

Chapitre 2 : cooccurrences consonantiques

Ce chapitre organisé autour des cooccurrences consonantiques cherche à faire le point sur le type de contraintes susceptibles d'appartenir à la proto-langue. Les processus d'assimilation et de dissimilation à distance sont présents dans les langues bantoues et ils s'y expriment à travers plusieurs règles synchroniques : la règle de Dahl, la règle de Meinhof et la règle d'assimilation nasale progressive. J'introduirai ce chapitre en résumant brièvement les règles qui mettent en évidence l'interdépendance des consonnes, ensuite je décrirai les contraintes perceptibles à travers l'étude de chaque proto-lexique. Enfin, dans une dernière partie intitulée « synthèse des observations », je chercherai à replacer les contraintes dans un cadre plus large d'organisation phonologique du lexique et à établir un lien avec les processus d'assimilation et de dissimilation synchroniquement actifs.

1 - *Introduction*

Règle de Dahl :

« *Dans certaines langues, si deux consonnes sourdes se succèdent, la première se sonorise lorsqu'elle se trouve dans l'environnement requis, lequel varie selon les langues... La règle de Dahl semble bien être une innovation, son aire de dispersion est délimitée et il n'y a pas de traces connues d'un mécanisme semblable en dehors de cette aire* », (cf. Bastin 1983).

Ce processus de dissimilation, qui s'applique toujours de la droite vers la gauche, a donc pour effet de modifier le voisement des consonnes. Davy et Nurse (1982) ont montré la diversité de cette règle dans les langues bantoues de l'est : en synchronie, elle peut ne subsister qu'en termes de traces ou bien se montrer active et agir à l'intérieur du radical, affecter les prefixes et même les suffixes. Elle peut toucher toutes les consonnes du système phonologique de la langue ou bien, comme en gikuyu, ne toucher qu'une consonne, en l'occurrence la vélaire non voisée. En diachronie, ces auteurs relèvent l'évolution des occlusives non voisées

de la proto-langue vers un réflexe voisé dont le conditionnement dépend du non voisement de la consonne suivante. D'après leur analyse, la consonne vélaire *k serait particulièrement affectée par ce conditionnement.

Règle de Meinhof :

Meinhof (1932) a le premier mis en lumière la dépendance de C₁ par rapport à C₂ pour le ganda et les contraintes distributionnelles qui en découlent, il parlait alors de loi du ganda. En 1935, Doke appellera cette même loi, "loi de Meinhof" et enfin Meeussen (1962) proposera de parler de règle de Meinhof puisque le phénomène serait géographiquement limité. Cette règle est exprimée comme suit par Bastin (1983) :

« Lorsqu'une séquence consonantique « nasale + sonore » est suivie d'une nasale simple ou d'une autre séquence « nasale + sonore », il arrive que la première séquence soit réalisée par une nasale double homorganique de la seconde consonne ... Il semble que ce phénomène soit d'origine proto-bantoue : on en trouve des traces du nord au sud du domaine bantou sauf à l'ouest, où des règles opposées de dissimilation consonantique se sont développées. »

Dans les langues du groupe B10 (Mouguiama-Daouda 1995), on trouve des traces de la règle de Meinhof dans le réflexe nasal des occlusives voisées. Schadeberg (2003) indique une autre application de cette règle en kwanyama (R21) où le processus de dissimilation atteint la mi-nasale de deuxième syllabe, celle-ci perd sa partie nasale lorsque la consonne initiale du thème est une mi-nasale. La finalité de la règle est identique en empêchant la présence de mi-nasales dans des syllabes successives.

Règle d'assimilation nasale progressive :

Ce processus synchronique, décrit par Greenberg (1951) pour des langues du centre et de l'ouest du domaine bantou, entraîne le remplacement de la consonne /d/ des suffixes de l'applicatif (*₁d) et du réversif (*₂d) en /n/ quand la consonne finale du radical est une nasale. Il ajoute un versant diachronique à cette règle en remarquant, pour les langues appliquant l'harmonie nasale sur les suffixes, que « ... words that contained in Proto-bantu the succession of consonants m-l is replaced by m-n as a regular rule of phonetic development ». Bastin (1983) suggère que « ce type d'assimilation nasale ... pourrait dater du proto-bantou », alors que Greenberg émettait un avis opposé « the phenomenon of nasal harmony does not go back to the proto-bantu period ».

Deux traits sont en cause dans ces règles : le trait de nasalité et le trait de voisement. Le trait de nasalité définit des règles supposées appartenir à la proto-langue or, si tel est le cas, on devrait trouver dans les reconstructions une forte corrélation entre consonnes nasales dans les cooccurrences $*C_1/*C_2$ et $*C_2/*C_3$. Quant au trait de voisement, les règles l'impliquent dans des processus synchroniques, mais nous verrons qu'il a un rôle déterminant dans les contraintes phonotactiques de la proto-langue, du moins telle qu'elle a été reconstruite.

2 - Analyse quantitative

L'analyse quantitative va prendre en compte les cooccurrences existant entre :

- Les consonnes occupant les positions $*C_1$ et $*C_2$: paragraphes 2.1.1, 2.1.2 et 2.2.1
- Les consonnes occupant les positions $*C_2$ et $*C_3$: paragraphes 2.1.3 et 2.2.2
- Les consonnes occupant les positions $*C_1$ et $*C_3$: paragraphes 2.1.4 et 2.2.3

Je préviens le lecteur que, dans la discussion qui suit, je présenterai les tableaux des cooccurrences consonantiques sous une forme qui privilégie la fréquence de chaque phonème. Les phonèmes sont classés en allant du plus fréquent au moins fréquent, ainsi je m'attends à trouver les occurrences les plus élevées dans la partie supérieure gauche du tableau. Toute occurrence faible ou absente dans cette partie sera considérée comme suspecte et fera l'objet d'une recherche plus approfondie, il en sera de même pour une occurrence élevée en dehors de cette zone. En annexe, ces mêmes cooccurrences sont consultables dans des tableaux respectant l'ordre phonologique classique d'ordonnancement des phonèmes.

Des premières reconstructions jusqu'à B.L.R.3, nous verrons que certaines cooccurrences restent totalement absentes de la proto-langue malgré la fréquence de distribution des phonèmes concernés. Si ces contraintes touchent le mode articulatoire des consonnes ou leur activité laryngale, il apparaîtra qu'elles s'entendent avant tout à l'intérieur d'un même lieu d'articulation et qu'elles ne les concernent pas tous de manière identique.

2.1 – Racines verbales

2.1.1 - Cooccurrences $*C_1/C_2$ des radicaux *-CVC- et *-CVCVC-

Je vais présenter les cooccurrences consonantiques reconstruites par Meeussen et par Guthrie dans les suites *-CVC- et *-CVCVC-²⁷, regroupées sous un seul tableau du fait de leur similarité.

2.1.1.1 - Distribution générale

Les reconstructions les plus fréquentes concernent des cooccurrences qui mettent en scène les phonèmes fréquents du corpus, mais pas systématiquement les plus fréquents, par exemple les suites *-tVd- et *-dVd- ont un pourcentage inférieur à celui attendu qui indique sûrement une tendance à la dissimilation que l'on retrouve pour les labiales entre elles et pour les vélaires entre elles.

La quantité élevée de cooccurrences pour les suites *-mVn- est due, en partie, à la règle d'assimilation nasale progressive où la consonne $d_2 \rightarrow n$ quand $*C_1$ est une nasale, au point que les trois étymons de type *-mVd- sont proposés par Meeussen avec un $*C_2$ incertain = d ou n²⁸. C'est la forme *-mVd- qui est retenue dans B.L.R.3 comme étant la plus ancienne, les racines *-mVn- sont considérées comme des variantes.

2.1.1.2 - Contraintes

Dans B.L.R.1 et dans les séries comparatives, des contraintes identiques apparaissent. Elles sont perceptibles au premier regard par une absence de cooccurrences, absence suspecte au vu de la distribution quantitative de chacun des phonèmes qui les composent : *-dVn-, *-bVm-, *-pVb-, *-gVk- et *-tVc-.

• *-dVn-, *-bVm-, soit une absence de cooccurrences entre une occlusive voisée et la nasale appartenant au même lieu d'articulation. On verra par la suite que la non-reconstruction des suites *-dVn- est tout à fait confirmée dans B.L.R.3.

On peut mettre en parallèle à la rareté de ces cooccurrences le fait que les cooccurrences entre une occlusive voisée et la mi-nasale appartenant au même lieu d'articulation, soit *-dVnd- et *-bVmmb- sont, d'un point de vue quantitatif, bien représentées parmi les reconstructions.

²⁷ *-CVC- tableaux BLR :61, Cs 62. *CVCVC- tableaux BLR :63, Cs 64.

²⁸ Ces trois étymons sont : -màd- (d/n) « *finish* » ; -mɛd- (d/n) « *grow, sprout* » ; -mid- (d/n) « *swallow, devour* ».

