

Aux origines de la parole :
des robots pour penser la complexité

Pierre-Yves Oudeyer

Les systèmes auto-organisés

Dans les ruches, les cellules ont une forme hexagonale parfaite qui n'a pas manqué d'attirer l'attention des chercheurs: comment les abeilles en sont-elles venues à bâtir de telles structures ? Un premier élément de réponse apparaît si l'on remarque que l'hexagone est la forme qui nécessite le moins de cire pour couvrir un plan avec des cellules d'une surface donnée. L'hexagone permet donc aux abeilles d'économiser de l'énergie métabolique, ce qui les aide à mieux survivre et à se répliquer plus efficacement que si elles construisaient d'autres formes.

L'explication néo-darwinienne classique suivante a alors été proposée : les abeilles ont dû commencer par construire des formes au hasard, puis par le jeu des mutations aléatoires de leur génome et de la sélection naturelle, petit à petit, des abeilles capables de construire des formes plus efficaces ont été sélectionnées, jusqu'à ce qu'on en arrive à l'hexagone parfait, forme optimale. Or, cette hypothèse sous-entend que la construction des hexagones est génétiquement déterminée et implique une complexité génomique exceptionnelle: le jeu des mutations aléatoires a-t-il pu être assez efficace pour produire à lui seul un tel génome?

L'explication par mutations aléatoires n'est sans doute pas suffisante. Aussi, D'Arcy Thompson, mathématicien et biologiste hors norme du début du 20^{ème} siècle, l'a-t-il complétée¹. Il s'est aperçu que lorsque des cellules de cire de forme pas trop tordue et de taille similaire étaient empilées et chauffées - comme elles le sont par le battement des ailes des abeilles - elles ont à peu près les mêmes propriétés physiques que des gouttes d'eau qu'on

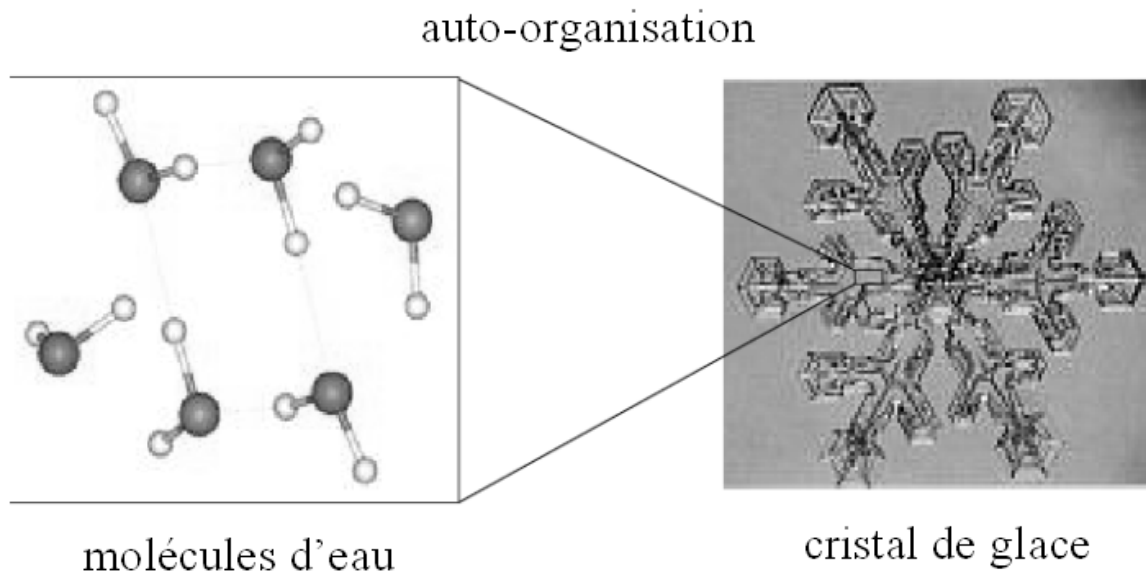
¹ Thompson, D'Arcy Wentworth; *On Growth and Forms* (1917); ré-édité par Cambridge University Press, 2000.

entasse les unes sur les autres. Et justement, il se trouve que quand on entasse des gouttes d'eau les unes sur les autres, elles prennent spontanément la forme d'hexagones !

Ainsi, D'Arcy Thompson a montré que la sélection naturelle n'a pas eu besoin de trouver des génomes qui pré-programmaient précisément la construction d'hexagones, mais seulement des génomes qui faisaient construire aux abeilles des cellules à peu près rondes, pas trop tordues, et à peu près de la même taille ... la physique faisait le reste ! Plus généralement, et bien que le terme n'existait pas encore, il avait surtout montré que des mécanismes physiques *auto-organisés* pouvaient contraindre l'espace des formes et faciliter grandement le travail de la sélection naturelle.

