

## Pour une linguistique du désordre et de la complexité

Louis-Jean Calvet  
Université de Provence  
[louis-jean.calvet@wanadoo.fr](mailto:louis-jean.calvet@wanadoo.fr)

### **Résumé**

*Une certaine crise de la linguistique se lit à travers différents indices, dont le moindre n'est pas son changement de nom (la linguistique laissant place aux sciences du langage), mais aussi son éclatement en différentes disciplines et sa disparition des medias (voir par exemple la difficulté à trouver aujourd'hui un rayon « linguistique » dans une librairie). Il s'agit là, cependant, de phénomènes de surface. Plus profondément, la linguistique se trouve dans une contradiction entre le caractère social de la « langue », assez généralement reconnu, et l'absence de référence au social dans les procédures de description. Elle a en outre du mal à se situer face à la complexité des situations et présente une tendance forte à vouloir ordonner ce qui semble souvent procéder du désordre. Or le désordre est partout : dans les « langues » (désordre interne, fruit de la variation) et entre les « langues » (désordre externe, babélien, lié en partie mais pas seulement à la mondialisation). Cette recherche de la mise en ordre, qui n'est pas en soi critiquable (après tout le rôle de la science est aussi de tenter de présenter de façon ordonnée le désordre des données ou des observables), nous mène parfois à simplifier considérablement les choses, voire à truquer (volontairement ou non) les données. Ce que j'appelle dans un passage de ce texte « la tentation de la loi » nous pousse à vouloir généraliser des tendances, des pratiques souvent erratiques. Or la notion de « langue » recouvre deux choses différentes : une construction abstraite produite par la linguistique d'une part, et d'autre part les pratiques et les représentations des locuteurs. Ainsi il nous est difficile de donner des réponses univoques à des questions apparemment simples, comme le nombre de langues parlées dans le monde, ou les langues les plus parlées, pour la simple raison que nous raisonnons en fonction de la*

L.J. Calvet

*« langue » du linguiste alors que les réponses se trouvent du côté des pratiques des locuteurs.*

*Face à ces difficultés il est possible de considérer que les « langues » sont des comportements tendanciels, les locuteurs fluctuant sans cesse autour d'un « noyau dur » ou entre plusieurs « noyaux durs » (dans les situations de plurilinguisme, de « langues mixtes », de « diglossie »), en se rapprochant parfois de l'un ou l'autre de ces noyaux, de ces formes tendanciennes.*

*S'ensuivent un certain nombre de propositions, que j'ai déjà présentées dans des publications antérieures mais que je réprécise ici, pour tenter d'ouvrir un débat sur ce que pourrait être une linguistique du désordre et de la complexité. Le cadre général en est la distinction entre une approche digitale et une approche analogique des observables, que je suggère non pas d'opposer mais de réunir dialectiquement.*

## **Introduction**

Claude Levi-Strauss écrivait en 1945 que la phonologie « n'a pas seulement renouvelé les perspectives linguistiques : une transformation de cette ampleur n'est pas limitée à une discipline particulière »<sup>1</sup>.

Et son *Anthropologie structurale*<sup>2</sup> sera marquée par cette vision enthousiaste et optimiste. Pour Pierre Bourdieu, un demi-siècle plus tard, Claude Levi-Strauss avait tenté de construire la science ethnologique « en usant notamment de la référence à la linguistique, alors à son zénith »<sup>3</sup>. Or le *zénith* est, comme on sait, un point culminant, c'est-à-dire qu'il a derrière lui une ascension et devant lui une lente chute jusqu'au *nadir*, son opposé. Pourtant les linguistes, peut-être grisés par cette image de science pilote que leur renvoyaient les autres sciences humaines, sont restés agrippés à cette idée de zénith, sans se rendre compte que leur science non seulement ne faisait plus rêver (ce qui n'est en soit pas très grave : là n'est pas nécessairement le but de la recherche)

---

<sup>1</sup> « L'analyse structurale en linguistique et en anthropologie », *Word* vol. 1, n°2, 1945, 1-21.

<sup>2</sup> Paris, Plon, 1958.

<sup>3</sup> Bourdieu P., *Esquisse pour une auto-analyse*, Paris, Raisons d'agir, 2004, 51.

L.J. Calvet

mais encore tournait en rond et plongeait chaque jour un peu plus dans une surdité grave face aux phénomènes sociaux. La linguistique, « science pilote », devenait un avion sans pilote, les linguistes avaient de moins en moins de choses à dire et les disaient d'une façon de plus en plus absconse, ce qui est la caractéristique de tous les dogmatismes.

Mais on aura remarqué que là où Claude Levi-Strauss parlait de la phonologie, Pierre Bourdieu parlait de *la* linguistique dans son ensemble. Depuis lors, *la* linguistique a officiellement disparu en France, remplacée dans les programmes des universités, les sections du CNRS et du CNU par les « sciences du langage ». Il y a là non pas une innovation mais deux : le remplacement de *langue* par *langage* d'une part (puisque la linguistique pourrait être définie comme *la science de la langue* - nous verrons plus loin que c'est à peu près la définition qu'en donne le Dictionnaire Robert). Et d'autre part, donc, le passage du singulier au pluriel, d'une science à des sciences. Non plus *la* science mais *les* sciences, du *langage* et non plus de *la langue*. Evaluer la seconde innovation impliquerait que nous disposions de définitions univoques de *la langue* et du *langage*, ce qui est loin d'être le cas. Disons simplement que ce passage de langue à langage élargit considérablement le domaine. Quant à la première innovation, elle va nous retenir plus longtemps.

Intervenant en 1991 aux « Assises nationales de la linguistique », Anne-Marie Houdebine s'interrogeait sur la « dilution de l'objet » qui se trouvait (ou se trouverait) derrière ce changement d'appellation. Le mot *dilution*, qui apparaît souvent dans son texte, ne concerne d'ailleurs pas seulement l'objet mais aussi la science, l'abandon de la langue selon elle étant lié à celui de la linguistique. Mais surtout elle posait de nombreuses questions qui me paraissent importantes. Laissons-lui la parole :

« Paysage pour le moins hétéroclite. Que se passe-t-il ? L'Objet de la Science (ici La Langue) se dilue-t-il ? Un autre Objet (disons vite Le Langage) se constitue-t-il ? Objet nouveau ? Réellement ? Comment ? De quel ordre, unitaire, homogène ou complexe, hétérogène que la science du XX<sup>ème</sup> siècle, celle qui a pu émettre le principe d'incertitude, la théorie du chaos, etc. peut accueillir ? Ou Objet hétéroclite, incertain,

L.J. Calvet

aux marges mal définies ? »<sup>4</sup>.

A ces questionnements sur l'objet de la science, elle en ajoutait d'autres sur sa fonction :

« Bref de ce champ hétéroclite (et pas encore hétérogène) que faisons-nous ? Que transmettons-nous dans nos cours, nos articles sous le nom de sciences du langage ? »<sup>5</sup>.

Et elle formulait une hypothèse concernant la fuite de la langue, hypothèse issue de sa double compétence (linguiste et psychanalyste) : la volonté de tout ramener ou réduire au social impliquait l'exclusion du sujet, du « fait que toute langue a comme une consistance propre qu'elle impose au sujet, à l'enfant qui y entre, en lui laissant des traces indélébiles »<sup>6</sup>.

Si l'abandon de la linguistique (au profit des sciences du langage) a impliqué l'abandon de la langue (ou la « dilution de l'Objet » comme l'écrit Houdebine), ce n'est cependant pas seulement par refus du freudisme mais d'abord par refus de certains dogmatismes (Martinet, Chomsky...) et plus largement par refus du structuralisme, sans que l'on propose grand-chose de consistant pour le remplacer. Car cet éclatement de la linguistique a mené à une atomisation du savoir : il est aujourd'hui difficile d'avoir une vue unifiée sur les petits bouts de territoires qui prolifèrent. L'espace d'une parenthèse, Houdebine distingue entre deux adjectifs aux sens très proches, *hétéroclite* et *hétérogène*, pour qualifier le champ des études linguistiques. On sait que des éléments hétéroclites appartiennent à des genres différents, tandis que des éléments hétérogènes sont de nature différente. Dans le premier cas le champ serait *éclaté* mais *unifiable* par un regard plus large (en d'autres termes la langue existerait mais serait approchée de différents points de vue), dans le second il n'aurait pas d'unité (et la langue n'existerait pas). Dans les cuisines des grands restaurants on peut trouver une équipe d'apprentis dont l'un ne fait que hacher de l'ail ou des oignons, l'autre est spécialisé dans le pelage des carottes, un troisième bat les œufs en

---

<sup>4</sup> Houdebine A-M., « La dilution de l'Objet », *Où en sont les sciences du langage dix ans après ?*, *BUSCILA*, Paris 1991, 135.

<sup>5</sup> op.cit. p. 138.

<sup>6</sup> op.cit. p. 139.

L.J. Calvet

neige et ne fait que ça, le quatrième poche les légumes, le cinquième prépare les fonds de sauces, etc. Mais au bout du compte il y a cependant une personne (le chef ou le client) qui goûte le plat, et peut juger du résultat final. Une linguistique (ou des sciences du langage) hétérogène(s) n'aurai(en)t même plus ce palais du gastronome... Chacun, dans sa petite baronnie, y développerait des analyses séparées, sans lien entre elles, de plus en plus incompréhensibles, sans aucun lien avec les pratiques sociales (et le social n'exclut pas pour moi l'inconscient mais l'inclut). Alors, hétéroclites ou hétérogènes les sciences du langage ? Je vais revenir un peu plus loin sur cette question.

Je voudrais en effet pour l'instant évoquer deux autres de leurs caractéristiques : l'abandon du social d'une part et d'autre part la complication de plus en plus grande dont elles témoignent à propos de domaines de plus en plus réduits. Mon problème n'est pas de rechercher une façon plus simple de s'exprimer (même si la clarté de l'expression et l'élégance des théories n'ont jamais fait de mal à personne) mais de tenter de remettre un édifice à l'endroit, de revenir « aux fondamentaux », c'est-à-dire à la communication humaine envisagée sous ses différents aspects, en particulier sous son aspect social. *Le facteur social est en effet la grande victime de l'évolution de la linguistique* qui, au fur et à mesure qu'elle s'est éloignée des intuitions d'Antoine Meillet par exemple, est devenue de plus en plus fermée sur elle-même, de plus en plus limitée à une vision étroite. J'aimerais rappeler ici, par quelques brèves citations, la position de Meillet. Lorsqu'il lit le *Cours de linguistique générale* publié par les auditeurs de Saussure, il en donne un compte-rendu très critique :

« En séparant le changement linguistique des conditions extérieures dont il dépend, Ferdinand de Saussure le prive de réalité; il le réduit à une abstraction qui est nécessairement inexplicable »<sup>7</sup>.

Pour sa part, influencé par les positions du sociologue Emile Durkheim, à la revue duquel il collabore, il considère la langue comme un « fait social » :

---

<sup>7</sup> « Compte rendu du *Cours de linguistique générale* de Ferdinand de Saussure », *Bulletin de la société linguistique de Paris*, 1916, 166.

L.J. Calvet

« Du fait que la langue est un fait social il résulte que la linguistique est une science sociale, et le seul élément variable auquel on puisse recourir pour rendre compte du changement linguistique est le changement social »<sup>8</sup>.

Cette notion de fait social est pour lui centrale, et il y revient plusieurs fois :

« Le langage est donc éminemment un fait social. En effet, il entre exactement dans la définition qu'a proposée Durkheim; une langue existe indépendamment de chacun des individus qui la parlent, et, bien qu'elle n'ait aucune réalité en dehors de la somme de ces individus, elle est cependant, de par sa généralité, extérieure à chacun d'eux : ce qui le montre, c'est qu'il ne dépend d'aucun d'eux de la changer et que toute déviation individuelle de l'usage provoque une réaction [...]. Les caractères d'extériorité à l'individu et de coercition par lesquels Durkheim définit le fait social apparaissent donc dans le langage avec la dernière évidence »<sup>9</sup>.

Ces prises de position n'ont, à mes yeux, pas pris une ride. Mais le problème n'est pas seulement de revenir à cette caractéristique sociale de la langue, il est aussi de se déterminer face à la contradiction permanente entre le désordre des données d'une part (les *faits linguistiques* ou les *actes de parole* que nous recueillons sont « désobéissants », ils n'entrent pas aisément dans des cadres pré-établis) et d'autre part la recherche forcenée d'un ordre sous-jacent qui caractérise certaines directions de recherche. D'un côté le désordre de la vie, de l'autre le rêve d'un ordre supérieur... J'ai posé plus haut la question de savoir si les sciences du langage devaient être considérées comme hétéroclites ou hétérogènes. Mais cette question en implique une autre : l'objet de ces sciences est-il hétéroclite ou hétérogène ? Ce qui pose à nouveau une question : quel est cet objet ? La *langue* répondront certains avec un air d'évidence. Cela n'est cependant pas si évident que ça. La notion de *langue* recouvre

---

<sup>8</sup> *Linguistique historique et linguistique générale*, 1965 [1921], 17.

<sup>9</sup> « Comment les mots changent de sens », *L'année sociologique* 1905-1906, repris dans *Linguistique historique et linguistique générale*, 230.

L.J. Calvet

en effet deux choses différentes : une construction abstraite, un produit de la linguistique (en ce sens la langue est une invention des philosophes puis des grammairiens et enfin des linguistes) d'une part, et d'autre part une représentation des locuteurs qui considèrent qu'ils parlent une langue donnée, différente de celles des peuples voisins (et en ce sens la langue est une pratique, une construction sociale). Et ces pratiques sont essentiellement *désordonnées*. Le problème est donc de savoir quel est le domaine d'étude de la linguistique (ou des sciences du langage), et s'il est possible d'avancer quelques propositions théoriques et méthodologiques face à toutes ces questions.

Vaste programme, certes, que je n'ai ici que l'intention d'initier. Mais, dans un premier temps, je voudrais jeter un regard critique sur cette linguistique maniaque de l'ordre, avant d'en venir à la formulation (ou à la reformulation) de quelques propositions susceptibles de nous aider à naviguer dans le désordre<sup>10</sup> et à tenter d'en rendre compte sans trop le déranger, c'est-à-dire sans vouloir absolument lui imposer un ordre venu d'ailleurs.

Cette tentation de l'ordre se retrouve partout, et par exemple dans l'idée qu'il y a filiation régulière d'une forme linguistique à une autre, qu'un dialecte *vient* toujours d'une forme précédente. Traitant du dialecte pamphylien, dont il est spécialiste, l'helléniste Claude Brixhe<sup>11</sup> met en question la *Stammbaumtheorie*, la croyance en la transmission automatique de la langue « maternelle ». Pour lui, « Intellectuellement satisfaisant, ce modèle répond à l'opinion commune, selon laquelle il y a un commencement à tout » (p.14). La métaphore de l'arbre implique en effet la mono genèse : il y a un latin et des langues romanes, un indo-européen et des langues indo-européennes, un Urbantu et des langues bantou, etc. Or, souligne Brixhe, aussi loin qu'on remonte dans le temps on trouve du plurilinguisme. Et la métaphore de la filiation devrait donc être remplacée par celle de l'héritage (on n'hérite pas que de ses parents). Cela le mène à proposer la notion de *coproduction* :

---

<sup>10</sup> Je prendrai le terme *désordre* en deux sens : un désordre externe, ou babélien, celui des langues du monde à l'heure de la mondialisation, et un désordre interne, résultat de la variation.

<sup>11</sup> « De la filiation à l'héritage, réflexion sur l'origine des langues et des dialectes », *Etudes anciennes* n°31, Nancy, 2006.

L.J. Calvet

« Etant donné que cette coproduction est le résultat d'interactions diverses, on ne peut envisager l'histoire d'un dialecte indépendamment de l'histoire de la communauté qui l'a utilisé, ni même indépendamment de l'écosystème où il a évolué »<sup>12</sup>.

Mais il dresse au passage un constat cruel de l'aveuglement des études anciennes qui a peut-être valeur plus générale :

-« Conceptuellement on en est resté à la fin du XIX<sup>ème</sup> siècle ou au début du XX<sup>ème</sup>, sans la moindre velléité de remise en cause, sans même l'ombre d'une interrogation sur la valeur de la méthode et des objectifs » (p. 8).

-« Paradoxalement, cet isolement favorise notre survie dinosaurienne en nous empêchant de prendre conscience de nos faiblesses, en même temps que le fonctionnement de notre microcosme assure notre légitimité institutionnelle, avec autoreproduction des enseignants et toute une série de procédures liturgiques rassurantes et homogénéisantes (colloques, mélanges, comptes rendus, honneurs divers) » (pp. 8-9).

En effet ce portrait au vitriol peut parfaitement s'appliquer à l'ensemble de la linguistique. Elle a d'abord institué un ordre de la langue, puis son ordre propre (l'ordre de la linguistique) et se trouve aujourd'hui confrontée à un désordre de la discipline sans nécessairement voir ou admettre que derrière ce désordre il y a peut-être celui de l'objet d'étude.

C'est donc autour de ces notions d'ordre et de désordre que je voudrais, pour commencer, tenter de réfléchir.

## **1. Une linguistique maniaque de l'ordre**

---

<sup>12</sup> « Situation, spécificités et contraintes de la dialectologie grecque », *Etudes anciennes* n°31, Nancy, 2006, 40.



L.J. Calvet

*1.1. D'Austerlitz à Waterloo*

Le 2 décembre 1805, dans une petite ville de Moravie, Napoléon remportait une écrasante victoire sur l'armée austro-russe : Austerlitz. Dix ans plus tard, le 18 juin 1815, près d'une petite ville belge, le même Napoléon était défait par les Prussiens de Blücher et les Anglais de Wellington : Waterloo. Deux écrivains, l'un français, Stendhal, l'autre russe, Tolstoï, ont décrit l'une ou l'autre de ces batailles. Dans *La chartreuse de Parme* (1839) Fabrice del Dongo traverse la bataille de Waterloo et, comme dans le film de Schoendorfer, *317<sup>ème</sup> section*, on ne voit que ce que voient les yeux du personnage. Aucun plan d'ensemble, aucune conscience de ce qui se passe... Fabrice cherche à rejoindre les combattants, il veut se battre, erre de gauche et de droite sans rien comprendre : bruits de canon, mouvements de troupe. Il entend crier « Vive l'empereur » :

« -C'est donc l'empereur qui a passé là ? dit-il à son voisin.  
-Eh ! Certainement, celui qui n'avait pas d'habit brodé.  
Comment ne l'avez-vous pas vu ? lui répondit le camarade  
avec bienveillance » (p. 63).

Plus tard, mêlé à la débâcle, il ne voit que désordre :

« Le tapage sur la grande route continuait, et avait duré toute la nuit : c'était comme le bruit d'un torrent entendu dans le lointain.  
-Ce sont comme des moutons qui se sauvent, dit Fabrice au caporal, d'un air naïf » (p. 71).

Et lorsqu'il en réchappe, il est bien incapable de raconter quoi que ce soit de la bataille : tout au plus peut-il dire ce qu'il a vécu, vu par le petit bout de la lorgnette.

