



***Journées du Patrimoine 2003  
vendredi 19 septembre, 20h30***

*Bayonne  
Auditorium C.C.I.*

*Conférence :  
« Langue et peuplement : le cas de la langue basque »*

*Intervention de Jean-Marie HOMBERT. Professeur des Universités à Lyon II. Directeur depuis 1994 du laboratoire « Dynamique du langage » (Lyon II). Coordonnateur du programme C.N.R.S. « Origine de l'Homme, du Langage et des Langues ».*

**Langues et gènes : les Basques, un cas à part ?**

**1**

L'exercice auquel nous avons été conviés n'est pas de ceux qui ont la préférence des chercheurs : ils n'aiment pas expliquer de manière simple et brève les résultats de leurs recherches. Il se trouve, en outre, que les questions posées sont extrêmement compliquées et qu'elles suscitent un intérêt passionné comme en témoigne le fait que vous soyez si nombreux, réunis ce soir. C'est pourquoi nous nous sommes mis à cinq pour répondre à ces questions et j'espère que vous ne serez pas déçus dans vos attentes.

Notre objectif commun est de vous montrer comment et à l'aide de quelles méthodes les différentes disciplines que nous représentons ici s'attachent à reconstituer sur de très longues périodes l'histoire des langues et des peuples qui parlent ces langues.

Vous allez constater immédiatement que certaines disciplines sont mieux équipées que d'autres pour traiter de certaines périodes historiques et que les connaissances acquises peuvent être assez précises pour certaines périodes chronologiques mais beaucoup moins pour d'autres. Il s'ensuit qu'à certaines questions que vous vous posez, nous pouvons apporter des réponses relativement assurées, qu'à d'autres nous ne pouvons livrer que des éléments partiels de réponse et qu'enfin, pour la majorité des questions qui vous intéressent, nous ne disposons pas à ce jour de réponses.

Il s'agit moins pour nous, en effet, de prétendre apporter à tout prix des solutions aux problèmes dont il sera question ce soir que de vous montrer de quelle façon nos disciplines travaillent pour essayer d'en trouver et sur quelles bases scientifiques ils s'appuient aujourd'hui dans leurs raisonnements. Vous comprendrez donc mieux pourquoi certaines questions attendent toujours leurs réponses et aussi pourquoi certaines réponses sont, à l'heure actuelle, mieux garanties que d'autres. En linguistique, comme dans toutes les sciences, on procède par hypothèses que l'on soumet ensuite à vérification. Il ne suffit pas que des hypothèses soient séduisantes ou seulement plausibles ; il faut qu'elles soient passées au crible des faits dont on dispose et qu'elles soient en mesure de résister à toutes les tentatives de réfutation. Bref mieux vaut en savoir peu, mais de façon sûre, qu'avancer des solutions qui ne pourront jamais être validées soit par les linguistes eux-mêmes soit par des représentants d'autres disciplines, comme la génétique, par exemple.

## 2

Nous allons parler de langues, de gènes, de populations et de peuples. Vous allez vérifier que, dans certains cas, il y a convergence dans les données produites par la linguistique, par la génétique, par l'histoire du peuplement et par l'anthropologie sociale et que, dans d'autres cas, il y a divergence. Le but du jeu, ce soir, est de faire comprendre comment ces quatre ordres de réalité que sont langues, gènes, populations et peuples peuvent tantôt s'articuler étroitement et former système tantôt paraître se divorcer en s'éloignant les uns des autres.

En résumé, qu'allons-nous faire ? Pour certains d'entre nous, raconter, en prenant 10 minutes chacun, les 100.000 dernières années. On admettra que c'est l'histoire la plus condensée jamais entendue : 10.000 ans à la minute ! J'insiste d'abord sur le fait que nous allons essayer d'être bien compris de vous et c'est pourquoi nous nous sommes promis de ne pas jargonner. Je ne suis pas sûr que nous puissions y arriver tout à fait. Qu'on n'hésite donc surtout pas à poser des questions et qu'on ne se laisse pas aller au découragement si nous utilisons parfois un langage technique qui ne vous paraîtra pas très clair.

Il est évident aussi que nous allons essayer de traiter ces différentes questions délicates de la manière la moins polémique possible. Le but du jeu n'est pas de susciter des réactions, favorables ou défavorables ; il est de dresser un état de nos connaissances. Que savons-nous ? Qu'ignorons-nous ? Que pouvons-nous espérer savoir ? Dans ce but, je vais introduire un certain nombre de généralités et de concepts de base concernant les langues. Il en ressortira, au minimum, une réponse très claire à la question qui sert de titre général à cette série d'exposés : « Les Basques, un cas à part ? ». Cette réponse est négative. Le basque est à l'évidence une langue passionnante à étudier ; les Basques forment, non moins évidemment, une population dont l'histoire offre un intérêt majeur.

Pour autant, il s'agit d'une langue parmi les 6000 langues parlées dans le monde, qui a ses spécificités comme n'importe quelle autre langue, de la même façon que la population basque en est une parmi une infinité d'autres, anciennes et actuelles, avec ses particularités comme n'importe quelle autre population. S'agissant de la langue basque, quelles sont ces spécificités ? C'est à tenter de les mettre en perspective que je vais m'employer.

