HAL
open science

Le rôle de la notion de substrat dans la représentation des langues créoles : application aux créoles seychellois et mauricien

Jean-Philippe Watbled

▶ **To cite this version:**

Jean-Philippe Watbled. Le rôle de la notion de substrat dans la représentation des langues créoles : application aux créoles seychellois et mauricien. Dodille N. Idées et représentations coloniales dans l'Océan Indien, Presses de l'Université Paris-Sorbonne, pp.247–266, 2009. hal-00905483

HAL Id: hal-00905483

https://hal.science/hal-00905483

Submitted on 17 Dec 2020

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

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

LE RÔLE DE LA NOTION DE SUBSTRAT DANS LA REPRÉSENTATION DES LANGUES CRÉOLES : APPLICATION AUX CRÉOLES SEYCHELLOIS ET MAURICIEN

Jean-Philippe Watbled
Université de La Réunion

La notion de substrat (définie ci-après) est à l'origine de débats parfois houleux dans le domaine de la créolistique. Le positionnement théorique sur cette question précise du substrat est souvent susceptible d'être le révélateur de la conception que l'on se fait de la genèse des parlers créoles en général, ou de tel parler créole en particulier.

Les partisans d'une origine mixte des créoles, autrement dit d'un métissage linguistique lors de la phase de créolisation, risquent fort d'être partisans de l'hypothèse d'un substrat lorsque le prétendu métissage n'est pas apparent ou l'est peu. En effet, soit le métissage est avéré et évident, et l'on ne peut que s'incliner devant les faits ; soit il n'est pas patent, et étant donné que le créole semble en surface, pourrait-on dire, dérivé pour l'essentiel d'une seule langue, normalement celle du colon, l'on est alors placé devant l'alternative suivante : ou bien l'on se contente de reconnaître les faits tels qu'ils sont offerts dans les données, ou bien l'on postule un substrat, qui sera la ou les langues du ou des peuples dominés lors de la colonisation. Reste alors à savoir quel est le substrat, et comment démontrer son existence et le bien-fondé de l'hypothèse.

Il serait bien risqué de s'engager pour ou contre l'hypothèse du substrat pour ce qui est du développement de tout parler créole, ou de toute langue de contact. Je me contenterai par conséquent ici d'examiner la question uniquement pour un choix restreint de langues : le créole seychellois et le créole mauricien.

Je commencerai naturellement par proposer une définition de la notion de substrat[1], notion centrale dans le problème qui nous occupe ici :

> En cas de contact de langues, la structure d'une langue « dominée » et qui est la première ayant existé dans l'aire de contact, joue le rôle de substrat pour la langue « dominante », en cas d'influence sous-jacente de la langue « dominée » sur la langue « dominante ».

L'influence est sous-jacente, autrement dit non évidente, et elle doit être mise au jour à l'aide de méthodes appropriées. On notera le rôle des adjectifs « dominant » et « dominé » dans cette définition. La notion de supériorité économique et politique joue en effet un rôle crucial[2].

On a pu ainsi, par exemple, envisager le rôle d'un substrat gaulois dans l'évolution du latin parlé en Gaule. Comme l'indique d'ailleurs très bien Martinet, « C'est à des romanistes qu'on doit les théories du substrat, du superstrat et de l'adstrat »[3]. Cette référence à Martinet fait apparaître deux autres notions, associées à celles de substrat : le superstrat et l'adstrat. Le superstrat est évidemment l'inverse du substrat : il s'agit dans ce cas de l'influence d'une langue « dominante » sur une ou plusieurs langues « dominées ». Des langues comme le français, l'anglais, ou encore l'arabe, par exemple, ont pu exercer une influence sur d'autres langues et y laisser des traces, en général majoritairement lexicales (voir le lexique d'origine arabe en espagnol, par exemple). Le superstrat s'observe donc surtout dans le cas des emprunts lexicaux massifs. On ne parle toutefois de superstrat que dans le cas où la langue loin de s'imposer complètement, a simplement contribué à la formation de la langue qui en a subi l'influence.

Les termes substrat et superstrat induisent, on l'a compris, soit un ordre chronologique en cas de contact de langues sur un terrain donné, soit une hiérarchie, que n'induit pas le terme adstrat, lequel laisse entendre qu'il peut y avoir égalité dans l'influence réciproque, ou coexistence dans une même période.

Le problème avec les deux nouveaux venus, le superstrat et l'adstrat, est que pour le premier l'influence se limite en général au lexique, autrement dit à la composante la plus poreuse, la plus perméable d'une langue, et que cela n'est

1 Voir par exemple Georges Mounin, *Dictionnaire de la linguistique*, Paris, PUF, 1974, p. 82-83.

2 Il est important de préciser que les expressions « dominée » et « dominante » appliquées aux langues sont bien entendu des métonymies, et que ce sont en réalité les populations qui sont, au sens strict, « dominées » ou « dominantes », notamment dans le cas de la colonisation et de l'esclavage.

3 André Martinet, *Économie des changements phonétiques*, Berne, A. Francke, 1955, p. 15-16.

idees_coloniales_C0.indb 246 10/07/2008 18:06:45

guère révélateur d'une situation spécifique : le contact entre communautés linguistiques entraîne forcément des emprunts, la question se ramenant finalement à la mesure du caractère plus ou moins massif des emprunts en question. En outre, il est facile d'observer que dans la situation actuelle, les emprunts à l'anglais sont importants dans de nombreuses langues, surtout dans le cadre de la mondialisation : ajouter à cela la notion de superstrat éclairerait-il davantage l'analyse du phénomène ?

En ce qui concerne le second terme ajouté, l'adstrat, on peut tout simplement se demander s'il existe[4], ce qui revient à mettre en question l'existence des situations d'égalité lors de contacts : on peut légitimement en douter, sauf lors d'emprunts isolés ou non massifs, mais à nouveau, quel intérêt à ajouter un terme pour un phénomène somme toute bien banal ?

Mais revenons à celle des notions qui nous intéresse ici (le substrat). Dès son invention, cette notion a exercé un attrait particulier et elle a pu être exploitée sans grand discernement. Je citerai à nouveau Martinet[5] :

> […] on a bien peu d'indications valables sur le processus selon lequel les dialectes et les langues « minoritaires » sont lentement étouffés par les langues plus puissantes et sur la façon dont ces dernières peuvent être diversement « teintées » par les idiomes qui disparaissent sous leur poussée.