Il a fallu attendre la théorie de la complexité, développée notamment par William Ross Ashby, Heinz von Foerster, Ilya Prigogine, Francesco Varela et René Thom, pour que le mot, et surtout le concept, d'auto-organisation entre véritablement dans le vocabulaire intellectuel qui permet aux scientifiques d'appréhender le monde. L'auto-organisation caractérise les systèmes dont les propriétés macroscopiques résultent de celles de leurs constituants tout en étant qualitativement différentes. On parle d'auto-organisation lorsque ces structures globales organisées apparaissent dans des systèmes qui ne possèdent au départ ni plan, ni modèle réduit pour les faire émerger.

Nous avons vu plus haut l'exemple des ruches des abeilles, mais on trouve des structures auto-organisées de même nature dans les nids de termites, dont la forme n'est ni codée ni connue par aucune des termites prise individuellement, mais apparaît de manière auto-organisée lors de leurs interactions. C'est aussi le cas de la formation des cristaux de glace à partir de molécules d'eau (fig. 1). C'est encore le cas de la formation des zébrures et des surfaces tachetées qui colorent la peau de nombreux animaux. Or ces phénomènes d'auto-organisation sont souvent compliqués à comprendre ou à prévoir intuitivement, et à formuler verbalement.



*Figure 1 : Le phénomène d'auto-organisation.
Les propriétés globales du cristal de glace sont qualitativement différentes des propriétés locales des molécules d'eau*

L'une des manières les plus efficaces aujourd'hui pour développer notre compréhension de la dynamique des systèmes auto-organisés est l'utilisation des ordinateurs ou des robots. En effet, ils permettent d'élaborer des modèles dont on connaît toutes les hypothèses, de les faire fonctionner, et d'en observer le comportement selon les valeurs des paramètres fixés dans le cadre des modèles.

Grâce à cette méthode, l'éthologie a fait un bon en avant dans les 15 dernières années avec les travaux de chercheurs comme Jean-Louis Deneubourg, Guy Theraulaz, Marco Dorigo ou Eric Bonabeau. Ils ont construit des simulations informatiques de sociétés d'insectes, basées sur le concept d'agents informatiques modélisant chaque insecte individuellement. Cela leur a permis d'établir des caractéristiques suffisantes du comportement et des capacités des insectes pour observer la formation de structures collectives, comme la construction des nids chez les termites, la formation de collectifs de chasse ou de recherche de nourriture chez les fourmis, la formation des bancs de poissons, la thermorégulation dans les ruches des abeilles ou la formation de structures sociales chez les guêpes. De manière générale, ils ont montré qu'il n'était souvent pas nécessaire que les insectes soient équipés de structures cognitives complexes pour que, collectivement, ils forment des structures complexes. Grâce à leurs simulations, ils sont même parvenus à faire des prédictions qui ont été vérifiées par la suite sur le terrain.

Les physiciens utilisent aussi de plus en plus l'ordinateur pour construire des simulations de systèmes complexes qui leurs permettent de développer leurs intuitions. En manipulant des automates cellulaires - sortes de grilles dont les cases peuvent être allumées ou éteintes et dont l'évolution dépend de l'état de leurs voisines selon des règles simples - ils ont découvert comment, à partir de structures soit complètement aléatoires soit complètement ordonnées, des motifs complexes avec des symmétries non-triviales pouvaient se former. Les exemples sont fort divers: les cristaux de glace, les distributions des avalanches dans les tas de sables ou dans les montagnes, les dunes dans le désert, les formes des deltas fluviaux, la formation des galaxies ou celle des polyèdres de bulles au pied des cascades.

Pour les physiciens, les automates cellulaires ne sont évidemment pas à proprement parler des modèles des cristaux de glace ou des avalanches, mais ils ont joué un rôle de métaphore et d'analogie qui a déclenché un renouvellement de la manière dont ils percevaient ces phénomènes, au point d'engendrer de nouveaux projets de grandes ampleurs, notamment au niveau européen avec le réseau « Exystence ».

Cependant, il doit être également possible d'utiliser les ordinateurs et les simulations à base d'agents non seulement pour nous aider à comprendre les phénomènes qui caractérisent l'auto-organisation de la matière, des structures biologiques simples, ou des sociétés d'insectes, mais aussi de les faire participer à l'étude des phénomènes qui caractérisent l'homme et ses sociétés. Bref, le temps est venu de faire entrer l'ordinateur et les robots parmi les outils des sciences humaines.

Or, il existe un domaine dans lequel leur utilisation est pleine de promesses : les origines du langage.

*Les origines du langage:
un champ de recherche florissant*

Nos lointains ancêtres ne produisaient que des grognements inarticulés; maintenant nous parlons. La question de savoir comment la parole est venue à l'être humain est l'une des questions les plus difficiles qui sont posées à la science. Alors qu'elle a été écartée de la scène scientifique pendant la presque totalité du 20^{ème} siècle, à la suite de la déclaration de la Société Linguistique de Paris qui la bannit de sa constitution, elle est redevenue le centre des

recherches de toute une communauté de scientifiques. La diversité des problématiques qui sont impliquées induit une forte pluri-disciplinarité : des linguistes, des anthropologues, des spécialistes de neurosciences, des primatologues, des psychologues, etc.

