes-nous déjà devenus des cybrgs ? \(Julien de Sanctis\)](#)



Le primat actuel de l'expérience sur la théorie – il existe des robots sociaux, mais pas de théorie du robot social – pousse à analyser la sociobotique en termes herméneutiques comme *activité productrice d'altérités nouvelles*. Car c'est bien de cela dont il s'agit : la robotique sociale accouche d'entités que nous pouvons qualifier spontanément d'*autres*, mais que l'appréhension phénoménologique pure, non raffinée par une description a posteriori, ne nous permet pas de singulariser. Une grande partie des efforts techniques en robotique sociale convergent vers la modélisation de comportements et, plus largement, de signaux phénoménologiques incitant à anthropomorphiser les sociobots. Est-ce à dire que ces autres nous apparaissent sur le mode de l'altérité humaine ?

Dans cet article, je m'intéresserais à cette question en convoquant le concept de *visage* développé par Levinas. Selon lui, la rencontre avec l'autre (humain) s'effectue dans la «visagité» qui est un mode éthique d'apparition d'autrui. Malgré les tropismes anthropomorphes de la robotique sociale, on ne saurait intégrer ses créations à la «visagité» telle que la définit le philosophe, à moins de trahir le concept initial en se cantonnant à une analyse phénoménologique trop superficielle. La «visagité» reste toutefois une analyse très intéressante pour penser l'altérité des robots et, surtout, ses effets sur l'humain.

Entre vulnérabilité et responsabilité

Levinas est un penseur incontournable pour qui s'intéresse à la question de l'altérité. On lui doit notamment de profondes analyses de la rencontre avec l'Autre et de ses implications éthiques dans ce qu'il nomme le *visage*. Le concept lévinassien de visage désigne une expérience méta-esthétique qui ne doit pas être confondue avec celle, quotidienne, du visage comme expression animée de la *face*. Au contraire, le visage d'autrui est ce qu'on expérimente, ce qu'on ressent, ce qui s'impose à nous indépendamment de la perception de son visage physique. Dans un entretien conduit par Philippe Nemo en 1984, Levinas déclare :

«Je me demande si l'on peut parler d'un regard tourné vers le visage, car le regard est connaissance, perception. Je pense plutôt que l'accès au visage est d'emblée éthique. C'est lorsque vous voyez un nez, un front, un menton, et que vous pouvez les décrire, que vous vous tournez vers autrui comme vers un objet. La meilleure façon de rencontrer autrui, c'est de ne même pas regarder la couleur de ses yeux.»

Comme j'ai pu l'expliquer [ailleurs](#), Levinas soutient que la philosophie n'est jamais parvenue à penser correctement l'altérité. Les grands penseurs se sont évertués à faire d'autrui l'objet d'une connaissance. Or, «toute connaissance revient à un retour identificatoire à soi [\[1\]](#)», c'est pourquoi tenter de *connaître* l'Autre, de comprendre ce qu'il est nous éloigne nécessairement de l'altérité authentique. La connaissance d'autrui n'est autre que l'expression du retour au Même, car elle ramène vers moi quelque chose qui m'échappe. C'est pourtant bien là le propre de l'altérité : (s')échapper. Le connaissable d'autrui est la part de Même qu'il offre en partage. Sa véritable altérité, ce qui le rend authentiquement autre est insaisissable, incompréhensible, absolument a-théorique, et il faut l'accepter. Le contact avec cet autrui invisible et indicible se fait dans la rencontre avec son Visage, mais un visage sans traits, sans contenu, un visage pure-présence qui déborde toute notion de forme. Le visage d'autrui n'est pas son visage en tant qu'il est un être singulier prénommé Yann, Camille ou Sylvain, mais uniquement en tant qu'il est *autre* que moi. C'est la trace de l'habitant de l'être, celle qui me fait deviner que derrière la chair, derrière l'expressivité de cette personne qui apparaît à mes sens, se cache un *sujet*. «Le visage, c'est l'irruption de l'autre dans mon existence (...), c'est l'exigence d'Autrui qui expulse le Moi de son tranquille repos [\[2\]](#)».



