



**HAL**  
open science

# Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique du contact

Joseph Jean François Nunez, Isabelle Léglise

► **To cite this version:**

Joseph Jean François Nunez, Isabelle Léglise. Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique du contact. Auzanneau Michelle; Bento Margaret; Leclère Malory. Espaces, mobilités et éducation plurilingues : éclairages d’Afrique ou d’ailleurs, Editions des archives contemporaines, 2017, 9782813002198. hal-01472768

**HAL Id: hal-01472768**

**<https://hal.science/hal-01472768>**

Submitted on 21 Feb 2017

**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.

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

## Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique du contact

Joseph Jean François NUNEZ (INALCO, SeDyL) et Isabelle LEGLISE (CNRS, UMR 8202 SeDyL)

### Introduction

Les mobilités et l'urbanisation croissante ou superdiversité ont pour conséquence le contact de répertoires linguistiques hétérogènes agissant comme ressources disponibles pour les locuteurs. Les pratiques langagières hétérogènes qui en résultent en sont de bons exemples. Si ce type de productions plurilingues est fréquent au Sénégal, il pose un ensemble de problèmes pour annoter et décrire les phénomènes langagiers à l'œuvre et beaucoup reste à faire dans le domaine de la linguistique de corpus plurilingues. Nous présentons ici quelques choix d'annotation des données plurilingues et montrons que les choix actuels disponibles renvoient tous, explicitement ou implicitement, à des *langues matrices* que les concepts de *pratiques langagières* ou *linguaging* devraient permettre de contourner.

Nous nous appuyons sur une méthode d'annotation dédiée aux corpus plurilingues (Léglise & Alby 2013, Vaillant & Léglise 2014) que nous appliquons à un corpus de pratiques langagières hétérogènes mêlant des éléments du créole casamançais, du français et du wolof, enregistrées auprès d'adultes originaires de Casamance et habitant les villes de Dakar et Ziguinchor.

La présentation de quelques caractéristiques du corpus nous permet de discuter deux oppositions binaires généralement admises en linguistique du contact : l'opposition entre langue matrice et langue insérée, d'une part qui s'est progressivement imposée comme une évidence dans les analyses, et la dichotomie entre switching et mixing d'autre part.

### 1. Contact de langues et pratiques langagières plurilingues au Sénégal

L'Afrique occidentale est connue pour son multilinguisme (voir en particulier Amuzu et Singler, 2014) qui est dû historiquement à l'émergence de plusieurs empires suite à l'éclatement de l'empire du Mali et dont le corollaire est l'hybridité culturelle. Du point de vue synchronique, l'urbanisation est le facteur le plus important (Lüpke 2010, 1) car beaucoup de populations quittent la campagne pour se retrouver dans des villes en perpétuelles mutations. Au Sénégal, les études consacrées au multilinguisme ont été réalisées pour la plupart à Dakar et à Ziguinchor, avec une focalisation sur Ziguinchor, région réputée la plus multilingue<sup>1</sup> du pays (Dreyfus & Juillard 2004 ; Lüpke 2013).

A Ziguinchor, des travaux ont montré que les jeunes utilisent quotidiennement plus de langues dans leurs pratiques langagières qu'à Dakar où le multilinguisme semble se maintenir grâce aux migrants installés dans des quartiers défavorisés (Juillard 1991b ; 2007). L'hétérogénéité linguistique et culturelle de la population de Ziguinchor est importante<sup>2</sup> et l'usage de plusieurs langues au sein des familles<sup>3</sup> a été documenté (Juillard 1991a). Juillard (1995) a également montré, grâce à une approche interactionnelle au sein de réseaux sociaux, que le plurilinguisme des individus et l'utilisation de l'anglais, du créole de Casamance, du diola, du français et du wolof dépendent généralement des personnes et de la situation de communication. Cette approche a été également adoptée par Lüpke (à paraître) à Agnack, en zone rurale ziguinchoroise. Cette dernière étude, qui s'ajoute à un certain nombre de travaux documentant également le plurilinguisme africain en milieu rural (cf.

---

<sup>1</sup> Ndecky (2011) met également en exergue l'hétérogénéité linguistique de la population de la ville de Goudomp en Haute Casamance dans la région de Sédhiou.

<sup>2</sup> La présence de certaines langues à Ziguinchor comme le wolof est liée à des facteurs attractifs, économiques, administratifs, migratoires et religieux (Juillard 1991a, 439-444).

<sup>3</sup> Ces familles sont généralement très composites parce que constituées de personnes se réclamant de groupes linguistiques différents.

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

en particulier les actes du colloque de Libreville de 2000), montre que la traditionnelle distinction entre ville et campagne, qui voit la ville comme étant le lieu de la manifestation du plurilinguisme par opposition à la campagne, ne tient pas. Le multilinguisme à Ziguinchor est, de plus, présenté par Dreyfus & Juillard (2004) et Juillard (1995) en termes de dichotomie entre centre et périphérie, le centre étant le lieu qui tend le plus vers des pratiques plus homogènes ou plus monolingues, et la périphérie apparaissant comme le lieu par excellence de la manifestation du plurilinguisme. Or, ces dernières années, certaines populations qui habitaient en zone périphérique ziguinchoroise se sont installées au centre de la ville tandis que d'autres ont fait le chemin inverse, modifiant ainsi la donne sociale et linguistique à la fois au centre-ville et en périphérie. Ces mouvements de populations bien qu'anciens (Mark 2002) se sont par ailleurs intensifiés avec l'extension de la ville, notamment vers le sud-est, ce qui a eu pour effet la diversification des répertoires linguistiques des populations installées au centre de la ville. Cette actualisation du plurilinguisme au sein des pratiques est favorisée par une attitude particulièrement positive des populations envers à la fois le multilinguisme sociétal et le plurilinguisme individuel<sup>4</sup>, attitude qui s'impose comme norme à Ziguinchor (Dreyfus & Juillard 2004 ; Lüpke 2013). Toutefois, la situation multilingue n'est pas homogène selon qu'on se trouve dans les quartiers du centre de la ville ou dans ceux de la périphérie.

Dans le domaine du contact de langues, les études se sont pour beaucoup concentrées, au Sénégal, sur le bilinguisme wolof-français (Ndao 1996 ; Juillard et al. 1994 ; Thiam 1994 ; McLaughlin 2001 ; 2008b ; Swigart 1992a). On assiste toutefois à l'émergence d'études incluant plus de deux langues, comme celles de Faye (2008) et de Nunez (2015). Si l'étude de Faye s'intéresse aux types d'alternance de langues et à leurs motivations du point de vue psycholinguistique, celle de Nunez s'intéresse à la manière dont trois langues s'agencent dans les pratiques langagières. Cobbinah (2010) pour sa part montre comment certaines catégories de classes nominales en baïnounck gubaher et diola banjal sont affectées par le contact de langues, tout comme McLaughlin (1997) avait montré la généralisation de la classe *b-* en wolof en raison du contact de langues et de processus de véhicularisation (Manessy 1995).

Dans son étude sur le wolof en contact avec le français dans la ville de Dakar, Thiam (1994) montre que les énoncés mixtes wolof-français sont produits le plus souvent par une population active que l'on peut associer à la classe moyenne. McLaughlin (2008b), conclut pour sa part que le français a été fossilisé dans le wolof, c'est-à-dire qu'on aurait des emprunts considérés comme « intégrés »<sup>5</sup> à un discours wolof. Or, la notion d'emprunt est souvent idéologique. Un même item peut être considéré comme un emprunt ou non selon que les locuteurs veulent converger ou diverger avec les orientations de leurs interlocuteurs (cf. la notion de convergence ou d'alignement (Giles, Coupland & Coupland 1991; Burt 1994; Du Bois 2002; Englebretson 2007). La notion d'emprunt devient alors fluctuante selon le contexte, les motivations du locuteur et selon l'interlocuteur. Ainsi un terme comme *simis* (chemise) est considéré par certains wolophones comme un emprunt au français chemise alors que d'autres le considèrent comme du wolof. Or, on sait que les chercheurs n'échappent pas à cette dimension idéologique (Gal & Irvine 1995). Les énoncés mixtes observés à Dakar où le wolof et le français alternent sont considérés généralement comme du « wolof urbain » parlé à Dakar et donc comme du « wolof de Dakar » (cf. Mc Laughlin 2001 ; Kiessling et Mous 2004), ce qui construit en négatif une représentation de langues « pures » ou plus authentiques, le français et le wolof (qui ne serait ni « urbain », ni « de Dakar »). Par ailleurs, Thiam (1994) a montré que ce qu'on appelle « wolof urbain » renvoyait à des réalités fort diverses. Le fait de labéliser ainsi ces énoncés mixtes laisse à penser qu'il s'agit de pratiques particulières, propres à la ville de Dakar, et qu'elles sont en partie extra-ordinaires, alors même que Swigart

---

<sup>4</sup> Nous entendons par multilinguisme, à la suite de nombreux travaux repris par le conseil de l'Europe, l'ensemble des langues parlées dans un milieu géographique donné et par plurilinguisme l'ensemble des langues présentes dans le répertoire d'un individu [http://www.coe.int/t/dg4/linguistic/division\\_fr.asp](http://www.coe.int/t/dg4/linguistic/division_fr.asp) (dernier accès 24.03.2016).