### 3

Commençons donc par cerner l'arrière-plan linguistique du problème suscité par la situation très particulière présentée par la langue basque. Une question fréquemment posée aux linguistes est celle-ci : « **Combien y a-t-il de langues parlées aujourd'hui dans le monde ?** ». La réponse apportée est la suivante: entre 3000 et 10000. Chacun peut constater l'imprécision de cette réponse. Elle peut étonner mais elle s'explique aisément. En premier lieu, nous manquons cruellement de données portant sur certaines aires linguistiques comme l'Amazonie ou la Nouvelle-Guinée. On parle un grand nombre de langues en Amazonie dont, pour certaines, nous ne connaissons à peu près rien. Quant à la Nouvelle Guinée qui est une île, l'hypothèse la plus couramment émise est qu'il y existe 800 langues différentes dont on est loin d'avoir fait l'inventaire. Ensuite les langues changent et se diversifient dans le temps. D'où la difficulté rencontrée pour en fixer le nombre. Une langue qui s'est beaucoup transformée est-elle toujours une et la même langue ou en est-elle devenue une autre ? Ce que nous savons, en revanche, c'est que si le nombre de langues change, il change dans le sens de la diminution. Les linguistes s'accordent pour estimer que, dans un siècle, 90 % des langues parlées aujourd'hui auront disparu. En d'autres termes, à supposer qu'il y ait 5000 langues parlées aujourd'hui dans le monde, l'on peut s'attendre à ce que, dans cent ans, elles ne soient plus que 500.

### 4

Une autre source d'incertitude, sur laquelle je souhaite insister, est la difficulté à proposer **une définition de ce qu'est une langue**. Comment distingue-t-on une langue d'un dialecte, un dialecte d'un patois ? Ce sont là des sujets qui ont alimenté d'interminables discussions entre linguistes et cela depuis fort longtemps. Voici quelques éléments de réponse simples à ces questions.

Comment différencie-t-on une langue d'un dialecte ? Les linguistes utilisent généralement *le critère d'intercompréhension*. On admet communément que le locuteur d'un dialecte comprend le locuteur de la langue source et qu'il est compris de ce dernier. Lorsqu'il y a intercompréhension entre deux individus,

on dit qu'ils parlent deux dialectes de la même langue. S'il n'y a pas intercompréhension, on dit qu'ils parlent deux langues différentes. Mais l'intercompréhension doit-elle être totale et doit-elle être immédiate ? Il est clair qu'il faut introduire une certaine graduation dans l'application du critère d'intercompréhension.

Prenons l'exemple du français et de l'italien. Le français et l'italien sont-ils deux dialectes de la même langue ou sont-ils deux langues différentes ? Si l'on mobilise le critère d'incompréhension, un locuteur français qui n'aurait jamais été mis au contact de l'italien comprend un certain pourcentage de ce qui est énoncé en italien. Quel est ce pourcentage ? 10%, 15% ? Combien en comprend-il après 8 jours d'exposition à l'italien ? 20%, 30% ? Après un mois ? Après trois mois ? On voit, par conséquent, que ce critère d'intercompréhension est un instrument dont l'emploi n'évite pas le vague ; il ne permet pas de faire vraiment la différence entre langue et dialecte...

Je vais vous donner un autre exemple de la difficulté à différencier une langue d'un dialecte. Il renvoie à ce phénomène que les linguistes appellent une chaîne dialectale. Soit un dialecte A ; il peut être compris par les locuteurs d'un dialecte voisin que l'on va désigner comme le dialecte B. A son tour, le dialecte B peut être compris par les locuteurs d'un dialecte C, etc. Il se passe donc qu'il peut y avoir intercompréhension entre une communauté de locuteurs et une seconde communauté de locuteurs, intercompréhension entre cette seconde communauté et une troisième sans que la première communauté de locuteurs soit nécessairement comprise par la troisième, celle qui parle le dialecte C. Qu'en conclure ? Avons-nous affaire, avec A et C, à deux langues différentes ou faut-il considérer A, B et C comme une succession de dialectes ? Répétons donc que la distinction entre langue et dialecte n'est pas simple à opérer. Une chose est sûre : il est vain de demander à un locuteur s'il parle le même dialecte que le locuteur d'une communauté voisine. En effet, la réponse qu'il va faire sera fortement influencée par des paramètres non linguistiques, de type «socio-linguistique» (existence d'une histoire vécue en commun, présence de croyances religieuses partagées, etc.). Il suffit de songer, à cet égard, à ce que pourraient dire les uns des autres, à propos de la langue qu'ils parlent, des informateurs serbes et des informateurs croates !