La rencontre avec le visage de l'autre est un *événement métaphysique*. Elle est foncièrement troublante car elle nous arrache à la quiétude du Même. Autrui communique avec moi sans avoir à me parler. Le Visage s'exprime même quand le visage est muet. Cette communication s'offre comme une révélation : elle raconte la misère, la nudité, la vulnérabilité foncière d'Autrui qui m'implore tout en exigeant de moi que je le soutienne, que je sois responsable de lui. La rencontre avec l'Autre n'a donc rien d'épistémologique, elle est d'emblée éthique. Pour Levinas, le Visage est toujours ambigu car la misère qu'il révèle invite au meurtre d'Autrui tout en l'interdisant. Rencontrer authentiquement Autrui, c'est donc faire face à une vulnérabilité qui me responsabilise en ce qu'elle me place immédiatement en situation de *choix originaire* : pour être, je me *dois* de répondre de cet Autre qui me fonde et m'implore. Nul doute que Levinas aurait vivement condamné les drones et ce qu'on appelle aujourd'hui les robots tueurs ou SALA (Systèmes d'Armes Létaux Autonomes) pour la déresponsabilisation réifiante dont ces technologies nous rendent coupables envers Autrui. Le meurtre interfacé a quelque chose du péché absolu car il s'accomplit dans un contexte de totale de *dé-visagéité* de l'Autre, de négation de sa visagéité qui justement fait de lui un Autre. Tuer par l'intermédiaire d'un drone ou d'un robot, c'est refuser l'altérité d'Autrui et décliner l'exhortation éthique fondatrice de mon rapport à lui comme à moi-même.

Conscience et visage

Ceci étant dit, on peut se demander si l'altérité robotique que nous souhaitons explorer rentre dans la sphère levinassienne du visage. A première vue, cela semble impossible car, n'étant pas mortel et, plus précisément, pas tuable, un robot apparaît comme ontologiquement étranger à la vulnérabilité. Est-ce à dire, à l'inverse, que tout être mortel est capable de visagéité ? Les animaux, les arbres, les végétaux en général ou même des bactéries sont-ils des autres de ce point de vue ? Autrement dit, la vulnérabilité se définit-elle, dans l'absolu, par l'inéluctabilité de la

mort doublée de sa possibilité permanente ou n'est-elle liée qu'à la conscience – ou, a minima, la sentience – de la mortalité ? Si la vulnérabilité se résume à la mortalité, alors chaque entité mortelle peut exprimer un visage. Mais si l'on définit la vulnérabilité à l'aune de la conscience plus ou moins claire qu'un être a de sa mortalité, alors le champ de la visagété se réduit considérablement.

Dans la visagété, Levinas lie de façon consubstantielle vulnérabilité et responsabilité. Malgré l'évidence, notons que, dans la description qu'il en donne, le visage émerge au sein d'une relation, ou plutôt d'une rencontre, entre des victimes et des tueurs potentiels. La vulnérabilité procède du choc virtuellement imminent entre mortalité et assassinat, entre possibilité et actualité du meurtre : l'autre est toujours vulnérable en tant qu'il est *face à quelqu'un*. Il est donc difficile de penser la visagété indépendamment d'une conscience capable de *vouloir*, de *choisir* de tuer et ainsi d'être responsable de ce choix. Le visage se manifeste *pour* celui qui a le pouvoir de tuer autrui. Voilà pourquoi, selon Levinas, la visagété est une exhortation à la responsabilité individuelle universelle, c'est-à-dire identique pour chaque homme. En bref, vulnérabilité et responsabilité sont prises dans la valse de leur corrélation, empêchant la visagété d'être élargie à l'ensemble des vivants. Parler de la visagété du zèbre face à la lionne n'aurait aucun sens car celle-ci ne peut accueillir la vulnérabilité du premier, elle n'en a pas le choix. Mais qu'en est-il de l'humain face à un animal sentient comme le chien, la baleine ou encore la vache ?

