CHAPITRE II

Les langues en contact

Il y aurait, à la surface du globe, entre 6 000 et 7 000 langues différentes et environ 200 pays. Un calcul simple nous montre qu'il y aurait théoriquement environ 30 langues par pays, et si la réalité n'est pas à ce point systématique (certains pays comptent moins de langues, d'autres beaucoup plus), il n'en demeure pas moins que le monde est plurilingue en chacun de ses points et que les communautés linguistiques se côtoient, se superposent sans cesse. Ce plurilinguisme fait que les langues sont constamment en contact. Le lieu de ces contacts peut être l'individu (bilingue, ou en situation d'acquisition) ou la communauté. Et le résultat de ces contacts est l'un des premiers objets d'étude de la sociolinguistique.

I. – Emprunts et interférences

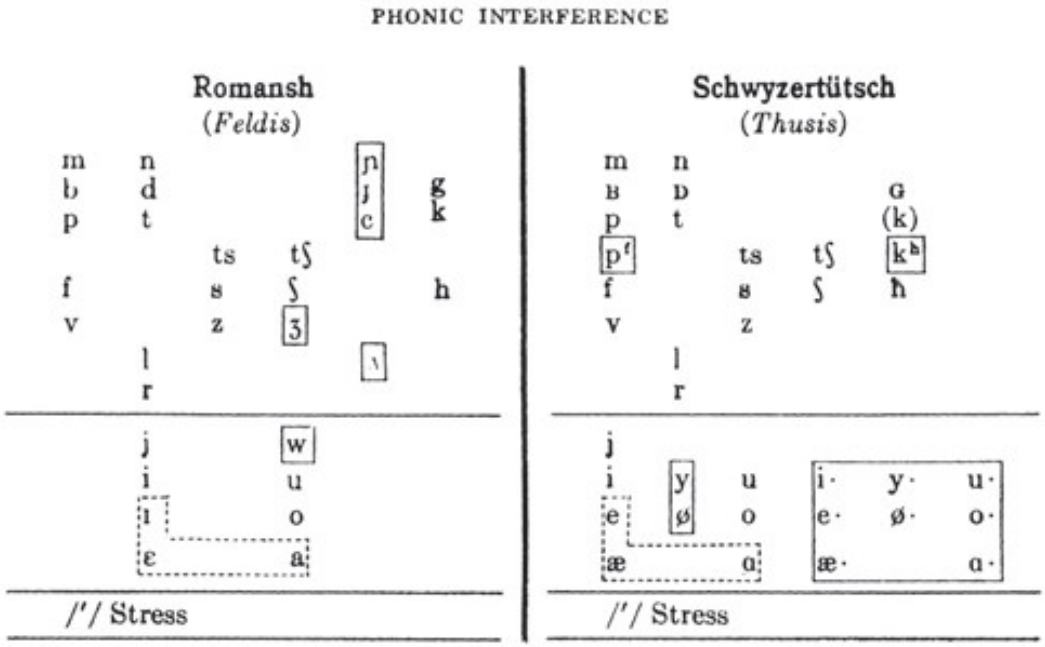
« Le mot interférence désigne un remaniement de structures qui résulte de l'introduction d'éléments étrangers dans les domaines les plus fortement structurés de la langue, comme l'ensemble du système phonologique, une grande partie de la morphologie et de la syntaxe et certains domaines du vocabulaire (parenté, couleur, temps, etc.)[1]. »

C'est ainsi qu'Uriel Weinreich définissait, en 1953, l'interférence dans son livre, *Languages in Contact*. Si cet ouvrage a marqué, et s'il est encore lu aujourd'hui,

1. Uriel Weinreich, *Languages in Contact*, New York, Publications of the Linguistic Circle of New York, n° 1, 1953 ; rééd. La Haye, Mouton, 1963, p. 1.

quarante ans après sa publication, c'est parce qu'il était le premier à faire le point, avec perspicacité et profondeur, sur les problèmes du bilinguisme. Mais la définition que nous venons de citer, qui pourrait s'appliquer au problème des langues en contact dans la société, sera uniquement utilisée par Weinreich en référence à l'individu bilingue. Il considérait en effet que des langues étaient dites en contact lorsqu'elles étaient utilisées alternativement par la même personne.

On peut distinguer trois types d'interférences : les interférences phoniques, les interférences syntaxiques et les interférences lexicales. Le tableau ci-dessous, emprunté à Weinreich, présente les phonèmes d'un dialecte alémanique parlé dans le village de Thusis (*schwyzertütsch*) et d'une variété de romanche parlée dans le village de Feldis (ces deux villages se trouvent dans les Grisons, en Suisse)[1].



L'opposition entre voyelles brèves et voyelles longues en schwyzertütsch pose des problèmes car, en romanche, les voyelles sont longues dans certains contextes et brèves dans d'autres. Il en résulte d'une

1. Uriel Weinreich, *ibid.*, p. 15.

part des confusions entre certains mots et d'autre part un « accent » romanche en schwyzertütsch. On trouvera un exemple semblable dans la difficulté que peuvent avoir les Français à réaliser la distinction anglaise entre le /i:/ long et le /i/ bref dans des mots comme *sheep* et *ship*, *sheet* et *shit*, etc.

Les interférences syntaxiques consistent à organiser la structure d'une phrase dans une langue B selon celle de la première langue A : ainsi un italianophone, sur le modèle courant de phrases comme *viene la pioggia* (« la pluie arrive ») ou *suona il telefono* (« le téléphone sonne »), pourra-t-il produire en français des phrases comme *sonne le téléphone*.