L'un des grands axes théoriques de la recherche sur les origines du langage considère qu'un certain nombre de ses propriétés ne s'explique que par la dynamique des interactions complexes des entités qui sont impliquées (les interactions entre les circuits neuronaux, le conduit vocal, l'oreille, mais aussi les interactions des individus qui les portent dans un environnement réel). C'est pourquoi en plus des linguistes, des psychologues, des anthropologues, des chercheurs en neurosciences, des généticiens et des physiologistes, les mathématiciens et les informaticiens/roboticiens ont désormais un rôle crucial dans cette recherche.

En effet, ainsi que nous l'avons dit plus haut, ils disposent d'outils nouveaux et indispensables pour aider à comprendre les phénomènes d'auto-organisation dans les systèmes complexes. Ces outils permettent de construire des modèles opérationnels pour décrire les interactions entre les sous-systèmes impliqués dans le langage.

Un modèle opérationnel est un système qui définit formellement l'ensemble de ses pré-suppositions et surtout qui permet de calculer ses conséquences, c'est-à-dire de prouver qu'il mène à un ensemble de conclusions données. Il existe deux grands types de modèles opérationnels. Le premier, celui utilisé par les mathématiciens et certains biologistes théoriciens, consiste à abstraire du phénomène du langage un certain nombre de variables et leurs lois d'évolution sous la forme d'équations mathématiques. Cela ressemble le plus souvent à des systèmes d'équations différentielles couplées, et bénéficie du cadre de la théorie des systèmes dynamiques. Le second type, qui permet de modéliser des phénomènes plus complexes que le premier, est celui utilisé par les chercheurs en intelligence artificielle: il consiste à construire des systèmes artificiels implantés dans des ordinateurs ou sur des robots. Ces systèmes artificiels sont composés de programmes qui le plus souvent prennent la forme d'agents artificiels dotés de cerveaux et de corps artificiels; on pourra les appeler robots même s'ils évoluent dans des environnements virtuels. Ces robots sont alors mis en interaction dans un environnement artificiel ou réel, et on peut étudier leur dynamique. C'est ce qu'on appelle la « méthode de l'artificiel ».

La méthode de l'artificiel

La construction de systèmes artificiels dans le cadre de la recherche sur les origines du langage et de l'évolution des langues bénéficie d'une popularité grandissante dans la communauté scientifique en tant qu'outil pour étudier les phénomènes du langage liés à l'interaction complexe de ses composants.

Il y a deux grands types d'utilisation de ces systèmes: 1) ils servent à évaluer la cohérence interne des théories verbales déjà proposées en clarifiant toutes les hypothèses et en vérifiant qu'elles mènent bien aux conclusions proposées (et, souvent, on découvre des failles dans les présupposés ainsi que dans les conclusions qui doivent être révisées) ; 2) ils servent à engendrer de nouvelles théories ou à explorer celles qui, souvent, apparaissent d'elles-mêmes quand on essaie tout simplement de construire un système artificiel qui reproduit les comportements de parole des humains.

Un certain nombre de résultats décisifs ont déjà été obtenu et ont permis d'ouvrir la voie à la résolution de questions jusque là sans réponses: la génération décentralisée de conventions lexicales et sémantiques dans des communautés de robots¹, la formation de répertoires partagés de voyelles ou de syllabes dans des sociétés d'agents, avec des propriétés de régularités structurelles qui ressemblent beaucoup à celles des langues humaines², la formation de structures syntaxiques conventionnalisées³ ou les conditions dans lesquelles la compositionnalité peut être sélectionnée⁴.

Il est important de noter que, dans le cadre de la recherche sur les origines du langage, cette méthodologie de l'artificiel est avant tout une *méthodologie exploratoire*. Elle s'insère dans une logique scientifique d'abduction, c'est-à-dire une logique dans laquelle on cherche des

¹ Steels, L.; *The synthetic modeling of language origins* in Evolution of Communication 1 (1997). Kaplan, F.; *La naissance d'une langue chez les robots*; Hermès, 2001.

² de Boer, B.; *The origins of vowel systems*, Oxford Linguistics, Oxford University Press, 2001. Oudeyer, P.-Y.; Origins and learnability of syllable systems, a cultural evolutionary model, in P. Collet, C. Fonlupt, J. H. E. L. M. S. (Ed.), *Artificial Evolution*; LNCS 2310 (2001); pp. 143–155.

³ Batali J.; Computational simulations of the emergence of grammar; in Hurford, J. R., Studdert-Kennedy, M. and Knight C., editors, *Approaches to the Evolution of Language: Social and Cognitive Bases*; Cambridge University Press; 1998.