Visagété simulée

Levinas n'a jamais tranché la question de la visagété animale – le risque aurait peut-être été d'instaurer une forme d'égalité, dommageable à ses yeux, entre l'humain et l'animal ? – mais il est certain qu'il aurait rejeté celle des robots. Pour lui, l'analyse confirmerait l'intuition initiale en la complétant : la rencontre avec un robot est étrangère au visage puisqu'elle ne révèle aucune vulnérabilité ni responsabilité corrélatrice. L'expérience de l'autre robotique ne m'expose pas à l'appel d'autrui, à ce vertige qui me place sans arrêt dans la position du choix éthique originaire. Je peux prendre soin d'un robot, je peux même projeter sur lui des sentiments d'affections tels que l'attachement, mais je ne peux pas le tuer : je ne peux que le détruire. Je ne peux donc pas être responsable de sa « mort » autrement que par une construction sociale, culturelle, religieuse ou psychologique (ce qui n'ôte rien à l'importance, notamment anthropologique, de ces constructions). La chose est foncièrement différente face à autrui : je peux commettre l'irréparable, l'irréversible, je peux le nier en tant qu'autre en m'arrogeant le droit de lui ôter la vie, en opposant à sa volonté de vivre ma volonté de le tuer. En somme, autrui peut être l'objet du mal conçu comme négation, perverse ou non, de son altérité infinie.

Cette expérience méta-esthétique de l'altérité comme éthique originaire n'est pas présente dans la rencontre avec le robot. Mais en est-on réellement sûr ? Depuis longtemps, roboticiens et designers au sens large s'appuient sur notre tendance naturelle à l'anthropomorphisme pour enrichir nos échanges avec les robots, aussi rudimentaires soient-ils. Longtemps associé à une

rémanence de la psychologie infantine, l'anthropomorphisme est aujourd'hui considéré comme une dimension fondamentale de l'esprit humain et représente une opportunité de développement. L'attribution d'intentionnalité et de comportements proches des nôtres aux robots est en effet l'une des composantes de leur bonne intégration à la société. Elle permet notamment de palier l'immaturation technique des artefacts en laissant la psychologie humaine se construire une représentation du robot comme interlocuteur au moins partiellement viable et favorise ainsi l'émergence d'un sentiment de proximité avec la machine. Ce sentiment peut s'épanouir dans un éventail d'émotions propices au développement de diverses formes d'attachements envers l'artefact.

Lire aussi : [Art robotique, un oxymore ?](#) (Charlie Renard)

Ces dernières années, de nombreux articles, académiques ou non, ont rendu compte des effets étonnants voire inquiétants que les robots peuvent avoir sur nos comportements [3]. L'une des études les plus importantes est menée par l'anthropologue et psychologue américaine Sherry Turkle. Son travail porte notamment sur la perception que les enfants et les personnes âgées ont des personnages numériques comme les tamagotchis ou des robots interactifs tels que les Furbies, les My Real Baby du fabricant de jouets Hasbro, ou encore Kismet, conçu par la chercheuse du MIT Cynthia Breazeal. Dans son livre *Alone Together* au sous-titre éloquent «Why We Expect More from Technology and Less from Each Other» [4], Turkle ne cache pas son inquiétude quant aux raisons et conséquences de la confusion ontologique entre ces artefacts et l'humain. Selon elle, le design n'est pas la seule cause de notre attachement aux machines : le sentiment de solitude est également un puissant moteur de projections d'affects sur des entités dénuées d'émotions. La chercheuse résume ses constats d'une phrase concise et percutante : «Ce que nous demandons aux robots est le pendant de ce dont nous manquons cruellement». Selon elle, la robotique interactive représente un tournant majeur de l'interaction humain-machine :

«Ces objets relationnels n'attendent pas que les enfants leur 'donnent vie', comme une poupée de chiffon ou un ours en peluche. Au contraire, ils se présentent comme déjà dotés de vie, déjà prêts pour nouer une relation. Ils promettent un engagement réciproque car, contrairement aux poupées traditionnelles, ils ne sont pas passifs. Ils formulent des exigences. Ils se présentent comme ayant leurs propres besoins et leur propre vie intérieure. Ils nous apprennent les gestes d'amour qui leur sont nécessaires pour s'épanouir. Pendant des décennies, nos ordinateurs nous ont demandé de *penser* avec eux : aujourd'hui, les ordinateurs et les robots, considérés comme sociaux, affectifs et relationnels, nous demandent *d'éprouver des émotions* pour eux et avec eux».

