

Jacques François

La genèse du langage et des langues

Retrouvez nos ouvrages sur
www.scienceshumaines.com
www.editions.scienceshumaines.com

Diffusion : Volumen
Distribution : Interforum

En application de la loi du 11 mars 1957, il est interdit de reproduire intégralement ou partiellement, par photocopie ou tout autre moyen, le présent ouvrage sans autorisation de l'éditeur ou du Centre français du droit de copie.

© **Sciences Humaines Éditions, 2017**

38, rue Rantheaume

BP 256, 89004 Auxerre Cedex

Tel. : 03 86 72 07 00 / Fax : 03 86 52 53 26

ISBN = 9782361064631

Jacques François

La genèse
du langage
et des langues

Éditions
SCIENCES
HUMAINES

INTRODUCTION

En ce début de ^{xxi} siècle, en dépit de la disparition irrémédiable de nombreuses langues due à la suprématie des langues mondiales, l'anglais et le chinois en priorité, les linguistes en comptent encore plus de 6 000, riches pour certaines d'entre elles de plusieurs dialectes. Pour mettre de l'ordre dans ce trésor linguistique, on est convenu de classer ces langues selon différents critères.

Un des premiers critères de classement d'une langue est son appartenance à une famille généalogique, c'est-à-dire ses affinités avec d'autres langues révélant une origine commune, une propriété que les linguistes ont mise en évidence pour les langues finno-ougriennes (essentiellement le finnois, le hongrois et l'estonien) dès le ^{xviii} siècle et pour les langues indo-européennes (principalement romanes, germaniques, slaves et celtiques), et sémitiques (hébreu, arabe, araméen, etc.) dès la première moitié du ^{xix} siècle.

Ensuite vient son appartenance à un type grammatical, une propriété indépendante de la précédente, également entrevue dès le début du ^{xix} siècle par Wilhelm von Humboldt pour les formes des mots (notamment dans les paradigmes de conjugaisons affectant les verbes et de déclinaisons affectant les noms et les adjectifs), mais développée essentiellement dans la seconde moitié du ^{xx} siècle.

À cela s'ajoute occasionnellement l'appartenance de la langue à une aire géographique, laquelle rapproche le lexique et plus rarement les structures de langues de familles généalogiques différentes, mais en contacts persistants. Tout au long du ^{xix} siècle,

les linguistes se sont ainsi demandé si l'albanais était une langue romane compte tenu de la prolifération des emprunts lexicaux à l'italien dans cette langue, avant de la considérer comme un groupe à elle seule.

Il faut se rendre à l'évidence : la question de l'origine commune ou indépendante des familles de langue (en termes techniques, leur monogénèse ou polygénèse), qui a fait couler beaucoup d'encre depuis deux siècles, reste actuellement sans solution, malgré les progrès considérables des méthodes de reconstruction des états passés des langues, y compris celles qui n'ont pas de tradition écrite.

La question de l'origine des familles de langues paraît à de nombreux chercheurs complètement distincte de celle de l'origine du langage, en tant que faculté humaine censée distinguer le genre humain de toutes les autres espèces animales. Et de fait, jusqu'aux alentours des années 1980, les chercheurs investis dans ces deux types d'investigations, des linguistes et philologues d'un côté, des spécialistes de biologie, psychologie et anthropologie évolutionnaires de l'autre, avaient peu d'occasions de confronter leurs vues. Cependant le tableau de la recherche sur les origines de la faculté de langage et de sa manifestation diversifiée à travers les langues du monde présentes et passées a beaucoup évolué à partir de la dernière décennie du siècle passé.

Trois facteurs ont joué un rôle déterminant :

- L'imagerie fonctionnelle et clinique du cerveau a permis de repenser la question ancienne de la localisation des « centres du langage » dans le cerveau et plus généralement de la « biologie du langage ».
- Le développement de la linguistique informatique et notamment la modélisation de l'évolution des systèmes dynamiques (dont le langage est un exemple flagrant) ont permis de construire des scénarios évolutionnaires compatibles avec les découvertes de l'archéologie préhistorique, en simulant l'interaction entre des

agents* multiples (représentant un locuteur et son interlocuteur, ou bien le transmetteur d'un savoir et le récepteur de ce savoir). – Les progrès de l'éthologie et de la psychologie animale, ainsi que l'émergence d'une nouvelle science, la biosémiotique*, ont mis en doute la spécificité humaine de nombreux comportements. On s'est ainsi aperçu que la distribution des fonctions cognitives et affectives entre les deux hémisphères du cerveau humain est partagée par la plupart des primates, et l'étude des vocalisations de certaines espèces d'oiseaux et de cétacés a révélé qu'elles peuvent présenter une grammaire et que, loin d'être limitées à des productions stéréotypées et innées, elles peuvent varier selon le contexte et faire l'objet d'un apprentissage.

Depuis le tournant du ^{xxi}e siècle, une communauté interdisciplinaire s'est constituée autour de la question de l'origine et de l'évolution du langage comme faculté psychologique (cognitive et affective) et des langues comme leur manifestation tangible. La grande question à laquelle ces chercheurs se sont attelés de concert est celle du poids respectif du substrat biologique, et notamment du génome avec l'effet de mutations génétiques sur les aptitudes linguistiques, et des comportements acquis, qu'il s'agisse du développement du langage chez les jeunes locuteurs ou de son évolution dans l'espèce humaine (respectivement l'ontogénèse et la phylogénèse linguistiques).

Ainsi, une nouvelle discipline a vu récemment le jour, la biologie évolutionnaire développementale (abrégée en Évo-Dévo) dont l'un des objectifs est de se représenter les modalités et les phases de l'évolution du langage au cours de la phylogénèse à partir des observations sur l'acquisition du langage chez l'enfant.

