

Etude des populations et singularité linguistique du Pays basque

Bernard Oyharçabal

► **To cite this version:**

Bernard Oyharçabal. Etude des populations et singularité linguistique du Pays basque. 2005.
<artxibo-00000017>

HAL Id: artxibo-00000017

<https://artxiker.ccsd.cnrs.fr/artxibo-00000017>

Submitted on 18 Sep 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Etude des populations et singularité linguistique du Pays Basque (B. Oyharçabal, CNRS, IKER)¹

Linguistique et génétique.

Il y a quelques jours, lors d'une conversation entre collègues où nous n'étions pas tous familiers des questions linguistiques, le point fut soulevé de la pertinence et de l'intérêt d'une étude génétique centrée sur un gène réputé intervenir dans le langage au sein d'une population ayant le même profil linguistique,² par exemple bascophone. Cette question était très surprenante, et, à l'évidence, révélait un manque d'informations certain concernant la nature de la faculté de langage et ce que l'on désigne parfois sous le terme de biolinguistique. Or, comme la conférence de ce jour associe dans sa thématique génétique et linguistique, je souhaiterais commencer mon intervention par un éclaircissement sur ce point.

Beaucoup de linguistes s'accordent aujourd'hui à considérer que la faculté de langage appartient en propre à l'espèce humaine, où chaque individu naît avec une capacité cognitive spécifique le rendant apte à apprendre durant sa petite enfance les langues, par simple immersion. Cette faculté est innée, et donc naturelle et universelle, même si l'on ignore tout, ou à peu près aujourd'hui, du soubassement génétique qui lui correspond, et que sa caractérisation demeure l'objet de nombreux débats.

Du point de vue du linguiste, cette faculté peut être décrite comme une grammaire universelle, et par nature elle ne connaît aucune restriction portant sur le type de langue naturelle pouvant l'activer. Par conséquent la part spécifique des langues vis-à-vis des autres langues naturelles, ce qui fait que chacune est différente de toutes les autres, n'est pas innée mais acquise. C'est la raison pour laquelle, l'enfant, quelle que soit l'appartenance ethnique ou l'origine géographique de ses géniteurs, apprend naturellement la langue ou les langues avec lesquelles il se trouve en contact au cours de ses premières années et grâce auxquelles il développe ses relations interpersonnelles. Il n'y a donc pas de dispositif génétique correspondant spécifiquement aux types de langues³ ou, encore moins bien sûr, aux langues particulières comme l'arabe, le basque, le chinois ou le français.

Cette précision étant apportée, et en venant à une approche sociohistorique des particularités linguistiques plus traditionnelle, pourquoi une collaboration scientifique entre généticiens et linguistes spécialistes d'une langue particulière est-elle envisagée aujourd'hui ? La réponse à cette question se trouve dans la thématique de l'histoire des populations. On estime en effet que la génétique des populations humaines et la linguistique sont deux domaines de recherche qui peuvent tous deux contribuer de façon significative à une meilleure compréhension de cette matière, laquelle, nous le savons, a depuis toujours intéressé les hommes. Pour la linguistique il s'agit d'une idée ancienne : cela fait fort

¹ Texte d'une intervention lors de la conférence-débat organisé par l'ICB et le CHCB à Bayonne le 15/04/2005 et qui associait un anthropologue généticien (Dr F. Bauduer), un médecin spécialiste de maladies génétiques (Dr J.A. Urtizbera), et l'auteur du présent texte, linguiste spécialiste de la langue basque. Titre de la conférence : « Particularités génétiques des Basques : applications et liens éventuels avec la langue ».

² Ce gène (FOXP32) a été découvert grâce à l'étude d'un groupe d'une trentaine de personnes appartenant à quatre générations de la même famille, dont la moitié souffrait, entre autres, de déficiences plus ou moins sévères concernant le langage et l'usage de la parole. Cette étude, qui a donné lieu à diverses interprétations, a été publiée dans *Nature* en 2001.