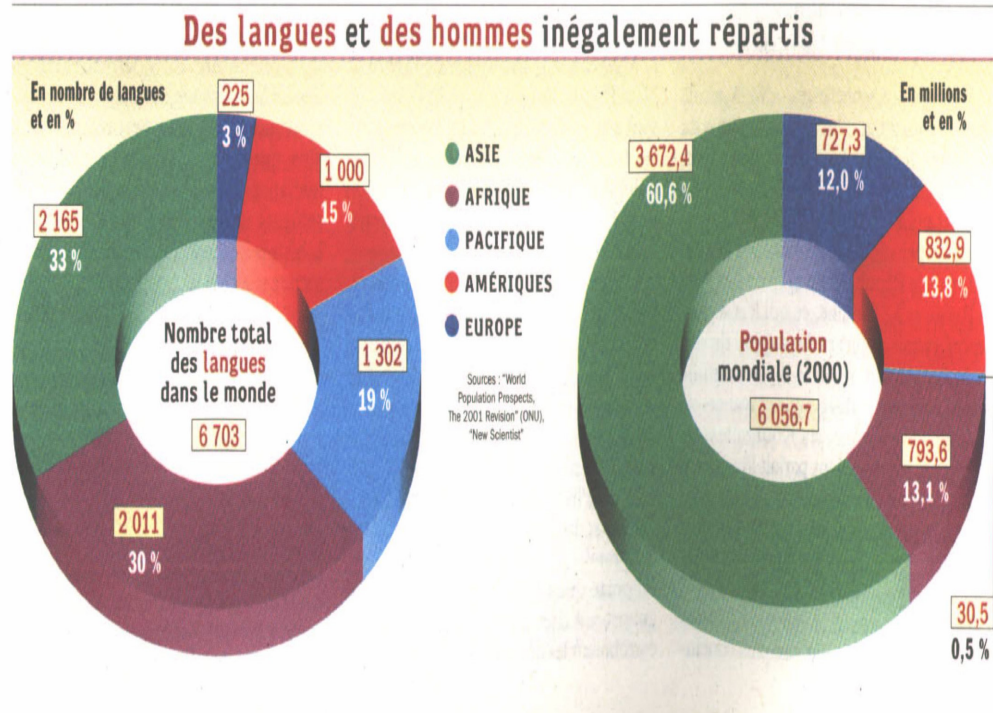
## 5

Si les linguistes sont bien en peine de dire exactement ce qu'est une langue, ils savent fort bien, en revanche, **ce qu'une langue n'est pas**.

- *Une langue n'a rien à voir avec un territoire national*. Affirmer qu'il existe entre 5000 ou 6000 langues parlées dans le monde ne présuppose

évidemment pas que l'on trouve dans le monde 5000 ou 6000 territoires nationaux distincts. Par conséquent, on peut très bien rencontrer une langue sans qu'à cette langue corresponde un territoire national clairement identifiable. Un espace linguistique n'est pas un espace national ; ses propriétés sont linguistiques et non politiques.

- *Une langue ne s'identifie absolument pas à un système d'écriture.* La plupart des langues parlées dans le monde aujourd'hui sont dépourvues de systèmes d'écriture. Elles n'en sont pas moins langues à part entière, c'est-à-dire structurées et organisées en langues. Et une langue n'est pas davantage « primitive » qu'une autre au prétexte qu'elle ne dispose pas d'un support d'écriture. Une langue qui s'écrit n'est pas moins « primitive » ou plus « avancée » qu'une langue seulement parlée. Il n'y a aucun lien entre le statut de langue et l'existence d'un système d'écriture.
- *Le statut de langue n'est pas davantage en rapport avec un nombre minimum de locuteurs.* Il existe aujourd'hui des langues qui sont parlées par moins de 10 locuteurs. Elles n'en sont pas moins des langues à stricte égalité avec celles parlées par plusieurs millions d'hommes et de femmes. (Et, évidemment, il n'y a aucun lien de quelque sorte que ce soit entre langue et niveau de développement technologique. Les Pygmées ont une langue et n'ont pas de fusée Ariane.) Voici qui m'amène à rappeler un point extrêmement important : il existe une très grande disparité entre les populations linguistiques (le nombre de locuteurs parlant une même langue). Juste deux chiffres pour frapper les esprits : plus de 50% de la population du globe parle l'une des 10 langues les plus parlées sur terre ; plus de 90% parle l'une des 100 langues les plus parlées. En d'autres termes, à supposer qu'il y ait 5000 langues, 10% de la population humaine parle l'une des 4900 autres langues restantes. Chacun peut donc constater qu'il y a un très grand nombre de langues qui ont très peu de locuteurs.



Il faut remarquer dans le graphique précédent la grande différence qui existe entre le nombre de langues parlées et l'importance des populations. On observera, par exemple, que l'Europe ne compte que 3% des langues parlées dans le monde alors que 12% de la population mondiale est européenne. On notera, dans le même temps, que l'Asie abrite 33% des langues répertoriées mais héberge 60% de la population actuelle du monde. Ce que montre ce graphique, c'est qu'il n'existe aucun lien entre le nombre des langues et la démographie des populations humaines.