On peut s'interroger sur l'attrait, je dirais même la séduction, qu'exerce la notion de substrat, non pas seulement sur certains linguistes, mais aussi et surtout sur le profane. La raison est sans doute à chercher dans l'association avec la notion de métissage[6]. L'indice de popularité de cette dernière est en effet très élevé dans le champ des études postcoloniales. Le substrat présente l'avantage supplémentaire de ne pas être décelable à première vue : il constitue donc un métissage d'autant plus intéressant qu'il faut un effort pour le mettre au jour. Il s'agit en quelque sorte d'un métissage étouffé et masqué. Rappelons aussi qu'il s'agirait, dans cette représentation, de l'apport des populations serviles. Ce serait rendre justice à l'histoire de ces peuples que de révéler le métissage sous-jacent, qui serait le résidu synchronique des langues serviles à l'origine d'un processus complexe de créolisation, réalisé conjointement avec les langues des colons.

4 Le dictionnaire de Mounin auquel il a été fait réference plus haut donne l'exemple du castillan et du catalan (p. 83).

5 *Économie…*, *op. cit.*, *p.* 193.

6 Concernant l'idée que les créoles seraient des langues « métisses » et pour un point de vue critique sur cette idée, voir Robert Chaudenson, *La Créolisation : théorie, applications, implications,* Paris, L'Harmattan, 2003, p. 54-55.

Le problème est l'absence de prudence que l'on peut constater régulièrement dans le recours à la notion de substrat. Martinet, toujours lui, notait déjà[7] que le terme de substrat « en vaudrait bien un autre, si on ne l'avait galvaudé dans des cas où l'on ne savait pratiquement rien de la langue de substrat […] ». C'est ici le lieu de faire intervenir le célèbre principe du rasoir d'Ockham. Ce principe, on le sait, tient son nom de Guillaume d'Ockham, philosophe scolastique et théologien (1280-1349). Ce maître du nominalisme en philosophie du langage est le premier à avoir formulé ce remarquable principe de méthodologie des sciences : « Entia non sunt multiplicanda præter necessitatem » (= il ne faut pas multiplier les entités sans nécessité). Dans le cas qui nous occupe, on pourrait dire, concernant les notions de superstrat et d'adstrat, que leur utilité reste à démontrer, mais surtout, concernant la notion de substrat, qu'elle ne devrait être invoquée que dans les cas où cela s'avère indispensable, c'est-à-dire quand les autres pistes ont été épuisées sans fournir de résultat satisfaisant. Il ne faut avoir recours à la notion de substrat que si l'on peut prouver que les développements et les faits observés sont beaucoup plus plausibles sous son influence éventuelle, ce qui suppose d'ailleurs que ce substrat soit connu et bien décrit. Or dans le domaine de la créolisation, c'est loin d'être souvent le cas. Ajoutons qu'il convient de ne pas commettre d'anachronismes : il ne s'agit pas de poser comme substrat l'état actuel des langues africaines ou malgaches, mais leur état, bien moins documenté, à l'époque de la mise en place du système colonial.

Comme l'ensemble constitué par les méthodes classiques de reconstruction interne et les méthodes structurales (ce dernier terme étant pris au sens large) a fait ses preuves, il importe de commencer par appliquer précisément ces méthodes sur le terrain créole. Je me propose ici de montrer le bien-fondé de cette démarche en l'illustrant avec quelques exemples de développement.

Avant d'entrer dans le vif du sujet, je me permettrai quelques considérations d'ordre général sur notre espèce. L'être humain est un animal parlant, mais aussi un sujet parlant, ce qui est sans doute pléonastique, car il n'est probablement de sujet que parlant. Je pense que l'observation indique qu'il possède un inconscient linguistique et qu'il est biologiquement programmé pour le langage : nous sommes des êtres sémiotiques par nature. Malheureusement, cette programmation n'est active et optimale que dans certaines conditions, et notamment seul le jeune enfant est réellement performant pour ce qui est de l'acquisition, qui est en fait une re-création permanente de la langue.

7 *Économie…*, *op. cit.*, *p.* 296.

idees_coloniales_C0.indb 248 10/07/2008 18:06:45

Un autre paradoxe est patent : autant nous sommes des virtuoses du langage grâce à cette programmation, autant l'activité linguistique fait partie intégrante de notre nature, et autant nous sommes aptes à proférer des inepties ou au mieux des inexactitudes dès que nous traitons, non d'un thème extralinguistique, mais du langage lui-même. Autrement dit, pour un extraterrestre venu d'une civilisation plus avancée et qui nous observerait, nous sommes impressionnants sur le plan linguistique, mais assez lamentables sur le plan métalinguistique. C'est que c'est presque toujours notre idéologie qui parle alors, et c'est une image réflechie dans un miroir déformant que nous exprimons en général. Pour couronner le tout, chacun a un avis en la matière, et le pense très vite autorisé.

Cette réflexion nous amène naturellement à la notion de représentation. Les représentations non scientifiques du langage sont donc problématiques dans la logique de ce qui précède, mais lorsqu'il s'agit de langues créoles, dont la genèse est étroitement liée à une histoire cruelle et inhumaine, elles méritent d'être examinées avec davantage encore d'attention, et d'être confrontées à ce que nous observons et savons de ces langues, si l'on se positionne comme un observateur et descripteur impartial, utilisant toutes les ressources de la meilleure linguistique et de méthodes éprouvées.

Comme exemple typique d'expression d'une représentation idéologique courante du langage, voici un extrait d'un petit manuel d'initiation au créole mauricien destiné aux touristes anglophones[8] :

> […] you need not feel apprehensive about making mistakes when speaking it – you just cannot make any ! This is a language with no grammatical or other rules, no verbs to learn, and you may write it down in any way you like ![9]

Cet extrait évoque en fait une idéologie bien plus courante qu'on ne serait tenté de le penser. On observe évidemment une confusion naïve entre grammaire codifiée et grammaire mentale, ou même, plus simplement, grammaire au sens où les linguistes entendent ce terme.

