



**HAL**  
open science

# L'assignation des noms aux classes nominales dans une perspective comparatiste: Le cas des langues bak (Niger-Congo, Atlantique)

Guillaume Segerer

## ► To cite this version:

Guillaume Segerer. L'assignation des noms aux classes nominales dans une perspective comparatiste: Le cas des langues bak (Niger-Congo, Atlantique). *Faits de langues*, 2022, 52 (2), pp.203-222. 10.1163/19589514-05202004. halshs-03509669

**HAL Id: halshs-03509669**

**<https://shs.hal.science/halshs-03509669>**

Submitted on 30 Aug 2022

**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.

# **L'assignation des noms aux classes nominales dans une perspective comparatiste : Le cas des langues BAK (Niger-Congo, Atlantique)**

**Guillaume Segerer**

*LLACAN (CNRS-INALCO)*

## **Résumé**

Cet article présente les résultats d'une expérience visant à comparer de manière très détaillée les stratégies d'assignation des noms aux classes nominales dans les langues Bak, un petit groupe de langues d'Afrique occidentale. Pas moins de 127 sources lexicales ont été examinées, concernant 21 zones dialectales. Nous avons pu mesurer la stabilité et la régularité de l'assignation des noms aux classes, que ce soit entre les 4 ensembles composant le groupe Bak ou au sein de chacun de ces ensembles. La méthode utilisée peut facilement être appliquée à d'autres groupes de langues.

**Mots-clés : langues atlantiques, langues Bak, joola, manjaku, balant, bayot, bijogo, classes nominales, assignation des noms aux classes**

## **Abstract**

This paper shows the result of a quantitative experiment aiming at comparing in great detail the noun class assignment strategies within the Bak subgroup of the Atlantic languages, a group of languages spoken on the westernmost coast of Africa. No less than 127 lexical sources covering 21 dialect zones have been examined. We have been able to measure the stability and regularity of noun class assignment between the 4 clusters of the Bak languages as well as within each one of them. The method used may easily be applied to other language groups.

**Keywords : Atlantic languages, Bak languages, Joola, Manjaku, Balant, Bayot, Bijogo, noun classes, noun class assignment**

## 1. Introduction

Les descriptions de systèmes de classes nominales donnent généralement l'impression que chaque nom 'appartient' à une classe, et que toute modification de cette classe entraîne une modification sémantique ou grammaticale, spécialement mais pas uniquement dans les domaines du nombre et de la dimension. Dans la plupart des cas cependant, c'est la combinaison d'une marque de classe et d'une base lexicale qui produit un lexème nominal. Certaines bases lexicales se prêtent à davantage de combinaisons que d'autres. La notion d'assignation du nom à une classe vient du fait que les langues concernées sont décrites dans d'autres langues (surtout européennes) et peut être considérée comme un effet direct de la traduction. Ainsi par exemple, en joola keeraak, la base lexicale *saana* peut-elle signifier 'fromager' (un grand arbre) en combinaison avec la classe H, mais signifie 'pirogue' (traditionnellement taillée dans un tronc de fromager) avec la classe B. S'il n'y a donc pas de classe par défaut<sup>1</sup> pour la base *saana*, il y a en revanche une classe par défaut pour le sens 'pirogue' (B) et une autre pour le sens 'fromager' (H). Au pluriel ces classes sont respectivement U et K. On dira donc, en faisant un raccourci, que le nom 'pirogue' est assigné à la classe B. Le présent article se propose donc d'examiner les stratégies d'assignation des noms aux classes dans quelques langues d'Afrique occidentale, dans une perspective comparatiste. On cherchera en particulier à répondre aux questions suivantes :

- Quelles régularités y a-t-il dans l'assignation des noms aux classes ...  
... entre langues d'un même ensemble ? ... entre langues d'ensembles différents ?
- Y a-t-il des notions plus stables que d'autres du point de vue de l'assignation des noms aux classes ?
- Y a-t-il des domaines sémantiques plus stables que d'autres du point de vue de l'assignation des noms aux classes ?
- L'assignation des noms aux classes est-elle un marqueur phylogénétique important ?

Ces questions seront abordées d'un point de vue quantitatif : pour les langues concernées, toutes les sources lexicales disponibles sont mises à contribution, et pour chacune des sources 131 notions ont été examinées, et les préfixes de classes correspondants ont été relevés.

## 2. Les langues Bak

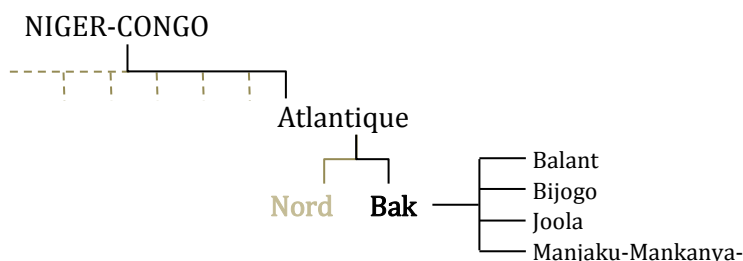
Ces langues sont parlées du sud de la Gambie au nord de la Guinée Bissau, le long de la côte atlantique de l'Afrique occidentale. Le groupe Bak, qui fait partie de la branche Atlantique du phylum Niger-Congo,

---

<sup>1</sup> L'expression 'par défaut' signifie ici 'singulier' pour les notions comptables, et 'non marqué pour la quantité ou la dimension' pour les autres.

comporte quatre ensembles relativement homogènes : Balant, Bijogo, Joola et Manjaku-Mankanya-Pepel.<sup>2</sup>

Ces langues sont caractérisées par des systèmes de classification nominale assez riches où l'accord en classe demeure très important. L'homogénéité à l'intérieur de chaque ensemble est importante également, ce qui signifie que l'inventaire des classes y est soit identique soit très proche. Cependant, il n'y a pas toujours intercompréhension entre les membres d'un ensemble.



### 3. Les sources

La source principale des données utilisées ici est la base de données en ligne RefLex, qui est quasiment exhaustive pour ce qui concerne les langues atlantiques en général et le groupe Bak en particulier. Quelques sources absentes de RefLex ont été utilisées lorsque j'y avait accès et qu'elles concernaient des parlers peu ou pas documentés dans RefLex<sup>3</sup>. Avant de détailler la méthode d'investigation et les résultats obtenus, une courte description des systèmes de chacun des quatre ensembles du groupe Bak est nécessaire, ainsi qu'une présentation des sources utilisées. Pour ce qui concerne les inventaires de classes, on a retenu à chaque fois le paradigme présentant la plus grande diversité, ce qui a conduit à privilégier pour trois ensembles (Joola, Manjaku-Mankanya-Pepel, Bijogo) les préfixes nominaux et dans le cas de l'ensemble Balant le paradigme des formes d'accord (Creissels et Biaye 2016). Les formes retenues sont notées en capitales, pour ne pas entrer dans le détail de la morphophonologie. Je rappelle que seules les classes de singulier ont été retenues pour cette étude. Les quelques exceptions (certains noms sont hors de la corrélation de nombre) sont signalées entre parenthèses. Les inventaires présentés ci-dessous ne sont pas extraits de descriptions grammaticales. Ils ont été obtenus à partir des données lexicales telles qu'elles figurent dans les sources originales. Il arrive assez fréquemment qu'il ne soit pas possible de déterminer la classe d'un nom d'après sa seule forme. Enfin, les

<sup>2</sup> J'utilise une majuscule initiale pour les noms des branches, groupes, ensembles, zones dialectales, mais les noms de langues sont entièrement en minuscules : ainsi, les parlers Bayot mais le bayot kugere.

<sup>3</sup> [www.reflex.cnrs.fr/Africa/](http://www.reflex.cnrs.fr/Africa/)

inventaires qui vont suivre ne tiennent pas compte des cas exceptionnels lorsque ceux-ci sont rares. Je m'explique : en bijogo kagbaaga par exemple 'nom' se dit *ŋa-kinno* (pl. *n-kinno*). Or le préfixe *ŋa-* est par ailleurs uniquement un préfixe de pluriel. Le cas du mot 'nom' est donc une exception. Mais comme cette exception se retrouve dans toutes les sources examinées, je retiens *ŊA* dans l'inventaire des marques de classes pour l'ensemble Bijogo. En revanche, le préfixe *i-* en manjaku est attesté sporadiquement dans quelques sources, mais pas suffisamment pour être retenu dans l'inventaire général de la zone Manjaku de l'ensemble Manjaku-Mankanya-Pepel. Cette question est discutée plus longuement ci-dessous pp. 6 et suivantes.

### *L'ensemble Joola*

C'est l'ensemble qui comprend le plus de variation dialectale, et également le plus de sources. Il est de tradition (Carlton & Rand 1993, Barry 1987, Doneux 1978) d'y inclure les parlers Bayot. On reviendra sur ce point dans l'analyse des résultats. Pour la présente étude, l'ensemble Joola est divisé entre les zones dialectales suivantes : Fogny, Banjal, Brin, Buluf, Kasa, Bliss, Keeraak, Ejamat, Karon, Kwaatay, Bayot. La nomenclature de ces zones dialectales n'est pas stable, et si je m'efforce d'utiliser les noms les plus courants, d'autres choix sont possibles (par exemple, le joola de Brin peut être appelé kujireray, et le joola banjal est souvent désigné par le terme eegima).