Le graphique suivant présente le « hit parade » des langues en fonction du nombre des locuteurs. Le classement peut surprendre. C'est ainsi que peu de gens savent que le bengali est la quatrième langue parlée dans le monde tandis que le français se situe très bas dans ce classement, battu –si l'on peut dire – par le wu.

### **Hit parade (langue maternelle) - en millions de locuteurs-**

- |                    |  |
|--------------------|--|
| ■ Chinois (885)    | ■ Japonais (125)   |
| ■ Espagnol (332)   | ■ Allemand (98)  |
| ■ Anglais (322)    | ■ Wu (77)  |
| ■ Bengali (189)    | ■ Javanais (75)  |
| ■ Hindi-Urdu (182) | ■ Coréen (75)  |
| ■ (Arabe (200) )   | ■ Français (75)  |
| ■ Russe (170)      |  |
| ■ Portugais (170)  | ■ ( <a href="http://www.sil.org/ethnologue">www.sil.org/ethnologue</a> ) |

## 6

Telle est donc la situation actuelle. Maintenant il nous faut remonter dans le temps et insister sur un point extrêmement important pour la bonne compréhension des exposés qui vont suivre : **les familles linguistiques.**

Les langues du monde encourent un processus constant de diversification interne et de transformation tout au long de leur histoire. Il n'existe aucune langue qui reste identique à elle-même dans le temps. Toutes les langues changent, quoique puissent décréter d'en haut académies et grammairiens là où ils existent, qu'elles disposent ou non d'un système d'écriture.

D'où il en résulte qu'une communauté linguistique, unifiée à un moment donné de son existence, va peu à peu développer des dialectes distincts. Et, si l'on se place à une échelle de temps suffisante pour permettre à ce processus d'aboutir à son terme, on assistera à l'émergence de langues différentes bien qu'issues de la même souche linguistique. Chacun des exposés qui vont suivre va insister sur ce point.

L'objectif du linguiste spécialisé en linguistique historique est de retracer l'histoire des familles linguistiques. Prenons un exemple extrêmement simple.

Le latin était une langue relativement unifiée ; il y avait seulement des dialectes. Deux mille ans plus tard, on se trouve face à l'espagnol, au portugais, au français, à l'italien, au roumain. Toutes ces langues sont des langues dérivées du latin et dérivées du latin selon des modalités différentes. Deux mille ans plus tard, la situation est donc celle de langues pour lesquelles il n'y a quasiment plus d'intercompréhension. Voilà ce que produisent « seulement » 2000 ans d'histoire ! Evidemment au fur et à mesure que le temps passe, le processus de diversification augmente. Toutes les langues s'inscrivent, par conséquent, dans une certaine histoire linguistique au long terme et, lorsqu'il y a plusieurs langues d'une même famille à un instant donné, aujourd'hui par exemple, on essaie de reconstruire l'histoire de cette famille et de regrouper les langues en une seule langue qu'on appelle la protolangue, la langue-mère.

C'est ce qu'on fait avec les langues dérivées du latin ; c'est ce à quoi s'appliquent plus généralement les linguistes avec la quasi totalité des langues parlées en Europe occidentale et centrale, de famille indo-européenne, famille à laquelle n'appartient pas le basque, et c'est là, disons-le immédiatement, tout l'intérêt de cette langue.

Les linguistes emploient des méthodes qu'on appelle les méthodes de linguistique comparative, de linguistique historique, fondées sur les évolutions linguistiques et la détection de correspondances entre langues évoluant. Dans le cadre de ces évolutions, les changements ne sont pas aléatoires. On peut observer des changements réguliers, c'est à dire des changements qui obéissent à des règles. Le travail du linguiste consiste à repérer ces changements réguliers, à détecter des correspondances, de manière à reconstruire les étapes anciennes des langues, même en cas d'absence de tout document écrit, ce qui est le cas pour la plupart des langues du monde. Reconstruire ces étapes antérieures, c'est un métier qui s'apprend à l'université, dans le cadre des études de linguistique.