Dans le domaine lexical, les interférences les plus simples sont celles qui consistent à tomber dans le piège des faux amis, lorsqu'un Anglais par exemple utilise en français le mot *instance* avec le sens de « exemple » qu'il a dans sa langue. On peut aussi rencontrer des traductions mot à mot : *estar direito* chez les Portugais des États-Unis traduisant directement l'anglais *to be right*, « avoir raison ». Ou encore des créations dans une langue sur le modèle de l'autre : le français du Québec regorge d'exemples de ce type, comme *vivoir* pour « salon » (anglais *living room*). Mais l'interférence lexicale est surtout fréquente lorsque les deux langues n'organisent pas de la même façon l'expérience vécue. On trouve ainsi en français d'Afrique un usage du verbe *gagner* avec un sens très large (« gagner » mais aussi « avoir », « posséder ») sur le modèle de certaines langues africaines qui n'ont qu'un verbe pour ces notions. Ainsi, une phrase comme *Ma femme a gagné petit* signifiera qu'elle a eu un enfant et non pas qu'elle l'a gagné dans une quelconque loterie…

Poussée au bout de sa logique, l'interférence lexicale peut produire l'emprunt : plutôt que de chercher dans sa langue un équivalent difficile à trouver d'un mot de l'autre langue, on utilise directement ce mot en

l'adaptant à sa propre prononciation. Au contraire de l'interférence, phénomène individuel, l'emprunt est un phénomène collectif : toutes les langues ont emprunté à des langues voisines, parfois de façon massive (c'est le cas de l'anglais empruntant au français une grande partie de son vocabulaire), au point que l'on peut assister, en retour, à des réactions de nationalisme linguistique. Ainsi au Québec, et dans une moindre mesure en France, s'est développé un mouvement officiel de lutte contre les emprunts. Nous y reviendrons dans le dernier chapitre de ce livre.

II. – Les langues approximatives

Ce plurilinguisme pose évidemment un problème différent lorsqu'un locuteur se retrouve dans une communauté dont il ne connaît pas la langue. Nous avons ici deux cas de figure : il peut s'agir d'une personne de passage (un touriste par exemple), qui tentera alors d'avoir recours à une tierce langue connue de lui et de la communauté environnante : il utilise dans ce cas ce qu'on appelle une *langue véhiculaire*, notion sur laquelle nous reviendrons plus bas. Mais il peut aussi s'agir d'une personne ayant l'intention de rester dans cette communauté, et il lui faudra bien alors, pour s'assimiler, acquérir la langue de la communauté d'accueil. C'est la situation dans laquelle se trouvent les travailleurs migrants, qui arrivent dans leur pays d'accueil sans en connaître la langue, ou en ne la sachant que peu, et qui sont forcés de l'acquérir sur le tas. Ce type d'acquisition est intéressant à analyser. Voici par exemple un court extrait de conversation avec une immigrée espagnole à Paris qui témoigne bien de ce phénomène :

« Vous l'aviez connue avant de venir en France ?

– Ah non ! Mais non, c'est porque yo habia metté une annonce sur un, journal *Figaro*, y elle me va escrir.

Et ma une otra petite qui travaille à Paris va me mener[1]. »

On y trouve une approximation de français dont les caractéristiques montrent bien l'origine linguistique de la locutrice :

- des mots espagnols dans le texte « français » : *porque* à la place de *pourquoi*, *otra* à la place de *autre* ;
- des mots inventés, produits par une interférence entre les deux langues : *escrir* à la place d'*écrire* (on dit en espagnol *escribir*) ;
- des mélanges syntaxiques : *yo habia metté* pour *j'avais mis*, où l'on trouve à la fois un segment espagnol et un mauvais usage du participe passé irrégulier du verbe mettre, etc.

Il peut aussi se produire que cette situation concerne non plus un individu mais un groupe social, aux prises avec un autre groupe dont il ne parle pas la langue et qui ne parle pas la sienne. S'il n'y a pas de tierce langue disponible, et si les deux groupes ont besoin de communiquer, ils vont s'inventer une autre forme de langue approximative, en général une langue mixte. Ainsi, on a parlé jusqu'au XIX^e^ siècle dans les ports de la mer Méditerranée la *lingua franca*, forme linguistique à base d'italien avec un vocabulaire empruntant en outre aux autres langues du pourtour méditerranéen. Molière, dans *Le Bourgeois gentilhomme* (acte IV, scène v), a recréé un passage en lingua franca. Il ne s'agit sans doute pas vraiment du sabir tel qu'il était parlé dans les ports de la Méditerranée, mais on y trouve cependant les caractéristiques de cette forme linguistique : les pronoms y sont ramenés à une seule forme (*ti* pour « tu » et « toi ») et les verbes sont tous à l'infinitif :

1. Christine de Heredia, « Le français parlé des migrants », *J'cause français, non*, Paris, La Découverte, 1983, p. 101.

Le texte de Molière	*Traduction*
Se ti sabir	Si tu sais
Ti respondir	Tu réponds
Se non sabir	Si tu ne sais pas
Tazir, tazir	Tu te tais
Mi star Mufti	Je suis Mufti
Ti qui star ci,	Toi, qui es-tu ?
Non intendir	Si tu ne comprends pas
Tazir, tazir	Tu te tais

Ces formes, qu'on appelle des *sabirs*, sont donc utilisées à l'origine entre des communautés n'ayant pas de langue commune mais entretenant par exemple des relations commerciales. Il s'agit d'un système extrêmement restreint : quelques structures syntaxiques et un vocabulaire limité à des besoins de communication particuliers. Lorsque ces formes couvrent des besoins de communication plus larges, que leur système syntaxique est plus étoffé, on parle de *pidgins*, le premier exemple en étant le *pidgin english* qui s'est développé dans les contacts commerciaux entre Anglais et Chinois le long des côtes de la mer de Chine, empruntant son vocabulaire à l'anglais et sa syntaxe au chinois (l'origine du mot *pidgin* serait d'ailleurs dans la déformation du mot anglais *business*, ce qui indiquerait bien la fonction sociale de cette forme linguistique). Ces formes approximatives, au contraire des formes individuelles que nous avons évoquées plus haut dans les situations d'acquisition, ne sont en général pas destinées à évoluer vers une meilleure pratique de la langue : elles sont simplement des auxiliaires, que l'on utilise dans une situation de contact.

III. – Mélanges de langues, alternances codiques et stratégies linguistiques

Lorsqu'un individu est aux prises avec deux langues qu'il utilise tour à tour, il arrive qu'elles se mélangent dans son discours et qu'il produise des énoncés « bilingues ».