Dans *La guerre et la paix* (1863-69) Tolstoï décrit la bataille d'Austerlitz d'une toute autre manière :

« Quand le soleil, complètement dégagé, eut inondé la campagne de son aveuglante clarté, Napoléon, comme s'il n'avait attendu que ce moment, déganta une de ses belles

L.J. Calvet

mains blanches, fit de son gant un geste aux maréchaux et donna l'ordre d'engager la bataille. Les maréchaux et leurs aides de camp galopèrent dans différentes directions et, au bout de quelques minutes, les forces principales de l'armée se portèrent rapidement vers le plateau de Pratzen, que les troupes russes abandonnaient de plus en plus pour gagner, vers la gauche, le ravin » (p. 347).

L'auteur nous décrit Koutouzov prenant son cheval à huit heures et se diriger vers Pratzen, le flanc droit de l'armée qui, à neuf heures, n'est pas encore engagé dans la bataille, situe géographiquement les différents corps d'armée, donne le nombre de pièces d'artillerie tombées entre les mains des Français. En bref, il voit la scène de haut, se situe du point de vue d'un observateur privilégié. Del Dongo a traversé à l'aveugle la bataille de Waterloo, à peine guidé par la plume de Stendhal, Tolstoï domine celle d'Austerlitz.

Les deux écrivains ont donc fait deux choix esthétiques antinomiques. Et ce traitement d'Austerlitz par l'un, de Waterloo par l'autre, ces positionnements différents, constituent une bonne illustration des positions que peut prendre le linguiste. Il peut en effet, à l'instar de Stendhal, se situer dans le grand désordre des pratiques linguistiques, des bruits, des sons, des mélanges, partir de ce désordre, le considérer comme une donnée de la situation. Il n'est d'ailleurs pas indifférent, de ce point de vue, que Stendhal décrive Fabrice comme un locuteur approximatif du français, qui parfois passe à l'italien, se trompe, et manque même pour cela d'être pris pour un espion.

Mais le linguiste peut à l'inverse, à l'instar de Tolstoï, décider qu'il domine la scène et individualiser puis nommer ce qu'il appelle des langues. Du même coup il supprime le désordre puisqu'il sait, lui, qu'il peut assigner telle pratique à telle langue, telle autre pratique à telle autre langue. Et il n'est pas non plus indifférent que le roman de Tolstoï commence en français : c'est dans une forme recherchée de cette langue qu'Anna Pavlova Scherer s'adresse au prince Basile, que la princesse Bolkonski s'exprime dans le salon d'Anna Pavlova, que cette dernière converse avec le prince André, etc. Et le passage du russe au français et du français au russe, censé donner de la couleur locale à la description de la cour du Tsar très francophone, distingue soigneusement entre les deux langues : ici pas de mélange, pas d'approximation, mais des langues

homogènes, autonomes.

Ainsi les deux romanciers illustrent-ils l'un l'ordre et l'autre le désordre non seulement dans leur traitement de la description d'une bataille mais aussi dans leur mise en scène des langues...

Didier de Robillard a donné un bel exemple de ce désordre des pratiques et des différentes façons de le traiter<sup>13</sup>. Il part d'une affiche manuscrite accolée sur le congélateur d'une épicerie de l'île Maurice et donnant une liste de produits et leur prix. On y lit successivement, de haut en bas :

« Items :  
Boneless mouton  
Flaps mouton  
Gigot mouton (sans os)  
Viande haché mouton  
Epaulette d'agneau  
Portion mouton  
Entré cotte d'agneau  
Butter fish  
Jurite  
Sa chrétien  
La perle »

Face à cette liste, poursuit Robillard, il y a deux points de vue possible. Celui du touriste qui, à de rares exceptions près (*jurite, sa chrétien, la perle*), va comprendre ce qui est écrit et l'assigner soit au français soit à l'anglais. Celui du Mauricien, généralement trilingue, qui comprendra tout sans nécessairement se demander en quelle langue est écrit tel ou tel mot, d'autant que certains sont polyvalents et se déplacent entre les langues. Un bel exemple est celui de *jurite*. Il existe en créoles mauricien et réunionnais le terme *zourite*, d'origine malgache, qui désigne le poulpe (et qui a donné à la Réunion une belle expression : *filer en zourite*). Or il y a une correspondance phonétique entre le

---

<sup>13</sup> Robillard D. (de), « Quand les langues font le mur lorsque les murs font peut-être les langues : *mobilis in mobile*, ou la linguistique de Nemo », *Revue de l'université de Moncton*, vol. 36, n°1, 2005.

français et ces créoles qui fait que lorsqu'on a un J d'un côté on a un Z de l'autre. Ainsi le prénom français Georges donne-t-il en créole Zorz. *Zourite* devrait donc, si cette correspondance s'appliquait, être prononcé en français \**Jourite*, mais il se dit en fait en français mauricien *ourite*, sans doute parce qu'on a entendu dans le Z initial la marque d'une liaison au pluriel et segmenté après le Z (des *zourites*>des *ourites*). Le rédacteur a donc produit une hypercorrection (*jourite*) et il l'a orthographié à l'anglaise (*jurit*), ce qui ne pose sans doute pas de problèmes au lecteur mauricien mais laisse perplexe l'étranger connaissant le français et l'anglais. Robillard en conclut que ce type d'affiche témoigne de la situation sociale et linguistique de l'île et, plus largement, que « ce qui est constitutif des langues, ce n'est ni leur matérialité, ni leur structure, mais la croyance dans leur existence, ancrée dans le besoin qu'elles existent ». C'est-à-dire qu'une idée de la langue, stable, homogène, hors contexte, l'emporte sur la complexité des situations.

On aura compris, sans qu'il soit besoin de le développer plus avant, le parallélisme entre les différentes réactions à ce type d'affiche et les traitements différents d'une bataille par Stendhal et Tolstoï : d'un côté un désordre assumé, abordé de l'intérieur, et qui participe à la construction du sens, de l'autre un ordre venu d'en haut, un sens préétabli, et le projet d'une *linguistique de l'ordre* qu'ici même Philippe Blanchet appelle *structurologistique* et Didier de Robillard *technolinguistique* : je crois que nous parlons à peu près de la même chose.

Face à la situation linguistique de l'île Maurice en effet, le linguiste ne peut d'abord que constater le « désordre », la difficulté à dire ce que les gens parlent (ou, dans le cas de l'affiche, écrivent). Sa première tâche n'est pas, dès lors, de faire le tri entre les différents codes, de les nommer, mais d'étudier comment les Mauriciens communiquent dans ce désordre, ou malgré ce désordre, voire même grâce à ce désordre. *Car ce qui peut, aux yeux des tenants de « l'ordre », sembler faire obstacle à la communication peut tout aussi bien la constituer.* L'histoire de l'île, la présence des Français d'abord, l'importation de main-d'œuvre africaine ou malgache, la constitution d'un créole, puis l'arrivée des Britanniques, l'importation de main-d'œuvre indienne, tout cela constitue un cadre général permettant de

L.J. Calvet

*comprendre* la généalogie de cette situation, dont l'environnement graphique témoigne en partie.

Il demeure bien sûr que plusieurs langues *tendanciennes* sont en présence, qu'il y a des *pôles linguistiques* que l'on peut nommer anglais, français, créole, bhojpuri, etc., ces pôles ne se situant pas sur une ligne droite mais se constituant plutôt en étoile, ou sur un cercle. Et c'est entre ces pôles que se situent les pratiques, se rapprochant de l'un d'entre eux dans certaines situations (formelles, familiales...), les mêlant dans d'autres situations. De la même façon qu'un locuteur du français (ou de toute autre langue) se rapproche selon les situations de communication du pôle normé, standardisé, ou du pôle familier, voire d'une forme « dialectale », on va dans une situation comme celle de Maurice se rapprocher parfois de l'une ou l'autre des langues tendanciennes, du français, de l'anglais ou du créole. De ce point de vue, face à l'affiche de Robillard, certains (les Mauriciens) comprennent ce qui est écrit (et achètent ce qu'ils veulent) tandis que d'autres (les linguistes) tentent de « mettre de l'ordre », de séparer entre les langues.

Certains en sont conscients et le disent. Le 8 octobre 2005, dans un colloque de sociolinguistique<sup>14</sup>, Catherine Miller présentait et commentait des enregistrements faits en 1981 dans des tribunaux de Juba, au sud Soudan. D'un côté, pour les besoins de son exposé, elle faisait un décompte des interventions en différentes langues (bari, arabe véhiculaire dit « Juba Arabic », arabe dialectal local, anglais, etc.), mais de l'autre elle expliquait que c'était elle qui, par souci de clarté, faisait cette distinction, que les locuteurs pour leur part passaient facilement d'une forme à l'autre et qu'il était même parfois difficile de dire exactement ce qui était parlé. Les extraits de corpus qu'elle avait distribués étaient ainsi parsemés de passages colorés en différentes teintes pour montrer ce qui, dans les déclarations des accusés ou des juges relevait plutôt du dialectal, plutôt de la forme véhiculaire voire parfois de l'arabe classique, et l'on voyait ainsi tout à la fois des langues tendanciennes entre lesquelles évoluaient les locuteurs, des comportements qui faisaient qu'ils se rapprochaient de l'un ou l'autre de ces pôles avec des effets (recherchés ?) variés mais surtout *de la*

---

<sup>14</sup> « Faits de langue et de cultures dans la dynamique interactionnelle », Paris 7-8 octobre 2005.

L.J. Calvet

*communication* grâce à des « langues », bien sûr, mais aussi malgré la pluralité de ces « langues ».

On a proposé pour rendre compte de ces pratiques les notions d'*alternance codique* et de *mélange de langues*, syntagmes directement traduits de l'anglais (*code switching*, *code mixing*) dont beaucoup semblent se satisfaire. L'alternance codique par exemple serait donc une pratique que les linguistes ont tenté de décrire, puis de mettre en équation, les différentes approches (variationniste, générative, conversationnelle...) donnant chacune leur point de vue, pratique dont la nomination est claire : elle implique l'existence de deux codes que l'on peut donc délimiter, décrire séparément, et entre lesquels le locuteur se déplace, passant de l'une à l'autre comme on allume ou éteint l'électricité (c'est le sens en anglais du verbe *to switch on* ou *off*). Les locuteurs font donc du *code switching* sans le savoir, comme monsieur Jourdain faisait de la prose, et le linguiste omniscient décrit avec pertinence leurs pratiques, leurs passages d'un code à l'autre.

Cette volonté de mettre de l'ordre dans le désordre peut paraître très satisfaisante : il y a un ordre des langues qui n'est pas mis en question et un désordre des pratiques, que l'on peut décrire par référence à cet ordre, en faisant éventuellement le tri dans les données, pour ignorer celles qui résisteraient à l'ordre. Mais l'ennui est que ces « alternances codiques » ou ces « mélanges de langues » ne laissent pas nécessairement les « langues » ou les « codes » indemnes, dans l'état où on les aurait trouvés avant usage. De la même façon que le plurilinguisme d'un pays n'implique pas celui de tous ses habitants, la présence de deux langues dans une niche écolinguistique n'implique pas le bilinguisme de tous les individus qui y vivent. Je pense en particulier à des situations dans lesquelles ce que le linguiste classera comme *code switching* ou *code mixing* est simplement une forme quotidienne de communication dont on peut, *historiquement*, souligner les origines diverses mais qui n'implique pas nécessairement la domination de ces origines par les locuteurs. Pour certaines couches sociales tunisiennes par exemple, le mélange arabe/français est un parler identitaire, convivial, qui implique bien sûr l'existence de l'arabe tunisien et du français de Tunisie mais peut être présenté comme un moyen terme dans un continuum entre deux langues tendanciennes. En parlant « arabe tunisien » on utilise des mots dont on ne sait pas, ou plus, qu'ils sont

d'origine française<sup>15</sup>, en parlant « français de Tunisie » on utilise des formes dont on ne sait pas qu'elles n'appartiennent pas au français hexagonal<sup>16</sup>, mais il y a tout un continuum entre ces deux *acclimations linguistiques* (nous verrons plus loin le sens que je donne à ce syntagme), continuum dont la notion de *code switching* ne rend pas forcément compte de la meilleure façon.

## 1.2. Un passage par les langues<sup>17</sup> « mixtes »

On dit dans le discours juridique que ce qu'une loi a fait, une loi peut le défaire. Les politiques linguistiques reposent sur un principe du même ordre, qui s'exprimerait ainsi : « on peut transformer ou modifier *in vitro* ce qui a été fait *in vivo* ». Je voudrais donc tenter d'analyser ce que ces politiques linguistiques peuvent, dans leurs réussites et leurs échecs, nous apprendre de la langue, ce que l'analyse de leurs réalisations nous enseigne. Or elles ont :

-tenté de faire des langues (les inventeurs de langues, comme l'espéranto ou l'ido, mais aussi les promoteurs de langues, les états qui font d'une des formes en présence la forme officielle)

-transformé le lexique des langues (depuis Kemal Ata Türk en Turquie jusqu'aux nombreuses commissions de terminologie)

-changé les fonctions des langues, les rapports entre les langues (promotion au statut de langues officielles, de langues régionales, de langues nationales, etc.)

-normalisé des langues, parfois de façon policière (lorsque des lois, comme en France la loi Toubon, permettent de poursuivre et de condamner les contrevenants...)

-fait disparaître des langues (les langues indiennes d'Amérique par exemple, immolées sur l'autel de l'anglais, de l'espagnol ou du portugais...).

Mais elles n'ont jamais, du moins à ma connaissance :

-modifié la phonologie des langues

---

<sup>15</sup> Par exemple le mot *zufri*, « voyou », venant du français *les ouvriers*.

<sup>16</sup> Par exemple *cycliste* pour « réparateur de bicyclettes », *taxiste* pour « chauffeur de taxi »...

<sup>17</sup> Je continue par facilité à utiliser le mot *langue* pour désigner ce que je considère comme un ensemble de pratiques et de représentations mouvantes.

L.J. Calvet

-transformé radicalement la syntaxe d'une langue.

Cela nous apprend-t-il quelque chose sur ce qu'il est convenu d'appeler langue (terme dont il est difficile de se débarrasser...)? La syntaxe et la phonologie, auxquelles les interventions humaines *in vitro* (les politiques linguistiques) semblent n'avoir jamais touché, constituent-elles une sorte de noyau dur, immuable ou mouvant? Et le lexique peut-il varier sans que ce noyau dur soit affecté? Derrière ces questions s'en profile une autre : à partir de quel moment, de quel degré de différenciation une langue cesse-t-elle d'être elle-même?

La question qui est ici posée est celle du changement linguistique, de la variation et de l'évolution, question qui a été largement polluée par la métaphore biologique selon laquelle les langues naissent, vivent et meurent. Ferdinand de Saussure, dans la troisième des conférences données en novembre 1891 à l'université de Genève, insistait pour sa part sur la continuité des langues :

« Sur l'impossibilité radicale non seulement de toute rupture, mais de tout soubresaut, dans la tradition continue de la langue depuis le premier jour même où une société humaine a parlé ; - sur ces différents points immédiatement évidents, qu'aucune langue ne peut *mourir*, si elle n'est violemment supprimée ; - qu'aucune n'a une vieillesse, et qu'aucune n'a une enfance, qu'enfin aucune langue nouvelle ne peut jamais naître sous le soleil [...] qu'il ne peut jamais y avoir sur le globe que continuation d'un idiome existant la veille, et toujours existant la veille, jusqu'à ce qu'on arrive à la nuit insondable des âges décidément antéhistoriques »<sup>18</sup>.

Dans les deux conférences précédentes, il martelait :

« Chaque individu emploie le lendemain le même idiome qu'il parlait la veille et cela s'est toujours vu. Il n'y a donc eu aucun jour où on ait pu dresser l'acte de décès de la langue latine, et il n'y a eu également aucun jour où on ait pu enregistrer la naissance de la langue française. Il n'est jamais arrivé que les gens de France se soient réveillés, en se disant

---

<sup>18</sup> *Ecrits de linguistique générale*, Paris, Gallimard, 2002, 163-164.



L.J. Calvet

*bonjour* en français, après s'être endormis la veille en se disant *bonne nuit* en latin ».

Et il n'envisageait l'hypothèse de la disparition d'une langue que dans des circonstances très particulières :

« Une langue ne peut pas mourir naturellement et de sa belle mort. Elle ne peut mourir que de mort violente. Le seul moyen qu'elle ait de cesser, c'est de se voir supprimée par force, par une cause tout à fait extérieure aux faits du langage. C'est à dire par l'extermination totale du peuple qui la parle, comme il arrivera prochainement pour les Peaux-Rouges d'Amérique du Nord. Ou bien par imposition d'un nouvel idiome appartenant à une race plus forte ; il faut généralement non seulement une domination politique, mais aussi une supériorité de civilisation, et souvent il faut la présence d'une langue *écrite* qu'on impose par l'Ecole, par l'Eglise, par l'administration... »,

affirmant que :

« Jamais on n'a signalé en effet sur le globe la naissance d'une langue nouvelle »<sup>19</sup>.

Ce qui pourrait apparaître comme un déni du changement s'explique bien sûr par sa volonté de construire une théorie de la langue, en synchronie. Et Saussure se piégeait d'ailleurs lui-même lorsque, faisant allusion à un Russe, Bogulawski, qui s'était pendant vingt fait régulièrement photographe, dans la même pose, le premier et le quinze de chaque mois, et qui exposait les 480 portraits ainsi réalisés, il commentait:

« Je n'ai pas besoin de vous dire que, si dans cette exposition on prenait sur la paroi deux photographies contiguës quelconques, on avait le même Bogulawski, mais que si l'on prenait le n°480 et le n°1 on avait deux Bogulawski. De même si l'on avait pu non pas photographe mais

---

<sup>19</sup> op. cit. pp. 152, 153, 154.

L.J. Calvet

phonographier au jour le jour dès l'origine tout ce qui a été exprimé en parole sur le globe ou sur une partie du globe, on aurait des images de langue toujours ressemblantes d'un jour à l'autre, mais considérablement différentes et parfois incalculablement différentes de 500 ans en 500 ans, ou même de 100 ans en 100 ans »<sup>20</sup>.

Sa comparaison le trahissait en effet car ces 480 photos s'étalant sur vingt ans témoignaient justement du vieillissement d'un individu, et le ramenaient à la métaphore biologique qu'il refusait. Mais son point de vue était que *continuité* et *mutabilité* sont deux principes permanents, que ces photos représentaient le même homme dans son évolution, et que la transformation de la langue n'entraîne pas en contradiction avec son unité fondamentale. En d'autres termes, les langues changent mais nous ne les voyons pas changer, et il nous faut décrire l'une des 480 photos avant de nous interroger sur ce qu'elle pourra devenir ou sur ce dont elle provient. C'est là le fondement de la distinction entre *synchronie* et *diachronie*, martelée depuis lors par la grande majorité des enseignants de linguistique et à peine modulée par la notion de *synchronie dynamique*.