<sup>5</sup> La différence entre emprunts intégrés, non intégrés et code-switching a fait couler beaucoup d'encre. La méthode d'annotation adoptée permet de ne pas entrer dans ce débat (cf. 3).

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

(1992a et 1992b) montrait qu'il s'agissait de pratiques ordinaires. Or, ce type de pratiques bilingues est observable à Ziguinchor aussi bien en contexte urbain que rural et dans d'autres régions du Sénégal.

Ainsi, bien que l'on dispose d'un certain nombre de travaux décrivant le multilinguisme sociétal sénégalais et son actualisation dans différentes villes, l'urbanisation croissante et les mobilités de population ont pour conséquence le contact – au niveau individuel – de répertoires linguistiques hétérogènes. Les pratiques langagières bilingues wolof-français décrites dans la littérature sont pour nous de bons exemples de l'utilisation, par les locuteurs, de l'ensemble des ressources linguistiques disponibles au sein de leurs répertoires, et ce, quelle que soit leur connaissance – importante ou limitée – (Jørgensen et al. 2011, 33) des langues en présence. Ces ressources sont mobilisables, dans l'interaction, au sein de pratiques langagières hétérogènes (Léglise 2013) qui résultent d'une activité de 'linguaging' (Jørgensen, ibid).

Or, d'une part l'analyse précise de pratiques langagières actuelles faisant intervenir plus de deux langues reste encore à réaliser pour de nombreux groupes, comme les créolophones casamançais. Et d'autre part, si ce type de productions plurilingues est fréquent au Sénégal, il pose un ensemble de problèmes pour annoter et décrire les phénomènes langagiers à l'œuvre car les normes d'annotation ont essentiellement été développées pour des corpus monolingues. Les parties suivantes montrent comment on peut en traiter dans le cadre d'une linguistique de corpus plurilingues en développement.

## 2. Un corpus de pratiques langagières hétérogènes

Le corpus sur lequel nous nous appuyons (Nunez 2005) a été recueilli chez des Casamançais plurilingues<sup>6</sup> vivant à Dakar, Thiès et Ziguinchor. Il est constitué de six heures d'enregistrements dans différentes situations et de notes prises à la volée. 19 personnes (12 femmes et 7 hommes) âgées de 15 à 80 ans ont pris part aux discussions. Ces personnes, dont le niveau de scolarisation varie du primaire au supérieur, ont des activités socio-économiques différentes. Voici trois exemples de locuteurs aux parcours différents :

1. YD est une femme d'une trentaine d'année. Née à Ziguinchor, sa famille a migré à Dakar par vagues successives. Ce n'est qu'en 1999 que toute sa famille s'est regroupée à Dakar. YD a deux enfants issus d'un mariage avec un locuteur du wolof, originaire du centre du Sénégal. YD est enseignante dans une école privée française. YD parle anglais, créole de Casamance, diola, français, mandingue et peul. Elle parle français à l'école avec ses élèves et avec la plupart de ses collègues. Le créole de Casamance, le français et le wolof sont les langues qu'elle utilise avec sa fratrie. Quand elle s'adresse à ses enfants, c'est le créole de Casamance et le français qu'elle utilise. Elle parle quasi exclusivement le créole de Casamance avec sa mère tandis qu'avec son mari, c'est le wolof et le français qui sont leurs langues de communication. Elle parle diola avec certaines de ses tantes et mandingue avec certains de ses amis. Le wolof est la langue de son voisinage à Dakar.
2. BB est une femme d'une quarantaine d'année née à Ziguinchor. Elle a migré à Thiès où elle vit avec son mari, un wolophone. BB a six enfants dont deux issus de ce mariage. Elle a fait des études jusqu'au niveau secondaire et parle créole de Casamance, diola, français, mandingue et wolof. Elle parle wolof et créole de Casamance à ses enfants. Lorsqu'elle s'adresse à ses frères et sœurs, c'est en créole de Casamance. Avec son mari, c'est en wolof qu'ils échangent. Le wolof est la principale langue de son voisinage à Thiès. Elle parle français tout comme mandingue qu'avec certains de ses amis.
3. TL est une ziguinchoroise âgée de 80 ans. Elle vit à Ziguinchor où elle est née. TL est à la retraite. Quand elle était en activité, elle a été, entre autres, dactylographe et responsable d'animation. Elle s'est mariée une fois avec un wolophone et la seconde avec un créolophone. TL a arrêté ses études au

---

<sup>6</sup> Ces locuteurs parlent généralement le créole de Casamance, le français, le wolof et d'autres langues locales.

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

niveau primaire et parle six langues : créole de Casamance, diola fogni, diola floup, français, mandingue et wolof. Elle parle créole de Casamance avec ses enfants et certains de ses amis. Elle parle créole de Casamance et wolof avec ses petits-enfants. Elle parle français, diola, wolof et mandingue avec certaines personnes de son entourage.

Ce corpus, constitué de pratiques spontanées enregistrées en famille ou entre amis, est un corpus plurilingue (Léglise et Alby 2013) où différentes langues alternent. Ce corpus n'est donc pas un *monolingual bias* au sens de Gafaranga & Torras (2002, 4) c'est-à-dire un corpus dans lequel des locuteurs plurilingues tiennent un discours unilingue ; comme on peut le constater dans l'extrait ci-dessous où trois langues alternent : le créole de Casamance (transcrit en caractère normal), le français (souligné) et le wolof (en gras). Cet extrait est une transcription 'brute' peu développée afin d'exemplifier uniquement l'alternance des trois langues avec un jeu sur les polices de caractère, la méthode d'annotation du corpus fera l'objet de la partie suivante.<sup>7</sup>

Exemple 1 : extrait de corpus brut

- 1 L1 : a-mi jardin d'enfant N ka na dépenser ciw mais sikola normal N na N na **waaw**  
*Moi je ne vais pas dépenser beaucoup (d'argent) pour l'école maternelle mais (pour) l'école normale [= primaire] je (le) ferai oui.*
- 2 L2 : **léegi fu ñu tollu nii sax dañoo bañ sax dañoo bloquer sax pour yeneen nationalité yi fu mu tollu nii français rekk la ñuy jël yeneen ak tubaab rekk jex na à part ki /moo tax ma doon laajte ndax Delphine**  
*À l'heure actuelle ils ont refusé, ils ont même bloqué (l'accès à cette école) aux autres nationalités actuellement. Ce ne sont que les Français et d'autres Blancs (qu'ils reçoivent), c'est la raison pour laquelle je demandais si (le fils de) Delphine (a été accepté).*
- 3 L1 : **amul** ☒☒☒ /  
*Il n'y en a pas* ☒☒☒
- 4 L2 : e kusanj e i ka tené e pa e lebá-l  
*Cette chose, cette, ils il n'y en a pas pour qu'ils l'amènent*

Ce corpus exhibe des pratiques langagières hétérogènes qui ont très peu été étudiées et codifiées dans les grands corpus (cf. Léglise et Alby 2013) et pour lesquelles nous avons développé une méthode d'annotation spécifique que nous présentons dans la partie suivante.

### 3. Méthode d'annotation des pratiques langagières hétérogènes

Ce corpus a été balisé sous xml<sup>8</sup> grâce à un éditeur, Jaxe<sup>9</sup>, adapté aux annotations choisies<sup>10</sup> et à un schéma de document spécifique pour corpus plurilingue<sup>11</sup> qui s'inspire de normes de la TEI (Text Encoding Initiative)<sup>12</sup>.

<sup>7</sup> Les numéros en début d'énoncé figurent les tours de parole, L1 et L2 renvoient aux locuteurs 1 et 2, la traduction libre est indiquée en italiques. Pour l'ensemble des conventions de transcription, voir l'annexe en fin de texte.

<sup>8</sup> « Extensible Markup Language » ou langage de balisage extensible, langage informatique qui s'est imposé comme norme en linguistique de corpus pour faciliter l'échange automatisé de contenus complexes (comme les corpus que nous utilisons en linguistique, ces textes « enrichis », c'est-à-dire associant le texte brut et des couches d'annotations) et l'interopérabilité entre systèmes d'informations hétérogènes.