- *-pVb-, *-gVk- : l'absence de cooccurrences entre les deux bilabiales et les deux vélaires, l'une non voisée l'autre voisée, est récurrente dans les lexiques et sera confirmée par B.L.R.3. car, pour les bilabiales par exemple on ne trouve que quelques séries régionales²⁹ et leur dispersion géographique se limite à l'est du domaine bantou.

En écho à ces absences, il me semble qu'il existe d'une part une certaine réticence à mettre en cooccurrence des occlusives appartenant au même lieu d'articulation. D'autre part, lorsque cela se produit les reconstructions ont favorisé les suites où les deux consonnes sont en harmonie de voisement. Cette remarque est particulièrement probante quand on compare le nombre de radicaux reconstruits avec deux occlusives identiques ($p_1/p_2 = 8$; $b_1/b_2 = 5$; $t_1/t_2 = 13$; $d_1/d_2 = 12$; $k_1/k_2 = 12$; $g_1/g_2 = 3$. Total=53) à celui où les occlusives s'opposent du point de vue du voisement ($p_1/b_2 = 0$; $b_1/p_2 = 1$; $t_1/d_2 = 9$; $d_1/t_2 = 4$; $k_1/g_2 = 2$; $g_1/k_2 = 0$. Total=16). Ainsi la non-reconstruction des suites *-pVb- et *-gVk- semble s'inscrire dans une règle plus générale que je développerai au paragraphe 3.2 de ce chapitre.

- *-tVc- : je retiens avec un peu de réserve cette cooccurrence, car le faible pourcentage de *c en position *C₂ peut relativiser le « trou » dans le tableau. Par ailleurs, cette cooccurrence s'avère rarissime dans B.L.R.3 où seulement deux verbes de type -CVCVC- la représentent mais ils restent à un niveau régional (*-t₂c₂d₂- « *jeter par terre* » Ce : LM, *-t₁c₁d₁- « *trembler* » NE : J) avec, pour le deuxième radical, une incertitude sur *C₁ et sur *C₂. Les contraintes entre ces deux occlusives non voisées alvéolaire et palatale seront à examiner de plus près.

²⁹ Il s'agit des étymons suivants : -pòb- « *be greedy* » JM; -pòb- « *be stupid* » GFM; -pùb- « *civer with cloth* » EJ ; -pàb- « *go astray, lose one's way* » J ; -péb- « *bleat* » J.

2.1.1.3 – Tableaux

B.L.R.1 : pourcentages *C ₁ / *C ₂ dans les verbes *-CVC- + *-CVCVC-																			
	*p	*t	*c	*k	*b	*d	*j	*g	*m	*n	*ŋ	*mb	*nd	*nj	*ŋg	*mp	*nt	*ŋk	
*p	1	1,4	0,5	2		2,4	0,1	0,1	0,5	0,8		0,8	1,3	0,1	0,9	0,1	0,13	0,13	
*t	0,6	1,6		1,6	0,6	1,1	0,6	0,6	1,6	0,5		1,6	1,8	0	1,4	0,1	0,63		
*c	0,3	0,1	0,5	1,9	1	1,4	0,3	0,3	0,6	0,9		0,8	0,5	0,8	1,4		0,25	10,8	
*k	0,6	1,8	0,8	1,5	0,5	3	0,3	0,3	1	0,8	0,5	1,9	0,5	0,1	1,6		0,38	0,13	
*b	0,1	0,6	0,1	0,5	0,6	2	0,6	0,4		0,4		1,3	0,8	0,1	1,1			8,66	
*d	0,9	0,5	0,3	0,9	1,1	1,5	0,3	0,9	2,1			1,9	1,3	0,1	2,3			13,9	
*j	0,9	1,3	0,4	1	0,8	1,5	0,5	0,9	0,9	1		0,6	0,1		1		0,25	11,2	
*g	0,1				0,8	2	0,4	0,4	0,4	0,8		1,3	0,6	0,5	0,8			7,9	
*m		0,4		0,1		0,4		0,1	0,4	1,3	0,4				0,3			3,26	
*n				0,1	0,1	0,1		0,1	0,1	0,6						0,25	1,51		
*ŋ						0,3			0,1						0,1			0,5	
	4,5	7,7	2,5	9,7	5,6	16	3	4	7,8	6,9	0,9	10	6,8	1,8	11	0,3	1,13	1	797

Tableau 65

Cs: occurrences *C ₁ /*C ₂ dans les verbes *-CVC-+*-CVCVC-																					
	d	ŋ g	mb	n d	m	k	n	t	b	g	p	c	ŋ k	n j	j	n t	j	y	ŋ	n c	
d	1,9	3,4	2	1,6	2,7	0,9		0,4	1,4	1,6	0,3	0,3								16,5	
k	3,1	1,5	1,5	1,2	1,5	1,3	0,6	1,1	1	0,3	0,6	0,6	0,1		0,2	0,2	0,11	0,11		15	
y	1,4	1,2	1,3	0,7	1,7	0,9	1,1	1,4	1,4	1	1,2	0,3	0,5	0,3		0,1	0,32	0,11		14,8	
t	1,2	1,6	1,8	1,8	0,6	0,9	0,6	1	0,6	0,7	0,7				0,1	0,4	0	0,11		12,2	
p	2,1	1,5	0,9	1,5	0,6	1,7	0,6	1,2			1,2	0,3								11,6	
b	2,8	1,1	1,1	1,3		0,6	0,5	0,9	0,3	0,5	0,1	0,1		0,2			0,11	0,11		9,68	
c	1,4	2,3	0,5	0,5	0,6	1,7	0,4	0,1	0,4	0,1	0,2	0,3		0,3					0,11	9,15	
g	1,9	0,1	0,5	0,4	0,1		1		0,5	0,4	0,1	0,2								5,32	
m	0,6	0,1			0,1	0,1	1,3	0,1		0,1		0,1			0,4					2,98	
n					0,2	0,1	0,9						0,2					0,11		1,49	
j	0,1	0,3			0,1				0,2											0,74	
ŋ		0,1			0,1															0,32	
ŋ		0,1			0,1															0,21	
	17	13	9,6	9	8,4	8,1	7,1	6,1	5,9	4,8	4,5	2,3	0,9	0,9	0,7	0,7	0,53	0,43	0,21	0,11	940

Tableau 66

2.1.2 - Cooccurrences $*C_1/C_2$ des radicaux $^-CVVC-$

2.1.2.1 - Dans B.L.R.1

Compte tenu du nombre élevé de $*d$, en position $*C_2$, je m'attends à ce que cette consonne participe aux cooccurrences maximales, or cela n'est vrai que partiellement puisqu'elle les partage avec $*k$. Ainsi les cooccurrences maximales sont atteintes par $*d$ en C_2 quand $*C_1$ est une consonne non voisée = $*t$, $*k$, et par $*k$ en C_2 quand $*C_1$ est une consonne voisée = $*b$, $*d$.

Par rapport aux $^-CVC(VC)-$ vus précédemment, les suites $^-bVVm-$ et $^-dVVn-$ sont possibles. Cependant leur analyse dans B.L.R.3 laisse un doute sur leur appartenance au proto-bantou.

$*-b \bar{u} \bar{u} m-$	<i>chase, hunt</i>	Variante zone A / Etymon PB $^*-g \bar{u} \bar{u} m-$
$*-d \bar{i} - a n-$	<i>dance, play</i>	Dérivé de $^*-d \bar{i} -$
$*-d \bar{u} \bar{u} n-$	<i>group, ascend</i>	Non étudiée dans B.L.R.3

Le cas des alvéolaires peut indiquer une différence dans les contraintes entre morphèmes. L'étymon pour « *dance* » laisse supposer que cette cooccurrence est possible entre un morphème lexical et un morphème grammatical alors que aucun radical n'est reconstruit avec cette distribution. D'autres dérivés existent avec le suffixe $^*-an-$.

Il n'apparaît pas dans les $^-CVVC-$, de distribution en lien avec le voisement de la consonne.