Le nombre important de sources est dû notamment au fait que l'ensemble Joola a suscité plusieurs études globales, avec à chaque fois des comparaisons lexicales, et donc des données lexicales (Kennedy 1961, Barry 1987, Carlton & Rand 1993). On a ainsi pas moins de 83 sources pour les 11 zones dialectales retenues. Evidemment, ces sources sont extrêmement variables, tant en quantité qu'en qualité. Les langues les mieux documentées à ce jour (au niveau lexical) sont le fogny, le kasa, le banjal, le gusilay, le keeraak et le kwaatay.

Les langues joola ont des systèmes de classes nominales très semblables sur le plan formel. Le tableau ci-dessous présente l'inventaire des préfixes rencontrés au cours de l'étude.

Fogny	A	B	BA		E	F	FA	J		(K)	KA	M	MA		(S)		U	
Banjal	A	B	BA		E	F	FA	J			KA	M			N	(S)	U	
Brin	A	B			E	F	FA	J			KA		MA		(S)		U	
Buluf	A	B	BA		E	F	FA	J	JA		KA	M			(S)		U	
Kasa	A	B	BA		E	H	HA	J		(K)	KA	M	MA		N	(S)	U	
Bliss	A	B	BA		E	H	HA	J			KA	M	MA		(S)		U	
Keeraak	A	B	BA		E	H	HA	J	JA		KA	M	MA		N	(S)	U	
Ejamat	A	B	BA		E	H	HA		JA	(K)	KA	M	MA		(S)		U	
Karon	A	B	PA		E	H	HA		CA		KA	M		N	(S)		U	
Kwaatay	A	B	BA	R	E	H	HA			(K)	KA	M	MA	N	N	(S)	TA	
Bayot	A	B	BA	D	E	F/H		J		(K)	KA	M			(T)		O	
<b>Joola</b>	<b>A</b>	<b>B</b>	<b>BA</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>FA</b>	<b>J</b>	<b>JA</b>	<b>(K)</b>	<b>KA</b>	<b>M</b>	<b>MA</b>	<b>N</b>	<b>N</b>	<b>(S)</b>	<b>TA</b>	<b>U</b>

Tableau 1 : Inventaire des préfixes de classes dans l'ensemble Joola

Les préfixes nominaux sont toujours de la forme V- ou CV-. Dans le tableau ci-dessus, les consonnes seules (B, F~H<sup>4</sup>, J, K, M, N, Ń, S) symbolisent des préfixes de forme CV- dont la voyelle est prévisible : -U pour les consonnes labiales et vélares (soit BU, FU~HU, KU, MU) et -I pour les autres (soit DI, JI, NI, ŃI, SI, TI). L'organisation du tableau ne présage pas des correspondances effectives entre les parlers. Ainsi, un nom en classe KA en Banjál ne sera pas obligatoirement en classe KA dans les autres zones dialectales. C'est précisément l'objet de la présente étude d'évaluer ces correspondances. Les classes notées entre parenthèses sont des classes de pluriel. Elles figurent dans le tableau du fait de la présence dans la liste lexicale des numéraux 'deux, 'trois' et 'quatre', qui en isolation sont cités avec ces classes.

### *L'ensemble Manjaku-Mankanya-Pepel*

Il s'agit d'un ensemble assez diversifié, mais contrairement à l'ensemble Joola, la documentation demeure assez pauvre. Comme son nom l'indique, cet ensemble est divisé en trois zones dialectales : Manjaku, Mankanya et Pepel. Pour le Manjaku, on dispose de sources importantes pour le parler babok qui est considéré comme la variété principale, c'est-à-dire à peu près comprise par tous les locuteurs manjaku (mais ce point est à vérifier). Pour les autres (Pecixe, Calequise, Pelundo, Costa de Baixo, Cur), on doit se contenter de listes lexicales assez limitées et parfois anciennes. Le dialecte cur est considéré comme divergent par tous les auteurs (Buis 1991, Doneux 1975). La zone Manjaku est représentée par 15 sources dans RefLex, auxquelles il faut ajouter une liste lexicale que j'ai recueillie à Paris en 2000 auprès d'un locuteur de Pelundo, non publiée. Les zones Mankanya et Pepel sont moins bien documentées, avec respectivement 8 et 6 sources, ce qui fait pour l'ensemble Manjaku un total de 30 sources.

L'inventaire des classes retenues pour la présente étude est ici représenté par les préfixes nominaux suivants :

Manjaku	B	BA	K	KA	M	NA	(G)		P	U
Mankanya	B	BA	K	KA	M	NA	(NG)	O	P	U
Pepel	B		K	KA	M	NA	(NG)	O	P	U
<b>M-M-P</b>	<b>B</b>	<b>BA</b>	<b>K</b>	<b>KA</b>	<b>M</b>	<b>NA</b>	<b>NG</b>	<b>O</b>	<b>P</b>	<b>U</b>

Tableau 2 : Inventaire des préfixes de classes dans l'ensemble Manjaku-Mankanya-Pepel

Ici, contrairement à ce qu'on a vu pour le Joola, les consonnes seules symbolisent soit des préfixes de la forme Cə- (pour les consonnes orales B, K et P), soit un m- syllabique (pour M). Les parenthèses

<sup>4</sup> La variation F~H dans les préfixes de classe reflète une correspondance phonétique régulière entre [f] et [h] entre, respectivement, les langues Joola 'orientales' (Fogny, Banjál, Brin, Buluf) et les langues Joola 'occidentales' (Kasa, Bliss, Keeraak, Ejamat, Karon, Kwaatay). La zone Bayot est un peu particulière de ce point de vue (voir ci-dessous p. 10)

indiquent la classe (de pluriel) utilisée pour les formes de citation isolées des numéraux ‘deux’, ‘trois’ et ‘quatre’.

### *L'ensemble Balant*

Cet ensemble est le moins bien documenté en ce qui concerne la dialectologie. Les deux grandes zones Ganja et Kentohe sont bien représentées au niveau lexical, avec pour la première les travaux pionniers de Geneviève Ndiaye-Corréard (1970) et le lexique important dû à Denis Creissels et Sékou Biaye (2016), et pour la seconde les travaux de Jean-Léonce Doneux (1984) et Chérif Mbodj (2011)<sup>5</sup>. Il existe un certain nombre d'appellations dont on ne sait pas toujours si elles désignent des dialectes différents ou si ce sont d'autres noms pour des parlars documentés par ailleurs : naga, mane, fora, fraas ~ fraasa ~ fraase<sup>6</sup>, soofa, pakra, hac. Naga et mane seraient proches du ganja, les autres seraient plutôt du kentohe (Mbodj 2011). Les parlars fraas et soofa sont en partie documentés, le fraas par Wilson (2007, 2008) et le soofa par Mbodj (2011). On dispose de 9 sources en tout pour le balant, qui sont rangées en trois zones (dont la validité n'est pas assurée) : Kentohe (dont soofa), Ganja, Autres. Ce dernier groupe contient les sources anciennes qui ne font pas mention de la variété dialectale : ‘Bulanda’ (Koelle 1854), ‘Balanta de Jabadá’<sup>7</sup> (Sousa Bella 1946).

L'inventaire des préfixes de classe du balant pose un problème spécifique. En effet, certains noms dépourvus de préfixe contrôlent des schèmes d'accords différents. Le choix de ne considérer que les préfixes nominaux trouve donc ici ses limites. Pour l'ensemble Balant, les symboles retenus tiennent donc compte des schèmes d'accords. Cependant, certaines sources sont très imprécises en ce qui concerne la morphologie, ce qui peut entraîner des erreurs dans l'assignation des classes aux noms.

Kentohe	B	F	HA	KƏ	(K)	U
Ganja	B	F	HA	GI	(G)	U
Autres	B	F	HA	KƏ	(K)	U
<b>Balant</b>	<b>B</b>	<b>F</b>	<b>HA</b>	<b>KƏ</b>	<b>(K)</b>	<b>U</b>

Tableau 3 : Inventaire des préfixes de classes dans l'ensemble Balant

Les formes sont presque identiques. On note seulement que G en ganja correspond à K dans les autres parlars Balant, ce qui semble une correspondance phonétique régulière : Kentohe *-bak* / Ganja *-bag* ‘porter sur l'épaule’ ; *K-boko* / *G-bɔgɔ* ‘mil’ ; *K-kək* / *G-gɛg* ‘chercher’.

<sup>5</sup> Il s'agit d'une thèse d'état, dont le volume II, consacré au lexique, nous est inaccessible. Les informations lexicales sont donc extraites du volume I, consacré à la description grammaticale. Ces données ne sont pas dans RefLex.