³ Depuis le XIX^{ème} siècle, les linguistes classent les langues en fonction de certains traits syntaxiques ou morphologiques; par exemple, on distingue traditionnellement les langues isolantes, agglutinantes, et flexionnelles, ou bien, encore, les langues à ordre O-V (objet - verbe) et les langues à ordre V-O (verbe - objet). Ces classements typologiques sont indépendants des relations de parenté génétique entre langues.

longtemps que l'on considère que la répartition des langues et des familles de langues dans l'espace nous apprend nombre d'informations sur la façon dont les populations ont structuré les continents à travers l'histoire. Pour la génétique, à l'inverse, il s'agit d'un domaine de recherche bien plus récent, en plein développement depuis notamment deux décennies.

Histoire des populations et hypothèses linguistiques sur la longue distance

Génétiens et linguistes ont dressé dans leur domaine respectif des classifications sous forme d'arbres de filiation. Ces classifications, par nature évidemment différentes, ont ensuite été comparées, la difficulté de cette dernière entreprise venant de ce que le degré de profondeur et de fiabilité des regroupements atteint dans la classification des populations grâce à la biologie moléculaire n'a pas de réel équivalent pour la classification historique des langues.

Le tableau ci-dessous illustre les correspondances qui peuvent être établies, entre les deux types de classification.

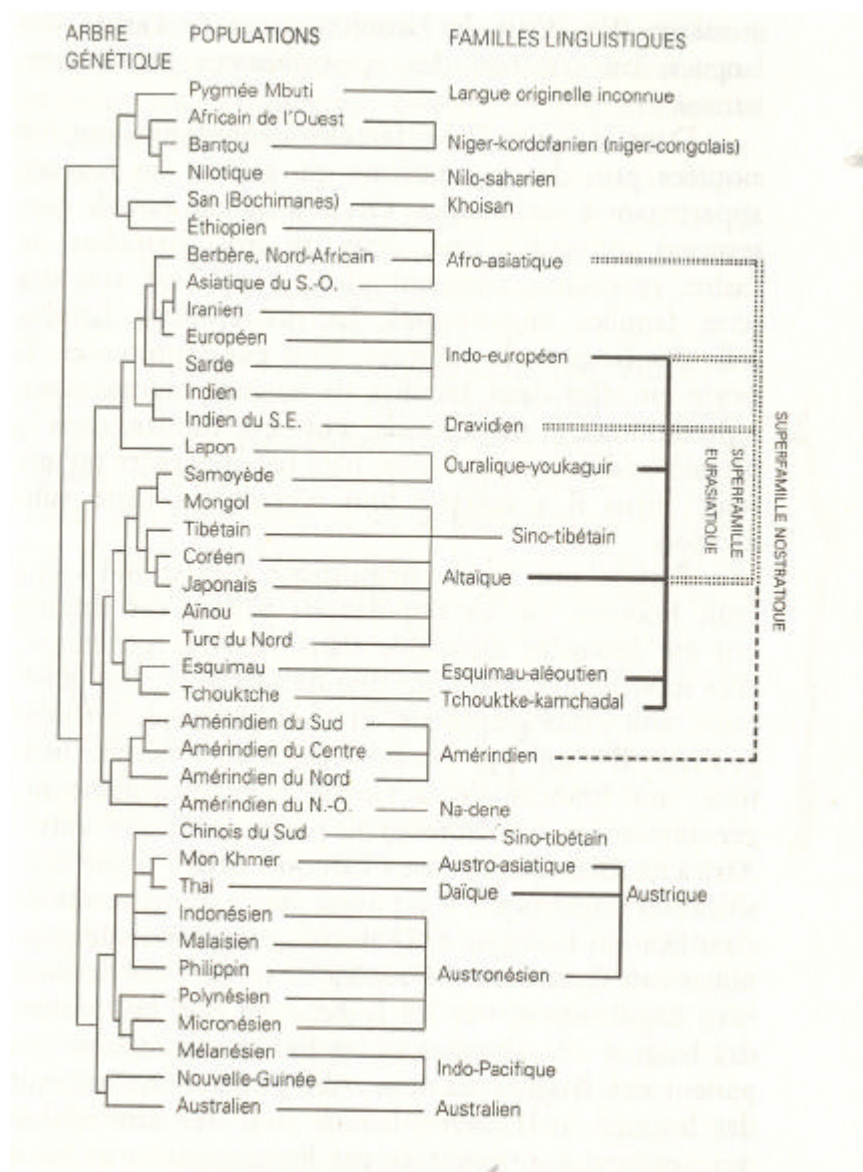
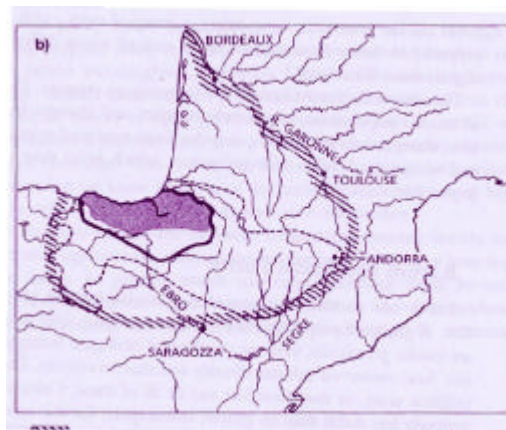