En d'autres termes, le tournant de la robotique sociale est celui du passage d'un rapport purement projectif aux artefacts à un rapport relationnel. Dans ce nouveau rapport, les robots sont

investis d'une dignité ontologique supérieure à celle de leurs prédécesseurs et à celle des objets non relationnels. Pour de nombreux enfants participants à l'étude, les robots sont «assez vivants [5]» pour être traités comme des vivants ; et ce phénomène touche aussi les adultes. Turkle cite l'exemple d'une expérience développée par l'artiste américaine Freedom Baird consistant à maintenir un Furby la tête en bas et à observer combien de temps les sujets testés parviennent à le maintenir dans cette position. Rapidement, la créature artificielle se met à gémir et, quand l'expérience se prolonge, à dire «Moi peur». «Quand ils commencent le test de la tête à l'envers, les adultes sont certains de deux choses : d'une part, que le Furby n'est qu'une machine, d'autre part, qu'ils ne sont pas sadiques. Pourtant, quand le test prend fin et qu'ils laissent derrière eux un Furby gémissant, ils font l'expérience d'un nouveau rapport à l'éthique». Et l'anthropologue de continuer : «Nous sommes arrivés au stade où nous considérons les objets numériques à la fois comme les créatures vivantes et comme des machines. Des plans très différents (celui de la voix, de l'animal de compagnie, de la machine, de l'ami) se rejoignent pour créer une situation dans laquelle le fait de savoir que le Furby est une machine ne change rien au sentiment qu'on peut la faire souffrir». Sur ce point, le témoignage d'une des participantes est particulièrement éloquent :

«[C'est très déplaisant], pas parce que je pense que le Furby a vraiment peur, mais parce que je refuse d'entendre de tels gémissements et de ne pas modifier mon comportement. J'ai l'impression que cette expérience pourrait profondément m'affecter si je la poursuivais [...] Ce n'est pas comme ça que je me comporte dans la vie [...] Et à ce moment-là, le Furby en vient à représenter la façon dont je traite toutes les créatures vivantes».

Ces mots sont riches de sens. Ce n'est pas l'artefact lui-même qui importe ici, mais sa force symbolique, sa capacité à nous projeter dans un rapport à l'altérité et, ainsi, à interroger nos actes, même si nous savons qu'il s'agit d'une machine. Nous avons beau savoir que le Furby n'est pas vivant au sens biologique du terme, qu'il n'est pas sentient et qu'il ne peut donc pas éprouver d'expérience subjective telles que la souffrance ou la détresse, nous nous comportons bien souvent comme si c'était le cas. Tout se passe alors comme si les robots pouvaient invoquer, en nous, le sentiment de la responsabilité. Si l'on en croit le témoignage précédent, cette responsabilité n'est pas vraiment celle de l'humain envers le robot, mais de l'humain envers l'idée de vulnérabilité. L'artefact n'est, dans cette histoire, qu'une sorte de catalyseur phénoménologique, un révélateur d'émotions et d'idées qui, en réalité, le transcendent. Ainsi l'un des grands pouvoirs de la robotique contemporaine est de parvenir à stimuler notre sens de l'éthique, du devoir même, alors que nous sommes conscients d'avoir affaire à une catégorie d'entités ordinairement exclues de ces sphères [5]. L'une des grandes questions éthiques soulevée par la sociobotique est celle des usages à venir de ce pouvoir.