<sup>6</sup> Synonyme de balant d'après Mbodj (2011) et Wilson (2007), variété manodj (équivalent à mane pour Mbodj, donc relevant de la zone Ganja) d'après Creissels & Biaye (2015), appartenant au domaine Kentohe pour Biaye (2005).

<sup>7</sup> Jabadá est situé sur la rive sud du Rio Geba, on y parle donc en principe la variété ‘Fora’ (Wilson 1961, Mbodj 2011).

Comme en Joola et en Manjaku-Mankanya-Pepel, les classes dont la forme comporte une consonne initiale n'ont en général pas de voyelle spécifiée (B, F, K), sinon -A ou  $\Theta$  qui peut être considéré comme la variante +ATR de A (il y a une exception avec GI en ganja). En revanche, et contrairement à ce qu'on a vu pour les deux ensembles précédents, les classes représentées par une consonne seule sont réellement réalisées comme telles, sans insertion d'une voyelle d'appui.

### *L'ensemble Bijogo*

Les langues Bijogo sont parlées dans l'Archipel des Bijagos, au large de la côte de Guinée Bissau. Une quinzaine d'îles sont habitées, et chacune a son parler propre. On reconnaît pour l'instant quatre zones dialectales : Kagbaaga (îles de Bubaque, Rubane et Soga), Kanyaaki (île de Canhabaque), Kajoko (îles d'Orango, Orangozinho, Uno, Unhocomo, Uracane) et Kamona (îles de Caravela et Caraxe). Cette dernière est la plus divergente. J'ai personnellement enquêté sur le kagbaaga (Segerer 2002) et très brièvement sur le kamona. W.A.A. Wilson est l'auteur de listes lexicales sur le kajoko, et L. Scantamburlo a travaillé sur le kanyaaki, ces deux dernières sources n'ayant pas été publiées. On a un total de 7 sources pour l'ensemble Bijogo, dont seulement deux (kagbaaga et kanyaaki) sont numériquement conséquentes.

Les préfixes de classe des parlers Bijogo sont remarquablement stables et similaires. Dans le Tableau 4 ci-dessous, les formes entre parenthèses signalent des classes habituellement dédiées au pluriel qu'on rencontre exceptionnellement au singulier (ainsi en kagbaaga, *ŋa-kinno* 'nom' - pl. *n-kinno*).

Kagbaaga	E	(I)	KA	KO	M	MO	NE	(ŊA)	ŊO	O	U
Kajoko	E	(I)	KA	KO	M	MO	NE	(ŊA)	ŊO	O	U
Kamona	E		KA	KO	M	MO	NE	(ŊA)	ŊO	O	U
Kanyaki	E	(I)	KA	KO	Ŋ <sup>8</sup>	MO	NE		ŊO	O	U
<b>Bijogo</b>	<b>E</b>	<b>(I)</b>	<b>KA</b>	<b>KO</b>	<b>M</b>	<b>MO</b>	<b>NE</b>	<b>(ŊA)</b>	<b>ŊO</b>	<b>O</b>	<b>U</b>

Tableau 4 : Inventaire des préfixes de classes dans l'ensemble Bijogo

Ici, comme en Joola, les formes CV ont comme voyelle soit A, soit une voyelle prévisible en fonction de la consonne : O pour les labiales (M) et vélaires (K, Ŋ), E pour N (qui en réalité est une copie de la voyelle radicale, mais que la reconstruction interne peut restituer). On ne peut pourtant pas symboliser les classes CV par la consonne seule à cause de la classe M qui est distincte de la classe MO et se réalise comme une nasale syllabique.

### *Synthèse sur les inventaires de marques de classes*

<sup>8</sup> Le lexique kanyaaki de Scantamburlo (1999, non publié) utilise systématiquement un ŋ- initial là où toutes les autres sources ont une nasale syllabique homorganique que pour diverses raisons j'ai choisi de représenter par M. Je considère le ŋ- du kanyaaki comme un choix de transcription et non une réalité phonétique.



Les dernières lignes des tableaux qui précèdent peuvent maintenant être mises en regard et comparées :

Joola	A	B	BA	D	E	F	FA	J	JA	(K)	KA	M		MA	N		N		(S)	TA	(U)
M-M-P	NA	B	BA			P				K	KA	M					NG		O		U
Balant	HA	B				F				(K)	KƏ										U
Bijogo	O				E					KO	KA	M	MO		NE	NA	NO			(I)	U

Tableau 5 : Inventaire des préfixe de classes des langues Bak

Le tableau 5 fait apparaître les ressemblances formelles entre les préfixes de classes des langues Bak. La première colonne (A/NA/HA/O, grisée) est particulière, puisqu'elle réunit les classes qui regroupent les noms désignant des êtres humains. La correspondance formelle n'est pas manifeste, mais on prévoit que ces classes seront utilisées pour les mêmes items de la liste lexicale.

Parmi les autres classes, seules trois (K, KA~KƏ, U) sont présentes dans chacun des groupes (mais U en Joola est une classe de pluriel) ; trois autres (B, F~P, M) sont attestées dans trois des quatre groupes ; quatre (BA, E, N~NE, NG~NO) concernent 2 groupes en enfin 11 sont uniques. Il importe de souligner que ce résultat provient d'un examen superficiel des formes et non d'une application rigoureuse de la méthode comparative. Ainsi par exemple, il existe des arguments convaincants pour rapprocher la classe S (= si-) des parlers Joola et la classe I des parlers Bijogo. En outre, la classe U en Bijogo est sans doute le résultat de la fusion de plusieurs proto-classes correspondant aux classes Joola U, B et F.

#### 4. Méthode

Dans un premier temps, j'ai compilé une liste de 131 notions, avec la double contrainte d'inclure un maximum de domaines sémantiques et que chaque domaine soit si possible représenté par plusieurs notions (évidemment, une notion peut appartenir à plusieurs domaines). Il fallait aussi que les notions représentées soient présentes dans les sources. La liste est donnée en annexe. Pour chacune des 127 sources, j'ai cherché dans RefLex chacune des 131 notions, dont j'ai relevé le préfixe de classe. J'ai choisi le préfixe le moins marqué : singulier pour les notions dénombrables, ou forme non marquée pour la dimension ou la quantité pour les notions concernées (eau / un peu d'eau par exemple). Les sources étant très inégales en quantité (nombre de mots disponibles) comme en qualité (formulation des traductions), sur les presque 17000 formes possibles, plus de 10000 ont été effectivement collectées. Une des difficultés majeures est venue des notions pour lesquelles la distinction générique / spécifique est importante. Il est relativement aisé d'identifier les termes pour 'arbre', 'oiseau' ou 'poisson' (encore que certaines sources traduisent simplement par 'arbre' des dizaines de termes désignant des espèces particulières), mais l'exercice peut être plus compliqué pour des notions comme 'maladie', 'calebasse' ou 'fourmi'. C'est pour minimiser l'effet de telles

difficultés que j'ai choisi de considérer la totalité des données disponibles, et non quelques sources sélectionnées.

Pour chaque ensemble (Bijogo, Balant, Joola, Manjaku-Mankanya-Pepel), les sources ont été groupées par zone dialectale (voir ci-dessus). Pour chaque zone dialectale, et pour chacune des 131 notions, le préfixe le plus fréquent a été relevé. En cas d'égalité, plusieurs préfixes pouvaient être sélectionnés. Les préfixes divergents ont été notés comme exceptions. Enfin, pour chacun des quatre ensembles, les zones dialectales ont été comparées et le préfixe le plus fréquent sélectionné. Ici aussi, on a relevé les exceptions. Les notions pour lesquelles il n'a pas été trouvé d'exception ont été marquées comme stables.

Les tableaux ci-dessous illustrent la méthode avec quelques données de l'ensemble Manjaku :

Notion	CL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
'maladie'	<b>P</b>	p					m	p		p						p
'œuf'	<b>P</b>	p	p	p		m	p	p		p			b		p	p
'graisse'	<b>M</b>	m			m		m	m		m						m
'nez'	<b>B</b>	b	b	b	b	b	b	b	b	b	b	b	b	b	b	b
'feuille'	<b>KA</b>	ka	ka	ka		(i)	ka	ka		ka			ka		ka	ka
'cœur'	<b>P;U</b>	p	u	p	k		p	u	u	p	u	p	u	u	p	p

Tableau 6 : Codage des classes de quelques notions pour les 15 sources Manjaku<sup>9</sup>

Comme on peut le voir dans le tableau 6, certaines notions sont bien documentées et stables du point de vue de la classification nominale ('nez'), d'autres sont bien documentées et instables ('cœur'), d'autres sont mal documentées et stables ('graisse'), d'autres enfin sont mal documentées et instables ('maladie'). Les exceptions sont signalées par un fond gris. Dans la colonne CL figure la forme qui a été retenue pour cette notion dans cette zone dialectale (ici la zone Manjaku de l'ensemble Manjaku-Mankanya-Pepel). Le tableau suivant (Tableau 7) reprend, pour les mêmes notions, les classes sélectionnées pour chacune des zones dialectales de l'ensemble M-M-P, et attribue à chaque notion la classe la plus fréquente. Si toutes les zones ont la même classe pour une notion, l'indice de stabilité est 1, 0 dans le cas contraire. De même, pour chaque cellule du tableau, il est facile de noter si la classe inscrite correspond ou non à la classe retenue pour la zone dialectale. Par exemple, dans le Tableau 7, pour la notion 'maladie', la zone Pepel, avec la classe O, fait exception, de même que dans le Tableau 6 ci-dessus, pour la même notion 'maladie', c'était cette fois la source 6 (en l'occurrence Buis 1990) qui faisait exception avec m, alors que toutes les autres sources Manjaku ont p. Il est ainsi possible, pour chaque source d'une part, puis pour chaque zone dialectale d'autre part, de mesurer un indice de régularité, qui est égal au nombre de cas conforme à la majorité divisé par le nombre de

<sup>9</sup> Les nombres renvoient aux 15 sources utilisées pour le Manjaku, dans l'ordre de leur mention dans l'annexe 2 (p. 19).

notions renseignées. Ces chiffres seront discutés dans les chapitres suivants.

