

Le cas des créoles français : mutualisation des ressources pour des dialectes apparentés

Pascal Vaillant, Université Paris XIII

Je m'appelle Pascal Vaillant, je travaille à l'université Paris 13, et je vais vous parler des langues créoles, sur lesquelles j'ai commencé à travailler lorsque j'étais enseignant-chercheur à l'université des Antilles et de la Guyane, d'abord en Martinique, puis en Guyane, dans l'équipe de recherche « Structure et Dynamiques des Langues » du CNRS, basée à Villejuif.

Je vais commencer par vous présenter quelques informations sur la nature de ces langues un peu particulières, en ce qu'elles forment un groupe de langues qui sont à la fois diverses mais comportent également de nombreux éléments communs, ce qui justifie une réflexion sur la nature de leur norme et sur l'intérêt de réfléchir à la mutualisation de ressources.

101

Histoire et genèse

Pour commencer, je voudrais vous présenter quelques éléments de compréhension historique et linguistique.

La situation de ces langues est assez particulière, car contrairement à beaucoup d'autres langues régionales de France métropolitaine, qu'elles soient de l'aire gallo-romane ou non, les créoles ne résultent pas du développement naturel lent, « écologique », de langues au fil des générations sans rupture brutale.

Les créoles se répartissent sur deux aires : l'aire Amériques-Caraïbes, et l'aire de l'océan Indien avec la Réunion, les îles Mascareignes, Maurice. Je vais, dans la suite de cet exposé, parler surtout de l'aire Amérique-Caraïbes, car c'est celle que je connais.

L'émergence des langues créoles dans ces régions est liée à une période historique particulière. Après la découverte du Nouveau Monde, l'Espagne et le Portugal se précipitent dans une ruée vers les nouvelles richesses. Le traité de Tordesillas, signé entre les souverains de l'Espagne et du Portugal en 1494, suite à l'initiative du pape Alexandre VI, répartissait le monde entre ces deux grandes puissances catholiques alliées à la papauté. L'Espagne s'est précipitée sur les mines d'argent et les grandes civilisations à piller (les empires inca et aztèque), et le Portugal sur les comptoirs côtiers où l'on pouvait installer des ports pour faire du commerce. Les pays européens qui étaient des puissances maritimes émergentes mais un peu en retard par rapport à l'Espagne et au Portugal, c'est-à-dire la France, l'Angleterre et les Pays-Bas, sont venus s'intercaler dans un petit coin, qui n'intéressait à l'époque ni les Portugais ni les Espagnols, sur ce qui est devenu les petites Antilles et la côte nord de l'Amérique du Sud (les trois Guyanes). C'est la raison pour laquelle aujourd'hui toute l'Amérique du Sud parle portugais ou espagnol, sauf ces trois petits pays qu'on appelle les Guyanes, et les îles Caraïbes.

102 Au départ, ces trois puissances maritimes du nord ont donc pris pied dans ce petit coin du monde, qu'elles ont commencé par arracher à l'emprise de l'Espagne, puis qu'elles se sont pendant un certain temps disputé entre elles. Une période d'économie agricole de plus en plus intensive a commencé, amenant à l'importation d'esclaves africains, une importation qui est devenue de plus en plus massive et qui a connu un siècle extrêmement intense entre 1750 et 1850 environ, au cours duquel beaucoup de gens étaient incorporés à la population à chaque génération sans avoir le temps de s'adapter. Quand des minorités arrivent dans un pays et apprennent la langue, il y a généralement une dose filée de personnes qui ont le temps de s'incorporer et d'apprendre une langue cible. Dans cette situation-là, il n'y avait plus d'équilibre. D'après une étude, réalisée par Jaques Arends, de ce qui s'est passé sur le plan démographique pendant ce siècle intense d'économie esclavagiste sur la colonie du Surinam, la moitié des gens présents à un instant T ne l'était pas dix ans auparavant. Le cas du Surinam est assez représentatif de l'évolution des autres colonies esclavagistes à la même époque. Il s'agit donc d'un arrivage massif de gens qui n'ont pas le temps de s'adapter, générant un laboratoire in vivo d'évolution linguistique assez original.