Dans ce type de représentation, tous les poncifs sont à l'œuvre : réduction de la langue à sa composante lexicale ; ne mérite d'être appelé proprement langue que ce qui est doté d'une écriture officiellement codifiée, avec une

8 Jacques K. Lee, *Mauritius : Its Creole Language, The Ultimate Creole Phrase Book & Dictionary*, London, Nautilus Publishing Co, 1999.

9 « [...] il n'y a pas lieu d'éprouver de l'appréhension à l'idée de faire des fautes – en parlant le créole – il est tout bonnement impossible d'en faire ! C'est une langue sans règles d'aucune sorte, grammaticales ou autres, il n'y a pas de verbes à apprendre, et vous pouvez l'écrire comme bon vous semble ».

graphie étymologique de préférence ; il n'y a pas de langue digne de ce nom en l'absence d'une codification officielle (confusion entre point de vue politique/idéologique, et point de vue scientifique).

Parmi les autres poncifs, illustrés en partie dans l'extrait cité, ou sous-entendus : on n'a de « grammaire » que si la flexion est complexe, que si les désinences verbales sont en grand nombre, et que si la langue connaît des règles d'accord sujet/verbe… Seules ces propriétés assurent, dans ce type de représentation, la dignité de la langue, qui doit par ailleurs être « littéraire », pour accéder au même statut que les langues dites de culture.

On observe, dans cette idéologie linguistique, l'omission ou l'occultation d'un fait simple et clair, qui est que tout ce qui se parle est langue et obéit à des règles, et que la seule question linguistique « identitaire » est en réalité celle des frontières : il faut se demander à quelle langue appartient tel ou tel lecte[10], sachant que tout lecte ressortit forcément à une langue.

L'idéologie critiquée ici est en outre typique des sociétés que j'appellerai graphocentristes, autrement dit celles qui mettent la question graphique au centre, et qui envisagent la langue avant tout sous une forme écrite, oubliant que cette forme écrite n'est qu'un code secondaire destiné à donner une image spatialisée d'une réalité foncièrement acoustique : les Grecs de l'époque classique, pourtant férus de philosophie du langage (comme Platon dans le *Cratyle*), avaient déjà du mal à définir les éléments distinctifs (phonèmes) autrement qu'en les nommant « lettres ».

On éprouve donc la plus grande difficulté à séparer le bon grain scientifique de l'ivraie idéologique, représentationnelle, sans sombrer non plus dans un scientisme naïf lui aussi. On peut en tout cas lutter contre les préjugés et l'ethnocentrisme, mais sans se méprendre : la scientificité se distingue du politiquement correct.

La difficulté est accrue du fait que parmi les préjugés séduisants peuvent se trouver des idées généreuses et humanistes, mais fausses sur le plan factuel ou observationnel, ou plus généralement, sur le plan scientifique. Citons comme exemple l'idée tenace que les langues s'influencent les unes les autres profondément, et que le contact serait le seul moteur du changement. Or les langues changent dans tous les cas, avec ou sans contact. Parmi les facteurs internes du changement linguistique, on peut citer, sans prétendre être exhaustif, le phénomène de la coarticulation, l'asymétrie des organes (non spécifiques) de la parole[11], la re-création permanente de la langue par l'enfant,

10 On appelle *lecte* toute variété d'une langue donnée.

11 André Martinet, *Économie des changements phonétiques*, Berne, Francke, 1955 (*passim*).

étant entendu que la langue est non pas une « chose », mais un ensemble de structures mentales permettant le discours.

Dans ce cadre, le problème est la séduction du contact, de l'influence et de l'emprunt, ainsi que le « désir » de métissage. En outre, l'enjeu est particulièrement important dans le cas des langues créoles, puisque sont impliquées les situations de colonisation, ainsi que les situations postcoloniales ou néocoloniales.

Bien loin de la caricature présentée plus haut, les linguistes professionnels ne sont toutefois pas à l'abri de descriptions ou assertions erronées, sans doute dues à leurs propres représentations *a priori.* Ainsi peut-on lire ceci sous la plume d'Henriette Walter[12] :

> À l'origine des créoles : le portugais ?
> Les créoles sont généralement définis comme des langues dont la structure grammaticale de base est d'origine africaine ou asiatique et le lexique d'origine européenne, ici portugais.

ou encore ceci, toujours sur les créoles[13] :

> Ce sont des langues à part entière, dont la structure grammaticale est proche de celle des langues africaines et dont le lexique est en très grande majorité d'origine européenne.

Ce point de vue n'est ni plus ni moins qu'une variante de la théorie de la relexification[14], forme extrême de substratisme, qui prétend que les structures grammaticales des créoles sont celles des langues serviles, dont on aurait simplement « remplacé » le lexique par celui de la langue des colons : c'est l'ethnocentrisme à l'envers, ou l'envers de l'ethnocentrisme.

On ne peut que regretter en outre la vacuité d'expressions vagues comme « structure grammaticale de base […] d'origine africaine ou asiatique ».

Loin de cela, l'attitude scientifique consiste à ne poser de substrat que là où il est nettement observable ou démontrable, à appliquer les méthodes de

12 Henriette Walter, *L'Aventure des langues en occident*, Paris, Robert Laffont, 1994 (Livre de poche, p. 254).

13 Henriette Walter, *Le Français d'ici, de là, de là-bas*, Paris, JC Lattès, 1998 (Livre de poche, p. 123).

14 L'une des principales avocates de cette théorie est Claire Lefebvre, qui postule que le créole haïtien serait du fongbe (langue africaine) relexifié en français : voir notamment Claire Lefebvre, *Creole Genesis and the Acquisition of Grammar, The Case of Haitian Creole*, Cambridge University Press, 1998. Pour une critique détaillée et convaincante de cette approche, voir Robert Chaudenson, *La Créolisation…*, *op. cit.*, p. 286 *sq.*

reconstruction et à privilégier la piste interne à chaque fois que c'est possible, en pratiquant l'analogie avec des faits historiques avérés[15].

Pour bien examiner le développement linguistique et le rôle d'un éventuel substrat, il convient également de ne pas confondre langue et lexique et de ne pas réduire la langue au lexique (voir *supra*). Au contraire, il faut prendre en compte les diverses composantes de toute langue. Le schéma que je propose est la mise en relation du plan de l'expression (phonétique) avec le plan du contenu (sémantique), la syntaxe jouant le rôle d'interface entre ces deux plans. Par ailleurs, la langue dispose d'un ensemble ouvert d'unités lexicales, ou lexies.

