

BEW II 5629 (1)

INSTITUT DES LANGUES ET CIVILISATIONS ORIENTALES

Par JS CD

**LE SHINGAZIDJA**

(grand-comorien)

**UNE LANGUE BANTU SOUS INFLUENCE ARABE**

Comment l'apport de nombreux mots d'origine arabe a modifié  
l'organisation des mots et des sons de la langue, suite  
d'un essai de dictionnaire etymologique  
complété d'index

vol I

**L'ORGANISATION DES MOTS ET DES SONS**

thèse de troisième cycle ancien régime

LANGUES AFRICAINES

présentée par

**MICHEL LAFON - OULD-AOUDIA**

sous la direction et la co-direction de  
Pierre ALEXANDRE, Professeur à l'INALCO  
Suzanne PLATIEL, chercheur au CNRS

DM 119351

Paris 1987



Michel LAFON-OULD AOUDIA  
Le shingazidja,  
une langue bantu sous influence arabe

- ERRATA -

La précipitation dans laquelle cette thèse a été achevée a entraîné un certain nombre d'erreurs et d'omissions, en rendant parfois la lecture difficile; il nous a paru un moindre mal de les signaler et d'en fournir au plus tôt un corrigé.

Nous prions les lecteurs d'excuser cette difficulté supplémentaire.

Nous avons repris intégralement le plan (vol. I), que l'omission de deux pages et l'inversion de deux autres rendent incompréhensible - il est paginé ici de i à xii - ainsi que la bibliographie (vol. III), qui contient de nombreuses inexactitudes.

La page 349 du lexique (lettre TSH) a été oubliée et est donc donnée.

Les autres corrections concernent exclusivement le premier volume; elles sont présentées dans l'ordre des pages; pour ne pas faire trop long, nous n'avons corrigé que les erreurs susceptibles de nuire à la compréhension, à l'exclusion des fautes d'orthographe ou de frappe.

Nous n'avons pas cherché à enrichir le lexique, malgré quelques nouvelles informations (nouveaux mots, précisions d'emploi ou de sens, etc), recueillies depuis; nous n'avons pas rectifié non plus les erreurs dans les graphies arabes des racines.

Une nouvelle édition de ce travail intègre toutes ces corrections.

volume I

première page de garde :

proverbe : lire *kapvana* (et non \**kapavana*)

deuxième page de garde (absente de la présente version):

Cette recherche a eu pour point de départ l'ATP 1951 du CNRS, dirigée par le Prof. Pierre ALEXANDRE; cette ATP a permis mes deux séjours aux Comores, en 1975 et 1977-8; outre moi-même, les bénéficiaires étaient Mlles Marie-Françoise ROMBI et Catherine BOZON-VERDURAS (cette dernière dans un premier temps seulement) et Mr Jean-Luc SIBERTIN-BLANC

Introduction générale

p 4, milieu de la page : lire carte 3 (et non 2)

p 11, avant dernier paragraphe : lire *culturelle* (et non *cultuelle*)

Méthodologie et définitions

p 32, fin du premier paragraphe, compléter la date de parution de l'ouvrage sur l'Expansion Bantoue: 1977

p 44, fin de page, *yatima* signifie orphelin de père (en face de *ndishe* orphelin de mère) ; *lamanda*, amende pénale

p 45, milieu de page, lire *urozi* (et non \**urosi*)

p 50, milieu de page, lire ... *étude des sons* (et non de phonologie dynamique); la *laryngale* arabe [q] (et non vélaire), la *vélaire* [k] (et non palatale)

p 51, fin de page, lire ... les dictionnaires *arabes* consultés

p 52, avant dernier paragraphe, compléter ... (voir *lexique et bibliographie* pour l'inventaire des sources)

p 58, fin du premier paragraphe, compléter ... quelques excursions dialectales au niveau du *lexique* grâce à DOZY ...

p 59, 6è ligne, compléter ... jusqu'à trouver une *étymologie satisfaisante* ; ligne ante-pénultième, lire ... le mot arabe (...) ne correspondra pas alors strictement au mot *shingazidja* (et non présentera alors ...)

p 60, 2è moitié de page, compléter ... dont le terme même *signifiant français*, Ng *farantsa*

p 62, dernière ligne, compléter ... et d'une liste d'étymologies orientales douteuses

p 71, fin de page, à propos de nb, ... (voir p 241 (et non en fin de paragraphe)

p 77, haut de page, ... -le notion de grandeur! (et non d'éloignement)

p 78, milieu et bas de page, lire AL-ANI (et non \*EL-ANI)

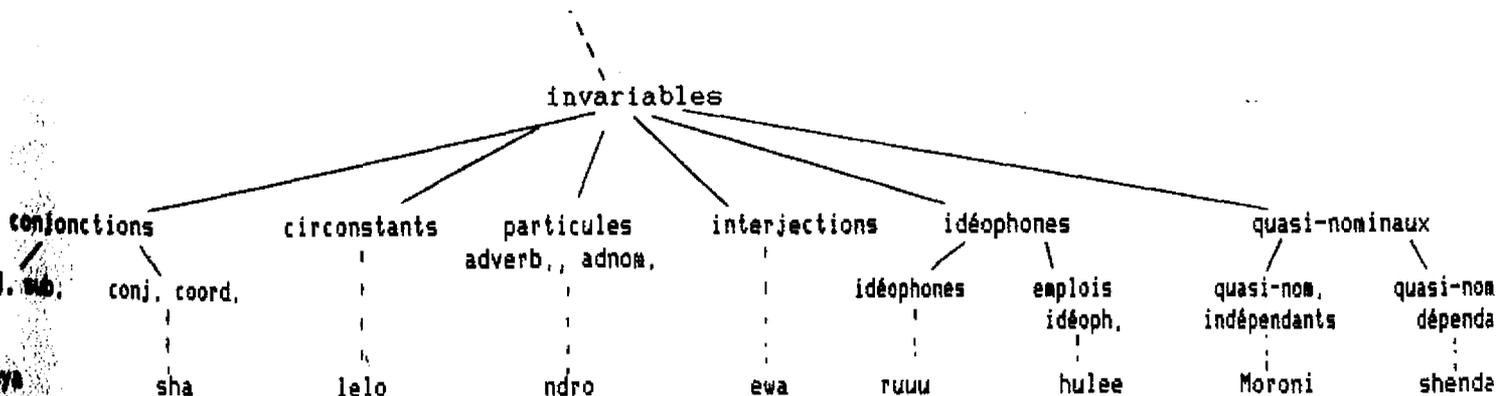
*En addenda au chapitre sur le bantu, signalons la parution récente (dont nous n'avons eu connaissance qu'après l'achèvement de notre travail), d'un article de DEREK-NURSE où celui-ci étudie la fréquence, dans différentes langues bantu de l'Est, de diverses évolutions phonétiques à partir du Proto-Bantu : les conditionnements que nous avons définis se retrouvent pour la plupart dans ces évolutions (DEREK-NURSE 1987 Toward a Typology of Diachronic Phonological Change in Bantu Languages, Revue de l'Association des Provinces Atlantiques, Canada, vol 9, pp 100 - 122)*

#### Constitution des mots

p 82 schéma : une accolade liant recensement organisé montre que la lexicographie est l'aboutissement d'un recensement organisé

p 89 bas de page, lire Itsandra (...) le 'thème' ainsi isolé ne paraît pas ...

p 91 il manque dans le tableau des invariables, l'extrême-droite, constituée des quasi-nominaux : voici le schéma des invariables dans son intégralité:



p 110, premier tiers: wamaliza, ils/elles ont terminé (et non sont ...)

p 118, premier tiers, renvoi manquant, rajouter pp 167 ; de même en bas, p 329

p 119, bas du deuxième tiers, renvoi manquant, rajouter p 259

p 121, milieu : lire **-huu, mfu** : **-fa** et **-hua** ont pour voyelle spécifique à l'accompli /u/ (ci-dessus p 102 et sur la voyelle de l'accompli, p 110) ; paragraphe en dessous, lire ... nous *parlons* alors de sème (et non nous partons)

p 130, milieu, une accolade joint |pvo-l, |ho-l et |mo-l, classes dites locatives

p 135, troisième tiers : lire ... identique à l'allomorphe dji- du préfixe de cette classe ; supprimer le renvoi (incomplet) sous **dzina**

p 136, deuxième paragraphe, compléter l'indication de page, 259

p 137, premier paragraphe, lire

**lekoko ngilo bwade**, la grand-mère est malade (plutôt que \*lekoko ngilo djiwade, qui serait très péjoratif, la personne tant en quelque sorte réifiée (un autre accord, que nous appelons notional, est aussi possible, qui fera mettre les éléments soumis à l'accord de classe, dont l'adjectif, en cl. 1 des animés: **lekoko ngilo mwade**; c'est sans doute la forme la moins connotée)

fin de page, lire

**buzi dume**, bouc |Ø\*-buzi#Ø\*-ume| /

**mabuzi maume**, des boucs |ma-Ø\*-buzi#ma-ume|

**djindru**, géant |dji-ndru| /

**madjindru**, des géants |ma-dji-ndru|

**daho**, maison, |Ø\*-laho| → **malaho**, des maisons |ma-laho|

→ **madaho**, de grandes maisons |ma-Ø\*-laho|

**paha**, chat |Ø\*-pvaha| → **mapvaha**, des chats |ma-pvaha|

→ **mapaha**, de grands chats |ma-Ø\*-pvaha|

p 138, deuxième paragraphe, lire

**dzi-** est sans doute le reflet normal de BC \*di\*, préf. cl. 5; **dji-** correspond lui au préfixe de la classe 21 augmentative \*gi (source bibliographique, ci-dessous p 261); comme par ailleurs le phonème /dz/ a été concurrencé par /dj/ introduit à la suite de l'influence arabe (voir p 326), il a pu se produire une confusion entre dj et dz, qui a entraîné une confusion des classes \*21 et \*5, ce qui a entraîné une valeur augmentative dans la classe 5 *tout en marginalisant dz*, lequel ne reste alors plus que dans des mots anciens, ce que montrent les quelques formes où il est attesté, notamment l'adjectif **-du**, noir, dont l'ancienneté est confirmée par les accords de cl. 1 & 3 (cf. p 131).

p 139, milieu, lire **mapaha** (et non \*mapvaha)

p 140, fin de page, compléter l'indication de renvoi p. 279

p 142, troisième tiers, compléter l'indication de renvoi p. 242

p 144, fin de page, lire **ndemevu** (et non \*demevu)

p 147, deuxième tiers, supprimer le renvoi (incomplet) sous **-andzih-**

- p 148, première ligne, lire pré-préf. cl. 15 ou préf. cl. 17  
 p 165, milieu de la page, suppléer l'indication de renvoi p. 253  
 p 166, fin de page, suppléer l'indication de renvoi p. 369  
 p 167, dernier paragraphe, lire

Quant à la seconde forme (\*iki\*), il faut remarquer que l'allomorphe shingazidja lui correspond dans la mesure où \*ki\* > sh, mais que la réflexion normale ici est une fricative, s (voir p. 256).