Il ne s'agit plus ici d'interférence mais, pourrait-on dire, de collage, du passage en un point du discours d'une langue à l'autre, que l'on appelle *mélange de langues* (sur l'anglais *code mixing*) ou *alternance codique* (sur l'anglais *code switching*), selon que le changement de langue se produit dans le cours d'une même phrase ou d'une phrase à l'autre. En voici un premier exemple, extrait d'une conversation entre des femmes espagnoles vivant en Suisse, à Neuchâtel : « Ahora, con cabronas de ponermelos en lo alto de la oficina, en lo alto de la mesa de la oficina ; sin explicacion y sin na ! ça va pas ou quoi ? Por quien se toma este imbecil que apesta a vaca, eh ? Y subo y digo, dice : bueno, je vais voir si je trouve, je monte tout de suite[1]. »

L'insertion de segments en français (*ça va pas ou quoi ? Je vais voir si je trouve, je monte tout de suite*) dans un discours en espagnol témoigne ici de la situation de contact de langues dans laquelle se trouve la locutrice et constitue selon l'auteur du « parler bilingue », un mélange de langues en fait, qui est bien sûr commun aux deux personnes : toutes deux sont espagnoles, toutes deux travaillent dans un contexte francophone, et l'alternance d'une langue à l'autre fonctionne le plus souvent comme citation d'un fragment de discours qui a été énoncé dans l'autre langue, ou comme façon d'ancrer le discours dans la réalité à laquelle il est fait référence : il n'y a pas, ici, de *stratégie* particulière.

L'alternance codique ou le mélange de langues peuvent répondre à des stratégies conversationnelles, faire du sens. Voici un exemple de conversation dans une famille d'origine italienne vivant au Canada anglophone. Les parents sont nés en Italie, les quatre enfants sont nés au Canada, et l'un d'entre eux, une fille, se trouve en France où elle poursuit ses études : toute la famille lui envoie une sorte

1. Jean-François de Pietro, « Vers une typologie des situations de contacts linguistiques », *Langage et Société*, n° 43, mars 1988.

de lettre orale, l'enregistrement sur une cassette d'une conversation collective à destination de l'absente :

Fille cadette. – E goes, « oh those Marines », dangerous « n ».

Frère. – Yup ; Stay away from i marins e tutti soldat (*rires*).

Fille cadette. – Et tut ji soldat (*rires*).

Père. – E mit sendemind nde…

Fille cadette. – Ah ! OK (Frère : rires).

Père. – E nen fa la Stupet la ma'Em a la fran^g pe fal : a deven'da kju smart envEc a'pu/… Anh.

Fille cadette. – Fa kju Stupet.

Père. – An (kju) keva ala skol e kju se devendEm le kos bon[1].

(Traduction : les passages en anglais sont en gras, ceux en italien sont en italiques.)

Fille cadette. – Il dit : oh, ces « Marines » sont dangereux.

Frère. – **Ouais, garde-toi de** « les marines et tous les soldats » (rires).

Fille cadette. – Et tous les soldats (rires).

Père. – Et sois sage, ne…

Fille cadette. – **Ah, OK !** (Frère : rires).

Père. – Et ne fais pas de bêtises. Nous l'envoyons en France pour qu'elle devienne plus **intelligente** au contraire, après…

Fille cadette. – Elle devient plus bête.

Père. – Ouais, plus on va à l'école et plus on devient de bonnes choses.

On voit que la fille cadette commence par citer, en anglais, une phrase que le père a proférée en italien (mais elle prononce le mot *dangerous* avec l'accent, italien, du père). Le frère enchaîne en anglais, mais cite la phrase du père en italien, phrase que reprend la fille, toujours en

1. Elena Silvestri, « Choix de langues et rôles discursifs dans une conversation familiale italo-canadienne », *Plurilinguismes*, n° 1, 1990, p. 75-90.

italien. Le père ne parle, tout au long de la conversation, que l'italien mais il introduit dans son discours un mot anglais (*smart*).

Ainsi, les changements de langue effectués par les enfants ont ici une fonction ironique : il s'agit chaque fois de se moquer du père, de mettre en scène linguistiquement son comportement, et l'alternance répond donc à une stratégie.

Voici encore un exemple d'alternance codique correspondant à ce que l'on appelle la *négociation* de la langue d'interaction[1]. Dans la ville de Montréal (Québec), majoritairement francophone, l'anglais est en telle progression que les francophones se défendent par toute une batterie de lois linguistiques. L'une des retombées de ces lois est que l'administration se doit d'être bilingue, et la situation est si complexe et si tendue que l'auteur note qu'acheter une paire de chaussettes est devenu un acte politique… Il faut, en effet, choisir la langue de communication sans imposer à l'autre une langue qu'il ne parle pas ou ne veut pas parler.

L'exemple ci-dessous est une conversation téléphonique entre la standardiste du service des rendez-vous d'un hôpital et une patiente :

Standardiste. – Central Booking, may I help you ?

Patiente. – Oui, allô ?

Standardiste. – Bureau de renseignement, est-ce que je peux vous aider ?

Patiente (passe au français).

Standardiste (reste au français).

Patiente (revient à l'anglais).

Standardiste (en anglais).

Patiente (revient au français).

Standardiste (en français).

1. Monica Heller, « Negociations of language choice in Montreal », in J. Gumperz, *Language and Social Identity*, Cambridge, Cambridge University Press, 1982, p. 108-118.

Patiente. – Êtes-vous française ou anglaise ?
Standardiste. – N'importe, j'suis ni l'une ni l'autre.
Patiente. – Mais…
Standardiste. – Ça ne fait rien.
(La communication se poursuit en français.)

Les commentaires de Monica Heller sur cette conversation peuvent être résumés ainsi :

– tout au long de l'échange, les deux locutrices font la preuve qu'elles dominent également l'anglais et le français ;
– mais dans sa première réplique (« oui, allô »), la patiente force la standardiste à répéter sa phrase, comme pour dire : « Nous ne pouvons pas avoir cette conversation sans savoir si nous décidons de parler anglais ou français. » Elle pourrait aussi bien demander « Parlez-vous français ? », et la standardiste pourrait répondre « Oui », ou « Un peu », ou demander à la patiente de parler lentement, ou encore appeler quelqu'un d'autre, se faire remplacer par un ou une francophone (elle est en effet anglophone et a spontanément décroché le téléphone en parlant anglais) ;
– la standardiste choisit en fait de suivre la patiente en français, celle-ci, peu satisfaite, passe à l'anglais puis demande à la standardiste quelle est sa langue. La standardiste refuse (en français) de répondre, et la conversation se poursuit en français. La *négociation* est alors terminée, et le choix d'une des langues signifie que la patiente a fait une demande explicite de parler français et que la standardiste estime que son français est suffisamment bon. La conversation peut alors se poursuivre.