Dans ce cadre théorique, on supposera que le lexique est la seule composante vraiment en prise sur le culturel (avec la néologie et les emprunts, notamment). La composante phonétique est totalement étrangère à cela ; quant aux structures grammaticales, elle sont de ce point de vue largement arbitraires, même si elles obéissent à des principes universels, et par conséquent non arbitraires. En bref, le moyen d'expression (le langage) n'est pas de même nature que le monde dont il nous permet de parler. La relation entre la langue ou le langage et l'extralinguistique est une relation d'hétérogénéité. Quant à l'universel que l'on observe dans le langage naturel, il tient sans doute autant à la nature de l'esprit humain qu'à des principes d'ordre fonctionnel.

Opposée à la conception théorique défendue ici, l'approche ultra-relativiste va de pair avec le goût pour le politiquement correct et la séduction exercée par l'idée d'un substrat constitué par les langues serviles.

En appliquant le principe du rasoir d'Ockham (voir *supra*), je considère que sur le plan grammatical, ainsi que sur le plan phonétique, il convient de privilégier la piste du développement interne de la langue de départ, en tenant compte des conditions sociales et historiques. Mon hypothèse est que, sur le plan phonétique, on peut au mieux postuler un substrat « négatif », dans le sens où il entraînerait la perte de phonèmes du français à la source du créole : le système serait ainsi soit préservé soit modifié, mais (dans le second cas de figure) uniquement dans le sens d'une réduction, puisqu'on ne note aucune « importation » de phonèmes. Nous verrons aussi que le recours à une théorie de la complexité complète cette hypothèse de manière fort convaincante. Sur le plan grammatical, les mots clés seront restructuration, réanalyse, re-création. En bref, dans cette hypothèse, il sera peu question de métissage linguistique.

15 Sur le terrain de la créolistique, Chaudenson est celui qui a le mieux dénoncé la substratomanie (je renvoie à l'ensemble de ses travaux, et plus spécialement à l'ouvrage mentionné en notes 5 et 12 ci-dessus).

idees_coloniales_C0.indb 252 10/07/2008 18:06:45

Le tok pisin est un pidgin à base anglaise parlé en Nouvelle-Guinée où il a statut de langue nationale officielle. Cette langue est particulièrement intéressante, car elle offre à mon sens un cas très net de substrat[16], qui révèle la différence profonde avec ce que l'on observe dans les créoles de l'océan Indien, langues pour lesquelles il est inutile d'avancer une hypothèse substratiste. En tok pisin, c'est le tolai, langue locale de Papouasie, qui joue le rôle de substrat. Le substrat se manifeste sous la forme d'une forte sensibilité à la notion grammaticale de transitivité. C'est ainsi que la forme pronominale objet *him* de l'anglais a pu être réinterprétée en tok pisin comme un suffixe verbal de transitivité, sous l'influence analogique du tolai. Voici un exemple :

yu nid-im ka (= angl. *you need a car*) = 'il te faut une voiture'

Ce qui frappe l'observateur est l'absence totale de faits troublants de ce genre, révélateurs d'un substrat, dans les créoles de l'océan Indien.

On remarquera que la situation du tok pisin est radicalement différente de celle des créoles de l'océan Indien : en effet, on peut penser que l'existence d'un substrat est plus probable lorsque le peuple dominé reste sur son territoire et n'est pas privé de sa langue ou de ses langues d'origine, ce qui est le cas pour les locuteurs du tok pisin, alors que les populations serviles ont été déplacées dans les îles de l'océan Indien avant la créolisation.

Dans une définition stricte (voir *supra*), on ne devrait d'ailleurs évoquer un substrat que lorsque les populations dominées ne sont pas déplacées. Dans les autres cas, on devrait se contenter d'examiner la question d'une éventuelle influence sous-jacente. Néanmoins, suivant un usage assez courant, je prendrai ici le terme substrat dans son acception la plus large, c'est-à-dire incluant le substrat en quelque sorte « déplacé » de son lieu d'origine.

Si les créoles de l'océan Indien s'étaient construits sous l'influence d'un substrat, resterait à se demander lequel. Il faudrait se tourner vers Madagascar et l'Afrique de l'est.

Les parlers malgaches ont un système morphosyntaxique dont on ne trouve aucune trace, sous quelque forme que ce soit, dans les créoles. On peut même affirmer qu'il n'y a pas plus opposé que le système syntaxique malgache et le système syntaxique créole. L'une des principales caractéristiques des langues austronésiennes ou malayo-polynésiennes[17], famille dont fait partie le

16 Voir William A. Foley, *The Papuan Languages of New Guinea*, Cambridge University Press, 1986, p. 36-40, John W.M. Verhaar, *Towards a Reference Grammar of Tok Pisin*, Honolulu, University of Hawai'i Press, 1995, p. 297-310, ainsi que Claude Hagège, *L'Homme de paroles*, Paris, Fayard, 1985 (Folio, p. 44).

17 Sur les origines du malgache, on peut lire avec profit Pierre R. Simon, *La Langue des ancêtres, Ny Fitenin-dRazana*, Paris, L'Harmattan, 2006.

malgache, est un très riche système de voix[18], avec des marqueurs intraverbaux de diathèse donnant une morphologie verbale très complexe : voix active, voix passive, voix bénéfactive, voix instrumentale, voix circonstancielle. Or les parlers créoles ont un système de voix pratiquement inexistant, et le verbe mauricien ou seychellois est strictement invariable, l'alternance bien connue entre formes longues et formes brèves étant régie par le contexte et non par des propriétés flexionnelles : un complément postverbal dans le même groupe rythmique entraîne la troncation de la voyelle finale de la base de nombreux verbes, comme dans l'exemple *i n don mwa sa liv* ('il m'a donné ce livre'), qui illustre la troncation du verbe *doné* : *doné* → *don*, avant l'objet direct dans ce cas. Il s'agit là de la seule variation possible du verbe, ce qui nous situe aux antipodes des systèmes malayo-polynésiens.

Les partisans du substrat en matière de grammaire devront nous expliquer comment l'une des principales propriétés des éventuelles langues de substrat, le système de diathèse en l'occurrence, ne se retrouve absolument pas dans les créoles.