Reste à ...

- p 168, troisième paragraphe, lire

Une autre possibilité serait que le causatif des termes d'emprunt soit composite: il s'agirait alors de l'association d'une extension de forme -ih- (reconstruite comme \*ik, à quoi GUTHRIE prête différentes valeurs) et du causatif \*i\* ; la séquence \*ik + \*i\* peut aboutir à Ng -ish- tout comme \*iki\*. Dans la mesure toutefois où nous n'avons pas d'indication claire qu'une telle composition se soit produite, et où la mention chez GUTHRIE de la série \*iki\* ne rend pas nécessaire cette hypothèse, nous nous en tenons à cette dernière (voir cependant ci-dessous),

- p 171, premier tiers, suppléer l'indication de renvoi p. 267  
 p 172, première ligne, lire

A l'appui de cette interprétation, les doubles causatifs, où l'on peut avoir, soit une forme avec -iz- / -is- / -ish-, construite sur une racine déjà causative (alternance), soit une succession -iz- + -iz-

- p 175, suppléer l'indication de renvoi p. 116  
 p 177, dans le schéma, remplacer le symbole -Z- (causatif) par -s-  
 p 182, milieu de page, lire

Nous avons signalé quelques autres mots où apparaît une séquence [uh], voire [ul] par chute du [h], que l'on pourrait rapporter à cette extension

- p 183, suppléer l'indication de renvoi p. 266

#### Les emprunts

- p 198, première liste de mots, diverses omissions dans les étymologies:

ar. *wirāta*, héritage  
 ar. *dāras.inī*, canelle  
 ar. *limūn*, citron  
 ar. *s'akka*, percer ou *s'aqqa*, fendre

plus bas

ar. *s'ari'a*, loi  
 ar. *bah.r*, mer

- p 199, deuxième moitié, omission de la traduction,  
 ar. *fä'ida*, gain

- p 201, 7è ligne avant la fin, lire ... intégration *par analogie phonétique*  
 p 203, milieu, lire ... chute de la consonne *finale* (et non désinentielle)  
 p 207, fin de page, suppléer l'indication de renvoi p. 202  
 p 210, deuxième liste, étymologies manquantes, lire

ar. *siwāk*  
 ar. *s'akka* ou *s'aqqa*  
 ar. *s'irka*

- p 211, troisième exemple, lire

ar. *miād*, →  
**miadi**, promesse(s) : cl. 4

D'où la construction du thème *l-adi*, notion de promesse, *confirmé par le verbe naturalisé -adiana, se promettre (cf p. 220)*

- p 215, milieu, lire ... causatif *-ish-* (-z-, -s-) < \**-iki\** (\*i\*)  
 p 220, dernier exemple, lire

**miadi** (cl. 4), promesses ou *-wadi*, arriver  
 → \*-ad- + -i- + -an- → *l-adian-l*, notion de s'engager!  
 → *-adiana*, s'engager

(ce dernier terme confirme le thème *-adi* extrait de **miadi** cf. p 211)

- p 222, deuxième exemple, rajouter l'étymologie:

\*-mkini < ar. *makkana*

- p 226, fin de la première moitié, lire

... ne peuvent se fonder sur des processus ...

- p 228, troisième paragraphe, lire

... spontanément (nous semble-t-il) par tous les locuteurs, *il existe des cas où l'apparemment les termes issus de la même racine arabe ne sera pas reconnu, ou pour lesquels l'attitude variera d'un locuteur à l'autre, selon sa propre réflexion sur sa langue, sa connaissance de l'arabe, l'intégration des termes en shingazidja, etc*

#### L'organisation des sons

- p 232, milieu de la page, lire ... thèse *en cours* de J-L SIBERTIN-BLANC

- p 242, deuxième moitié, citation de WELMERS, lire [ɣw]

- p 255, première ligne, lire

... le swahili a ici *-kesha*, rester éveillé, et *kesho*, demain, qui n'expliquent pas la forme shingazidja

deuxième tiers, compléter ... elision complète (voir étude syllabique *infra et inventaire des réalisations p 245-6*)

dernière ligne, lire ... dans le fonds bantu (supprimer voir étude syllabique)

p 256, milieu, lire

... On peut imaginer le processus suivant: une évolution  $ki^* \rightarrow s$ , puis, avec la réduction du système vocalique, l'intégration de cette séquence parmi les séquences \*k + voyelle d'avant, d'où  $*ki^* \rightarrow sh$  (pour le causatif -ish- < \*-iki\*, cela suppose une formation ultérieure - cette succession des faits est celle supposée par DEREK-NURSE "palatalisation of velars appears to have been very common (...) in most cases it follows chronologically [Bantu spirantization], DEREK-NURSE 1987 : 105, signalé ci-dessus).

p 261, fin du premier tiers, lire ... et par GUTHRIE sous la forme \*gi

p 264, fin du deuxième tiers, compléter

... continue ou occlusive (à l'exception notable de -bwana / -djibwana, ci-dessus)

p 265, fin du premier tiers, suppléer l'indication de renvoi, p 296

p 270, fin du premier tiers, suppléer l'indication de renvoi, p 269 & 119

p 271, milieu, lire cl. 10 (et non 10a)

p 273, milieu, lire ... notre hypothèse sur (et non à)

p 285, fin des exemples, lire ndziwa, pigeon (et non + ndziwwa)

p 287, milieu, suppléer l'indication de renvoi, p 135

p 305, troisième tiers, lire AL-ANI (et non \*EL-ANI)

p 307, tableau i), supprimer les sons shingazidja sont encerclés ; tableau ii), supprimer les consonnes arabes sont encadrées

p 312, liste B), manque } ; liste C), manque { ; avant-dernier paragraphe, lire [b] → [b], [d] → [d]

p 335, milieu, lire pers. wai, partie (et non ar. pers. wai)

p 336, deuxième moitié, lire ar. silāh (et non \*sil`h

p 345, dernière ligne, lire ... rien n'oblige à (et non \* ne justifie)

p 349, point 5, lire

la correspondance Ng est i-, hi- selon les lexèmes:

Ng hitswa, 'tête' = st. swa. kichwa

Ng itranda, 'lit' = st. swa. kitanda (ci-dessus)

p 353, deuxième moitié, suppléer l'indication de renvoi, p 138-9

p 357, manque dans toute la page le symbole du 'o' ouvert, [ɔ], ainsi que celui de 'sh' en milieu de page ([ʃɔbo])

p 360, première ligne, suppléer l'indication de renvoi, p 258

p 368, premier tiers, compléter l'indication de renvoi, p 368

p 372, troisième tiers, ajouter (dawl) à côté de ar. *daul*

p 377, premier paragraphe, compléter l'indication de renvoi, p 202

p 379, milieu de page, ajouter un renvoi :

peut-être ce mot est-il venu par le swahili *kibiriti* (cf. p 203); plus bas, à propos des hésitations d'intégration en classe 9 ou 11, ajouter cf. p 204

p 391, milieu de page, compléter l'indication de renvoi, (voir voyelles p 372)

p 392, premier paragraphe, lire ... CVC<sup>1</sup>C<sup>2</sup>V (et non CVC<sup>1</sup>°2V)

p 398, compléter l'indication de renvoi, p 291

p 400, premier paragraphe, lire J-L SIBERTIN-BLANC 1980 (et non 1981)

A ma mère

kapavana haraka na ndro

"le vite est ennemi du bien"

## REMERCIEMENTS

Au seuil de ce travail, je tiens à assurer de ma reconnaissance tous ceux sans qui je n'aurais pu l'accomplir.

A mes amis comoriens tout d'abord, qui m'ont fait aimer leur pays et leur culture, et qui ont eu la patience de m'instruire de leur langue; il serait trop long de les nommer tous, mais je tiens à citer Mme Swafia NIRGANI, et Mr Abudu ASSOUNANI, Abdillahi NZE MWINYI, Ali SAADI, ainsi qu' AHMED-CHANANGA, sans oublier Ali SALEH qui le premier m'a fait connaître l'existence des Comores;

je tiens aussi à remercier Mme Suzy PLATIEL, qui m'a guidé et soutenu au long d'un parcours parfois difficile; il est peu de dire que ce travail lui doit beaucoup;

toute ma reconnaissance aussi à Mr Pierre VERIN qui a su tendre la main au naufragé et le ramener à terre, enfin à ma femme, pour sa patience et son aide, et à tous mes amis, pour m'avoir parlé d'autre chose pendant ces années.

Pour la partie bantu, j'ai bénéficié des conseils et suggestions souvent éclairantes de Gérard PHILIPPSON, qu'il en soit remercié !

Antoine LONNET a eu la gentillesse de relire avec moi les termes d'origine arabe du lexique et de m'éviter ainsi de trop nombreuses erreurs.

Bien entendu, toutes les inexactitudes que l'on trouverait sont miennes.

### *Mouvement général de la thèse*

La thèse est construite sur le postulat que les particularités et les irrégularités du comorien ne peuvent s'expliquer qu'en prenant en compte la dualité fondamentale qui le constitue (l'influence arabe dans un contexte culturel particulier sur un fonds bantu): certains comportements ne se vérifient que pour le lexique d'origine bantu, d'autres, que pour les emprunts orientaux, etc...

La thèse est présentée de façon à faire ressortir cette dualité: dans chaque partie, est distingué ce qui est imputable au bantu (d'après GUTHRIE) de ce qui est imputable à l'apport arabe. C'est ce que j'appelle 'étude différentielle'.

Chaque fois que cela est possible, le lexique fait apparaître les étymons bantu et les étymologies pour les emprunts.

L'idée n'a rien de neuf; l'originalité réside peut-être dans la tentative d'application systématique; je crois que prendre en compte la dualité bantu-arabe est nécessaire pour expliciter au moins certains aspects de la langue et paraît susceptible d'éclairer la manière dont la langue s'est constituée.

Plan général

LE SHINGAZIDJA ou GRAND-COMORIEN  
UNE LANGUE BANTU  
SOUS INFLUENCE ARABE

vol I

I INTRODUCTION GENERALE : MILIEU NATUREL ET HUMAIN 1 - 22

II ETUDE LINGUISTIQUE 22 - 400

Introduction 23 - 80

Méthodologie - approche différentielle 23 - 28

Les composants de la langue (1) : le bantu : 29 - 41

Généralités

La critériologie bantu

Modalité d'application au shingazidja

Les composants de la langue (2) : l'apport exogène 42 - 62

Spécificité et importance de l'apport arabe

L'arabe, comme source ou média de passage des emprunts orientaux

L'hypothèse historique : l'arabe comme langue véhiculaire ?