Tableau de correspondance des classifications génétique des populations et des langues (L. Cavalli-Sforza, *Langues, peuples et gènes*, O. Jacob, 1996, p. 225)

On remarquera que l'arbre génétique (partie gauche du tableau) est complet puisque l'on atteint un point de départ unique pour la population de l'Homme moderne (en Afrique il y a plus de 100 000 ans), alors que tel n'est pas le cas pour l'arbre des familles linguistiques (partie droite du tableau), qui ne remontent pas à ce qui constituerait une hypothétique langue mère unique. Par ailleurs, les regroupements les plus hauts proposés dans cette classification linguistique (notamment les macrofamilles eurasiatique ou nostratique), ou même intermédiaires (par exemple, l'amérindien) sont, en réalité, très contestés par de nombreux linguistes spécialistes.⁴

Quoi qu'il en soit, pour en venir au cas du Pays Basque, nous sommes en présence d'un cas fréquemment cité dans la littérature comme illustrant une correspondance frappante entre les deux types de classification. Il fait en effet coïncider dans le contexte européen deux singularités, l'une linguistique et l'autre génétique, comme le montrent les deux cartes ci-dessous :



En traits hachurés la limite d'extension maximale du basque
(L. Cavalli-Sforza, *Qui sommes-nous*, Flammarion, 1994, p. 322)

⁴ La possibilité d'identifier certains éléments minimaux d'une hypothétique langue originelle unique, et de regrouper les langues du monde (généralement évaluées à 6000) en une douzaine de macrofamilles, est défendue par des linguistes utilisant la méthode dite de comparaison multilatérale. Mais cette méthode est généralement contestée dans la communauté des linguistes. Les linguistes spécialistes (c'est-à-dire dont la compétence en matière de reconstruction historique est restreinte à un groupe de langues limité) qui étudient le passé des langues suivent généralement les voies de la linguistique historique classique, lesquelles, fondées en particulier sur les correspondances phonologiques entre les langues comparées, ne permettent pas d'établir de relations entre langues au-delà de quelques milliers d'années de séparation (nombre variable selon les cas). Ceci rend très difficile l'apparement des langues isolées comme le basque, étant entendu qu'en linguistique historique la notion d'isolat définit non pas une propriété intrinsèque des langues, mais simplement une limitation inhérente à la méthode et/ou un manque de connaissance; en effet, il n'existe pas de procédure permettant d'établir scientifiquement l'absence de relation de parenté entre deux langues. Le nombre des langues isolées est très variable selon les régions du monde : sur le continent eurasiatique, elles sont peu nombreuses car certaines grandes familles linguistiques s'y sont imposées, notamment en Europe même ; à l'inverse, elles sont très nombreuses par exemple sur le continent américain, bien que les multilatéralistes considèrent qu'il est possible de les regrouper en trois macrofamilles.



