

Principales hypothèses sur l'origine et la diffusion des langues

Bibliographie

Ruhlen M. (1997), L'origine des langues, Paris, Belin.

Pinker S. (1999), L'instinct du langage, Paris, Odile Jacob.

Introduction

L'origine des langues est une des questions les plus litigieuses qui a donné lieu à plusieurs discussions et réponses contradictoires. Le présent travail présente des hypothèses sur l'origine du langage humain pour en venir ensuite à l'origine et dispersion des familles des langues et des langues séparées, en présentant les données linguistiques à côté d'autres données, notamment génétiques, pour dresser l'arbre généalogique de l'origine et de l'expansion d'Homo sapiens.

Les langues ont-elles la même origine ou non ? Parmi les diverses versions une des plus importantes est celle de Chomsky. Sa théorie basée sur les universaux linguistiques innés a été révolutionnaire en corroborant l'hypothèse sur l'unicité et l'origine commune des langues. Depuis de nombreux linguistes, parmi lesquels Greenberg, ont établi des racines et des propriétés grammaticales, phonologiques et sémantiques communes dans les langues du monde entier qui ne varient pas librement.

Cela implique-t-il que les langues subissent la contrainte de la structure du cerveau ? La question de l'origine du langage humain en général (l'évolution de la capacité de parler chez nos ancêtres biologiques) intervient ici. Les réponses traitent plusieurs aspects de cette question et le plus souvent ne se contredisent pas, comme, par exemple, les scénarios biologique ou culturel avec les thèses de la monogénèse (restreinte ou large) ou de la polygénèse.

Origine du langage humain

Polygénèse : émergence simultanée du langage à des endroits différents.

Monogénèse : Le langage serait né une fois, toutes les langues existantes descendant de cette protolangue et en conservant certaines caractéristiques. Mais cela ne suffit pas pour expliquer un autre aspect du langage, les universaux, car il pourrait y avoir des ruptures dans la transmission du langage à travers les générations. En outre, on ne peut pas toujours voir des corrélations entre les différences linguistiques.

Une thèse alternative propose la version que les langues sont le reflet d'universaux de pensée (traitement d'information) qui ne sont pas spécifiques au langage.

On distingue monogénèse large qui signifie que toutes les langues existantes et disparues ont été des formes altérées de la langue unique originelle, et monogénèse restreinte, selon laquelle il peut avoir existé des langues apparues indépendamment puis disparues et qui seraient parlées par d'autres hominidés.

Selon le modèle de monogénèse , dit «africain», l'espèce Homo sapiens est née en Afrique il y a cent mille ans et a émigré en Eurasie en remplaçant d'autres hominidés. Le second modèle de polygénèse, modèle d'«évolution multirégionale» ne considère pas l'Afrique comme le berceau d'Homo sapiens mais soutient que cette espèce a évolué à partir des populations d'Homo erectus en Eurasie indépendamment dans les régions et à des époques différentes.

Le scénario culturel rattache l'origine du langage à l'expansion culturelle, « l'explosion sapiens », considérant le langage comme un artefact culturel développé au cours des cent mille dernières années où la culture humaine est devenue plus complexe. Certains proposent que cet avantage des ancêtres humains serait la cause de la disparition des Néandertaliens.

Le scénario biologique, avec ses versions variées, considère le langage comme le résultat d'une longue évolution des hominidés.

Une version proposée par certains psychologues est que les modifications des organes vocaux et de la circuiterie neuronale sont les seuls aspects du langage qui ont évolué dans notre espèce. Cette théorie suppose qu'il existe quelques aptitudes

d'apprentissage, générales dans le règne animal, qui fonctionnent avec une grande efficacité chez l'être humain. D'où les nombreuses expériences d'apprendre les singes à parler comme les humains, qui ont tous échoué.

La théorie de l'origine gestuelle, selon laquelle le langage aurait évolué à partir des gestes accompagnés de grognements et d'autres sons, s'ajoute à la précédente.