En revanche, dans le cas suivant, la communication arrive à son terme avant que la négociation n'ait pu déboucher sur un accord. Cet échange a été enregistré au bar d'un hôtel, en Crète. Un client (qui parlait français avec

sa femme au moment où arrive le garçon) s'adresse au garçon en grec :

Client. – Kalispera (« Bonsoir »).

Le garçon lui répond en français, et les répliques vont alors alterner le grec et le français :

Garçon. – Bonsoir monsieur.

Client. – Echete Ouzo ? (« Avez-vous de l'ouzo ? »)

Garçon. – De l'ouzo, bien sûr monsieur.

Client. – Dia Ouza, parakalo (« deux ouzo s'il vous plaît » avec une faute de grec : dia au lieu de dio).

Garçon. – Dio ? (« deux ? » : le garçon reprend l'adjectif numéral sous sa forme correcte).

Client. – Nai, dio (« oui, deux » : le client a accepté la correction).

Garçon. – Tout de suite Monsieur.

Cette interaction peut paraître paradoxale puisque chacun, jusqu'au bout, y parle la langue de l'autre (le Grec n'intervient qu'une fois en grec, pour corriger une faute faite par le client) et que nous avons là un exemple presque caricatural d'alternance codique. L'interaction est trop courte pour que l'on puisse juger de la capacité de l'un et l'autre des interlocuteurs à aller plus loin dans une conversation dans l'une ou l'autre des langues. Mais il est clair que le garçon veut montrer ses compétences « professionnelles » en français et que le client tient à montrer qu'il peut parler grec. Ainsi, lorsqu'il dit Echete ouzo ?, il ne demande pas seulement s'il y a de l'ouzo (il y en a, bien sûr : l'ouzo est en Grèce la boisson nationale…), il montre en même temps qu'il peut poser cette question en grec (sachant très bien que dans ce bar d'un hôtel international le garçon comprendra le français ou l'anglais). De son côté le garçon pourrait se contenter d'un échange en grec : il comprend parfaitement ce que lui dit le client, ses réponses le montrent. Mais, en répondant à la question citée plus haut : « De l'ouzo, bien sûr Monsieur », il dit bien sûr qu'il y a de l'ouzo (ce qui est évident), il montre en même temps qu'il a compris le grec du client et il

indique surtout qu'il a repéré son accent français et qu'il préfère parler français ou refuse de parler grec.

Dans cette courte séquence il se passe donc beaucoup plus de choses que la simple commande de deux boissons : un conflit de rôles se joue sur le choix de la langue d'échange, et l'échange va à son terme sans qu'aucun des interlocuteurs ne cède sur ce point. Mais si personne ne l'emporte, le garçon a cependant marqué un point symbolique en corrigeant la faute du client en grec alors qu'il n'a pas, lui, fait de faute en français.

Voici maintenant une dernière situation de communication plurilingue : un colloque sur la langue galicienne, réuni au printemps 1991 dans une petite ville de Galice, où il y avait, outre une cinquantaine de participants galiciens, quatre invités étrangers :

- un Belge, de première langue flamande, parlant également le français, l'allemand, l'anglais, l'espagnol, et pratiquant une « approximation de galicien » en jouant sur la forme phonétique de son espagnol, produisant quelque chose comme de l'espagnol prononcé à la portugaise ;
- un Français 1 ne parlant que le français ;
- un Français 2 parlant l'espagnol, l'italien et l'anglais ;
- un Italien parlant le français, l'anglais et l'espagnol.

Entre eux, les quatre invités parlaient français. Dans le cadre du colloque, les Galiciens parlaient exclusivement galicien, le Belge parlait son « approximation de galicien », les trois autres parlaient français. Mais hors du colloque, dans les cafés ou les restaurants, les choses étaient différentes. Le Français 1 parlait français, bien sûr, sa seule langue, et tout le monde lui parlait en français. Les Galiciens, selon leur maîtrise du français, parlaient français ou espagnol aux trois autres invités étrangers, le Français 2 et l'Italien parlaient espagnol ou français aux Galiciens, français et parfois italien entre eux, le Belge parlait également français ou espagnol (réservant son

« approximation de galicien » aux situations formelles du colloque). C'est-à-dire que nous avions là des comportements linguistiques dictés soit par la nécessité (parler la seule langue que l'on domine : c'est le cas du Français 1), soit par des stratégies plus complexes : pour les Galiciens, refuser de parler espagnol au colloque consistait à montrer leurs positions politiques (ils étaient tous militants de leur langue) et, pour le Belge, parler son approximation de galicien revenait à afficher son soutien à la cause des Galiciens (je parle votre langue, je suis de votre côté…).

Mélanges de langues et alternances codiques peuvent donc avoir des fonctions diverses. Dans l'exemple italo-canadien il s'agissait de se moquer gentiment du père, dans l'exemple québécois il s'agissait de décider d'un commun accord quelle serait la langue de l'interaction, dans l'exemple grec chacun voulait démontrer sa compétence dans la langue de l'autre, etc., mais dans tous les cas le contact de langues produit des situations dans lesquelles le passage d'une langue à l'autre revêt une signification sociale. La standardiste et la patiente parviennent finalement à un accord (implicite), le client et le garçon de café n'y parviennent pas, les participants galiciens du colloque imposent leur langue aux invités étrangers, et chaque fois, la communication se produit malgré le plurilinguisme, ou plutôt sous forme de *gestion* du plurilinguisme. Mais le bilinguisme social n'est pas toujours aussi harmonieux, il peut également être conflictuel.

IV. – Le laboratoire créole

Le contact de langues ne produit pas seulement des interférences, des alternances et des stratégies : il engendre surtout un problème de communication sociale. Nous avons vu un type de réponse à ce problème sous forme de langues approximatives (sabir, pidgin) qui ont pour caractéristique de n'être la langue première de personne.

Mais certaines situations sociologiques font que les langues premières perdent de leur efficacité communicationnelle, lorsque les populations sont à ce point mêlées que personne ne parle la langue de l'autre. C'est par exemple ce qui s'est produit dans les déplacements d'esclaves d'Afrique vers les îles : d'origines différentes, mélangés sur les plantations, les Noirs ne pouvaient plus communiquer avec leurs langues premières et ont dû se créer une langue approximative, un pidgin.