**Carte génétique des populations d' Europe de l'Ouest d'après Bertranpetit
(L. Cavalli-Sforza, *Qui sommes-nous*, Flammarion, 1994, p. 323)**

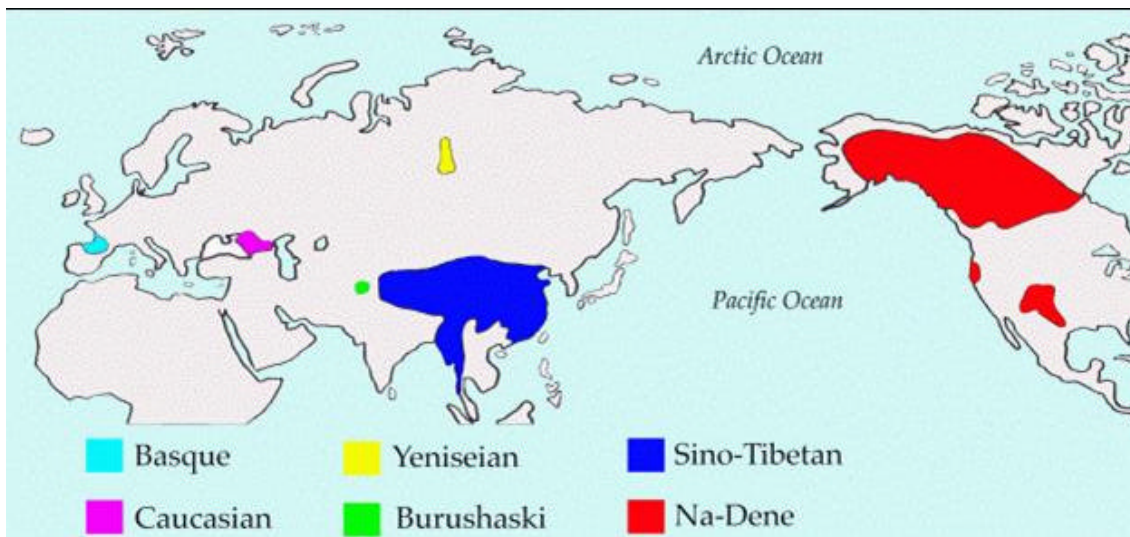
Sur la carte du haut, la zone aquitano-ibérique (en traits hachurés) correspondant *grosso modo* au maximum d'extension de la langue basque que l'on peut dresser pour le passé connu, c'est-à-dire de l'aire d'extension de la seule langue pré indo-européenne actuellement existant sur cette partie du continent. La carte suivante est une carte génétique des populations européennes : elle fait clairement apparaître un centrage sur cette même zone aquitano-ibérique. Bien entendu, cette coïncidence a frappé les esprits, et donné lieu à diverses interprétations. Si vous le voulez bien, nous allons examiner les hypothèses linguistiques qui ont fait l'objet d'une mise en résonance avec les indications apportées par les généticiens des populations.

En fait, jusqu'ici, ce sont les études linguistiques portant sur la longue distance, c'est-à-dire s'efforçant d'établir des relations avec des états de langues remontant loin dans le passé, qui ont principalement utilisé les informations apportées par les anthropologues biologistes. Deux hypothèses défendues au cours de ces dernières années doivent être mentionnées :

- L'une est l'hypothèse na-déné-caucasique, qui pourrait s'appuyer sur la thèse selon laquelle la population basque serait le prolongement des premières populations de l'Homme moderne en Europe.
- L'autre est l'hypothèse vasconne, qui serait le correspondant linguistique de la thèse selon laquelle le peuplement de l'Europe nord alpine après la dernière glaciation aurait été effectué par des populations vasconnes.

L'hypothèse na-déné-caucasique

L'hypothèse na-déné-caucasique a été développée dans le cadre de la méthode de linguistique historique dite multilatérale. Elle est illustrée par la carte ci-dessous, qui montre la répartition des membres de cette macrofamille (par ce terme on désigne un ensemble de langues ou de familles de langues supposées être issues d'une même langue plus ancienne, car présentant plus d'affinités entre elles qu'avec les autres familles extérieures à ladite macrofamille ; dans le cadre méthodologique où cette hypothèse est formulée, on compte une douzaine de ces macrofamilles pour le monde entier).