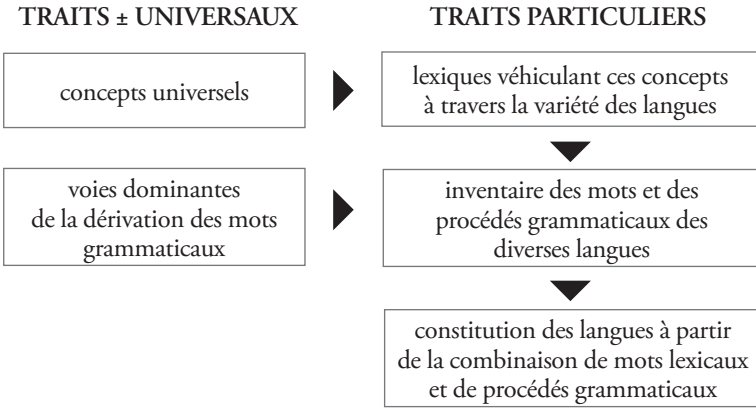
En marge de la biosémiotique évoquée plus haut, une autre discipline, la biolinguistique, s'est constituée, qui défend la thèse téméraire selon laquelle une mutation génétique aléatoire aurait permis à un hominidé, il y a quelque 50 000 ans selon les uns, 200 000 ans selon d'autres, d'acquérir une capacité intellectuelle

nouvelle, celle de combiner logiquement des représentations mentales, capacité qui, transmise à sa progéniture, aurait assuré à un clan un avantage intellectuel tel que les premiers hommes dotés de ce cerveau reconfiguré auraient fini par supplanter toutes les autres sociétés d'hominidés.

Certains chercheurs voient dans cette supériorité intellectuelle l'un des facteurs qui ont permis à l'espèce *Homo sapiens* de supplanter les néandertaliens dans l'Europe de la dernière glaciation, alors même que le cerveau de ces derniers était globalement plus volumineux.

Reste le grand écart entre l'étude de l'origine de chaque famille de langue séparément et celle de l'origine de la faculté de langage. Le constat d'un espace béant entre ces deux phases évolutives, l'une historique, l'autre préhistorique, a laissé la place, depuis le tournant du ^{xxi}^e siècle, à une vision innovante de la genèse de ce que les encyclopédistes du ^{xviii}^e siècle appelaient la « grammaire générale ». Le propos est cependant complètement renouvelé : il ne s'agit plus de rechercher dans le latin et le grec classique des traits grammaticaux susceptibles d'avoir une valeur universelle, mais de prendre pleinement en compte le caractère incarné des lexiques, c'est-à-dire la fondation de chacun d'eux à partir, d'une part, de représentations issues du corps humain (notamment celle de l'espace en termes de haut, bas, avant, arrière, gauche et droite) et de ses capacités (par exemple de mouvement avant, arrière, ascendant ou descendant) et, d'autre part, de la perception du temps et de la causalité.

Les linguistes spécialistes de la variété des langues ont pu dégager les voies dominantes de la dérivation des nouveaux mots grammaticaux à partir des mots lexicaux, ou plutôt des concepts de base communs aux différentes sociétés et cultures. Cet inventaire des voies de la grammaticalisation* permet certainement de relier l'origine des langues et celle du langage, ce que résume la figure ci-après.



Les chapitres 1 à 5 qui suivent sont centrés sur l'émergence de la faculté de langage. Ils abordent en premier les conditions de l'émergence de la parole (chap.1) et le langage en tant que fonction cognitive (chap.2) dotée d'un substrat génétique (chap.3), avant d'évoquer en détail divers aspects de la recherche interdisciplinaire sur la genèse du langage, à savoir les apports de la psychologie et de l'anthropologie évolutionnaires associés à des découvertes archéologiques (chap.4), puis ceux de la sémiotique*, de la philosophie évolutionnaire et de la modélisation informatique (chap.5). Le chapitre 6 traite de l'activité de symbolisation, dont le bioanthropologue Terrence Deacon estime qu'elle permet de définir une espèce indépendamment de tout critère anatomique actuellement identifiable, *Homo symbolicus*, et qui a conduit, sans doute par le processus de sélection relâchée*, à une assimilation génétique.

Les chapitres 7 et 8 portent enfin sur l'origine et l'évolution des langues. Le chapitre 7 traite de l'origine des classes de mots et de leurs combinaisons, lesquelles ont graduellement attribué un statut grammatical aux mots répétés le plus fréquemment. Nous avons là le chaînon manquant entre la genèse du langage en tant

que faculté universelle et la genèse des langues particulières, à savoir le processus universel par lequel les protolangages* ont pu donner lieu à de véritables langues structurées. Le chapitre 8 recense, pour conclure, les méthodes qui, depuis le XIX^e siècle, ont permis à la linguistique de développer une composante historique et finalement évolutionnaire. Au final, l'épilogue suggère que la recherche sur les trois niveaux généalogiques, celui de la parole et de la faculté de langage, celui des mécanismes généraux de la grammaticalisation et celui de la diversification des langues, réclame une vision de l'enchaînement de ces processus comme un monumental système adaptatif complexe*. Seule une telle perspective permettra de disposer progressivement les pièces du puzzle qui se laisse formuler ainsi : comment la communication d'abord orale, puis écrite, diffusée par l'imprimerie et démultipliée à l'infini par la technologie informatique, a-t-elle pu se dégager de la gangue informe des protolangages et gérer efficacement le triple défi de l'intercompréhension, de la mémorisation et de l'émergence de la pensée logique ?

Les mots suivis d'un astérisque sont explicités dans le glossaire qui figure en fin d'ouvrage.