Si l'on se dirige à présent vers l'Afrique, on sera bien en peine de décider quel pourrait être le substrat envisageable. Il est toutefois intéressant de donner au lecteur un aperçu du fonctionnement du swahili, la plus importante des langues de l'Afrique de l'est[19]. Dans cette langue typique du système bantou, les noms se rangent dans différentes classes signalées par des préfixes. Chaque classe se reconnaît ainsi au préfixe de singulier et au préfixe de pluriel. On a par exemple (avec les préfixes suivis d'un trait d'union) :

m-toto ('enfant'), *wa-toto* ('enfants') ; dérivé : *u-toto* ('enfance')
m-tu ('homme'), *wa-tu* ('hommes') ; dérivé : *u-tu* ('masculinité')
ji-tu ('homme très grand'), *ma-ji-tu* ('hommes très grands')
ki-tu ('chose'), *vi-tu* ('choses').

À nouveau, rien de tel, bien au contraire, dans les créoles de l'océan Indien, langues dans lesquelles les noms sont invariables. Or cela va d'autant plus à l'encontre de l'hypothèse d'un quelconque substrat de ce genre, que ce système de classes est si puissant qu'il fonctionne pour les noms étrangers

18 Voir Régis Rajemisa-Raolison, *Grammaire malgache*, Tananarive, Fianarantsoa, 1969, ch. V, Bakoly Domenichini-Ramiaramanana, *Le Malgache, essai de description sommaire*, Paris, SELAF, 1977, ou encore Narivelo Rajaonarimanana, *Grammaire moderne de la langue malgache*, Paris, L'asiathèque, 1995, p. 51-55.

19 Sur les langues africaines, on peut consulter Denis Creissels, *Description des langues négro-africaines et théorie syntaxique*, Grenoble, Ellug, 1991, et Bernd Heine & Derek Nurse (éd.), *African Languages, An Introduction*, Cambridge University Press, 2000. Sur le swahili, voir Benji Wald, *Swahili and the Bantu Languages*, dans Bernard Comrie (éd.), *The World's Major Languages*, London/Sydney, Croom Helm, 1987, p. 991-1014, et plus spécialement p. 1013 pour le problème traité ici.

empruntés par le swahili. Ainsi, le mot swahili pour « livre » est *kitabu*, manifestement emprunté à l'arabe *kitaab.* La syllabe initiale *ki* de ce nom arabe a été interprétée, sur des bases phonologiques[20], comme le marqueur de classe que l'on a dans l'exemple *ki-tu* ci-dessus. On a ainsi en swahili le paradigme suivant :

ki-tabu ('livre'), *vi-tabu* ('livres')

La forme de pluriel *vi-tabu* montre bien que le terme arabe a été reconstruit par analogie sur le modèle de *ki-tu, vi-tu*, et sa base est devenue *tabu*, ce qui n'est absolument pas le cas en arabe, langue sémitique avec racines consonantiques : la racine de ce nom en arabe[21] est en fait triconsonantique *k…t…b.* Le principe d'organisation morphologique sémitique est donc complètement passé inaperçu lors de la « naturalisation » du nom arabe en swahili.

L'exemple qui vient d'être présenté illustre typiquement les cas d'influence sous-jacente lors de la restructuration d'un terme emprunté en swahili et révèle l'importance cruciale du système de classes nominales dans ce type de langue. Or toute influence de ce phénomène typique du swahili et des langues apparentées est rigoureusement absente dans les produits de la créolisation.

Dans le domaine des pronoms personnels et des paradigmes verbaux, la démonstration donne les mêmes résultats négatifs. Voici un fragment du système des pronoms personnels en swahili :

Formes isolées :

mimi ('moi'), *wewe* ('toi'), *yeye* ('lui', 'elle'),

sisi ('nous'), *ninyi* ('vous'), *wao* ('eux', 'elles')

Préfixes sujets :

1sg : *ni-*, 2sg : *u-*, 3sg : *a-*

1pl : *tu-*, 2pl : *m-*, 3pl : *wa-*

Le swahili présente ici deux paradigmes : des pronoms isolés, c'est-à-dire disjoints (non conjoints au verbe) et des préfixes pronominaux, qui sont des indices pronominaux conjoints au verbe plutôt que de vrais pronoms.

Le verbe sémantiquement équivalent de 'lire' offre le paradigme suivant :

• Présent :

1sg : *ni-na-soma*, 2sg : *u-na-soma*, 3sg : *a-na-soma*

1pl : *tu-na-soma*, 2pl : *m-na-soma*, 3pl : *wa-na-soma*

• Passé :

20 Voir Gerrit J. Dimmendaal, *Morphology*, ch. 7, p. 189 de l'ouvrage collectif édité par Bernd Heine & Derek Nurse (voir note 17 ci-dessus).

21 Sur la structure morphologique dans les langues sémitiques, voir John J. McCarthy, *Formal Problems in Semitic Phonology and Morphology*, PhD Dissertation, M.I.T., 1979, ou Robert Hetzron (éd.), *The Semitic Languages*, London/New York, Routledge, 1997.

1sg : *ni-li-soma*, 2sg : *u-li-soma*, 3sg : *a-li-soma*
1pl : *tu-li-soma*, 2pl : *m-li-soma*, 3pl : *wa-li-soma*
• Futur :
1sg : *ni-ta-soma*, 2sg : *u-ta-soma*, 3sg : *a-ta-soma*
1pl : *tu-ta-soma*, 2pl : *m-ta-soma*, 3pl : *wa-ta-soma*

Or le français possède lui aussi des séries disjointes (pronoms) et des séries conjointes (indices) :

Formes disjointes :

moi, toi, lui, elle, ça, nous, vous, eux, elles.

Formes conjointes sujets :

je, tu, il, elle, on, nous, vous, ils, elles.

Il ne faut pas se laisser abuser par les cas d'homonymie pour *nous, vous, elle, elles,* car on a effectivement deux séries :

moi je dis… toi tu dis… lui il dit… elle elle dit…

Le dernier exemple montre que l'on a bien deux elle distincts à la suite : le premier est disjoint (série moi, toi…) et le second est conjoint (série je, tu…). Les formes disjointes sont d'authentiques pronoms, comme les formes isolées du swahili, tandis que les forme conjointes sont en fait des clitiques, ou quasi-affixes, qui sont des indices pronominaux plutôt que des pronoms[22], exactement comme les préfixes sujets du swahili.