Le problème des sources arabes : le choix de l'arabe classique

Remarques sur les étymologies arabes

Autres langues sources d'emprunts

Variété décrite 63 - 69

Niveau état de langue; phonostyle

Composition du corpus

Norme ou forme de référence

Conventions de notation et divers 70 - 79

1) Shingazidja

1-1) Signes

a) symboles orthographiques

b) élocution délibérée

1-2) Mode de citation des unités

a) Mode de citation des lexèmes

b) Mode de citation des constituants syntaxiques

2) Arabe : système de transcription

Conventions diverses

## I

**La constitution des mots 81 - 231**

Programme 81- 84

*Plan du discours : constituants syntaxiques*

*Plan de la langue : monèmes*

**A - INVENTAIRE DES PARTIES DU DISCOURS OU  
CONSTITUANTS SYNTAXIQUES**

85 - 92

1) Modèle théorique

2) Inventaire général des classes de constituants syntaxiques

**A Termes variables**

1 *Nominaux*

2 *Verbaux*

3 *Nominaux-verbaux*

**B Termes invariables**

4 *Conjonctions*

5 *Circonstants*

6 *Particules*

7 *Interjections*

8 *Idéophones*

9 *Quasi-nominaux*

Remarques

## I-2 Emprunts naturalisés

### 1) Intégration par assimilation à des thèmes

#### 1-1) Intégration morphologique

- a) *Affectation de préfixes sur des bases sémantiques*
- b) *Intégration par analogie phonétique*
- c) *Conjonction des facteurs sémantique et analogique*
- d) *Remotivation des classes*

*Remarque : fluctuations significatives ou non significatives*

#### 1-2) Des thèmes, des thèmes

### 2) Intégration par assimilation à des racines

#### a) existence d'un support lexical attesté

##### a-1) formations à un degré d'extension

- i) applicatif
- ii) intensif
- iii) causatif
- iv) réciproque
- v) passif
- vi) moyen-neutre

##### a-2) formations à deux degrés d'extension

- i) combinaison facultative
- ii) combinaison obligatoire

#### b) absence de support lexical attesté

- i) co-existence de plusieurs formes étendues
- ii) une seule forme étendue

#### c) conséquences lexicologiques

II Aspect lexicologique

224 - 231

rapprochements à base étymologique

I I

L'organisation des sons 232 - 400

DEFINITION DES CONTEXTES D'APPARITION DES REALISATIONS 234 - 236

I - CONSONNES

A INVENTAIRE GENERALE 237 - 248

a) consonnes simples

b) articulations complexes:

- 1) prénasalisées;
- 2) vélarisées
- 3) consonnes hors système

remarques à propos de l'inventaire

B APPROCHE DIFFERENTIELLE 249 - 338

B - 1 FONDS BANTU 251 - 302

1) correspondances régulières BC → Ng

Illustrations

- 1) \*kV<sup>avant</sup> : Ng fricative palatale
- 2) \*CV<sup>ternées</sup> → Ng fricatives (affriquées)
- 3) \*C → Ng occlusives ou continues,
  - 3-a) \*V<sup>ternées</sup> C<sup>occlus.</sup> → Ng occlusives
    - i) éléments morphologiques
    - ii) éléments lexicaux
  - 3-b) \*VC<sup>occlus</sup>V → Ng continues (occlusives)
    - i) position interne lexicale
    - ii) position initiale lexicale (intervocalique)
      - nominaux dépendants
      - nominaux indépendants & verbes
  - 3-c) \*#C<sup>occlus</sup> → Ng continues
    - i) préfixes
    - ii) invariables
    - iii) série Ng ha < BC \*ka
- 4) affriquées BC → Ng affriquées

5) \*y + Ng Ø

6) \*NC + Ng prénasalisées

i) position interne lexicale

ii) position initiale de mot

7) \*N + Ng nasales

8) vélarisées shingazidja de fonds bantu

Remarques sur l'établissement des correspondances

**résultats**

a) distribution des reflets bantu

b) inventaire

B - 2 APPORT ARABE

302 -338

1) inventaire phonétique de l'arabe

2) réalisations des consonnes arabes en shingazidja

a) correspondances phonétiques régulières

b) Tableau

3) Intégration de l'apport arabe

3-1) Comparaison entre les reflets bantu et les réalisations de consonnes arabes

a) tableau des reflets bantu

i) segments lexicaux

ii) segments non-lexicaux

b) tableau d'ensemble (reflets bantu & réalisations des consonnes arabes)

i) articulations simples

ii) articulations complexes

3-2) Analyse des modalités d'intégration des réalisations 'arabes'

A) réalisations spécifiques (réalisations 'arabes')

i) réalisations maintenant leur spécificité

ii) réalisations spécifiques connaissant toujours une variante

**B) phonèmes arabes coïncidant avec des reflète bantou : phonèmes  
'COMMUNS'**

- α) 'phonologisation' des 'variantes combinatoires'
- γ) arrêt d'évolutions en cours : palatales et affriquées /dj/, /dz/ (/tsh/, /ts/)
- γ) /t/ : introduction probable d'un phonème
  - i) relation /t/, /ts/, /tr/
  - ii) dissymétrie sourdes + sonores

**4) Conséquences de l'influence arabe**

**De quelques correspondances remarquables**

**a) correspondances arabe → shingazidja montrant une adaptation des emprunts**

- 1) consonnes se conformant à l'alternance consonantique
- 2) continues initiales h < k, q
- 3) séquences étymologiques nasale + consonne

**b) correspondances atypiques**

- 1) vélarisées Ng < consonnes arabes non emphatiques
- 2) divers

**C RECAPITULATION**

**CONSONNES SANS SOURCE RECONNUE**

339 - 347

**1) Récapitulation**

- i) consonnes simples
- ii) séries complexes: prénasalisées, vélarisées
- iii) hors-système

**2) Consonnes sans source reconnue**

- a) tsh (tshw); ntsh
- b) prénasalisées initiales: np, ntr
- c) prénasalisées vélarisées: nbw, npbw, nk, ngw

**ANNEXES**

347 - 355

**1) Comparatisme bantou**

**2) Emprunts aux langues européennes**

1) le système vocalique de l'arabe2) intégration des réalisations arabes dans le système bantu

- a) voyelles isotimbres brèves ou longues → voyelles isotimbres
- b) successions ayant donné des voyelles intermédiaires ou des successions vocaliques

3) Insertion de voyelles épanthétiques

- a) conditionnement en position finale
- b) conditionnement en position interne

4) conséquences de l'influence arabe sur le système vocalique

- a) modification de la fréquence
- b) modification de la distribution

## III SYLLABE ET STRUCTURE SYLLABIQUE

381 - 394

## A PRESENTATION GENERALE

381 - 383

1) syllabe de type /CV/2) syllabe de type /V/3) syllabe de type -m-Autres types de syllabes

## B ETUDE DIFFERENTIELLE

## B - 1 FONDS BANTU

384 - 388

1) structure syllabique du bantu commun2) structure syllabique des mots de fonds bantu

- a) Ng /CV/ & /V/ < \*CV, \*CVV
  - a-1) mots composés de syllabes \*CV réalisées [CV]
  - a-2) mots comportant une syllabe \*CV réalisée [V] ou [CV]
  - a-3) mots comportant une syllabe \*CV toujours réalisée [V]
- b) \*m → Ng /m/

## II VOYELLES

355 - 380

### A PRESENTATION GENERALE

355 - 364

#### Inventaire phonologique général

##### Réalisations

##### 1) Généralités : aperture

##### 2) longueur

- i) allongement stylistique à valeur expressive
- ii) réalisation de /V/ en variante
- iii) Réalisation de [CV:] en variante

##### 3) nasalisation

- i) interjection négative an-an
- ii) nasalisation conditionnée

##### 4) variation de timbres

##### 5) arrondissement : réalisation [i:]

##### 6) Successions vocaliques

- i) dévocalisation
- ii) élision
- iii) coalescence

### B ETUDE DIFFERENTIELLE

364 - 380

#### B - 1 FONDS BANTU

364 - 369

##### 1) inventaire vocalique du BC

##### 2) correspondances BC → Ng

##### 3) distribution

##### a) position initiale

- i) voyelles appartenant à des éléments grammaticaux
- i) voyelles appartenant à des éléments lexicaux

##### b) position interne et finale

1) système syllabique de l'arabe

2) intégration des syllabes arabes en shingazidia

a) adaptation

i) syllabes arabes ouvertes CV et CV<sup>h</sup> → Ng /CV/ = [CV]

ii) syllabes arabes fermées CVC et CV<sup>h</sup>C, CVCC et CVC-CV → Ng /CV-CV/

iii) syllabes arabes CV → Ng /V/

b) maintien de mots comportant des syllabes fermées

Conclusion générale sur la syllabe

## INTRODUCTION GENERALE

*Où l'on présente l'île et ses habitants*

Le professeur VERIN remarquait un jour que la plupart des travaux universitaires consacrés aux Comores commençaient par la localisation géographique de l'archipel; nous ne saurions nous écarter de cette tradition débutante.

Après avoir situé la Grande-Comore ou Ngazidja et précisé les conditions écologiques qui y prévalent, nous essaierons de présenter brièvement la pré-histoire ou les origines du peuplement en ce qu'ils éclairent le fond culturel.

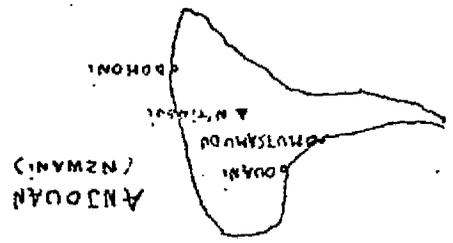
#### Situation géographique et écologique

Cette présentation doit beaucoup à l'ouvrage clair et documenté de BATTISTINI & VERIN 1984, *Géographie des Comores*, réédité en 1986 sous le titre *Les Comores*, auquel nous renvoyons pour les détails et des illustrations très parlantes.

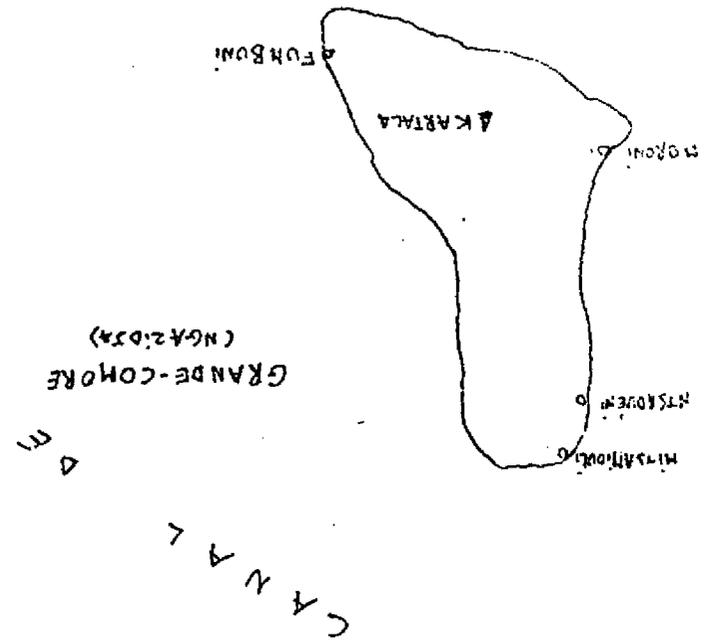
Donc, il était une fois une île de l'Océan Indien, située dans le canal de Mozambique, à mi-distance entre la côte est-africaine et Madagascar, et qu'on appelait Grande-Comore ou Ngazidja; cette île, la plus grande de l'archipel des Comores, couvre 1024 km<sup>2</sup>, et est actuellement peuplée de 189128 h., soit une densité de 185 h. au km<sup>2</sup>

Seules sont densément occupées la bande littorale et les hauteurs avoisinantes; l'intérieur, montagneux, longtemps difficile d'accès, avec le redoutable volcan Karthala et les massifs de La Grille et du Mbadjini qui le prolongent est peu propice à l'installation humaine (voir carte 1).