Chapitre 1

L'émergence de la parole

S'interroger sur les conditions et les modalités de l'émergence de la parole dans l'espèce humaine revient à répertorier et à ordonner une série de points d'interrogations : qui ?, quand ?, où ?, comment ?, pourquoi et pour quoi faire ? Il va de soi que l'accès direct à des réponses argumentées à ce catalogue de questions nous est actuellement (et peut-être pour toujours) interdit. Est-ce une raison suffisante pour en faire l'économie et se contenter de ce que nous pouvons reconstituer de l'histoire des langues ? Certaines données empruntées à la génétique des populations, à l'histoire de la terre et du climat, à l'anthropobiologie*, à l'étude de l'acquisition de la langue maternelle (domaine de la linguistique développementale) et à celle des cultures des dernières populations de chasseurs-cueilleurs (un secteur de l'anthropologie culturelle) permettent de lever un voile du mystère, au risque bien entendu du désaccord entre spécialistes sur la valeur des preuves indirectes* qui peuvent y être appliquées pour l'apparition des premières vocalisations* assimilables à de la parole.

Qui ?

Les chercheurs s'entendent pour considérer que la parole n'a pu émerger que dans l'une des espèces du genre *Homo*. L'un des arguments pour cette limitation vient de l'observation de sociétés de chimpanzés et de bonobos : les primatolo-

gues¹ estiment que si ces grands singes anthropomorphes ne se communiquent des indications ou des requêtes que par la gestuelle et limitent leurs vocalisations à l'expression des émotions, c'est parce que l'échange d'informations y est peu utile. Les individus vivent continuellement en groupe et chacun est informé de ce que les autres font et sans doute ont à l'esprit, l'univers social est clos et paisible à condition de respecter les pratiques de toilettage mutuel. Inversement, les anthropopithèques qui ont quitté la forêt pour pratiquer la chasse dans la savane et sont devenus omnivores ont développé à la fois des outils meurtriers et des vocalisations informatives en raison de leur dispersion dans un environnement où ils ne pouvaient plus coordonner visuellement leurs mouvements. Quant à la question subsidiaire (« quelle espèce du genre *Homo* a pu pratiquer quel type de parole? »), elle est liée à la suivante.

Quand et où?

Les bioanthropologues évoquent deux migrations à partir de l'Afrique à deux époques très distantes l'une de l'autre. La première est le fait de l'espèce *Homo erectus*, il y a environ deux millions d'années, la seconde celle de l'espèce *Homo sapiens* il y a environ cent mille ans. C'est l'hypothèse dite *out of Africa* ou monocentriste, l'Afrique étant la zone de départ des deux migrations. Une seconde hypothèse, dite pluricentriste, suggère que les *Homo erectus* dispersés à travers l'Eurasie ont connu des évolutions analogues ayant engendré différents groupes d'*Homo sapiens*. Cette seconde hypothèse est plus spéculative, puisqu'elle implique un parallélisme évolutif d'une probabilité douteuse. Une troisième hypothèse, dite réticulée (soit : en réseau) suggère une hybridation génétique entre les *Homo erectus* et néandertaliens résidant en Eurasie depuis deux millions d'années (et supposés avoir évolué) et les nou-

veaux migrants du genre *Homo sapiens*. Il est difficile actuellement de trancher, mais du point de vue linguistique, les deux premières hypothèses, monocentriste et pluricentriste, entrent en résonance avec la question de la monogénèse* ou de la polygénèse* des langues.

Dans le premier cas de figure, celui de l'apparition de l'espèce *Homo sapiens* en Afrique entre la première et la seconde migration (hypothèse monocentriste), les nouveaux migrants sont censés avoir supplanté les espèces présentes d'*Homo pre-sapiens*, notamment les néandertaliens, en raison de leur faculté d'intelligence individuelle et sociale supérieure. L'hypothèse de la monogénèse africaine des langues passées et présentes est alors avantagée.

Dans le second cas, celui de l'apparition dispersée de l'espèce *Homo sapiens* en Afrique et en Eurasie, la seconde migration, celle des *Homo sapiens* africains, aurait eu un impact linguistique moindre, car on peut supposer que les uns et les autres avaient déjà dépassé à cette époque le stade du protolangage* et l'hypothèse de la polygénèse des langues, encore élémentaires, des premiers hommes d'Afrique et d'Eurasie gagne en probabilité. Dans cette hypothèse, il est probable que ces langues avaient des degrés de grammaticalisation, et donc d'efficacité variable. Le maniement alternatif de deux langues grammaticalisées à des degrés divers a été possible dès lors que les usagers de la première sont parvenus à se rapprocher intellectuellement de ceux de la seconde pour pratiquer un jeu analogue sur les concepts, consistant par exemple à passer du concret à l'abstrait, à comparer un concept superordonné et un concept subordonné pour choisir entre le premier et le second, et à construire une représentation mentale complexe en combinant, hiérarchisant et emboîtant les concepts.

À titre d'exemple élémentaire, si un usager X de la langue L_1 plus avancée sur cette voie s'adresse à un usager Y de la langue

L₂ en employant les mots de L₂ pour déclarer : *L'orage menace*, Y peut comprendre les mots employés sans savoir que faire de la combinaison « orage » + « menacer », car pour lui une menace vient obligatoirement d'un ennemi. Mais si X associe à cet énoncé une gestuelle et une mimique effrayante, il est probable que Y pourra associer la représentation des éclairs, du tonnerre et des inondations avec les signaux iconiques fournis par son interlocuteur, et pratiquer ainsi une inférence de cet ordre :

Quand un ennemi me fait peur il menace/il est menaçant
 → OR les orages me font peur
 → DONC un orage qui me fait peur menace

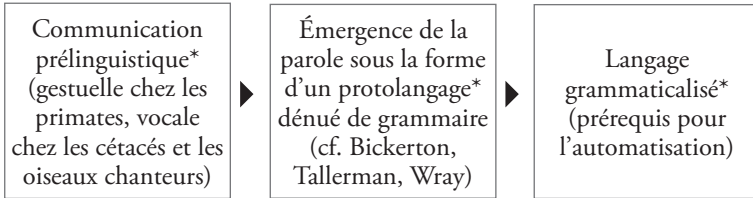
Comment ?