Notion	Manjaku	Mankanya	Pepel	CL	stabilité
'maladie'	P	P	O	<b>P</b>	<b>0</b>
'œuf'	P	P	P	<b>P</b>	<b>1</b>
'graisse'	M	M	M	<b>M</b>	<b>1</b>
'nez'	B	B	B	<b>B</b>	<b>1</b>
'feuille'	KA	KA	KA	<b>KA</b>	<b>1</b>
'cœur'	P ; U	P	O	<b>P ; O</b>	<b>0</b>

Tableau 7 : Codage des classes de quelques notions pour les 3 zones Manjaku-Mankanya-Pepel

Le même procédé est mis en œuvre pour chaque zone dialectale de chacun des quatre ensembles. Pour les notions présentées aux tableaux 6 et 7, le résultat est le suivant (Tableau 8) :

Notion	Joola	M-M-P	Balant	Bijogo	quadruplet
'maladie'	KA	P		ŃO	KA-P- -ŃO
'œuf'	F	P	F	NE	F-P-F-NE
'graisse'	M	M		KO	M-M- -KO
'nez'	E	B	B	ŃO	E-B-B-ŃO
'feuille'	KA	KA	KƏ	KO	KA-KA-KƏ-KO
'cœur'	E	P ; O	F	E	E-P/O-F-E

Tableau 8 : Classes préférées de quelques notions pour le groupe Bak

Contrairement aux étapes précédentes, cette fois on ne cherche pas à déterminer une 'classe préférée' pour chaque notion. On se contente d'enregistrer la série des quatre 'classes préférées' pour les quatre ensembles. Ces quadruplets pourront ensuite être listés et comparés. Nous avons maintenant toutes les données nécessaires à l'analyse. Mais avant de s'intéresser aux résultats globaux, il convient de se pencher à nouveau sur chacun des ensembles. Comme ils sont très différents pour ce qui concerne le nombre de sources, on doit s'attendre à des niveaux d'homogénéité variés. En effet, plus on a de sources pour un ensemble donné, plus on peut s'attendre à rencontrer des exceptions. Il est évident que les 11 zones dialectales et les 83 sources de l'ensemble Joola doivent être plus diversifiés que les 9 sources et les trois zones dialectales de l'ensemble Balant.

## 5. Stabilité et régularité

Pour chaque ensemble, les indices de stabilité et de régularité vont être présentés et discutés. Rappelons que l'indice de stabilité concerne un ensemble et est défini comme le rapport entre le nombre de cas stables (toutes les zones ont la même classe pour une notion donnée) et le nombre de notions renseignées. L'indice de régularité concerne une zone dialectale par rapport à son ensemble et est défini comme le rapport entre le nombre de cas conformes au cas général (la classe sélectionnée pour l'ensemble) et le nombre de notions renseignées.

Dans un sens, on peut dire que l'indice de régularité est relatif alors que l'indice de stabilité est absolu.

### ***L'ensemble Joola***

Cet ensemble est plutôt bien documenté, même si au sein des 83 sources examinées, seule une dizaine comptent plus de 1000 mots. L'ensemble affiche un indice de stabilité de 35%, ce qui est faible comparé aux autres ensembles (voir ci-dessous). Cette faible stabilité s'explique en partie par le grand nombre de classes différentes, ce qui augmente mécaniquement la possibilité de rencontrer des exceptions. Une autre explication tient aussi à l'inclusion, dans l'ensemble Joola, de la zone Bayot. Les parlers bayot sont en fait assez divergents par rapport au Joola, à la fois d'un point de vue phonologique et lexical. Leur inclusion dans l'ensemble Joola est due au fait que malgré les divergences, les ressemblances demeurent plus fortes entre un parler Bayot et un parler Joola qu'avec n'importe quel autre langage. Cependant j'ai présenté ailleurs des arguments pour considérer le Bayot comme un ensemble distinct (Segerer 2016). Et il s'avère que si l'on écarte la zone Bayot de l'ensemble Joola, l'indice de stabilité passe à 55%. Les indices de régularité sont très intéressants également. Les parlers Joola 'classiques' présentent des indices de régularité qui vont de 93% (zone Ejamat) à 100% (zone Buluf). Les parlers considérés comme 'atypiques' (zones Karon et Kwaatay) ont des indices de régularité plus faibles, mais encore de 84% et 82% respectivement. En revanche, la zone Bayot affiche un indice de régularité de seulement 61%. Mais voyons maintenant les autres ensembles.

### ***L'ensemble Manjaku-Mankanya-Pepel***

L'indice de stabilité global est de 70%, ce qui est beaucoup plus que l'ensemble Joola, même privé de la zone Bayot. Il faut dire que le nombre de sources (29) et le nombre de classes (10) sont aussi beaucoup plus faibles. Les indices de régularité des zones Manjaku, Mankanya et Pepel sont respectivement de 98%, 91% et 95%. Il s'agit donc d'un ensemble que l'on peut qualifier de très homogène.

### ***L'ensemble Balant***

Les langues Balant sont peu nombreuses, les sources également et seules trois zones dialectales ont été retenues, dont une n'est pas vraiment une zone. En outre, avec seulement 6 classes retenues, la probabilité d'exceptions est minimale. On doit donc s'attendre à un fort indice de stabilité. Celui-ci est de 71%, ce qui est certes le maximum, mais à peine supérieur à l'ensemble Manjaku-Mankanya-Pepel. Les indices de régularité sont de 98% pour le Kentohe, 93% pour le Ganja et 85% pour la troisième zone, dont la documentation est la moins fiable, avec un grand nombre de noms dont il est impossible de déterminer la classe.

### ***L'ensemble Bijogo***

L'indice de stabilité global est seulement de 53%, alors que le nombre de sources est très faible (6). Sans surprise, ce sont les parlars de la zone Kamona qui sont les plus divergents, avec un indice de régularité de 87%.

***Synthèse des mesures de stabilité et de régularité***

	Joola	J. sans Bayot	M-M-P	Balant	Bijogo
indice de stabilité	35%	55%	70%	71%	53%
nbre de zones	11	10	3	3	4
indice de régularité	61%-100%	82%-100%	91%- 98%	85%-98%	87%-97%
nbre de sources (total)	83	75	29	9	6
nbre de classes	18	18	10	6	11
zone la plus 'régulière'	Buluf	Buluf	Manjaku	Kentohe	Kajoko
sources	4	4	15	3	3

Tableau 9 : stabilité et régularité dans le groupe Bak

## 6. Résultats et discussion

D'une manière générale, il est raisonnable de penser que plus le nombre de zones dialectales est important, plus l'indice de stabilité devrait être faible. L'établissement de zones dialectales est généralement basé sur des calculs lexicostatistiques, et donc, par définition, des zones différentes présentent des différences lexicales notables. Toutefois, l'exercice proposé ici ne compare pas le lexique, mais les classes nominales. Des langues proches peuvent utiliser pour une même notion des bases lexicales différentes avec des classes identiques, ou l'inverse. Il y a ainsi en théorie quatre types de cas qui correspondent à différents scénarios d'évolution, puisque l'on est parti du principe que ces quatre ensembles linguistiques forment une unité généalogique :

1. Les bases lexicales sont apparentées et les marques de classe aussi : il peut s'agir d'un héritage commun ou d'un emprunt sans reclassement.

2. Les bases lexicales sont différentes mais les marques de classe sont identiques : il peut s'agir d'une innovation ou d'un emprunt avec maintien de la classe de départ.

3. Les bases lexicales sont apparentées et les marques de classe sont différentes : il s'agit donc d'un reclassement, qui peut avoir pour origine la disparition d'une classe, un changement de valeur sémantique associée à la classe, ou le remplacement d'un terme par un terme apparenté (le diminutif par exemple, qui est formé par substitution de classe).