Ce qui est frappant dans ce cas est que les seuls pronoms qui ont laissé des traces comme sujets (ou autres fonctions) dans les créoles sont les formes disjointes. Voici le système des pronoms personnels en seychellois et en mauricien[23] :

	Seychellois		Mauricien	
	formes faibles	formes fortes	formes faibles	formes fortes
1sg	*mõ*	*mwa*	*mo*	*mwa*
2sg	*ou*	*ou*	*to*	*twa*
3sg	*i*	*li*	*li*	*li*
1pl	*nou*	*nou*	*nou*	*nou*
2pl	*zot*	*zot*	*zot*	*zot*
3pl	*zot*	*zot*	*zot*	*zot*

22 Pour d'excellentes mises au point sur cette question, voir Denis Creissels, *Éléments de syntaxe générale*, Paris, PUF, 1995, p. 22 *sq.*, et Christian Touratier, *Analyse et théorie syntaxique*, Aix-en-Provence, Publications de l'Université de Provence, 2005, ch. 5.

23 Pour le mauricien, voir Philip Baker, *Kreol, A Description of Mauritian Creole*, London, C. Hurst & Company, 1972, et pour le seychellois, voir Chris Corne, *Seychelles Creole Grammar*, Tübingen, TBL Verlag Gunter Narr, 1977.

Les formes faibles apparaissent en position préverbale et les formes fortes en position postverbale. Il est clair en tout cas que les deux séries ont pour source les formes disjointes du français :

moi	> *mo* (MAU)	> *mõ* (SEYCH)
vous	> *ou*	
lui	> *li*	
nous	> *nou*	
vous-autres	> *zot*	
eux-autres	> *zot*	

Force est de reconnaître qu'il n'est absolument pas nécessaire de postuler un substrat et que les méthodes de reconstruction interne fonctionnent parfaitement ici. La forme *moi* /mwa/ est préservée en position forte[24], i.e. postverbale, mais en position faible, i.e. préverbale, elle devient /mo/, par fusion de la semi-voyelle /w/ et de la voyelle /a/ : /o/ a en effet le caractère labial de /w/ et un degré d'aperture moyen intermédiaire entre celui de /w/ (degré haut) et celui de /a/ (degré bas). Ensuite, en seychellois, la nasale bilabiale /m/ entraîne une nasalisation de la voyelle : /mõ/. La forme *ou* /u/ résulte simplement de l'effacement de l'initiale labiodentale voisée /v/, consonne particulièrement fragile dans de nombreuses langues. Les formes /twa/ et /to/ dérivent de *toi* exactement comme /mwa/ et /mo/ dérivent de *moi.* La forme /li/ résulte simplement de la chute de la semi-voyelle précédant le /i/. La forme *nou* /nu/ vient de *nous* /nu/ (inchangé). Enfin, la forme *zot* /zot/, prononcée [z⊃t], vient de la réduction de *vous-autres* et *eux-autres* à la seconde syllabe, la première, faible parce qu'inaccentuée, s'étant effacée : [vu'z⊃t] > [z⊃t], [ø'z⊃t] > [z⊃t].

La disparition des formes conjointes du français (*je, tu, il…*) affaiblit considérablement l'hypothèse du substrat (africain en l'occurrence), étant donné la remarquable convergence entre la langue des colons et une langue comme le swahili (ou autre). Or cette convergence n'a joué strictement aucun rôle. On pourrait objecter que l'on ne sait pas quelles langues ont servi de substrat, mais cette objection se retourne précisément contre les substratistes.

Avant de rechercher la présence d'un hypothétique substrat, il est préférable (voir Martinet, *Economie…*, *op. cit.*) d'appliquer les méthodes et les principes de la linguistique générale dans un cadre comparatiste : recherche des universaux, inventaire des processus phonétiques naturels, étude de

24 Voir en annexe la liste des symboles utilisés dans les transcriptions phonémiques.

la complexité linguistique devraient être prioritaires. L'acquisition de la langue doit être vue comme un processus de re-création, de ré-analyse et de restructuration permanente obéissant à des principes universels.

La méthode permet par exemple de reconstruire le système des auxiliaires en créole. Prenons le cas du créole mauricien. On a les possibilités suivantes pour le verbe *mãze* ('manger')[25] :

sujet + verbe :	*mo mãze*
sujet + imperfectif + verbe :	*mo pe mãze*
sujet + accompli + verbe :	*mo fin mãze*
sujet + prospectif + verbe :	*mo pou mãze*
sujet + passé + verbe :	*mo ti mãze*
sujet + passé + imperfectif + verbe :	*mo ti pe mãze*
sujet + passé + accompli + verbe :	*mo ti fin mãze, mo ti'n mãze*
sujet + passé + prospectif + verbe :	*mo ti pou mãze*

Les auxiliaires du mauricien sont des particules préverbales qui sont d'authentiques déterminants verbaux, ce qui n'est pas le cas des auxiliaires (*avoir, être*) du français, lesquels prennent en charge le mode, le temps et l'accord. Les particules préverbales sont ici strictement invariables et sont dépendantes du verbe. Leur source est facile à retrouver : l'imperfectif *pe* vient de (*être*) *après* ; l'accompli *fin* /fin/ vient de *finir* ; le prospectif *pou* vient de (*être*) *pour* ; enfin, le passé *ti* vient d'une forme d'imparfait de *être* : *l'était, était* > *té* > *ti.*

On pourrait être tenté de voir dans l'absence de verbe copule (= verbe *être*) en mauricien et en seychellois l'influence d'un substrat malgache ou africain. Voici un exemple en seychellois[26] : *banan i n mir komela* ('les bananes sont mûres à présent'). On voit dans cet énoncé que l'accompli (*n* < *fin*) est suivi directement de l'adjectif (*mir*), sans que la copule ne s'interpose. Deux objections s'opposent ici à l'idée d'un substrat. La première est que la copule de type *être* n'apporte rien sur le plan sémantique : dans notre exemple, c'est l'adjectif qui a une valeur prédicative, que ce soit en seychellois ou en français, donc qu'il y ait ou non un verbe copule. En effet, dans des langues comme le français ou l'anglais, la copule *être* ou *be* sert essentiellement à porter les