B.L.R.1 : pourcentages *C ₁ / *C ₂ dans les verbes *-CVVC-												
	*d	*k	*m	*t	*n	*p	*g	*ŋg	*b	*n j	*j	*ŋ
*p	3,67	1,83	3,67	0,92	1,83	2,75	2,75	0,92		1,83		0,92
*t	5,5	3,67	0,92	4,59	0,92		0,92	0,92	1,83			19,3
*b	4,59	5,5	0,92			1,83	0,92	0,92	0,92		0,92	16,5
*d	4,59	5,5	0,92	2,75	1,83	0,92						16,5
*k	5,5	2,75	0,92	0,92	0,92							11
*c	0,92	1,83	1,83									4,59
*n	0,92			0,92	2,75							4,59
*g	1,83		0,92									2,75
*m			0,92				0,92	0,92				2,75
*j	0,92											0,92
	28,4	21,1	11	10,1	8,26	5,5	5,5	3,67	2,75	1,83	0,92	0,92
												/ 109

Tableau 67

2.1.2.2 - Dans les Cs

Cs: pourcentages *C ₁ / *C ₂ dans les verbes *-CVVC-																	
	*d	*ŋg	*k	*t	*p	*n	*nd	*mb	*m	*b	*g	*c	*j	*ŋ	*n t	*ŋk	
*t	5,45	1,82	2,42	1,82	0,61	1,82	2,42	0,61	0,61	1,82	1,21		0,61			21,2	
*d	4,85	2,42	2,42	2,42	0,61	1,21	0,61	1,82	0,61							17	
*b	2,42	1,21	3,03	1,21	1,21		2,42	1,21		0,61		1,21				14,5	
*p	1,82	1,21		0,61	4,85	1,82	0,61		1,82		0,61					13,3	
*k	4,85	1,21	1,21	2,42	0,61	1,21		0,61				1,21				13,3	
*c	1,82	1,21	0,61	0,61				1,21	1,21	0,61		2,42	0,61			10,3	
*n	1,21	1,21	0,61	0,61		1,21								0,61	0,61	6,06	
*m		0,61									1,21					1,82	
*g									0,61	0,61						1,21	
*j										0,61						0,61	
*y											0,61					0,61	
	22,4	10,9	10,3	9,7	7,88	7,27	6,06	5,45	4,85	4,24	3,64	2,42	2,42	1,21	0,61	0,61	/165

Tableau 68

Les séries comparatives appellent un commentaire pour la cooccurrence *-pVVp- car la quantité élevée mérite vérification. Les huit Cs comptées dans ce contexte sont les suivantes :

Cs 1448	-p à a p-	carry (child) on back	Série partielle
Cs 1449	-p à a p-	bear child	Série partielle
Cs 1489	-p è è p-	blow (as wind)	Etymon PB
		fan	
		winnow	
Cs 1490	-p ì e p-	smoke (tobacco)	INC / Etymon PB : -p e e p-
Cs 1545	-p í v p-	become hot	Refusé
Cs 1560	-p ì i p-	become black	Dérivé
Cs 1605	-p ò v p-	blow as wind	Etymon PB
Cs 1623	-p ù u p-	blow as wind	Variante de la Cs 1605

Ce contexte nous met face à une distribution dont la quantité élevée n'est pas une caractéristique de la proto-langue mais le reflet de la méthodologie de Guthrie et en particulier de sa définition des séries comparatives. On perçoit bien la relation

sémantique qui existe entre les Cs 1448 et 1449, et entre les Cs 1489 et 1490, de même que les doublets vocaliques des Cs 1605 et 1623 et peut-être 1489. Dans le cadre d'un travail de reconstruction, les doublets vocaliques et sémantiques ne devraient pas figurer au même niveau et diminuer d'autant la quantité de cooccurrences.

Le pourcentage élevé de *d en position *C₂ amène cette consonne à participer à la plupart des cooccurrences maximales, les seules consonnes *C₁ qui ne partagent pas leur plus fort pourcentage avec *d sont les bilabiales *p et *b. *b a un peu plus de cooccurrences avec *k et *p avec *p

Le tableau des *-CVVC- se distingue de celui des *-CVC- à travers les pourcentages accordés aux suites *-tVVd- et *-dVVd- qui font ici partie des plus élevés alors que dans les *-CVC- ils étaient en dessous de ce que l'on pouvait attendre. Dans le même ordre d'idée, les suites *-cVVc-, que l'on ne trouve que dans les Cs, sont les seules occurrences possibles en position C₂ pour *c. Ceci pourrait être à priori suspect, mais en définitive elles appartiennent pour trois³⁰ d'entre elles à la proto-langue. Il semble qu'une corrélation se dessine entre consonnes identiques dans le schème *CVVC-.

2.1.3 PB X

2.1.3.1 - Radicaux *-CVC-

La suite la plus fréquente de ce tableau *-bVd- met en cooccurrence deux consonnes voisées. Les consonnes initiales qui ne partagent pas leur plus fort pourcentage avec *d sont les deux alvéolaires *t et *d et les deux palatales *y et *c. Dans les Cs les suites *-pVmb- atteignaient un pourcentage tout à fait adapté au nombre de chacun des phonèmes or, elles disparaissent totalement du PB X. Dans B.L.R.3, les suites *-pVmb- sont bien confirmées comme appartenant à la proto-langue à travers trois Cs de Guthrie et une nouvelle reconstruction de Tervuren. Une nouvelle fois, la démarche de Guthrie entre Cs et PB X s'avère trop sélective. Il y a dans ce tableau une quantité élevée de cooccurrences entre b₁ et mb₂, c'est même pour *mb le contexte le plus fréquent et le pourcentage est relativement important pour cette partie du tableau, elles sont confirmées en quantité élevée.

³⁰ Cs242, Cs 305, Cs 401.

2.1.3.2 - Radicaux *-CVVC-

Les verbes *-CVVC- ne présentent pas de distribution particulière par rapport aux Cs. Les cooccurrences les plus fréquentes se passent entre les phonèmes les plus fréquents bien qu'elles appartiennent au même lieu d'articulation. Je tiens à préciser que sur cinq racines de type *-dVVd-, quatre sont dans un contexte à voyelles isotimbres, c'est également le seul contexte où la cooccurrence maximale se produit entre deux consonnes identiques. Lorsque *p est en position *C₂ la suite vocalique précédente est, dans tous les cas, isotimbre.

2.1.3.3 - Radicaux *-CVCVC-

La cooccurrence la plus fréquente se présente sous la forme : *-gVdVC-, mais encore une fois, cette situation n'est qu'apparente car parmi les cinq reconstructions concernées, deux sont synonymes et appartiennent à la même racine (cf. Cs 758 et 759), ce qui est totalement confirmé dans B.L.R.2. De ce fait on peut dire que *-gVdVC- et *-pVkVC- obtiennent des pourcentages comparables.

En comparant ce tableau à celui des Cs je constate plusieurs différences : - La consonne *k prend en PB X le rang un. – Les mi-nasales (= 16,4%) sont moins nombreuses que dans les autres corpus pour la même suite.

2.1.2 - Cooccurrences *C₂/*C₃

La question principale est de savoir si les tendances et les contraintes perçues entre les syllabes du radical se retrouvent lorsqu'une frontière morphologique est supposée.

2.1.2.2 - Distribution générale

Tout d'abord, dans B.L.R.1 le pourcentage maximal est reconstruit pour les suites *-CVmbVd-, *-CVkVd- et non pour *-CVdVd-, consonnes qui présentent pourtant le plus grand nombre d'occurrences dans les deux positions *C₂ et *C₃. Il y a deux types de suites reconstruites avec *C₂=*C₃ : *-CVdVd- et *-CVkVk-, ce qui est du à la forte productivité des suffixes reconstruits avec ces deux consonnes. Pour les autres consonnes, cette situation ne se produit jamais.

Les observations faites pour les reconstructions de B.L.R.1 sont en grande partie valables pour les Cs : le pourcentage maximal est atteint par la suite *-CVdVk-, la suite *- CVdVd- permet un nombre correct de cooccurrences, mais en-deçà de ce que l'on était en droit d'attendre compte-tenu du total « ligne » et du total

« colonne ». La suite $^{*-}CVkVk-$ est extrêmement rare, de plus dans les Cs cette rareté touche la suite $^{*-}CV\mathfrak{ng}Vk-$. Finalement, les contraintes restent du côté des consonnes appartenant au même lieu d'articulation.

2.1.2.2 - Contraintes

Les cases noires qui font apparaître les trous dans la distribution mettent bien en évidence que le contexte touché est celui où les deux consonnes appartiennent au même lieu d'articulation. On retrouve des contraintes décrites dans les cooccurrences entre $^{*}C_1$ et $^{*}C_2$: d_2/n_3 , d_2/t_3 , b_2/m_3 , g_2/k_3 , auxquelles s'ajoutent : mb_2/m_3 , mb_2/b_3 .

Dans les cooccurrences entre $^{*}C_1$ et $^{*}C_2$ on ne trouvait jamais dans deux syllabes successives la consonne vélaire $^{*}g$ avec une consonne non voisée. Dans le tableau des séries comparatives, ces cooccurrences sont possibles. Leur analyse dans B.L.R.3 conduit à refuser les trois items $^{*-}CVgVt-$ au profit d'un étymon $^{*-}CVkVt-$, la raison invoquée étant toujours la règle de Dahl. Par contre l'étymon $^{*-}CVgVp-$ est accepté comme dérivé d'un nom : $^{*-}j\mathfrak{o}gvp-$ « fear Der. De $^{*-}j\mathfrak{o}g\mathfrak{a}$.