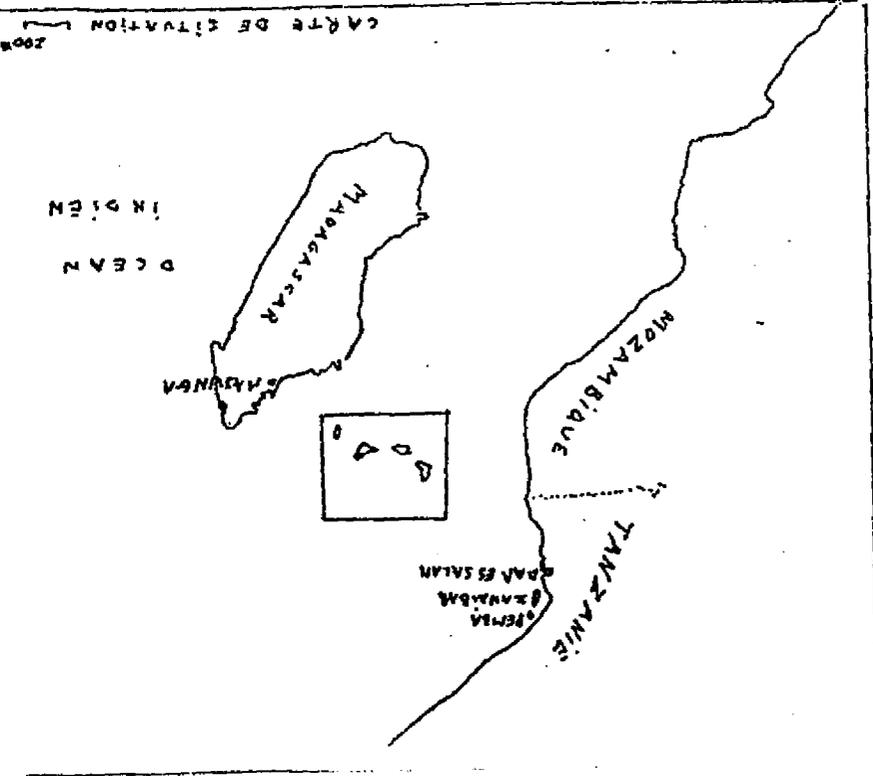
MAYOTTE (MAHORE)



M O Z A M B I Q U E

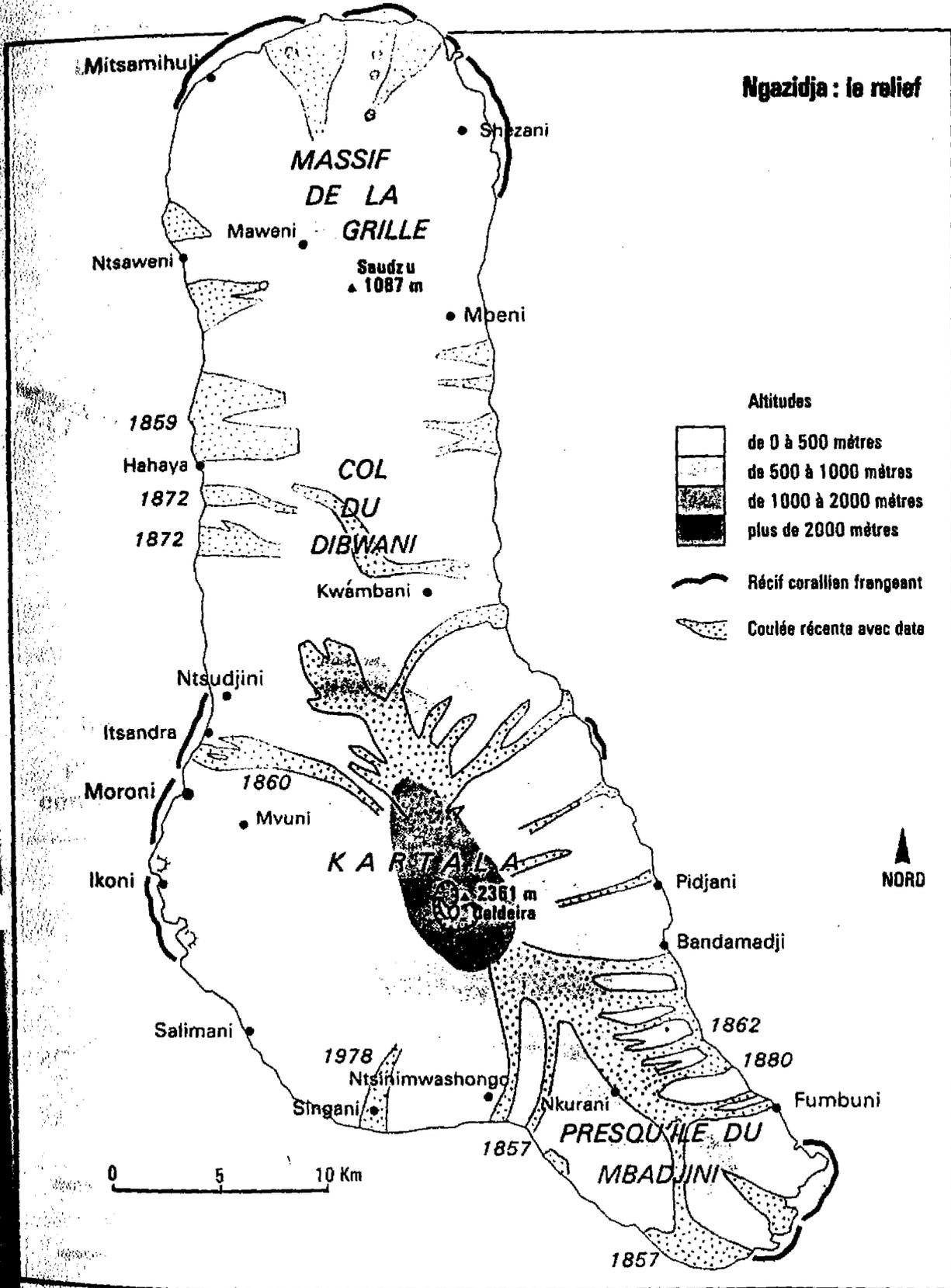


2-22



CARTE DE SITUATION 200M

### Ngazidja : le relief



(Vern & Battistini, op. cit., 24)

CAATE n° 2

L'île connaît peu de plaines, la bande côtière étant très peu large; les seules étendues planes sont en altitude.

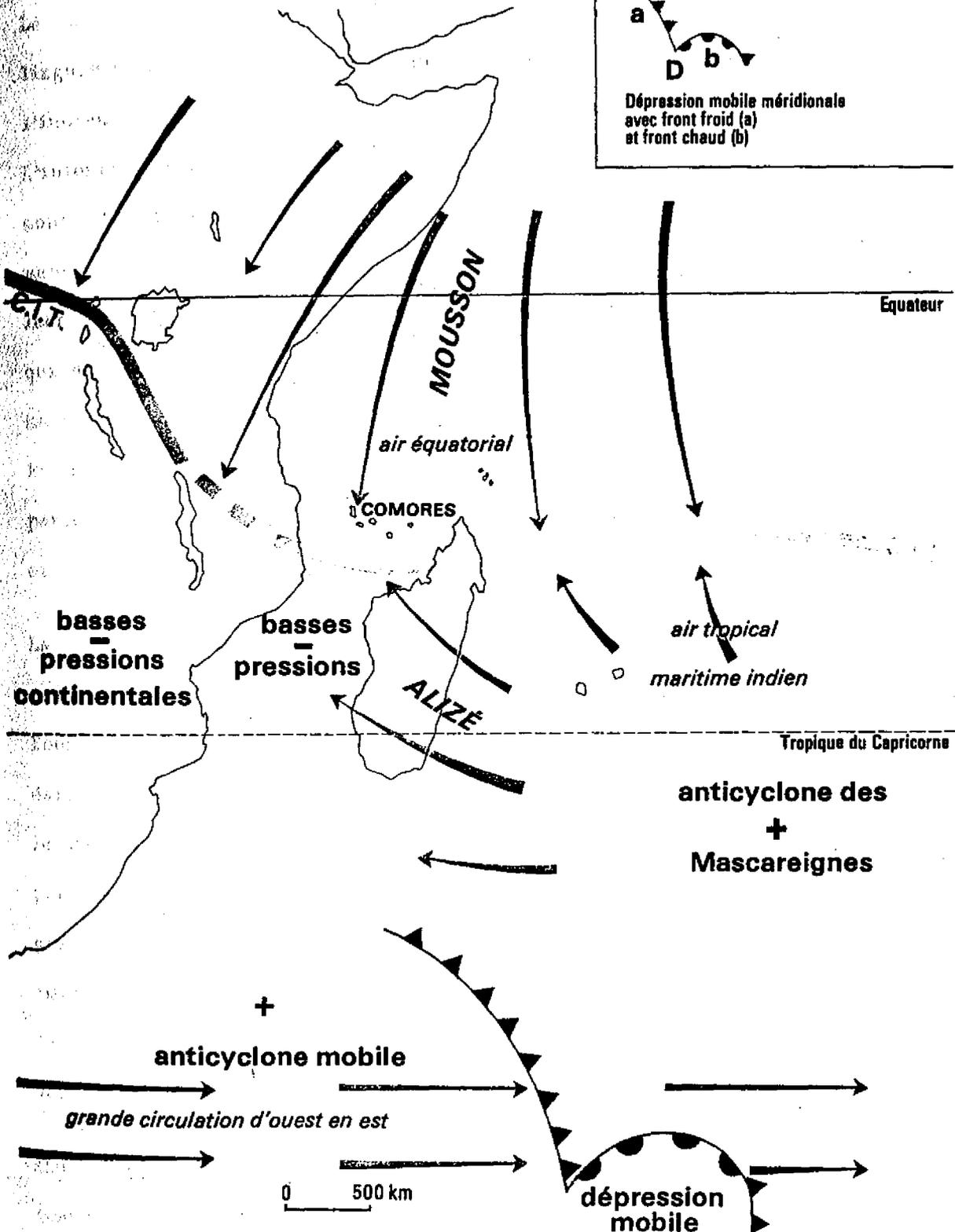
La latitude (entre l'équateur et le tropique du Capricorne), l'insularité et le relief expliquent un climat tropical, chaud et humide, aux chutes de pluie abondantes et fréquentes; l'île est soumise au régime des moussons, avec une saison chaude d'octobre/novembre à avril (*hasihazi*, saison des pluies, mousson qui souffle du Nord-Est, de l'Inde vers la côte africaine), et une saison 'fraîche' et plus sèche, de mai à septembre/octobre (*husi*, mousson soufflant du Sud, de l'Afrique orientale vers l'Inde). A ces moussons, il convient d'ajouter les alizés et contre-alizés, pour dresser un tableau complet des vents qui placent les Comores au carrefour de diverses routes (voir carte 2).