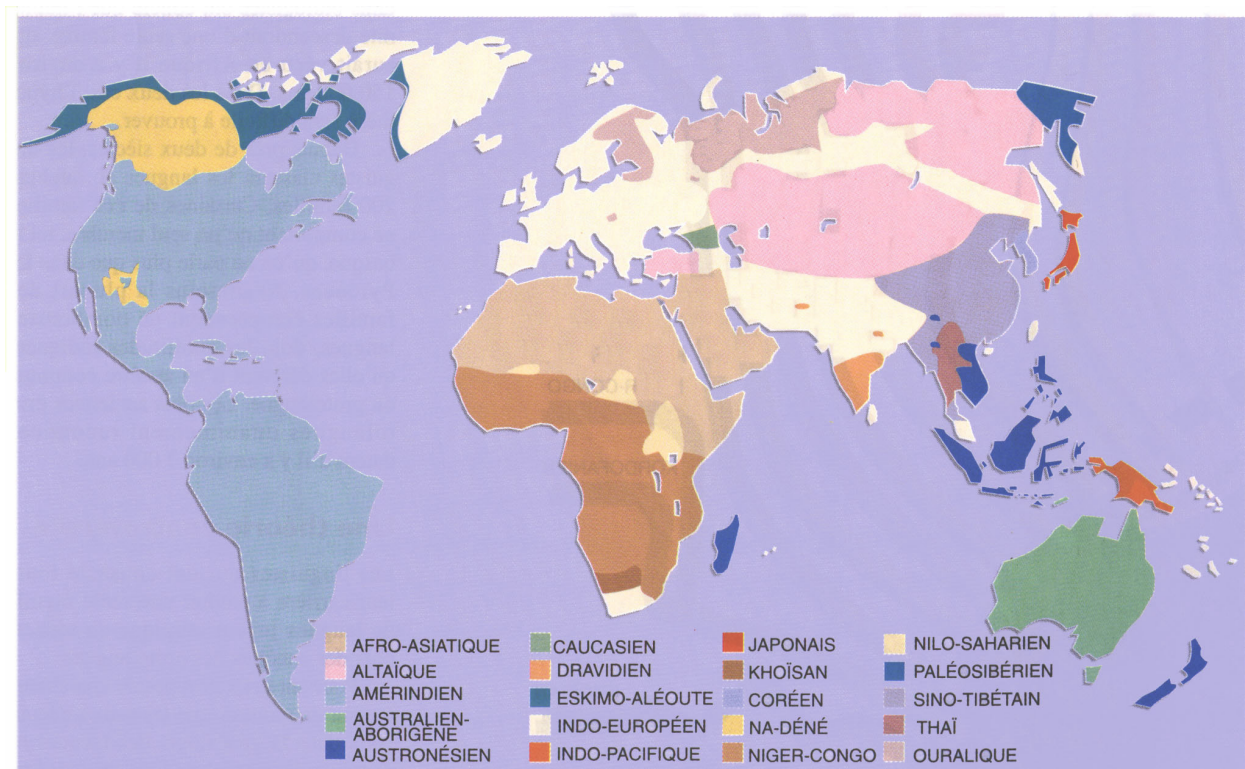
Puisque la diversification linguistique est fonction du temps qui passe, le nombre de familles reconstruites dépend évidemment de la profondeur historique. Par exemple, si vous avez déjà entendu parler un québécois français, vous remarquez que son français est différent du français parlé en France. Cette différence correspond à une divergence linguistique datant d'environ 5 siècles. Comparons maintenant l'italien et le français ; la différence constatée correspond dans ce cas précis à une divergence linguistique remontant à environ 20 siècles. En procédant sur ces bases, si l'on remonte à une profondeur historique d'environ 2000 ans, les 6000 langues parlées aujourd'hui peuvent être regroupées en approximativement 300 familles différentes. Si l'on remonte à une profondeur historique d'environ 5000 ans, les 6000 langues parlées aujourd'hui correspondent à environ 50 familles. Si on remonte à 10000 ans et plus, on en arrive à une douzaine de familles. Il est évident que plus on se place en amont dans le temps, plus les données utilisées pour trancher sont délicates à



manier et moins il y a consensus parmi les chercheurs qui se livrent à ce genre de travaux.

Pour être plus précis, les spécialistes de linguistique historique seraient globalement d'accord en ce qui concerne la première phase (l'hypothèse de 300 groupes linguistiques, il y a 2000 ans, rencontre un certain consensus) ; en revanche, des divergences s'expriment pour l'étape située à 5000 ans, et un désaccord massif se fait jour dans la communauté scientifique dès lors qu'on évoque l'étape située à 10 000 ans et au delà. Par ailleurs, la connaissance que l'on peut avoir de ces différents groupements de langues est très inégale. Nous en savons bien plus sur certains que sur d'autres ; ce qui biaise sans doute les opérations de traitement des données.

### Répartition géographique des familles linguistiques



**2. LES FAMILLES DE LANGUES VIVANTES DANS LE MONDE.** Cette carte représente la classification de la plupart des familles de langues, qui sont généralement reconnues par tous les linguistes. Cependant, certaines familles, comme celle des langues amérindiennes, sont controversées : on a représenté les langues amérindiennes partout où les langues indo-européennes récemment adoptées, comme l'anglais, le français ou le portugais, ont provoqué la disparition d'une famille de langues.

Cette carte, difficile à lire, est destinée à montrer la répartition d'une vingtaine de familles linguistiques mais elle illustre aussi la difficulté rencontrée à distribuer certaines familles linguistiques dans le temps et dans l'espace. Le cas, par exemple, de la famille des langues amérindiennes reste controversé ; c'est ainsi que, sur cette carte, on a implanté des langues amérindiennes partout où les

langues indo-européennes, récemment adoptées, comme l'anglais, le français, l'espagnol ou le portugais, ont provoqué la disparition d'une famille de langues : ces langues disparues sont postulées appartenir à la famille des langues amérindiennes. On observera néanmoins que ces familles sont à peu près également réparties sur l'ensemble de la planète. La carte permet, par exemple, de remarquer que, sur le continent africain, il existe 4 familles différentes tandis qu'en Europe on trouve une seule grande famille, la famille indo-européenne. La seule langue qui ne fait pas partie de cette famille d'Europe occidentale est le basque.