4. Les bases lexicales sont différentes, les marques de classe également : il peut s'agir ici aussi d'une innovation, ou d'un emprunt avec reclassement.

Les quatre cas ci-dessus sont bien attestés dans les langues Bak. Les exemples suivants sont tous choisis dans l'ensemble Joola<sup>10</sup> :

Source	'eau'	préfixe	base
Barry 1987 : Kuricaakay (Ky)	<i>mu'mɛl</i>	M-	<i>-mɛl</i>
Koelle 1854 : I-A-2 - Fílham	<i>mómɛl</i>	M-	<i>-mɛl</i>
Sapir 1970 : joola fogny	<i>mumɛl</i>	M-	<i>-mɛl</i>
Carlton & Rand 1994 : JIG (fogny kalounaye)	<i>mmɛl(ám)</i>	M-	<i>-mɛl</i>
La Courbe in d'Avezac 1845 : Floupe	<i>moumɛl</i>	M-	<i>-mɛl</i>

Tableau 10 : bases lexicales identiques et classes identiques en Fogny (Joola)

Source	'sang'	préfixe	base
Barry 1987 : Kuricaakay	<i>ə'sim</i>	A-	<i>-sim</i>
Koelle 1854 : I-A-2 - Fílham	<i>fúsim</i>	F-	<i>-sim</i>
Sapir 1970 : joola fogny	<i>fasím</i>	FA-	<i>-sim</i>
Carlton & Rand 1994 : fogny kalounaye (JIG)	<i>kəsim</i>	KA-	<i>-sim</i>
La Courbe (?) in d'Avezac 1845 : Floupe	<i>moussin</i>	M-	<i>-sin</i>

Tableau 11 : bases lexicales identiques et classes différentes en Fogny (Joola)

Source	'feuille'	préfixe	base
Sapir 1970 : joola fogny	<i>katoj</i>	KA-	<i>-toj</i>
Carlton & Rand 1994 : bliss (JOG)	<i>kəsus</i>	KA-	<i>-sus</i>
Carlton & Rand 1994 : buluf (JIM)	<i>kāfāl</i>	KA-	<i>-fāl</i>
Carlton & Rand 1994 : buluf (MLN)	<i>kānɛ:n</i>	KA-	<i>-nɛɛn</i>
Barry 1987 : Karoon	<i>ka'hatal</i>	KA-	<i>-hatal</i>

Tableau 12 : bases lexicales différentes et classes identiques en Joola

Source	'dos'	préfixe	base
Sapir 1970 : joola fogny	<i>balámuk</i>	BA-	<i>-ləmuk</i>
Diatta 2006 : joola bliss	<i>bú-sól</i>	B-	<i>-sól</i>
Wintz 1909 : joola kasa	<i>əsuñ</i>	E-	<i>-suŋ</i>
Barry 1987 : Elana	<i>fú'mob</i>	F-	<i>-mɔb</i>

Tableau 13 : Bases lexicales différentes et classes différentes en Joola

Les tableaux ci-dessus donnent des exemples en Joola, mais on pourrait facilement en présenter d'autres pour les autres ensembles. En revanche, lorsqu'il s'agit de comparer les ensembles entre eux, les choses se compliquent. On a vu ci-dessus (Tableau 5 p. 8) que seules trois formes sont directement comparables dans les quatre ensembles. On doit donc chercher des 'correspondances régulières' de classe. Pour ce faire, on utilise les 'quadruplets' tels qu'ils ont été définis ci-dessus. Un quadruplet est composé des quatre marques de classes sélectionnées pour une notion, correspondant aux quatre ensembles du groupe Bak, toujours dans l'ordre Joola/M-M-P/Balant/Bijogo. Quelques précisions importantes sont nécessaires avant d'examiner les résultats :

1. dans chacun des ensembles, on a vu que certaines marques de classe peuvent être réduites à une consonne seule, qui sera réalisée de

<sup>10</sup> La transcription est celle de la source originale pour le terme d'origine mais elle est harmonisée pour les colonnes 'préfixe' et 'base'.

manière variable suivant les cas. Dans la plupart des cas, à côté des ces ‘classes consonantiques’, on trouve des classes de la forme CA. Sans vouloir trancher le débat ancien sur la réalité du ‘post-préfixe’ postulé par J. L. Doneux (1975), on observe dans certains cas une variabilité entre les formes à classe C et les formes à classe CA. Le tableau suivant illustre ce phénomène avec le cas du mot ‘faim’ en Joola :

Source	‘faim’	Classe
Carlton & Rand 1994 : bliss (JOG)	<i>buc<sup>h</sup>ar</i>	B
Carlton & Rand 1994 : kasa (ELK)	<i>buc<sup>h</sup>aar</i>	B
Sapir 1965 : gusilay	<i>bacar</i>	BA
Weiss 1939 : fogny	<i>bacar</i>	BA
Segerer 2020 : keeraak	<i>bacaar</i>	BA
Carlton & Rand 1994 : ejamat (YOU)	<i>buc<sup>h</sup>aar</i>	B

Tableau 14 : variation B/BA en Joola

Lorsqu’il n’est pas vraiment possible de choisir une forme en C ou la forme correspondante en CA, on notera ‘C(A)’ au lieu de ‘C;CA’ comme pour les autres cas indécis.

2. Lorsqu’un quadruplet comporte un cas indécis, on considère qu’il y a en fait deux quadruplets : par exemple, pour ‘genou’, le quadruplet ‘F;KA-P-F-KO’, qui comporte une indécision en Joola (F;KA) est scindé en deux quadruplets ‘F-P-F-KO’ et ‘KA-P-F-KO’. Ce cas concerne 34 des 131 notions de la liste. Lorsqu’un quadruplet comporte deux cas indécis, il en résulte quatre quadruplets. Par exemple, ‘herbe’ est représenté par le quadruplet ‘KA;MA-KA;U-KƏ-KA’ qui donne les quatre quadruplets suivants : ‘KA-KA-KƏ-KA’, ‘KA-U-KƏ-KA’, ‘MA-KA-KƏ-KA’ et ‘MA-U-KƏ-KA’. De tels cas sont rares (8 sur 131). Aucun quadruplet ne comporte plus de deux cas indécis. Par suite de la multiplication des quadruplets indécis, les 131 notions de la liste produisent 189 quadruplets.

3. Certains quadruplets sont incomplets, lorsqu’il n’a pas été possible d’attribuer une classe (voire deux) à une notion pour un ensemble donné, soit parce que les données étaient insuffisantes, soit parce que la variété ne permettait pas de choisir. Ces 12 ‘triplets’ ont peu de poids et demeurent valides pour la comparaison.

A l’issue de la procédure de sélection des classes représentatives de chaque ensemble pour chacune des notions, nous avons maintenant 189 quadruplets. Chaque ensemble ne contribue pas de la même façon à leur diversité. Pour l’ensemble Joola, 17 préfixes ont été retenus ; pour l’ensemble M-M-P, 12 ; pour le Balant, 8 et 11 pour le Bijogo. Le nombre total de quadruplets théoriquement possibles s’élève donc à  $17 \times 12 \times 8 \times 11 = 17952$ . En réalité cependant, on sait que tous les préfixes n’ont pas la même probabilité d’occurrence. Typiquement, les préfixes dédiés aux pluriels des noms sont par définition très rares dans le corpus. Et les préfixes de la forme C(A) ne sont pas des préfixes, mais signalent des hésitations entre C et CA. Si l’on tient compte de ces corrections, le nombre de quadruplets possibles est de 6864. Si les classes étaient distribuées au hasard, la probabilité de trouver plusieurs fois le même quadruplet dans les 189 occurrences serait

infime. Mais les classes ne sont pas distribuées au hasard, comme nous allons le voir maintenant.

### Les quadruplets fréquents

On a évoqué ci-dessus le cas de la 'classe des humains'. Dans la majorité des langues à classes du phylum Niger-Congo, on trouve une classe spécialisée pour les noms désignant des êtres humains, et en particuliers les notions générales telles que 'personne', 'femme', 'homme'. Outre ces 3 notions, notre liste comporte pour cette catégorie sémantique les notions 'enfant' et 'chef'. Quatre de ces cinq éléments sont affectés du même quadruplet A-NA-HA-O, le dernier, 'enfant', étant pourvu du quadruplet 'compatible' A-NA-U-O. La différence est causé par le Balant, qui a ici la classe U. Il faut préciser de nouveau que le Balant est un peu particulier, puisque certains noms présentent des schèmes d'accord mixtes. C'est notamment le cas pour le nom 'enfant', *mbùùtá* en ganja, dont une partie des accords est en classe U et l'autre en classe HA.