L'élévation rapide des altitudes entraîne une différenciation végétale entre "les bas", domaine des cocoteraies, et "les hauts", où subsiste encore la forêt, en cours de défrichement; à cette différenciation écologique s'est ajoutée une spécialisation culturelle: les cultures de rente, destinées à l'exportation, occupent les zones littorales (vanille, ylang-ylang, girofle, cocoteraies de plantations souvent abandonnées, etc) et les cultures vivrières occupent les abords immédiats des villages ou sont repoussées sur les hauteurs (cocotiers et bananiers, à la place si importante dans la civilisation comorienne, entre lesquels sont cultivés manioc, maïs, embrevades, songes ou taros, etc); sur les pentes les plus élevées, se trouvent des pâturages.

La pêche est restée longtemps une activité traditionnelle, réservée à certaines communautés.

On n'a pas découvert sur la Grande-Comore (jusqu'à présent) de ressource minérale.

### Les grands centres d'action météorologiques: situation en saison chaude



(Vern & BaHötter, op. cit. 52) (CARTE n° 3)

La population est répartie dans des agglomérations relativement isolées, le plus souvent côtières ou installées à proximité immédiate de la côte; les villageois tirent leur subsistance de l'agriculture et de la mer.

Ces communautés villageoises de la Grande-Comore ont longtemps vécu dans une économie peu monétarisée, en autarcie plus ou moins totale, permise par la luxuriance du climat, la fertilité des sols, la variété des espèces naturelles, le poisson, et entretenue par la difficulté des communications due au relief.

Les trois autres îles qui, avec Ngazidja, forment l'archipel des Comores, Anjouan ou Nzuani, Mayotte ou Maore, Mohéli ou Mwali, partagent, dans l'ensemble, les mêmes caractéristiques écologiques.

#### La situation linguistique de l'archipel

Deux familles de langues sont représentées aux Comores parmi les langues 'autochtones': des langues ou parlers bantu, sur les quatre îles, et un parler malgache à Mayotte.

Sur le plan de la recherche linguistique, l'archipel est resté, entre le début du siècle et les années 1970, en friche - peu d'études lui ont été consacrées après avoir pourtant retenu l'attention de swahilisants au début du siècle (SACLEUX en particulier) et l'on commence à peine à mieux connaître ces langues.

Le malgache de Mayotte a été décrit par GUEUNIER, avec la publication récente d'un lexique très complet (GUEUNIER 1976 & 1986).

Parmi les langues ou parlers bantu, le shimaore, parlé à Mayotte, fait l'objet d'une thèse soutenue en 1981 (ROMBI); le shindzuani, parlé à Anjouan est en cours de description (AHMED-CHAMANGA); quasiment rien n'a été fait sur le

mohélien, et nous espérons que le grand-comorien ou shingazidja sera mieux connu avec notre travail.

Le caractère récent et partiel de la documentation explique que la situation linguistique de l'archipel et les relations des parlers entre eux soient pour l'instant peu claires.

L'intercompréhension est réalisée sans nul doute entre maore et anjouanais, qui sont très proches; il n'en est pas de même entre shingazidja et maore/anjouanais: cela dépend des locuteurs, de leur faculté d'adaptation et de leur attitude linguistique (le particularisme insulaire est très fort); pour ce qui est du mohélien, sur lequel nous n'avons pas d'information, il semble qu'il se situe à mi-chemin entre le grand-comorien et l'anjouanais.

En dépit d'une indéniable unité sur le plan culturel, les différences entre les parlers de l'archipel sont telles qu'elles ne permettent pas une étude globale et précise.

La question du statut de ces parlers les uns par rapport aux autres reste entière: dialectes d'une même langue ou langues distinctes mais apparentées ?

Nous n'entrerons dans ce débat; nous étudions le grand-comorien et lui seul dans sa spécificité, et nous utiliserons le terme de 'langue' ou de parler pour le désigner. L'objet de ce travail est avant tout de décrire la langue parlée par les Grand-Comoriens et le fait que la Grande-Comore soit l'île la plus grande et la plus peuplée de l'archipel est fortuit; chaque parler a la même importance sur le plan linguistique, quel que soit le nombre de ses locuteurs.

Ainsi, le système orthographique utilisé a été conçu pour pouvoir transcrire les différents parlers, sans en privilégier un (voir plus loin).

Outre ces langues, le français, le malgache et l'arabe sont pratiqués aux Comores: le français est la langue de l'administration, de la vie publique, de

l'école; il ouvre la voie de la promotion sociale; sa diffusion est en rapport avec celle de la scolarisation et de l'émigration - de nombreux Comoriens sont émigrés en France ou à la Réunion; depuis longtemps, des Comoriens ont émigré vers Madagascar, où certains sont installés depuis plusieurs générations; le retour d'un grand nombre d'entre eux en 1976, à la suite des émeutes de Majunga, a favorisé la pénétration en comorien de termes malgaches; l'arabe enfin, outre son rôle dans la formation de la langue et son prestige de langue religieuse, est depuis 1978 langue officielle avec le français, des étudiants sont envoyés dans les pays arabes, un système scolaire utilisant l'arabe comme véhicule a été organisé parallèlement au système français.

#### Origine du peuplement et formation culturelle

Pour une société comme celle des Comores, l'histoire se confond, jusqu'à la période récente, avec celle de son peuplement; la Grande-Comore se différencie sans doute peu du restant de l'archipel.

Les découvertes archéologiques récentes (en attendant d'autres trouvailles ? ) permettent de penser que l'île fut occupée dès les X-XI<sup>è</sup> siècles (fouilles de Mwali et de Maore); il est difficile de déterminer l'origine des populations que ces fouilles appellent, mais il paraît établi qu'elles comprenaient des Musulmans et qu'elles entretenaient des relations au moins commerciales avec la côte sud-arabique (poteries 'sassano-islamiques', BATTISTINI & VERIN 1984 : 51).

La navigation côtière et le régime des vents rendent ces relations maritimes plausibles.

On sait par ailleurs que se sont installés sur les îles, directement ou après escale sur la côte africaine, des 'Shiraziens', marchands négociants originaires de Shiraz, port du golfe persique.

Si l'on a longtemps pu tenir pour acquis que les 'Shiraziens', accompagnés d'esclaves africains, et déjà sans doute dans une phase avancée de métissage avec eux, furent les premiers occupants, cette conception a évolué: ainsi, dans un article ancien, Basil DAVIDSON pouvait-il attribuer aux 'Arabes' la civilisation citadine de la côte est-africaine; la carte ci-jointe, extraite du même texte, illustre les routes suivies:

*"having established trading stations along the coast of India, south-east Asia and southern China, Arabs, from the Persian Gulf States began to enlarge a long existing but hitherto minor interest in African trade. Founded early in the ninth century, trading stations along the East coast soon grew into substantial ports. Some of these, whether on the coast itself or on nearby islands, became fine stone-built cities after A.D 1300"*  
(Basil DAVIDSON 1978 : 9).

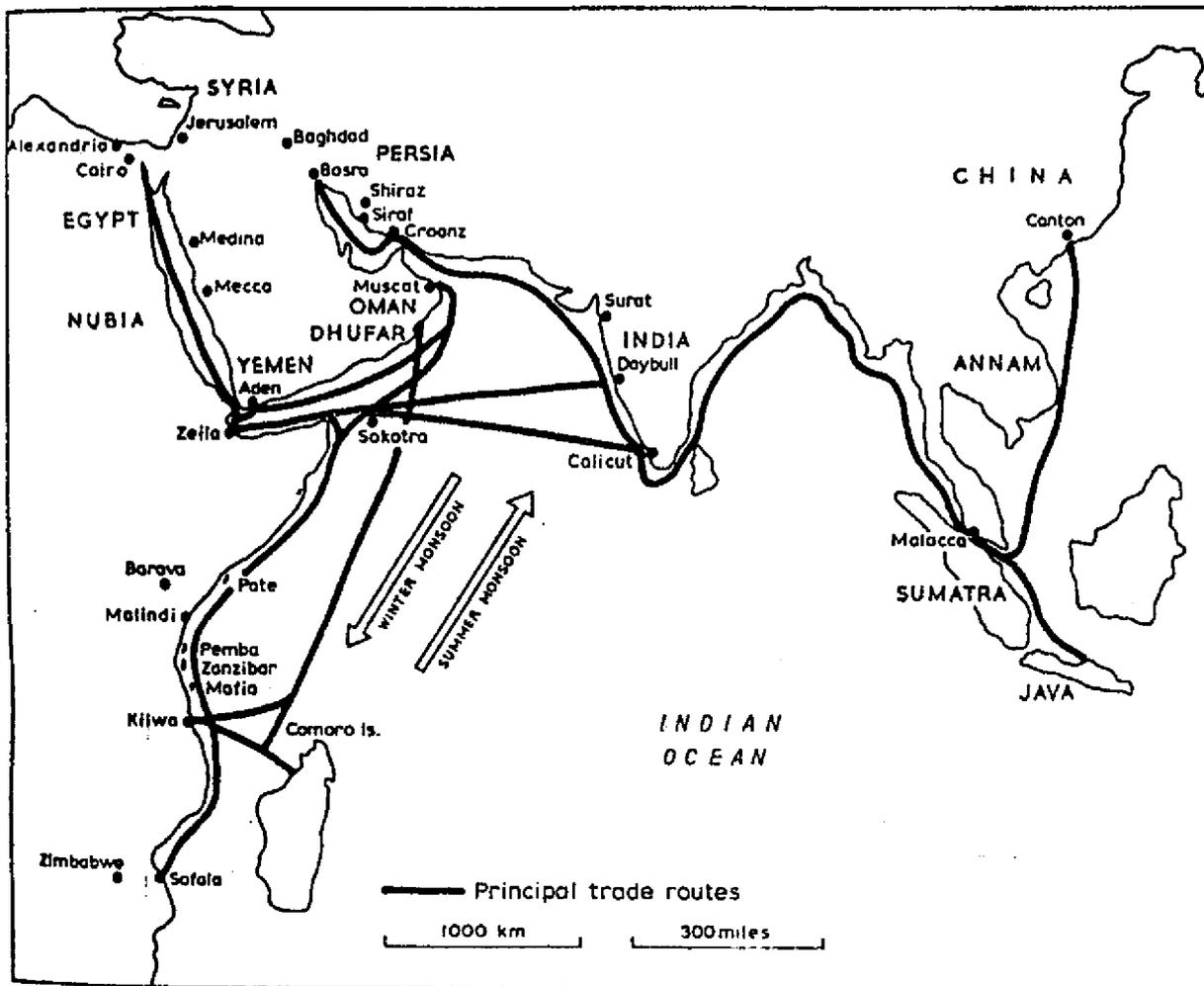
Au vu des acquis archéologiques et linguistiques récents, il paraît tout à fait vraisemblable que ces immigrants trouvèrent sur place des populations connaissant un certain degré de développement, sans doute établies dans des cités, à qui ils se mêlèrent; ces populations étaient alors vraisemblablement à dominante africaine bantu, même si des éléments arabes y étaient aussi présents, et les sites archéologiques relevés doivent leur être imputés.