Carte de la macrofamille na-déné-caucasique selon M. Ruhlen
(disponible sur le site <http://ehl.santafe.edu/maps5.htm>)

Comme on le voit, la macrofamille na-déné-caucasique est extraordinairement dispersée à travers le continent eurasiatique et l'Amérique du nord et regroupe six composants, comprenant outre le basque, les langues du Caucase du nord, le burushaski au pied de l'Himalaya, le ket en Sibérie centrale (seule langue iénissienne encore vivante bien que très menacée), les langues sino-tibétaines et les langues indiennes na-déné du Canada et du sud des Etats-Unis. Bien qu'en aucune façon confirmée par une correspondance de classification dans les populations, cette hypothèse a pourtant été considérée par ses défenseurs (Ruhlen 1994) comme s'accommodant du scénario selon lequel les populations basques seraient apparentées aux premières populations européennes. La référence temporelle nous amènerait en ce cas à une époque antérieure à la grande glaciation laquelle a débuté il y a une vingtaine de milliers d'années. Combien de temps avec cette période? Plusieurs milliers d'années, peut-être au moment où les premières populations européennes s'installèrent sur le continent eurasiatique, avant que n'arrivent les populations qui donneront naissance à la macrofamille des langues eurasiatiques (largement dominantes ensuite sur ces territoires). En tout état de cause un temps suffisamment long pour que toute trace de parenté génétique ait disparue, alors que, selon les défenseurs de cette hypothèse, comme Ruhlen, la parenté linguistique est supposée avoir été mieux conservée.

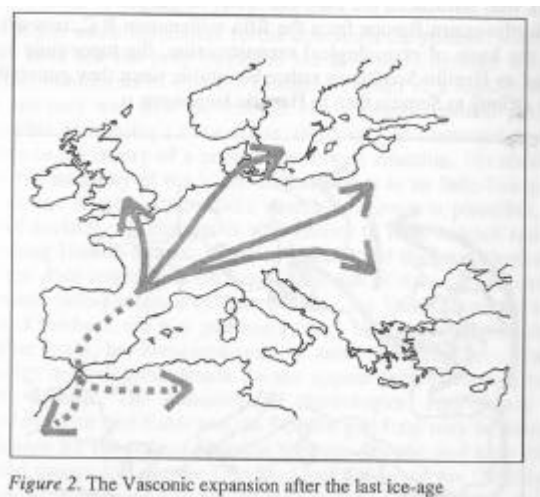
Ceux qui dans le public avaient assisté à la conférence de septembre 2003 (consultable sur le site de l'ICB) savent que cette classification n'a pas résisté à l'examen des spécialistes, en particulier des bascologues. Dix ans après le vif débat organisé par la revue *Mother Tongue* à ce propos, on peut dire que le rattachement du basque à cet ensemble (où en fait les langues isolées du continent eurasiatique n'ayant pu être classées dans les autres grandes familles – à l'exception des langues kartvèles dans le Caucase du sud – sont regroupées) n'a pas bénéficié du minimum de consensus nécessaire pour être retenu comme fondé. En effet, les bases linguistiques de cette classification sont trop fragiles, lorsqu'elles ne sont pas tout simplement erronées. En fait, aux yeux de la plupart des linguistes, les distances de temps impliquées par cette hypothèse sont absolument inatteignables par la linguistique historique.

L'hypothèse du substrat vascon

L'hypothèse vasconne, qui a bénéficié d'un important écho public à la suite d'une reprise dans les diverses éditions européennes de la revue *Scientific American* (*Pour la Science* en France) en 2002, pose moins de problèmes épistémologiques que la précédente, car elle ne

revendique pas une méthode de comparaison différente de celle habituellement mise en œuvre en linguistique historique classique. Cette hypothèse s'intègre à une analyse de la situation linguistique de l'Europe pré-indo-européenne. Pour la comprendre, il faut avoir à l'esprit qu'en dehors du basque et, à l'est, des langues finno-ougriennes (estonien, finnois, hongrois), l'ensemble des langues parlées en Europe appartient à la même famille linguistique appelée indo-européenne, car également présente (notamment) en Inde. Un linguiste allemand, professeur de linguistique à l'Université de Munich, et spécialiste des langues germaniques et de typologie, T. Vennemann, a proposé dans une série d'articles publiés au cours de la dernière douzaine d'années,⁵ de distinguer deux substrats linguistiques au sein de l'ensemble indo-européen de l'Europe de l'Ouest : l'un vascon, appelé ainsi par référence au basque, et l'autre sémite.⁶ Le terme de substrat est employé en linguistique pour désigner la langue remplacée en cas de substitution de langues, car ce substrat peut influencer ou apparaître parfois sous forme de traces ou de reliquats dans la langue de remplacement.