Le mode d'émergence des créoles, liée au commerce triangulaire et au trafic des esclaves, est encore très discuté dans la communauté scientifique. Les linguistes, en effet, ne sont pas tous d'accord sur l'origine des pidgins et des créoles (deux hypothèses s'affrontent, l'hypothèse monogénétique et l'hypothèse polygénétique) et sur leurs processus de formation. Pour certains, un créole est un pidgin devenu langue vernaculaire (c'est-à-dire la langue première d'une communauté), ayant donc un lexique beaucoup plus étendu, une syntaxe plus élaborée et des domaines d'usage variés. Il serait alors caractérisé par un vocabulaire emprunté à la langue dominante, celle des planteurs, et une syntaxe fondée sur celle des langues africaines. D'autres soulignent qu'aucune description n'a pu vraiment prouver les rapports entre la grammaire des créoles et celles des langues africaines, et penchent plutôt pour l'hypothèse d'une *approximation d'approximation*. C'est la thèse de Robert Chaudenson qui, se fondant essentiellement sur le créole de la Réunion, soutient avec des arguments convaincants que, dans un premier temps, les esclaves, peu nombreux et vivant relativement proches de leurs maîtres, ont acquis un français sommaire (« une approximation de français ») et que, dans un second temps, le nombre d'esclaves se multipliant, les nouveaux venus ont appris le « français » auprès des esclaves plus anciens (acquérant donc une « approximation d'approximation »). Se fondant sur une analyse méticuleuse de l'histoire du peuplement de la Réunion, il voit donc dans l'histoire des

créoles trois phases. Une première phase d'installation tout d'abord : « L'importance numérique, économique et sociale du groupe blanc me donne de plus en plus à penser que cette phase a dû être beaucoup moins caractérisée par l'apparition d'un pidgin que par la réalisation d'approximations du français par des locuteurs qui, par ailleurs, conservaient, pour partie, l'usage de leur langue d'origine. »

La deuxième phase « commence avec le développement des cultures coloniales (café ou canne à sucre) qui entraînent des besoins de main-d'œuvre considérables et de très fortes immigrations qui réduisent très sensiblement le pourcentage des Blancs dans la population totale ». Au cours de cette phase, les nouveaux arrivants, qui travaillent sur les plantations, n'ont que peu de contacts avec les Blancs : ils sont encadrés par les premiers esclaves qui sont soit domestiques soit contremaîtres et leur transmettent leurs rudiments de français. Et c'est au cours de la troisième phase que le créole va s'établir définitivement comme un code séparé du français, au sein d'une relation diglossique (voir plus bas)[1].

Face à cela, Derek Bickerton avance une autre hypothèse fondée sur l'existence d'un « bioprogramme » inné en chaque individu qui va s'activer et donner naissance à un créole dans les situations sociales que nous avons esquissées et lorsque la langue dominante est imparfaitement transmise[2]. Le problème n'a pas fini d'être débattu, et la relative jeunesse des langues créoles fait que leur étude est extrêmement importante pour comprendre la genèse du langage. C'est pourquoi Claude Hagège a parlé de « laboratoire créole », expression que j'ai reprise en titre de ce paragraphe.

---

1. Robert Chaudenson, *Créole et enseignement du français*, Paris, L'Harmattan, 1989, p. 164-166.
2. Derek Bickerton, *Roots of Language*, Ann Arbor, Karoma Publishers, 1981.

Parmi les nombreux créoles parlés à travers le monde, il faut signaler ceux qui ont pour origine lexicale :

– l'anglais (à Hawaï, à la Jamaïque, en Mélanésie où on l'appelle beach-la-mar, bislama, bèchez-de-mer, à Sainte-Lucie, etc.) ;
– le français (en Haïti, à la Guadeloupe, à la Martinique, en Guyane, aux Seychelles, à la Réunion, etc.) ;
– l'espagnol (à Porto Rico, etc.) ;
– le portugais (dans les îles du Cap-Vert, etc.).

Malgré leur extrême variété et leurs différences importantes, les créoles manifestent parfois des traits communs. Par exemple, le redoublement emphatique des formes verbales :

– *sé manjé m ap manjé* (créole des Antilles françaises),
– *a nyam mi a nyam* (créole de la Jamaïque),
– *come mi ta come* (papamientu),

avec le même sens, « je suis en train de manger », et mot à mot « question de manger je suis en train de manger ».

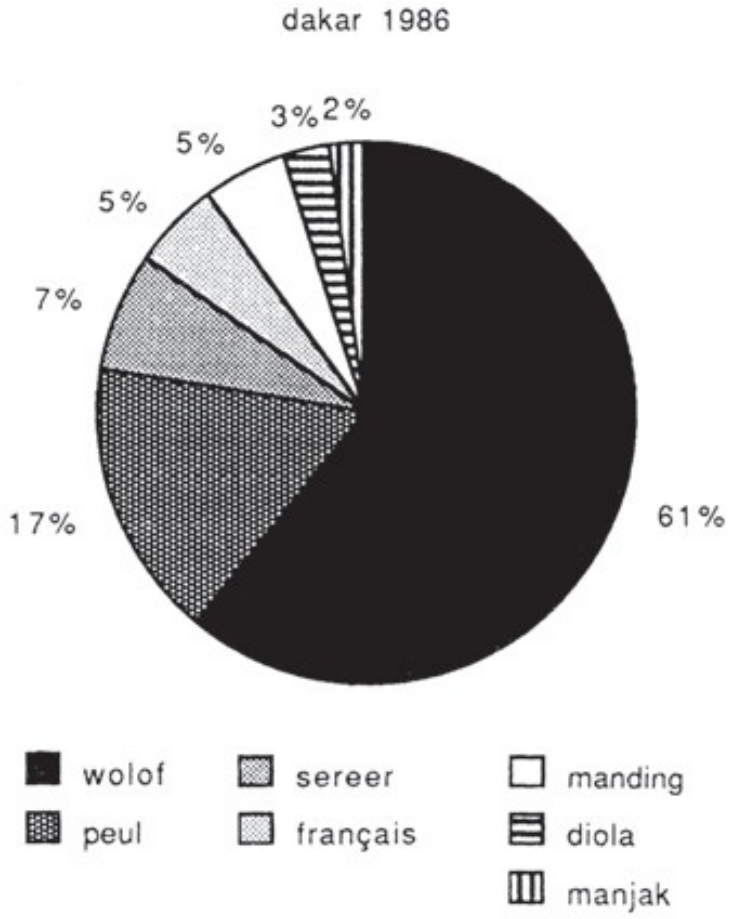
Quoi qu'il en soit, le créole est aujourd'hui une langue comme les autres, dont la seule caractéristique spécifique est dans son mode d'émergence particulier. Longtemps méprisés, considérés comme des formes inférieures, et à ce titre n'ayant pas accès aux fonctions officielles (enseignement, administration), les créoles sont aujourd'hui parfois promus au rang de langue officielle (aux Seychelles) et utilisés à titre expérimental dans l'enseignement (aux Antilles françaises et en Haïti).