Pourtant, sans parler de la variation, perceptible, qui est l'un des moteurs du changement, nous sommes régulièrement confrontés à des faits qu'il est difficile de considérer autrement que comme du changement, qui est parfois rapide, du *changement à vue* pourrions-nous dire. Prenons un exemple simple, celui de la relexification. C'est Pieter Muysken qui, en 1981<sup>21</sup>, a décrit ce qu'il appelle une « *media lingua* », appelée par ses locuteurs *utilla ingiru* (« petit quichua »), qui tire ses catégories fonctionnelles du quichua et son lexique de l'espagnol. Le quichua est ainsi *relexifié* par l'espagnol, présentant une syntaxe indiscutablement quichua et un lexique majoritairement d'origine espagnole. L'exemple suivant illustre bien le processus :

Media lingua :

---

<sup>20</sup> op.cit. pp 156-57.

<sup>21</sup> « Halfway between Quichua and Spanish : The case for relexification », in Highfield A., Valdman A. dir., *Historicity and Variation in Creole studies*, Ann Arbor, Karoma.

L.J. Calvet

*Unu fabur-ta pidi-nga-bu bin-xu-ni*

Une faveur demander venir

Quichua :

*Shuk fabur-da maña-nga-bu shamu-xu-ni*

Espagnol :

*vengo para pedirun favor*

On voit comment les mots espagnols *un*, *favor*, *venir*, *pedir* sont adaptés à la phonétique du quichua et insérés dans une syntaxe et une morphologie du quichua. *Unu* (espagnol *un*) remplace *shuk*, *pidi* (espagnol *pedir*) remplace *maña*, *bin* (espagnol *venir*) remplace *shamu*, mais le verbe est conjugué « à la quichua » et l'ordre des mots est quichua. Nous avons là du « réel », mais un réel qui procède résolument du désordre (ici à la fois interne et externe : voir note n°10). A quelle langue la phrase *Unu fabur-ta pidi-nga-bu bin-xu-ni* doit-elle être assignée ? S'agit-il de quichua ? D'espagnol ? De la manifestation en synchronie de quelque chose à apparaître en diachronie et peut-être amené à remplacer l'espagnol et le quichua ? En outre, quelle langue pensent parler les locuteurs qui produisent une telle phrase ? Et ceux qui l'entendent ? Car la présentation en trois lignes que j'ai donnée (*media lengua*, quichua, espagnol), reprise de Muysken, est déjà une façon de mettre de l'ordre dans une situation désordonnée, manifestation d'un besoin de classer qui relève dit-on de l'esprit scientifique mais aussi peut-être de la psychanalyse.

Ce qui est sûr, c'est que derrière la « *media lengua* », « *halfway between Quichua and Spanish* » pour reprendre le titre de l'article de Muysken, il y a des locuteurs qui se trouvent à mi-chemin non pas entre deux langues mais d'abord entre deux communautés, entre deux statuts, entre deux cultures. Certains ajouteront qu'ils sont en voie d'assimilation sociale et linguistique, et que la *media lengua* est un chaînon dans une dérive du quichua vers l'espagnol, mais il est tout aussi possible que cette *media lengua* soit amenée à se substituer au quichua et/ou à l'espagnol, ou encore que cette situation perdure, que l'espagnol et le quichua ne soient que des langues tendanciennes, comme le créole, le français et l'anglais dans l'exemple de l'île Maurice, et que la *media lengua* ne soient qu'une des possibilités intermédiaires qu'un linguiste

européen passant par là a promu au statut de « lingua ». Certains pourraient encore dire qu'une langue à ce point relexifiée est menacée de disparition, que le linguiste peut d'ailleurs élaborer une *symptomatologie* de cette situation : la fluctuation phonologique constitue l'un de ces symptômes, la relexification en constituerait un autre.

Il demeure qu'il y a là des pratiques réelles, incontournables. Mais se pose alors la question de savoir si les locuteurs sont conscients de parler cette « media lingua », ou s'ils pensent parler quichua ou espagnol. Il y a un argument fort en faveur de la première hypothèse, c'est qu'ils nomment cette forme. En effet, *nommer* est une façon de faire exister, nommer une *pratique* c'est la différencier d'autres pratiques, et donc baptiser *utilla inguru* une certaine façon de parler le quichua (ou l'espagnol) c'est décider qu'il ne s'agit pas (ou plus) de quichua. Mais il reste à savoir si cette différenciation qui donne en quelque sorte naissance à quelque chose de nouveau, ni quichua ni espagnol, est le fait des locuteurs de cette forme ou celui d'observateurs attachés aux normes et qui veulent stigmatiser cette forme. Nous pourrions tout aussi bien considérer que ces appellations, *media lingua*, *utilla inguru*, renvoient à des *champs de variation* entre des langues tendancielle, dans lesquels il y a des productions possibles et d'autres impossibles (ou du moins des productions acceptables ou non aux yeux des locuteurs), de la même façon que le champ de dispersion d'un phonème : au delà d'une certaine limite phonétique il ne joue plus son rôle, ou ne joue plus le même rôle... Ceci reviendrait à considérer les « langues » non plus comme des objets dont le linguiste décrirait la structure mais *comme d'immenses variables dont nous ne percevons bien sûr que des variantes*. Dès lors, si les locuteurs perçoivent à un moment donné dans le champ de variation d'une langue tendancielle quelque chose qu'ils ont besoin de nommer différemment (non pas sans doute la « media lingua » du linguiste européen mais le « petit quichua » des locuteurs quichua), ce serait le signe que quelque chose est en train de se passer, que le désordre serait en évolution vers un autre désordre.

Il existe dans la littérature linguistique une autre langue « mixte » qu'il peut être intéressant d'analyser, le *mitchif*, parlée dans le Manitoba, le long de la rivière Rouge par des descendants de métis de trappeurs français et de femmes cree.

L.J. Calvet

Voici tout d'abord le début d'un petit conte dans cette forme linguistique, *Les trois ours*. Les éléments issus du français sont en romaine, ceux issus du cri sont en italique :

**Trwa lii n-oor**  
**(Les trois ours)**

**Mafwee, *kayaash* il avec trwa lii n-oor. Li Paapaa, la Maaman *ekwa* li pchi**

Ma foi, autrefois, il (y) avait trois les ours. Le papa, la maman et le petit

**beebii, *oohin* lii n-oor. *Keetahtawee* en maten, la maaman *awa* li porij *kiishisham*.**

bébé ces les ours. Une fois, un matin, la maman ce le gruaou faisait.

**Li ootmiil *kii- kiishisham* ooma li manzhii, li breekfist.**

Le gruaou passé - cuire ce le manger le petit déjeuner.

***Kii- wiyashtaashoow* maaman. Tut parii *kii- ashtaw* dan la tab.**

passé - mettre - la - table maman. Tout prêt passé - mettre dans la table.

**Trwa lii z-asyet plen li ootmiil. Aben, tut *ashtaw*, *itweey***

Trois les assiettes pleines le gruaou. Et-bien, tout (est en) place (elle) dit,

**li manzhii. *Chii- patapiwak mweehchi ee- wii- miichishouchik...***

le manger. Ils-se-sont-assis juste (prêts) pour-volitionnel-manger [...].

Et voici la traduction de ce texte en français courant :

« Ma foi, il y avait une fois trois ours. Le papa, la maman et le petit bébé, c'était les ours. Une fois, un matin, la maman préparait du gruaou. Elle cuisait ce gruaou pour manger, pour le petit-déjeuner. La maman a mis la table. Elle a tout mis sur la table. Trois assiettes pleines de gruaou. Et bien, tout le manger est en place, dit-elle. Ils se sont assis, prêts pour manger [...] ».

On remarque qu'à part quelques rares emprunts lexicaux à l'anglais (ootmiil = hot meal, breekfist = breakfast) on a

L.J. Calvet

systématiquement des substantifs et des adjectifs d'origine française et des verbes cree. En fait la réalité est beaucoup plus étrange. Le Mitchif combine en effet des syntagmes nominaux français, avec une morphologie française, et des syntagmes verbaux cree, avec une morphologie cree. Comment faut-il classer cette forme ? S'agit-il d'une alternance codique systématique, syntagme (nominal) après syntagme (verbal) ? L'auteur de la présentation<sup>22</sup> du conte ci-dessus insiste sur le fait que :

« Dans la création d'une langue mixte comme le mitchif, les locuteurs sont toujours des bilingues et maîtrisent généralement les deux langues en présence. Contrairement au créole, cette langue mixte ne résulte pas d'un effort de la part d'un groupe à apprendre la « langue de l'Autre », mais plutôt du développement d'un nouvel idiome afin de se distancier de l'un et l'autre groupe. Pour les Métis, ils voulaient sensiblement se distancier tant des autochtones cris que des Blancs francophones ».

Et ceci nous incite à réfléchir en termes de fonction plutôt qu'en termes de forme. Nous sommes en effet ici confrontés moins à une forme mixte (même si l'exemple ci-dessus peut parfaitement être présenté en ces termes) qu'au produit d'une fonction *identitaire*, qu'il faut ici opposer à la fonction *véhiculaire*, la première construisant un plus petit dénominateur commun là où la seconde cherche au contraire un plus grand dénominateur commun. Quelques linguistes générativistes de Montréal ont tenté de présenter le créole haïtien comme du fon (une langue du Bénin) relexifié par le français, du fon dont seuls les signifiants seraient français, « du fon à la forme phonétique près » comme l'a écrit l'un des tenants de cette théorie. Ils n'ont pas convaincu grand monde, pour des raisons qui ont largement été développées ailleurs<sup>23</sup>, mais ce qui importe ici c'est que la distinction entre ces deux fonctions, la fonction *identitaire* et la fonction *véhiculaire*, nous permet peut-être de distinguer entre des formes linguistiques comme le mitchif et d'autres comme les créoles, dans le premier cas des « langues

---

<sup>22</sup> <http://www.tlfq.ulaval.ca/axl/index.shtml>

<sup>23</sup> Voir le numéro 8 de la revue *Plurilinguismes, Créolistique et grammaire générative*, 1994.

mixtes » et dans le second cas des « langues de contact ». La langue mixte coexisterait avec les deux langues dont elle tire ses formes et aurait une fonction identitaire (et nous sommes alors dans le cas d'une *mixité comme situation*) tandis que la langue de contact, à fonction véhiculaire, remplacerait l'autre langue (et nous serions alors dans le cas d'une *mixité comme résultat*). J'entends ici par *mixité comme situation* le fait que deux langues, A et B, coexistent en synchronie, et que les locuteurs puissent produire des énoncés mixtes, des variantes à mi-chemin entre des « langues tendanciennes », mais dans ce cas A et B continuent d'exister séparément (« exister » dans les pratiques, dans les représentations, dans la conscience des locuteurs) et c'est un troisième terme qui nous retient, comme la *media lingua* ou le *mitchif*. Et j'entends par *mixité comme résultat* le fait que deux langues A et B (ou des langues A, B, C, D, E, etc.) qui ont coexisté se soient fondues pour laisser la place à une nouvelle forme<sup>24</sup>.

### 1.3. Du désordre à l'ordre : l'obsession du linguiste

Nous verrons plus loin que dans sa volonté de mettre de l'ordre dans le désordre la linguistique a parfois tendance à dériver vers le rêve d'un ordre supérieur, d'une langue originelle, voire d'un « *intelligent design* ». Un regard sur l'histoire récente, celle pour laquelle nous avons des traces tangibles, nous donne pourtant une autre leçon.

Le français dit-on, vient du latin. Mais qu'est-ce que ces « le » (le français, le latin) impliquent ? *L'Appendix Probi* nous montre qu'il n'y avait pas « le » latin mais des latins et nous savons qu'il n'y a pas « le » français mais des français. Nous pourrions donc réécrire la proposition selon laquelle *le français vient du latin* de la façon suivante : un champ de variation « latin » a évolué vers un champ de variation « français », ou encore une variable « latin » a évolué vers une variable « français ». Et comme tout le monde sait que les locuteurs ne se sont pas couchés un soir en parlant latin pour se réveiller le lendemain en parlant français, il s'agit donc d'un désordre qui a lentement évolué vers un autre désordre.

---

<sup>24</sup> Voir L.-J. Calvet, « La mixité : situation ou résultat ? Fonction identitaire ou véhiculaire ? », conférence présentée à Tunis, université de la Manouba, 16 février 2006.

L.J. Calvet

Pour cerner l'éventuelle productivité de ces idées de « langues tendanciennes », de « champ de variation », de « désordre », tentons de soupeser la seconde question.

Le général Faidherbe (1828-1889) qui commença sa carrière en Algérie nous a laissé des exemples de « sabir », dont celui-ci :

Moi meskine, toi donnar sordi  
« Je suis pauvre, donne-moi de l'argent »

Sbanioul chapar bourrico, andar labrisou  
« L'Espagnol a volé un âne, il ira en prison »

Quand moi gagner drahem, moi achetir moukère  
« Quand j'aurai gagné de l'argent, j'achèterai une femme »<sup>25</sup>

On peut bien sûr souligner dans ce micro-corpus la présence de mots arabes (*meskine*), espagnols (*andar*, *moukère*) ou la manifestation d'interférences phonologiques avec l'arabe (par exemple *sbanioul* pour espagnol et *labrisou* pour la prison témoignant de l'absence de /p/ et de nasales en arabe). Mais ces exemples ont autre chose à nous dire. Faidherbe signale comme en passant une chose intéressante dont il ne soupçonne peut-être pas l'importance, le fait que, « *en se servant de ce langage, le troupiier est persuadé qu'il parle arabe, et l'Arabe est persuadé qu'il parle français* ». Ni le troupiier ni l'Arabe ne baptiseraient donc cette pratique « sabir », ils sont tous les deux dans des langues tendanciennes, à la recherche d'une approximation d'arabe ou d'une approximation de français comme dirait Robert Chaudenson, ces deux recherches ou ces deux langues tendanciennes convergeant vers quelque chose que d'autres baptisent « sabir ». Nous sommes alors confrontés ici à deux choses différentes, des pratiques ou des productions d'une part, qui constituent une volonté de communication, et d'autre part une classification par un observateur (« ces pratiques sont du sabir »). On voit que nous ne sommes pas très loin de l'exemple Mauricien présenté plus haut (la liste d'items sur un congélateur), que nous avons dans tous

---

<sup>25</sup> Louis-Léon-César Faidherbe, 1884, "L'Alliance française pour la propagation de la langue française dans les colonies et les pays étrangers", in *Revue scientifique*, 3<sup>ème</sup> série, 7.



L.J. Calvet

les cas des langues *tendanciennes* et des pratiques tendant vers l'une ou l'autre de ces langues.

Qu'est-ce que ces exemples, et leur mise en perspective romanesque (Stendhal, Tolstoï), signifient pour ce qu'il est convenu d'appeler la linguistique ? Commençons par une évidence : les sciences sont des pratiques sociales, des produits d'un moment de l'histoire, et le scientifique, le chercheur, peu importe comment on l'appelle, est lié à son temps, à son époque, à la société dans laquelle il travaille et qu'il accompagne, ainsi qu'à son histoire personnelle. Il est donc légitime de s'interroger sur les rapports entre cette science qu'est ou que veut être la linguistique et le temps qui l'a vue naître, ce que j'avais en 1974, dans *Linguistique et colonialisme*, exprimé de façon un peu abrupte :

« 'Chaque siècle a la grammaire de sa philosophie', écrivait Antoine Meillet. Cette proposition, on l'aura compris, nous paraît très incomplète et, par souci de simplification, c'est par la suivante que nous la remplacerons pour conclure : chaque société a la linguistique de ses rapports de production » (p. 39).

L'histoire de cette science (ou l'histoire de la volonté de faire exister cette science, ou encore l'histoire de la volonté de transformer en science des discours épilinguistiques divers) est récente : à peine un siècle. Et l'on peut comprendre que pour asseoir la scientificité d'un discours, il fallait lui donner un objet indiscutable, cernable, homogène : c'est ce qui explique en grande partie la posture choisie par Ferdinand de Saussure lorsqu'il enseigna ce qui est devenu le *Cours de linguistique générale*. Mais si comprendre, c'est expliquer, ce n'est pas pour autant admettre. Car une autre approche était possible et nous devons simplement constater qu'elle n'a pas été utilisée : la linguistique embryonnaire a décidé qu'il fallait croire aux langues. Pourquoi pas, si cela devait être un stade transitoire ? L'ennui, c'est que la plupart des linguistes y croient toujours. Après avoir mis au point des instruments d'analyse (phonologie, syntaxe...) ils sont entrés dans un cercle vicieux, s'intéressant à des aspects de plus en plus étroits, microscopiques, et perdant en cours de route la réalité des situations de communications. La grammaire générative est, de ce point de vue, une caricature : plus de locuteurs (sinon un « locuteur idéal », le plus souvent ramené aux intuitions du linguiste qui décrit ainsi sa propre pratique et fait

l'économie du travail de terrain et de la constitution de corpus), plus de groupes sociaux, plus de société, mais une mécanique que l'on s'applique à démonter, pour montrer comment elle fonctionne. Certains linguistes font ainsi penser à ces soldats sans bataille qui passent leur temps à démonter, graisser puis remonter leur fusil-mitrailleur. Mais, même si les mots peuvent parfois tuer, les *langues tendanciennes* dont j'ai parlé ne sont ni des fusils mitrailleurs ni des mécaniques bien huilées, elles ne sont pas des ensembles homogènes et réguliers, et ceux qui le pensent travaillent en fait sur des artefacts, des objets produits par la description afin de conforter les méthodes de description, sur ce que j'ai appelé des O.L.N.I., des Objets Linguistiques Non Identifiés. L'artefact est ainsi devenu, dans l'esprit de certains linguistes, une réalité, et la linguistique a perdu ses marques, elle stagne, elle qui fut dans les années 1960 et 1970 une science pilote. Confrontée à un paradoxe (ou à une contradiction) qui consiste à reconnaître le plus souvent le caractère social de la langue et à faire abstraction du social dans ses procédures de description, elle semble être entrée en crise.

Une tendance récente risque d'ailleurs de mener cette quête forcenée de l'homogène sur des chemins inattendus. On voit en effet percer, dans le débat sur l'origine des langues, une sorte d'alliance clandestine avec les tenants de l'*intelligent design*, qui se gardent bien de souligner en quoi leurs positions vont contre les théories de Darwin mais n'en développent pas moins un discours créationniste. La paléontologue française Anne Dambricourt a par exemple récemment développé une théorie selon laquelle l'évolution de l'homme n'était pas le fruit de l'adaptation au milieu mais plutôt celui d'une « inside story », une mutation de l'ADN et une courbure corrélative de l'os sphénoïde correspondant à un arrondissement du crâne. Le lien entre cette théorie et l'« intelligent design » développé par la droite religieuse américaine est évident et son auteur s'en cache à peine : il s'agit là d'une machine de guerre anti-darwiniste qui ne dit pas son nom.