D'une manière générale, il ne faut pas s'attendre à trouver un grand nombre de quadruplets identiques, dans la mesure où le nombre des quadruplets 'disponibles' est très petit devant le nombre de quadruplets possibles. C'est pourquoi il convient de porter une attention particulière aux quadruplets 'compatibles'. L'expression désigne ici des quadruplets qui ne diffèrent que par un élément. Ainsi, pour le quadruplet A-NA-HA-O, il y a quatre quadruplets compatibles, que l'on obtient en remplaçant tour à tour chaque élément par un point d'interrogation qui symbolise n'importe quel classe : ?-NA-HA-O, A-?-HA-O, A-NA-?-O et A-NA-HA-?. Il est intéressant de voir que le seul quadruplet compatible attesté est celui du mot 'enfant'. La seule autre notion désignant un être humain est 'bébé', qui fait souvent appel à une classe de diminutif.

13 quadruplets concernent plus de 2 notions. Certaines notions peuvent être associés à plusieurs quadruplets en cas de division (voir ci-dessus), mais aussi parce que l'on relève aussi les quadruplets compatibles mais non rigoureusement identiques. Le Tableau 14 ci-dessous présente les quadruplets les plus fréquents avec les notions qui leurs sont associées. Comme on peut s'y attendre, les correspondances partielles sont plus fréquentes que les correspondances complètes. Sur les 189 quadruplets, 94 (soit 50%) sont 'réguliers', au sens où ils s'appliquent à plus de 2 notions. Si l'on considère plutôt les 131 notions de la liste de départ, 76 d'entre elles (soit 58%) sont associées à un quadruplet 'régulier'. Pour autant, les quadruplets 'réguliers' ne reflètent pas tous une réalité étymologique.

nb (+ comp.)	Quadruplet	Notions <sup>11</sup>
--------------	------------	-----------------------

<sup>11</sup> Les chiffres placés après les notions signalent les quadruplets obtenus après division, comme expliqué ci-dessus. Les parenthèse signalent les quadruplets compatibles. Une même notion peut être compatible avec plusieurs quadruplets :



11 (+21)	E-U-U-E	chèvre 2, chien, crocodile, fourmi, hippopotame, pou, poulet, puce, vache, viande 1, village 3 (animal, araignée 1, cadavre, chat 1, chat 2, cheval, chèvre 1, corps, couteau, éléphant, mouche, moustique 1, oiseau, ombre 4, pluie, poisson 1, serpent, soleil 1, viande 2, village 1, village 4, vent)
5 (+12)	KA-KA-K(A)-K(A)	peau, plume, aile, oreille, feuille, (cou 1, pied 2, lance 2, branche, nuage, main/bras, calébasse, poussière, cheveu 1, herbe 1, herbe 2, herbe 3)
2 (+13)	F-P-F-NE	œuf, sein, (pierre, banane, cuisse 1, cuisse 2, genou 1, jour 1, langue, nombril 2, œil, queue 2, sang, termitière 1, ventre 2)
1 (+10)	E-P-F-E	cœur 1, (cadavre, cœur 2, jambe, jour 1, pied 3, pierre, os 1, graine 1, un 1, village 1)
4 (+5)	B-B-B-U	baobab, chemin, dos 1, pirogue 2, (arbre, mortier, pirogue 1, dos 2, tête)
2 (+7)	F-U-U-K(A)	bouc 2, ver, (bouc 1, coq 2, éléphant, foie 2, guerre, termitière 2, viande 2)
2 (+5)	KA-P-F-K	corne, os 2, (dent, genou 2, jour 2, marmite 2, pied 4)
1 (+6)	E-U-F-K	étoile, (cendre 1, cendre 2, chapeau 1, graine 2, moustique 2, un 2)
0 (+7)	E-P-U-KA	(cadavre, fleur 2, maison, nombril 1, sable, village 1, village 2)
4 (+1)	A-NA-HA-O	chef, femme, homme, personne, (enfant)
3 (+1)	M-M-B-M	cerveau, huile de palme, lait, (eau)
1 (+3)	E-U-B-NE	année 1, (abeille, année 2, terre 2)
2 (+1)	M-P-U-M	miel, sel, (vin de palme)

Tableau 15 : Quadruplets 'fréquents'

Les quadruplets KA-KA-K(A)-K(A), B-B-B-U et M-M-B-M présentent à la fois des ressemblances formelles et une certaine unité sémantique. Le premier regroupe des parties du corps ('aile', 'main', cheveu', 'plume', etc.) et de végétaux ('branche', 'feuille', 'herbe') pour 14 des 17 notions associées. En ce qui concerne le quadruplet B-B-B-U, le U du Bijogo est en fait très probablement un ancien B : le mot kagbaaga *bu* 'tête' fait une partie de ses accords en *b*-. Ce quadruplet pourrait se réécrire B-B-B-\*B, retrouvant ainsi une unité formelle complète. Enfin, le quadruplet M-M-B-M regroupe quelques noms de liquides ou assimilés ('cerveau' est presque toujours dans la classe des liquides dans les langues de la région). Le balant étant l'une des rares langues Niger-Congo ne possédant pas de classe nasale, on ne s'étonne pas de voir ici un B. On se souvient (cf. Tableau 5 p. 8) que les seules classes attestées dans chacun des ensembles sont K, KA et U. On peut donc

---

'pierre : E-P-F-NE' est compatible avec les quadruplets F-P-F-NE et E-P-F-E. De tels cas ne sont toutefois pas très fréquents.

s'étonner de ne pas trouver de quadruplet U-U-U-U dans le tableau ci-dessus. En fait, la classe U du Joola étant une classe de pluriel, elle ne figure ici que parce qu'elle est utilisée pour certaines notions collectives (en l'occurrence 'bois à brûler'). Il est donc logique qu'elle ne corresponde pas aux classes U des autres ensembles<sup>12</sup>. Ainsi, les quadruplets 'étymologiques' que l'on observe sont en fait les seuls que l'on pouvait s'attendre à trouver.

Il reste quelques cas dont la fréquence demande des explications. Le quadruplet E-U-U-E rassemble 32 notions, dont 20 désignent des animaux. Sur les 131 notions de notre liste, 24 désignent des animaux (on inclut les notions 'animal', 'oiseau', 'poisson'). Dans l'ensemble Manjaku-Mankanya-Pepel, toutes ces notions sans exception sont en classe U (resp. O en Pepel). En Balant, les animés non humains ont des schèmes d'accords mixtes combinant les marques des classes HA et U. En Joola, la grande majorité des noms d'animaux est en classe E (pl. SI), comme en Bijogo (pl. en I, c'est une correspondance phonétique régulière : joola *s* / bijogo  $\emptyset$ ). Ainsi, comme les noms d'animaux présentent une grande régularité du point de vue des classes dans chacun des ensembles, il est naturel que cette 'correspondance' soit elle aussi régulière. Parmi les notions affectées de quadruplets compatibles avec E-U-U-E, 4 sont du type E-U-U-ŃO avec les valeurs suivantes : 'oiseau', 'animal', 'poisson' et 'serpent', c'est-à-dire des dénominations génériques. C'est une particularité du système de classe du bijogo de ranger tous les termes génériques dans la classe ŃO. On y trouve aussi les mots 'arbre' et 'chose' par exemple. La classe ŃO se trouve être aussi la classe unique des infinitifs.

Le quadruplet F-P-F-NE réunit 15 notions (dont 2 seulement ont le quadruplet complet), avec ici aussi une majorité de noms de parties du corps : 'œuf', 'sein', 'cuisse', 'genou', 'langue', 'nombril', 'œil', 'queue', 'sang', 'ventre'. On peut supposer qu'il y a ici une véritable correspondance entre Joola, Manjaku-Mankanya-Pepel et Balant. L'ensemble Bijogo a perdu la classe \*P/F et a reclassé une partie des noms dans la classe NE, qui est par ailleurs une classe de diminutifs. C'est le cas pour 'œuf', 'œil', 'langue', 'nombril', 'sang' et 'ventre'. On observe par ailleurs (mais ce n'est pas apparent dans notre liste lexicale) que les classes F du Joola, P du Manjaku-Mankanya-Pepel et F du Balant hébergent préférentiellement les noms de fruits, même si ce n'est pas systématique. En Bijogo (du moins en kagbaaga) les noms des fruits sont dans plusieurs classes, dont la classe NE est une des principales. Cette 'polysémie' entre les valeurs 'fruit de' et 'diminutif' peut sembler naturelle mais elle n'est attestée ni en Balant, ni en Joola, ni en M-M-P, et l'on peut ajouter que l'autre classe d'accueil des noms de fruits en bijogo est la classe KA, qui peut avoir par ailleurs une valeur d'augmentatif (qu'on trouve aussi pour la classe F en Joola).