2.1.2.3 - Tableaux

	Cs: pourcentages *C ₂ / *C ₃ dans les verbes *-CVCVC-										
	*d	*k	*m	*t	*n	*b	*ŋg	*c	*g	*j	*p
*d	5,61	5,61	3,27	0,47		1,4	1,87		0,93		19,2
*k	6,07	0,47	2,34	4,21	0,47	0,93		0,93			15,4
*mb	6,07	1,4		1,4	0,47						9,35
*t	1,87	1,87	2,8			0,93		0,47			7,94
*m	4,67	1,87			0,93						7,48
*ŋg	4,21	0,47	0,47	0,47	0,93					0,47	7,48
*p	1,87	2,34	0,93	0,93	0				0,47		6,54
*g	4,21		1,4		0,93						6,54
*b	3,27	0,93		0,93							5,14
*n	1,87	2,8	0,47								5,14
*j	2,8	1,4									4,21
*nd	1,4	0,93									2,34
*c	0,47		0,47		0,47						1,4
*p	0,47	0,47									0,93
*nj		0,47	0,47						0,47		0,93
	44,9	21	12,6	8,41	4,21	3,27	1,87	1,4	1,4	0,47	0,47 /214

Tableau 69

Cs: pourcentages *C ₂ / *C ₃ dans les verbes *-CVCVC-										
	d	k	m	t	n	p	ŋg	b	g	c
d	5,5	9,1	3			0,3	1,2	0,3	0,3	
ŋg	7	0,9	1,2	0,3	1,8	0,3			0,3	12
k	5,2	0,3	0,9	2,7	0,6				0,3	10
m	4	2,7		1,2	0,6					8,5
mb	4,3	2,7		0,6	0,3					7,9
t	2,4	2,4	2,1		0,6			0,3		7,9
nd	3,6	2,7	0,6							7
b	3	2,4	0,3	0,6						6,4
n	1,5	3,6	0,9			0,3				6,4
p	2,1	1,8		0,6		0,3			0,6	5,5
g	2,4	0,3	0,9	0,9		0,3				4,9
c	0,9		0,3							1,2
ŋ	0,6	0,3								0,9
j	0,3						0,3			0,6
n j	0,3	0,3								0,6
y	0,3									0,3
ŋk	0,3									0,3
	44	30	10	7	4	1,5	1,2	0,9	0,9	0,6
										/329

Tableau 70

2.1.4 - Cooccurrences $*C_1/*C_3$

2.1.4.1 - Contraintes

La suite qui domine ici le tableau met en cooccurrence l'alvéolaire non voisée avec l'alvéolaire voisée soit la suite $*-tVCVd-$. Ce pourcentage paraît particulièrement élevé si on se place du point de vue du nombre total de $*t$ en position initiale ou si on le compare au pourcentage de la ligne suivante (2,8%). La configuration générale du tableau montre une répartition en lien avec la fréquence de chaque phonème sans pointer de corrélation particulière.

La suite $*-kVCVk-$ est paradoxalement moins fréquente que la suite $*-gVCVk-$. Est-ce qu'il y a à chercher un lien de cause à effet entre ces suites ? Les données de B.L.R.3 rétablissent une proportion cohérente par rapport au nombre total et le déséquilibre disparaît (réflexe indirecte ? données insuffisantes ?).

Dans les Cs, mises à part les suites à initiale $*b$ et $*m$, toutes les autres totalisent leur maximum d'occurrences avec la consonne $*d$ en position C_3 et le pourcentage maximal est obtenu avec les deux consonnes de rang 1 : $*y$ pour les C_1 et $*d$ pour les C_3 .

La suite $*-dVCVd-$ atteint un pourcentage particulièrement élevé qui distingue les Cs de B.L.R.1. Cette alvéolaire est la seule consonne à appartenir à des reconstructions où $*C_1=*C_2=*C_3$, les cas sont rares mais ils figurent dans B.L.R.3 dans des suites $*-dVVd-Vd-(V)$.

La suite $*-cVCVd-$ se trouve dans une situation similaire à celle qu'elle occupait dans B.L.R.1, avec un pourcentage à la fois supérieur à celui attendu et présentant un écart par rapport aux autres pourcentages de la ligne. Les Cs de forme $*-cVCVd-$ obtiennent un degré de fiabilité qui les situe à un niveau de reconstruction générale ou régionale, contrairement aux étymons de B.L.R.1 qui étaient en majeure partie, pour cette forme, considérés comme des séries localisées.

On perçoit bien que dans les cooccurrences entre $*C_1$ et $*C_3$, il n'est plus du tout question de contraintes de lieu d'articulation et les cooccurrences absentes jusqu'alors comme b/m ou d/n deviennent possibles.

2.1.3.2 - Tableaux

B.L.R.1 : pourcentages C_1/C_3 dans les verbes *-CVCVC- ³¹											
	*d	*k	*m	*t	*n	*b	*ŋg	*g	*c	*j	*n
*j	8,33	5,09	1,85	1,39	0,46	1,85		0,46			19,4
*k	6,94	1,39	4,17	1,39	0,46	0,93	0,93	0,46		0,46	17,1
*t	10,2	2,78	1,85	1,39	0,46						16,7
*p	4,63	3,24	0,93	2,31	0,46				0,93		12,5
*d	5,09	2,31	1,85	0,46	0,93						10,6
*c	3,24	1,85	0,93	0,93	0,46						7,41
*b	2,78	1,39	0,46	0,93			0,93	0,46			6,94
*g	1,85	2,31	0,93	0,46	0,93	0,46					6,94
*m		0,46			0,46					0,46	1,39
*n	0,46				0,46						0,93
	43,5	20,8	13	9,26	5,09	3,24	1,85	1,39	0,93	0,46	0,46
											/216

Tableau 71

Cs: occurrences C_1/C_3 dans les verbes *-CVCVC-											
	*d	*k	*m	*t	*n	*p	*ŋg	*b	*g	*c	
*y	8,71	7,21	1,5	1,5	1,5	0,9		0,9	0,3		22,5
*k	6,91	2,7	2,4	2,1	0,3	0,3	0,9				15,6
*d	7,81	3,9	1,2	1,2	0,3					0,3	14,7
*t	5,71	3,3	1,5	0,6	0,6						11,7
*p	3,9	3,3	1,5	1,2	0,6					0,3	10,8
*c	4,8	1,8	0,3		0,9						7,81
*b	2,4	3,3	1,2				0,3		0,3		7,51
*g	3	1,8	0,6						0,3		5,71
*m	0,6	1,5									2,1
*n	0,3			0,3		0,3					0,9
*j	0,3	0,3									0,6
	44,4	29,1	10,2	6,91	4,2	1,5	1,2	0,9	0,9	0,6	/333

Tableau 72

³¹ Ce tableau compte quatre alternances de * C_1 et une de * C_3 .

2.2 - Thèmes nominaux

2.2.1 - Cooccurrences $*C_1/*C_2$

2.2.1.1 - Dans B.L.R.1

Je présenterai le tableau des cooccurrences consonantiques uniquement dans les suites *-CVCV car les *-CVCVCV, les *-CVVCV et les *-CVCVV sont trop peu nombreuses dans ce corpus pour être analysables en fonction des cooccurrences C_1/C_2 , je vérifierai simplement si ces suites modifient la distribution et leur tableau sera consultable en annexe.

B.L.R.1 : pourcentages $*C_1/*C_2$ dans les noms *-CVCV ³²																	
	*d	*ŋg	*mb	*k	*nd	*b	*m	*g	*t	*n	*p	*j	*n j	*c	*ŋ	*n t	
*k	3,81	0,79	1,57	1,44	2,1	0,79	0,79	0,52	1,18	0,52	1,05	0,26	0,39	0,66			
*j	2,23	1,71	0,92	1,31	0,13	1,18	1,05	1,44	0,79	1,05	0,13	1,18	0,39	0,26	0,13	0,13	
*b	4,07	1,71	1,44	0,26	0,92	0,79	0,26	1,31	0,52	0,39		0,79	0,52	0,26			
*d	2,62	2,36	1,44	0,66	0,79	0,92	1,44	1,05	0,39	0,13	0,13						
*t	1,84	1,84	1,97	1,05	1,57	0,39	0,92		1,05	0,39	0,13			0,13	0,26		
*g	2,89	1,18	1,71		1,44	1,44	0,66	0,39	0,13	0,52		0,13	0,52			11	
*p	2,23	1,31	0,66	1,97	0,79		0,26	0,39	0,79	0,52	0,79			0,26		9,97	
*c	1,57	2,36	1,05	1,31	0,26	0,26		0,26	0,26	0,79	0,66		0,39	0,13	0,13	9,45	
*n	0,13	0,13		0,13						0,66						1,05	
*ŋ		0,26		0,26			0,39		0,13							1,05	
*m	0,39		0,13		0,13									0	0,13	0,79	
	21,8	13,6	10,9	8,4	8,14	5,77	5,77	5,38	5,25	4,99	2,89	2,36	2,23	1,71	0,66	0,13	762

Tableau 73

Distribution générale :

Dans les *-CVCV comptabilisés ci-dessus, les cooccurrences les plus fréquentes sont représentées par la bilabiale *b en $*C_1$ avec l'alvéolaire *d en $*C_2$. Dans le même ordre d'idée, les cooccurrences $*g_1 - *d_2$ atteignent un total de 3% ce qui pourrait indiquer une harmonisation de voisement, d'autant que *g ne compte aucune occurrence avec les non voisées *k, *t, *p. La cooccurrence entre *g et *t,

³² Ce tableau compte trente alternances de $*C_1$ ou de $*C_2$ pour un total de 732 étymons nominaux.

signalée dans le tableau, correspond au thème $^*g\ddot{\text{v}}t\text{i}$ « *neck* » refusé dans B.L.R.3 car explicable avec la règle de Dahl.