<sup>9</sup> Jaxe est un éditeur xml cf. <http://jaxe.sourceforge.net/fr/> dont les auteurs sont D. Guillaume, S. Ayadi, B. Tasche, O. Kykal, C. Dedieu, L. Guillon, B. Delacretaz, S. Kitschke.

<sup>10</sup> Les niveaux d'annotation ont été définis dans le cadre du projet ANR JCJC 09 CLAPOTY. Ces choix sont également partagés par les participants au projet LABEX EFL LC1 (Language contact and change) financé par ANR/CGI.

Nunez J. J. F. & Légise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

L'unité de base des transcriptions est le tour de parole. Ces tours de paroles sont visibles soit par l'initiale du locuteur, soit par le signe « – » voire par un numéro permettant de citer le passage en question avec plus de précisions. L'exemple suivant en est une illustration.

Exemple 2 : Visualisation d'un extrait de corpus annoté

094.	L1 :	<u>mais</u>	bebé	soŋ	bebé	soŋ	<u>continuer</u>	bebé	ke / (bruit)
<b>A [ C ] 1</b>		mais.CO	boire	seulement	boire	seulement	continuer	boire	DEM.D2
		CONJ	V	ADV	V	ADV	V	V	DET
		<i>Mais bois seulement, continue de boire ce</i>							
095.	L2 :	a-mi	N	ka	ta	<u>sauter</u>	misiñu		
<b>A [ C ] 2</b>		TOP-1SG.SBJ	1SG.SBJ	NEG	AS.HAB	sauter	médicament		
		PRN	PRN	PRT	PRT	V	N		
		<i>Moi je ne reste pas (des jours sans prendre des) médicaments</i>							
096.	L1 :	misiñu	k-e		falá-bu	(bruit)			
<b>A 1</b>		médicament	que.REL.OBJ-3PL.SBJ		AS.PRF.dire-2SG.OBJ				
		N	CONJ-PRN		V				
		<i>Le médicament qu'on t'a prescrit</i>							
097.	L2 :	a-mi	N	ka	ta	<u>sauter</u>	misiñu		
<b>A [ C ] 2</b>		TOP-1SG.SBJ	1SG.SBJ	NEG	AS.HAB	sauter	médicament		
		PRN	PRN	PRT	PRT	V	N		
		<i>Moi je ne reste pas (des jours sans prendre mes) médicaments</i>							

Cet exemple propose, sur la ligne de transcription principale, une représentation visuelle (ou visualisation) des langues de l'énoncé grâce à un jeu sur la police de caractère pour identifier le français (souligné), le wolof (gras) et le créole (times normal). On note, dans cet exemple, que seules deux langues alternent : le créole et le français (quelques insertions). La seconde ligne de transcription propose une transcription juxtalinéaire avec des gloses morphosyntaxiques<sup>13</sup> désormais classiques pour tout exemple linguistique, la troisième ligne de transcription propose l'identification des parties du discours correspondantes à chaque unité linguistique<sup>14</sup>. Enfin, la quatrième ligne propose une traduction libre en italiques.

Sur la base des informations recueillies pour chaque tour de parole, un codage est automatiquement proposé sous le numéro de ligne d'énoncé de corpus. Ce codage suit la proposition de Auer (1995) qui attribue des lettres aux langues en fonction de leur ordre d'apparition et s'appuie sur le principe de la séquentialité des échanges. A priori, le choix non marqué du locuteur consiste à suivre le choix de mélange de langues initié au tour précédent : on voit que c'est le cas ici, L1 s'exprime majoritairement en créole en 094 avec quelques insertions de français (mais, continuer), ce qui est codé A [C] et L2 s'aligne à ce choix en 095.

L'exemple 2 montre un cas classique où la notion de langue matrice proposée par Myers-Scotton (2002) pourrait facilement s'appliquer. Devant ces tours de parole majoritairement en créole, on n'a pas trop d'hésitation, en effet, même si les énoncés sont hétérogènes en partie, le créole est non seulement la langue majoritaire mais aussi celle qui a la contribution la plus importante sur le plan morphosyntaxique. Les items de

<sup>11</sup> Pour une présentation du schéma de document, voir Vaillant et Légise (2014).

<sup>12</sup> (<http://www.tei-c.org>)

<sup>13</sup> Elles s'appuient sur les gloses de Leipzig (<https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf> dernière consultation 24.03.2016) et sur quelques adaptations.

<sup>14</sup> La totalité des abréviations utilisées est fournie en annexe.

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

français sont très limités . Il s'agit, en l'occurrence, d'une conjonction de coordination qui permet d'articuler le discours – un cas relativement classique de discours bilingues – et trois verbes non conjugués. D'autres cas, en revanche sont plus complexes, comme l'exemple 3.

Exemple 3 : Visualisation d'une prise de parole multilingue

<b>ma</b>	<b>a-bo</b>	<b>maam</b>	<b>Gisepa</b>	<b>i</b>	<b>moré</b>	<b>na</b>	<b>cent deux</b>	<b>ans</b>	<b>non ///</b>
mais.CO	EMP.TOP-2SG	grand.parent	Gisepa	3SG.SBJ	AS.PRF.mourir	dans.PREP.LOC	cent.deux	an	non.EN.ASRT
CONJ	PRN	N	N	PRN	V	ADP	DET	N	PRT

*Mais toi, grand-mère Gisepa n'est-elle pas morte à cent deux ans non ?*

On voit, dans l'exemple 3, plusieurs langues apparaître dans la prise de parole d'un même locuteur puisqu'elle commence par une conjonction et un pronom en créole de Casamance, se poursuit par un terme de parenté wolof, puis par un pronom, un verbe et une préposition en créole de Casamance et se termine par un groupe nominal en français suivi par la particule interrogative 'non'. Plutôt que d'attribuer de manière arbitraire et problématique une langue matrice à un tel énoncé, nous considérons ce tour comme « multilingue ». Afin de visualiser ce multilinguisme au sein même de l'énoncé, un surlignage jaune apparaît sur l'écran de visualisation des exemples mais il n'est pas visible ici du fait de l'impression en noir et blanc. Nous l'avons rendu ici par un encadré pointillé. Par ailleurs, à l'intérieur de ce tour de parole multilingue, nous pouvons dissocier des « segments » associables à telle ou telle langue : l'exemple 4 montre ainsi les balises<sup>15</sup> xml de la prise de parole qui est visualisée en 3 et l'annotation d'un attribut « langue » à ces segments de corpus.

Exemple 4 : Balisage xml d'une prise de parole multilingue

<sup>15</sup> Dans la syntaxe xml, n'importe quel élément est précédé et suivi pas une balise (de début et de fin). Ainsi, tout ce qui concerne la prise de parole e95 commence après la balise « Prise de parole 'e95 < » et se termine avant la balise « > Prise de parole 'e95 ». Les attributs de cette prise de parole sont le locuteur concerné, le texte de cette prise de parole (et les éventuelles langues concernées), sa traduction juxtalinéaire, l'annotation des parties du discours correspondantes et enfin sa traduction libre. A l'intérieur du texte, nous avons également des indications sur les langues présentes et toutes sortes d'informations importantes (ici par exemple la notation d'une pause). Nous renvoyons à Vaillant et Léglise (2014) pour une présentation du schéma de document xml.

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

The screenshot displays a software interface for XML annotation of a multilingual speech sample. The interface is organized into several sections, each with a dropdown menu and a checkmark icon:

- Prise de parole 'e95'**: The main speech segment.
- Locuteur 'L2' Locuteur**: The speaker's name and language code.
- Texte de la prise de parole 'mul'**: The original multilingual text: "ma a-bo maam".
- Segment langue 'pov'**: Annotations for the Casamance Creole segments: "ma" and "a-bo".
- Segment langue 'wol'**: Annotations for the Wolof segment: "maam".
- Segment langue 'fra'**: Annotations for the French segment: "cent deux", "ans", "non".
- Pause 'longue'**: A long pause annotation.
- Traduction juxtalinéaire**: A line-by-line translation: "mais CO EMP.TOP-2SG grand.parent" and "Gisepa 3SG.SBJ AS.PRF.mourir dans PREP.LOC" and "cent.deux an non.EN.ASRT".
- Étiquettes partie du discours**: Part-of-speech tags: "CONJ", "PRN", "N", "PRN", "ADP", "DET", "N", "PRT".
- Traduction libre**: A free translation: "Mais toi, grand-mère Gisepa n'est-elle pas morte à cent deux ans non ?".
- Prise de parole 'e95'**: The end of the speech segment.

Nous voyons ici concrètement le balisage xml qui est proposé pour cette prise de parole multilingue : on y trouve des segments en créole de Casamance (identifiés par les balises « segment langue 'pov' »), en wolof (balises « segment langue 'wol' ») et en français (« segment langue 'fra' »).