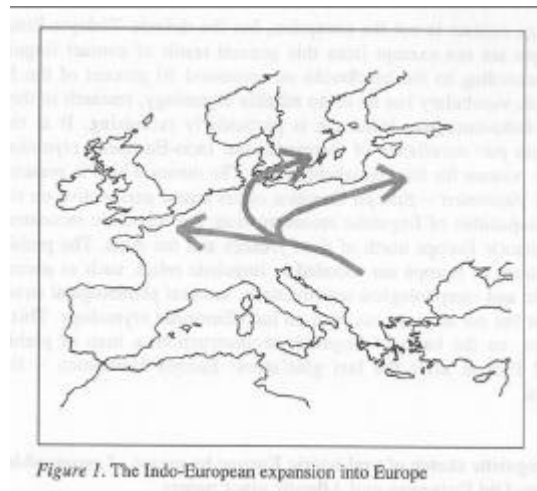
L'hypothèse substratique vasconne du point de vue des populations correspond à l'idée selon laquelle le peuplement de l'Europe nord-alpine se serait effectué à partir du refuge ibéro-aquitain à la fin de la glaciation. Cette vue peut être illustrée dans les deux cartes suivantes. La première illustre l'expansion vasconne depuis la zone de refuge aquitano-ibérique après la dernière glaciation; la seconde montre l'expansion indo-européenne produite quelques milliers d'années après la précédente et en provenance de l'est.



Carte de l'expansion vasconne à la fin de la dernière glaciation
(T. Vennemann, *Europa Vasconica - Europa Semitica*, Mouton de Gruyter, 2003, p. XIV)

⁵ La plupart de ces articles sont regroupés dans un gros volume publié en 2003.

⁶ Le substrat sémitique correspondrait à un phénomène culturel que les généticiens des populations considèrent comme indépendant de l'histoire des populations du point de vue de leur discipline : la civilisation des mégalithes, bien connue sur la bordure atlantique du continent européen et dans diverses îles méditerranéennes. Pour Vennemann, les porteurs de cette civilisation, en provenance d'Afrique du Nord, parlaient une langue chamito-sémitique, dont les traces substratiques en Europe seraient demeurées repérables, notamment dans les langues celtiques. Les langues sémitiques actuelles seraient la continuation de cette proto-langue que Vennemann appelle sémitidique.



Carte de l'expansion indo-européenne en Europe
 (T. Vennemann, *Europa Vasconica - Europa Semitica*, Mouton de Gruyter, 2003, p. XV)

Les arguments linguistiques apportés en faveur de la thèse substratique vasconne reposent notamment sur la toponymie.⁷ L'idée, en ce qui concerne la toponymie, est la suivante : les populations vasconnes ayant peuplé les territoires d'Europe devenus habitables après la glaciation dénommèrent dans leur langue les noms de lieux remarquables. Lors de la substitution linguistique au bénéfice de l'indo-européen quelques milliers d'années plus tard, ces toponymes, et notamment les hydronymes, auraient été souvent conservés comme il arrive fréquemment dans les cas de substitution linguistique. Or, depuis les années 1950, une série d'hydronymes européens, présentant du point de vue de leur forme linguistique une allure particulière, a été discutée chez les spécialistes de l'indo-européen. La proposition de Vennemann consiste à analyser ces noms à partir d'un substrat vascon. Les éléments évoqués en faveur de cette conception sont, notamment, la présence dans ce corpus hydronymique d'une suffixation agglutinante, l'emploi d'une finale *-a* semblable au déterminant de même forme du basque, une très grande fréquence de voyelles à l'initiale, ou encore une série de noms à initiale en *iz-*, correspondant (selon cette vue, inspirée d'une analyse ancienne en bascologie) à une racine attestée en basque et signifiant 'eau' (cf. *izotz* 'gelée, glace' > *iz-otz*, lit. 'eau froide').