Cette hypothèse s'appuie sur une interprétation des traditions locales, de la proto-histoire linguistique malgache et sur une vision globale de la pré-histoire de la région depuis l'aube des temps historiques.

• Les traditions et généalogies locales comoriennes référant à la constitution de la société, certes à visée idéologique, racontent l'arrivée sur l'île de princes arabes, et leur alliance avec les filles de chefs locaux, assurant ainsi à leurs descendants une double légitimité, conforme aux deux coutumes, arabe par les hommes, bantu par les femmes (ci-dessous); VERIN, s'appuyant sur MARTIN 1968 a montré le parti qu'on pouvait tirer, en restant prudent, de ces traditions orales: elles réfèrent à une période antérieure à la venue des 'Arabes' (c'est-à-dire des

BACKGROUND TO THE CONTINENT

*Africa in Historical Perspective*



The Indian Ocean and its trade in about A.D. 1200.

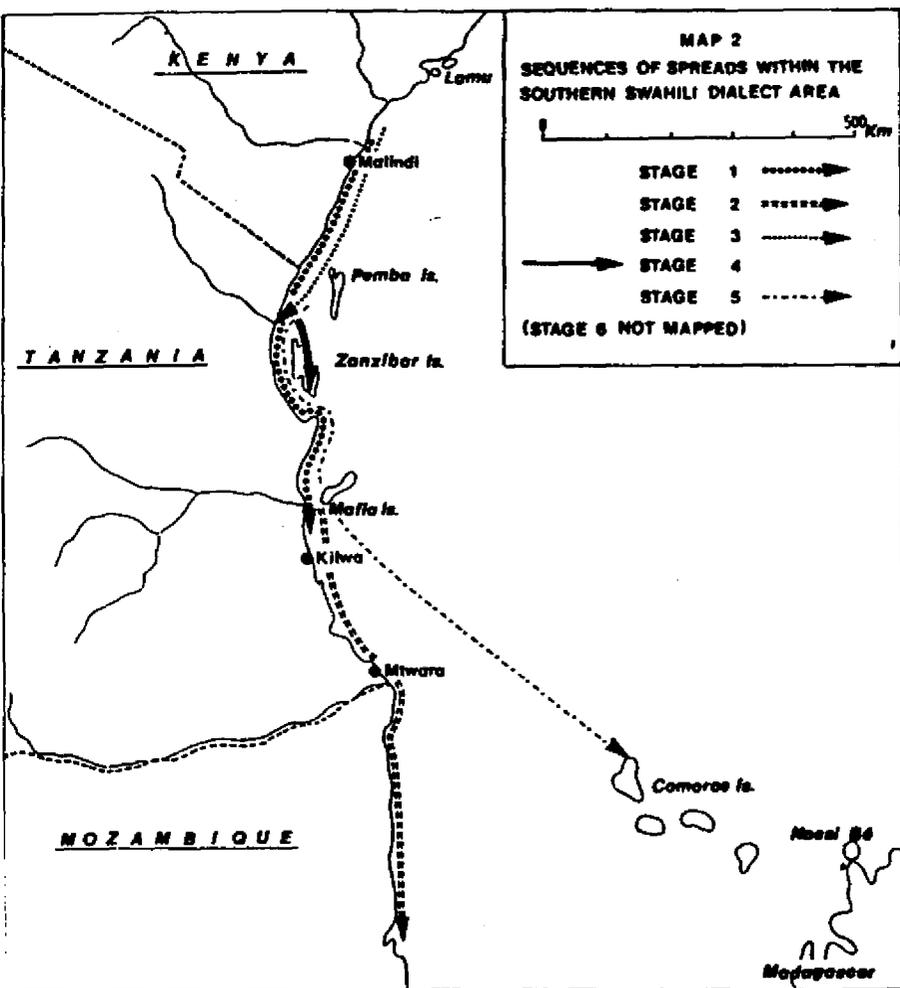
Shiraziens) et à l'islamisation et montrent l'existence de sociétés organisées bantu dès avant la venue des Shiraziens: "Ces *beja* [terme désignant des propriétaires de terres avant l'arrivée des Arabes] n'étaient pas de simples occupants du terroir. Ils s'étaient déjà organisés [en chefferie *mafe*, institution purement africaine] (VERIN 1982 : 193).

• La présence à Madagascar d'un substrat bantu ancien avant l'arrivée des premiers immigrants indonésiens paraît à présent établie: DAHL suppose d'ailleurs que cette population devait appartenir à la même branche de la famille bantu que les occupants des Comores (DAHL, à paraître); SIMON nomme 'actéo-bantu' ('bantu du rivage') ces populations qu'il rattache aux populations que NURSE a appelées 'sabaki' (du nom d'une rivière du Nord Kenya, où se situerait leur lieu de dispersion) (SIMON 1987 )

• Enfin, Mark HORTON avance, dans un exposé récent et novateur ré-interprétant les données archéologiques, qu'une voie maritime ("*the Swahili corridor*") existait dès le début de l'ère chrétienne entre la côte arabe et la côte est-africaine, et qu'elle était prolongée par des courants d'échanges réguliers avec les populations de l'intérieur africain; ces contacts expliqueraient, dès le IX<sup>e</sup> siècle, la diffusion d'éléments culturels et économiques le long de la côte jusqu'en Afrique méridionale: HORTON note "*it is possible to argue that the 'late' iron age in southern Africa developed (...) as a result of the local development of (...) societies who had maritime contact with the North. Along this corridor, cattle and pottery styles were transmitted and perhaps small populations came to settle*" (HORTON 1986 : 18); sur la côte africaine et les îles voisines, ces contacts auraient donné naissance, avant l'arrivée plus considérable d'Arabes shiraziens et la généralisation subséquente de l'islamisation, à une civilisation citadine, de comptoirs commerciaux, à dominante africaine, vraisemblablement païenne mais comportant des éléments musulmans, où se retrouvent sans doute les *beja* et *mafe* cités plus haut.

S'appuyant sur les travaux d'histoire culturelle de C. EHRET, HORTON voit dans ces populations des groupes couchitiques (voire indonésiens, malgré l'absence d'argument archéologique).

Intégrant cette même hypothèse à des recherches de linguistique historique, DEREK-NURSE estime que dès les environs du IX<sup>e</sup> siècle, ces populations sont bantu, soit que les éléments bantu recouvrent les populations



(DEREK. NCASE 1982:227-228)

in view of the wide lack of primary data, we can only speculate about the possible stages of a sequence of development in the SD area. Based on the combined available linguistic and archaeological picture, we might posit (see map 2):

- Stage 1. An early SD community comes south from the northern Sabaki homeland in the ninth century (see fn. 21), possibly attracted by the rich trade network known to have existed along the southern coast at that time. If it had a single dispersal point, this might have been the northern Tanzania/Southern Kenya coast/Pemba/Zanzibar/Bagamoyo area or Kilwa/Songo Mnara/Sanje ya Kati/Mafia.
- Stage 2. Shortly thereafter, or at the same time, the ancestral Mwaní community moves south to the coast and islands of Mozambique.
- Stage 3. Early Comorian speakers come south from the northern Sabaki homeland. They interact with the early SD community/communities along the Tanzanian coast or islands. Since all the SD (except Mwaní and Unguja) and Comorian show the same verbal innovations, the SD communities are likely to have been compact or adjacent. There is

particularly heavy contact between Comorian and Yumba/Jambiani.

- Stage 4. Yumba and Jambiani separate, Yumba maintaining contact with early Comorian. Stages 4 and 6 may be interchangeable.
- Stage 5. The "Comorians" reach the Comoros by the tenth or eleventh century A.D. (see fn. 19), because Al-Idrisi reports boat-using, partly Islamised Comorian-speakers in situ by 1150 A.D.
- Stage 6. Other SD communities (Mtang'ata/Lugha ya zamani, Tumbatu, Pemba, rest of Hadimu-Makunduchi, Unguja, Mafia, Mgao, etc.) spread along the coast and the islands over the next millenium. Their high level of internal similarity suggests the possibility of a relatively recent and rapid spread and/or the maintaining of a high level of contact between their communities.

antérieures, soit que ces populations aient été bantu depuis leur origine: "there is archeological evidence for a common cultural [bantu] complex from approximately Muqdisho to southern Mozambique, also the Comoro Islands and parts of Madagascar. (...) starting in the ninth century A.D. (...). The elements of [this] cultural complex are added gradually (...) to the twelfth century." (DEREK-NURSE 1982 : 240-1). Il poursuit en dressant, à l'aide de données issues de la comparaison des langues, un tableau de l'itinéraire de ces populations, qui vont rejoindre, depuis leur foyer situé sur la côte kenyane, à l'embouchure de la rivière *sabaki* leurs localisations actuelles. Nous reproduisons la carte de DEREK-NURSE et sa légende (on peut regretter toutefois que son étude ne comporte que peu de données comoriennes et ne distingue pas clairement entre les parlers).

Des populations bantu qui connaissaient déjà un développement culturel certain se seraient ainsi trouvées aux Comores à l'arrivée des "Shiraziens".

Cette vaste et pénétrante perspective reprend d'ailleurs une hypothèse formulée parmi d'autres par Nevill CHITTICK:

"It seems to me that there are three solutions to the question of the spread of the Negroid "bantu" in (...) eastern Tanzania and northeastern Mozambique. (...) The second solution is that the Bantu people moved up from the south but were overcome by a cattle-keeping people, perhaps Kushitic, who came from the north and who established themselves as a ruling class over the bantu, while adopting the Bantu language. (...)" (CHITTICK 1975 : 28)

Cette hypothèse reste à être étayée par des arguments archéologiques et linguistiques plus détaillés, en particulier sur les relations entre le comorien et le substrat bantu malgache; elle est à même de fournir une explication aux deux strates qui apparaissent en shingazidja sur le plan linguistique, comme nous le verrons, distinguant un fonds bantu ancien, et de rendre compte de la grande divergence qui existe entre le shingazidja et le swahili, malgré le cours de l'histoire qui les a rapprochés.