La communication prélinguistique des premiers représentants du genre *Homo* était probablement constituée de syllabes autonomes pointant sur des choses, des personnes et des lieux et évoquant des événements. Le protolangage avancé des *Homo erectus* comportait probablement des chaînes de mots articulés dans un ordre fixe et significatif. Il pourrait s'agir de sortes de mots hypercomposés attestés dans les actuelles langues polysynthétiques².

Le langage grammaticalisé qui en est dérivé s'est caractérisé par l'aptitude à combiner des mots lexicaux (seulement des noms et des verbes à l'origine) et des mots grammaticaux (conjonctions, pronoms, déterminants, prépositions et postpositions) selon des schémas réguliers, et sans doute par l'association de sens différents à des groupes syntaxiques consti-

2. Les langues dites polysynthétiques, notamment amérindiennes, pratiquent l'agglutination de morphèmes* de types divers, lexicaux et grammaticaux, en formant des mots très complexes susceptibles de correspondre en français à une phrase élémentaire. B. Lee Whorf (*Linguistique et anthropologie*, 1969) en donne un exemple en nootka, une langue amérindienne parlée sur l'île de Vancouver : *tlim'sh-ya-'is-ita-'iyl-ma* (bouill-i-mang-eurs-aller_chercher-il_fait) qu'il reformule en « il (ou quelqu'un) va chercher (invite) des mangeur de nourriture cuite ».

tués de morphèmes placés dans un ordre différent (c'est, par exemple, l'origine de la distinction entre l'adjectif épithète et l'adjectif attribut et entre les voix active et passive³). Le diagramme ci-dessous résume cette évolution :



Pourquoi et pour quoi faire ?

L'articulation de signes linguistiques enchaînés de manière fixe et ordonnés hiérarchiquement de manière complexe a permis de fonder des sociétés aptes à gérer les périls externes venus des prédateurs (la parole est indispensable à la coordination de la défense et à la préparation d'entreprises collectives, p.ex. une migration⁴), les périls internes en cas de conflit (la langue est un outil de paix sociale⁵) ainsi que la distribution équilibrée du pouvoir (les hommes étant essentiellement occupés à transmettre leurs gènes, ce sont les femmes qui organisent l'espace social).

Les gestes articulatoires à la source de la parole

La question des conditions de l'émergence de la parole humaine a donné lieu à deux hypothèses, formulées à un quart

3. Certaines langues, comme le kimbundu, une langue bantoue de l'Angola (citée par Talmy Givón) forment une voix passive intermédiaire entre les voix actives et passives telles que nous les connaissons dans les langues européennes, sous la forme d'une construction active à sujet indéfini combinée à un complément d'agent : *Nzua a-mu-mono kwa meme* (John ils-l'ont-vu par moi) → *John a été vu par moi*.

4. Ch. Coupé & J.M. Hombert (2005), « Les premières traversées maritimes : une fenêtre sur les cultures et les langues de la préhistoire ». J.M. Hombert (dir. : *L'origine du langage et des langues*, chap.5), Fayard ; J.L. Dessalles (2000), *Aux origines du langage : Une histoire naturelle de la parole*, Hermès Sciences.

5. R. Dunbar (1998), *Grooming, Gossip and the Evolution of Language*. Harvard University Press.

de siècle de distance, qui restent actuellement en concurrence. La première, émise en 1968, se dit « formaliste et universaliste »⁶, la seconde date du tournant du ^{xxi}^e siècle et se dit « fonctionnaliste et auto-organisatrice »⁷.

Selon la première hypothèse, nous disposons dans notre patrimoine génétique d'une triple prédisposition acquise au fil de l'évolution de l'espèce *Homo sapiens*. Nous sommes prêts à :

1. distinguer des traits phonologiques (découverts en 1939 par N. Troubetzkoy)⁸
2. articuler, dans l'opération de production, des phonèmes composés d'un bouquet de traits concernant le passage ou l'obturation de l'air expiré par la bouche ou le nez, la position des organes phonatoires, l'intervention des cordes vocales, etc.
3. identifier ces mêmes phonèmes* dans l'opération de réception.

En fait il semble acquis que ces deux dernières opérations sélectionnent d'abord des syllabes et dans un second temps les phonèmes qui les constituent.

La seconde hypothèse, plus récente, fait valoir qu'il n'est plus nécessaire de s'imaginer les phonèmes et les traits phonologiques comme des axiomes innés. Car ces derniers « émergent de contraintes antérieures sur la perception, l'articulation et l'apprentissage, conformément à des principes biologiques généraux d'auto-organisation » (M. Studdert-Kennedy).

Un exemple emprunté à la chimie permettra de mieux comprendre l'hypothèse de l'auto-organisation des unités linguistiques. En chimie, deux atomes (plus exactement deux ions) comme le sodium (Na⁺) et le chlore (Cl⁻) se combinent pour former du chlorure de sodium Na⁺Cl⁻ (le sel commun).

6. N. Chomsky & M. Halle (1968), *The Sound Pattern of English*. MIT-Press.

7. P. MacNeilage (2008), *The Origin of Speech*, Oxford University Press.

8. N. Troubetzkoy (1949), *Principes de phonologie*, Paris, Klincksieck.

Par ailleurs l'ion sodium Na^+ se combine aussi avec l'ion HO^- pour former de l'hydroxyde de sodium (la soude caustique), tandis que l'ion chlore Cl^- se combine avec l'ion hydrogène H^+ pour former l'acide chlorhydrique H^+Cl^- .