Contraintes :

Des contraintes se manifestent par des absences de cooccurrences et également par des pourcentages inférieurs à ceux que l'on pourrait attendre ; ainsi, on retiendra qu'il n'y a aucune cooccurrence entre *p_1 et *b_2 , entre *b_1 et *p_2 , entre *g_1 et *k_2 et également entre *t_1 et *g_2 , entre *g_1 et *p_2 .

Du point de vue des restrictions, le pourcentage des cooccurrences entre *k_1 et *ng_2 , avec seulement 0,8% est particulièrement faible puisque *k_1 occupe le premier rang et *ng_2 le deuxième. Avec un peu plus de réserves, je relèverai *j_1 et *nd_2 ainsi que *c_1 et *d_2 , *t_1 et *d_2 .

La suite $^*-\mathbf{dVnV}$ est représentée par le nominal $^*-d\text{vna}$ « *lower abdomen* », que je n'ai pas retrouvé sous cette forme dans B.L.R.3, mais sous la forme $^*-n\text{̄nā}$ en tant qu'étyomon. La suite à initiale *d correspond plus sûrement à une variante locale et d'après les données sur les verbes, la forme avec *d à l'initiale pourrait provenir de la zone A ?

2.2.1.2 - Séries comparatives

Les cooccurrences seront présentées sous forme de tableau pour les suites *-CVCV et *-CVVCV.

2.2.1.2.1 - Thèmes *-CVCV

Cs : pourcentages ${}^*C_1/{}^*C_2$ dans les noms *-CVCV ³³																				
d	${}^\eta g$	*mb	*nd	*k	*n	*b	*m	*t	*g	*p	*c	*y	${}^*n j$	*j	*n	*nc	*nt	${}^*\eta k$		
*y	2,14	1,32	0,81	0,71	1,73	1,93	1,02	1,42	1,02	1,02	0,2	0,71	0,71	0,41	0,51	0,31	0,61		16,6	
*k	3,05	0,51	0,92	1,63	0,41	0,51	0,92	0,61	1,02	0,41	1,02	0,61	0,31	0,2	0,1		0,1		12,3	
*d	1,93	2,24	1,42	1,22	0,51		1,32	1,42	0,31	1,22	0,41								12	
*b	4,37	0,92	1,12	0,51	0,2	0,81	0,61	0,31	0,31	1,12		0,41	0,31	0,41	0,1				11,5	
*g	3,46	1,02	0,92	1,12	0,1	0,61	1,42	0,31	0,1	0,2	0,1		0,51	0,2					10,1	
*p	2,24	1,02	0,61	0,51	1,42	0,51		0,1	0,92	0,31	0,41	0,41	0,1	0,51		0,1			9,16	
*c	1,32	2,03	0,92	0,51	1,42	0,51	0,2	0,2	0,41	0,41	0,61	0,41	0,1	0,1					9,16	
*t	1,12	1,22	1,02	1,22	0,81	0,41	0,41	0,81	0,51	0,1	0,1					0,1			7,83	
*j	1,42	0,1	0,41	0,2	0,2	0,2	0,51		0,1	0,41				0,1					3,66	
*n	0,1	0,81	0,2	0,1	0,41	0,41		0,31	0,41				0,1		0,1				2,95	
*n	0,1	0,2			0,31	0,81		0,31				0,1					0,1		1,93	
*m	0,51		0,1			0,41			0,1			0,1	0,1			0,1			1,42	
${}^*\eta g$	0,1	0,1									0,1		0,1						0,41	
*mb		0,1	0,1	0,1															0,31	
${}^*\eta$			0,1					0,1											0,2	
${}^*\eta k$				0,1					0,1										0,2	
*nd	0,1																		0,1	
*nc	0,1																		0,1	
*nj	0,1																		0,1	
	22,2	11,6	8,65	7,93	7,53	7,12	6,41	5,9	5,29	5,19	2,95	2,64	2,34	2,03	0,71	0,61	0,61	0,2	0,1	/983

Tableau 74

Distribution générale :

Je rappellerai d'une part, la place prépondérante de la palatale *y dans les Cs puisqu'elle occupe le premier rang des suites -CVCV et d'autre part, celle du

³³ Dans ce total de 983 cooccurrences sont incluses 7 alternances de *C_2 et 35 alternances de *C_1 .

phonème *d en *C₂ car il obtient un pourcentage près de deux fois supérieur à celui de *ŋg₂. Cette quantité élevée de d₂ se retrouve dans ses cooccurrences avec les autres phonèmes et tout spécialement lorsque les voisées *b et *g sont en position initiale. En effet, 39% des *b en *C₁ ont *d en *C₂ (20% des *d en *C₂ sont après *b) et 34% des *g en *C₁ ont *d en *C₂. Par contre cela ne se vérifie pas quand *d est en *C₁ car alors la cooccurrence maximale est avec *ŋg en *C₂. Il en est de même pour l'alvéolaire non voisée *t et la palatale *c qui connaissent plus d'occurrences avec les mi-nasales qu'avec *d. L'organisation globale ressemble à celle perçue pour les racines verbales.

Contraintes :

Du côté des restrictions, on retiendra le très faible pourcentage des cooccurrences entre *k₁ et *ng₂. Cette restriction entre une non-voisée et une mi-nasale n'apparaît que pour les vélaires, elle ne se retrouve pas avec les autres lieux d'articulation, contrairement à ce que j'avais observé pour les verbes.

Toujours à propos des vélaires, une seule Cs autorise la cooccurrence *g₁ - *k₂ : Cs 828 *-g ī k v « *foyer* », mais elle est refusée dans B.L.R.3 au profit de l'étymon *- j ī k v, donc on peut considérer que cette cooccurrence n'est pas autorisée dans les thèmes *-CVCV.

Enfin, deux types de cooccurrences sont absents, il s'agit de *d₁/*n₂ de *p₁/*b₂ et de *b₁/*p₂. Ces trous dans le système seront discutés au cours de la synthèse, car ils se constatent dans tous les corpus analysés.

2.2.1.2.2 - Thèmes *-CVVCV

Cs : pourcentages $*C_1/*C_2$ dans les noms *-CVVCV																
	$*d$	$*mb$	$*\eta g$	$*t$	$*n$	$*nd$	$*p$	$*k$	$*b$	$*g$	$*y$	$*j$	$*m$	$*nj$	$*c$	
$*k$	5,1	1,02	1,02			4,08	2,04	4,08	1,02	1,02	1,02				20,4	
$*t$	2,04	1,02		6,12	1,02			2,04	1,02						13,3	
$*d$	4,08	2,04	5,1	2,04						1,02					14,3	
$*b$	6,12	1,02	1,02			1,02			1,02			2,04		1,02	13,3	
$*p$		1,02				1,02	4,08			1,02			1,02		8,16	
$*c$	1,02	1,02	2,04		2,04	1,02							1,02		8,16	
$*g$	1,02	2,04	1,02		2,04	1,02									7,14	
$*j$	3,06	1,02													4,08	
$*m$											2,04		2,04		4,08	
$*n$					3,06										3,06	
$*y$	2,04														2,04	
$*p$	2,04														2,04	
	26,5	10,2	10,2	8,16	8,16	8,16	6,12	6,12	3,06	3,06	3,06	2,04	2,04	2,04	1,02	/98

Tableau 75

Toutes les séries comparatives de ce tableau, construites avec les mi-nasales $*mb$ $*nd$ $*nj$ en position C_2 et précédées par une voyelle double isotimbre ont un doublet avec voyelle simple. Les séries comparatives construites avec $*\eta g$ présentent la même situation à l'exception de trois d'entre elles : les Cs 552, 599 et 600 qui sont composées de voyelles hétérotimbres. Guthrie a appliqué sur l'ensemble du lexique son principe d'allongement devant les mi-nasales. Parallèlement il n'y a que trois *-CVVCV qui partagent le même numéro de série avec un *-CVCV en dehors du contexte mi-nasal, parmi elles deux sont refusées au profit de l'étymon à voyelle simple, la troisième est une variante (Cs 12881/2 *-má a y ñ) dont les réflexes sont ambigus du point de vue de la quantité vocalique.