Comment comprendre ce qui se passe dans une telle situation d'évolution linguistique « chaotique » ?

La description proposée par Robert Chaudenson pour le créole réunionnais s'adapte à notre cas. La langue cible est le français. Il ne s'agit pas du français d'aujourd'hui tel que standardisé par l'Académie Française, mais sûrement une *koïnè* de français oraux de différentes régions – à l'origine surtout des dialectes de l'ouest de la France, mais on a vu au fil du 18^e siècle beaucoup d'arrivants de région parisienne et de Picardie notamment. Puis est venue cette période d'intense traite esclavagiste où la langue a perdu pied, puisque les gens qui arrivaient n'avaient pas le temps d'apprendre. Lorsqu'un esclave arrive sur une plantation, il ne suit pas de cours de langue, il est mis directement au travail, et les gens avec lesquels il est en contact ne sont pas des francophones : ce sont des gens qui eux-mêmes sont arrivés quelques années avant lui et qui n'ont pas eu le temps d'apprendre la langue. Après un siècle de ce cycle-là, la langue perd pied sur les bases de l'évolution éco-linguistique « naturelle ». Ce dérapage rapide et massif fait que si le créole martiniquais d'aujourd'hui est, en un certain sens, apparenté au français, de la même manière que l'on peut dire que le français est apparenté au latin, il l'est à la suite d'une dérive dans l'évolution du système et des caractéristiques typologiques qui les rendent extrêmement différents. Pourtant un siècle d'évolution seulement sépare le créole et le français, alors qu'il a fallu dix-sept siècles pour que le latin devienne du français.

103

Aire linguistique

Le contexte historique est intéressant pour comprendre la genèse de ces langues et leurs caractéristiques. Par ailleurs, il ne faut pas perdre de vue le fait qu'elles sont apparentées, parce que le contact entre les différentes îles et colonies n'a pas été rompu (des colons allaient d'une île à l'autre avec leurs esclaves), mais que la situation était quand même celle d'un archipel. Ce n'est pas un continuum dialectal comme sur un continent, où les gens gardent contact d'un village à l'autre. Nous sommes dans le cas d'îles, des communautés se forment, convergent à certains moments, puis divergent et restent longtemps dans des situations divergentes. La question de l'existence d'une véritable famille linguistique fait elle-même

débat : certains parlent des créoles comme des différents dialectes d'une même langue, alors que d'autres, à l'extrême, affirment que même au sein de la zone Amérique-Caraïbes, les différents créoles résultent d'évolutions entièrement parallèles. Une position raisonnable consiste à considérer que l'on est en présence d'une aire dialectale, qui n'est certes pas comparable à ce que l'on peut appeler un continuum dialectal comme dans le cas des aires des langues d'oc, mais qui constitue malgré tout une répartition géographique d'une famille de dialectes apparentés. C'est par exemple l'avis du linguiste allemand Stefan Pfänder, qui a étudié cette aire linguistique qu'il a baptisé la *Creolia* – en référence à la *Romania*, comme les linguistes appellent l'aire des langues romanes. Je précise au passage que pour des linguistes, le terme « dialecte » n'a pas la moindre connotation péjorative : il désigne simplement une langue, considérée comparativement à d'autres langues d'un groupe, et n'implique aucun jugement de valeur ou hiérarchisation.

104

En ce qui concerne l'aire que je connais, la zone Amériques-Caraïbes, des créoles sont parlés dans quatre ex-colonies françaises, cinq si on compte l'île de la Trinité, où le créole est en voie de disparition : Haïti – qui bien qu'indépendant depuis 1804, continue à avoir le créole et le français comme langues officielles ; et trois territoires qui sont encore aujourd'hui dans le giron de la République, et qui sont la Guyane Française, la Martinique et la Guadeloupe. Le créole à base française des Petites Antilles est par ailleurs parlé dans des îles voisines de la Martinique et de la Guadeloupe, qui sont la Dominique et Sainte-Lucie (un temps des colonies anglaises, aujourd'hui des pays indépendants). Pour simplifier le tableau, nous en resterons à quatre variantes principales : Haïti, Guadeloupe, Martinique, Guyane.