La liste lexicale qui a servi de base à cette étude a été préparée de façon relativement arbitraire, en essayant de combiner du vocabulaire de base et des notions plus spécifiques à la région où les langues Bak

<sup>12</sup> On peut ajouter qu'en bijogo, U est une classe de singulier pour certains noms (avec pl. en ŃA), de pluriel pour d'autres (avec sing. en KA).

sont parlées. Il s'avère que sur les 131 notions retenues, 70 font partie de la liste Swadesh 200 (préférée à la liste de 100 mots puisqu'on ne compte que les noms). Si l'on prend comme base les 189 notions incluant des doublons (voir plus haut) ce nombre passe à 106. Dans les deux cas, c'est un peu plus de la moitié. Parmi ces 106 quadruplets, 63 peuvent être considérés comme 'réguliers'. Et dans les 83 quadruplets correspondant à des notions n'appartenant pas au lexique de base (tel que défini par la liste Swadesh 200), 54 peuvent être considérés comme 'réguliers'. La proportion de quadruplets réguliers dans le lexique de base est donc de  $63/106 = 59\%$ , et dans le lexique 'culturel' cette proportion est de  $54/83 = 65\%$ . Ainsi, on ne peut pas dire que le lexique de base soit plus stable du point de vue de l'assignation des noms aux classes nominales.

## 7. Conclusion

Les langues Bak ne représentent pas, on le savait, un ensemble très homogène. On note toutefois, dans le domaine de la classification nominale, une tendance générale, que chaque ensemble met en œuvre à sa manière, à simplifier les marques de classe en neutralisant leur élément vocalique. Cette tendance est à rapprocher des cas proches mais non semblables de langues de la branche Nord de la famille Atlantique : ainsi, le wolof a complètement perdu les voyelles des anciennes marques de classes (seules les consonnes participent encore au système de classification), alors que les langues du groupe Tenda ont tendance à perdre les consonnes. Les langues Bak maintiennent globalement des préfixes de forme CV, mais les langues Balant ont déjà amorcé une simplification massive puisque la plupart des préfixes sont non vocaliques, et que par ailleurs le nombre même des classes est nettement inférieur à ce qu'on trouve dans les trois autres ensembles du groupe.

D'autre part, l'assignation des noms aux classes dans les langues Bak est intéressante pour la comparaison. En effet, les quadruplets récurrents (ou 'réguliers'), c'est-à-dire les séquences de 4 marques de classes correspondant respectivement aux préfixes 'reconstruits' pour les ensembles Joola, Manjaku-Mankanya-Pepel, Balant et Bijogo ne sont pas répartis au hasard. Si les ressemblances formelles sont rares en raison de la diversité des systèmes (rappelons qu'on a retenu 18 classes pour l'ensemble Joola contre 6 pour l'ensemble Balant), on observe en revanche des régularités dans le domaine sémantique. Autrement dit, les langues étudiées ont tendance à classer les noms suivant des critères sémantiques comparables, mais avec du matériel différent. Les domaines sémantiques les plus stables sont 'classiques' dans le contexte Niger-Congo : humains, animaux, parties du corps, liquides/masses. De ce point de vue, on ne peut donc pas dire que la comparaison des stratégies d'assignation des noms aux classes exprime un signal phylogénétique fort.

On a pu proposer en passant quelques scénarios de réorganisation de certains systèmes de classification nominale. La méthode employée

ici peut suggérer quelques pistes pour comprendre comment chaque système suit sa propre voie. En effet, les contraintes internes se combinent avec des tendances de fond (notamment sémantiques) mais la connaissance de ces dernières seule ne peut pas permettre d'expliquer les mécanismes à l'œuvre.

## 8. Références

- Barry, Abdoulaye. 1987. *The Joola languages: subgrouping and reconstruction*. SOAS, University of London (PhD thesis).
- Buis, Pierre. 1990. *Essai sur la langue manjako de la zone de Bassarel*. Bissau : Instituto Nacional de Estudos e Pesquisa (INEP).
- Carlton, Elizabeth M. & Sharon R. Rand. 1993. *Enquête sociolinguistique sur les langues Diolas de basse Casamance*. Dakar : SIL.
- Creissels, Denis & Konstantin Pozdniakov (éds). 2015. *Les classes nominales dans les langues atlantiques*. Köln : Rüdiger Köppe Verlag.
- Creissels, Denis & Séckou Biaye (2016). *Le balant ganja : Phonologie, morphosyntaxe, liste lexicale, textes*. Dakar : IFAN.
- Dabo, Mamadou. 2013. *La morphosyntaxe de l'édammé (langue atlantique jóla du nord-ouest de la Guinée-Bissau)*. Dakar : Université Cheikh Anta Diop (thèse de doctorat).
- Doneux, Jean Léonce. 1975. Hypothèses pour la comparative des langues atlantiques. *Africana Linguistica* VI, Tervuren, pp. 41-129.
- Doneux, Jean Léonce. 1975. *Lexique manjaku*. Dakar : CLAD.
- Doneux, Jean Léonce. 1978. Les liens historiques entre les langues du Sénégal. *Réalités africaines et langues française* 7, pp. 6-55.
- Doneux, Jean Léonce, Albano Mendes & Armando Tchoba dos Santos Pereira. 1984. *Lexique balante-français*. Ms.
- Kennedy, Alistair M. 1961. *Dialectes de diola : listes de mots*. Ms.
- M'Bodj, Chérif. 2011. *Description synchronique du balante so:fa (Guinée-Bissau)*. Dakar : Université Cheikh Anta Diop.
- N'Diaye-Corréard, Geneviève. 1970. *Etudes fca ou balante (dialecte ganja): système des classes nominales, lexique fca-français, et textes*. Paris : SELAF.
- Segerer, Guillaume. 2016. Some Hypotheses About Possible Isolates within the Atlantic Branch of the Niger-Congo Phylum. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, vol. 36. Berkeley : Berkeley Linguistics Society. pp. 368-381.
- Segerer, Guillaume (2002). *La langue bijogo de Bubaque (Guinée Bissau)*. Louvain, Paris : Peeters.
- Segerer Guillaume & Sébastien Flavier, 2011-2020 *RefLex: Reference Lexicon of African Languages*, Version 2.0. Paris, Lyon. <http://www.reflex.cnrs.fr/Africa/>
- Segerer Guillaume & Jouni Maho, 2004-2020. *WebBall: Web Bibliography of African Languages and Linguistics*. Paris. <http://www.reflex.cnrs.fr/Lexiques/webball/index.html>
- Wilson, William André Auquier. 1961. Outline of the Balanta language. *African language studies* 2, pp. 139-168.
- Wilson, William André Auquier. 2007. *Guinea Languages of the Atlantic Group*. Frankfurt : Peter Lang.
- Wilson, William André Auquier. 2008. *Guthrie's wordlist for Atlantic languages, 1958-59*. Ms.

**Annexe 1 : liste des notions**

1. abcès	34. vache	67. cheval	100. corde
2. animal	35. crocodile	68. maison	101. sel
3. fourmi	36. jour	69. faim	102. sable
4. bras/main	37. maladie	70. fer	103. graine
5. cendre	38. chien	71. genou	104. ombrage
6. bébé	39. saison sèche	72. couteau	105. épaule
7. dos	40. poussière	73. feuille	106. peau
8. banane	41. oreille	74. jambe	107. ciel
9. baobab	42. terre	75. foie	108. fumée
10. abeille	43. œuf	76. pou	109. serpent
11. ventre	44. éléphant	77. homme	110. lance
12. oiseau	45. œil	78. viande	111. araignée
13. sang	46. graisse	79. lait	112. étoile
14. corps	47. plume	80. lune	113. pierre
15. os	48. feu	81. mortier	114. soleil
16. arc	49. bois à brûler	82. moustique	115. queue
17. cerveau	50. poisson	83. bouche	116. termitière
18. branche	51. cinq	84. nom	117. cuisse
19. sein	52. puce	85. nombril	118. trois
20. bouc	53. fleur	86. cou	119. langue
21. brousse	54. mouche	87. nuit	120. dent
22. fesse	55. pied	88. nez	121. arbre
23. calebasse	56. quatre	89. un	122. deux
24. pirogue	57. fruit	90. huile de palme	123. village
25. chat	58. chèvre	91. vin de palme	124. guerre
26. caméléon	59. herbe	92. chemin	125. eau
27. charbon	60. cheveu	93. personne	126. vent
28. poulet	61. chapeau	94. pluie	127. aile
29. chef	62. tête	95. saison des pluies	128. femme
30. enfant	63. cœur	96. anneau	129. ver
31. nuage	64. hippopotame	97. rivière	130. année
32. marmite	65. miel	98. coq	131. honte
33. cadavre	66. corne	99. racine	

**Annexe 2 : liste des sources utilisées**

Il serait trop long de citer intégralement l'ensemble des sources utilisées. Pour conserver l'ensemble de l'information, j'ai utilisé une forme de citation abrégée, suivie d'un nombre qui permet de voir la référence complète sur la base de données Webball<sup>13</sup>, en le copiant dans le champ de recherche intitulé 'Ref'. La mention (n.p.) signales les cas où j'ai utilisé mes notes de terrain personnelles. La plupart des

<sup>13</sup> [www.reflex.cnrs.fr/Lexiques/webball/index.html](http://www.reflex.cnrs.fr/Lexiques/webball/index.html).

sources sont également accessibles dans RefLex<sup>14</sup>. Enfin, le code Glottolog<sup>15</sup> est indiqué entre crochets.