Le nombre global d'occurrences n'est pas très élevé et le tableau devient plus délicat à analyser. Les cooccurrences notées absentes en *-CVCV, le sont également en *-CVVCV. Une particularité de cette suite tient à la réduplication avec allongement de $*V_1$ et cette réduplication se produit lorsque les consonnes non voisées $*p$, $*t$ et $*k$ sont en position $*C_2$.

2.2.2 - Cooccurrences $*C_2/*C_3$ dans les noms³⁴

Les consonnes occupant la position $*C_3$ des noms sont en nombre faible, quelque soit le corpus étudié. Les cooccurrences entre les $*C_2$ et les $*C_3$ présentent le même type de contraintes qu'entre les $*C_1$ et les $*C_2$, à savoir une tendance à la dissimilation des lieux d'articulation. On la perçoit bien à travers les deux cooccurrences les plus fréquentes : $^{*-}CV\mathbf{d}V\mathbf{\eta}gV$ et $^{*-}CV\mathbf{\eta}gV\mathbf{d}V$ (les voir en détail, voir aussi d-d et comparer avec le C1).

Les suites $^{*-}CV(V)\mathbf{d}V\mathbf{d}V$ sont, dans trois cas sur quatre, des dérivés construits à partir d'un radical $^{*-}CVC-VC-$ ou $^{*-}CVVC-VC-$, ce qui ne représente pas du tout le reste du corpus, car dans les autres contextes les déverbatifs sont rares.

Les cooccurrences notées absentes entre $*C_1$ et $*C_2$ le sont toujours entre $*C_2$ et $*C_3$ car la suite $^{*-}CV\mathbf{d}V\mathbf{n}V$ correspond à la Cs 988 $^{*-}kádì-anà$ « *girl* » qui est composée de deux morphèmes lexicaux.

2.2.3 - Cooccurrences $*C_1/*C_3$ dans les noms³⁵

³⁴ Tableaux 78, 79

³⁵ tableaux 76 77.

3 - Synthèse des observations

On vient de constater qu'il existe, dans les cooccurrences consonantiques, des contraintes récurrentes à travers l'ensemble des reconstructions. Ces contraintes touchent les traits de nasalité et de voisement des consonnes, mais leurs conséquences sur les cooccurrences s'entendent pour un même lieu d'articulation et varient - de rares à absentes - en fonction de ce dernier. Il y aurait alors un degré dans l'étendue de ces contraintes, degré qui pourrait nous indiquer une direction de changement.

Ces traits de nasalité et de voisement sont les deux traits dont la distribution est régie par les règles synchroniques que j'ai décrites en introduction (règles de Dahl, de Meinhof, de Greenberg). Dans les reconstructions, les contraintes sur le trait de nasalité montrent des similarités avec celles présentes dans les langues actuelles, alors que les contraintes sur le trait de voisement s'en éloignent. Comment comprendre cette situation, implique-t-elle les mêmes processus et génère-t-elle les mêmes distributions ?

Je propose de commencer par un résumé des contraintes reconstruites par Meeussen et Guthrie. Avant de tirer des conclusions, je vérifierai dans B.L.R.3 si ces contraintes sont confirmées ou infirmées. Ensuite, à partir de généralisations obtenues en comparant les classes de consonnes, je chercherai la place que prennent ces contraintes dans l'organisation phonologique du proto-lexique. Enfin, je ferai un lien avec les processus synchroniques d'assimilation et de dissimilation, avec comme objectif de montrer que si certains processus étaient déjà en place en proto-bantou, d'autres, comme la règle de Dahl, sont plus récents mais trouvent leur motivation dans les distributions marquées de la proto-langue.

3.1 – Résumé des contraintes

3.1.1 - Contrainte de voisement

3.1.1.1 - Les bilabiales

Quelle que soit la catégorie grammaticale à laquelle appartient l'étymon, aucune des reconstructions de Guthrie ni de Meeussen n'a la bilabiale non voisée *p en position C₁ avec la bilabiale voisée *b en position C₂. Guthrie pose, pour cette cooccurrence, deux séries partielles qui sont issues de la même racine, comme on peut le constater ci-dessous :

ps 401 -p^vb- *être* sotCe NE : FGM

ps 402 -p^vbá sot SW Ce : HMN

Vingt fiches sont disponibles dans B.L.R.3 pour ce contexte, mais le résultat reste identique puisque cette cooccurrence n'appartient qu'à des radicaux dont les attestations localisées se situent principalement dans la zone J, ce qui laisserait supposer un développement plus récent. Une recherche de cognats³⁶ dans les langues de l'ouest et du sud du domaine pourrait s'avérer intéressante.

Si on poursuit l'observation de ce contexte entre bilabiales en s'interrogeant sur la validité des cooccurrences entre la consonne voisée *b en C₁ et la non voisée *p en C₂, le seul substantif disponible appartient aux séries partielles, mais il est refusé par B.L.R.3 car il s'explique par l'application de la règle de Dahl : ps 19 *-b ē p ò « *froid* » et c'est l'étymon *-p ē p ò qui reçoit un statut de reconstruction générale.

Par contre, dans les verbes, les rares étymons qui figuraient dans les Cs et dans B.L.R.1, sont bien validés par B.L.R.3 :

Cs 171-b ó p- *lier* **Etymon** PB

Cs 94 -b ē e p- *mentir* **Etymon** PB

La cooccurrence b₁ / p₂ appartient à la proto-langue dans les racines verbales, avec un thème nominal dérivé de la Cs 94. En définitive, la seule cooccurrence qui reste absente est celle qui implique /p/ en *C₁ et /b/ en *C₂.

3.1.1.2 - Les vélaires

La cooccurrence g₁ / k₂ est absente des verbes et représentée dans un seul nom : *-g ī k ē « *sauterelle* » avec des cognats dans les zones J S. Les données de B.L.R.3 maintiennent un statut de **variante** au thème -g ī k ē et un statut d'**étymon** pour la forme avec deux consonnes voisées *-g ī g ē.

Dans B.L.R.3, trois fiches font état de cette cooccurrence : la variante, que je viens de citer, à laquelle s'ajoutent deux substantifs dont l'un est un composé (*-gia-kadi), et l'autre est refusé (*-g ī k ò « *fireplace* »). La cooccurrence g₁ / k₂ reste totalement absente de la proto-langue et même des reconstructions régionales puisque aucune autre fiche n'y figure.

Parallèlement à cette non-reconstruction, les tableaux indiquaient également pour *g un évitement vis-à-vis de toutes les consonnes non voisées. Que deviennent ces rares cooccurrences dans les données plus récentes de B.L.R.3 ?

- *-gVp(V)(C) : sur quatre fiches, deux sont les variantes d'une reconstruction comportant une autre consonne en position *C₁, une est refusée à cause de la règle

³⁶ Une première recherche a été effectuée dans ce sens sur les langues du groupe B10, mais n'a donné aucun résultat.

de Dahl, la quatrième est localisée au NW et n'est rattachée à aucun autre étymon :

ps 245	<i>*-g ūpá</i>	<i>insect, tick,</i>	Refusée, étymon PB <i>*-k ūpá</i>
Cs 970	<i>*-g áp í</i>	<i>paddle, 9/10</i>	Variante de l'étymon PB <i>*-k áp í</i>
Cs 833	<i>*-g ìp a g-</i>	<i>kill</i>	Variante SW de l'étymon PB <i>*-y ìp a g-</i>
Tv	<i>*-g íp-</i>	<i>suffocate</i>	NW : C

• **-gVt(V)(C)* : le nombre de fiches est supérieure puisque la requête en fait apparaître vingt-trois, mais dix-sept restent très localisées, cinq sont refusées et enfin une variante peut-être soumise à la règle de Dahl

Tv	<i>*-g à à t ì</i>	<i>bread</i>	Variante NE de l'étymon PB <i>*-k á á t í</i>
----	--------------------	--------------	---

L'évitement entre la vélaire voisée **g* et une consonne non voisée en position **C₂* se confirme pleinement. Il appartient bien à la proto-langue telle qu'elle a été reconstruite de nos jours.