Or il en va de même pour les théories sur l'origine des langues et la « nouvelle synthèse » que l'on voit se manifester depuis quelques années. Merritt Ruhlen d'une part qui propose vingt sept racines communes à toutes les langues du monde, racines qui proviendraient donc d'une langue originelle commune, hypothèse bien entendu invérifiable et qui relève plus de l'idéologie que de la reconstruction.

L.J. Calvet

Luigi Cavalli-Sforza d'autre part qui s'appuie sur l'analyse de l'ADN mitochondrial pour tenter de démontrer l'existence d'une correspondance entre généalogie linguistique et généalogie génétique. Renfrew qui privilégie une seule technologie humaine, l'agriculture, pour tenter de démontrer l'existence d'une corrélation entre les progrès de l'agriculture et la multiplication des langues. Et le linguiste français Jean-Marie Hombert qui cautionne le tout en parlant d'une « surprenante corrélation entre les classifications s'appuyant sur des marqueurs linguistiques et celles obtenues à partir de marqueurs génétiques »<sup>26</sup>. La recherche de l'*ordre* change alors de genre, ou de niveau, c'est d'un ordre supérieur qu'il s'agit désormais, celui qui a programmé l'évolution de la courbure de l'os sphénoïde, ou qui a implanté quelque part chez nos ancêtres le noyau d'une langue commune. J'ai longuement critiqué dans mes *Essais de linguistiques* cette « nouvelle synthèse » autoproclamée, et Sylvain Auroux a de son côté des arguments forts :

« Les hongrois sont génétiquement apparentés à leurs voisins européens, leur langue (famille finno-ougrienne) ne l'est pas ; les pygmées n'ont rien à voir génétiquement avec leurs voisins bantous, mais ils parlent des langues bantou »

et une conclusion sans appel :

« Il y a plus d'effets d'annonce dans la 'nouvelle synthèse' que de vrais résultats scientifiques »<sup>27</sup>.

Mais l'idée d'effets d'annonce, si elle permet de comprendre l'écho donné par la grande presse, ou la presse de vulgarisation, à cette « nouvelle synthèse », ne suffit pas à expliquer pourquoi certains linguistes se sont précipités sur ce leurre. Encore une fois, en effet, c'est la belle régularité d'une pseudo théorie qui l'a emporté sur la complexité des faits. Les Brésiliens sont pratiquement tous locuteurs du portugais, mais ils sont génétiquement apparentés les uns aux Portugais, les autres aux Japonais, d'autres encore aux Africains ou aux Indiens d'Amazonie,

---

<sup>26</sup> Hombert J.-M., « Qui de la langue ou du gène influence l'autre ? », *Sciences et avenir* n°125, janvier 2001, 47.

<sup>27</sup> Auroux S., « Les embarras de l'origine des langues », *Marges Linguistiques* n°11, mai 2006, 85.

et il est donc douteux que les lusophones du Brésil présentent le même ADN. Cela s'explique bien sûr par une histoire relativement récente, tout au plus cinq siècles. Mais il n'y a aucune raison de penser que ce type de brassage de population n'a pas toujours été la règle. Même s'il y avait eu à un moment de l'histoire un lien entre gènes et langues (ce qui reste à prouver), le désordre engendré par l'histoire elle-même le rendrait aujourd'hui illisible.

#### 1.4. *La tentation de la loi*

Dans cette tentative constante et désespérée de mettre de l'ordre dans le désordre, de refuser de prendre en compte ce qui est la caractéristique principale des pratiques linguistiques, c'est-à-dire leur constante variabilité, il faut évoquer la question des lois.

Une loi scientifique est la traduction (nécessairement un peu tautologique) de ce qu'on observe, de certaines régularités (comme la loi de la gravitation universelle, la loi du mouvement périodique des pendules, etc.). A partir d'observations on formule donc des hypothèses qui, si elles sont « universelles », confirmées par d'autres observations, sont considérées comme des « lois », l'idéal étant de rendre compte du plus grand nombre possible de faits observables par le plus petit nombre possible de lois. Corrélativement, pour être acceptable, une loi doit pouvoir être *falsifiable*. Cela ne signifie naturellement pas qu'elle soit immédiatement invalidée par des contre-exemples, mais qu'elle se prête à une démarche de vérification, que sa formulation permette une telle démarche.

Le mot *loi* apparaît en linguistique au XIX<sup>ème</sup> siècle, à propos de l'évolution phonétique. Jacob Grimm propose en 1822 d'expliquer les correspondances entre les langues germaniques par une série de mutations : les aspirées de l'indo-européen devenant des non-aspirées, les sonores devenant des sourdes et les sourdes devenant des aspirées. Derrière ces « lois » il y a donc l'idée qu'un phonème d'une langue donnée, dans un environnement phonétique donné et à une époque donnée, connaît le même changement dans tous les mots de la langue. Par exemple, un /k/ initial latin restera toujours un /k/ en italien, en espagnol, en provençal, etc., mais passera à /ch/ en français. Ainsi l'évolution phonétique de la première syllabe du latin *caballus* vers le

L.J. Calvet

français *cheval* est généralisable dans cette langue (*camelus* a donné en français *chameau*, *campus* a donné *champ*, *capellus* a donné *cheveu*, *calvus* a donné *chauve*, etc.) et dans d'autres langues romanes (la consonne initiale de *caballus* reste la même dans l'espagnol *caballo*, le portugais *cavalho*, etc.).

Cette idée de «lois phonétiques» permettait donc à la fois de présenter de façon élégante l'évolution par exemple du latin vers les langues romanes (c'est-à-dire tout simplement de mettre en forme des faits observables) et de reconstruire des langues disparues comme l'indo-européen : tautologie dans le premier cas, hypothèse dans le second. Et ces lois, bien sûr, n'étaient pas prédictives, elles ne nous disaient pas ce que deviendraient dans l'avenir, même dans un avenir proche, les sons d'une langue. Encore une fois, dans une science d'observation comme la linguistique (que j'oppose ici aux sciences expérimentales), une loi n'est que la mise en forme hypothétique d'une série de faits observés. Et la force de ces lois phonétiques venait du fait qu'elles étaient *falsifiables* : il était possible de les passer au filtre d'un grand nombre d'exemples, de vérifier par exemple dans les différentes langues romanes que le /k/ latin avait bien le devenir que disait la loi.

Peut-être grisés par le succès de la phonologie (voir plus haut ce que j'écrivais sur le rôle de Claude Levi-Strauss dans la construction du mythe de la linguistique comme science pilote) certains sont partis en quête de nouvelles lois, en faisant du terme une notion à géométrie variable. André Martinet par exemple évoquait souvent dans ses séminaires ce qu'il appelait la « loi du moindre effort » pour présenter l'évolution du lexique (haplogogie, chute des finales...) et pour expliquer les changements syntaxiques (la disparition des marques casuelles impliquant une réorganisation de la phrase). Mais le terme *loi* était ici légèrement abusif. Si nous considérons que l'une des fonctions des langues est la communication, alors la théorie de l'information nous dicte une hypothèse raisonnable : le début d'un mot est plus « préservé » de l'usure que sa fin. Avant qu'un locuteur ouvre la bouche, nous pouvons attendre de lui une infinité de messages. Mais dès lors qu'il a prononcé une syllabe, il a effectué un choix entre toutes les syllabes possibles de la langue française, entre toutes les combinaisons possibles de consonnes et de voyelles dans le cadre des combinaisons acceptables (CV, CCV, V, VC...). Cette première syllabe est donc chargée d'une

énorme information. Considérons qu'il a prononcé *mé*. Les choix offerts à lui (ou à l'auditeur qui tente d'anticiper ce qu'il va dire) sont alors contraints : il ne peut poursuivre qu'avec la seconde syllabe de tous les mots français qui débutent par *mé*. C'est-à-dire que l'information apportée par cette première syllabe est énorme, mais que la seconde syllabe ne constituera un choix qu'entre environ trois cents possibilités. Si cette seconde syllabe est *tro*, alors *métro* ne peut se poursuivre que dans moins d'une dizaine de mots différents (*métrologie, métrologique, métromanie, métronome, métropole, métropolitain, métropolite, métrorrhagie*). Si la troisième syllabe est *po*, alors *métropo* n'ouvre que trois possibilités, *métropolite, métropolitain, métropole*. Dans les cinq syllabes de *métropolitain*, S1 S2 S3 S4 S5, la première, S1, apporte donc beaucoup plus d'information que S2 qui en apporte plus que S3 qui en apporte plus que S4 qui en apporte plus que S5. Si nous prenons donc en compte la fonction communicative, alors S1 a plus de poids que S2 qui a plus de poids que S3 qui a plus de poids que S4 qui a plus de poids que S5<sup>28</sup>. Et il est logique que la « loi du moindre effort » commence par la fin du mot, par S5... Mais si nous prenons en compte la fonction cryptique dont témoignent certains usages, alors il se passera l'inverse : c'est S1 qui sera l'objet des modifications ou de la suppression (on peut songer ici au verlan, qui dans un mot de deux syllabes transforme S1 S2 en S2 S1, ou à un code comme le *largonji* qui va transformer un mot de deux syllabes, C1V1 C2 V2 en LV1 C2Ji, *jargon* donnant *largonji*). Apocope et aphérèse correspondent donc à deux fonctions différentes, mais il en est d'autres possibles. La fonction identitaire, la volonté de marquer ses racines, peut par exemple avoir des effets sur le signifiant des signes linguistiques. Un Marseillais prononcera ainsi *aïoli* avec un accent tonique sur l'avant-dernière syllabe, là où un Parisien le mettra sur la dernière, *aïoli*. Et le même Marseillais mettra le même accent sur la même syllabe des mots d'origine provençale se terminant par i, **p**ati, **b**abi, allant parfois jusqu'à faire la même chose, par plaisanterie, avec des mots qui ne sont pas d'origine provençale : *bon appétit*...

Il est donc difficile de parler de « loi du moindre effort » comme principe de changement ou d'évolution, puisqu'il y a d'autres principes (d'autres fonctions) qui sous-tendent le changement. Et nous sommes ici

---

<sup>28</sup> On peut ici étudier le système de réservation de billets sur les machines de la SNCF.

confrontés à la question suivante : *lois ou tendance* ? En fait, la variation constante des pratiques linguistiques interdit, lorsque l'on prend réellement en considération cette diversité, de parler de lois. Nous ne pouvons discerner que des tendances, ce qui devrait nous inciter à une certaine prudence, ou à une certaine modestie, et à conclure que *la linguistique n'est pas tout à fait une science exacte...*

Mais, encore une fois, le choix de l'ordre, ou la volonté de mettre de l'ordre dans le désordre, nous pousse parfois à des classements abusifs. Le verlan ou le largonji seront classés comme marginaux, l'apocope sera considérée comme la règle, etc. : la tentation de la loi implique que l'on rejette une partie des faits observables. La grammaire générative atteint, en la matière, des sommets, pratiquant le *principe de la cafétéria*. De la même façon que le client qui, son plateau en main, choisit les plats qui lui conviennent, ceux dont il a envie, le générativiste choisit les exemples linguistiques qui lui conviennent ou qui confortent sa théorie et rejette les autres. Ce pourquoi les exemples qu'elle propose sont, je l'ai dit, essentiellement des OLNI, des « objets linguistiques non identifiés ».

On peut certes opposer à cette image du *principe de la cafétéria* un fait tout simple. Le client, son plateau en main, choisit certes ce dont il a envie, comme certains linguistes sélectionnent les exemples qui les arrangent, mais il choisit le plus souvent en fonction d'une *syntaxe* : en France par exemple il prendra un hors-d'œuvre, un plat et un dessert. Mais rien ne l'empêche de prendre seulement deux ou trois hors-d'œuvre, ou de prendre deux plats, ou encore six desserts. Et si ces comportements gastronomiques sont statistiquement extrêmement minoritaires, nul n'a le droit de les écarter de son champ d'observation, ou de les qualifier d'*agrammaticaux*. La *syntaxe* du repas, cependant, est une tendance, parfois forte, mais pas une loi. Il est statistiquement vrai, dans la culture française, que le client aura tendance à constituer un « syntagme gastronomique » en choisissant dans trois paradigmes une entrée, un plat et un dessert. Mais ceci est falsifié, dans la culture française, par la pratique du buffet, dans d'autres cultures, parfois importées en France, cette syntaxe est remplacée par d'autres pratiques, les mezzés, le plat unique, etc. Encore une fois, la tentation de l'ordre, de la loi, voudrait que la syntaxe du repas soit constituée par une succession immuable, un hors-d'œuvre, un plat et un dessert (précédé ou pas d'un

L.J. Calvet

fromage), mais les pratiques s'inscrivent parfois (souvent ?) en faux contre cette idée. Et la réalité a la peau dure...

### 1.5. Combien de langues ?

Malgré cette recherche permanente de l'ordre sous le désordre, la linguistique apparaît parfois comme elle-même désarmée, impuissante, face à des questions apparemment anodines.

On sait (ou on admet) qu'il y a environ sept mille langues sur la surface du globe. Ce décompte n'est pas une « découverte » mais plutôt une création de la science. En effet, tant que la langue n'était pas définie comme *orale*, tant que ce qui n'avait pas de grammaire écrite, de corpus littéraire, etc., n'était pas considéré comme *langue*, le nombre de « langues » était beaucoup plus restreint, le reste étant considéré comme « dialectes », « parlers » voire comme « patois ». Nous sommes ici confrontés à un processus que l'on peut présenter historiquement, en simplifiant beaucoup, en deux étapes principales :

- Tout d'abord ce que Sylvain Auroux a appelé la *grammatisation*, lorsqu'il écrit par exemple :

« La création d'*outillages linguistiques* (depuis l'écriture, jusqu'aux grammaires et aux dictionnaires) ou *grammatisation* a considérablement changé l'écologie de la communication. Les grandes langues de culture sont en quelque sorte des artefacts, des produits de l'outillage dans un contexte particulier, celui du monolinguisme des Etats-Nations »<sup>29</sup>.

Cette première étape a constitué une véritable révolution, créant dans l'ensemble des pratiques communicatives orales une distinction entre celles d'entre elles, peu nombreuses, qui étaient aussi écrites, et les autres.

- La seconde étape se situe au XX<sup>ème</sup> siècle, au moment où la

---

<sup>29</sup> Auroux S., « Introduction », *L'hyperlangue brésilienne*, *Langages* n°130, juin 1998.



L.J. Calvet

linguistique définit la langue comme *orale*. Dès lors la distinction opérée par la *grammatisation* est relativisée, les anciens dialectes sont considérés comme des langues dont le décompte est modifié, et l'on peut annoncer l'existence de plusieurs milliers de « langues ».

On voit donc que la *grammatisation* puis l'accent mis sur l'oralité constituent deux moments qui ont joué ici un rôle fondamental : le nombre des « langues » du monde est un sous-produit de la linguistique, un artefact, puisqu'avant de compter il faut bien décider de ce qui « mérite » d'être compté. Et, en décidant de considérer les langues comme des pratiques *orales*, la linguistique en a *ipso facto* augmenté considérablement le nombre. Ceci appartient bien sûr à l'histoire : c'est la théorie linguistique qui a produit le décompte des langues du monde sur lequel nous sommes tous aujourd'hui à peu près d'accord et à partir duquel nous posons par exemple des problèmes de politique linguistique ou des questions sur les retombées linguistiques de la mondialisation.

Philippe Blanchet en donne un exemple intéressant dans ce numéro, lorsqu'il met en évidence le fait que deux approches (qu'il nomme « structurilinguistique » et « sociolinguistique ») identifient face aux mêmes pratiques l'une l'*occitan* et l'autre le *provençal*, le *niçois*, le *patois*, le *gascon*, etc., et débouchent sur des politiques linguistiques différentes. De la même façon, lorsque l'on parle de *langues internationales*, on songe le plus souvent à celles qui sont reconnues comme langues officielles ou de travail dans les organisations internationales (le français, l'anglais, l'arabe, le russe, le chinois, l'espagnol...) et l'on oublie celles qui sont utilisées par plusieurs nations (le swahili, le malais, le bambara, etc.) et sont donc également « internationales ». C'est le point de vue qui crée l'objet et ici la linguistique (ou du moins un certain état de la linguistique) qui crée les langues.

Mais nous n'en avons pas fini pour autant avec cette question. On s'accorde donc à considérer qu'il existe aujourd'hui environ sept mille langues. Le site du Summer Institute of Linguistics, *Ethnologue*, en comptabilisait 6703 en 2001, 6800 en 2003 et 6912 en 2006. Cette augmentation régulière n'est bien sûr pas le produit d'une soudaine multiplication des langues (il n'est pas à exclure qu'il en apparaisse de nouvelles, mais pas deux cents en cinq ans) mais témoigne plutôt de

changements dans les grilles d'analyse du SIL, qui classe comme langues différentes ce qu'il classait auparavant comme dialectes d'une même langue. Reste une autre question : quelles sont, parmi toutes ces langues, celles qui sont les plus parlées ?

On trouvera ci-dessous trois classements provenant du SIL des douze premières langues du monde (par ordre décroissant de nombre de locuteurs). La première colonne (1992) provient du site *Principales langues du monde*<sup>30</sup> qui utilise les données du SIL, la seconde (1996) du site de l'université Laval qui reprend les chiffres du SIL et la troisième (2006) a été établie par mes soins en août 2006 à partir des chiffres donnés par *Ethnologue* pour chacune des langues concernées. Nous avons donc là une sorte de « top 12 » des langues du monde établi par le même organisme à trois dates différentes, et j'ai simplement ajouté entre parenthèses le nombre de locuteurs (en millions), toujours selon le SIL.

	1992	1996	2006
1	Chinois (885)	Mandarin (885)	Mandarin (867)
2	Anglais (450)	Espagnol (358)	Espagnol (322)
3	Hindi-ourdou (333)	Anglais (322)	Anglais (309)
4	Espagnol (266)	Arabe classique (200)	Arabe standard (206)
5	Portugais (175)	Bengali (189)	Hindi (180)
6	Bengali (162)	Hindi (182)	Portugais (177)
7	Russe (153)	Portugais (170)	Bengali (171)
8	Arabe (150)	Russe (170)	Russe (145)
9	Japonais (126)	Japonais (125)	Japonais (122)
10	Français (122)	Allemand (98)	Allemand (95)
11	Allemand (118)	Coréen (78)	Coréen (67)
12	Wu (77)	Français (77)	Français (64)

On constate dans ce tableau un certain nombre de variations d'une colonne à l'autre.