25 Voir Philip Baker, *Kreol, A Description of Mauritian Creole*, *op. cit.* p. 107.
26 Voir Chris Corne, *Seychelles Creole Grammar*, *op. cit.*, p. 62, à qui l'exemple est emprunté.

idees_coloniales_C0.indb 258 10/07/2008 18:06:46

marqueurs de mode/temps ou, plus généralement, les indices flexionnels[27]. Or après la créolisation, le verbe mauricien ou seychellois a perdu toute trace de flexion : il est donc simplement inutile de conserver une unité sémantiquement vide et ce qui n'était qu'un simple support de flexion n'est plus indispensable dans un système où l'auxiliation s'est généralisée et où la flexion a disparu. Le maintien d'une forme de copule dans certains contextes où elle est syntaxiquement requise vient à l'appui de cette explication : *kot ou liv i été* ('où est ton livre ?'). Dans cet énoncé interrogatif, la copule se maintient sous la forme *été.* On peut penser que ce maintien s'explique par l'absence de groupe syntaxique à la fois placé après la copule et en relation avec celle-ci : c'est en effet l'interrogatif *kot*, antéposé, qui est en relation avec la copule.

On voit bien que l'hypothèse du substrat est totalement inutile ici (application du rasoir d'Ockham : voir *supra*).

La méthode préconisée permet également d'expliquer de manière naturelle et pertinente l'évolution des systèmes phonologiques. Pour ces derniers, je propose d'appliquer la théorie de la marque[28], ou théorie de la complexité. Cette théorie telle que je la conçois postule que dans une opposition dichotomique, l'un des deux termes de cette opposition est plus complexe que l'autre. Autrement dit, il n'y a pas égalité entre les deux termes. Une voyelle nasalisée /õ/, par exemple, est plus complexe que sa contrepartie non nasalisée /o/. Une consonne voisée /b/ est plus complexe que sa partenaire non voisée /p/. On dit que le terme complexe de l'opposition est marqué, et que l'autre est non marqué. La complexité peut aussi résulter de la combinaison de traits, et non simplement de la présence de tel ou tel trait. Ainsi le français possède des voyelles complexes en ce sens que la combinaison de traits est complexe : il s'agit des voyelles /y/ (*lune*), /ø/ (*peu*), /œ/ (*peur*). Ces voyelles palatales labiales sont nettement marquées par rapport aux voyelles palatales non labiales /i/ (*lit*), /e/ (*thé*), /ɛ/ (*taie*), ou par rapport aux voyelles rétractées labiales comme /u/ (*loup*) ou /o/ (*haut*). Cette complexité vient de la combinaison de la labialité (lèvres arrondies) et de la palatalité (articulation antérieure, dans la région du palais dur) : les voyelles non marquées de ce point de vue sont soit palatales, mais non labiales, soit rétractées si elles sont labiales.

27 Sur cette question de la copule et pour un point de vue analogue, voir John Lyons, *Introduction to Theoretical Linguistics*, Cambridge University Press, 1968, p. 322-323.

28 Sur la théorie de la marque (angl. *markedness*), voir par exemple Roman Jakobson & Linda Waugh, *The Sound Shape of Language*, Brighton, Harvester Press, 1979, *passim*.

Que constate-t-on ? que ce sont précisément les voyelles marquées qui ont disparu en mauricien et en seychellois : /y/ > /i/, /ø, œ/ > /e/. Cette évolution va tout à fait dans le sens de ce que l'on constate lors de la comparaison des systèmes dans un cadre typologique : si une langue possède des voyelles palatales labiales, elle possède forcément des palatales non labiales et les rétractées labiales, mais l'inverse n'est pas vrai. On songe ici à un système comme celui de l'espagnol, avec ses cinq voyelles : une voyelle centrale basse /a/, deux voyelles palatales non labiales /i/, /e/, et deux voyelles rétractées labiales /u/, /o/. Ce n'est absolument pas un hasard si les seules voyelles françaises (non nasalisées) qui subsistent en mauricien et en seychellois sont précisément ces cinq unités.

Pour ce qui est des consonnes, le principe est exactement le même. Le français possède deux sifflantes : /s/ et /z/ (*casse* /kas/ ~ *case* /kaz/), mais aussi deux chuintantes, /š/ et /ž/ (*cache* /kaš/ ~ *cage* /kaž/). Les sifflantes sont des consonnes sulcales[29] laminales alvéolaires, tandis que les chuintantes sont des consonnes sulcales laminales palato-alvéolaires, articulées plus en arrière que les sifflantes. Les chuintantes sont nettement marquées par rapport aux sifflantes, en raison de leur articulation palato-alvéolaire. On constate que dans les langues qui les possèdent, les chuintantes sont acquises plus tardivement et avec plus de difficulté que les sifflantes. On constate également qu'il existe des langues qui ne connaissent pas l'opposition entre sifflantes et chuintantes. Dans ce cas, on a soit une articulation sifflante de type [s], soit une articulation intermédiaire entre le [s] et le [š] français, mais on risque peu de rencontrer une langue possédant une chuintante palato-alvéolaire, mais pas de sifflante. Or que s'est-il passé dans les créoles qui nous intéressent ici ? Les chuintantes ont disparu, en se transformant en sifflantes : /š/ > /s/, /ž/ > /z/ : ainsi le français *chauffer* donne *sofé* /sofe/ et *jour* donne *zour* /zuR/ par exemple.

En résumé, la théorie de la marque permet de prédire les points de fragilité du système et par conséquent les changements qui risquent de se produire, changements qui ont effectivement eu lieu. Il est vrai que les principes de la théorie de la marque peuvent ici être renforcés par ce que j'ai appelé plus haut un substrat « négatif ». Je veux dire par là qu'il existe deux types de complexité : (i) une complexité inhérente, qui est celle qui a été décrite plus haut ; (ii) une complexité relative, qui résulte de la confrontation entre la langue première et la langue qu'il s'agit d'acquérir. Prenons un exemple : un locuteur germanophone possède dans sa langue des voyelles palatales labiales comme phonèmes, et il n'éprouvera donc aucune difficulté à les acquérir

29 Ce terme signifie que la langue est creusée en sillon pour l'articulation de ces consonnes.

idees_coloniales_C0.indb 260 10/07/2008 18:06:46

en français. En revanche, un locuteur hispanophone ou arabophone ne les possède pas dans sa langue : il cherchera inconsciemment à les transformer en voyelles rétractées labiales (/y/ > /u/) ou en voyelles palatales non labiales (/y/ > /i/). Autre exemple : un Grec, qui ne possède pas de /š/ dans sa langue, risque de transformer le /š/ français en /s/.