Mais cela ne doit pas entraîner une minimisation de l'apport shirazien, qui commence sans doute au XIII<sup>e</sup> et va aller croissant; ces Shiraziens vont se fondre, non sans influencer profondément sur la culture et la langue, avec cette première population bantu.

Si CHARPANTIER écrit "[parmi les Shiraziens] l'élément arabe et peut-être juif était probablement aussi important que l'élément persan" (CHARPANTIER 1977 : 2), il est très généralement accepté que ce terme est "une allusion à l'origine géographique de certaines grandes familles plutôt qu'à leur origine ethnique, arabe plutôt qu'iranienne (ALEXANDRE 1981 : 322), et que les Shiraziens sont majoritairement des Arabes (voir également CHITTICK 1975 : 41).

Ces affirmations sont en fait basées essentiellement sur l'observation linguistique: swahili comme shingazidja comportent un grand nombre de termes d'emprunt 'orientaux', dont la majeure partie sont d'origine arabe ou montrent un passage par l'arabe, comme nous le verrons plus loin.

Une fois les Shiraziens installés à la Grande-Comore, le réseau d'échanges entre la côte sud-arabique et le golfe, d'une part, la côte est-africaine et le Nord de Madagascar, d'autre part, s'intensifia; cela détermina

*"l'apogée de la civilisation chirazienne de Zanzibar: c'est dans ce milieu avant tout commerçant vivant en symbiose avec les zandj, leurs enfants métissés et leurs esclaves, sensible à des oppositions politiques violentes et ostracisantes ainsi qu'à des conflits religieux graves que vers le XIV<sup>e</sup> siècle la civilisation chirazienne de Zanzibar connaîtra son apogée, nourrissant alors des relations constantes avec l'archipel des Comores et la côte septentrionale de Madagascar (...)" (CHARPANTIER 1977 : 3).*

On trouve là sans nul doute, dans cette culture shirazienne, tournée vers ses origines orientales, dans le souvenir de cet âge d'or une fois cette période révolue, le fondement du prestige de l'arabité aux Comores, qui rejaillit de si forte façon sur les attitudes linguistiques (voir paragraphe suivant).

Le même auteur continue en évoquant

*une voie de communication directe entre Zanzibar et l'Inde qui passe par les Comores, le Nord de Madagascar, les Seychelles et les Maldives"*

faisant allusion aux *wadibili* des chroniques swahili (du nom sans doute de la ville indienne de *Dabol* ou *Daybul*, signalé entre autres par MOLLAT 1974 : 57), dont le rôle fut de moindre importance.

L'arrivée des Européens - la découverte de la route du Cap par Vasco de Gama date de 1497-98 et la brève installation aux Comores des Portugais, de 1500-1505 - ne modifia pas, dans un premier temps, les circuits commerciaux traditionnels autour de l'Océan Indien Occidental ni les rapports de force: ainsi que l'écrit A. TOUSSAINT, *"la navigation continua de suivre pendant longtemps les routes périphériques qu'elle suivait depuis des siècles"* (TOUSSAINT 1960 : 11)

La reconquête arabe, à partir de la chute de l'Ormuz portugais (1622), qui devait culminer avec le transfert à Zanzibar du sultanat d'Oman en 1841, fut d'une autre importance pour l'archipel. La Grande-Comore est alors morcelée en divers sultanats, sans qu'un pouvoir unique ne s'établisse sur l'île; cette faiblesse facilita l'intervention du sultan de Zanzibar, alors même que l'île n'était pas sous sa dépendance directe; de là date sans doute la pression sur la langue du swahili, lui-même alors soumis à une forte influence arabe; le swahili, en graphie arabe, devint la langue officielle à la Grande-Comore et était le véhicule des traités politiques ou commerciaux (voir SACLEUX 1906 : IV).

La fin de cette période est marquée par les incursions dévastatrices des pirates malgaches, auxquelles mirent un terme le protectorat puis la

colonisation françaises (dès 1886) (sur cette période, on se reportera à l'oeuvre sans doute exhaustive de MARTIN 1983).

La situation culturelle actuelle, résultat de l'histoire

La société grand-comorienne actuelle et ses modes de représentation sont à l'image de l'histoire de ce peuplement métissé: ALEXANDRE a parlé de "pot-pourri ethnique", associant des *zandj* et des éléments essentiellement arabes, mais aussi persans, chinois, indiens, etc (ALEXANDRE 1981 b : 322).

Sur le plan culturel, ces deux pôles sont perceptibles dans presque chaque aspect de la vie sociale; nous en donnons quelques exemples, en raison de leur forte répercussion linguistique:

• l'Islam ou l'arabité :

Ces deux termes, d'un point de vue comorien, sont sinon synonymes, du moins indissociables: qui dit arabe dit musulman, et les Musulmans par excellence sont les Arabes ...

L'ensemble de la population se revendique de l'Islam, à tel point qu'on peut reprendre l'affirmation du président Ali SOILIH, sans doute dite ironiquement mais au demeurant incontestable, *kavvatsi mkomor utso isilamu, kavvatsi !*, "de Comorien qui ne soit musulman, il n'y en a point !".

Ngazidja appartient à l'Islam sunnite de rite shaféite, utilisant comme source du droit le *Minhadj at-Talibin* (voir les commentaires et recueils de jurisprudence de Paul GUY); des confréries sont actives (*shadhulia*, *kadiriya*, etc); BOULINIER a rappelé qu'il ne faut pas négliger entièrement "une empreinte *shi'ite*", perceptible selon lui dans certains thèmes culturels, notamment l'importance de l'eau, qu'il met en relation avec l'origine iranienne de certains immigrants (BOULINIER 1976 : 57).

La culture musulmane est acquise à l'école coranique, dont tous les enfants grand-comoriens fréquentent le premier niveau au moins (paa la shio, école du livre, où la lecture et l'écriture arabes sont enseignées à l'aide du kurwasa, sorte d'abécédaire, composé d'une sélection de sourates coraniques) (voir DE BRYE 1985, sur l'importance de l'école coranique dans l'insertion sociale de l'enfant comorien). Cela explique la diffusion de la graphie arabe spontanée du comorien.

L'Islam est omni-présent, il représente la légitimation nécessaire de toute chose, sans quoi rien n'est valable; en raison de l'identification entre arabe et Islam, tout ce qui est ou paraît arabe est hautement valorisé; ainsi l'aristocratie a toujours recherché l'alliance matrimoniale avec des Arabes d'Arabie, "phénomène qui renforçait, de son poids propre, la stratification de la société et le primat de son caractère arabe" (CHARPANTIER op. cit., 5).

Mais l'Islam comorien, comme nous le verrons, a su s'ouvrir pour intégrer d'autres valeurs et les reprendre à son compte.

• les traditions africaines

La part africaine de la culture ne peut être aussi aisément isolée que l'aspect arabe ou islamique, mais on lui rapportera par hypothèse les coutumes non liées à l'Islam (même si elles sont actuellement intégrées à l'Islam comorien):

\* la hiérarchie sociale est fondée sur le degré d'accomplissement du grand mariage *ada*, que tout homme doit faire pour parvenir au statut d'homme de bien (voir sous *mdji* et *ada* dans le lexique); les dépenses nécessaires à l'accomplissement du *ada*, croissantes, sont d'ailleurs une des raisons majeures de l'émigration des Grand-comoriens.

\* le rôle prépondérant des femmes est apparent dans la transmission utérine d'un 'matrimoine' incessible (le *manyahule*), comprenant souvent une maison, qui

sera la demeure du couple, ainsi que dans la succession par les femmes dans les clans royaux;

x l'importance symbolique dans l'identité grand-comorienne de ce qui touche à la terre et à l'élevage du bétail, surtout des boeufs, qui représentent la richesse et constituent un élément essentiel dans toutes les cérémonies traditionnelles (peut-être est-ce là une piste vers les ancêtres couchitiques ? )

Il est fascinant de constater les analogies, sur le plan du mariage et de la place de la femme, entre les coutumes comoriennes que nous présentons comme d'origine africaine et *"les tendances matrilineaires"* décrites pour l'Arabie anté-islamique par Fatima MERNISSI, où *"l'enfant appartient au groupe de la mère (...) et le cadre géographique du mariage est uxorioca."* (Fatima MERNISSI 1983 : 73); utopie, coïncidence ou convergence significative ?

L'apprentissage des coutumes se fait chaque jour, par l'immersion des jeunes dans une vie familiale et sociale qui en est imprégnée: dès cinq - six ans, le jeune garçon sera intégré à un groupe d'enfants de son âge, qui a ses lieux de réunion, ses règles, à qui sont dévolues certaines tâches et qu'il quittera pour rejoindre le groupe supérieur, etc.

• une synthèse culturelle:

Mais la société comorienne n'est pas duale, elle a su mêler ces deux références pour constituer une synthèse culturelle originale: si, de par sa nature totalisante, aux Comores comme ailleurs, l'Islam tend à organiser la vie sociale dans toutes ses dimensions, il a su s'adapter aux conditions locales en incorporant de nombreuses pratiques et croyances 'para-musulmanes', remontant sans doute aux traditions africaines, qui se voient ainsi légitimées et valorisées.

Nous avons déjà signalé que les traditions historiques mêlaient les deux aristocraties, faisant épouser aux princes arabes les filles des chefs locaux,

caractéristiques du pouvoir selon la tradition africaine: cette attitude a perduré, somme toute, dans l'aristocratie.

La reprise à son compte par l'Islam des coutumes apparaît aussi dans la structuration sociale découlant du *ada*: l'accomplissement du *ada* permet non seulement au 'grand marié' de bénéficier à son tour des morceaux de choix dans les festins et de participer pleinement à la vie du village mais il ouvre l'accès aux privilèges des notables religieux, concrétisés notamment par une place éminente à la mosquée; le sommet de la hiérarchie est atteint par celui qui a accompli toutes ses obligations coutumières, mais aussi par celui qui a fait le pèlerinage ou qui est considéré comme très éclairé dans la religion; tels Obélix, les descendants du Prophète, les *sharifu*, peuvent se dispenser des obligations du *ada* sans déroger, eu égard à leur origine bénie : les deux façons de s'élever ou d'occuper des places importantes dans la hiérarchie sociale, la voie traditionnelle et la voie religieuse, se rejoignent.

De même, la divination et la magie, qui utilisent des grimoires en arabe, font, au moins en apparence, intervenir l'Islam, à travers des versets coraniques, le choix de certaines prières, etc

Cela apparaît aussi dans la coloration musulmane donnée à une pratique aussi peu conforme à l'Islam que le *manyahule* ('matrimoine'), avec une certaine difficulté toutefois, ainsi qu'en témoigne la discussion d'un arrêt du Grand Cadi de la Grande-Comore, rapporté par Jean MAS d'après Paul GUY (MAS 1979 : 103-179).