Enfin la recherche sur les familles linguistiques se heurte à deux difficultés majeures : l'existence de langues dites isolées (comme, par exemple, le basque), et d'une infinité de langues qui ont disparu sans laisser de descendance (ce qu'on appelle les langues mortes).

## 7

Qu'appelle-t-on une **langue isolée** ?

Il y a deux possibilités lorsqu'on travaille dans une zone déterminée et que l'on rencontre une langue qui n'est en aucune façon apparentée linguistiquement avec les langues voisines :

- il peut se faire que l'arrivée de cette langue dans cette zone géographique soit extrêmement récente. Il y avait auparavant un groupe humain unifié du point de vue linguistique. Un autre groupe arrive de l'extérieur ; il s'installe au beau milieu de cette zone géographique. Il en résulte que la langue parlée par ce groupe présente une apparence de langue isolée, mais si le linguiste élargit son enquête à l'échelle d'une aire linguistique beaucoup plus étendue, il constatera sans doute que cette langue entretient une parenté avec une ou des langues parlées ailleurs, plus loin.
- il existe une autre possibilité, représentant un cas de figure beaucoup plus intéressant pour le linguiste : c'est lorsque cette langue se trouve être la dernière descendante d'une famille de langues. Impossible de l'apparenter avec certitude à aucune autre langue. Dans ce cas, une langue isolée n'est assurément pas une langue « orpheline », car elle s'inscrit à coup sûr dans la filiation d'une langue mère, mais c'est une langue dont on ne connaît aucune « parente », même lointaine. Elle est, si l'on peut dire, seule au monde « de sa famille ».

## 8

Venons-en au cas de la langue basque. C'est précisément une langue isolée en ce sens qu'il n'est pas possible à ce jour de l'apparenter à une autre langue. Ce n'est pas faute d'efforts ! Depuis plusieurs décennies, les linguistes tentent de rattacher le basque à des familles linguistiques différentes : langues du Caucase, langues amérindiennes d'Amérique du nord, langues sino-tibétaines, langues na-déné, langues d'Afrique (langues niger-congo, langues khoisan du sud de l'Afrique), langues isolées ou disparues comme l'étrusque, le picte ou le burushaski. On pourrait presque dire qu'on a essayé de relier le basque à toutes les familles linguistiques.

Aucune de ces entreprises comparatives n'est convaincante. La conclusion à en tirer est celle-ci : on ne parvient pas à trouver, où que ce soit sur la planète, une langue qui soit linguistiquement apparentée au basque ; la langue basque est la dernière descendante d'une famille linguistique dont l'identité n'a pas encore été trouvée. En d'autres termes encore, aucune langue parlée aujourd'hui ne présente un ancêtre linguistique commun avec la langue basque **dont l'existence serait postérieure aux huit derniers millénaires.**

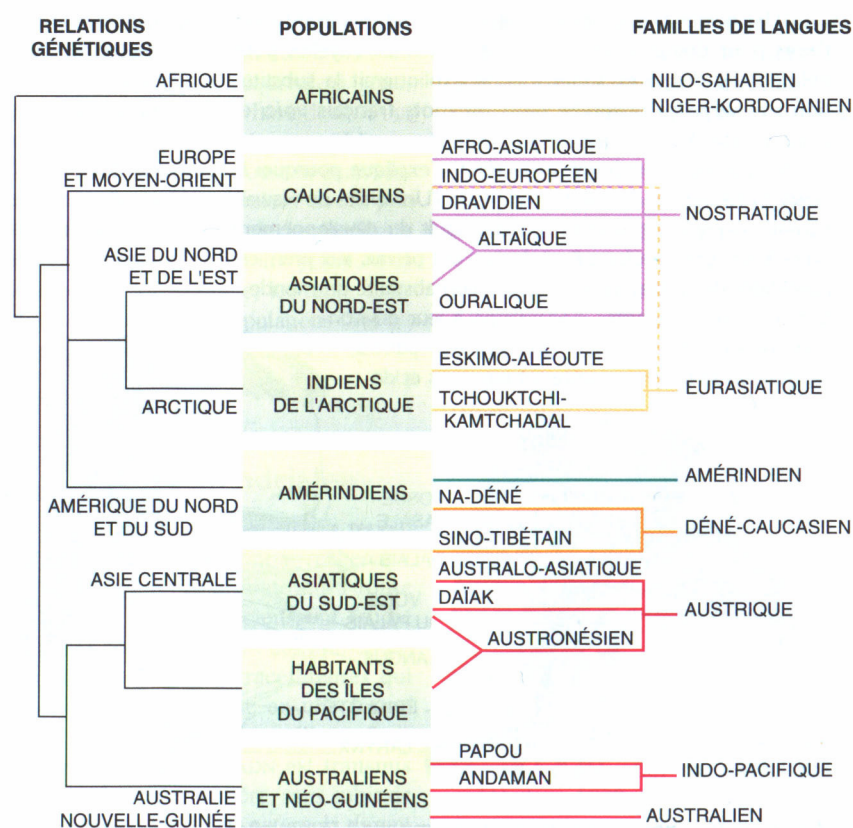
La précision est d'importance. Elle renvoie à ce fait que la méthodologie linguistique, telle qu'elle a été développée et appliquée jusqu'à aujourd'hui, nous permet d'obtenir des résultats relativement fiables jusqu'à moins 8000 ans. En utilisant les approches et les techniques de la linguistique historique, il se révèle impossible de trouver une langue apparentée au basque. Il convient d'en déduire que, si une langue aujourd'hui parlée est apparentée au basque, elle s'en est séparée il y a plus de 8000 ans. Peut-être, dans un avenir plus ou moins proche, des méthodes d'investigation seront-elles mises au point qui permettront d'apporter la démonstration que le basque est apparenté à des langues parlées quelque part sur notre planète. Toujours est-il qu'à l'heure actuelle cette démonstration est impossible. Faudrait-il abdiquer et renoncer à en savoir plus ? Assurément pas. C'est ici qu'intervient tout l'intérêt de la recherche interdisciplinaire dont j'ai parlé au début et, notamment, le lien à consolider encore entre génétique et linguistique.