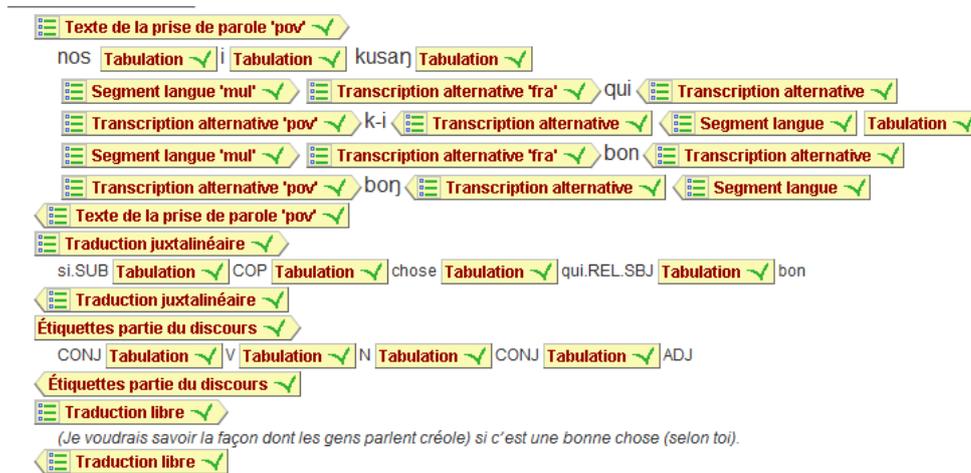
On peut lire cet extrait ainsi : au sein de la prise de parole e95 à laquelle on attribue un nom de locuteur « L2 », le texte de la prise de parole reçoit comme attribut « multilingue », et chaque segment de corpus reçoit un attribut linguistique : par exemple, *ma* et *abo* sont attribuables à du créole de Casamance ('pov' dans le code ISO<sup>16</sup>) alors que *maam* est attribuable à du wolof par exemple ('wol' code ISO). Une tabulation permet d'isoler chaque item. Par ailleurs, on peut décomposer les items comme *a-bo* afin de montrer la compositionnalité de l'élément emphatique de topicalisation et la deuxième personne du singulier, ce que la traduction juxtalinéaire notée en dessous permet justement de préciser (cf. ligne 2 dans l'exemple 3).

Par ailleurs, quand un même item existe dans au moins deux des langues impliquées dans notre corpus, nous le notons sous l'éditeur Jaxe au moyen de l'étiquette « multiple » en proposant une transcription alternative de chacune des langues à laquelle cet item pourrait appartenir (Léglise et Alby, 2013). L'exemple 5, avec sa visualisation en 6, en est une illustration.

Exemple 5 : Annotation et transcriptions alternatives de segments multilingues

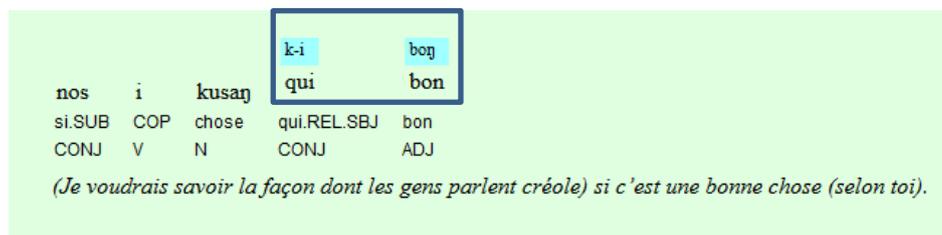
<sup>16</sup> code-langue ISO 639 (cf. [http://www.iso.org/iso/fr/home/standards/language\\_codes.htm](http://www.iso.org/iso/fr/home/standards/language_codes.htm).)

Nunez J. J. F. & Légise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.



Dans cet exemple, les items *ki/qui* et *boŋ/bon* pourraient appartenir au créole de Casamance et au français. Une transcription alternative indiquant chacune des langues auxquelles ces items en question sont susceptibles d'appartenir est proposée. C'est ainsi que sont annotées "fra" pour français et "pov"<sup>17</sup> pour créole de Casamance. La visualisation de cet extrait est présentée dans l'exemple 6.

Exemple 6 : Visualisation des transcriptions alternatives d'éléments multilingues



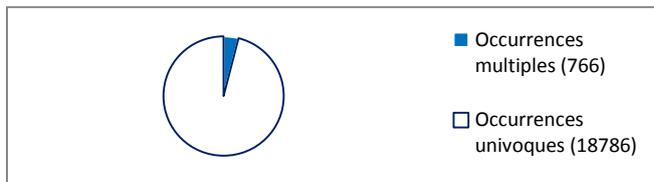
Ainsi, là où des chercheurs comme Cobbinah (2010) cherchent à savoir ce qui est emprunt ou pas dans le baïnounck Gubaher et la direction de cet emprunt<sup>18</sup>, les outils que nous avons développés nous permettent de proposer plusieurs interprétations possibles en fonction des langues auxquelles ces items pourraient appartenir. Ces items sont de fait « flottants » (Ledegen 2012) entre ces langues. Ainsi, près de 4% de l'ensemble du corpus est flottant (ou attribuable à plusieurs langues) comme on le voit sur le graphique 1 ci-dessous (766 occurrences 'multiples' sur les 19552 formes). Cela illustre par ailleurs des frontières floues entre les langues et l'utilisation préférentielle parfois, par les locuteurs plurilingues, de formes non différenciées (Légise, à paraître).

Graphique 1. Proportion d'occurrences « multiples » dans le corpus

<sup>17</sup> Le créole de Casamance et celui de Guinée-Bissau sont regroupés sous la même étiquette par Ethnologue, ce qui pose problème du fait que ces deux créoles sont différents. Nous utilisons ici cette étiquette pour renvoyer uniquement au créole de Casamance.

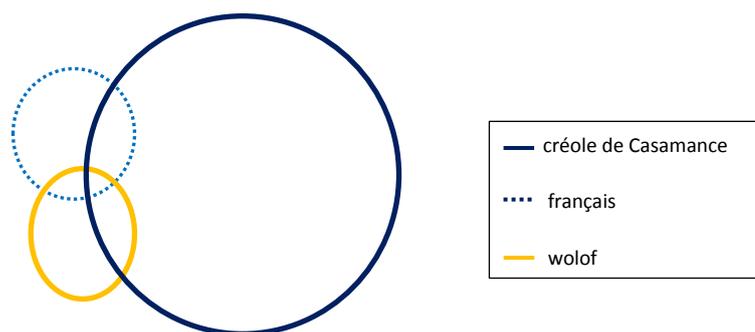
<sup>18</sup> C'est-à-dire d'où vient l'emprunt : dans le cas d'un terme comme *cabi* (*la clé*) du baïnounck Gubaher par exemple, qui vient du portugais (*chave*), la question est de savoir s'il a été emprunté directement au portugais ou au créole de Casamance (*cabi*) ou en passant par une autre langue africaine comme le wolof (*cabi*) (Cobbinah 2010, 185).

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.



Le graphique 2, ci-dessous, illustre par ailleurs les intersections réelles constatées dans le corpus. A côté des items univoques dans les trois langues majoritaires, créole de Casamance (cercle noir), français (cercle pointillé) et wolof (cercle gris), on observe quatre types d'intersections qui sont représentées sur le graphique 2 : 1,24% des cas sont des items pouvant être catégorisés comme du français ou du créole de Casamance (à l'intersection des cercles pointillé et noir), 1,16% du créole de Casamance ou du wolof, 0,81% du français ou du wolof. Enfin, on a également des items pouvant être catégorisés à la fois comme du français, du créole de Casamance ou du wolof, cela concerne 0,70% du corpus (à l'intersection des trois cercles). Bien que faibles, ces chiffres montrent tout de même que les items pouvant être catégorisés comme du français ou du créole de Casamance sont les plus nombreux, ce qui pourrait s'expliquer par la proximité lexicale entre ces deux langues ; leur lexique est en effet en grande partie dérivé du latin<sup>19</sup>.

**Graphique 2. Visualisation des intersections entre langues dans le corpus**



Les chiffres du sous-corpus sur lesquels ces statistiques ont été réalisées sont les suivants : items en créole de Casamançais 14061, items en français 3269, items en wolof 2833. Il y a trois types d'intersections dans le corpus : les items attribuables à la fois à du créole casamançais et du wolof (au nombre de 227), les items attribuables à la fois à du français et du wolof (au nombre de 159), et les items attribuables à la fois à du français et du créole casamançais (au nombre de 243, comme dans l'exemple 7). Enfin, 137 items ont été annotés comme « multiples » et attribuables aux trois langues, le français, le créole de Casamance et le wolof, comme dans l'exemple 8.