La version que Pinker propose en éliminant ces thèses est d'attribuer les universaux linguistiques à un instinct universel du langage et ne pas les réduire à l'histoire, ou à la cognition : il y a un seul et même plan au-dessous de la surface des langues qui implique l'universalité des principes et des propriétés de base, à partir duquel les paramètres modifient l'aspect superficiel d'une langue.

Pinker affirme, comme Chomsky et d'autres linguistes, que « cet instinct du langage uniquement humain est incompatible avec la théorie darwinienne moderne du langage, dans laquelle des systèmes biologiques complexes apparaissent par l'accumulation progressive sur des générations de mutations génétiques aléatoires favorisant une meilleure reproduction », mais que cet instinct, comme tous les autres, a évolué en raison de la sélection naturelle comme Darwin l'a proposé. Ce n'est pas non plus un produit culturel qui s'apprend mais une adaptation ou instinct biologique, ou comme s'exprime Darwin, « une tendance instinctive à acquérir un art (celui du langage) »

L'affirmation que le langage humain est extrêmement différent de la communication animale et est unique dans le règne animal semble être en contradiction avec celle de Darwin sur le caractère graduel des modifications. En fait la bonne théorie de l'évolution montre que l'homme et le singe ont évolué d'un ancêtre commun éteint déjà, et le langage serait apparu après la séparation des hommes et des singes.

Les raisons de diversité et parenté des langues

Ressemblances entre diverses langues- Depuis la découverte du sanskrit en 1786 par W. Jones de nombreuses études ont été entreprises pour trouver des ressemblances dans les vocabulaires des langues afin de les classer et montrer leur parenté. La classification des langues à partir des ressemblances fonctionnent pour la raison que l'association de la forme d'un mot et de son sens est arbitraire, comme l'a montré Saussure, ce qui implique qu'il est peu probable que les nombreuses ressemblances des mots soient par hasard. Ruhlen parle de trois principes qui expliquent l'existence de formes semblables ayant le même sens dans diverses langues : la convergence (les mots originellement différents se ressemblent accidentellement), l'emprunt (il s'agit des mots hors du vocabulaire de base, empruntés en même temps que la chose elle-même, souvent liés aux événements historiques) et l'origine commune.

Différences entre les langues : Une autre question qui se pose concerne la diversité linguistique : si le plan de base du langage est inné et fixe dans toute espèce, pourquoi l'ensemble ne l'est-il pas ? Comme on l'a déjà vu, la réponse de Darwin est, comme pour la formation des espèces distinctes, le développement graduel.

Pinker parle de trois facteurs qui ont pour effet la différenciation des langues : la variation ou l'innovation linguistique (il y en a de nombreuses sources, comme les mots forgés, les changements sémantiques, la réanalyse, etc), l'hérédité ou l'aptitude à apprendre (Un trait étant au départ un produit d'apprentissage peut devenir inné à cause de la pression sélective si l'environnement est stable, comme c'est le cas pour beaucoup de propriétés et mécanismes linguistiques. Pourtant le besoin et l'avantage d'apprendre les parties variables d'une langue, comme le vocabulaire ou la grammaire, afin de pouvoir partager le même code, restent toujours parce que les génomes humains changent) et l'isolement par la migration et les barrières sociales dont le résultat est, par exemple, le fait que les innovations s'accumulent séparément. Ainsi, beaucoup de prétendus américanismes ont été en fait apportés d'Angleterre, où ils ont disparu ensuite.

Evolution des langues - Les facteurs proposés par Ruhlen pour expliquer l'évolution des langues sont : recherche de la facilité d'articulation (tendance à simplifier les

sons), l'amour envers un certain degré de changement, surtout chez les jeunes (nouveaux mots et tournures empruntés, p. ex : « coul », en anglais - « dog » a remplacé « hound »).

Origine des langues et diffusion des familles de langues

Ces trois aspects des langues amènent-ils à penser que les langues s'évaluent à partir d'une protolangue ?

Jusqu'à une certaine époque l'indo-européen était considéré comme la forme la plus évoluée du langage, les autres langues étant des stades primitifs.