⁴ Kirby, S.; *Spontaneous evolution of linguistic structure - an iterated learning model of the emergence of regularity and irregularity*; IEEE Transactions on Evolutionary Computation 5 (2); 2001; 102–110.

prémises qui peuvent mener à une conclusion donnée (au contraire de la déduction dans laquelle on cherche les conclusions auxquelles peuvent mener des prémisses donnés).

Le mot "modèle" a ici un sens différent de son acceptation traditionnelle. Selon cette dernière, modéliser consiste à observer un phénomène naturel puis à essayer d'en abstraire les mécanismes et les variables fondamentales pour construire à partir d'elles un formalisme capable de prédire précisément la réalité. Dans le cas qui nous intéresse, il s'agit plutôt de s'interroger qualitativement sur les types de mécanismes que la nature a pu mettre en œuvre pour résoudre tel ou tel problème. En effet, le langage est un phénomène tellement complexe que la simple observation ne permet pas de *déduire* des mécanismes explicatifs. Au contraire, il est nécessaire d'avoir au préalable une bonne conceptualisation de l'espace des mécanismes et des hypothèses qui pourraient expliquer les phénomènes complexes du langage. Et c'est là le rôle des systèmes artificiels, ceux qu'on appelle parfois "modèles": développer notre intuition sur les dynamiques de formation du langage, et ébaucher l'espace des hypothèses.

Il ne s'agit donc pas d'établir directement la liste des mécanismes responsables de l'origine de tel ou tel aspect du langage. L'objectif est plus modestement d'essayer de faire une liste des candidats possibles, de contraindre l'espace des hypothèses, en particulier en montrant des exemples de mécanismes qui sont suffisants et des exemples de mécanismes qui ne sont pas nécessaires.

Le code de la parole

Pendant de longues années, j'ai travaillé, non sur le problème général de l'origine du langage proprement dit, mais sur celui posé par l'une de ses composantes essentielles, la parole, c'est-à-dire les systèmes de sons en tant que véhicules et supports physiques du langage (au même titre par exemple que les signes visuels, qui sont le véhicule et le support physique dans les langues des signes). Je voudrais dans ce qui suit illustrer par cet exemple ce qui vient d'être dit sur l'utilisation des systèmes artificiels, en en montrant un qui a été construit dans le but de faire progresser la réflexion et la conceptualisation sur la question de l'émergence de la parole.

Les humains ont un système de vocalisations complexe. Celles-ci sont digitales et compositionnelles, c'est-à-dire qu'elles sont construites à partir d'unités élémentaires systématiquement recombinaisonnées, puis ré-utilisées dans les vocalisations. Ces unités sont

présentes à plusieurs niveaux (e.g. les primitives motrices d'obstruction du flux de l'air dans le conduit vocal, qu'on appelle gestes ; les coordinations de gestes, que l'on appelle phonèmes et dont font partie les consonnes et les voyelles ; les syllabes). Alors que l'espace articulatoire est continu et permet potentiellement une infinité de gestes et de phonèmes, chaque langue discrétise cet espace à sa manière en utilisant un répertoire de gestes et de phonèmes à la fois petit et fini.

En outre, malgré la grande diversité de ces unités dans les langues du monde, on y rencontre en même temps de fortes régularités (par exemple, la fréquence élevée du système à cinq voyelles *e, i, o, a, u*). La manière dont les unités sont combinées est aussi très particulière: d'une part toutes les séquences de phonèmes ne sont pas autorisées dans une langue donnée, d'autre part l'ensemble des combinaisons de phonèmes est organisé en types génériques. Cette organisation en types génériques veut dire qu'on peut, par exemple, résumer les combinaisons de phonèmes autorisées en Japonais pour former des syllabes ("moras" plus exactement) par les types "CV/CVC/VC", où par exemple "CV" est un type qui désigne les syllabes composées de deux emplacements, avec dans le premier emplacement uniquement des phonèmes de la catégorie que l'on appelle "consonnes", alors que dans le second emplacement seuls les phonèmes de la catégorie "voyelles" sont autorisés.

En outre, il faut remarquer que la parole est un code conventionnel. Alors qu'il y a des régularités statistiques au travers des langues humaines, chaque communauté linguistique possède sa propre manière de catégoriser les sons, et son propre répertoire de règles de combinaisons de ces sons. Par exemple, les Japonais n'entendent pas la différence entre le *r* de *read* et le *l* de *lead* en anglais. Comment alors une communauté linguistique en arrive-t-elle à former un code qui est partagé par tous ses membres, sans qu'il n'y ait de contrôle supervisé global ?

Depuis les travaux de de Boer ou de Kaplan, on sait comment un nouveau son ou un nouveau mot peut se propager et être accepté dans une population donnée. Mais ces mécanismes de négociation, encore appelés "dynamiques du consensus", font appel à la pré-existence de conventions et d'interactions linguistiques. Ils concernent donc plutôt la formation et l'évolution des langues, mais ne proposent pas de solution quant à l'origine du langage. En effet, quand il n'y avait pas déjà de systèmes de communication conventionnels, comment sont apparues les premières conventions de la parole ?