V. – Les langues véhiculaires

Quelle que soit la théorie explicative de l'origine des créoles qui s'avérera, nous avons vu que leur émergence implique deux choses : un groupe dominant et minoritaire (et la langue de ce groupe) d'une part, une majorité

d'esclaves dominés d'autre part n'ayant plus de langue commune.

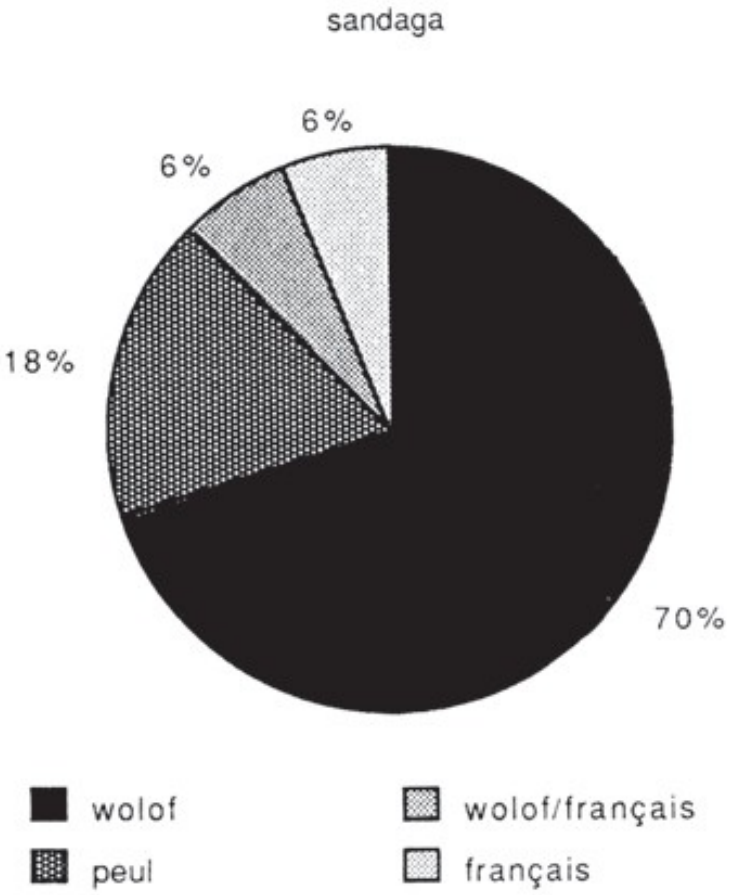
Mais il est d'autres situations dans lesquelles le plurilinguisme crée des difficultés de communication entre des groupes homogènes, ayant leurs propres langues, mais ayant du mal à communiquer entre eux. Nous allons prendre un exemple urbain d'une situation de ce type, celui de la capitale du Sénégal, Dakar. Selon une enquête effectuée en 1986 dans les écoles de cette ville, on y trouvait sept langues principales (il s'agit des langues premières, ou maternelles) (ci-dessous)[1].



À chacune de ces langues correspondent bien sûr des familles, des quartiers parfois, les locuteurs des langues minoritaires viennent de régions où ces langues sont majoritaires (le diola en Casamance, le peul dans la région du

1. Enquête non publiée de Martine Dreyfus, Dakar, 1986.

fleuve, à la frontière avec la Mauritanie, etc.), et la communication interne est donc assurée en peul, en diola ou en manjak. Mais que se passe-t-il lorsque des locuteurs du wolof, du peul et du diola se rencontrent ? En quelle langue vont-ils communiquer ? Une seconde enquête, menée sur un marché central de la ville, le marché de Sandaga, nous montre que trois langues seulement sont utilisées dans les interactions commerciales, et que le wolof y domine largement.



C'est-à-dire que des gens dont le wolof n'est pas la première langue l'utilisent pour communiquer avec d'autres personnes qui n'ont pas la même première langue qu'eux. C'est la définition d'une langue véhiculaire : *une langue utilisée pour la communication entre des groupes qui n'ont pas la même première langue*. Les exemples abondent : le swahili traversant l'Afrique des côtes est de l'Afrique jusqu'au Zaïre, le quichua dans la Cordillère des Andes, le sango en Centre-Afrique, le bambara/dioula en Afrique

de l'Ouest, etc.[1]. Dans tous ces cas, l'émergence d'une langue véhiculaire est la réponse que la pratique sociale et communicative des locuteurs donne au problème posé par le plurilinguisme de la communauté. Cette réponse peut s'incarner dans deux formes différentes :

– la langue véhiculaire peut être la langue d'un des groupes en présence (par exemple, le wolof au Sénégal, le bambara au Mali, etc.) ;
– la langue véhiculaire peut être une langue créée, langue composite empruntant aux différents codes en présence (par exemple, le munukutuba au Congo)[2].

Pour cerner l'importance de la fonction véhiculaire d'une langue, on calcule un *taux de véhicularité*, c'est-à-dire le rapport entre les locuteurs de cette langue et ceux qui l'ont pour langue première. Ainsi, une langue utilisée dans une communauté d'un million d'habitants dont 300 000 l'ont pour première langue aura un taux de véhicularité beaucoup plus important (70 %) qu'une langue utilisée dans une communauté d'un million d'habitants dont 700 000 l'ont pour première langue (30 %).