Traits caractéristiques des créoles : similitudes et divergences

Nous sommes dans le cadre de langues apparentées et dans une certaine mesure inter-compréhensibles, malgré des différences phonologiques et syntaxiques que je vais un peu expliciter.

En ce qui concerne les caractéristiques de ces langues, le lexique est à base française à environ 90 %, selon les estimations de Marie-Christine Hazaël-Massieux. En revanche, au niveau typologique et au niveau grammatical,

ce n'est plus du français. L'ordre y est certes aussi Sujet-Verbe-Objet, mais ce sont des langues isolantes, avec des mots invariables, une grammaire très analytique, pas de morphologie flexionnelle, et pour ainsi dire pas de morphologie dérivationnelle à part quelques réintroductions récentes. Par ailleurs, on a une caractéristique très intéressante que le créole partage en tant que langue isolante avec d'autres langues comme le chinois par exemple, qui est la très grande plasticité des catégories morpho-syntaxiques : un mot peut sans problème être utilisé en tant que verbe dans une phrase, en tant que nom entre un déterminant et un adjectif dans une autre, etc.

Pour illustrer des caractéristiques communes et des caractéristiques divergentes de ces langues, je vais vous présenter un exemple de modélisation des parties communes et des parties divergentes, d'abord avec le système du marquage du temps et de l'aspect pour l'expression du verbe, puis avec le système de détermination dans le groupe nominal.

Groupe verbal

Le groupe verbal est organisé autour du prédicat de la phrase (appelons-le ici, pour simplifier, le verbe, bien qu'il puisse parfois s'agir d'une autre catégorie de mots). La phrase typique, dans sa forme minimale, comprend donc un sujet et un verbe, par exemple « nous dansé ».

105

- Par défaut, un pronom et un verbe, sans autre mot, expriment l'aspect accompli : *nou dansé* = « nous avons dansé ».
- La particule *ka* permet d'exprimer l'imperfectif (action fréquente ou répétée) ou le progressif (action vue dans son développement) : *nou ka dansé* = « nous sommes en train de danser ».
- La particule *ké* exprime l'aspect prospectif, le futur : *nou ké dansé* = « nous danserons ».
- La particule *té* sert à exprimer le passé : *nou té dansé* = « nous avons dansé ».
- Il est possible de combiner *té* et *ka* pour exprimer l'imperfectif dans le passé : *nou té ka dansé* = « nous étions en train de danser ».
- La combinaison de *té* et *ké* donne le sens d'un conditionnel : « Nou té ké dansé », « nous danserions » (le conditionnel est ici exprimé comme une possibilité future vue depuis un point situé dans le passé ; étymologiquement, c'est d'ailleurs la même image qui a donné naissance au conditionnel des langues romanes).

- On peut enfin même combiner les trois pour exprimer un conditionnel imperfectif : *nou té ké ka dansé* – comme on l’entend dans une chanson de zouk très célèbre : *kolé séré nou té ké ka dansé* – : « nous serions en train de danser aujourd’hui » (sous-entendu : si nous n’avions pas rompu il y a vingt ans).

On a donc là toutes les combinaisons possibles d’un système à trois unités, dont chacune peut alterner entre l’aspect zéro (non-exprimé) et l’aspect particule invariable (exprimée). Ces trois unités peuvent se combiner pour donner tous les effets de sens possibles.

<i>nou</i>	∅	∅	∅	<i>dansé</i>
	<i>té</i>	<i>ké</i>	<i>ka</i>	
	Passé	Prospectif	Imperfectif	

Groupe nominal

106

Pour le groupe nominal, il y a comme en ancien français la possibilité d’avoir un nom sans article avec une valeur générique : *moun* = « la personne », « les personnes », « les gens ». Il y a un article indéfini antéposé comme en français, et un article défini qui, dans tous ces créoles, est postposé. Il se présente sous une forme de base *la*, et vient probablement étymologiquement d’un déictique (« là »).

Au-delà de ces quelques points communs, la construction des groupes nominaux montre plusieurs divergences dans les différents créoles. L’une d’entre elles réside dans le démonstratif : il est antéposé et se combine avec le défini en guyanais, alors qu’en martiniquais il est postposé, et se combine également avec l’article défini, mais en s’amalgamant avec lui.