C'est à cette dernière question en particulier que je me suis intéressé. Elle est évidemment liée à celle de la formation des langues, car il s'agit de comprendre comment un code de la parole a pu être formé pour constituer la base des toutes premières langues. La similarité entre la question de la formation et l'évolution des systèmes de paroles et celle de l'origine de la parole se matérialise donc par le fait que dans les deux cas, la réponse doit être un mécanisme qui permet effectivement la formation d'un système de parole ayant les propriétés fondamentales de la parole humaine telle qu'on la connaît aujourd'hui. La différence principale entre les deux questions réside dans les propriétés qui doivent caractériser le mécanisme que l'on cherche. Pour la question de l'origine de la parole, on doit en particulier chercher un mécanisme explicatif qui ne présuppose ni l'existence de conventions linguistiques, ni l'existence de structures cognitives spécifiques au langage (cela impliquerait qu'on a affaire à des individus qui parlent déjà, et donc pour lesquels le langage est déjà apparu).

*L'auto-organisation au secours
de la sélection naturelle*

Il est donc naturel de se demander d'où vient cette organisation de la parole et comment un tel code conventionnel et partagé a pu se former dans une société d'agents qui ne disposaient pas déjà de conventions. Deux types de réponses doivent être apportés. Le premier type est une réponse fonctionnelle : il établit la fonction des systèmes sonores, et montre que les systèmes qui ont l'organisation que nous avons décrite sont efficaces pour remplir cette fonction. Cela a par exemple été proposé par Bjorn Lindblom¹ qui a montré que les régularités statistiques des répertoires de phonèmes peuvent être prédites en recherchant les systèmes de vocalisations les plus efficaces. Ce type de réponse est nécessaire, mais non suffisant : il ne permet pas d'expliquer *comment* l'évolution (génétique ou culturelle) pourrait avoir trouvé cette structure quasi-optimale, ni *comment* une communauté linguistique fait le "choix" d'une solution particulière parmi les nombreuses solutions quasi-optimales. En particulier, il se peut que la recherche darwinienne "naïve" avec des mutations aléatoires ne soit pas suffisamment

¹ Lindblom, B.; Phonological units as adaptive emergents of lexical development; in Ferguson, Menn, Stoel-Gammon (Eds.); *Phonological Development: Models, Research, Implications*. York Press, Timonium, MD, 1992; pp. 565–604.

efficace pour trouver des structures complexes comme celles de la parole : l'espace de recherche est trop grand¹.

C'est pourquoi un second type de réponse est nécessaire : il faut aussi trouver le moyen d'établir comment la sélection naturelle a trouvé ces structures. On peut pour cela montrer que l'auto-organisation est susceptible, dans ce cas précis, de contraindre l'espace de recherche et d'aider la sélection naturelle. Il suffit de montrer qu'un système beaucoup plus simple s'auto-organise spontanément *en formant la structure que l'on cherche à expliquer*.

En fait, nous reprenons pour la question de l'origine de la parole la même structure argumentative que celle de D'Arcy Thompson à propos des formes hexagonales des cellules de cire dans les ruches des abeilles. Nous allons donc présenter maintenant un tel système et montrer comment des prémisses relativement simples d'un point de vue évolutionnaire peuvent conduire à la formation auto-organisée de codes de la parole.

Le système artificiel

Techniquement, le système artificiel est basé sur le couplage de dispositifs nerveux sensori-moteurs génériques qui sont câblés aléatoirement au départ et "implantés dans la tête" des agents artificiels. Les agents disposent d'une oreille artificielle, capable de transformer un signal acoustique en impulsions nerveuses qui stimulent les neurones d'une carte de neurones artificiels perceptuels. Ils disposent aussi d'une carte de neurones moteurs dont l'activation produit des mouvements d'un modèle du conduit vocal, qui lui même produit une onde acoustique. Les cartes nerveuses (perceptuelle et motrice) sont totalement connectées entre elles. La figure 2 résume cette architecture.

¹ Une argumentation élaborée est développée par Philip Ball dans son livre *The Self-Made Tapestry*, Oxford University Press.