## 9

Il se trouve, en effet, que le débat sur la langue basque, qui ne mobilisait jusqu'à une époque très récente que d'étroits cercles linguistiques, a été relancé par les travaux d'un généticien. Il s'agit, évidemment, de Luca Cavalli-Sforza, un savant d'origine italienne qui a mené l'essentiel de sa carrière à l'université de

Stanford en Californie. A partir de ses travaux en génétique, portant notamment sur les marqueurs sanguins, il a pu mettre en évidence l'existence d'une certaine convergence entre la classification des populations, sur la base de marqueurs génétiques, et la classification des langues, sur la base de marqueurs linguistiques. En d'autres termes, on observe une congruence certaine entre les résultats d'une classification génétique et ceux d'une classification linguistique. Cela n'est pas étonnant dans la mesure où, lorsque des populations se divisent, langues et gènes encourent un processus parallèle de différenciation.

### Classification linguistique vs. classification génétique



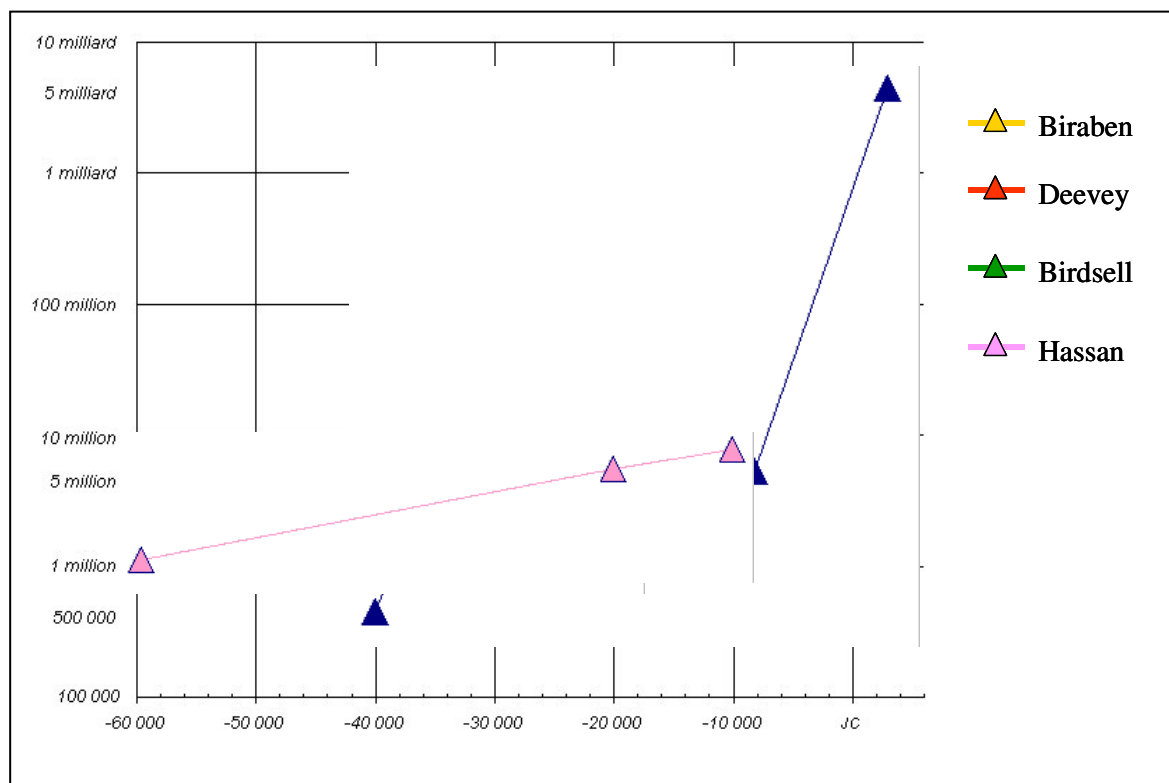
**3. LES GÈNES ET LES LANGUES** évoluent sans doute parallèlement, car ils se différencient simultanément lorsque des populations se divisent. La distance génétique entre deux populations est représentée par leur distance à l'embranchement le plus proche (*à gauche*) ; leur distance linguistique est indiquée de même par l'arbre des familles et des superfamilles de langues (*à droite*).