Il est évident que dans de tels cas, la langue première sert de substrat (au sens large du terme), mais de manière négative, autrement dit en faisant disparaître un type d'articulation. On peut penser qu'un substrat de type « négatif » a été à l'œuvre lors de la créolisation : on peut l'interpréter comme la faible fréquence ou même l'absence des phonèmes marqués (voyelles palatales labiales, consonnes chuintantes)[30].

Ce cas du substrat « négatif » est à distinguer nettement du substrat « positif », qui consiste à introduire des types articulatoires inconnus de la langue dominante : or on ne constate rien de tel en mauricien et en seychellois.

Il est vrai que l'on peut considérer que ces parlers ont introduit au niveau phonétique une série complémentaire de voyelles non nasalisées : [i∂], [e∂], [u:] ou [u∂], [⊃:] ou [⊃∂], [α:]. Toutefois, ce jeu de nouvelles voyelles résulte non pas de l'influence d'un substrat quelconque, mais de l'affaiblissement du phonème /R/ en position postvocalique, avant consonne ou pause :

/iR/ → [i∂]
/eR/ → [e∂]
/uR/ → [u:], [u∂]
/oR/ → [⊃:], [⊃∂]
/aR/ → [α:].

La semi-vocalisation ou vocalisation du /R/ dans ce contexte entraîne soit un allongement de la voyelle qui le précède, soit une diphtongaison, ou à la fois allongement et rétraction dans le cas de /a/. Exemples : *sir* [si∂] ('certain'), *ter* [te∂] ('terre') *zour* [zu∂] ('jour'), *lepor* [lep⊃:] ('port'), *tar* [tα:] ('tard'). Toutefois il ne saurait être question de considérer que l'on a affaire à de nouveaux phonèmes : il s'agit simplement de réalisations phonétiques de séquences voyelle + /R/, et les représentations phonémiques des termes cités sont clairement /siR/, /teR/, /zuR/, /lepoR/, /taR/[31]. Les alternances justifient

30 Ce n'est pas un hasard si la présence ou l'absence des voyelles palatales labiales et des chuintantes constitue sur le terrain réunionnais la ligne de partage entre une variété de créole réunionnais plus francisée, possédant ces phonèmes, et une variété moins francisée, ne les possédant pas. Cette ligne de partage est absente du mauricien et du seychellois, tous deux bien plus proches de la variété réunionnaise moins francisée pour ce qui est du système phonologique.

31 J'ai utilisé par convention le symbole /R/ pour un phonème dont l'articulation de base est une spirante uvulaire.

cette analyse. Ainsi, par exemple, la forme courte du verbe *tiré*, signifiant *tirer*, donne phonétiquement [tiə], tandis que sa forme longue se prononce [tiRe] à peu près comme le verbe français *tirer* : on voit bien que la représentation phonémique de la base verbale est /tiR/.

L'affaiblissement du /R/ postvocalique (en position finale de syllabe) est un processus naturel particulièrement fréquent dans les langues qui possèdent une consonne rhotique, uvulaire notamment (ce qui est le cas dans les créoles étudiés ici). On l'observe dans une langue comme le danois, par exemple, qui possède aussi une spirante uvulaire particulièrement faible en position postvocalique. Il s'agit donc d'un processus résultant de facteurs purement internes correspondant à une tendance générale dans les langues[32].

Toutes ces observations révèlent que les processus à l'origine des créoles ne présentent aucune spécificité, mais obéissent en fait à des principes naturels et universels.

J'ai essayé de montrer que les méthodes de reconstruction appliquées aux créoles de l'océan Indien ne laissent pas une grande place aux substrats. La reconstruction interne, la mise en évidence de la dynamique de la langue, l'application d'une théorie bien conçue de la complexité, ainsi que de la naturalité des processus linguistiques, révèlent que bien souvent, et notamment dans le cas des créoles de l'océan Indien, il n'est ni utile ni nécessaire d'avoir recours à la notion de substrat.

Cette conclusion ne peut que décevoir les amoureux du métissage linguistique, mais il est des erreurs tenaces qui continueront longtemps de séduire : rien de tel en effet que la séduction d'une belle erreur. Toutefois, il faut se garder de tout dogmatisme : on peut en effet mettre en relation les faits linguistiques observés avec, par exemple, l'histoire du contact et des migrations, ou la privation ou non des langues de départ, entre autres paramètres. Tous ces facteurs favorisent ou non le rôle du substrat. L'essentiel est de ne pas s'obstiner à le voir là où il brille par son absence.

32 Il est inutile d'y voir une influence de l'anglais, dont le /r/ est d'ailleurs postalvéolaire, et non uvulaire. Outre qu'il s'agit de processus très courants dans les langues les plus diverses, on observe en effet des processus très proches en créole réunionnais, terrain sur lequel l'anglais ne peut être invoqué. L'influence de l'anglais en mauricien et en seychellois se limite essentiellement au lexique.

Annexe : symboles

Les symboles suivants ont été utilisés pour les transcriptions phonétiques ou phonémiques :

[i] comme dans *lit*
[e] comme dans *ré*
[ɛ] comme dans *sept*
[u] comme dans *sou*
[o] comme dans *sot*
[⊃] comme dans *sotte*
[a] comme dans *patte*
[α] comme dans *pâte* (en français standard)
[y] comme dans *lu*
[ø] comme dans *peu*
[œ] comme dans *peur*
[∂] comme dans la finale de l'anglais *cinema*
[w] comme dans *oiseau*
[v] comme dans *vous*
[s] comme dans *casse*
[z] comme dans *case*
[š] comme dans *cache*
[ž] comme dans *cage*
[R] comme dans *roue*

Les deux points après un symbole vocalique signalent la longueur.

Le tilde signale la nasalisation d'une voyelle : /õ/ (*bon*), /ã/ (*banc*).