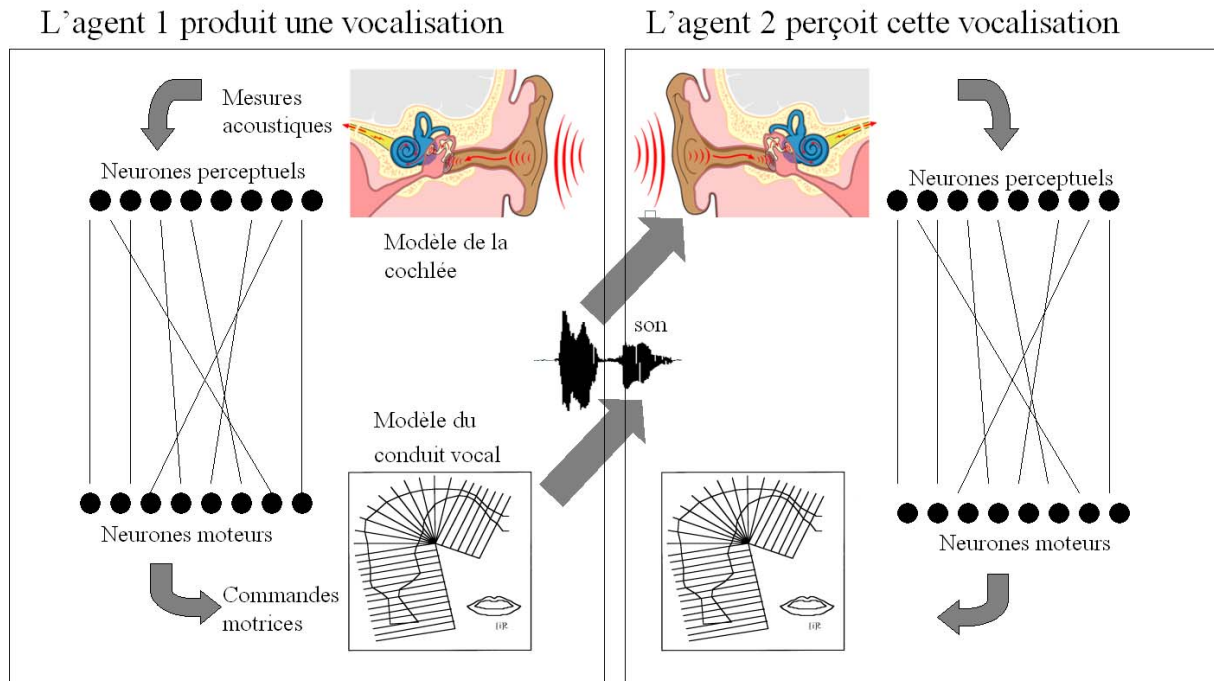


Figure 2: L'architecture d'un agent du système artificiel.

Les agents sont dotés d'une oreille artificielle, d'un conduit vocal artificiel, et d'un cerveau artificiel qui couple ces deux organes. Les agents sont eux-mêmes couplés par l'environnement commun dans lequel ils évoluent : ils perçoivent les vocalisations de leurs voisins.

Les neurones s'adaptent aux stimuli par sensibilisation : leur dynamique est telle que si un stimulus S est perçu, alors ils sont modifiés de telle manière que si l'on présente le même stimulus S juste après ils répondront encore plus. Les connections entre les deux cartes de neurones évoluent selon la loi de Hebb : celles qui relient des neurones qui sont souvent activés en même temps deviennent plus fortes, et celles qui relient des neurones dont l'activité n'est pas corrélée deviennent plus faibles. Ces connections sont aléatoires au début, et grâce au babillage des agents, elles s'organisent de telle manière que l'agent devient capable de trouver les commandes motrices correspondant à un son qu'il "entend".

Les connections entre les deux cartes perceptuelles sont aussi telles que la distribution des sons codés par la carte perceptuelle (celle des sons perçus) reste à peu près la même que la distribution des sons codés par la carte motrice (celle des sons produits). Autrement dit, l'architecture nerveuse de l'agent est telle qu'il a tendance à produire la même distribution de sons que celle qu'il entend. Par contre, il est important de noter que les agents ne reproduisent jamais un son qu'ils viennent d'entendre, et ne stockent pas explicitement un son qu'ils entendent pour le répéter plus tard : en bref, ils ne s'imitent pas. En fait, ils ne disposent

d'aucun moyen de coordination sociale. Ils sont disposés dans un environnement virtuel dans lequel ils se baladent aléatoirement. A des moments aléatoires, ils activent au hasard les neurones de leurs cartes motrices, ce qui produit un son qui est entendu par eux-mêmes et par les agents qui sont à côté d'eux.