De même, deux phonèmes comme la consonne occlusive bilabiale⁹ sonore [b] et la voyelle centrale ouverte [a] se combinent en français dans la syllabe [ba], laquelle peut figurer dans un mot (ex. *BATEAU*, *raBAT*) ou constituer un mot (ex. *le BAS*) ou un morphème constitutif d'un mot (ex. *le contre-BAS*), ce que M. Studdert-Kennedy résume ainsi : « À un certain niveau les unités retiennent leur intégrité tout en se combinant pour fournir des structures superordonnées avec des éventails de fonctions qualitativement différents par rapport à leurs composants. »

En effet, les usagers du français sont aptes, d'une part, à opposer d'un côté [ba] et [pa] (*le bas* ≠ *le pas*) et d'un autre côté [ba] et [bô] (*le bas* ≠ *le beau*). D'autre part, ils parviennent à identifier [ba] comme la première syllabe de *bateau* ou la seconde de *rabat*, [pa] comme la première syllabe de *parent* et la seconde de *compas*, et [bo] comme la première syllabe de *beauté* ou la seconde de *rabot*.

Bien que les trois syllabes [ba, bo, pa] soient immédiatement perçues, dans un contexte approprié, comme des noms monosyllabiques (*bas*, *beau*, *pas*), cette aptitude à distinguer les syllabes apparentées à ces mots prouve que l'identité des deux constituants de rang élémentaire (les phonèmes) est préservée dans la reconnaissance des constituants de rang intermédiaire (les syllabes).

Cela prouve aussi qu'à l'échelon supérieur l'identité des syllabes est préservée dans la reconnaissance des mots comme le montre le tableau ci-après.

9. Consonne interrompant le passage de l'air expiré par clôture des deux lèvres .

CHIMIE			LINGUISTIQUE	
molécule	Na ⁺ Cl ⁻ (chlorure de sodium)	↔	<i>bas</i> /ba/, <i>bateau</i> / bato/, <i>badaud</i> / bado/	mot
atome	Na (sodium)	↔	/ba/	syllabe
constituants subatomiques	11 protons ¹	↔	[b] consonne occlusive bilabiale sonore [a] voyelle centrale ouverte	phonèmes

1- Le nombre des électrons varie en fonction de l'ionisation.

Cette organisation hiérarchique assure que – sauf en cas d’ambiguïté – les productions imprécises d’un locuteur sont rangées automatiquement par son interlocuteur dans des classes structurées de sons linguistiques (des classes de syllabes au niveau intermédiaire et des classes de phonèmes au niveau sous-jacent).

Selon la première hypothèse formaliste et universaliste, les constituants primitifs des phonèmes et des syllabes sont les traits phonologiques supposés innés, alors que selon la seconde, fonctionnaliste et auto-organisatrice, ce sont les gestes articulatoires*, et chez les hominidés l’évolution a sélectionné (sous la pression de conditions sociales et cognitives complexes) la « machinerie vocale » la mieux adaptée à l’exécution rapide des gestes articulatoires, malgré la nécessité de préserver le fonctionnement primaire d’organes originellement destinés à l’absorption d’aliments.

L’évolution des structures anatomiques du cerveau a sélectionné les gestes articulatoires de la parole ainsi que d’autres types de gestes à fonction communicative, notamment ceux de la main et de la face. Les mammifères, et plus particulièrement les primates, ont commencé par développer des applica-

tions de la motricité orale à la succion, au léchage, à la mastication et à la déglutition. Ces capacités, qui assuraient déjà la gestion équilibrée de deux besoins primaires, la respiration et l'alimentation, étaient le prérequis pour le développement des gestes phonétiques (par exaptation*). Ainsi la production des consonnes bilabiales sourde [p] et sonore [b] présuppose la mobilité des lèvres qui intervient déjà dans l'aptitude innée à la succion chez les nouveau-nés des espèces de mammifères. L'évolution a sélectionné dans l'espace phonatoire plusieurs systèmes moteurs différents sous contrôle cérébral, notamment ceux de l'apex (l'extrémité de la langue), de son tronc, du voile du palais et du larynx, et chaque geste phonétique résulte de la combinaison particulière de valeurs de ces différents systèmes neuromoteurs. Il est probable que leur organisation, par exemple la répartition généralement symétrique des systèmes vocaliques entre voyelles antérieures (formées dans la bouche) et postérieures (formées contre le voile du palais, dans le pharynx et le larynx), s'est graduellement affinée sous l'effet du mimétisme vocal, lequel a sans doute bénéficié de l'intervention des neurones-miroirs¹⁰.

Les sons de la langue forment système

On sait, depuis les expérimentations de Jacques Mehler dans les années 1980, que, dans un premier temps, les jeunes enfants sont capables de distinguer de nombreux sons linguistiques qui ne font pas partie du bain linguistique auquel ils sont exposés. Mais, à partir du moment où ils prennent garde à cet environnement de productions orales, ils perdent cette capacité pour se concentrer sur le repérage de chaînes de sons dotées d'un sens, dans les énoncés qui leur sont adressés ou dont ils sont témoins, donc des signes linguistiques.

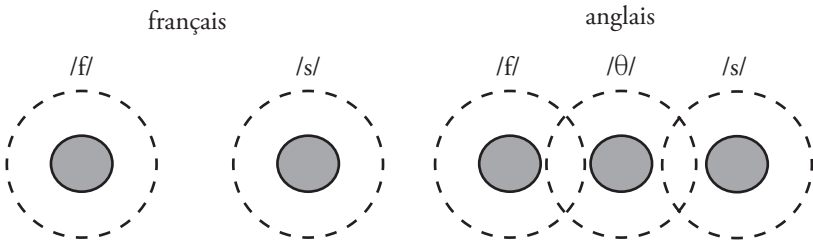
10. Arbib M. (2012), *How the Brain got Language : The Mirror system Hypothesis*. Oxford University Press.