Exemple 7 : Visualisation d'items attribuables à la fois à du français et à du créole casamançais

<sup>19</sup> Une grande partie du lexique du Créole de Casamance vient en effet du portugais.

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

là	qui		
la	k-i	bay	baŋ
là.D2	REL.OBJ-3SG.SBJ	AS.PRF.aller	PST
ADV	PRN	V	PRT

*C'est là-bas qu'il était parti*

Exemple 8 : Visualisation d'items attribuables à la fois à du français, du créole de Casamance et à du wolof

			paa
			pá(papa)
			pā(pai, papé)
e	falá	unsoŋ	/
3PL.SBJ	AS.PRF.dire	un.CARD	père
PRN	V	DET	N

*Ils ont dit qu'un vieux (y est allé)*

Dans l'exemple 7, l'adverbe la/là et le relatif ki/qui sont attribuables à la fois à du français et à du créole casamançais, il serait artificiel de décider a priori, même si l'environnement créole pourrait nous influencer et nous amener à catégoriser tout cet énoncé comme « créole casamançais ». De la même manière, dans l'exemple 8, l'item pa est attribuable à la fois à du français (dans ce cas il s'agit d'une forme tronquée de papa), à du créole casamançais (dans ce cas il s'agit d'une forme tronquée de papé ou pai) et à du wolof (dans ce cas il s'agit de paa). La visualisation verticale des trois items – comme un paradigme de possibles – que nous renforçons ici par un encadrement pour mieux les repérer nous permet d'indiquer ces trois possibilités de lecture.

Par ailleurs, plutôt que de parler d'emprunt<sup>20</sup>, de calque, de transfert, de code-switching, de réorganisation syntaxique, de grammaticalisation etc., toutes sortes de phénomènes auxquels la linguistique de contact s'intéresse habituellement mais qui posent des problèmes de sur-interprétation dès lors qu'on veut les annoter pas à pas, nous avons choisi d'utiliser le terme de « phénomènes remarquables » pour qualifier ce qui nous semble digne d'un intérêt particulier :

Nous utilisons « remarquable » dans les deux sens de l'adjectif : soit les phénomènes observés sortent de l'ordinaire (de la langue ordinaire) – et nous partons d'un sentiment d'écart par rapport à la forme attendue ou de référence (Léglise, 2012) pour qualifier la forme observée de « remarquable », c'est-à-dire digne d'un intérêt particulier, soit les phénomènes observés nous paraissent exemplaires de phénomènes connus et bien décrits dans la littérature sur le contact de langues – et nous partons d'un sentiment de fréquence ou d'exemplarité (Léglise et Alby 2013, 106).

Ce type d'approche permet d'éviter

<sup>20</sup> Pour la notion d'emprunt notamment, comme évoqué plus haut, cela est dû aux raisons idéologiques en particulier sur le terrain casamançais.

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

tous les termes particulièrement foisonnants dans le domaine du contact de langues et très souvent contradictoires (d'un auteur à l'autre ou d'un cadre à l'autre). Il permet [également] de ne pas entrer dans des débats terminologiques comme ceux qui existent sur la distinction entre emprunt et code switching par exemple ou encore calque, interférence ou transfert, etc. (Léglise & Alby 2013, 106-107).

Les phénomènes intéressants observés dans le corpus sont ainsi regroupés en méta-catégories selon leur comportement ou leurs caractéristiques. Nous distinguons des phénomènes intéressants au niveau morphosyntaxique (Phénomènes Remarquables Morphosyntaxiques, PREMS), au niveau interactionnel (PRINT, Phénomènes Remarquables INTeractionnels), au niveau discursif (PREDISC, Phénomènes Remarquables DISCursifs) (cf. Léglise & Alby (2013) pour plus de détails). Les exemples suivants illustrent des phénomènes remarquables au niveau morphosyntaxique.

Exemple 9 : Visualisation d'un mixing au niveau morphosyntaxique impliquant trois langues (PREMS)

<u>parce que</u>	N	na	<u>misère-<b>loo</b>-i-l</u>	tok	i	moré
parce.que	1SG.SBJ	AS.NPRF	misère-CAUS-VT-3SG.OBJ	jusque	3SG.SBJ	AS.PRF.mourir
CONJ	PRN	PRT	V	CONJ	PRN	V

*Parce que je vais lui faire de la misère jusqu'à ce qu'il meure*

Dans l'exemple 9 où on note la présence d'éléments en créole de Casamance (caractère normal), français (souligné) et wolof (gras), un terme particulièrement remarquable nous semble être la suite : « misère-**loo**-i-l ». Nous avons ici la combinaison d'un nom français (misère), d'un morphème wolof (-loo) servant à exprimer le causatif et d'une voyelle thématique (-i) ainsi que d'un pronom personnel objet de 3e personne du singulier (-l) en créole de Casamance. Nous avons ici un cas de *mixing* des trois langues, particulièrement complexe, au niveau morphologique.

Exemple 10 : recatégorisation d'un élément multilingue entre français et wolof (PREMS)

L2 :

<u>mais</u>	<u>graw-ul</u>
mais	grave-NEG
CONJ	ADJ-PRT

*Mais ce n'est pas grave*

Dans l'exemple 10 où sont impliqués des éléments en français (souligné) et en wolof (gras), ce qui nous semble remarquable c'est le mot « graw-ul », qui apparaît comme composé d'un adjectif d'origine française (grave) et d'une marque de négation wolof (-ul). La combinaison de ces deux éléments montre que les frontières établies entre les langues sont parfois floues dans les pratiques langagières des locuteurs car on serait bien en peine d'identifier une langue unique pour cet élément. Cette combinaison entraîne par ailleurs une recatégorisation grammaticale. En effet, de l'adjectif français **grave**, on est arrivé à un verbe mixte français-wolof. Ce genre de combinaison est considéré par McLaughlin (2008, 728) comme étant un cas d'emprunt verbal. Ici encore, pour

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

les raisons épistémologiques évoquées plus haut, nous préférons éviter le terme d'emprunt<sup>21</sup> et insister sur la réalité plurilingue intrinsèque des items mêmes au sein des pratiques langagières hétérogènes.

#### 4. Conséquences pour la linguistique de contact et discussion

Le domaine de la linguistique de contact (Thomason 2001, Winford 2003), particulièrement fécond ces quinze dernières années, propose un certain nombre de concepts et cadres théoriques que ce type de pratiques langagières hétérogènes, avec le système d'annotation ainsi choisi, permet de discuter. Selon Myers-Scotton (1993 ; 1997 ; 2002) par exemple, dans un énoncé impliquant au moins deux langues les éléments grammaticaux sont produits dans une langue et les éléments lexicaux dans une autre langue. La langue qui produit les éléments grammaticaux est la langue matrice (*Matrix Language*) tandis que celle qui produit les éléments lexicaux est la langue insérée (*Embedded Language*). En proposant ce cadre théorique, Myers-Scotton a établi une distinction fonctionnelle des langues qui est sensée fonctionner pour tout discours bilingue ou pour toutes les pratiques langagières hétérogènes. Elle estime par ailleurs que les morphèmes grammaticaux tardifs externes (*outsider late system morphemes*) ne sont fournis que par la langue matrice. Enfin, ce *Matrix Language Frame (MLF) model* s'appuie sur un principe de structure universelle (*Universal Structure Principle (USP)*) qui attribue une langue matrice parmi les langues en contact (Myers-Scotton & Jake, à paraître).

Des développements récents ont permis aux auteurs d'affiner leur modèle, sur des bases psycholinguistiques, en dissociant : les morphèmes activés au niveau conceptuel (*conceptually-activated morphemes*) et les morphèmes dépendant au niveau structurel (*structurally-assigned morphemes*) qui auraient des niveaux d'accessibilité différents dans le lexique mental, ce qui entrainerait une partition entre morphèmes entiers (*content morphemes*) et morphèmes grammaticaux précoces (*early system morphemes*). Pour Myers-Scotton & Jake (à paraître) les morphèmes entiers sont constitués de noms et de verbes alors que les morphèmes grammaticaux précoces sont par exemple des marques de pluriel et n'apparaissent qu'avec un noyau comme les déterminants (dont les articles). Les morphèmes dépendants au niveau structurel<sup>22</sup>, eux, sont également subdivisés en deux catégories : les morphèmes grammaticaux tardifs transitoires (*bridge late system morphemes*) et les morphèmes grammaticaux tardifs externes (*outsider late system morphemes*). Les premiers types évoqués sont souvent constitués de prépositions tandis que les seconds types sont constitués de marques de l'accord entre le sujet et le prédicat d'une part et, d'autre part, les marques de TAM ou de cas pour les langues à cas.