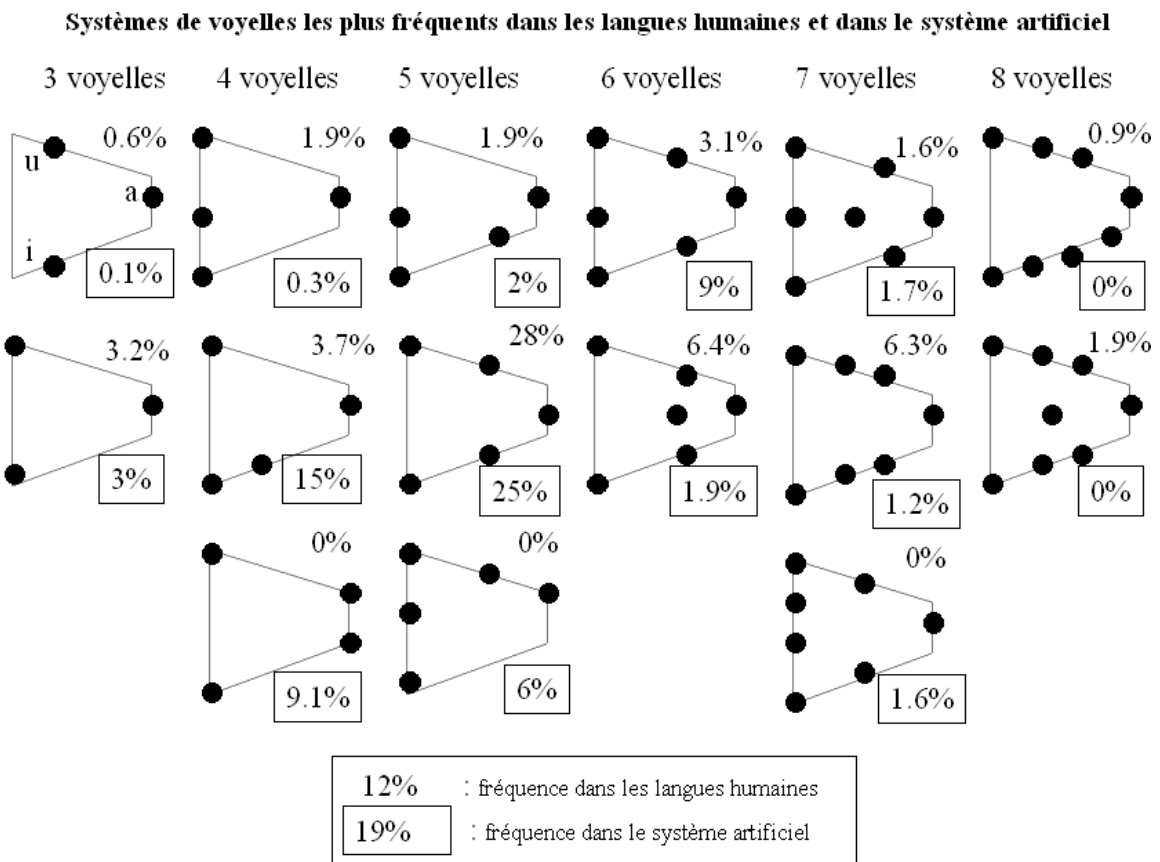
Tous les neurones des cartes motrices et perceptuelles sont initialement aléatoires et uniformes. Cela veut dire que leurs vocalisations sont holistiques (globales) et inarticulées: l'espace continu des configurations aléatoires est utilisé uniformément. Comme tous les agents produisent la même distribution de vocalisation d'après le mécanisme de couplage que j'ai décrit dans le paragraphe précédent, cette situation initiale est donc en état d'équilibre.

Seulement, si l'on fait fonctionner la simulation, on s'aperçoit que cet équilibre n'est pas stable. En effet, il y a du bruit - de la "stochasticité" - qui fait que par hasard et de temps en temps, certains types de vocalisations seront prononcées plus souvent que d'autres. Or, le mécanisme de couplage décrit plus haut introduit une boucle de rétroaction positive: ces déviations de la moyenne sont amplifiées lorsqu'elles sont assez grandes, et la symétrie du système se casse. Les cartes de neurones s'auto-organisent alors en groupes concentrés de neurones, codant pour des configurations acoustiques et articulatoires très précises dans l'espace des vocalisations.

En bref, l'espace continu des vocalisations a été discrétisé. Les vocalisations que les agents produisent ne sont plus holistiques, mais digitales: elles sont systématiquement construites par la mise en séquence de quelques configurations clés, que l'on peut alors appeler phonèmes. On voit apparaître le codage phonémique qu'on a décrit plus haut pour les langues japonaise ou anglaise. En outre, le « code phonémique » qui apparaît est le même chez tous les agents d'une même simulation, alors qu'il est différent d'une simulation à l'autre. On observe donc la formation d'une "convention culturelle", qui peut être diverse d'un groupe à l'autre.

Cependant, l'ensemble des systèmes formés est caractérisé par des régularités statistiques: par exemple, en utilisant un modèle réaliste de la production des voyelles, le système artificiel permet de prévoir quels sont les systèmes de voyelles les plus communs dans les langues humaines (voir fig. 3). De plus, on peut montrer comment ces systèmes font apparaître des règles phonotactiques et une organisation en types génériques des combinaisons de phonèmes

qui sont analogues lorsqu'elles sont partagées par une même communauté d'agents, ou qui sont différentes dans des communautés différentes.



Fi

gure 3

Comparaison entre la distribution des systèmes de voyelles apparaissant dans le système artificiel et celle des systèmes de voyelles dans les langues humaines (d'après la base de donnée UPSID (Maddieson, 1984)).