Pour le pluriel, on a un article défini postposé pluriel en guyanais, *moun yan* (= « les gens »), mais on a une particule pré-nominale « sé » en martiniquais et en guadeloupéen. Il s’agit d’un morphème « transparent », c’est-à-dire : qui ne sert qu’à exprimer une seule valeur, le pluriel. Il se combine donc avec d’autres articles ou des démonstratifs.

Les différentes possibilités pour la construction du groupe nominal sont illustrées dans le tableau ci-dessous.

		haït.	guad.	mart.	guya.	français
Degré générique		<i>moun</i>	<i>moun</i>	<i>moun</i>	<i>moun</i>	personne (humain)
Singulier	indéfini	<i>yon moun</i>	<i>on moun</i>	<i>an moun</i>	<i>roun moun</i>	une personne
	défini	<i>moun nan</i>	<i>moun la</i>	<i>moun lan</i>	<i>moun an</i>	la personne
	démonstratif	<i>moun sa a</i>	<i>moun lasa</i>	<i>moun tala</i>	<i>sa moun an</i>	cette personne
Pluriel	indéfini	<i>moun</i>	<i>moun</i>	<i>moun</i>	<i>moun</i>	des personnes
	défini	<i>moun yo</i>	<i>sé moun la</i>	<i>sé moun lan</i>	<i>moun yan</i>	les personnes
	démonstratif	<i>moun sa yo</i>	<i>sé moun lasa</i>	<i>sé moun tala</i>	<i>sa moun yan</i>	ces personnes

107

Modélisation de la langue et mutualisation des ressources

Pour faire entrer ces langues dans l'ère des technologies de l'information, il faut les « mettre en boîte » dans un ordinateur. Il s'agit concrètement d'un travail de développement de ressources numériques (grammaires, lexiques) qui est très coûteux en temps et en ressources. Le point principal que je souhaite défendre dans cet exposé est que tout l'enjeu, dans le cas de langues apparentées, est de savoir mettre en commun une partie de ce travail.

En ce qui concerne la grammaire, ce que je propose est une modélisation de la langue qui permette de ne pas refaire ce travail autant de fois qu'il y a de dialectes à représenter. On l'a vu, les différents créoles ont des structures communes, et des structures spécifiques. Il s'agit de pouvoir les considérer comme une seule langue lorsque l'on représente leurs structures communes, et comme des langues différentes lorsque l'on représente leurs structures spécifiques. Comment faire ? Une approche qui a déjà été testée, avec succès, dans le domaine des technologies

de la langue, est celle des « méta-grammaires » : elle consiste à créer un étage abstrait de description des structures communes, qui engendre ensuite, dans un processus automatique, des grammaires partiellement incomplètes de chacune des langues ; on complète ensuite chacune de ces grammaires en y développant les parties spécifiques. Cependant, cette approche aboutit encore, au final, à un résultat consistant en quatre grammaires distinctes et étanches. Je pense que nous pouvons encore faire mieux, en développant une grammaire *modulaire*, c'est-à-dire une grammaire en « cercles concentriques », dont certaines parties seraient communes à plusieurs langues et d'autres spécifiques, et qui coexisteraient dans le même processus informatique.

108

J'ai proposé un prototype de ce type de grammaire modulaire. J'y utilise un formalisme bien connu (appelé TAG) qui représente chaque construction élémentaire d'un mot (par exemple la manière d'utiliser un article, ou de construire la phrase autour d'un verbe) comme un petit « arbre syntaxique » élémentaire, ces arbres pouvant ensuite se combiner entre eux. En général, les paramètres permettant de combiner ces petits arbres sont des paramètres internes à une langue (la catégorie grammaticale d'un mot, le genre, le nombre, la personne, etc.). La nouveauté dans le modèle proposé est que l'on ajoute, à ces paramètres, un paramètre qui représente tout simplement la langue en cours d'utilisation. On a donc une grammaire avec des arbres qui modélisent des aspects communs de ces différentes langues, et quelques arbres grammaticaux qui vont nous permettre d'exprimer des choses en spécifiant le paramètre de langue lorsqu'il s'agit de structures spécifiques à une langue ou à un sous-ensemble de langues. Tant que les structures sont communes, on laisse le paramètre de langue non-instancié. Ainsi, les règles de production du passé en utilisant la particule *té*, par exemple, font partie du noyau commun aux quatre variétés créoles. Lorsqu'il s'agit en revanche de faire comprendre à l'ordinateur que tel mot, ou telle construction, n'est valable que pour l'un des dialectes (par exemple l'article défini pluriel du guyanais), on indique la valeur obligatoire de ce paramètre de langue. À tout instant, ces différents petits arbres coexistent dans l'ordinateur : c'est le processus de construction qui va déterminer, en fonction de la langue utilisée, lesquels sont utilisables et combinables entre eux, et lesquels ne le sont pas.