Pour lire le graphique précédent, il faut se garder en mémoire le fait que la distance génétique entre deux populations est représentée sur des arborescences par leur distance par rapport à l'embranchement le plus proche tandis que la distance linguistique est indiquée par la position sur l'arbre des familles et des super-familles. Dans le cadre d'une lecture très sommaire de ce genre de

graphique, il suffit de noter que plus les branches sont courtes, plus les populations génétiques et les langues sont proches.

Un coup d'œil même rapide permet d'observer une certaine ressemblance entre le profil des populations et celui des langues. Les adversaires de tout rapprochement entre données génétiques et linguistiques, par position de principe, insistent sur les dissemblances. Ceux qui sont favorables au croisement des données font, au contraire, remarquer la présence de ressemblances.

Peut-être entendrons-nous s'exprimer, au cours de la discussion qui va suivre, les tenants du verre à moitié vide et les supporters du verre à moitié plein. Pourtant, il est, à mon sens, remarquable qu'il y ait des ressemblances entre les positions respectives des populations et des langues. On pourrait, en effet, s'attendre à n'en trouver aucune puisque, bien évidemment, la langue qu'un homme parle n'est pas déterminée par ses gènes. L'idée qu'on puisse parler la langue de ses gènes est absurde ! Supposons un homme dont le père et la mère sont des aborigènes d'Australie ; il est né et a été élevé dans la banlieue parisienne. Vivant dans un environnement parisien et non en Australie, il parlera le français avec un accent parisien et pas du tout avec ce qui serait l'accent d'un aborigène australien. Les gènes d'un individu n'entretiennent aucun rapport de causalité avec la langue dans laquelle il s'exprime. Compte tenu de ce fait d'évidence, il est remarquable, à mon sens, de noter l'existence d'une certaine correspondance entre gènes et langues. Cette correspondance réclame à tout le moins une explication.



Le graphique précédent présente l'évolution des populations sur le plan démographique depuis 60.000 ans. Sur l'axe horizontal, la chronologie ; sur l'axe vertical, les populations.

- Le premier phénomène à observer est celui-ci : il n'y avait que 10 millions d'individus sur terre il y a 10.000 ans. On remarquera, au passage, que 10.000 ans pour notre espèce, celle d'*Homo sapiens*, qui remonte à environ 100.000 ou 150.000 ans, c'est hier ! Hier, c'est-à-dire l'apparition de l'agriculture. Nous sommes aujourd'hui 6 milliards. Il y a 10000 ans, la population de l'ensemble de la planète était à peu près équivalente à celle de la région parisienne aujourd'hui.
- Maintenant, remontons un peu dans le temps : lorsque *homo sapiens* arrive en Europe Occidentale, il y a environ entre 40.000 et 50.000 ans, la population de la planète était d'environ 1 million d'individus, soit la population de Lyon aujourd'hui. A cette époque, correspondant aux premières migrations d'*homo sapiens*, les densités de population étaient, par conséquent, extrêmement faibles. Il existait des groupes de chasseurs-cueilleurs, comptant peut-être entre 15 et 30 individus, pourvus de la capacité langagière mais qui avaient peu l'occasion de communiquer avec des individus d'autres groupes en raison de la très faible densité de population et de la taille des espaces de circulation. Lorsque ces groupes se déplaçaient, ils transportaient évidemment leurs gènes, mais aussi leurs langues. Puisque rares étaient les rencontres entre groupes, rares étaient contacts de langues et puisqu'on mélangeait très peu les gènes, les langues ne se mélangeaient donc pas.
- Opérons encore un grand saut en arrière, de 50.000 ans aussi : avec les données de génétique des populations qui vont vous être présentées par Luis Quintana, vous allez pouvoir constater que toutes les recherches conduites durant les dernières années mènent aux conclusions suivantes : 1) les *Homo sapiens* ont une origine africaine, 2) ils sont sortis de l'Afrique il y a environ 100000 ans, donc relativement récemment. C'est à cette date qu'un premier ensemble de population a quitté l'Afrique en direction de l'Asie du Sud-Est tandis qu'un second ensemble de population a migré d'abord au Proche Orient pour gagner ensuite l'Europe Occidentale. Et cette dernière vague de populations en provenance d'Afrique est arrivée en Europe Occidentale il y a 40.000 ou 50.000 ans.

Ce qui amène à poser la question suivante, à laquelle je ne répondrai pas pour ménager un certain suspense : les populations basques d'aujourd'hui descendent-elles directement des premiers arrivants dans cette partie de l'extrême ouest de l'Europe occidentale, il y a environ 40000 ans ?

Jean-Marie Hombert