En résumé, l'opposition de voisement est limitée (cf. §3.2), voire empêchée, entre deux consonnes appartenant à des syllabes successives. Cette contrainte s'exprime entre consonnes bilabiales et de façon plus large avec la consonne vélaire voisée quel que soit le lieu d'articulation de la consonne suivante à partir du moment où cette dernière est non voisée. Ceci aboutit à la non-reconstruction des séquences suivantes :

p₁/b₂, p₂/b₃
g₁/k₂, g₂/k₃
g₁/t₂, g₂/t₃
g₁/p₂

3.1.2 - Contrainte de nasalité

3.1.2.1 - Les bilabiales

La cooccurrence *b₁/m₂* demeure absente des étymons verbaux de Guthrie et de Meeussen. Par contre, la série partielle qu'avait proposée Guthrie prend dans B.L.R.3 un statut d'étymon : ps 48 **-b ūm-* « *frapper, tuer* ». Les thèmes nominaux³⁷,

³⁷ Cs 159, 228, 229.

présents dans les lexiques étudiés, sont eux aussi confirmés à un niveau de proto-langue, ce qui permet de rétablir cette cooccurrence pour les bilabiales.

3.1.2.2 - Les alvéolaires

Le phonème *d ne se trouve pas en *C₁ avec *n en *C₂ et, dans le cas des alvéolaires, cette restriction touche tous les étymons, c'est-à-dire les deux catégories nominale et verbale et qui plus est dans tous les proto-lexiques. Les rares entrées qui figurent dans B.L.R.3 le sont à titre de dérivées, rattachées à un étymon de forme *-CV-. La nasale *n appartient alors au morphème grammatical *-an-, ayant valeur de réciproque.

Enfin un radical est retenu sous la forme *-dùn-, il est considéré comme une variante de la zone A et est rattaché à l'étymon *-nùn- « être vieux ». Cette variante sous-entend qu'en zone A, un processus de dissimilation entre consonnes nasales a été, et est peut-être encore, actif. Schadeberg (2002) parle de cette variante de zone A avec un *C₁ non nasal comme « *part of the unsolved nasality puzzle* ». Une autre fiche existe, mais elle reste à un stade régional avec une incertitude sur *C₁ : *-dìn- (J L) « *hésiter, ignorer* ».

Cette contrainte touche également la cooccurrence *C₁ = n et *C₂ = d, elle n'était pas évidente dans les Cs ni dans B.L.R.1, mais une requête dans B.L.R.3 établit son absence dans les noms et dans les verbes à voyelle simple, car les quelques items proposés sont les variantes d'un étymon où les deux consonnes sont nasales. La cooccurrence *C₁ = n et *C₂ = d n'est pour l'instant reconstruite que dans deux suites *-CVVC-.

Ce type de contrainte entre une occlusive voisée et la nasale de même lieu d'articulation toucherait uniquement les alvéolaires voisées. Parallèlement, la cooccurrence entre l'alvéolaire non voisée *t et la nasale *n ne connaît aucune restriction. Les bilabiales admettent ce type de cooccurrences, même si elles restent rares. Pour les vélaires, cette cooccurrence est inenvisageable, vu que la nasale vélaire n'appartient pas au proto-système.

À propos des cooccurrences consonantiques en luganda, Katamba et Hyman (1991) observent : « ...consonants sharing the same place of articulation, but differing only in nasality, do not freely co-occur in lexical roots ». Dans cette langue, les contraintes exprimées s'entendent pour un lieu d'articulation défini et touchent l'ensemble du système ; comme si la contrainte, apparente pour les alvéolaires en proto-bantou, avait affecté l'ensemble du lexique de cette langue. Un autre point commun apparaît entre cette langue et les reconstructions, il est dans le nombre

élevé de cooccurrences entre occlusives voisées et mi-nasales, ce qui semble contrebalancer la contrainte énoncée.

Pourquoi les alvéolaires ? Cette restriction entre l'occlusive voisée *d et la nasale du même lieu d'articulation n'est pas sans rappeler la règle d'assimilation nasale progressive de Greenberg, puisqu'elle concerne les consonnes alvéolaires. Mais la règle de Greenberg s'applique entre la consonne nasale du radical et la consonne *d du suffixe qui devient alors *n, c'est-à-dire entre un morphème lexical et un morphème grammatical. Il est vrai par ailleurs que la cooccurrence *d/*n n'est pas reconstruite entre les consonnes *C₂ et *C₃ et qu'elle figure dans les reconstructions seulement entre *C₁ et *C₃ des *-CV-VC.

d₁/n₂, d₂/n₃

3.1.2.3 - Autre contrainte

Lorsque l'alvéolaire non voisée *t est en *C₁, la palatale *c est absente de la position *C₂. Cette situation est maintenue dans B.L.R.3, où encore une fois, les trois fiches contenant cette cooccurrence ont une distribution limitée (il y a cinq fiches, mais deux d'entre elles ont un *C₁ incertain : /c/ ou /t/). C'est le seul cas de contrainte entre deux phonèmes ne partageant pas le même lieu d'articulation. Le phonème /c/ fonctionnerait-il avec les alvéolaires ? En tout état de cause, il partage avec elles le trait coronal. Meeussen (1965) avait remarqué que "...*un certain nombre de correspondances attestent une hésitation entre /t/ et /c/*". Il semble que les alternances entre les réflexes de *t et de *c trouvent leur explication dans l'évolution de *t qui, après la voyelle fermée *i, a évolué vers une forme identique à celle du réflexe régulier de *c (Bastin 1983), d'où les risques de confusion et les hésitations des reconstructeurs. Cette confusion des réflexes ne nous dit rien de la contrainte relevée.

t₁/c₂

3.2 - Organisation phonologique du lexique

Une question s'avère maintenant importante à poser, c'est celle de la place que prennent ces contraintes dans le lexique. Sont-elles révélatrices d'une règle plus générale ou bien restent-elles ponctuelles et traces d'un processus qui nous échapperait ?

Avec cet objectif, j'ai effectué des regroupements de consonnes pour observer les éventuelles relations de dépendance du point de vue du lieu d'articulation, de l'activité laryngale et du mode articulatoire puisque des tendances de cet ordre, semblaient se profiler dans les tableaux.

Pour chaque regroupement observé, j'ai utilisé le test du chi-2 avec une lecture à deux niveaux ; La distribution des cooccurrences consonantiques est-elle ou non significative ? Si oui, quel type de cooccurrences contribue à rendre la distribution significative ?

La plupart des tableaux ont un caractère significatif, et ceux qui ne le sont pas ressemblent trait pour trait aux fréquences attendues !

3.2.1 - Lieux d'articulation

Les consonnes ont été regroupées en respectant les quatre lieux d'articulation du système phonologique proto-bantou. Le test du chi-2 appliqué à ces données s'est révélé significatif pour les *-CVC- ($p = 0,001$), les *-CVCVC- ($p = 0,05$ dans B.L.R.1 et $p = 0,001$ dans les Cs) et *-CVCV ($p < 0,001$). Ce test souligne l'existence de relations de dépendance où le hasard n'a pas sa place, quelle que soit l'interprétation que l'on attribue à ces valeurs.

Les déséquilibres significatifs sont comparables d'un type de suite à l'autre et ils se font dans le sens d'une **différenciation** des lieux :

- Déséquilibre en excès pour :
 - Les cooccurrences entre labiales et alvéolaires
 - Les cooccurrences entre alvéolaires et labiales
- Déséquilibre en défaut pour :
 - Les cooccurrences entre labiales
 - Les cooccurrences entre vélaires
 - Les cooccurrences entre alvéolaires et palatales

L'évitement des alvéolaires avec le groupe des palatales sous-entend la pertinence du trait coronal car alors alvéolaires et palatales appartiendraient au même groupe et respecteraient une organisation du lexique en termes de différenciation des lieux d'articulation.

Ces relations de dépendance ne sont pas sans rappeler le principe du contour obligatoire (OCP) qui évite de mettre côté à côté deux segments identiques (cf.

Roca and Johnson 1999). Pour Hyman et katamba, c'est l'OCP qui motiverait les contraintes du luganda.

Pour les *-CVVC-, je n'ai pas pu utiliser le test de chi-2 car les conditions d'application n'étaient pas respectées³⁸. Je m'appuie sur une simple comparaison entre les fréquences observées et les fréquences théoriques pour indiquer que, dans chaque corpus, il n'est plus question de dissimilation entre les lieux d'articulation. Les corrélations entre *C₁ et *C₂ deviennent semblables aux corrélations perçues entre *C₁ et *C₃, ce qui va dans le sens de l'hypothèse probable où une consonne se serait amuïe, donnant ainsi naissance aux *-CVVC-. De ce fait on peut soulever la question de l'influence d'un contexte *C₁ = *C₃ sur l'amuïssemement de *C₂.

3.2.2 - Activité laryngale

les consonnes ont été réparties en deux groupes³⁹ selon la valeur qu'elles prennent pour le trait de voisement.

La distribution consonantique, dans les thèmes nominaux *-CVCV et dans les radicaux *-CVC- des Cs et de B.L.R.1, s'avère franchement significative ($p < 0,001$). Deux types de cooccurrences contribuent à ce caractère significatif :

- Le nombre élevé de cooccurrences entre les consonnes non voisées.
- Le nombre faible de cooccurrences entre les consonnes initiales voisées et les consonnes C₂ non voisées.