Si jamais les éléments de la langue insérée sont longs, les auteurs les considèrent comme des îlots de langue insérée (*Embedded Language Island*) du moment où ils fonctionnent comme des syntagmes dont les règles répondent à la norme de la langue intégrée. Il y aurait dès lors une relation de dépendance entre ces îlots de langue intégrée et le reste de l'énoncé.

Or, le système d'annotation et les outils de recherche automatique que nous avons développés dans le cadre du projet ANR CLAPOTY et qui ont servi à annoter les données présentées dans cette étude nous permettent de montrer que, dans ce type de pratiques langagières hétérogènes, les éléments grammaticaux sont fournis par toutes les langues en présence (créole de Casamance, français et wolof). Il n'y a pas de distinction fonctionnelle des langues (éléments grammaticaux vs. éléments lexicaux). On a ainsi pu mettre en évidence une absence de frontières dans la structuration grammaticale des productions linguistiques en contexte

---

<sup>21</sup> Considérer « grawul » comme un emprunt au français signifierait par exemple l'annoter comme un item wolof (d'origine française). D'une part cela fait disparaître l'hétérogénéité intrinsèque de cet item en synchronie et d'autre part cela constitue un choix (identifier une langue unique pour cet élément) que nous ne souhaitons pas réaliser pour notre part. Nous préférons insister justement sur sa possible multiple appartenance.

<sup>22</sup> Ce sont des morphèmes les moins activés du point de vue conceptuel et qui sont dépendants au niveau structurel.

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

multilingue chez les créolophones casamançais. Les morphèmes grammaticaux des langues impliquées sont combinés avec des noms réalisés dans d'autres langues au sein du même syntagme nominal, ce qui va à l'encontre de la théorie de Myers-Scotton. Les trois exemples suivants en sont des illustrations.

Exemple 11 : combinaison d'un nom français avec un possessif créole

<b>A</b>	i	yel	ki	na	pagá
	COP	3SG.TON	que.REL.SBJ	AS.NPRF	payer
	V	PRN	CONJ	PRT	V
	<i>C'est lui qui va payer</i>				

<b>B</b>	pa	bo	bu	<u>billet</u>
	pour.PREP	2SG.TON	POSS.2SG	Billet
	ADP	PRN	DET	N
	<i>ton billet pour toi</i>			

Dans l'exemple 11, le déterminant possessif créole casamançais « bu », un élément grammatical, est combiné avec un nom « billet », réalisé en français. Dans ce cas, on pourrait considérer que « billet » est classiquement un « emprunt » (ou un switch) au sein d'un énoncé totalement en créole.

Exemple 12 : combinaison d'un nom français avec un déterminant wolof dans un environnement créole

i	misti <sup>23</sup>	<u>dévotion</u>	<b>b-i</b>
3SG.SBJ	AS.PRF.vouloir	dévotion	CLF.SG-ART.DEF.D1
PRN	V	N	DET
	<i>Il aime la dévotion (que tu voues envers lui).</i>		

Dans l'exemple 12, l'article défini wolof « bi », un élément grammatical, est combiné avec un nom réalisé en français « dévotion ». Ici, alors qu'on a un syntagme verbal en créole en début d'énoncé, le syntagme nominal est mixte, avec un élément lexical en français et un déterminant wolof. On voit bien qu'il est difficile d'isoler une langue matrice fournissant tous les éléments grammaticaux à l'énoncé. Faudrait-il ici considérer qu'on a un énoncé en créole avec un « îlot » inséré qui serait un mixte de wolof et de français ?<sup>24</sup> ou un îlot de wolof ayant préalablement emprunté au français ? ou deux « emprunts » (ou switch) successifs, l'un au wolof, l'autre au français ? Ces interprétations possibles nous semblent de bons exemples de sur-interprétation à partir des données – qui, elles, illustrent seulement l'hétérogénéité et la co-présence d'éléments susceptibles d'être attribués à trois langues au sein d'un même énoncé (un pronom et un verbe en créole, un nom en français et un déterminant en wolof).

Exemple 13 : combinaison d'un déterminant français avec un nom créole

<sup>23</sup> Cette forme est celle du créole de Guinée-Bissau et est réalisée par certains Casamançais. La forme du créole de Casamance est **mesté**.

<sup>24</sup> Toutefois, parler d'îlot dans ce cas est problématique car les îlots sont sensés être l'insertion d'éléments d'une seule langue au sein d'une langue matrice.

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

i	beŋ	<u>cou</u> -í-l	<u>même</u>	diya
3SG.SBJ	AS.PRF.venir	coup-VT-3SG.OBJ	même	jour
PRN	V	V	DET	N

*Il lui a (ensuite) fait un (sale) coup le même jour.*

Dans l'exemple 13, où trois items sont attribuables à du créole et deux sont attribuables (partiellement ou totalement) à du français, le groupe nominal est constitué d'un déterminant qualificatif français « même », donc un élément grammatical, et d'un nom réalisé en créole casamançais « diya ». Ici, on voit bien que le matériel grammatical n'est pas fourni exclusivement par le créole – qui pourrait être considéré comme la langue « matrice » - mais également par le français, « l'emprunt » (ou le switch) qui apparaît n'étant pas réalisé au niveau lexical mais pour un déterminant. Par ailleurs, le second verbe de cet énoncé présente une forme mixte intéressante, composée d'un nom français verbalisé et comprenant une marque de personne en créole précédée par une voyelle thématique.

La théorie de Myers-Scotton a été assez discutée, notamment par Macswan (2005a) pour qui, pour des raisons théoriques et empiriques, ce modèle doit être rejeté<sup>25</sup>. Toutefois, comme nous l'avons montré dans la partie précédente, elle continue d'influencer très largement les travaux en linguistique de contact et en linguistique de corpus – puisque les systèmes d'annotation de type TEI induisent une langue matrice pour les énoncés.

Auer (1999) a de son côté proposé une typologie très intéressante des discours bilingues, présentée sous forme de continuum allant du code-switching à la fusion de lectures en passant par le code-mixing. Toutefois elle n'est pas satisfaisante car Auer, comme beaucoup d'autres auteurs, semble ne prendre en compte que des cas où seulement deux langues sont en contact. Le corpus qui sert à cette étude résiste à une caractérisation simple par le moyen de cette typologie. Si on essaie toutefois de l'appliquer, on remarque que certaines parties relèvent du code-switching insertionnel. Ce dernier, selon Auer, correspond à l'insertion d'un élément lexical (*content word*) –souvent un nom, un verbe et plus rarement un adjectif ou un adverbe–dans un passage environnant dans l'autre langue (c'est le cas notamment de l'exemple 11 ci-dessus). Les indices prosodiques – tels que les insistances ou les pauses de transition<sup>26</sup> permettant de passer d'une langue à l'autre, donc de déclencher le *code-switching*– et les marqueurs discursifs (tels que les commentaires métalinguistiques et les hésitations) peuvent servir à souligner la juxtaposition et à la transformer en un phénomène localement significatif, c'est-à-dire en un phénomène qui a, aux yeux des locuteurs, une signification.

Mais la majeure partie du corpus relève plutôt, du « language mixing » c'est-à-dire de situations où la juxtaposition des langues doit avoir, selon Auer, du sens pour les participants à un niveau global, où le mélange est un type de communication récurrent, et où ce mélange est en fait le langage de l'interaction. Gafaranga & Torras (2002, 11-12) proposent, en s'appuyant en partie sur l'avis des bilingues sur leur façon de parler, deux grandes catégories à l'intérieur du phénomène de l'alternance de langue : « language alternation itself as the medium » et « language alternation as deviance ». Le premier correspond à la possibilité d'utiliser deux langues au sein de la même conversation. Les interactants ne font aucune différence entre l'utilisation de ces deux langues car ceci constitue une façon normale de parler. Le second, lui, constitue un écart par rapport au medium qu'utilisent les interactants. Le code-switching est ainsi défini, plutôt que toute occurrence de deux langues dans la même conversation, comme une façon de parler non orientée par les participants et ne

<sup>25</sup> Nous renvoyons également à Macswan (2005b) pour plus de détails.

<sup>26</sup> C'est-à-dire les pauses comprises à la jonction des deux langues insérées.

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

nécessitant pas de réparation. En d'autres termes, le code-switching est un cas de distinction interactionnelle qui se situe au niveau du choix langagier (Gafaranga et Torras 2002, 19-20).

Les pratiques langagières hétérogènes constituant ce corpus relèvent majoritairement du « language alternation itself as the medium » qui correspond au language mixing de Auer et qui peut dans certains cas être identifié comme une variété plurilingue même si, dans notre cas, les locuteurs n'ont pas forcément de termes pour nommer ce mixing. Le multilinguisme est bien vu par la population créolophone et par les Casamançais de manière générale, en particulier à Ziguinchor où les attitudes des locuteurs envers le multilinguisme et le plurilinguisme se traduisent par des pratiques plurilingues. Comme nous l'avons vu dans la première partie de ce texte, s'exprimer au moyen de pratiques hétérogènes constitue un moyen d'intégration sociale réussie, et, au-delà, une revendication identitaire, celle d'appartenir à une région multilingue par opposition au reste du territoire national. Les Casamançais se positionnent ainsi en tant que plurilingues devant leurs interlocuteurs issus d'autres localités.