Je dois dire que ces thèses n'ont pas été bien accueillies chez la plupart des spécialistes, notamment bascologues, comme L. Trask et J. Lakarra, qui en ont fait une critique très sévère. Par exemple, l'analyse d'un suffixe *-a* qui se retrouverait dans l'article du basque contemporain est difficilement admissible puisque ce dernier morphème est certainement de formation bien plus tardive car issu d'un démonstratif à une époque postérieure au début de l'ère (les textes aquitains ne le font pas apparaître). De même, l'existence d'un radical basque *iz* signifiant 'eau' est sérieusement contestée depuis la critique portée à cette analyse par Michelena il y a plus de quarante ans. Ces contestations, solidement argumentées, n'ont pas reçu de réponse et l'hypothèse vasconne n'est pas considérée comme valide par les spécialistes du basque. Ce fait est d'autant plus significatif que l'hypothèse du substrat vascon a été également critiquée dans le cadre indo-européaniste.

⁷ D'autres éléments sont également évoqués, comme le comptage vigésimal (par vingtaine) dans certaines langues indo-européennes (parfois à titre sporadique, comme en français), l'accentuation sur la première syllabe (celte, german, italique), des formes syntaxiques plaçant le modifieur à gauche du modifié (sauf pour l'adjectif par rapport au nom), et diverses étymologies. Mais techniquement ces divers arguments sont discutables (accentuation du basque ancien) ou peu pertinents (mode de comptage, contraintes d'ordre syntaxique mentionnées) et ils sont loin de former un argumentaire globalement convaincant aux yeux des bascologues.

Outre la pauvreté de ses arguments linguistiques, cette thèse a un autre point faible : elle présuppose, de fait, une certaine unité linguistique de l'Europe pré indo-européenne. Vennemann rejette tout paysage linguistique de l'Europe post-glaciaire du type patchwork comme proposé par Trask (1977) et il explique l'homogénéité vasconne qu'il défend par le scénario de repeuplement des territoires européens demeurés inhabitables durant la glaciation. Or, même en se situant dans ce cadre, rien ne permet de penser que tel fut le paysage linguistique de l'Europe du Sud. La situation rencontrée au début de notre ère, dans cette partie du continent sud de l'Europe de l'ouest, zone sur laquelle nous avons plus d'informations, indique même le contraire.

Les essais de ces dernières années confirment que, contrairement aux espérances qu'avaient exprimées leurs promoteurs, les limites de la linguistique historique ne peuvent être surmontées par la simple évocation d'hypothèses audacieuses, même lorsqu'elles sont cohérentes avec les hypothèses des généticiens des populations, comme dans le cas de l'hypothèse vasconne. Ces limites s'expriment sur deux plans:

- L'impossibilité de remonter très loin dans le temps pour établir d'éventuelles parentés, en particulier dans le cas des langues isolées.
- L'absence d'indications sur la diversité et les modes d'évolution des langues dans les conditions de l'Europe du paléolithique. Nous n'avons pas à l'heure actuelle de modèle satisfaisant représentant la façon dont les langues pouvaient évoluer dans un tel environnement.

Une nouvelle problématique en histoire des populations

Il semble bien qu'il faille se résoudre à admettre une différence radicale entre la génétique et la linguistique quant à la profondeur atteignable dans les classifications historiques. Est-ce à dire pour autant qu'il faille renoncer, dans les cas tels que celui de la langue basque, à une collaboration avec les historiens des populations pour comprendre le passé ? Pas nécessairement, car les questions intéressant l'histoire de la langue et des populations basques sont loin de se limiter à l'identification de parentés linguistiques lointaines. Cette dernière problématique dans une large mesure est, en l'état actuel des connaissances (et certainement pour longtemps), hors de portée d'une véritable validation scientifique, et les travaux réalisés au cours de ces dernières années, qui ont tous abouti à des résultats décevants, n'ont fait que confirmer ce point. A l'inverse, la question des conditions de la continuité de la langue basque durant la période ayant suivi la néolithisation de cette partie du continent constitue un thème plus prometteur, et, à mes yeux, plus intéressant d'histoire des populations.

Dépourvue de tout prestige culturel, parlée par des populations n'ayant jamais constitué jusqu'à l'époque récente, une entité administrative ou politique consciente de cette personnalité linguistique et en mesure par là même d'organiser sa promotion, placée en contact direct et permanent avec des populations allophones culturellement dominantes, la langue basque aurait dû dix fois disparaître, comme le firent les langues voisines, parlées éventuellement par des populations plus nombreuses, culturellement et socialement plus fortes. Il en est advenu autrement. Comment expliquer ce fait ?