Dans tous les cas, il est intéressant d'étudier les rapports entre forme et fonction que nous révèle le phénomène véhiculaire. Ainsi, les travaux de Paul Nzété au Congo[3] et de Ndiassé Thiam au Sénégal[4] montrent que le lingala dans le premier cas, le wolof dans le second cas, se « simplifient » en fonction véhiculaire. La notion de simplification n'est bien entendu pas très scientifique, nous l'utilisons ici pour désigner le fait que la langue véhiculaire voit son système grammatical se réduire,

1. Voir Louis-Jean Calvet, *Les Langues véhiculaires*, Paris, Puf, « Que sais-je ? », n° 1916, 1981 (épuisé).
2. *Ibid.*, p. 78.
3. Paul Nzété, *Le Lingala de la chanson zaïro-congolaise de variétés*, thèse d'État, université René-Descartes, Paris, 1991.
4. Ndiassé Thiam, « L'évolution du wolof véhiculaire en milieu urbain sénégalais ; le contexte dakarois », *Plurilinguismes*, n° 2, 1990.

se régulariser. Ainsi le système de classes de ces deux langues est-il plus complexe en milieu rural (où elles sont surtout langues premières) qu'en milieu urbain, où elles sont surtout langues véhiculaires. Et le fait que la *fonction* d'une langue puisse avoir des retombées sur sa *forme* est une des découvertes fondamentales de la sociolinguistique.

VI. – La diglossie et les conflits linguistiques

Nous avons vu que le bilinguisme était pour Weinreich un phénomène individuel. C'est au bilinguisme social que va s'attaquer Ferguson lorsque, dans un article de 1959, il lance le concept de *diglossie*, coexistence dans une même communauté de deux formes linguistiques qu'il baptise « variété basse » et « variété haute »[1]. Pour l'illustrer, il prend quatre exemples : les situations arabophones (dialecte/arabe classique), la Grèce (demotiki/katharevoussa), Haïti (créole/français) et la partie germanophone de la Suisse (suisse allemand/hochdeutch). Et les situations de diglossie sont pour lui caractérisées par un ensemble de traits dont voici la liste :

- une répartition fonctionnelle des usages : on utilise la variété haute à l'église, dans les lettres, dans les discours, à l'université, etc., tandis qu'on utilise la variété basse dans les conversations familières, dans la littérature populaire, etc. ;
- le fait que la variété haute jouisse d'un prestige social dont ne jouit pas la variété basse ;
- le fait que la variété haute ait été utilisée pour produire une littérature reconnue et admirée ;
- le fait que la variété basse soit acquise « naturellement » (c'est la première langue des locuteurs) tandis que la variété haute est acquise à l'école ;

1. Charles Ferguson, « Diglossia », *Word*, vol. 15, 1959 ; cité ici dans Pier Paolo Giglioli, *Language and Social Context*, *op. cit.*, 1972.

– le fait que la variété haute soit fortement standardisée (grammaires, dictionnaires, etc.) ;
– le fait que la situation de diglossie soit stable, qu'elle puisse durer plusieurs siècles ;
– le fait que ces deux variétés d'une même langue, liées par une relation génétique, aient une grammaire, un lexique et une phonologie relativement divergents.

Tout cela lui permet de définir la diglossie comme « une situation linguistique relativement stable dans laquelle, outre les formes dialectales de la langue (qui peuvent inclure un standard, ou des standards régionaux), existe une variété superposée très divergente, hautement codifiée (souvent grammaticalement plus complexe), véhiculant un ensemble de littérature écrite vaste et respecté […], qui est surtout étudiée dans l'éducation formelle, utilisée à l'écrit ou dans un oral formel mais n'est utilisée pour la conversation ordinaire dans aucune partie de la communauté[1] ».

Quelques années plus tard, Joshua Fishman reprend le problème en élargissant la notion de diglossie[2]. Il distingue d'abord entre le bilinguisme, fait individuel, qui relève de la psycholinguistique, et la diglossie, phénomène social, puis ajoute qu'il peut y avoir diglossie entre plus de deux codes et, surtout, que ces codes n'ont pas besoin d'avoir une origine commune, une relation génétique. C'est-à-dire que n'importe quelle situation coloniale par exemple, mettant en présence une langue européenne et une langue africaine, relève de la diglossie. Restent les rapports entre bilinguisme et diglossie, que Fishman structure dans un tableau à double entrée.

Nous avons donc, selon Fishman, quatre situations polaires :

1. Charles Ferguson, *ibid.*, p. 245.
2. Joshua Fishman, « Bilingualism with and without diglossia, diglossia with and without bilingualism », *Journal of Social Issues*, vol. 23, n° 2, 1967.

		Diglossie	
		+	–
Bilinguisme	+	1 bilinguisme et diglossie	2 bilinguisme sans diglossie
	–	3 diglossie sans bilinguisme	4 ni diglossie, ni bilinguisme

1 *Bilinguisme et diglossie* : tous les membres de la communauté connaissent la forme haute et la forme basse. C'est le cas du Paraguay (espagnol et guarani).

2 *Bilinguisme sans diglossie* : il y a de nombreux individus bilingues dans une société, mais on n'utilise pas les formes linguistiques pour des usages spécifiques. Ce serait le cas de situations instables, de situations en transition entre une diglossie et une autre organisation de la communauté linguistique.

3 *Diglossie sans bilinguisme* : dans une communauté sociale il y a répartition fonctionnelle des usages entre deux langues, mais un groupe ne parle que la forme haute tandis que l'autre ne parle que la forme basse. Fishman cite ici le cas de la Russie tsariste (la noblesse parlait français, le peuple russe).

4 *Ni diglossie ni bilinguisme* : il n'y a qu'une seule langue, et cette situation n'est imaginable que dans une toute petite communauté.

La notion de *diglossie* a eu un important écho dans la sociolinguistique naissante, avant de prêter le flanc à un certain nombre de critiques, venant en particulier des chercheurs travaillant sur les créoles et sur le bilinguisme hispanique (surtout les sociolinguistes catalans). En effet, Ferguson comme Fishman avaient tendance à sous-estimer les conflits dont témoignent les situations de diglossie. Lorsque Ferguson introduisait la *stabilité* dans la définition du phénomène, il laissait entendre que ces situations pouvaient être harmonieuses et durables. Or la diglossie, tout au contraire, est en perpétuelle évolution.