La façon dont les langues contemporaines évoluent nous aide à remonter loin dans le passé et à établir leurs origines. Les travaux les plus exhaustifs réalisés dans ce domaine appartiennent à Greenberg et M. Ruhlen, qui ont montré que les frontières linguistiques transcendent souvent les frontières géographiques *ou politiques* et ont classé les langues en familles et ces familles en des groupes plus grands jusqu'à monter vers une proto-langue originelle et commune pour toutes les langues.

La comparaison entre des familles a permis à Greenberg et à Ruhlen de présenter les groupes ou familles suivantes pour les langues du monde entier : khoisan, nigéro-kordofanien, nilo-saharien, afro-asiatique, dravidien, kartvélien, eurasiatique, déné-caucasien, austro-asiatique, indo-pacifique, australien, amérindien.

L'Afrique présente peut-être la plus grande diversité linguistique à la surface du monde, mais toutes les **langues africaines** appartiennent en tout, selon Greenberg, à 4 familles : khoisan, nigéro-kordofanien (branches kordofanien et nigéro-congolais), nilo-saharien et afro-asiatique. Cette classification dépassant le plafond de six mille ans a révolutionné la linguistique africaine, mais a contribué à expliquer la répartition à travers l'Afrique des gènes humains et des découvertes archéologiques.

Le nilo-saharien est beaucoup plus hétérogène et par suite plus ancien que l'indo-européen. Il a été proposé par le linguiste Gregersen d'unir cette famille et le nigéro-kordofanien dans une famille congo-saharienne. Cette hypothèse n'a pas été

confirmée par les linguistes mais s'accorde avec la distribution des gènes à travers l'Afrique et pourrait être exacte, selon les théories génétiques.

Les langues khoisanes sont parlées en Afrique du Sud, mais certains indices montrent que cette famille était autrefois plus répandue. Greenberg a rattaché à cette famille 2 langues isolées parlées en Afrique de l'Est, le sandawe et le hadza qui utilisent des clicks particuliers à la famille khoisane mais d'une origine inconnue. En dépit de ce fait, Greenberg et Ruhlen ont trouvé des racines reliant cette famille aux autres langues du monde.

La famille nigéro-kordofanienne couvre une grande surface mais la majeure partie de ce territoire a été installée au cours des deux derniers millénaires à la suite de l'expansion bantoue. Les plupart des branches de son groupe nigéro-congolais se trouvent en Afrique de l'Ouest où il faut chercher le foyer originel de cette famille, conclusion qui est confirmée par la récente expansion bantoue.

La famille afro-asiatique inclut, entre autre, le tchadique, ce qui a provoqué de désaccords car les locuteurs de cette langues sont negroïdes, alors que les autres locuteurs de la familles sont caucasoïdes, mais les faits linguistiques reliant cette langue avec la famille afro-asiatique sont sans appel. Les éthiopiéens parlant ces langues sont génétiquement associés aux Africains subsahariens, ce que Cavalli-Sforza explique par le métissage très ancien entre des Africains et des Caucasoïdes

En Eurasie on trouve, entre autres, les familles eurasiatique et kartvélienne, la position de celle-ci n'est pas encore déterminée. Greenberg soutient qu'elle est apparentée à l'eurasiatique de manière lointaine, tandis que les nostraticistes l'incluent dans cette famille. La parenté d'eurasiatique, de kartvélien et de l'afro-asiatique est plus claire, mais le degré de leurs rapports est à découvrir.

Un autre groupe d'Eurasie, le dene-caucasien, est plus archaïque que l'eurasiatique et comprend des langues et familles bien divergentes et géographiquement isolées comme le basque, le caucasien, le burushaski, le iénisséen, le sino-tibétain et, en Amérique du Nord, le na-dene. Le linguiste Bengston a proposé d'unir le basque, le burushaski et le caucasien dans un sous-groupe. Cavalli-Sforza a montré que les basques ont des particularités génétiques par rapport aux autres populations

européennes et dans ses travaux les familles dene-caucasiennes divergent génétiquement (le tibétain s'intègre au groupe d'Asie du Nord, le chinois - au groupe austro-asiatique, le na-dene - au groupe amérindien). Ruhlen explique cela par le fait que la migration dene-caucasienne est plus ancienne que l'expansion eurasiatique et a été submergée par celle-ci en ne conservant que peu de langues dans des endroits isolés et moins accessibles. Il y a eu ensuite un courant d'échange génétique entre ces groupes isolés et les peuples indo-européens, mais les langues préservent une image plus nette de l'état antérieur que ne le font les gènes.