## **Joola**

### ***Fogny [jola1263]***

Coly et al. 2008 : Diola Fogny (26601)

La Courbe (?) in d'Avezac 1845 : Floupe (25106)

Sapir 1970 : joola fogny (24177)

Weiss 1939 : joola fogny (20832)

Kennedy 1961 : kufoji (24642)

Koelle 1854 : I-A-2 - Fílham (24178)

Barry 1987 : Fogny (Fy) (1624)

Carlton & Rand 1994 : BAI (fogny) (22993)

Carlton & Rand 1994 : SUE (fogny) (22993)

Carlton & Rand 1994 : SND (fogny) (22993)

Carlton & Rand 1994 : JNK (fogny kombo) (22993)

Carlton & Rand 1994 : JNL (fogny kombo) (22993)

Carlton & Rand 1994 : BAR (fogny kombo) (22993)

Carlton & Rand 1994 : JIG (fogny kalounaye) (22993)

Barry 1987 : Kuricaakay (Ky) (1624)

Barry 1987 : Tenduck (TK) (1624)

Barry 1987 : Kanyobong (Kb) (1624)

Carlton & Rand 1994 : KOU (narang) (22993)

Barry 1987 : Jaŋ (Jŋ) (1624)

### ***Banjai [band1340]***

Bassène 2006 : joola banjai (24044)

Kennedy 1961 : bandial (24642)

Carlton & Rand 1994 : ENM (banjai) (22993)

Koelle 1854 : I-A-1 - Fúlup (24178)

Barry 1987 : Banjaal (BL) (1624)

Sapir 1965 : Joola Gusilay (24176)

Barry 1987 : Gusiilaay (Gy) (1624)

Carlton & Rand 1994 : TSL (gusilay) (22993)

GIE CLOA & SIL 2018 : gusiilaay (27549)

Barry 1987 : Afinyaam (Af) (1624)

Carlton & Rand 1994 : AFF (gusilay) (22993)

### ***Brin [elun1241]***

Barry 1987 : Kuluunay (KI) (1624)

Carlton & Rand 1994 : BRN (elun) (22993)

### ***Buluf [mlom1238]***

Barry 1987 : Elana (El) (1624)

Barry 1987 : Gulompaay (Gp) (1624)

Carlton & Rand 1994 : JIM (buluf) (22993)

Carlton & Rand 1994 : MLN (buluf) (22993)

### ***Kasa [jola1262]***

Wintz 1909 : joola kasa (24326)

Kennedy 1961 : casa (24642)

Barry 1987 : Huluf (HF) (1624)

Kennedy 1961 : huluf (24642)

Carlton & Rand 1994 : OUK (kasa huluf) (22993)

Carlton & Rand 1994 : EMY (kasa huluf) (22993)

Carlton & Rand 1994 : KAH (kasa huluf) (22993)

Wilson 2007 : husuy (24179)

Carlton & Rand 1994 : JAK (kasa selek) (22993)

Carlton & Rand 1994 : NMB (kasa ayun) (22993)

Carlton & Rand 1994 : ELK (kasa fluvial) (22993)

Barry 1987 : Esuulælu (Es) (1624)

Carlton & Rand 1994 : MLS (kasa esuulaalur) (22993)

<sup>14</sup> [www.reflex.cnrs.fr/Africa](http://www.reflex.cnrs.fr/Africa).

<sup>15</sup> <http://glottolog.org/>

Sambou 1979 : joola kasa  
esuulaaluur (24051)

***Bliss [blis1238]***

Barry 1987 : Blis (Bs) (1624)  
Carlton & Rand 1994 : NMN~NIO  
(bliss) (22993)  
Carlton & Rand 1994 : JOG (bliss)  
(22993)  
Diatta 2006 : joola bliss (25125)

***Keeraak [kera1248]***

Segerer 2020 : keeraak (n.p.)  
Kennedy 1961 : her (24642)  
Barry 1987 : Keeraku (Kr) (1624)  
Carlton & Rand 1994 : KAB (her)  
(22993)

***Ejamat [ejam1238]***

Carlton & Rand 1994 : KHM (ejamat)  
(22993)  
Carlton & Rand 1994 : YOU (ejamat)  
(22993)  
Carlton & Rand 1994 : EFK (ejamat)  
(22993)  
Barry 1987 : Efok (EF) (1624)  
Assine 2008 : kusanuay (26086)  
Wilson 2007 : felup (24179)

***Karon [karo1294]***

Wilson 2007 : karon (24179)  
Barry 1987 : Karoon (Kn) (1624)  
Carlton & Rand 1994 : HIL (karon)  
(22993)  
Sambou 2007 : Karon (24932)

***Kwaatay [kuwa1246]***

Payne 2000 : joola kwaatay (25092)  
Payne 1992 : joola kwaatay (23073)  
Kennedy 1961 : kwatay (24642)  
Barry 1987 : Kuwaatay (Kw) (1624)  
Carlton & Rand 1994 : JEM (kwaatay)  
(22993)  
Tastevin 1936 : joola dyamberin  
(19409)  
Wilson 2007 : diembering (24179)

***Bayot [bayo1262]***

Diagne 2009 : bayot kugere (24931)  
Barry 1987 : Bayoot (Bt) (1624)  
Tastevin 1936 : ba yot (19409)  
Carlton & Rand 1994 : KAI (bayot)  
(22993)  
Kennedy 1961 : bayot/Kailou  
(24642)  
Barry 1987 : Kuxinge (Kñ) (1624)  
Carlton & Rand 1994 : JAL (essin)  
(22993)  
Barry 1987 : Gubaare (Gb) (1624)

***Manjaku-Mankanya-Pepel***

---

***Manjaku [mand1419]***

Doneux 1975 : manjaku bok (5648)  
Wilson 2008 : Manjaku (Babok)  
(25289)  
Wilson 2007 : manjaco baboque  
(24179)  
Tastevin 1936 : ma n'jak (19409)  
Koelle 1854 : I-B-4 - Kányöp (24178)  
Buis 1990 : manjaku de Bassarel  
(3443)  
Segerer 2000 : manjaku Ulund (n.p.)  
Wilson 2008 : Manjaku (Calequisse)  
(25289)  
Carreira & Basso Marques 1947 :  
manjaku (Costa de Baixo) (3797)  
Wilson 2008 : Manjaku (Pecixe)  
(25289)  
Wilson 2007 : manjaco pecixe  
(24179)  
Wilson 2008 : Manjaku (Churo)  
(25289)  
Basso Marques 1947 : manjaku  
(24074)  
Wilson 2007 : manjaco churo  
(24179)  
Carreira & Basso Marques 1947 :  
manjaku (Churo)(3797)

***Mankanya [mank1251]***  
Gaved & Stammers 2004 : mancagne  
(25198)  
Koelle 1854 : I-B-2 - Sárâr (24178)  
Trifkovic 1969 : mancagne (19919)  
Tastevin 1936 : ma n'kap (19409)

Wilson 2007 : mancanha (24179)  
 Wilson 2008 : Mankanya (Bola)  
 (25289)  
 Koelle 1854 : I-B-1 - Bóla (24178)  
 Basso Marques 1947 : Brame  
 (24074)

***Pepel [pape1239]***

Ndao 2010 : pepel (25199)  
 Wilson 2007 : papel safim (24179)  
 Wilson 2007 : papel biombo (24179)  
 Wilson 2008 : Pepel (Biombo)  
 (25289)  
 Koelle 1854 : I-B-3 - Pèpèl (24178)  
 Basso Marques 1947 : Pepel (24074)

**Balant**

---

***Kentohe [bala1301]***

Doneux & al 1984 : balante kentohe  
 (24048)  
 Wilson 2007 : balanta K (24179)  
 Wilson 2008 : balanta (fraase)  
 (25289)  
 Mbodj 2011 : balante so:fa (27365)

***Ganja [bala1302]***

Ndiaye-Corréard 1970 : balante-fca  
 (ganja) (14769)

Creissels & Biaye 2015 : balante  
 ganja (26594)  
 Tastevin 1936 : ba lant (19409)

***Autres***

Koelle 1854 : XII-A-b-3 - Bulánda  
 (24178)  
 Sousa Bella 1946 : balante de Jabada  
 (18828)

**Bijogo**

---

***Kagbaaga [kagb1241]***

Segerer 2002 : bijogo-kagbaaga  
 (18390)

***Kajoko [kajo1240]***

Koelle 1854 : XII-A-a-2b - (Bidşógo)  
 Wün (24178)  
 Wilson 2007 : bijago (24179)  
 Wilson 2008 : Bijago (Orango)  
 (25289)

***Kamona [kamo1256]***

Segerer 1998 : bijogo-kamona (n.p.)  
 Koelle 1854 : XII-A-a-2a - (Bidşógo)  
 Ân·kâras (24178)

***Kanyaki [anha1238]***

Scantamburlo 1999 : bijogo de  
 Canhabaque (25064)