Les unes concernent le nombre de locuteurs attribués aux langues. On voit qu'ils peuvent être à la baisse ou à la hausse et que ces changements ne peuvent pas correspondre à des variations de population (sauf peut-être pour le japonais : la population du Japon a tendance à

<sup>30</sup> <http://alis.isoc.org/langues/grandes.htm>

décroître) mais plutôt à des rectifications des grilles d'analyse. Le cas de l'anglais et du français sont de ce point de vue significatifs. Crédité de 450 millions de locuteurs en 1992, l'anglais est ramené à 322 millions en 1996 et à 309 millions en 2006, tandis que le français passe de 122 millions de locuteurs à 77 puis à 64. Dans les deux cas il paraît évident que l'on est passé de la prise en compte de tous les locuteurs de ces langues à la seule prise en compte de ceux qui les ont comme langues premières. Et il semble s'être passé la même chose avec le russe, l'allemand, le hindi, etc. C'est-à-dire qu'en quelques années, face à une même question (« quelles sont les langues les plus parlées dans le monde ? ») on a choisi de répondre à partir de critères différents : les plus parlées comme langues véhiculaires puis comme langues premières. Ce faisant, on voit que le décompte du SIL modifie l'ordre des langues et, de façon un peu inattendue, fait passer l'espagnol avant l'anglais. Mais il faut en même temps souligner que ces chiffres ne correspondent pas à ceux d'autres évaluations. A la même date (août 2006) l'Union Latine par exemple donne sur son site 375 millions de locuteurs de l'espagnol (contre 322 pour le SIL), 130 millions pour le français (contre 64 selon le SIL) et 630 millions pour l'anglais (face aux 309 du SIL). Et derrière ces variations qui ne sont pas mineures apparaissent donc des questions importantes concernant les critères selon lesquels on classe les langues.

Une seconde variation concerne ce qui est baptisé en 1992 hindi-ourdou et devient ensuite hindi, passant de 333 à 182 puis à 180 millions de locuteurs. Cette chute spectaculaire s'explique par deux facteurs différents. D'une part hindi et ourdou sont désormais considérés par le SIL comme deux langues séparées. Le hindi a, selon le site *Ethnologue*, 180 millions de locuteurs en Inde et quelques centaines de milliers dans différents pays, l'ourdou en ayant pour sa part 10 millions au Pakistan et 62 millions en Inde. En additionnant ces chiffres nous obtenons environ 255 millions de locuteurs, c'est-à-dire moins que le chiffre de 1992 (333), puisqu'il s'agit là de locuteurs langue première. Le site annonce par ailleurs 120 millions de locuteurs du hindi seconde langue, ce qui nous mène alors à 375 millions, explicables par la croissance démographique. Mais ces approximations font problème. Le sort réservé au hindi en 2006 donne peut-être une image faussée de la situation sociolinguistique de l'Inde, où cette langue joue un rôle véhiculaire non négligeable. Par ailleurs, le même site annonce à l'entrée

*hindi* que «Hindi, or Hindi-Urdu, has four varieties... », reprenant donc l'appellation de 1992 abandonnée depuis, ces quatre variétés étant le hindi, l'ourdou, le dakhini et le rekhta. Et à l'entrée *ourdou* on apprend que la langue est parlée au Pakistan, en Inde et dans une vingtaine d'autres pays, et on lit en outre : «Intelligible with Hindi, but has formal vocabulary borrowed from Arabic and Persian. Dakhini is freer of Persian and Arabic loans than Urdu. Rekhta is a form of Urdu used in poetry ». On ne comprend alors plus rien. Le rekhta par exemple est classé d'un côté comme dialecte du hindi et de l'autre côté comme forme poétique de l'ourdou qui lui-même apparaît d'un côté comme langue et de l'autre comme dialecte du hindi... Ajoutons, pour couronner le tout, que selon l'encyclopédie en ligne Wikipédia, 360 millions de gens ont le hindi pour première langue et 480 le parlent (alors que les chiffres du SIL sont respectivement 180 et 300...).

Une troisième variante intéressante concerne le sort réservé à l'arabe car non seulement son nombre de locuteurs augmente mais son nom change : *arabe* puis *arabe classique* puis *arabe standard*. La seconde appellation, en 1996, semble paradoxale : comment une langue « classique » pourrait-elle avoir des locuteurs ? Mais la troisième ne l'est pas moins. En effet, *Ethnologue* distingue en fait entre une quarantaine d'arabes différents (l'égyptien avec 44 millions de locuteurs, l'algérien avec 20 millions, le marocain avec 18 millions, le soudanais avec 15 millions, etc., position d'ailleurs tout à fait acceptable lorsqu'on connaît la situation de l'arabe) auxquels il ajoute « standard arabic » en précisant « 206.000.000 first-language speakers of all Arabic varieties ». C'est-à-dire qu'il crée un « arabe standard » par addition des locuteurs de tous les différents arabes.

Restons-en là, car ces quelques exemples nous montrent une grande hésitation, ou de nombreuses fluctuations, dans la façon de répondre à des questions comme :

- *Qu'est-ce qu'une langue ?* (Y a-t-il un arabe ou des arabes ? L'hindi et l'ourdou sont-ils deux formes de la même langue ou des langues séparées ?)

- *Qu'est-ce que parler une langue ?* (Les locuteurs langue seconde doivent-ils être comptabilisés ?)

et donc

- *Qu'est-ce qu'une communauté linguistique ?*

Mon propos n'est pas ici de critiquer le Summer Institute of Linguistics (il y a bien d'autres raisons de le critiquer, politiques celles-ci, qui n'ont rien à faire ici), mais de montrer tout simplement que la « réalité » n'est pas facile à appréhender, que l'on choisit de façon parfois arbitraire entre différentes possibilités (par exemple hindi et ourdou ou hindi-ourdou) et que cela peut mener à des contradictions. Encore une fois, le désordre des données et la volonté de l'ordre... J'ai surtout insisté dans ce qui précède sur des problèmes comptables, mais cette tension entre le désordre des données et la volonté d'y mettre de l'ordre se manifeste dans bien d'autres domaines, de la syntaxe à la phonologie en passant par la sémantique et quelques autres petites choses.

## **2. Vers un désordre assumé**

### *2.1 Le piège de la sociolinguistique*

Cette linguistique de l'homogène a donc eu besoin, pour survivre, d'homogénéifier aussi la société. Le lexique témoigne parfois de cette tendance. Dans les vieux états, aux frontières établies depuis longtemps, on aligne le nom des langues sur ceux du pays et des citoyens. En France il y a des Français qui parlent français, en Espagne des Espagnols qui parlent espagnol, en Allemagne des Allemands qui parlent allemand, en Italie des Italiens qui parlent italien... Comme si l'on ne parlait que l'espagnol en Espagne, que le français en France ! Comme si l'on pouvait passer sous silence non seulement les autres langues de ces territoires, endogènes ou issues des migrations, mais aussi, surtout peut-être, les différentes formes de la même « langue ». Pris au piège qu'elle s'était elle-même tendu, la linguistique n'a pu que se renforcer dans son autisme, déléguant ce qui relevait à ses yeux d'une approche périphérique à la « sociolinguistique ». Et le résultat est là : le délitement d'une science qui fut prometteuse, un enfermement dans une approche à courte vue, avec la complicité des « sociolinguistes » qui, assumant cette

L.J. Calvet

approche « périphérique » justifiaient du même coup l'existence d'une linguistique « centrale ». Je parle ici pour ma propre chapelle, étant comme d'autres tombé dans le piège en tentant de développer une critique de la linguistique « dure » en me situant dans ses périphéries et non pas en son centre. Disons que j'essaie, depuis quelques années, de m'en guérir.

Ce partage des tâches, imposé par les uns et accepté par les autres, a eu plusieurs retombées négatives. D'une part, il justifiait et confortait une certaine forme de linguistique aveugle aux déterminations sociales : occupez-vous de vos phonèmes, de vos syntagmes, de votre intonation, de vos verbes, nous nous occupons du reste... D'autre part il enfermait du même coup les « sociolinguistes » dans un rôle marginal. En outre il confirmait que la langue pouvait être un objet homogène, étudiable comme objet. Mais surtout, il rendait impossible une science linguistique allant jusqu'au bout de l'affirmation, maintes fois répétées, que la langue était un « fait » ou un « produit » ou une « institution » sociale.

En d'autres termes, après le suicide de la linguistique (après Austerlitz, Waterloo, ou après le zénith le nadir), suicide dans lequel, en France, le CNRS a joué un rôle central en privilégiant des programmes de recherches de moins en moins liés aux pratiques linguistiques réelles, favorisant des lubies ou des modes sous la pression de différents « lobbies ». Ce partage des tâches programmait à terme le suicide de la « sociolinguistique » que cette linguistique avait suscitée pour pouvoir survivre vaille que vaille.

Il faudrait aussi évaluer ici la stratégie d'un des phares de la sociolinguistique, William Labov. Lors de la réunion de 1964 à l'UCLA que l'on considère comme la « naissance » de la sociolinguistique et qui s'est en tout cas prolongée par la publications d'actes portant pour la première fois le titre de « sociolinguistique » (Dell Hymes, 1967), les différents participants étaient dans leur grande majorité réunis par leur opposition aux théories de Chomsky. Le plus simple est ici de citer un article que j'ai consacré à cette réunion et pour lequel j'avais interrogé la plupart des témoins encore vivants de cette réunion :

« Si les participants de la réunion constituent un groupe ou un réseau, leur situation est en fait plus défensive qu'offensive, car ils se sentent menacés par une théorie en

L.J. Calvet

pleine ascension, celle de Noam Chomsky. William Samarin écrit : « *I would have at that time considered myself simply a linguist, aware, however, that my kind of linguistics was already in danger of being washed away by Noam Chomsky and his transformational-generative kind of linguistics* » (« Je me serais simplement considéré à cette époque comme un linguiste, conscient cependant que ma linguistique était déjà menacée d'être balayée par Chomsky et sa linguistique transformationnelle-générative »). John Gumperz m'explique qu'il voulait se démarquer de ceux qui mettaient l'accent principalement sur les structures linguistiques, pensant pour sa part qu'il fallait mettre l'accent sur les communautés linguistiques et sur leur répertoire. Seul Dell Hymes m'écrit qu'il profite de sa présence à l'UCLA pour aller écouter une conférence de Chomsky, ce qui ne l'empêche pas de dire à propos de la réunion « *It was a good representation, good range of interests in language beyond formal grammar* » (« C'était une bonne représentation, un bon échantillon des intérêts pour la langue au delà de la grammaire formelle ») et de déclarer dans un autre email (9 février 1999) : « *At the same time, as you may know, I was writing about 'communicative competence' in reaction to Chomsky's use of 'competence'* » (« A la même époque, comme vous le savez peut-être, j'écrivais sur la 'compétence communicative' en réaction à l'usage que Chomsky faisait de 'compétence' »).

C'est-à-dire que Noam Chomsky, qui n'est bien sûr jamais cité dans les textes de 1964, et de façon plus générale la grammaire générative, planent comme une ombre sur la réunion. C'est contre la menace d'une linguistique de plus en plus formelle et de moins en moins « humaine » que ces gens se réunissent, et l'on peut prendre l'hypothèse que la sociolinguistique est née en partie *contre* le générativisme, contre une certaine idée de la linguistique s'éloignant de plus en plus du contexte social<sup>31</sup>. »

Les participants à cette réunion étaient au nombre de vingt-six

---

<sup>31</sup> « Aux origines de la sociolinguistique, la conférence de sociolinguistique de l'UCLA (1964) », *Langage et Société* n°88, 1999, 45-46.

dont treize avait présenté une communication. Mais Labov avait parmi eux un statut un peu particulier : il était là parce qu'Uriel Weinreich, son directeur de thèse, était malade (il devait mourir trois ans plus tard, à quarante et un ans) et l'avait fait inviter à sa place. Né en 1927, du même âge que Bright (1928), Friedrich (1927), Garfinckel (1929), Hymes (1927) ou Samarin (1926), Labov était cependant plus « jeune » qu'eux, du point de vue académique, car il avait travaillé pendant dix ans comme ingénieur chimiste avant de se lancer dans la linguistique et venait à peine de terminer son master. Il avait donc le même âge que Hymes, mais celui-ci était déjà connu et cité tandis que Labov était l'étudiant prometteur d'Uriel Weinreich (qui n'avait qu'un an de plus que lui). Or Weinreich, contrairement nous l'avons vu à la majorité des participants à cette réunion, était séduit par le générativisme. Il avait tenté à la fin des années 1950 de faire publier dans la revue *Word* le premier texte de Chomsky (Martinet, directeur de *Word*, s'y était refusé), il avait lui-même travaillé en sémantique du point de vue générativiste et Labov, son étudiant, était un peu sur les mêmes positions. Une partie de son œuvre (en gros ce que Pierre Encrevé a appelé le « second Labov ») montre qu'il s'est en fait rêvé comme le complément sociolinguistique de Chomsky, comme celui qui apporterait à la grammaire générative l'aspect social qui lui manquait. Il se trouve que Chomsky l'a proprement envoyé sur les roses en lui opposant une fin de non recevoir hautaine. Interrogé en 1977 par Mitsou Ronat, il déclare à propos de Labov et de la sociolinguistique :

« Le seul type de propositions qui soit venu de la dite sociolinguistique, c'est que le discours d'un individu ne consiste pas en l'interaction de systèmes idéaux, mais en un seul système, avec quelques variantes marginales. Si c'est cela, ce n'est pas très intéressant [...] l'existence d'une discipline nommée « socio-linguistique » reste pour moi chose obscure »<sup>32</sup>.

Labov s'est alors trouvé piégé. Plutôt que de chercher à construire cette sociolinguistique qui, disait-il, était **la** linguistique, il avait tenté d'être reconnu par les générativistes, d'être admis dans leur groupe de pensée, et il était orphelin de leur reconnaissance. Disons que sa

---

<sup>32</sup> *Dialogues avec Mitsou Ronat*, Paris, Flammarion, 1977, 72 et 74.



L.J. Calvet

stratégie théorique avait échoué. C'est pourquoi je plaide depuis longtemps contre un tel partage des tâches et pour l'idée que rien de linguistique ne doit être étranger à la « sociolinguistique », qu'elle doit *aussi* entrer dans le détail des descriptions formelles, mais d'un autre point de vue, pour devenir enfin *la* linguistique, dans le cadre par exemple de ce que j'ai appelé une approche allant de l'analogique au digital (voir plus loin).

Tout cela est bel et bon, dira-t-on, mais que faire et comment le faire ? Nous sommes en effet confrontés à deux problèmes difficilement séparables : celui d'une théorie nouvelle à élaborer, celui de descriptions concrètes à proposer. Ce qui implique du même coup que nous y voyions clair sur les rapports entre théorie et descriptions. Faut-il accumuler des descriptions et en tirer une théorie ? Ou à l'inverse sortir une théorie *ex nihilo* et montrer comment, ô surprise, elle s'applique merveilleusement aux situations ? Le marxiste non repent que je suis, qui pense que les faits ont la vie dure et qu'il faut tenter de faire « l'analyse concrète d'une situation concrète », penche bien sûr pour la première solution, d'autant que la seconde risque de nous mener à de nouveaux artefacts. C'est de l'analyse concrète de situations concrètes qu'il nous faut partir, pour en tirer patiemment des éléments de théorisation. C'est par exemple ce que j'ai tenté de faire, dans un domaine limité, à propos des villes plurilingues, les décrivant pour chercher ce que l'on pourrait en tirer dans la théorisation de l'émergence des langues véhiculaires, ce qu'une sociolinguistique urbaine pouvait nous apprendre sur l'évolution des situations, sur l'apparition de langues d'unification, etc.

Ceci ne doit pas pour autant nous empêcher de pénétrer sur le terrain que les linguistes de l'homogène se réservent, bien au contraire : il n'est pas question de laisser le phonème ou la théorie du signe de côté, mais au contraire de les revisiter d'un autre point de vue, avec un autre positionnement et d'autres interrogations. Le phénomène véhiculaire par exemple que nous pouvons analyser dans les villes plurilingues a des retombées formelles : le wolof urbain de Dakar est de plus en plus différent du wolof rural, le bambara véhiculaire se différencie du bambara langue première, etc. C'est donc l'ensemble des faits observables qu'il convient de prendre en compte, mais différemment. Et cette différence de point de vue est, elle, du domaine de l'heuristique. Il

L.J. Calvet

s'agit de jeter sur des faits par nature désordonnés un regard qui ne cherche pas nécessairement à y mettre de l'ordre. Je vais donc dans les pages suivantes présenter essentiellement des résumés de propositions que j'ai déjà présentées ailleurs et qui toutes se situent dans le cadre d'une écologie des langues qu'avec quelques autres (Mufwene, Mülhassler...) j'ai tenté de développer. Il sera temps, ensuite, de nous demander s'il est possible d'aller plus loin sur le plan théorique.

## 2.2. *Quelques petites questions*

La linguistique dont a hérité ma génération, celle qu'enseignaient André Martinet, Roman Jakobson, Gleason et beaucoup d'autres, était elle-même héritière d'une vision élaborée au XIX<sup>ème</sup> siècle, en plein romantisme. Selon cette vision, la langue était un organisme vivant, reflétant la pensée et les valeurs d'une nation ou d'un peuple, et dotée d'une structure propre. Elle était le garant d'une identité collective, transmise à la fois par la mère (d'où la « langue maternelle », « lingua materna », « muttersprache », « mothertongue », etc.) et par le père (la patrie, le patrimoine, qui lui sont étymologiquement liées). Et elle se devait donc d'être immuable, toute modification risquant de mettre en péril l'identité du groupe, de la nation. Tout ceci, bien sûr, est généralement considéré aujourd'hui comme un peu vieillot. Mais que pouvons-nous proposer à la place ?

Commençons donc par une « petite » question : *qu'est-ce qu'une langue ?*

Bien des linguistes ne se posent pas cette question : la langue est ce qu'ils décrivent, un mécanisme dont il convient de montrer le fonctionnement, une donnée indiscutable dont certains se considèrent comme dépositaires. C'est par exemple le cas des chomskyens qui, je l'ai dit, se prennent comme leur propre informateur, leur intuition linguistique suffisant, pensent-ils, à fonder leur approche. Laissons de côté ce comportement caricatural. D'autres linguistes sont plus modestes, ne se considèrent pas comme LE locuteur idéal dont l'intuition résoudrait tous les problèmes de corpus et vont chercher ailleurs, chez les autres, de quoi alimenter leurs descriptions. C'est-à-dire qu'ils travaillent sur corpus, qu'ils enregistrent avec parfois des précautions méthodologiques louables des données qu'ils vont ensuite

L.J. Calvet

analyser. Mais, s'ils sont honnêtes, et si leur corpus est suffisamment large, ils se rendent vite compte que les données recueillies ne vont pas toutes dans le même sens, qu'une hypothèse à peine esquissée ou qu'une loi à peine émise se trouvent invalidées par de nouveaux exemples. Il nous faut donc accepter de considérer la « langue » comme un ensemble de *pratiques*, souvent désordonnées, et comme un ensemble de *représentations*, celles des locuteurs, bien sûr, mais aussi celles des linguistes, et chercher à comprendre comment fonctionne ce désordre. C'est-à-dire que, plutôt que de se rassurer (et de conforter les modèles) en se reposant sur des arguments du type *l'exception qui confirme la règle*, il conviendrait de considérer les différentes variantes non pas comme des exceptions mais comme des faits, tout simplement, des faits dont il faut rendre compte, comme de tous les autres.