Pour les autres cooccurrences, l'écart entre les fréquences théoriques et les fréquences observées n'est pas significatif d'un point de vue statistique, mais leur distribution va dans le même sens, c'est-à-dire vers une **harmonie de voisement**. Cela est perceptible du fait que l'on compte pour les suites composées de deux consonnes voisées plus de cooccurrences et pour les suites composées d'une consonne non voisée en *C₁ et d'une voisée en *C₂, des occurrences inférieures au nombre prédit.

³⁸ Un minimum de cinq occurrences est nécessaire .

³⁹ J'ai inclus les mi-nasales non voisées dans le groupe 1 des consonnes non voisées. J'ai effectué les calculs avec et sans ces mi-nasales, en fait leur quantité est trop faible pour avoir une quelconque influence sur la totalité du tableau. Leur distribution dans les *-CVC- de B.L.R.1 est intéressante à retenir car 15 occurrences sur 19 sont après un C₁ non voisé.

Les verbes n'ont pas tous le même comportement ; les *-CVC-, comme les noms, présentent des corrélations significatives en privilégiant une suite de deux consonnes non voisées et en contraignant les suites où le *C₁ est voisé et le *C₂ non voisé.

Dans les *-CVVC-, la distribution consonantique est strictement identique à celle prédicté par les fréquences théoriques ! Enfin dans les *-CVCVC-, la distribution du trait de voisement suit l'idée d'une harmonie sans pour autant être statistiquement significative.

3.2.3 - Mode articulatoire

Les consonnes ont été réparties en trois groupes : occlusives, nasales et mi-nasales, puis en quatre groupes dans le but de distinguer les palatales /*y *j *c/ des occlusives, puisque leur mode serait plutôt fricatif ou affriqué, mais non occlusif, contrairement à ce que laisse supposer le symbole qui les représente. La séparation des palatales n'a aucun caractère significatif.

Tous les tableaux sont statistiquement significatifs du fait d'un ou de deux types de cooccurrences qui mettent en lumière des relations de dépendance :

- Déséquilibre en excès entre consonnes nasales dans les noms et dans les verbes.
- Déséquilibre en défaut entre consonnes nasales et mi-nasales dans les verbes.

Un déséquilibre intéressant est présent dans B.L.R.1 : les cooccurrences avec les mi-nasales non voisées sont significatives du fait d'une quantité faible avec les occlusives voisées et d'une quantité élevée avec les non voisées.

En définitive, les contraintes distributionnelles reconstruites en proto-bantou vont dans le sens d'une **différenciation** des lieux d'articulation et d'une **harmonie** de voisement et de nasalité. Cette distribution fonctionne pour les cooccurrences *C₁/*C₂ et *C₂/*C₃, soit entre consonnes appartenant à des syllabes successives. Si on combine ces deux contraintes, on s'attend à ce que les étymons, construits avec un *C₁ et un *C₂ de même lieu d'articulation et en opposition de voisement ou de nasalité, soient limités en nombre, voire absents. C'est effectivement ce qui se produit, puisque les cooccurrences ne respectant pas cette organisation, ne concernent pas plus de 3,5% des reconstructions et font figure d'éléments marqués. Trois des suites non reconstruites *-pVb(V), *-gVk(V), *-dVn(V)

respectent ces tendances et peuvent s'interpréter comme la conséquence directe de cette organisation phonologique du lexique. Par contre, elles n'expliquent en rien la non-reconstruction des suites $*-tVc(V)$, car elle s'oppose doublement à une organisation du proto-lexique alliant différenciation des lieux et harmonie de voisement ; c'est donc une autre motivation qui la sous-tend.

Meeussen avait bien précisé qu'il aurait pu reconstruire les fricatives /s, z/ en lieu et place des palatales /c, j/. La non-reconstruction de consonnes fricatives en proto-bantou est un fait surprenant essentiellement pour deux raisons :

- D'une part, ce fait est contraire à la typologie des langues du monde. Dans la base UPSID⁴⁰ 93,5% des systèmes phonologiques ont des fricatives et dans la base BDPROTO⁴¹ 95,5% des proto-systèmes phonologiques en possèdent. Dans les deux bases les (proto-) systèmes phonologiques ont en moyenne quatre fricatives⁴², et 2,5 lieux d'articulation : alvéolaire et glottal. Bien entendu la typologie est un point de vue qui ne peut à lui seul justifier un choix, mais les données sont suffisamment franches pour qu'elles invitent à examiner attentivement la question des fricatives en proto-bantou.
- D'autre part, la palatale $*c$ aurait un réflexe fricatif alvéolaire lié au contexte nasal (Pn9) dans les langues actuelles : $*c \rightarrow s$, à tel point qu'une reconstruction sous la forme $*s$ serait justifiée.

Typologie et études comparatives des langues bantoues adoptent une direction commune. L'absence de cooccurrences entre $*t$ et $*c$ peut devenir un argument supplémentaire pour proposer une fricative alvéolaire en lieu et place de l'occlusive palatale.

Dans le contexte où l'exclusion est en lien avec le voisement, pourquoi du côté des labiales, la cooccurrence affectée touche la suite non voisée - voisée et du côté des vélaires la suite voisée - non voisée, plutôt que l'inverse ? On peut proposer une première explication en lien avec la phonétique générale dès lors que les deux consonnes, qui étaient attendues en position C₁, sont des consonnes décrites pour être les plus sensibles au changement au sein de leur lieu d'articulation respectif⁴³,

⁴⁰ UPSID : UCLA Phonological Segment Inventory Database.

⁴¹ BDPROTO : base de données sur les proto-systèmes phonologiques, cf. Marsico (1999).

⁴² Ces chiffres m'ont été transmis par Egidio Marsico (laboratoire DDL).

⁴³ Dans la série des bilabiales quand une occlusive disparaît c'est la non voisée, dans la série des

cette fragilité a pu accentuer le processus. Par ailleurs, la suite ${}^*-\mathbf{b}V(V)\mathbf{p}$ - a bien été signalée comme rare... Les labiales et les vélaires seraient les plus touchées par cette contrainte, puis les alvéolaires.

3.3 - Conclusion

3.3.1 - Contraintes et règles d'assimilation nasale

Les contraintes reconstruites autour du trait de nasalité peuvent-elles supporter le lien avec la règle d'assimilation nasale de Greenberg ?

Des ressemblances existent avec la règle de Greenberg, dans la mesure où il est question d'un processus d'assimilation nasale et que cette règle touche les alvéolaires.

Par contre en proto-bantou, ce processus s'appliquerait quelle que soit la place occupée par la nasale dans le thème. Contrairement à ce que Greenberg avait formulé cette règle agirait de manière progressive mais aussi régressive, et elle ne serait pas limitée à la frontière de morphèmes et toucherait le thème nominal et verbal. L'implication de cette règle dans le lexique serait plus forte, on aurait peut-être subi une évolution qui limiterait son domaine d'application aux suffixes.

3.3.2 - Contraintes et règle de Dahl

Si maintenant on établit un parallèle entre l'organisation du proto-lexique et les contraintes engendrées par la règle de Dahl, on perçoit que ces dernières impulsent une direction opposée car, en empêchant deux consonnes non voisées de se trouver dans des syllabes successives, elles entraînent une **dissimilation de voisement**. Avec cette règle, les cooccurrences $C[-\text{vois}]/C[-\text{vois}] \rightarrow C[+\text{vois}]/C[-\text{vois}]$, ce qui génère directement la cooccurrence g_1/k_2 , mais également g_1/p_2 , et g_1/t_2 .

Il est probable que l'application de la règle de Dahl vient combler les distributions marquées du proto-bantou, et que les « trous » dans le système ont favorisé son apparition. Si la règle de Dahl a bien puisé sa motivation dans la structure du proto-lexique, alors on s'attend à ce qu'elle touche progressivement les différents lieux d'articulation et que le lieu d'articulation de C_1 soit déterminant. Ainsi, elle devrait être particulièrement active avec les consonnes vélaires. Les cooccurrences

vélaires quand une occlusive disparaît c'est la voisée. Les raisons invoquées relèvent du domaine de l'acoustique.

entre la consonne vélaire voisée et les consonnes non voisées étant absentes de la proto-langue, elles peuvent émerger directement par application de cette règle. Ensuite, en s'appuyant sur les pourcentages reconstruits on peut déduire que les labiales seront atteintes dans un deuxième temps puis, en dernier lieu, les alvéolaires.

Cette hypothèse se vérifie en partie dans les langues étudiées par Davy et Nurse (1982) car, dans ces langues, lorsque le processus de dissimilation ne touche qu'une partie du système la consonne vélaire est systématiquement soumise à la règle de Dahl. On serait alors dans un cadre de processus en chaîne, où cette règle trouverait sa motivation dans les distributions marquées de la langue-mère.