Le cas de la Grèce, que Ferguson prenait comme l'un de ses exemples, est ainsi, trente ans après, complètement modifié : la variété « basse » de Ferguson, le grec démotique, est aujourd'hui langue officielle et l'ancienne variété « haute » ne sera bientôt qu'une langue morte. De façon plus générale, l'histoire nous montre que très souvent l'avenir des variétés « basses » est de devenir variété « haute » (ce fut le cas des langues romanes, français, espagnol, italien, etc., face au latin). C. Ferguson est d'ailleurs revenu de façon critique sur son texte dans un article de 1991, « Diglossia revisited » (*Southwest Journal of Linguistics*, vol. 10, n° 1).

On a ainsi l'impression que le succès du concept de diglossie s'explique par le moment historique où il fut lancé. À l'époque des indépendances africaines, de nombreux pays étaient aux prises avec une situation linguistique complexe : plurilinguisme d'une part et prédominance officielle de la langue coloniale d'autre part. Donnant un cadre théorique à cette situation, la diglossie tendait à la présenter comme normale, stable, à gommer le conflit linguistique dont elle témoignait, à justifier en quelque sorte qu'on n'y change rien (ce qui fut d'ailleurs le cas dans la plupart des pays décolonisés). Ces rapports entre science et idéologie ne sont pas chose rare, et nous présenterons dans le chapitre VI les problèmes qu'ils peuvent poser dans le cadre d'une politique linguistique.

VII. – La sociolinguistique urbaine

La ville est le lieu par excellence de ces contacts de langues. L'urbanisation et les migrations font en effet converger vers les grandes cités des groupes de locuteurs qui viennent avec leurs langues et créent ainsi du plurilinguisme avant, parfois, de s'assimiler à la langue dominante.

Ces situations ont mené à une sorte d'urbanisation de la linguistique, à des études de terrain que l'on a classées sous l'étiquette générale de « sociolinguistique urbaine » et qui peuvent se ramener à trois grands courants.

Le premier concerne l'analyse des rapports entre les langues dans les villes plurilingues. Les études portent ici soit sur le corpus (la forme des langues dans la ville, les effets de l'urbanisation sur les langues par le biais d'emprunts, de régularisation des formes irrégulières, etc.), soit sur le statut (les rapports entre les langues, sur les marchés par exemple, l'apparition de langues véhiculaires), soit sur les deux (on lira sur ce point la description que j'ai donnée de la situation linguistique de la ville d'Alexandrie, en Égypte, dans *Essais de linguistique*[1]), s'intéressant donc à la gestion *in vivo* du plurilinguisme.

Le deuxième courant concerne la ville définie non pas par son éventuel plurilinguisme, mais par sa « mise en mots », par l'appropriation des lieux à travers la langue, avec un accent mis sur l'analyse du discours et plus récemment une approche interdisciplinaire, en particulier en relation avec la géographie sociale. Cette approche, impulsée par les travaux de Thierry Bulot[2], repose sur l'idée que l'espace n'est pas une donnée mais une construction sociale, que l'action humaine a une dimension spatiale, et que les discours sur la ville modifient la perception du réel urbain, qu'ils finissent par devenir *la ville*. De ce point de vue, il faut signaler les travaux effectués sur les villes du Maghreb, en particulier sur la différence entre *urbains* et *citadins*[3], ou sur la façon dont leurs habitants

1. Louis-Jean Calvet, *Essais de linguistique*, Paris, Plon, 2004.
2. Thierry Bulot, « La sociolinguistique urbaine : une sociolinguistique de crise ? Premières considérations », in *Lieux de ville et identité*, t. I, Paris, L'Harmattan, « Marges linguistiques », 2004.
3. Voir Leila Messaoudi, « Parlers citadins, parlers urbains. Quelles différences ? », Cádiz, 5th Aida Conference, 2002 ; et Thierry Bulot, Leila

se nomment (autodésignation) et nomment les autres (hétérodésignation)[1].

Le troisième courant concerne la ville considérée comme productrice lexicale : de nombreuses études portent par exemple sur le langage des jeunes dans les cités, les banlieues, sur le « verlan » pour ce qui concerne le domaine français, et sur les rapports entre ces comportements linguistiques et les problèmes d'intégration.

On peut s'interroger sur le bien-fondé d'une classification spécifique, sur l'intérêt de distinguer dans l'ensemble de la linguistique une « sociolinguistique urbaine », et nous verrons au chapitre v que ces phénomènes peuvent être mieux décrits et expliqués dans le cadre d'une analyse allant « de l'analogique au digital ». Mais il demeure que la ville sera dans les décennies à venir un enjeu linguistique important. Le taux d'urbanisation est en effet en croissance constante dans l'ensemble du monde. Pour la période récente, nous avons par exemple les chiffres suivants : 29,4 % d'urbanisation en 1950 (c'est-à-dire que 29,4 % de la population mondiale vivait dans les villes), 37 % en 1970, 43,6 % en 1990, 48,2 % en 2000. Et la hiérarchie des mégapoles est elle-même en mutation. New York, Tokyo, Shanghai ou Mexico étaient en tête à la fin du xx^{e} siècle.

Or, selon les projections pour 2015, c'est-à-dire à court terme, si Tokyo restera la première mégapole du monde, avec 26 millions d'habitants, elle sera suivie de Bombay, Lagos, Dacca, São Paulo, Karachi, Mexico, New York, Jakarta, Calcutta, Delhi, etc. On trouve dans cette liste un grand nombre de villes plurilingues et, les zones les plus riches en langues (en Afrique, en Asie) étant en voie

---

Messaoudi (dir.), *Sociolinguistique urbaine, frontières et territoires*, Cortil-Wodon, Éditions Modulaires Européennes, 2003.

1. Ibtissem Chachou, « L'auto-désignation et l'hétéro-désignation comme procédés langagiers de ségrégation urbaine : le cas de la ville algérienne de Mostaganem », *Synergies Algérie*, n° 15, 2012.

d'urbanisation rapide, on peut prévoir que le nombre de langues y diminuera. Nous avons là un indicateur fiable, un paramètre dont la validité a été maintes fois testée : la ville, et en particulier la capitale, est une grande dévoreuse de langues, elle attire des ruraux ou des provinciaux qui viennent à la fois y gagner leur vie et y perdre en quelques générations leurs langues. De ce point de vue, la linguistique (ou la sociolinguistique) urbaine (ou s'intéressant à la ville) devrait dans l'avenir se pencher sur ces mégapoles, car c'est là que se manifestent le plus les contacts de langues, c'est là que se jouera dans l'avenir à la fois la disparition de langues et l'éventuelle apparition de nouvelles langues.