Pour le locuteur débutant, les sons constitutifs de signes linguistiques se présentent comme un continuum sans démarcations évidentes et impliquant différentes dimensions. Ainsi pour les consonnes, le passage de l'air expiré peut être interrompu entre les lèvres (/p/, /b/) ou entre la lèvre inférieure et les dents supérieures (/t/, /d/) ou entre la langue et le voile du palais (/k/, /g/) ou bien il peut être simplement canalisé entre les dents (/f/, /v/), entre la langue et les dents supérieures (/s/, /z/) ou entre la langue et le palais (/ʃ/, /ʒ/), etc. Chacune de ces paires de sons se distingue en fonction de l'intervention (consonnes sonores) ou pas (consonnes sourdes) des cordes vocales situées dans le larynx. En outre l'air peut être expiré par le pharynx et le nez, produisant des consonnes nasales (/m/, /n/, /ɲ/ comme dans *bagne*).

Au fil de leur évolution, les différentes langues tirent parti de certaines distinctions et en négligent d'autres, ce qui peut entraîner des difficultés dans l'acquisition d'une seconde langue, car l'apprenant doit d'abord s'entraîner à discerner la distinction nouvelle de manière à pouvoir la reproduire. Un exemple permettra de comprendre comment ce continuum des productions phonétiques se structure en fonction des distinctions opératoires, c'est-à-dire qui permettent de ne pas confondre deux signes linguistiques de sens différent. Le français distingue par exemple les consonnes fricatives (ou constrictives, avec canalisation du passage de l'air) sourdes /f/ et /s/ (sans vibration des cordes vocales) *fort* [fɔʁ] ~ *sort* [sɔʁ], tandis que l'anglais distingue une troisième consonne fricative sourde /θ/ entre ces deux premières : *fin* (nageoire) /fɪn/ ~ *thin* (fin, adj.) /θɪn/ ~ *sin* (péché) /sɪn/.

La figure ci-après représente schématiquement, pour chaque consonne du français à gauche et de l'anglais à droite, l'espace flou de leur production (les articulations approximatives, délimitées par un cercle discontinu) et leur valeur focale (l'articu-

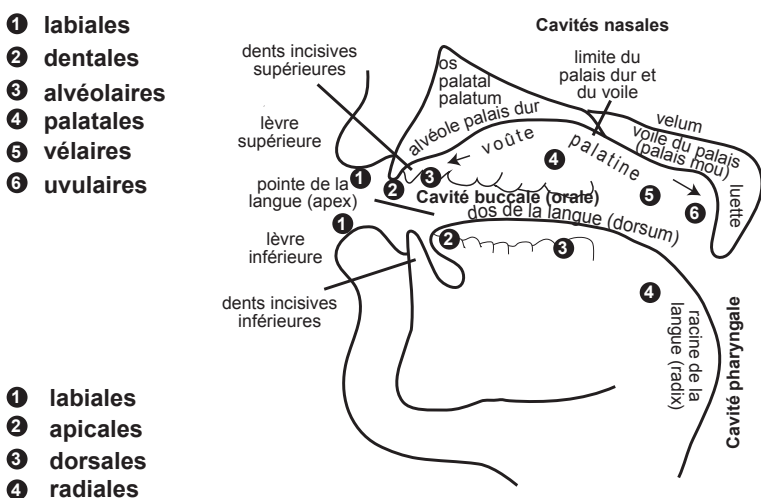
lation parfaitement distinctive). En français, à moins que le locuteur souffre d'une défaillance articulatoire (par exemple à la suite d'une déformation du palais ou dans la période intermédiaire entre la chute des dents de lait et la croissance des dents définitives), la confusion entre /f/ et /s/ est improbable, car les espaces flous d'articulation des deux consonnes fricatives n'ont pas de partie commune. En revanche, en anglais /f/ et /s/ peuvent se confondre avec /θ/ si l'espace flou d'articulation de chacune des trois consonnes est aussi étendu qu'en français pour /f/ et /s/ en raison de superpositions entre ces espaces. Ce que représentent les cercles discontinus de la partie droite de la figure, ce sont donc les espaces flous de /f/, /θ/ et /s/, non pas dans l'usage des anglophones, mais dans l'articulation et l'audition d'un francophone débutant dans l'apprentissage de l'anglais. Son entraînement consistera essentiellement à réduire l'espace flou de chacune des trois consonnes jusqu'à ce qu'ils ne se recouvrent nulle part, évitant ainsi les confusions d'abord auditives et ensuite articulatoires.



Les phonéticiens identifient une succession de points d'articulation des sons linguistiques. Si le flux d'air expiré est interrompu, la consonne est dite occlusive (ou plosive), s'il est étroitement canalisé, la consonne est dite fricative, s'il l'est moins étroitement, on est en présence d'une semi-voyelle (/w/, /j/ en français comme dans *oie* /wa/ et *ail* /aj/). Enfin si le canal est largement ouvert, c'est une voyelle qui est produite.

La figure¹¹ ci-dessous, qui représente la vue du côté droit de la partie orale de l'espace phonatoire, identifie six points d'articulation supérieurs concernant la lèvre supérieure, la dentition supérieure et l'espace alvéolaire, derrière les dents, la voûte du palais, son voile et son extrémité postérieure, la luette, allant des propriétés labiales à gauche aux uvulaires (formés au niveau de la luette) à droite, et quatre points d'articulation inférieure concernant la lèvre inférieure, la dentition inférieure et les trois composantes de la langue, son extrémité (l'apex), son dos et sa racine, ce qui distingue des localisations labiales, apicales, dorsales et radiales.