Et ceci nous mène à une autre « petite » question : *Qu'est-ce que la linguistique ?*

Emile Littré, dans son dictionnaire, en donnait au XIX<sup>ème</sup> siècle la définition suivante : « L'étude des langues considérées dans leurs principes, dans leurs rapports et en tant qu'un produit involontaire de l'instinct humain ». Lorsqu'on la compare avec celles que nous proposons aujourd'hui les dictionnaires modernes, on se rend compte que Saussure est passé par là. Le *Robert* par exemple la définit par une citation du maître de Genève : « Science qui a pour objet 'la langue envisagée en elle-même et pour elle-même' ». Nous savons aujourd'hui que cette phrase n'a jamais été prononcée par Saussure (ou du moins qu'il n'y en a aucune trace dans les sources manuscrites du *Cours*), mais qu'importe : en un siècle on a évacué l'être humain (même si la notion d'instinct, avancée par Littré et reprise récemment par S. Pinker, est discutable). Et lorsque l'on met plutôt l'accent, comme c'est mon cas, sur les *pratiques*, sur le fait que les « langues » n'existent que par les gens qui les parlent, cette définition est inacceptable et doit donc être revue. Faut-il moderniser la définition de Littré en disant que la linguistique est quelque chose comme « l'étude des pratiques communicatives des êtres humains considérées dans leurs principes et dans leurs rapports », ce qui aurait l'avantage (ou le désavantage) de l'élargir à la sémiologie ? Ou faut-il parler désormais *des* « sciences du langage », ce qui a le désavantage (ou l'avantage) de faire éclater *une* approche en une multitude de micro-sciences ?

L.J. Calvet

A ces questions s'en ajoutent d'autres. Toute science a besoin de pouvoir répondre de façon univoque à une interrogation sur les limites de son intervention, elle doit pouvoir dire : « ceci appartient, ou n'appartient pas, à mon objet d'études ». Mais elle doit aussi présenter une cohérence entre la définition de son objet et les méthodes d'analyse et de description qu'elle se donne. Dire par exemple que la langue est un fait social, un produit social ou une institution sociale implique que son étude (la « linguistique ») soit une science sociale et qu'elle s'en donne les moyens.

Reprenons ces deux points.

*Une science doit présenter une cohérence entre la définition de son objet et ses méthodes d'analyse.* Cela signifierait pour la linguistique qu'elle se situe résolument sur le terrain social, qu'elle refuse la division entre des petits domaines considérés comme « périphériques » (sociolinguistique, psycholinguistique, ethnolinguistique...) opposés à LA linguistique « centrale ». Car cette distinction entre le centre et les périphéries est devenue la pierre de touche de bien des aveuglements et de bien des exclusions. C'est elle en particulier qui justifie la distinction entre linguistique interne et linguistique externe, et je tenterai de convaincre que ces deux approches doivent être nécessairement liées.

*Une science doit pouvoir dire « ceci appartient, ou n'appartient pas, à mon objet d'études ».* Cela signifierait, pour la linguistique, qu'elle puisse dire « ceci est une langue, ceci n'en n'est pas », qu'elle mette de l'ordre dans des situations désordonnées, en un mot qu'elle puisse extraire des pratiques un certain nombre de « structures », les langues.

Or il y a une certaine contradiction entre ces deux exigences. Se situer par exemple sur le terrain social, se confronter à la complexité des choses, tenter de lier « histoire interne » et « histoire externe », implique à la fois une démarche empirique (récolter le plus de faits possibles, ne rien exclure) et une démarche inductive (remonter éventuellement de ces faits vers des lois générales dont nous avons vu plus haut qu'il fallait s'en méfier). Je me propose donc dans un premier temps non pas de donner une théorie livrée clefs en main mais de revisiter un certain

L.J. Calvet

nombre d'avancées que j'ai pu suggérer en partant à la fois de *pratiques*, de données, et d'une critique de la façon dont les théories linguistiques en rendaient compte.

### 2.3. *Quelques principes heuristiques*

Je voudrais donc tenter de mettre de l'ordre non pas dans les faits linguistiques que je considère, on l'aura compris, comme désordonnés de façon inhérente, mais dans un certain nombre de propositions que j'ai pu avancer depuis quelques années.

#### *De l'analogique au digital*

En 1999, nous étions intervenus, avec Lia Varela, dans un débat sur les frontières entre *sociolinguistique*, *sociologie du langage* et *linguistique*<sup>33</sup>. Notre propos était d'abord de montrer que les syntagmes *sociolinguistique* et *sociologie du langage* étaient apparus séparément, le premier en anglais et le second en français, pour désigner la même chose<sup>34</sup>. Ainsi lorsqu'en 1966 Dell Hymes rédige un rapport sur la sociolinguistique pour le « Social Science Research Council » américain, il indique que cette discipline est enseignée sous les noms de *linguistique*, *sociolinguistique*, *langue et culture*, *sociologie du langage* ou *comportement linguistique*. Sociolinguistique et sociologie du langage vont ensuite s'imposer, mais toujours avec le même sens. Certains auteurs ont d'ailleurs hésité entre l'un et l'autre, comme Joshua Fishman qui publiera successivement *Readings in the sociology of language* (en 1968) puis *Sociolinguistics* (en 1970), pour s'en tenir ensuite à la première appellation. L'existence de ces deux appellations va cependant pousser les chercheurs à leur donner un contenu différent, *sociolinguistique* en venant à désigner plutôt une branche de la linguistique et *sociologie du langage* plutôt une branche de la sociologie.

---

<sup>33</sup> Calvet L.-J., Varela L., « De l'analogique au digital. A propos de sociologie du langage et/ou sociolinguistique et/ou linguistique », *Langage et Société* n°89, septembre 1999.

<sup>34</sup> Pour être exact, il faut signaler J.O. Hertzler, « Towards a Sociologie of langage », *Social Forces* 32, 1953, mais le terme ne fut pas repris en anglais jusqu'au début des années 1960, tandis que Marcel Cohen l'imposait en français.

L.J. Calvet

C'est-à-dire que l'existence de deux signifiants avait entraîné sans doute par souci de cohérence la distinction entre deux signifiés différents.

Notre propos était de prendre le contre-pied de cette division tripartite (*linguistique* ou *sociolinguistique* ou *sociologie du langage*), en montrant que la linguistique avait, pour se constituer en science, délibérément choisi la voie digitale (du latin *digitus*, « doigt » sur lequel on compte), posant que les unités de la langue étaient comptables, discrètes. La phonologie est, bien sûr, le résultat le plus achevé de cette approche, faisant l'inventaire des phonèmes d'une langue, mais négligeant parfois des phénomènes comme l'intonation qui participent pourtant à la construction du sens. Face à cette pensée digitale il en est une autre, analogique, que l'on considère généralement comme moins précise ou moins « scientifique ».

Ce passage du continu au discontinu, du non-discret au discret, définit d'une certaine façon le projet de la linguistique et explique certains de ses échecs : la phonologie est évidemment un traitement digital des sons du langage, mais la sémantique ne peut pas croire à la digitalité des signifiés. Il est donc possible d'imaginer, face au versant linguistique de la société, une approche allant de l'analogique au digital, dans un continuum dont les pôles extrêmes peuvent être baptisés de différentes façons : analogique (ou continu, ou non discret) d'un côté, digital (ou discontinu, ou discret) de l'autre.

**Analogique**  
**(ou continu, ou non discret)**

**Digital**  
**(ou**  
**discontinu**  
**ou discret)**

Sociologie du langage..... ....sociolinguistique

Indices

Signification

Langage (s)

linguistique

signaux

communicati

on

langues

Etc.

Nous avons dans la colonne de droite les pôles « digitaux » (la linguistique, qui étudie des signaux et non pas des indices, dans le cadre de la communication et donc des langues) et dans la colonne de gauche les pôles « analogiques ». L'analyse procéderait de la gauche vers la droite, par un effet de zoom, une variation de focale allant du « flou » vers le « net » ou d'une vue large vers un gros plan. Ferdinand de Saussure postulait une science, la sémiologie, qui étudierait « la vie des signes au sein de la vie sociale » (CLG p. 33), et nous pourrions ici, en le paraphrasant, postuler une *sémiologie des langues* étudiant la vie des langues au sein de la vie sociale, science dont la « linguistique » constituerait le pôle digital. Ce qui nous mène vers un protocole de description plus vaste et plus complexe que celui des linguistiques actuelles, vers un changement de paradigme que j'ai essayé d'exemplifier à propos de la ville d'Alexandrie, dans mes *Essais de linguistique*.

*Le signe linguistique revisité du point de vue analogique/digital*

Le point sur lequel ce changement de paradigme a des retombées les plus nettes est celui de la théorie du signe. La conception saussurienne du signe linguistique condense en effet une bonne partie des critiques que l'on peut faire à cette linguistique de l'ordre. Lorsqu'il utilise l'image de la feuille de papier dont un côté représente le signifiant et l'autre côté le signifié, expliquant qu'en coupant la feuille avec des ciseaux on coupe en même temps les deux côtés, donc le signifiant et le signifié, Saussure postule bien entendu que les deux faces du signe sont indissociables, intimement liées et du même ordre. J'ai proposé ailleurs<sup>35</sup> une critique de cette vision, du point de vue de l'espace dans lequel cette feuille de papier a été perçue (celui de la géométrie euclidienne), des rapports entre linguistique et psychanalyse (Freud, Lacan, Benveniste et l'énantiosémie), etc., et je me contenterai donc ici de rappeler ma conclusion : *le signifiant est du côté du digital et le signifié du côté de l'analogique*.

---

<sup>35</sup> L.-J. Calvet, *Essais de linguistique*, Paris, Plon, 2004.

Cette proposition a de nombreuses répercussions, par exemple sur l'analyse que nous pouvons faire de la variation et du changement. L'histoire d'un signe (son « étymologie ») est ainsi constituée par deux histoires séparées, qui ne sont pas nécessairement parallèles, l'une phonique et l'autre sémantique. Ainsi la phonétique historique nous explique comment le latin *tripalium* a évolué vers le français *travail*, l'espagnol *trabajo* ou le portugais *travalho*. Nous sommes ici dans le domaine des lois phonétiques, dont nous avons vu qu'elles étaient généralisables, qu'elles mettaient en forme un grand nombre d'observations et se prêtaient à vérification. Mais le mot *tripalium* ne signifiait pas « travail » en latin, il désignait un instrument de torture. Travail se disait *labor* (qui a donné l'italien *lavoro*) et l'histoire de l'évolution sémantique de « instrument de torture » à « travail » est d'un tout autre ordre que l'histoire de l'évolution phonétique de *tripalium* à *travail*, *trabajo* ou *travalho*... De la même façon, l'évolution phonétique qui mène du latin *caballus* à l'espagnol *cavallo*, au français *cheval* ou au portugais *cavalho* est une chose, mais cheval se disait en latin *equus* et *caballus* désignait un mauvais cheval. Or si tous le /k/ latin à l'initiale donnent /k/ en italien, espagnol, portugais et /ch/ en français, tous les noms de mammifères dans les langues romanes ne viennent pas d'un terme latin désignant un mammifère défectueux. Nous avons donc bien là deux histoires de type très différent, l'une qui concerne l'évolution régulière d'unités digitales (les sons) et l'autre qui concerne l'évolution irrégulière d'unités analogiques (les sens), ce qui met fortement en question la théorie saussurienne de l'indissociabilité du signe, du parallélisme du signifiant et du signifié.

De la même façon, m'inspirant des travaux de Pierre Guiraud sur le vocabulaire argotique, j'ai suggéré dans un *Que Sais-Je ?* sur l'argot que l'on pouvait rendre compte de certains procédés néologiques à l'aide de *matrices sémantiques*. Les désignations argotiques de l'argent viennent par exemple majoritairement en français de mots signifiant la nourriture (la principale exception étant un emprunt à l'arabe : *flouze*) : *galette*, *pognon*, *oseille*, *avoine*, *fric*, etc. J'avais donc suggéré qu'à partir d'une équation du type ARGENT = NOURRITURE (ou si l'on préfère argent = ce qui permet d'acheter de la nourriture) nous avons une matrice permettant de « prédire » d'autres créations : *caviar* ou *couscous* par exemple pourraient un jour dans des dialectes sociaux



différents signifier « argent ». En outre, cette vision est confirmée par l'histoire (par exemple les rapports étymologiques entre *sel* et *salaire*, entre *denrée* et *denier*...) et se vérifie dans d'autres langues : *dough* (pâte à pain) ou *bread* (pain) désignent l'argent en anglais, comme *psomi* (pain) en grec, *grano* en italien ou *pasta* en espagnol... Mais il demeure que l'on dit aussi *flouze* en français, qu'en espagnol d'Argentine on dit *lana*, qu'en hongrois on dit *dohany* mot emprunté à l'arabe qui désigne en cette langue le tabac, etc.

Dans le même ordre d'idées, il peut sembler élégant d'expliquer qu'un *poulet* est en argot celui qui « picore des renseignements », donc un policier. Mais alors pourquoi un *pigeon*, qui picore tout autant, désigne-t-il une dupe ? Parce qu'on le plume ? Parfait, mais on plume aussi les poulets... Mes belles *matrices sémantiques* fonctionnent donc parfois mais peuvent aussi parfois être falsifiées : il y a sans cesse du désordre, de l'imprévisible dans cette créativité sémantique, une vieille carpe devient un cheval, une galette de l'argent, mais le /k/ latin donne toujours /ch/ en français. Encore une fois l'évolution d'unités digitales comme les sons peut être présentée comme régulière, mais celle des sens est irrégulière, moins prévisible, et les signifiés sont donc du côté de l'analogique.

De ce point de vue *l'énantiosémie* (le fait qu'un même signifiant puisse renvoyer à deux signifiés opposés, comme en français *hôte* ou *louer*), qui a fait couler beaucoup d'encre et a fait « monter au créneau » des linguistes comme Benveniste et Hagège, tous les deux pour nier le phénomène (« la langue ne peut pas se contredire », a même écrit Hagège...), ne doit être considérée ni comme une invention d'illuminés ni comme un accident statistiquement peu représentatif des faits de langues mais simplement comme *un fait rendu possible par le signe linguistique*. Si peuvent apparaître des énantiosèmes (et leur existence est difficilement contestable) c'est parce que si le signifiant est bien du côté du digital (ou analysable de façon à peu près cohérente en termes digitaux) les signifiés sont, eux, du côté de l'analogique. L'énantiosémie n'est plus alors ce qui vient mettre en cause une certaine conception du signe, l'un des dogmes saussuriens, mais *un cas limite de la polysémie, qu'une autre conception du signe explique parfaitement*. Une langue qui n'aurait que des énantiosèmes serait certes malaisée à utiliser, mais il demeure qu'ils existent et que la théorie du signe qui vient d'être

L.J. Calvet

développée rend parfaitement compte du phénomène en le remettant à sa place.

*Hôte/parasite, proie/prédateur*

Un autre principe, qui tombe sous le sens, a des retombées heuristiques non négligeables. Il est difficile de contester qu'il n'y a pas de langues sans locuteurs, ce pourquoi, dans le cadre d'une métaphore biologique, on parle de *langues mortes* (celles qui n'ont plus de locuteurs). Cette évidence qu'il n'y a pas de langues sans locuteurs nous suggère que l'évolution quantitative des langues doit pouvoir être étudiée en fonction de l'évolution démographique de la société. Nous poserons donc l'existence de deux populations, celle des êtres humains et celle des langues, et considérerons qu'elles entretiennent des rapports de type hôte/parasite : il n'y a pas de langues sans êtres humains, les langues « vivent » aux dépens des hommes et leur nombre est donc fonction de la population humaine. Cette fonction n'est ni mécaniste ni directe. Une simple observation nous montre en effet que la population humaine n'a cessé de croître, de façon exponentielle, et qu'elle continuera de croître encore pendant de nombreuses années, tandis que la population des langues s'est stabilisée et commence à décroître<sup>36</sup>.

Il est possible de présenter cette différence entre le devenir des deux populations considérées en considérant les rapports entre les langues comme des rapports de type proie/prédateur. Il y a là un principe qui peut apparaître comme darwinien, mais qui correspond en même temps à ce que j'ai appelé naguère la glottophagie<sup>37</sup>, ou la guerre des langues<sup>38</sup>, phénomènes qui se vérifient largement dans l'histoire : les groupes humains n'ont cessé de changer de langue, d'imposer leur langue à des populations dominées, d'adopter la langue de populations dominantes. En termes de génétique des populations, il s'agit d'une compétition, qui peut prendre deux formes :

---

<sup>36</sup> Pour une présentation plus détaillée de cette approche, voir L.-J. Calvet, *Essais de linguistique*.

<sup>37</sup> Calvet L.-J., *Linguistique et colonialisme, petit traité de glottophagie*, Paris, Payot, 1974.

<sup>38</sup> Calvet L.-J., *La guerre des langues et les politiques linguistiques*, Paris, Payot, 1987.

L.J. Calvet

-Une *compétition par exploitation*, lorsque les populations ne sont pas en relation directe mais exploitent des ressources communes.

-Une *compétition par interférence*, lorsque les populations sont en relation directe et que l'une empêche l'autre ou les autres d'accéder aux ressources,

l'idée générale étant que « deux ou plusieurs espèces présentant des modes d'utilisation des ressources identiques ne peuvent coexister dans un environnement stable, la plus apte éliminant les autres »<sup>39</sup>.

Nous pouvons alors proposer un modèle de croissance qui rende compte des rapports entre ces deux populations, des rapports entre la démographie humaine et la démographie linguistique.

Toutes les études écologiques nous montrent que les prédateurs se concentrent en général là où il y a le plus de proies, ce qui implique de façon un peu paradoxale que les proies potentielles sont moins menacées lorsqu'elles sont isolées. Mais il ne s'agit que d'un paradoxe apparent. De ce point de vue en effet, plus la densité des langues est grande et plus elles sont menacées, plus elles sont des proies potentielles pour des prédateurs aux aguets. C'est-à-dire, pour revenir à notre sujet, que le plurilinguisme génère l'émergence de formes véhiculaires qui peuvent à terme menacer les autres langues présentes dans la même niche écolinguistique. Ce qui explique à la fois que l'urbanisation soit un facteur d'unification linguistique (la concentration des langues de migrants en ville les rend fragiles face à la langue dominante) et que les 225 langues du Cameroun par exemple soient plus menacées par le français et l'anglais, les deux langues officielles, que la langue unique du Rwanda ou du Burundi ou que le guarani face à l'espagnol au Paraguay.