Pour répondre, on évoque souvent l'isolement géographique : le basque se serait maintenu parce que parlé sur une zone du continent particulièrement difficile d'accès. Malgré la présence d'une zone montagneuse où effectivement le basque a été mieux préservé de la romanisation, ceci est globalement inexact : la zone aquitano-ibérique n'a pas été au cours des derniers millénaires une zone d'accès difficile, mais au contraire une zone de passage et de contacts continuels avec, depuis deux mille cinq cents ans, des civilisations aussi riches et diverses que celles des ibères, celtes, romains, germains, arabes, ... de même qu'au cours du dernier millénaire, les populations bascophones se sont trouvées intégrées au sein de

structures étatiques, parmi les plus puissantes en Europe, soutenant activement deux langues romanes, prestigieuses et à vocation hégémonique et impériale précoce.

L'idée de territoires basques géographiquement isolés et linguistiquement homogènes, inatteignables par les influences extérieures, encore évoquée par certains, est inexacte. Au demeurant, l'érosion et la réduction de la zone bascophone depuis le début de notre ère témoigne de la réalité de ces influences, de même que l'adoption par les populations basques, sans grandes difficultés semble-t-il, des mêmes modes de vie, de croyance et de production des biens que les autres populations ouest-européennes.

C'est ici que les informations apportées par les géogénéticiens s'avèrent intrigantes car elles nous indiquent que la continuité de la langue est associée à la présence de caractéristiques provenant des populations européennes anciennes. Cette coïncidence est soulignée par les anthropologues généticiens, qui lient la conservation parmi ces populations de ces traits résiduels considérés comme paléolithiques, au facteur linguistique : la langue aurait constitué un élément structurant de comportements endogamiques de ces populations. L'explication pourtant n'est pas vraiment satisfaisante. En premier lieu, on ne voit pas pourquoi le facteur linguistique aurait eu cet effet uniquement dans le cas des populations basques. En second lieu, il semble que l'argument soit frappé de circularité : en effet, en l'absence de structures socio-politiques faisant intervenir d'autres facteurs, la langue n'ayant pu être préservée que par la simple transmission familiale, on explique généralement le maintien de la langue par les tendances endogamiques des mêmes populations. Pour sortir de cette circularité, il paraît inévitable de prendre en compte d'autres éléments explicatifs, appartenant eux à la sphère de l'anthropologie culturelle (organisation socio-familiale notamment), même si l'on a dispose de peu d'informations à ce sujet.⁸

La question des conditions de la pérennité de la langue et de la relation de cette continuité avec l'évolution de la population demeure irrésolue à ce jour, et c'est pour l'aborder que l'on envisage aujourd'hui une collaboration associant notamment anthropologues généticiens et linguistes, sur une base moins aléatoire que pour la question des origines ou des parentés lointaines de la langue basque. Autour du thème général et central de la dynamique langue – population depuis le néolithique jusqu'à la fin du premier millénaire, diverses questions annexes, d'un intérêt certain pour le linguiste, pourraient être abordées comme, par exemple, la chronologisation de la formation des dialectes du basque (de création récente, au haut Moyen Age, selon les linguistes, bien plus anciens car renvoyant aux divisions tribales antérieures à la romanisation, selon les anthropologues à la suite de Caro-Baroja), ou l'existence d'une hétérogénéité des populations basco-pyrénéennes, indépendantes ou non des divisions linguistiques anciennes.⁹

⁸ Une voie d'exploration pourrait être celle proposée par Etchebarne dans une étude de 2003 (DEA d'études basques, CIEB, Bayonne). En s'inspirant de la démarche de Benveniste pour les langues indo-européennes, elle consiste en une analyse du vocabulaire de la famille en basque, et s'appuie sur la typologie des organisations socio-familiales dressée en anthropologie.

⁹ C'est une thématique peu traitée jusqu'ici. Cependant, pour le Pays basque sud, il convient de mentionner une publication d'Iriondo & al. (2003), déjà évoquée précédemment par F. Bauduer dans son exposé, et selon laquelle les populations basques au-delà de la Bidassoa seraient structurées en fonction des partitions dialectales est-ouest, sans que la frontière géographique marquée par la ligne de partage des eaux ne joue aucun rôle significatif.