La famille austro-asiatique d'Asie du Sud-Est a 4 sous-familles : l'austro-asiatique, le miao-yao, le daïque et l'austro-nésien. Le foyer originel de ce dernier est probablement à Taiwan où les peuples austronésiens ont migré il y a 6000 ans. Elle comprend plus de mille langues et a le même âge que l'indo-européen.

Les langues non austronésiennes de Nouvelle-Guinée appartiennent à une famille plus ancienne, l'indo-pacifique, selon Greenberg.

Les langues aborigènes d'Australie font partie de la famille australienne, ancienne de 40 000 ans.

En Amérique Greenberg a proposé de distinguer 3 familles, ce qui a été confirmé par les données génétiques : l'eskimo-aléoute, le na-dene et l'amérindien qui sont le reflet de 3 vagues migratoires venues d'Asie : la première a amené le proto-amérindien dont dérivent les autres langues amérindiennes. Des généticiens dirigés par Douglas Wallace ont constaté que les populations na-denes et amérindiennes dérivent des populations asiatiques et sont entrées en Amérique séparément à des époques différentes.

Les travaux des généticiens jettent de la lumière sur la question de l'ordre dans lequel ces divers groupes ont divergé du tronc originel en relevant des corrélations entre les classifications des populations et des langues., comme par exemple, la répartition des peuples aborigènes d'Amérique dans les familles définies par Greenberg, ce qui n'était pas admis par plusieurs linguistes.

Les familles indo-européenne, afro-asiatique, dravidienne, altaïque et eskimo-aléoute et des langues koréenne et japonaise sont regroupées par Greenberg et Ruhlen dans un groupe provenant d'une protoprotolangue nostratique. Mais dans les travaux de

Cavalli-Sforza les branches asiatiques eurasiatiques sont plus proches aux familles américaines na-dene et amérinde.

Un autre groupe des généticiens dirigé par Allan Wilson a découvert que le plus profond clivage de la population humaine sépare les africains subsahariennes du reste du monde il y a environ deux cents mille ans, ce qui était prouvé aussi par Cavalli-Sforza par d'autres moyens.

En résumant les données linguistique, génétiques et archéologiques il semble que le modèle de monogénèse africain, où apparaît la plus grande diversité linguistique, est plus proche à la vérité. La diffusion des humains à travers la Planète se caractérise par 3 migrations essentielles dans lesquelles un rôle majeur joue le développement d'agriculture qui était encore inconnue il y a dix mille ans : 1) la dispersion des Bantous à travers l'Afrique australe, 2) l'expansion austronésienne (de Chine méridionale il y a \approx 8000 ans vers l'Indonésie et les îles Pacifiques), et 3) partie de l'Anatolie il y a \approx 8000 également qui serait l'origine des indo-européens, selon Colin Renfrew. La ramification initiale de cette migration était l'expansion dene-caucasienne en direction à la fois est et ouest du foyer proche-oriental, qui aurait peuplé l'Europe et aussi donné naissance en Asie à la famille sino-tibétaine. Ensuite la migration eurasiatique, originaire du Moyen-Orient, a recouvert cette famille en ne laissant subsister que le basque, le burushaski en Pakistan et le sino-tibétain. Les familles iénisséienne et na-dene sont liées plutôt aux mouvements secondaires que l'expansion dene-caucasienne originelle. Ensuite de cette population eurasiénne un groupe donne naissance à la famille afro-asiatique, et un autre pénètre en Amérique il y a \approx 12000 ans. Les populations eurasiatiques restées en Eurasie ont dispersé en donnant naissance aux familles indo-européen (fin du néolithique), ouralien, altaïque, tchouktchi-kamtchatkien et eskimo-aléoute en Amérique du Nord.

Geneve, 2008