Les points d'articulation



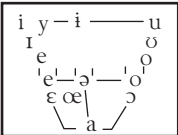
Pour les consonnes postérieures, on se contente de noter l'articulation supérieure, car la mobilité de la langue est limitée : une articulation palatale implique le dos de la langue et une articulation uvulaire implique sa racine. Pour les

consonnes antérieures en revanche, on note les deux points d'articulation supérieure et inférieure. Ainsi, en français /p/ et /b/ sont des consonnes bilabiales (interrompant le flux de l'air expiré entre les deux lèvres) et les consonnes fricatives /f/ et /v/ sont labio-dentales, c'est-à-dire que le flux de l'air est canalisé entre la lèvre inférieure et les dents supérieures, avec vibration des cordes vocales pour /v/ et sans vibration pour /f/.

Question : découle-t-il de cette multiplicité de points d'articulations et de traits distinctifs complémentaires que les systèmes de consonnes, voyelles et semi-voyelles des langues du monde présentent une variété infinie de composants et de combinaisons de ces composants ? Il est vrai que l'alphabet phonétique international répertorie une foule de sons linguistiques aptes à distinguer les signes linguistiques à travers la variété des langues. Cependant des chercheurs ont pu mettre en évidence certaines propriétés universelles de ces systèmes. Le groupe de recherche « Communication parlée » du CNRS basé à Grenoble a notamment dégagé les propriétés de répartition des voyelles dans les systèmes vocaliques de plusieurs centaines de langues modernes. Il est à noter que bien que l'espace de production des voyelles soit représenté sous la forme d'un trapèze renversé, dans tous les systèmes représentés, seule la position centrale de la base est occupée. Le /a/ y occupe donc toujours la position centrale, c'est la voyelle la plus ouverte.

Comme la forme de base d'où dérivent tous les autres systèmes est un triangle isocèle disposé sur la pointe, on constate que le premier système dans toutes les configurations – à l'exception de la dernière à neuf voyelles – regroupe une proportion plus importante de langues. Les systèmes à cinq voyelles représentent le meilleur équilibre avec 134 langues et parmi celles-ci le premier système en regroupe 65,7%, soit 88 langues sur un total de 451 (environ 1/5).

Les chercheurs de Grenoble ont donc raison de parler de tendances universelles : il va de soi que toutes les langues du monde ne peuvent pas avoir convergé vers un seul système vocalique, mais elles l'ont clairement fait selon un même principe génératif constitué d'un triangle vocalique de base occupant tout l'espace : /i/ (voyelle fermée antérieure), /u/ (voyelle fermée postérieure) et /a/ (voyelle centrale ouverte) et d'un enrichissement interne affectant symétriquement la partie antérieure (entre /i/ et /a/) et la partie postérieure (entre /u/ et /a/)¹².

nombre de voyelles nombre de langues			Les propriétés des différents systèmes vocaliques
3	20		C'est le système vocalique le plus simple avec trois voyelles /i/ - /a/ - /u/ réparties aux extrémités de l'espace disponible et donc parfaitement symétrique entre la partie antérieure et la partie postérieure.
4	32		Ce système à quatre voyelles présente une organisation intermédiaire entre le précédent et le suivant avec une asymétrie au profit de la partie antérieure (le /e/ fermé).
5	134		Dans le principal système à cinq voyelles, l'équilibre est rétabli avec le /o/ occupant la position symétrique du /e/ dans la partie postérieure.

12. Le tableau ci-dessous est un fragment du tableau 5 de N. Vallée, L.J. Boë et M. Stefanuto (1999) « Typologies phonologiques et tendances universelles. Approche substantialiste » <https://linx.revues.org/863>

Le chiffre en bas à droite de chaque figure indique la proportion des langues corpus répartissant les trois à neuf voyelles de cette manière.

6	72		30.6	Le premier système à six voyelles est analogue au système précédent avec l'ajout du <i>e</i> muet, le schwa ¹ /ə/.
7	49		40.8	Le principal système à sept voyelles est parfaitement symétrique avec l'ajout du <i>e</i> ouvert /ɛ/ à l'avant et du o ouvert /ɔ/ à l'arrière.
8	34		17.6	Le premier système à huit voyelles résulte de la combinaison des deux systèmes précédents : il a la même organisation de base que le système à sept voyelles en y ajoutant le schwa /ə/ du système à six voyelles.
9	35		20	Enfin le premier système à neuf voyelles enrichit à nouveau les deux parties avant et arrière par l'ajout de deux variantes moins fermées respectivement du /i/ à l'avant et du /u/ à l'arrière.

1. Schwa, en hébreu, désigne le vide. En linguistique, il désigne une voyelle neutre, qui peut, à l'occasion, être négligée (comme dans « sam'di »).

Les mouvements cycliques de la mâchoire ont produit des syllabes

Dans les années 1960, avec l'ouvrage de référence de Noam Chomsky et Morris Halle sur les patrons phonologiques de l'anglais¹³, la phonologie, dédiée aux différents systèmes organisant les sons linguistiques sélectionnés par chaque langue, a pris un caractère formel, calculatoire et dénué de perspective évolutionnaire. À partir des années 1990, une nouvelle génération de phonologues investis dans les sciences cognitives,

13. N. Chomsky & M. Halle (1968), *op.cit.*, note 6.

notamment Michael Studdert-Kennedy et Peter MacNeilage, ont plaidé au contraire en faveur d'une réorientation évolutionnaire de la phonologie. Cette entreprise, qui vise à réintégrer la phonologie dans le champ des sciences cognitives, se dit conforme à « l'intuition normale que la parole a évolué du simple au complexe au lieu d'avoir développé instantanément la forme mentale qui est aujourd'hui à sa base » (MacNeilage, critiquant Chomsky & Halle et l'école dite de « phonologie générative »).