Les locuteurs utilisent des ressources linguistiques qui sont à leur disposition sans se soucier des « normes » monolingues de chacune de ces langues telles qu'elles ont été établies par les linguistes descriptivistes notamment. Jørgensen (et al., 2011, et Varga 2011) suggèrent ainsi le terme de « polylinguaging » pour rendre compte de la façon fluide dont les locuteurs plurilingues parlent dans la vie quotidienne. Le polylinguaging est l'utilisation de traits attribuables à différentes langues même si les locuteurs ne connaissent que très peu de ces traits. Si les créolophones de Casamance ont tendance à bien connaître le créole et relativement le wolof et le français (en fonction de leur scolarisation et de leur degré d'exposition à ces deux dernières langues), on observe également des items attribuables à d'autres langues telles que l'anglais ou le portugais auxquelles ils ne sont pas spécialement exposés. Toutefois, le terme de « polylinguaging » semble surtout réservé à l'utilisation de termes issus de langues variées au sein d'un groupe – comme au sein d'un groupe adolescent multiculturel par exemple où il est « cool » de s'exprimer en faisant feu de toutes voix, de toutes ressources linguistiques.

## **Conclusion**

Le multilinguisme sénégalais a fait depuis longtemps l'objet de nombreuses descriptions sociolinguistiques qui ont montré comment ce multilinguisme – et le plurilinguisme de ses habitants – se jouait différemment en fonction d'oppositions telles que « tradition vs. modernité » ou « rural vs. urbain » à la fois dans la présence des langues sur ce territoire et dans les pratiques langagières des populations. En Casamance, les travaux de Juillard (1995) ont été pionniers dans l'étude des pratiques langagières plurilingues et ont permis de montrer l'intrication de dimensions globales et locales dans l'actualisation de ces pratiques. On sait désormais que le multilinguisme casamançais se traduit notamment par des pratiques langagières plurilingues, que nous nommons hétérogènes ou fluides, et différentes en outre d'un locuteur à un autre.

Toutefois, les mobilités constantes de population, la scolarisation et l'urbanisation massive dans les villes africaines entraînent des modifications des situations jusqu'alors étudiées et les pratiques langagières actuelles des populations locales, en constante évolution, restent encore à documenter et à analyser dans leur détail. La méthode d'annotation de ces pratiques, présentée ici, offre des outils intéressants pour révéler l'hétérogénéité des pratiques de ces populations vivant en situation de superdiversité. Ils permettent d'insister sur la fluidité de ces pratiques hétérogènes et sur le sens de ces mélanges, localement et globalement, posant les locuteurs casamançais en tant que « plurilingues ».

Appliquée à des corpus plurilingues sénégalais, ces outils permettent en outre de discuter des catégories circulant dans le domaine de la linguistique de contact qui ont surtout été établies pour des situations bilingues. L'hétérogénéité des pratiques langagières – au sens où elles exhibent des formes susceptibles

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

d'appartenir à plusieurs langues, styles et variétés, vient ici déranger ces catégories binaires (langue matrice vs. langue insérée, code-switching vs. code-mixing par exemple).

Par ailleurs, les corpus sur lesquels nous nous appuyons permettent de documenter des évolutions sociales et sociolinguistiques actuelles, telles qu'une modification de la dichotomie entre centre et périphérie décrite par Dreyfus & Juillard (2004) par exemple. En effet, des enregistrements dans les quartiers centraux de Ziguinchor, tout comme dans ceux qui sont à sa périphérie, révèlent le dynamisme et la complexité du plurilinguisme. Les populations passent spontanément d'une langue à une autre. La mobilité des populations entre ces deux pôles de la ville se traduit ainsi d'un côté par le renouvellement de l'environnement multilingue qui dépend de ces mobilités tant sur un plan global que local et d'un autre côté par la complexification des répertoires linguistiques des locuteurs installés au centre de la ville qui, en plus de leurs langues familiales s'approprient des véhiculaires urbains et des éléments des langues de leurs voisins. Il en découle des pratiques langagières hétérogènes, éminemment plurilingues en ce qu'elles mixent des éléments susceptibles d'être attribués à plusieurs langues : ici majoritairement le créole casamançais, le français et le wolof, mais parfois également, l'anglais, le baïnounck, le mancagne, le mandingue etc. À une échelle macro, Dreyfus & Juillard (2004b) présentaient également Dakar comme étant le centre et Ziguinchor la périphérie. Là aussi, la situation semble avoir évolué. En effet, une comparaison des données que nous avons recueillies à Dakar et à Ziguinchor, dans une perspective micro sociolinguistique, révèle le même plurilinguisme des populations concernées. Dans ce cas de figure aussi, la mobilité des populations à Dakar<sup>27</sup> semble jouer un rôle prépondérant dans les recompositions à l'œuvre.

## Références citées

- Auer, Peter. 1999. « From codeswitching via language mixing to fused lects : Toward a dynamic typology of bilingual speech ». *The International Journal of Bilingualism* 4 (3): 309-32.
- Amuzu, Evershed K. & Singler, John Victor. 2014. « Introduction of Codeswitching in West Africa ». *International Journal of Bilingualism* 18 (4): 329-45.
- Burt, Susan Meredith. 1994. "Code Choice in Intercultural Conversation: Speech Accommodation Theory and Pragmatics." *Pragmatics* 4 (4): 535–59.
- Cobbinah, Alexander. 2010. « The Casamance as an area of intense language contact. The case of Bainouk Gubaher ». In *Multilingualism and language contact in West Africa: towards a holistic perspective*, édité par Fiederike Lüpke & Mary Chambers. *Journal of language contact*.THEMA 3 (3): 175-201
- Du Bois, John W. 2002. "Stance and Consequence." *Annual Meeting of the American Anthropological Association, New Orleans*.
- Dreyfus, Martine & Juillard, Caroline. 2001. « Le jeu de l'alternance dans la vie quotidienne des jeunes scolarisés à Dakar et à Ziguinchor (Sénégal) ». *Cahiers d'Études africaines* 163-164: 667-96.
- Dreyfus, Martine & Juillard, Caroline. 2004. *Le plurilinguisme au Sénégal. Langues et identités en devenir*. Paris: Karthala.
- Englebretson, Robert. 2007. *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*. John Benjamins Publishing.
- Faye, Pépin. 2008. « Etude du discours mixte et du code-switching français-seereer-wolof : approches sociolinguistique et psycholinguistique ». Thèse de doctorat, Dakar: Université Cheikh Anta Diop.
- Gal, Susan, & Judith T. Irvine. 1995. "The Boundaries of Languages and Disciplines : How Ideologies Construct Difference." *Social Research* 82 (4): 967–1001.
- Gafaranga, Joseph & Torras, Maria-Carme. 2002. « Interactional otherness: Towards a redefinition of codeswitching ». *International Journal of Bilingualism* 6 (1): 1-22.

---

<sup>27</sup> Cf. également Lüpke (2013) sur ce sujet.

- Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.
- Giles, Howard, Nikolas Coupland, & Justine Coupland. 1991. "Accommodation Theory: Communication, Context, and Consequence." In *Contexts of Accommodation. Studies in Emotion and Social Interaction*. Cambridge University Press.
- Jørgensen, Jens Normann, Karrebaek, M. S., Madsen, L. M & Møller, J. S. 2011. « Polylinguaging in Superdiversity ». In *Language and Superdiversities*, édité par Gabriele Alex et Christiane Kofri, 13: 23-37. 2. Paris/Göttingen: UNESCO and MPIMMG.
- Jørgensen, Jens Normann & Varga, Somogy. 2011. « Norms and Practices of Polylingual Behaviour: A Sociolinguistic Model ». *Journal of Estonian and Finno-Ugric Linguistics* 2 (2): 49-68.
- Juillard, Caroline. 1991a. « Comportements et attitudes de la jeunesse face au multilinguisme en Casamance (Sénégal) ». *Cahiers des Sciences humaines* 27 (3-4): 433-56.
- . 1991b. « Le plurilinguisme au quotidien : Ziguinchor au Sénégal ». *Afrique contemporaine* 158: 31-52.
- . 1995. *Sociolinguistique urbaine : la vie des langues à Ziguinchor*. Paris: CNRS Editions.
- . 2005a. « Hétérogénéité des plurilinguismes en Afrique à partir du terrain sénégalais ». *La Linguistique* 41 (2): 23-36.
- . 2007. « Le plurilinguisme, objet de la sociolinguistique descriptive ». *Langage et Société* 3-4 (121-122): 235-45.
- Juillard, Caroline, Moreau, Marie-Louise & Thiam, Ndiassé. 1994. « Leur wolof dit-il qui ils sont ? La perception des appartenances régionales et ethniques au travers du wolof urbain parlé par les adolescents ». *Langage et Société* 68: 35-62.
- Kiessling, Roland & Mous, Maarten. 2004. « Urban Youth Languages in Africa ». *Anthropological Linguistics* 46 (3): 303-341.
- Ledegen, Gudrun. 2012. « Prédicats « flottants » entre le créole acrolectal et le français à la Réunion : exploration d'une zone ambiguë ». In *Changement linguistique et langues en contact. Approches plurielles du domaine prédictif*, édité par Claudine Chamoreau & Laurence Goury, 251-70. Paris: CNRS Editions.
- Léglise, Isabelle. 2012. « Variations autour du verbe et de ses pronoms objets en français parlé en Guyane : rôle du contact de langues et de la variation intrasystémique ». In *Changement linguistique et langues en contact*, édité par Claudine Chamoreau & Laurence Goury, 203-30. Paris: CNRS Editions.
- . 2013. « Multilinguisme, variation, contact. Des pratiques langagières sur le terrain à l'analyse de corpus hétérogènes. » Thèse d'habilitation, Paris: INALCO.
- . à paraître, « Pratiques langagières plurilingues et frontières de langues », in Auzanneau, M. & Greco L., *Dessiner les frontières*, Lyon : ENS Editions.
- Léglise, Isabelle & Alby, Sophie. 2013. « Les corpus plurilingues, entre linguistique de corpus et linguistique de contact : réflexions et méthodes issues du projet CLAPOTY ». *Faits de Langues* 41: 95-122.
- Lüpke, Friederike. À paraître. « Pure fiction - essentialist ideologies meet hybrid practices. A view from Agnack ». Édité par Mandana Seyfeddinipur. *Language documentation and conservation*
- Lüpke, Fiederike. 2010. « Multilingualism and language contact in West Africa: towards a holistic perspective » In *Multilingualism and language contact in West Africa: towards a holistic perspective*, édité par Fiederike Lüpke & Mary Chambers. *Journal of language contact*.THEMA 3 (3): 1-11.
- Lüpke, Fiederike. 2013. « Multilingualism on the ground » In *Repertoires and choices in African languages*, édité par Fiederike Lüpke & Anne Storch. Berlin/New York :Mouton de Guyter: 13-76 .
- Mackey, W. F. 1976. *Bilinguisme et contact des langues*. Paris: Klincksieck.
- Macswan, Jeff. 2005a. « Codeswitching and generative grammar : A critique of the MLF model and some remarks on "modified minimalism" ». *Bilingualism: Language and Cognition* 8 (1): 1-22.
- . 2005b. « Remarks on Jake, Myers-Scotton and Gross's reponse: There is no "Matrix Language" ». *Bilingualism: Language and Cognition* 8 (3): 277-84.
- Manessy, Gabriel. 1995. *Créoles, pidgins, variétés véhiculaires. Procès et genèse*. Paris: CNRS.
- Mark, Peter. 2002. *"Portuguese" style and Luso-African identity*. Bloomington and Indianapolis: Indiana University Press.
- McLaughlin, Fiona. 1997. « Noun classification in Wolof: When affixes are not renewed ». *Studies in African Linguistics* 26 (1):1-28.
- .2001. « Dakar Wolof and the Configuration of an Urban Identity ». *Journal of African Cultural Studies* 14 (2): 153-72.

Nunez J. J. F. & Légise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

———. 2008b. « On the origins of urban Wolof: Evidence from Louis Descemet's 1864 phrase book ». *Langage in Society* 37 (5): 713-735.

Moreau, Marie-Louise. 1994a. « Ombres et lumière d'une expansion linguistique. Les attitudes des Diola et des Peul d'Oussouye à l'égard du wolof ». *Langage et Société*, 68-88.

Myers-Scotton, Carol. 1993. *Social Motivations for Codeswitching: Evidence from Africa*. Oxford, UK: Clarendon Press.

———. 1997. *Duelling languages*. Clarendon Press. Oxford.

———. 2002. *Contact linguistics, bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.

Myers-Scotton, Carol & Jake, Janice L. À paraître. « Cross-language asymmetries in codeswitching patterns: Implications for bilingual language production ». In *The Cambridge Handbook of Bilingual Processing*, édité par John W. Schwieter. Cambridge, UK: Cambridge University Press.

Ndao, Papa Alioune. 1996. « Contact de langues au Sénégal. Etude du code switching wolof-français en milieu urbain : Approches linguistique, sociolinguistique et pragmatique. » Thèse de doctorat d'Etat, Dakar: UCAD.

Ndecky, Albinou. 2011. « Pratiques et représentations des parlers mancagnes de Goudomp ». Thèse de doctorat, Amiens: Université Jules Verne.

Nunez, Joseph Jean François. 2015. « L'alternance entre créole afro-portugais de Casamance, français et wolof au Sénégal : une contribution trilingue à l'étude du contact de langues. » Thèse de doctorat, Paris/Dakar: INALCO/UCAD.

Sankoff, David, Poplack, Shana & Vanniarajan, Swathi. 1991. « The empirical study of code-switching ». In *Papers for the Symposium on Code-switching in Bilingual Studies: Theory, Significance and Perspectives*, 181-206. Strasbourg: European Science Foundation.

Swigart, Leigh. 1992a. « Two codes or one ? The insiders' view and the description of codeswitching in Dakar » in *Codeswitching*, édité par Carol M. Eastman. *Journal of multilingual and multicultural development*, 13 (1 et 2) : 83-102. Clevedon. England/Philadelphia. Penn/Adelaide. Multilingual Matters.

———. 1992b. « Practice and perception: Language use and attitudes in Dakar. » Thèse de doctorat, University of Washington.

Thiam, Ndiassé. 1994. « La variation sociolinguistique du code mixte wolof-français à Dakar : une première approche. » *Langage et Société* 68: 11-33.

Thomason, Sarah G. 2001. *Language Contact: an introduction*. Edinburg: Edinburg University Press.

Vaillant, Pascal & Légise, Isabelle. 2014. « À la croisée des langues : annotation et fouille de corpus plurilingue Numéro spécial Fouille de données et humanités numériques ». *Revue des Nouvelles Technologies de l'Information* 2: 81-100.

Winford, Donald. 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell.

Zentella, Ana Celia. 1997. *Growing up bilingual: Puerto Rican children in New York*. Oxford: Blackwell Publishers.

## Sigles, abréviations et convention de transcription

### Abréviations grammaticales

1, 2, 3	Première, deuxième, troisième personne
ADJ	Adjectif
ADP	Adposition
ADV	Adverbe
ART	Article
AS	Aspect
ASRT	Assertif
CAUS	Causatif
CLF	Classificateur
CO	Coordinateur
CONJ	Conjonction
COP	Copule
D1, D2	[déictique] proximal, [déictique] distal

Nunez J. J. F. & Léglise I., 2017, Ce que les pratiques langagières plurilingues au Sénégal disent à la linguistique de contact, in M. Auzanneau, M. Bento & M. Leclère (eds), *Espaces, mobilités et éducation plurilingues : éclairages d'Afrique ou d'ailleurs*, Editions des archives contemporaines, 99-119.

DEF	Défini
DEM	Démonstratif
DET	Déterminant
EMP	Emphatique
HAB	Habituel
IMP	Impératif
LOC	Locatif
N	Nom
NEG	Négation
NPRF	Imperfectif, Inaccompli
OBJ	Objet
PL	Pluriel
POSS	Possessif
PREP	Préposition
PRF	Perfectif, Accompli
PRN	Pronom
PRT	Particule
PRT EN	Particule énonciative
PST	Passé
REL	Relatif
SBJ	Sujet
SG	Singulier
SUB	Subordination
TON	Tonique
TOP	Topique
V	Verbe
VT	voyelle thématique

### Autres

Fra	Attribuable au français
Mul	attribuable à plusieurs langues
Pov	Attribuable au créole de Casamance
PREMS	phénomènes remarquables morphosyntaxiques
PREDISC	phénomènes remarquables DISCursifs
PRINT	phénomènes remarquables INTeractionnels
Wol	Attribuable au wolof
XML	langage extensible de balisage ( <i>extensible markup language</i> )
ZIG	Ziguinchor

### Conventions de transcription

☒☒☒	passage inaudible
///	pause longue
//	pause moyenne
/	pause courte
-	frontière de morphème
.	combinaison d'éléments du métalangage