L'altérité est ailleurs

Concluons franchement : le robot n'a pas de visage, du moins pas au sens strict développé par

Levinas. La visag  t   qu'il d  crit est celle *de l'homme envers l'homme en tant qu'homme*. Toutefois, nous venons de voir que les sociobots peuvent provoquer un sentiment proche de l'appel   thique, de cette responsabilit   originaire r  v  l  e dans le visage. Qu'on se rassure toutefois : bien qu'il semble de plus en plus pressant de r  viser nos cat  gories ontologiques dominantes, je n'invite pas    la confusion entre humain (ou animal) et robot. Et j'insiste : tuer – ou faire souffrir – un homme (ou un animal) reste un acte *infiniment* plus grave que d  truire un robot. L'embarras philosophique provoqu   par les questions classiques, mais non moins   pineuses, de la distinction entre l'humain et la machine, entre le vivant et le non-vivant, nous conduit bien souvent sur le terrain de l'authenticit   au sens de correspondance entre fond et forme.

Prenons un exemple : lorsque le Furby dit «Moi peur», a-t-il r  ellement, authentiquement peur ? La r  ponse la plus intuitive (et peut   tre la plus culturelle ?)    cette question est non car,   tant d  nu   de sentience, le Furby ne peut que *simuler* la peur et non la ressentir *effectivement*. Les sociobots sont techniquement capables de simuler toutes sortes d'  motions et d'  tats int  rieurs, mais incapables de les ressentir. C'est le propre de la simulation au sens ici pr  sent   : rester dans l'apparence, dans la superficie du ph  nom  ne esth  tique. Bien entendu, l'humain est lui aussi un tr  s bon simulateur, mais ses simulations sont *intentionn  es*, c'est-  -dire qu'elles correspondent    une intention, qu'elles ont un but. Les simulations humaines sont dissimulatrices, ce sont des *dis-simulations*. Le robot, lui, simule avant tout pour s'int  grer au tissu social en donnant l'impression qu'il est un interlocuteur viable. Pour cette raison, on ne peut consid  rer que maintenir un Furby et un   tre sentient la t  te en bas sont des actes   quivalents.

Lire aussi : [G  nther Anders : l'obsolescence de l'homme](#) (Didier Durmarque)

Ceci   tant dit, le mode d'existence simulationniste des sociobots ne d  l  gitime pas le r  le singulier qu'ils peuvent jouer dans la vaste sph  re de l'alt  rit  , ici analys  e    l'aune de l'  thique. Je pense que la question la plus pressante n'est pas tant de savoir quelles sont les diff  rences ontologiques pr  cises qui les s  parent du genre humain, mais plut  t d'analyser l'effet qu'ils ont sur nous. Comme en t  moigne l'exp  rience du Furby, les robots sociaux ont une force symbolique mat  rialis  e, *encorpor  e* renvoyant    l'id  e de responsabilit   face    l'alt  rit  . C'est donc dans une forme de visag  t   g  n  rique que ces artefacts nous entraînent : nous ne sommes pas responsables envers eux, mais envers la vuln  rabilit   et,    travers elle, envers la responsabilit   elle-m  me. Ainsi per  u, le sociobot appara  t plus que comme un vecteur de r  flexivit     thique accrue pour l'humain.

Dans cet article, je n'ai pas cherch      caract  riser directement l'alt  rit   sp  cifique du robot social, mais    questionner son lien avec une partie de l'analyse que Levinas a de l'alt  rit   humaine. Ce faisant, nous avons vu que la rencontre avec le sociobot en tant qu'entit   poss  dant de nombreux attributs allophaniques peut convoquer en nous un sens de l'  thique qui ne s'attache pas tant aux artefacts qu'   l'  thique elle-m  me. Ce ph  nom  ne est peut-  tre l'une des portes d'entr  es possibles dans l'analyse de l'alt  rit   propre aux robots : le symbolisme.

[1] SALMON Maria, «La trace dans le visage de l'autre», *Sens-Dessous*, 2012/1, n° 10.

[2] *Ibid*, p.107.

[3] Citons, par exemple, le cas très discuté en psychologie de certains soldats américains prêts à risquer leur vie pour «sauver» leur robot démineur qu'ils considèrent comme un frère d'armes.

[4] Traduit en français : *Seuls ensemble. De plus en plus de technologies, de moins ne moins de relations humaines* (éd. L'Echappée, 2015).

[5] Précisons que cette exclusion est propre au monde occidental et que les rapports humains au monde et aux objets sont bien différents selon les cultures comme en atteste l'anthropologie depuis bien longtemps.