Ce qui est amusant, pour me diriger vers la conclusion, c'est que le même problème de variation se pose au niveau du lexique. Dans le lexique, le vocabulaire commun est important, ne serait-ce que parce que nous avons affaire à des langues basées sur du vocabulaire français, qui ont subi des processus d'érosion et d'évolution comparables dans des zones géographiques comparables. Mais en même temps, tous les dictionnaires actuels de créoles – je me limite à parler de ceux que je connais, de l'aire Amérique-Caraïbes – résultent d'initiatives locales et n'ont pas été construits en commun : le dictionnaire de Barthélémy pour le guyanais ; celui de Frank pour le saint-lucien ; deux dictionnaires pour le guadeloupéen : celui de Pouillet, Telchid et Ludwig et celui de Tourneux et Barbotin ; et Confiant pour le martiniquais. Peut-on les rassembler en essayant de construire une norme sur la base de méta-phonèmes, comme on l'a fait pour d'autres aires dialectales, pour le breton avec l'étude des variations qui ont servi de base à l'élaboration de la norme *Kerne-Leon-Treger* (KLT) ? Ce n'est pas si évident que cela. Il y a quelques méta-phonèmes systématiques, mais surtout beaucoup de variation. La publication très récente d'un atlas linguistique des petites Antilles, résultant d'une enquête réalisée sous la direction de Guylaine Brun-Trigaud et de Jean Le Dû, montre qu'en réalité beaucoup de mots sont acceptés en compréhension dans plusieurs aires dialectales, donc dans plusieurs sites de ces territoires. J'ai simplifié en disant Guadeloupe-Martinique-Guyane, mais on ne parle pas de la même manière à l'ouest et au sud de la Martinique, et on voit apparaître parfois des répartitions de fréquences qui ne sont pas exactement les mêmes. Mais il y a surtout beaucoup de variation, au sein même des différentes communautés.

109

Cela pose la question, pour conclure sur les questions et les perspectives, d'abord de la difficulté de la détermination d'une norme, et ensuite de la pertinence de la mutualisation, qui est un peu je crois ce que les technologies de la langue pourraient apporter à ce type de groupe de langues apparentées.

La difficulté de la norme est de savoir de quelle norme on parle, et qui la fixe. A-t-on affaire à un ou plusieurs créoles ? Pour se limiter à l'aire Amériques-Caraïbes pour laquelle on est sûr qu'il s'agit de langues apparentées, certaines personnes, à un moment donné, ont eu envie de dire « le créole ». Il y a eu un débat autour des années 2000-2001, car le ministre de l'Éducation avait décidé de créer un CAPES de créole afin

de l'introduire comme enseignement de langue régionale au lycée. Il y a des gens qui ont insisté pour que l'on dise « un » créole. En même temps, d'autres arguments pourraient tout aussi bien démontrer que l'on devrait parler « des » créoles. « Le créole » a même été utilisé pour englober aussi les créoles de l'océan indien, ce qui linguistiquement est une absurdité, mais vous savez tous que dans la gestion des termes liés à la dénomination des langues il y a au moins autant de politique que de linguistique. Ce qui un jour est du serbo-croate devient vingt ans plus tard une langue serbe et une langue croate qui n'ont officiellement plus rien à voir. Alors, a-t-on affaire à une langue unique avec des variantes ? À plusieurs langues apparentées ? Y a-t-il un continuum ?