Ces chercheurs entendent démontrer que la syllabe a évolué à partir de contraintes affectant les mouvements de la mâchoire dans les activités primaires liées à l'absorption des aliments : la mastication, la succion et le léchage. Les consonnes et les voyelles sont nées de « l'alternance rythmique générée par les mouvements d'élévation et de dépression de la mâchoire ». Le babil du jeune enfant est constitué d'alternances entre l'ouverture et la fermeture de la bouche, lesquelles produisent des articulations qui préfigurent les syllabes. Les chercheurs ont observé que l'acquisition de la parole passe successivement par deux stades de développement.

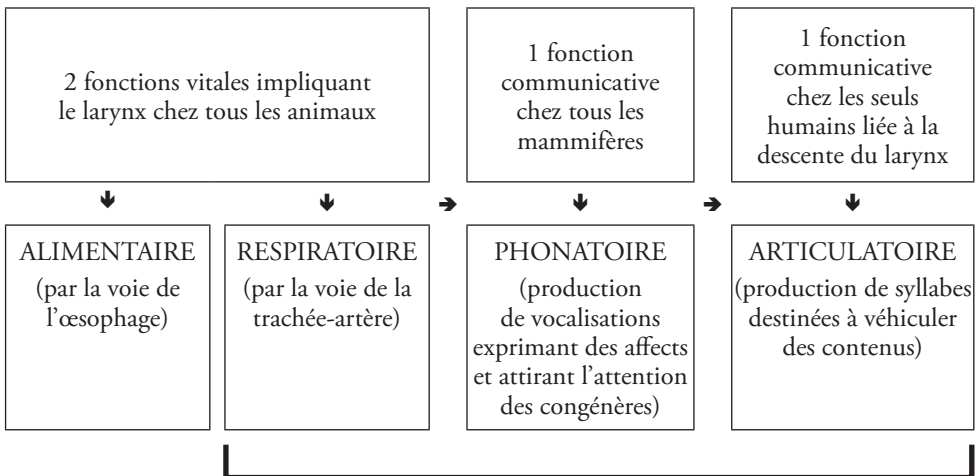
Le premier est celui des « cadres » (*frame stage*) : l'apprenti locuteur cadre le format syllabique des futurs sons linguistiques (prioritairement CV : consonne-voyelle).

Le second est celui de l'association des cadres à des contenus (*frame-content stage*) : l'inventaire des sons s'accroît, l'enfant acquiert la capacité d'enchaîner les syllabes et les premiers mots sont créés.

Ce faisant, conformément à l'hypothèse défendue par le biologiste allemand Ernst Haeckel en 1867, le jeune enfant semble récapituler la genèse des syllabes au cours de l'évolution de l'espèce. Dans sa forme actuelle, cette hypothèse est au moins partiellement innéiste : le mécanisme à la base de la production des syllabes ferait partie de notre patrimoine géné-

tique, car les patterns de formation des syllabes observés dans le babil se révèlent plus universels que les formats syllabiques rencontrés dans les langues et ils ne peuvent donc pas être appris à partir de l'environnement.

Quatre fonctions de l'espace oral se sont successivement cumulées au fil de l'évolution du vivant. Les deux premières sont la fonction d'absorption des aliments et celle de respiration. Chez les mammifères, il s'y ajoute la fonction phonatoire, concernant la production de vocalisations émotives et informatives. Jusqu'aux primates supérieurs, le larynx est structuré de manière à empêcher les aliments de prendre la direction des bronches, mais la descente du larynx – caractéristique de l'espèce humaine, mais qui ne touche le jeune enfant qu'à partir de la phase du babil – a permis l'articulation des sons linguistiques par un processus dit d'exaptation* (l'espace de la bouche, du pharynx et du larynx est exploité pour l'exercice de la fonction surajoutée), dont l'un des effets délétères est



Les trois systèmes de contrôles moteurs à l'origine de la création des premières syllabes

le risque nouveau de s'étrangler, c'est-à-dire de laisser passer des bribes d'aliments dans la trachée-artère. Le schéma page précédente résume cette évolution. Les quatre fonctions représentées sont générées par des systèmes cérébraux de contrôle moteur et les trois systèmes de contrôle respiratoire, phonatoire et articulaire représentent trois stades évolutifs qui ont permis la création des premières syllabes.

Les premiers hommes faisant usage de la parole ont donc réussi à combiner au mieux les trois fonctions respiratoire, phonatoire et articulaire. Le facteur constant est resté une propriété biomécanique fondamentale, l'inertie, laquelle a modelé les relations entre consonnes et voyelles constitutives de la syllabe. Le système articulaire contrôle l'aptitude de la mâchoire à produire des cycles alternant les voyelles et les consonnes. MacNeillage considère que ce système est aussi ancien que le système phonatoire (de production de vocalisations) mais qu'il n'est devenu intéressant pour la communication que lorsque nos ancêtres primates ont développé la communication visuofaciale, soit la production et la compréhension des mimiques.

La première exploitation de ce système qui se destinait spécifiquement à la parole a consisté à regrouper le système articulaire et les deux autres systèmes préalables pour obtenir les protosyllabes. Le cycle universel consonne-voyelle est constitué d'une phase de constriction de faible amplitude (une consonne) et d'une phase de haute amplitude dénuée de constriction (une voyelle). Il est présent dans toutes les langues et domine le processus d'acquisition de la parole dès son apparition.

À l'origine (dans le développement linguistique des locuteurs en devenir et par hypothèse dans l'évolution humaine) la syllabe, base du système de codage de la parole, a pu être étendue dans le temps et donc fortifiée par son redoublement.