Dans cette histoire parallèle des deux populations il y a donc eu d'abord une augmentation : l'augmentation de la population humaine génère celle de la population linguistique (en termes simples, plus il y avait des gens et plus il y avait de langues), pour la raison évidente que

---

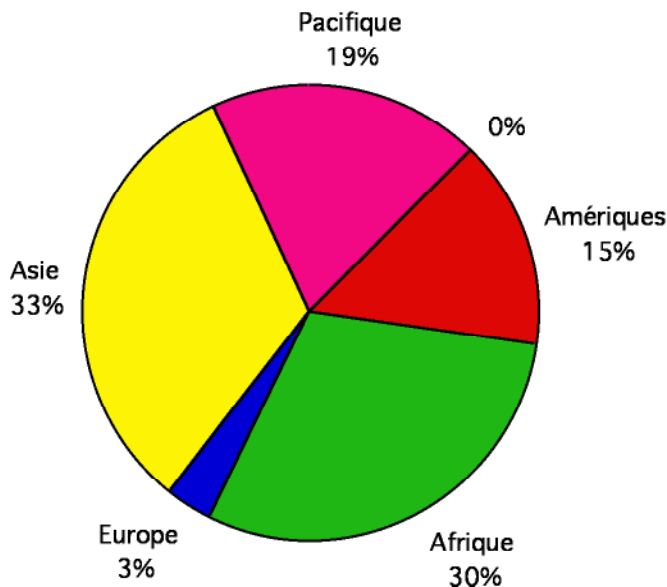
<sup>39</sup> Hardin G., « The competitive exclusion principle », *Science* n°131, 1960, cité par Barbault, op.cit. p 41.

L.J. Calvet

les groupes se séparaient pour survivre. La croissance démographique et la dispersion des populations s'accompagnaient d'une fragmentation dialectale et de l'augmentation du nombre de langues. *Nous avons là le produit du rapport de type hôte/parasite entre les deux populations.*

Puis la concentration des populations, la multiplication des communications, l'urbanisation, et surtout l'émergence d'Etats centralisateurs vont entraîner un ralentissement de la progression de la population linguistique. Le rôle des états apparaît clairement lorsque l'on considère la répartition des langues dans l'espace.

### Les langues parlées dans le monde



Nous voyons en effet dans le graphique ci-dessus que 3% des 6700 langues environ parlées dans le monde le sont en Europe, où se trouvent les états les plus anciens, et que les continents plus « riches » en langues (l'Asie, l'Afrique) sont ceux où l'on trouve les états les plus récents (à l'exception bien sûr pour l'Asie de la Chine et du Japon).

Ce ralentissement de la croissance de la population linguistique du monde ne peut à terme mener qu'à une réduction du nombre de langues, après la période de stabilisation que nous connaissons aujourd'hui. *Et*

L.J. Calvet

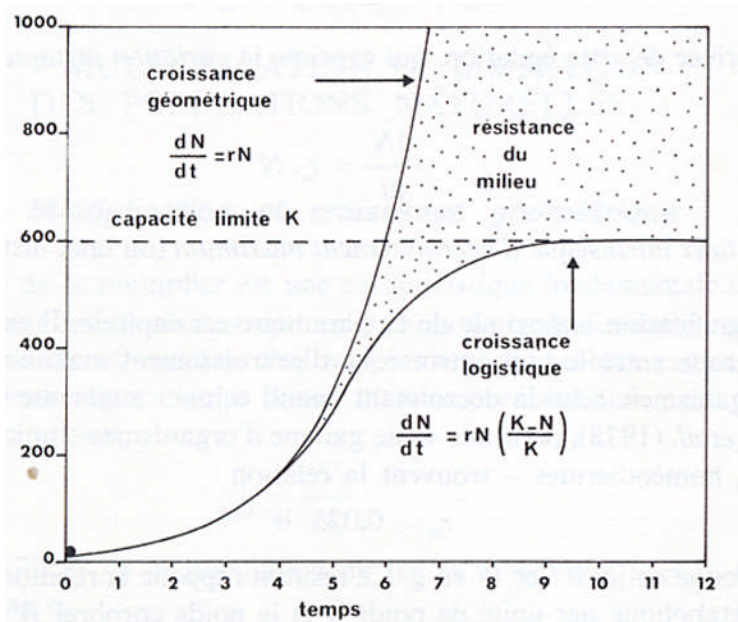
*nous avons là le produit de la compétition au sein de la population des langues, le produit du rapport proie/prédateur. J'ai proposé pour rendre compte de cette relation d'utiliser ce que Barbault a appelé un « modèle logistique de croissance ». Une population, à partir de son état initial, connaît d'abord une croissance exponentielle (ou géométrique), jusqu'à ce qu'elle atteigne une biomasse maximale<sup>40</sup>. Pour Barbault, « le taux d'accroissement réel diminue puis devient nul, le surpeuplement provoquant, soit une diminution de la natalité (et de l'immigration), soit une augmentation de la mortalité (et de l'émigration), soit les deux à la fois »<sup>41</sup>. La croissance d'une population ne serait donc pas géométrique mais logistique, comme le montre la figure ci-dessous, restreinte par une « capacité limite K » que Barbault définit comme la « *capacité biotique* du milieu pour la population considérée »<sup>42</sup> En d'autres termes, dans des conditions optimales, lorsqu'elle ne rencontre aucune limitation spatiale ou alimentaire, une population peut avoir une croissance géométrique, mais elle finit toujours par rencontrer dans son expansion une résistance du milieu qui va la limiter.*

---

<sup>40</sup> Voir par exemple Daget J., « Biodémographie », *Encyclopedia Universalis*, Paris, 1993, tome 4, pp 138-140.

<sup>41</sup> Barbault R., *Ecologie des populations et des peuplements*, Paris, Masson, 1981, 12.

<sup>42</sup> op.cit. p.13.



La population humaine est, depuis l'origine, en constante augmentation et les projections à long terme donnent une augmentation continue au moins jusqu'en 2150. C'est-à-dire qu'elle est encore dans sa phase de croissance géométrique. En revanche, la population linguistique semble avoir atteint sa biomasse maximale, la résistance du milieu, et se trouver maintenant dans sa phase de croissance logistique. Ce qui signifie que dans le schéma ci-dessus la courbe de gauche représenterait l'évolution de la population humaine tandis que l'évolution de la population linguistique serait mieux représentée par celle de droite. L'histoire quantitative des langues du monde répond ainsi à un *modèle logistique de croissance* : la population des langues est confrontée à la résistance du milieu, sa croissance est stoppée et sa courbe pourrait marquer une baisse.

Mais tout ceci n'empêche pas que si la population des langues est en diminution et si nous pouvons en rendre compte de façon élégante par ce modèle, *des langues continuent d'apparaître*. L'hindoustani en Inde a éclaté en deux formes, le hindi et l'ourdou, en ex-Yougoslavie plus personne ne déclare parler le serbo-croate mais on se revendique plutôt du serbe, du croate ou du bosniaque. Bien sûr les différences entre ces formes sont minimes, ne gênent en rien l'intercompréhension, mais

L.J. Calvet

l'action des représentations (ici générées par le nationalisme ou une quête identitaire) sur les pratiques peut, à terme, mener à une différenciation maximale. C'est-à-dire que des événements extérieurs et non prédictibles (des oppositions politiques ou religieuses dans les deux exemples précédents) peuvent venir contrarier le modèle proposé. Encore une fois, ce que présente mon modèle est une *tendance* et non pas une *loi*, même si la tendance est actuellement forte. Et les contradictions que je relevais plus haut dans la présentation par le SIL de l'arabe, de l'ourdou ou du hindi, viennent du fait que *ne sont pas prises en compte les pratiques et les représentations des locuteurs, mais seulement les représentations des linguistes.*

#### *Acclimatement/acclimatation*

On distingue en écologie, à partir du verbe *s'acclimater*, deux processus différents, *l'acclimatement* et *l'acclimatation*. L'acclimatement est le fait, pour une espèce déplacée, de survivre, un animal ou un végétal que l'on transporte d'un climat à un autre s'adaptant à de nouvelles conditions de vie. L'acclimatation est pour sa part le fait, pour les mêmes espèces déplacées, non seulement de survivre mais aussi de se reproduire.

Il en va de même pour les langues, que déplacent leurs locuteurs. Elles peuvent survivre un temps, coexister avec d'autres langues dans une niché écolinguistique donnée, mais elles peuvent aussi prendre racine. Et de la même façon qu'un ours polaire transporté sous les tropiques devrait pour s'adapter perdre une partie de sa graisse et de sa fourrure, les langues s'adaptent, prennent en quelque sorte des couleurs locales. Le néerlandais par exemple a connu en Indonésie une période d'acclimatement puis il a disparu : la reproduction correspond ici à la transmission d'une génération à l'autre. En revanche le portugais est passé au Brésil par un processus d'acclimatation, il a éliminé une grande partie de la population linguistique pour prendre sa place, mais en même temps il s'est en partie transformé. Car une espèce qui s'acclimate se transforme nécessairement, s'adapte aux conditions locales, en particulier aux conditions climatiques, et il en va de même pour les langues (même si, bien sûr, le climat ne joue ici aucun rôle). L'acclimatation de l'espagnol à Cuba ou en Argentine a entraîné des

L.J. Calvet

modifications à la fois de la phonologie, du lexique et de la syntaxe, au point que certains parlent aujourd'hui de l'espagnol de Cuba (ou d'Argentine, etc.), voire du cubain (ou de l'argentin, etc.). De la même façon l'anglais parlé en Angleterre n'est pas tout à fait le même que celui parlé aux USA ou en Australie, etc. Et bien que nous ne sachions pas si le français connaîtra en Afrique une acclimatation ou s'il restera au stade de l'acclimatation avant de disparaître, nous pouvons d'ores et déjà constater des modifications, des adaptations, qui font qu'on ne parle plus tout à fait la même langue dans les rues de Bamako (Mali), d'Abidjan (Côte d'Ivoire) ou de Libreville (Gabon) que dans celles de Paris.

Encore une fois nous sommes incapables de dégager ici des lois. Face au résultat d'une acclimatation linguistique nous pouvons, *a posteriori*, évoquer des facteurs externes (emprunts, interférences...) ou internes (dynamique du changement, tendances à la régularisation...) qui expliquent en partie une évolution, mais nous ne pouvons pas prévoir à coup sûr cette évolution. Là où un chimiste par exemple peut dire qu'une base ajoutée à un acide donne un sel et de l'eau, énonçant une loi que des milliers d'expériences confirmeront, un linguiste en est réduit à observer ce qui se passe, à le décrire, à tenter de le comprendre, sans pouvoir en tirer pour sa part des lois.

### *Les langues comme comportements tendanciels*

J'ai plusieurs fois dans ce texte souligné que, dans nos tentatives de compréhension des données linguistiques, nous étions confrontés plus à des *tendances* qu'à des *lois*. C'est là une caractéristique des sciences sociales dont les objets d'étude n'ont pas les comportements réguliers, voire immuables, des objets étudiés par les sciences exactes... Et j'ai par ailleurs, en particulier à propos d'un bref exemple concernant l'île Maurice, parlé de *langues tendanciels*. Cette rencontre n'est pas fortuite, car le désordre que j'ai évoqué ne concerne pas seulement les rapports entre des objets décrits, le comportement de ces objets ou de leurs utilisateurs, il concerne les objets eux-mêmes qui se caractérisent par un constant bouillonnement. Lorsque que j'ai avancé, dans *Pour une écologie des langues du monde*, que les langues n'existaient pas, que nous n'avions en fait que des pratiques et des représentations, il ne



s'agissait nullement d'une boutade mais d'une réalité pour moi indiscutable, incontournable. La vieille idée selon laquelle une langue est un *instrument de communication* pourrait nous pousser à considérer que, comme un marteau que l'on sort de sa boîte à outils pour bricoler et que l'on range ensuite, ou comme une guitare que l'on prend lorsqu'on veut en jouer pour ensuite la remettre en place, la langue s'utilise lorsqu'on en a besoin, se range ensuite. Elle aurait ainsi un *mode d'emploi*, que les linguistes essaieraient de formuler, ou une *méthode d'utilisation* : grammaires, dictionnaires... Mais nous savons tous qu'il en va autrement, même si cela ne nous arrange pas nécessairement. Ce que nous baptisons variation, variation inhérente, évolution, dialectalisation, etc., nous montre bien que contrairement à un marteau ou à une guitare, la langue n'est pas un objet physique, que l'on peut montrer, isoler, exposer. J'ai parlé d'artéfact, d'objet inventé ou construit par les moyens de description, mais cet « objet de l'art » est antérieur à l'intervention du descripteur, bien en amont de la linguistique, il est sans cesse produit et transformé par les locuteurs eux-mêmes, dans leurs pratiques et leurs représentations.

L'exemple de ce qu'on a appelé la *diglossie* peut être ici éclairant. Lorsque Charles Ferguson baptise et décrit ce phénomène en 1959, il considère qu'il s'agit là d'une situation « stable » entre deux formes d'une même langue ou entre deux langues ayant la même origine, la *variété haute* et la *variété basse*, entre lesquelles il y a une répartition fonctionnelle des usages. Signalons au passage que le cas de la Grèce (l'un des quatre exemples de Ferguson) dément l'idée de situation stable : des choix politiques après le retour à la démocratie ont fait que la katharevoussa a laissé place à la démotiki... Ce qui compte surtout, c'est que la diglossie ainsi considérée était un bel exemple de « construction sociolinguistique », reposant sur l'hypothèse de l'existence de langues homogènes et séparées... Ferguson avait remarqué, alors qu'il travaillait au Caire, que dans les journaux égyptiens on utilisait plutôt l'arabe classique pour certains types d'articles et plutôt l'arabe égyptien pour les bandes dessinées ou d'autres types d'articles, et il en avait tiré une théorie en ajoutant à ses données trois autres exemples de situations, celles d'Haïti (français/créole), de la Grèce (katharevoussa/démotiki) et de la Suisse alémanique (Hochdeutsch/Suisse allemand). Il y avait donc chaque fois pour lui deux formes différentes, et selon les situations on parlait plutôt la variété haute ou la variété basse, la katharevoussa plutôt

L.J. Calvet

que la démotiki, l'arabe classique plutôt que l'arabe dialectal, le créole haïtien plutôt que le français...

Ce modèle a été critiqué par des linguistes locuteurs de langues dominées, comme le catalan par exemple, expliquant que la diglossie n'était pas une organisation harmonieuse des rapports entre deux langues coexistant dans une même niche, mais la traduction de rapports de force. Ils ne mettaient ainsi pas en cause les fondements du modèle, ils le complétaient.

Lorsque Joshua Fishman, en 1967, propose de considérer qu'il peut y avoir diglossie entre des langues non apparentées génétiquement (par exemple entre espagnol et guarani au Paraguay) il remet d'un certain point de vue les pendules à l'heure, en insistant sur les fonctions et non plus sur les formes. Mais il vidait du même coup le modèle de son contenu : pour Ferguson, la diglossie était un fait linguistique, pour Fishman elle était un fait social.

Or ce fait linguistique, ou plutôt cette postulation d'une situation linguistique dont on pouvait tirer un modèle, consistait à introduire une segmentation digitale dans une continuité analogique, ce qui pose quelques problèmes. Les choses sont en effet beaucoup plus compliquées. Certains créolistes ont aujourd'hui tendance à présenter les faits globalement considérés comme « un créole » comme un continuum allant d'un basilecte à un acrolecte en passant par un mésolecte. D'autres considèrent que le créole est une langue à part entière. D'autres encore considèrent qu'il s'agit d'une forme de la langue « lexificatrice » (français, anglais, portugais...). Il en va de même pour l'arabe : les arabisants ne considèrent guère aujourd'hui les choses en termes binaires (classique/dialectal) et parlent par exemple d'arabe moyen (*wusta*), etc. Et les travaux de Peter Hawkins<sup>43</sup> sur le grec ont suggéré qu'il n'y avait pas deux formes normées (la variété haute et la variété basse) mais plutôt une forme stable, la *démotiki* (variété basse, dans la terminologie de Ferguson) et « a range of intermediate variants, representing different degrees of or stages of purification » (le choix de ce dernier mot, *purification*, s'explique par le fait que le nom de la variété haute grecque, *katharevousa*, vient d'un verbe qui signifie « épurer »). Ici

---

<sup>43</sup> Hawkins P., « Diglossia revisited », *Language sciences* n°5, 1983, 1.

L.J. Calvet

encore donc, et plus nettement qu'ailleurs, non pas deux pôles différenciés mais plutôt un continuum.

De la même façon, lorsqu'André Martinet croyait résoudre le problème de la variation sociale dans la langue en faisant référence à ce qu'il appelait des *niveaux de langue*, que d'autres parlent de *registres* de langue, nous sommes face à la volonté d'introduire une fois de plus du discontinu dans le continu, ou du digital dans l'analogique. Ni ces *niveaux* ni ces *registres* ni même les deux *variétés* de la diglossie ne sont pourtant identifiables, il est impossible de les décrire (niveau un, deux, trois...) en les définissant par des traits linguistiques univoques, impossible de dire que tel niveau est caractérisé par tel ensemble de faits syntaxiques, lexicaux ou phonologiques. Certaines formes de palatalisation et l'usage du verlan par exemple pourraient être considérés comme la marque d'un parler « jeunes de la banlieue parisienne ». Mais la palatalisation gagne les jeunes de la bourgeoisie, le verlan est utilisé par tout le monde ou presque et, une fois encore, nous sommes confrontés à des tendances dans le cadre d'une continuité.

Il nous faut donc concevoir les langues comme des constructions sociales mouvantes, des pratiques fluctuant sans cesse autour de ce que nous pourrions appeler des « noyaux durs », des pratiques en constant changement, en constant mouvement. Mais en mouvement vers quoi ? Et dans quelles limites, sous quelles contraintes ?

La notion d'homéostat peut ici être utile. Au milieu des années 1960, Lluís Aracil, voulant poser les bases de ce qui est aujourd'hui connu en Catalogne comme la « normalisation linguistique », avait distingué entre les actions portant sur « le développement des fonctions socio-culturelles de la langue » et celles portant sur la « réorganisation des fonctions linguistiques de la société »<sup>44</sup>. M'inspirant de lui, j'ai proposé de considérer que toute situation linguistique est le produit sans cesse mouvant d'un rapport homéostatique, d'une autorégulation, entre les *besoins linguistiques de la société* et les *fonctions sociales des langues*, c'est-à-dire les *réponses sociales que pouvaient apporter les langues aux demandes linguistiques de la société*. Cette vision des

---

<sup>44</sup> Aracil L., 1965, *Conflit linguistique et normalisation linguistique dans l'Europe nouvelle*, Nancy, C.E.U.

choses peut être utilisée dans le domaine des politiques linguistiques, à la fois pour formuler les problèmes auxquels sont confrontées des sociétés et pour esquisser des solutions. Mais, de façon plus large, elle permet également de rendre compte des pratiques linguistiques qui, dans leur désordre, ou leur bouillonnement, témoignent à la fois d'une recherche de cohésion au sein d'un groupe et d'une volonté d'exclusion. Nous avons d'un côté une tendance grégaire (on construit ou on conforte le groupe en instaurant les conditions d'une possibilité de « communication » en son sein) et de l'autre, corrélativement, une tendance différenciatrice (le groupe ainsi constitué ou conforté est séparé d'autres groupes constitués ou confortés par d'autres pratiques linguistiques). Les besoins linguistiques du groupe sont donc satisfaits par un travail de la forme linguistique : les pratiques impriment sur la langue, consciemment ou pas, des volontés de différenciation ou de conformité, de rupture ou de convergence. Ces volontés peuvent parfois être divergentes. La tendance à la grégarité par exemple peut entrer en conflit ou en contradiction avec d'autres tendances, en particulier une tendance à la véhicularité. Les deux tendances composent alors entre elles : par exemple on privilégie l'une des formes ou on en crée une nouvelle pour assurer la fonction véhiculaire. Dans certaines situations, la recherche d'une communication entre différents groupes l'emporte sur la recherche de la communication au sein du groupe, et des formes grégaires peuvent être évincées. Il s'agit là de phénomènes transitoires, qui peuvent évoluer de différentes façons et qui constituent toujours des solutions *ad hoc*, que parfois le groupe, ou la mémoire sociale, tendent à durcir, que le pouvoir dans sa volonté de légiférer tend à standardiser et que le linguiste dans sa volonté de décrire comme objet invariant tend à figer. Il s'agit alors là de formes linguistiques institutionnalisées, le plus souvent de langues écrites, congelées, qui n'ont plus rien à voir avec les pratiques réelles continuant, de leur côté, leur évolution et leurs variations.