110

Une autre question est celle de la position de la norme au sein d'un continuum post-créole. C'est une notion qui existe dans les études créoles depuis 1971 : on distingue un *acrolecte*, qui est la langue vers laquelle les gens tendent quand ils font partie des couches sociales favorisées, qui est en général la langue de base, l'anglais dans les pays anglophones, le français dans les pays francophones, et un *basilecte* qui est le créole le plus « rustique », celui que l'on parle dans les campagnes. C'est une conception certes un peu simplificatrice, que j'énonce comme cela pour faire vite ; il y a eu différents modèles. Mais dans tous les cas une question centrale demeure : si l'on doit fixer une norme, où la fixe-t-on ? On ne peut pas fixer le français, car cela n'aurait aucun intérêt. Certaines personnes essaient de la fixer au basilecte le plus rustique possible, mais souvent les locuteurs de la langue elle-même ne s'y reconnaissent plus. Il y a donc des enjeux politiques et idéologiques à cette normalisation, ne serait-ce qu'au niveau de l'orthographe : est-ce qu'on veut favoriser l'autonomie de la langue par rapport au français, en faisant un choix de déviance, ou est-ce qu'on veut favoriser parfois un critère de lisibilité ? La question suscite un véritable débat, qui draine des enjeux à la fois idéologiques, linguistiques et didactiques. Fixer une norme proche de celle du français tend à favoriser la confusion, à ne pas aider ceux qui veulent prendre à bras le corps le problème de l'écriture en créole à réussir à franchir le pas. En outre, elle entretient, auprès d'un certain public, l'idée erronée que le créole est du mauvais français. Mais par ailleurs, dans des pays comme la Martinique, la Guadeloupe ou la Guyane, où les gens ont tous appris à lire et à écrire en français avant d'apprendre à lire et à écrire en créole, même si le créole est la langue première, la lisibilité n'est parfois

pas améliorée si l'on fait un choix de représentation orthographique qui est trop éloigné du français.

Pour terminer, la pertinence de la mutualisation des ressources, à mon sens, est importante, et je crois qu'on va en reparler dans les exposés suivants. Lorsqu'on a des langues qui se ressemblent, pourquoi refaire 100% du travail quand une partie du lexique et de la grammaire sont communes ? La mutualisation permet la représentation d'un système hétérogène — on rend compte non pas d'une seule langue, mais d'un ensemble de dialectes, ce qui n'est pas inintéressant en soi — ; non seulement elle permet un gain en termes de taille de description et de ressources de travail, mais en outre elle permet de prendre en compte la question du plurilinguisme, ce qui m'amène à mon mot de conclusion.

Où sont les gens qui parlent créole aujourd'hui ? Nombre d'entre eux sont plurilingues, pas seulement créole-français, mais de plusieurs créoles, et ils ne sont plus seulement dans les îles des Antilles, ils sont aussi dans l'agglomération parisienne, lyonnaise, bordelaise, toulousaine, nantaise, rouennaise, etc. Il y a une nouvelle *koïnnè* créole : si vous écoutez dans le métro, les gens parlent un créole qui n'est plus du créole martiniquais, mais du créole de seconde génération, qui est presque plus vivant que celui qui est promu dans les milieux universitaires en Martinique. C'est un véritable enjeu de recherche qui est très intéressant.

111

Pour conclure, je souhaite citer Amin Maalouf, sur le fait que les technologies de la langue peuvent être à la fois une menace mais aussi une opportunité pour les langues minoritaires, dans *Les Identités Meurtrières* (1998) :
« Je ne doute pas que la mondialisation menace la diversité culturelle, en particulier la diversité des langues et des modes de vie ; je suis même persuadé que cette menace est infiniment plus grave que par le passé [...]. Seulement, le monde d'aujourd'hui donne aussi à ceux qui veulent préserver les cultures menacées les moyens de se défendre. Au lieu de décliner et de disparaître dans l'indifférence comme ce fut le cas depuis des siècles, ces cultures ont désormais la possibilité de se battre pour leur survie ; ne serait-il pas absurde de ne pas en user ? »