Les systèmes de voyelles sont représentés sur le triangle vocalique, dont la dimension horizontale correspond au premier formant, et la dimension verticale au second formant effectif. On observe que les systèmes les plus fréquents engendrés par les agents artificiels sont aussi les plus fréquents chez les humains, en particulier le système à 5 voyelles symétrique /a,e,i,o,u/ avec 25 pourcent dans les systèmes artificiels et 28 pourcent dans les langues humaines.

La construction de ce système artificiel a permis de remettre en cause certains dogmes particuliers dans les théories qui tentent d'expliquer la parole. Ainsi, il est couramment admis que le codage phonémique de la parole, c'est-à-dire son caractère discret (alors que l'espace articulatoire est continu) est une conséquence des non-linéarités qui caractérisent le système qui transforme des mouvements du conduit vocal en ondes acoustiques et en perceptions

sonores. C'est par exemple ce que propose la théorie quantale de la parole développée par Kenneth Stevens. Or le système que j'ai présenté permet de montrer que ces non-linéarités ne sont pas nécessaires à l'existence des phonèmes : des codes de la parole discrets et combinatoires peuvent apparaître avec des conduits vocaux et des systèmes de perceptions linéaires. Par contre, la préférence statistiques pour certains types de phonèmes par rapport à d'autres est elle expliquée grâce à ces propriétés de non-linéarités.

De manière plus générale, le mécanisme auto-organisé de ce système apparaît comme un complément nécessaire au mécanisme néo-darwinien classique pour expliquer l'origine de la parole. Il est donc compatible avec un scénario néo-darwinien classique dans lequel l'environnement favorise la réplication des individus capables de communiquer. Dans ce scénario, notre système artificiel joue le même rôle que les lois de la physique des gouttes d'eau dans l'explication des formes des cellules de cire des abeilles: il permet de comprendre comment des mécanismes auto-organisés ont pu faciliter le travail de la sélection naturelle en contraignant l'espace des formes.

En effet, en construisant ce système artificiel, nous avons montré que pour qu'un code de la parole soit formé, la sélection naturelle n'a pas eu nécessairement à trouver des génomes qui pré-programment la formation de structures cérébrales précises et spécifiques qui contiendraient à la naissance le plan de formation d'un système de parole discret et combinatoire. Nous avons aussi montré qu'il n'était pas nécessaire de présupposer l'existence des capacités d'interactions sociales coordonnées, pas plus que celle de la capacité d'imitation explicite ou des motivations pour communiquer, des conventions linguistiques, ni des modèles mentaux de l'autre. Au contraire, il est suffisant que les agents possèdent des connexions génériques entre la modalité acoustique et la modalité motrice vocale (comme celles qui permettent d'apprendre la coordination entre par exemple les mouvements de la main et la vision) et que les neurones qui composent chaque carte nerveuse évoluent par sensibilisation aux stimuli. On voit donc plus facilement maintenant que le pas que l'évolution a dû faire pour permettre à nos ancêtres de passer des grognements inarticulés à la parole n'a peut être pas été si grand.

Remerciements: Ce travail a été rendu possible grâce au soutien de Luc Steels et du Sony Computer Science Laboratory, Paris, qui l'a financé.

Bibliographie:

Pour une description détaillée du système artificiel que j'ai présenté, voir :

Oudeyer, P-Y. (2005) *The Self-Organization of Speech Sounds*, Journal of Theoretical Biology, Volume 233, Issue 3, pp.435—449.

Pour plus d'informations et des documents multimédias illustrant cette recherche, voir :

<http://www.csl.sony.fr/~py>

Pour une vision d'ensemble de la recherche sur les origine de la parole, voir :

Lindblom, B.; Phonological units as adaptive emergents of lexical development; in Ferguson, Menn, Stoel-Gammon (Eds.); *Phonological Development: Models, Research, Implications*. York Press, Timonium, MD, 1992; pp. 565–604.

de Boer, B.; *The origins of vowel systems*, Oxford Linguistics, Oxford University Press, 2001.

Oudeyer, P.-Y.; Origins and learnability of syllable systems, a cultural evolutionary model, in P. Collet, C. Fonlupt, J. H. E. L. M. S. (Ed.), *Artificial Evolution*; LNCS 2310 (2001); pp. 143–155

Studdert-Kennedy, M. and Goldstein, L.; Launching language: The gestural origin of discrete infinity. In M.H. Christiansen and S. Kirby, editors, *Language Evolution: The States of the Art*. Oxford University Press, 2003.

Pour le rôle de l'auto-organisation dans l'origine des formes en biologie, et en particulier les relations entre auto-organisation et la théorie néo-darwinienne de la sélection naturelle, voir :

Kauffman S.; *At Home in the Universe: The Search for Laws of Self-Organization and Complexity*, Oxford University Press, 1996.

Ball, P.; *The self-made tapestry, Pattern formation in nature*, Oxford University Press, 2001.

D'arcy Thompson; *On Growth and Form*, Cambridge University Press 2000.