Pour dire les choses de façon triviale, les pratiques linguistiques constituent donc une façon de se débrouiller pour se faire comprendre de ceux auxquels nous sommes socialement liés (les membres d'un groupe, d'une communauté), pour *établir une relation* avec eux, ce qui implique que nous ne soyons pas compris par les membres d'un autre groupe, d'une autre communauté, que nous *coupons la relation* avec eux. Les facteurs intervenant sur ces pratiques sont d'ordres divers, parfois

L.J. Calvet

contradictoires, et le résultat en est le grand désordre que nous avons évoqué. Mais ce « bordel ambiant » que constituent les langues n'est ni une tare ni un atout, il est tout simplement une donnée, le produit de la recherche d'une solution linguistique à un problème social, ou le versant linguistique de la constitution d'un groupe social, quelle que soit la taille ou la nature de ce groupe. Ce processus concerne en effet aussi bien une bande de jeunes qui se construit une forme linguistique cryptique qu'une nation en voie de constitution. Et il s'observe aussi bien dans des groupes clos qu'aux intersections entre groupes. Ainsi, dans les zones frontalières, des formes linguistiques mixtes peuvent apparaître (par exemple le *portunhol* entre le Brésil, l'Argentine et l'Uruguay), qui peuvent être transitoires ou perdurer.

La langue est donc le résultat de comportements tendanciels, choisis ou imposés par la situation, le résultat d'une constante recherche, elle est une *forme tendancielle*, et la linguistique, qui veut en rendre compte, apparaît alors comme une *science tendancielle*, une pulsion vers la mise en équation de comportements qui résistent aux lois, qui n'ont pas seulement un ordre propre, interne (même si certains facteurs d'évolution sont internes), mais également des déterminations externes permanentes. Il nous faut alors en prendre acte, et en tirer les conclusions qui s'imposent pour nos procédés heuristiques.

Mais c'est là que commencent les difficultés car il n'y a pas de réponse simple, carrée, digitale, à ce problème. Il serait par exemple erroné de vouloir prendre le contre-pied systématique de la désinvolture des générativistes face aux pratiques réelles en sacrifiant le corpus, qui est certes indispensable, mais à condition de ne pas en faire le *nec plus ultra* des données. Didier de Robillard a, sur ce point crucial, des mots très durs à l'endroit des linguistes transformés en ventriloques faisant parler leur terrain :

« Il n'en demeure pas moins que la richesse des corpus ne peut que s'appauvrir considérablement quand ce que le lecteur final en voit est réduit à quelques fragments dont le sens n'est guère accessible sans la contextualisation et les abondantes et indispensables gloses du chercheur. Or cela est profondément contradictoire : comment une approche contextualisante peut-elle aboutir au fait que l'on soit obligé de compter sur le récit de contextualisation qu'en fait le chercheur, ce qui

L.J. Calvet

revient à un constat d'échec : on voulait que le lecteur fasse l'expérience directe et saisissante du contexte, et on aboutit, souvent sans assumer pleinement ce rôle, sans l'explicitier (Qui suis-je, moi, le traducteur ? Pourquoi est-ce que je traduis ?), à le lui traduire, de manière inévitablement non neutre, par bribes, qui le laissent sur sa faim ».

Reprenant une formule d'Antony Lodge, Philippe Blanchet (2005) pour sa part avançait sur ce point une proposition qui choquera peut-être les « linguistes de terrain » (ceux qui scandent « c'est pas dans mon corpus, c'est pas dans mon corpus », travaillant inlassablement sur un texte enregistré parfois depuis de longues années) mais me paraît frappée au coin du bon sens :

« Je dirais qu'en termes d'observables, 'tout est bon à prendre, ça dépend de ce qu'on en fait', puisque les observables réunis en corpus et présentés dans un discours scientifique ne fondent pas prioritairement la connaissance mais en témoignent en alimentant le processus interprétatif ».

Il prônait en fait une sorte d'ethno-sociolinguistique dans laquelle rien ne peut remplacer la connaissance du terrain, dans laquelle aussi le corpus n'est pas une condition *sine qua non* de la scientificité mais plutôt un exemplier, les exemples étant proposés non pas par ce qu'ils sont *représentatifs* mais parce qu'ils sont *significatifs*. J'ai parfois des divergences avec Philippe Blanchet, en particulier pour ce qui concerne l'usage des statistiques. Certaines des enquêtes que j'ai faites (par exemple sur des marchés plurilingues) m'ont convaincu de l'importance d'une approche chiffrée, même si cela ne signifie nullement qu'il ne faille prendre en compte que les pratiques majoritaires. L'utilisation de tris croisés est également éclairante, même si cela ne signifie nullement qu'il faille y voir *la vérité* sur les faits observés. Blanchet pour sa part n'est pas nécessairement convaincu par ces instruments (ce qui ne nous a pas empêché de collaborer à une analyse des langues de France dans le recensement de 1999). Mais je considère avec lui que des faits notés à la volée, des indications marginales, une affiche comme celle relevée par Didier de Robillard que j'ai citée plus haut et surtout la connaissance du « terrain », de l'histoire de sa constitution, sont parfois beaucoup plus lourds en signification qu'un bon corpus réuni dans des conditions

L.J. Calvet

indiscutables mais nécessairement limité.

L'opposition entre données *significatives* et données *représentatives* me paraît donc ici fondamentale. Une enquête par questionnaire doit évidemment utiliser un échantillon de population représentatif. Mais face à la complexité profonde des faits que nous observons et dont nous essayons de rendre compte, c'est aussi par la signification, les indices, la sociologie du langage ou, si Blanchet le désire, une ethno-sociolinguistique (il s'agit plus simplement pour moi d'une linguistique analogique, ou d'une approche analogique des faits de langue) qu'il nous faut commencer, dans cette approche en zoom que j'ai proposée, de l'analogique au digital. Tout est bon à prendre, en effet, et la tentative de digitalisation de certains observables viendra toujours assez tôt. Notre travail ne peut être ici qu'un *récit*, ponctué de pôles digitaux lorsque les données s'y prêtent, de la façon dont s'établissent, ou tentent de s'établir, des relations entre des êtres sociaux.

Mais, même si cette approche est plus complexe que la projection d'une grille préétablie sur des pratiques, il nous faut pour cela beaucoup de modestie. Il nous faut supporter parfois la frustration face à des faits qui n'entrent dans aucun système ou invalident les systèmes provisoires que nous avons pris comme hypothèses descriptives. Il nous faut conserver tout ce que l'observation nous donne à voir et à entendre, sans rien évacuer au nom de je ne sais quel ordre transcendantal. Il nous faut garder en tête ce qu'écrivait Roland Barthes, dans son *Système de la mode*, lorsqu'il soulignait que dans le discours sur un code (en l'occurrence ici le code de la route) il y a toujours la « parole intimidante » du moniteur d'auto-école :

« Je reçois sans doute un message objectif : *le rouge est le signe de l'interdiction* (à preuve que mes actes pourront s'y conformer), mais ce que je vis réellement, c'est la parole de mon moniteur, c'est sa phraséologie [...]. La dissociation des deux systèmes ne peut être que théorique ou expérimentale : elle ne correspond à aucune situation réelle ; car il est très rare que devant une parole intimidante (qui est toujours une parole connotée), on ait le pouvoir de séparer *in petto* le

L.J. Calvet

message dénoté (le contenu du discours) du message connoté (l'intimidation) »<sup>45</sup>.

Face à la complexité sociale dont les faits linguistiques ne sont bien sûr qu'un aspect, il nous faut surtout accepter de nous placer au ras des pâquerettes, comme Fabrice del Dongo à Waterloo, plutôt que de s'autoproclamer observateur omniscient comme Tolstoï face à Austerlitz.

Je sais que ceci pourrait être lu comme une abdication : Quoi ! La linguistique ne serait pas une science !

Ou comme une régression : Saussure a constitué la linguistique en science et voilà qu'on nous propose d'oublier les lois, les structures, les oppositions binaires, pour nous pencher sur du désordre ! Voilà qu'on prétend marier le digital à l'analogique ! Le net au flou !

Il va de soi que je récusé par avance ces lectures, que je ne suggère ni d'abdiquer ni de reculer. L'histoire a parfois d'étranges détours et le progrès n'est pas toujours là où on pense le trouver. Un amphithéâtre romain n'est jamais que l'accolement de deux théâtres grecs, deux demi-cercles faisant dès lors un cercle. Progrès ? Peut-être, mais guère bouleversant... En revanche *Guernica*, la toile de Pablo Picasso, rend caducs tous les modes d'expressions « engagés » qui l'ont précédée. Je veux dire par là que l'addition de savoirs ou de techniques antérieures n'a que rarement bouleversé un paradigme, que l'amphithéâtre romain ne révolutionne guère l'architecture grecque. Je veux dire aussi que l'accumulation des regards parcellaires (je pense bien sûr aux différentes « sciences du langage ») peut produire de l'information ou de l'étrangeté, plus rarement une révolution scientifique. Les certitudes hautaines de la phonologie ajoutées au souk des différentes théories syntaxiques, aux balbutiements de la sémantique, aux éclairs de la sémiologie, aux intuitions de l'analyse interactionnelle, à la vacuité prétentieuse des approches cognitives ou des grammaires formelles et à différents autres dogmatismes, ne constituent en rien, par accumulation, une analyse de la complexité sociale et ne témoignent que de la coexistence de petites baronnies disciplinaires.

---

<sup>45</sup> op.cit. p. 43.



L.J. Calvet

Ayant enseigné la linguistique depuis bientôt quarante ans dans deux universités françaises, à la Sorbonne et à Aix-en-Provence, je suis frappé par le fait que, bien souvent, la mise au point des programmes de linguistique ne se fonde pas sur ce que devraient savoir les étudiants, sur une vision globale de la linguistique, mais sur ce que savent ou veulent enseigner les enseignants. En outre une tendance forte à l'endogamie (c'est-à-dire à recruter sur place ses propres étudiants le plus tôt possible après leur soutenance de thèse) a au moins deux retombées négatives. D'une part elle est en train d'appauvrir considérablement la recherche en créant des groupes dans lesquels il n'y a guère de discussions scientifiques puisqu'on recrute des enseignants du même bord théorique que leurs « patrons », qui vont donc continuer à creuser le même sillon, sans grandes confrontations théoriques avec d'autres courants de pensées. D'autre part elle multiplie l'éclatement de la science. De jeunes enseignants-chercheurs qui, en général à l'instigation de leur directeur de thèse, ont travaillé dans un domaine étroit, arrivent sur le marché sans avoir une vision globale des problèmes linguistiques, incapables de proposer aux étudiants un panorama de la science, de son histoire, de ses conflits. Dans ses *Essais de linguistique générale*, Roman Jakobson reprenait deux fois la même formule :

« Ainsi nous nous rendons de mieux en mieux compte que notre but suprême, c'est l'observation du langage dans toute sa complexité. Je dirai, paraphrasant Térence : *Linguista sum : linguistici nihil a me alienum puto* »<sup>46</sup>.

« Je suis linguiste et rien de linguistique ne m'est étranger » : bien peu de gens tiennent aujourd'hui ce langage, et j'ai tendance à considérer que c'est bien dommage. Ce qui me mène, pour finir, à une dernière métaphore.

Il existe une médecine générale, des médecins généralistes, et différentes spécialités : un médecin de famille peut vous envoyer consulter un dermatologue, un pneumologue, un oncologue, un gynécologue, lorsqu'il l'estime nécessaire, mais il demeure que le corps humain est un tout, qui a sa science, la médecine, et que la médecine elle

---

<sup>46</sup> *Essais de linguistique générale*, Paris, 1963, 27 et 248.

aussi est un tout, composé de généralistes et de spécialités plus pointues. Les différences entre médecine et linguistique sont nombreuses. La « langue » tout d'abord, nous l'avons vu, n'est pas un objet aussi cerné que le corps humain, et il n'est même pas certain qu'il soit cernable. En outre la médecine intervient toujours sur le corps, alors que la linguistique intervient rarement sur la langue (sauf dans le cas des politiques linguistiques). Mais la différence fondamentale entre les deux ensembles me paraît être ailleurs, dans les rapports entre généralistes et spécialistes, qui ne sont pas du même type dans les deux « sciences ». Dans le cas de la médecine ils sont **complémentaires, centripètes**, dans le cas de la linguistique ils sont **éclatés, centrifuges**, ce qui rend difficile l'élaboration d'une théorie de la complexité sociale vue sous son aspect linguistique.

Il peut paraître paradoxal, après avoir longuement insisté sur le désordre, de plaider *in fine* pour la linguistique générale (et donc d'un certain point de vue contre les sciences du langage), car c'est bien de cela qu'il s'agit. C'est pourtant mon propos, manifesté depuis longtemps, lorsque j'ai commencé à écrire *(socio)linguistique* et non pas *sociolinguistique*, avec l'idée que le segment entre parenthèses devrait un jour tomber comme une feuille morte. Pour *une* linguistique du désordre et de la complexité, donc. Qu'il nous faut élaborer avec modestie, en considérant modestement que nous ne pratiquons pas une science exacte mais une science tendancielle, c'est-à-dire une tentative de compréhension et de récit toujours à la poursuite d'une « vérité » mouvante qui sans cesse nous glisse entre les doigts.

### *Bibliographie*

- Aracil L.V., 1965, *Conflit linguistique et normalisation linguistique dans l'Europe nouvelle*, Nancy, C.E.U.
- Auroux S., 1998, « Introduction », *L'hyperlangue brésilienne, Langages* n°130.
- Auroux S., 2006, « Les embarras de l'origine des langues », *Marges Linguistiques* n°11, 85.
- Barbault R., 1981, *Écologie des populations et des peuplements*, Paris, Masson.
- Blanchet P., 2005, « Sur le statut épistémologique de la notion de 'corpus' dans un cadre ethno-sociolinguistique », in *Actes du*

L.J. Calvet

*colloque Faits de langues et de cultures dans la dynamique interactionnelle*, Paris, 2005.

- Bourdieu P., 2004, *Esquisse pour une auto-analyse*, Paris, Raisons d’agir.
- Brixhe C., 2006, « De la filiation à l’héritage, réflexion sur l’origine des langues et des dialectes », *Etudes anciennes* n°31, Nancy.
- Brixhe C., 2006, « Situation, spécificités et contraintes de la dialectologie grecque », *Etudes anciennes* n°31, Nancy, 40.
- Calvet L.-J., 1974, *Linguistique et colonialisme, petit traité de glottophagie*, Paris, Payot.
- Calvet L.-J., 1987, *La guerre des langues et les politiques linguistiques*, Paris, Payot.
- Calvet L.-J., 1999, « Aux origines de la sociolinguistique, la Conférence de Sociolinguistique de l’UCLA (1964) », *Langage et Société* n°88, 45-46.
- Calvet L.-J., 2004, *Essais de linguistique*, Paris, Plon.
- Calvet L.-J., 2006, « La mixité : situation ou résultat ? Fonction identitaire ou véhiculaire ? », Conférence présentée à Tunis, Tunis, Université de la Manouba, 16 février 2006.
- Calvet L.-J., Varela L., 1999, « De l’analogique au digital. À propos de sociologie du langage et/ou sociolinguistique et/ou linguistique », *Langage et Société* n°89.
- Chomsky N., 1977, *Dialogues avec Mitsou Ronat*, Paris, Flammarion, 72 et 74.
- Daget J., 1993, « Biodémographie », *Encyclopedia Universalis*, Paris, tome 4, 138-140.
- Durkheim E., 1905-1906, « Comment les mots changent de sens », *L’année sociologique* 1905-1906, repris dans *Linguistique historique et linguistique générale*.
- Durkheim E., 1965 [1921], *Linguistique historique et linguistique générale*, Paris, Klincksieck.
- Faidherbe L.-L.-C., 1884, « L’Alliance française pour la propagation de la langue française dans les colonies et les pays étrangers », *Revue scientifique* 3, 7.
- Hardin G., 1960, « The competitive exclusion principle », *Science* n°131.
- Hawkins P., 1983, « Diglossia revisited », *Language sciences* n°5, 1.
- Hertzler J.-O., 1953, « Towards a **Sociologie of langage** », *Social Forces* 32.

L.J. Calvet

- Highfield A., Valdman A., dir., 1981, « Halfway between Quichua and Spanish : The case for relexification », in *Historicity and Variation in Creole studies*, Ann Arbor, Karoma Publishers.
- Hombert J.-M., 2001, « Qui de la langue ou du gène influence l'autre ? », *Sciences et avenir* n°125, 47.
- Houdebine A.-M., 1991, « La dilution de l'Objet », *Où en sont les sciences du langage dix ans après ? BUSCILA*, Paris, 135.
- Hymes D., 1966, *Sociolinguistics*, The Hague, Paris, Mouton.
- Jakobson R., 1963, *Essais de linguistique générale*, Paris, éd. Minuit.
- Levi-Strauss C., 1945, « L'analyse structurale en linguistique et en anthropologie », *Word* vol. 1, n°2, 1-21.
- Levi-Strauss C., 1958, *Anthropologie structurale*, Paris, Plon.
- Meillet A., 1916, « Compte rendu du Cours de linguistique générale de Ferdinand de Saussure », *Bulletin de la société linguistique de Paris*, 166.
- Miller C., 2005, « Faits de langue et de cultures dans la dynamique interactionnelle », Colloque de sociolinguistique, Paris, 7-8 octobre 2005.
- Muysken P., 1981, « Halfway between Quichua and Spanish : The case for relexification », in Highfield A., Valdman A. dir., *Historicity and Variation in Creole studies*, Ann Arbor, Karoma Publishers.
- Revue Plurilinguismes, n°8, 1994, *Créolistique et grammaire générative*.
- Robillard D. (de), 2005, « Quand les langues font le mur lorsque les murs font peut-être les langues : *mobilis in mobile*, ou la linguistique de Nemo », *Revue de l'université de Moncton*, vol. 36 n°1.
- Saussure F. (de), 1891/2002, *Écrits de linguistique générale*, Paris, Gallimard. [Bouquet (S.) & Engler (R.), (éds.)].
- Saussure F. (de), 1916/1972, *Cours de linguistique générale*, Paris, Payot [Édition critique préparée par Tullio de Mauro].

L.J. Calvet