Cela aboutit à une société qui, jusqu'à l'indépendance, apparaît à de nombreux observateurs comme conformiste, immobile, figée dans des habitudes et des pesanteurs, dues entre autres à cette double légitimation sociale où les pratiques coutumières, recevant la sanction de la religion, deviennent

intouchables et paralysantes, où tout acte doit pouvoir recevoir une double approbation, coutumière et religieuse, où seuls ont voix au chapitre les anciens qui ont montré, par leur vie et en payant de leurs personnes, leur respect des coutumes et où toute innovation, on s'en doute, pose problème.

Ce 'double héritage' est une dimension d'emblée évidente dans la langue: le shingazidja montre une influence arabe importante, sur tous les plans, quel que soit le niveau de langue, avec toutefois des niveaux plus ou moins arabisés (nous parlerons de 'phonostyles'); le choix du niveau de langue peut ainsi servir de facteur d'identification sociale.

#### La période actuelle et ses effets sur la langue

L'indépendance (1975) fut suivie d'une période 'révolutionnaire' agitée, controversée (1975-78), qui allait jeter les Comores dans le monde moderne en provoquant, sous l'impulsion d'Ali SOILIH I une révolution culturelle et mentale dont les Comores vivent toujours les conséquences : dans le projet de développement conçu par Ali SOILIH I, la langue tenait un rôle fondamental, et c'est un des domaines où son action, finalement, reste sensible.

Ali SOILIH I voulut unifier le comorien, en faire la langue de l'état, et promouvoir la culture comorienne: lui-même orateur remarquable, il utilisa le shingazidja pour exprimer toutes choses qui étaient habituellement dites en français, créant ou popularisant des néologismes ou des emprunts (à l'arabe comme au français ou au swahili), ré-employant des termes anciens dans de nouvelles significations, introduisant parfois des termes provenant des autres parlers de l'archipel, montrant, par son exemple, que le shingazidja était apte à rendre la vie moderne (les discours d'Ali SOILIH I doivent faire l'objet d'une publication, LAFON - VERIN, à paraître; on trouvera l'histoire de cette période,

ainsi que la traduction d'un discours d'Ali SOILHI, dans Emmanuel VERIN, sous presse).

Ali SOILHI stimula la collecte des traditions orales, en créant un service de collecte et d'enregistrement et en les faisant diffuser à la radio (certains des textes recueillis alors ont malheureusement été perdus ou détruits). Cette impulsion décisive devait plus ou moins se poursuivre après sa chute.

À son initiative, fut menée une campagne d'alphabétisation en caractères latins, qui se répandit en peu de temps dans toutes les couches de la population, malgré les réticences des tenants des caractères arabes (sur ce point, LAFON sous presse); elle fut mise en pratique dans certains secteurs administratifs; après son renversement, la graphie latine et plus généralement l'usage officiel du comorien furent interrompus voire interdits, identifiés qu'ils étaient au régime d'Ali SOILHI; mais cette première tentative reste la référence obligatoire de toute instrumentalisation de la langue.

On ne s'étonnera donc pas, notre enquête ayant été effectuée pendant cette période, de trouver dans le lexique notamment des termes qui ont été introduits par Ali SOILHI et qui sont depuis passés dans la langue.

\*

\*

\*

L'aspect culturel essentiel pour notre propos est la prédominance idéologique affectant tout ce qui est réputé arabe, quelle que soit son origine réelle; nous avons pu rapporter ce trait à la période shirazienne; cette prédominance

proclamée de l'Islam est telle que l'on a pu parler, pour définir la société comorienne, d'une société 'bantu d'idéal islamique'.

De cette dichotomie bantu - arabe que l'on retrouve sur le plan linguistique, découle l'organisation de notre travail, fondé sur la distinction entre un substrat bantu et des apports arabes; l'histoire récente explique l'influence du swahili, à laquelle nous ferons parfois allusion.

METHODOLOGIE ET DEFINITIONS

Où l'on expose l'organisation du travail et l'on  
introduit  
quelques définitions

## METHODOLOGIE

## APPROCHE DIFFERENTIELLE

Cette étude comporte deux parties : en premier lieu, une description du fonctionnement de la langue, sur les plans morphologique et phonologique; puis une seconde partie, qui est une tentative de dictionnaire étymologique et méthodique shingazidja-français.

La nécessité de ces deux parties, du moins d'une présentation grammaticale au lexique, tient à la spécificité du shingazidja : celui-ci est, comme nous l'avons dit, une langue bantou qui a été fortement influencée par l'arabe, lors même de sa constitution. On retrouve ainsi au niveau de la langue les deux composantes fondamentales arabe et africaine du peuplement et de la culture comoriens.

Cette influence apparaît, non seulement à travers la présence en shingazidja d'un grand nombre d'emprunts lexicaux, mais aussi aux plans phonologique, morphologique et syntaxique. C'est cette particularité qui a commandé notre mode d'analyse, que nous avons appelé approche différentielle et qui est associée à une conception de l'histoire de la langue.

Notre hypothèse en effet est que l'on ne peut comprendre le fonctionnement du shingazidja sans tenir compte de ce double héritage, mêlant fonds bantou et apport arabe.

Cette approche est fondée sur l'étymologie, et sur la constatation de la spécificité des emprunts; à sa base, deux postulats, justifiés par l'histoire:

- le shingazidja est une langue bantu qui s'est constituée en acceptant des apports de langues dites 'orientales'; parmi ces langues, l'arabe, à lui seul, suffit à représenter l'apport oriental;
- seule la prise en compte de ces influences orientales permet d'explicitier le fonctionnement du shingazidja.

L'aspect étymologique va de pair avec une hypothèse historique et génétique, qui la rend plausible; c'est sans doute le propre de ce qui touche à l'étymologie, en tant que discipline de la linguistique, que les hypothèses ou les modèles proposés ne puissent se concevoir sans relation à l'histoire.

Toute étymologie - et nous pensons autant aux emprunts qu'aux reconstructions bantu (voir plus bas) - implique une hypothèse historique; rappelons ce qu'en disait TURGOT, dans l'article "Etymologie" de l'Encyclopédie (cité par P. GUIRAUD 1979 : 31-32, qui en souligne l'actualité):

*"Une étymologie est une supposition, elle ne reçoit un caractère de vérité et de certitude que de sa comparaison avec les faits connus; du nombre de circonstances de ces faits qu'elle explique; des probabilités qui en naissent et que la critique apprécie (...). On doit vérifier par l'histoire des conquêtes et des migrations des peuples, du commerce, des arts (...) les étymologies qu'on établit sur les mélanges des peuples et des langues".*

D'une part, l'histoire fournit le cadre général qui rend plausible l'approche différentielle: elle nous apprend que des populations d'origine africaine bantu et des populations d'origine orientale 'shirazienne' ont été en contact, depuis des siècles, sur la côte orientale d'Afrique et dans les îles proches, et que ces contacts sont à l'origine d'une civilisation originale, dont les langues telles que le swahili, le shingazidja, sont les effets linguistiques; d'autre part, si, dans le modèle d'analyse que nous proposons, l'hypothèse qui privilégie, parmi les langues orientales, l'arabe, est d'abord fondée sur les

données linguistiques, avec une visée avant tout pragmatique, et ne saurait recevoir de lecture historique directe, il n'en demeure pas moins qu'elle s'inscrit dans une vision historique d'ensemble, ce dont elle se trouvera renforcée: nous proposerons ainsi une interprétation et une explication historique et sociologique de la pré-éminence de l'arabe sur le plan linguistique, alors même que les données historiques n'excluent pas l'influence directe sur le shingazidja du persan, voire d'autres langues orientales, influence dont on verra d'ailleurs que les faits linguistiques eux-mêmes confirment la possibilité.

Dans les différents aspects de la langue que nous décrirons, nous nous efforcerons donc de démêler ce qui provient de chacun de ces deux apports constitutifs, et de décrire la manière dont ils se marient pour aboutir au shingazidja actuel, la question essentielle étant de pouvoir attribuer une origine aux différents éléments de la langue.

La nécessité de ne pas se limiter à la synchronie dans la description d'une langue de ce type a été revendiquée par POLOME dans sa description du swahili:

*"Though its data [il s'agit de l'introduction à sa description du swahili] have essentially been elaborated from a synchronic point of view, a diachronic perspective has been introduced whenever it appears to throw some light on complex descriptive problems"* (POLOME 1967 : IX).

Dans une perspective plus générale, les premiers travaux structuralistes ont depuis longtemps reconnu l'artificiel des délimitations trop impératives de champ: étudiant les langues slaves, les membres du Cercle de Prague écrivaient:

*"On ne saurait poser de barrières infranchissables entre les méthodes synchronique et diachronique (...) la description synchronique ne peut pas non plus exclure absolument la notion d'évolution, car même dans un secteur envisagé synchroniquement existe la conscience du stade en voie de disparition, du stade présent et du stade en formation; les éléments*

*stylistiques sentis comme archaïsmes, en second lieu la distinction entre formes productives et non productives sont des faits de diachronie que l'on ne saurait éliminer de la linguistique synchronique"* (Travaux du cercle linguistique de Prague, thèse 1929 : 23-24).

Cette approche 'différentielle' correspond donc à la construction d'un 'modèle' qui devrait permettre la compréhension du fonctionnement actuel de la langue; ce modèle doit être jugé avant tout dans son résultat, dans son pouvoir explicatif, dans son aptitude éventuelle à rendre compte de la langue.

De l'exposé qui précède, on ne doit pas déduire que bantou et arabe soient, pour le shingazidja, sur le même plan.

En effet, quand nous posons le shingazidja comme le résultat de contacts entre langues bantou et arabe, nous ne voulons pas dire que la relation du shingazidja, au bantou et à l'arabe, soit symétrique: d'une part bantou et arabe sont des réalités de nature différente, et d'autre part, le shingazidja est une langue bantou (ce n'est pas un dialecte de l'arabe ni un parler sémitique), qui a été influencée par l'arabe: cela situe le composant bantou comme interne (aboutissement d'évolutions internes), et le composant arabe comme externe (résultat d'apports externes, essentiellement d'emprunts lexicaux massifs); ainsi, l'emprunt devant se prouver, tout ce qui en shingazidja n'est pas identifié comme d'origine exogène, arabe ou autre, et ne présente pas de raison particulière d'être suspecté de l'être, a vocation à être considéré, provisoirement en tout cas, comme autochtone, natif, 'indigène', bantou, que cet élément puisse ou non être rattaché à un étymon bantou concret.

L'élément bantou est donc substrat, et l'élément arabe, adstrat - ce qui rejoint l'hypothèse historique d'une population bantou antérieure à la venue des Shiraziens.

Toutefois, nous concevons le shingazidja comme le résultat d'un processus où l'arabe a joué un rôle majeur; lorsque, à partir du shingazidja actuel et du comparatisme bantu, nous postulons un ensemble de faits 'bantu', dégagé de toute influence non-bantu reconnue (nous parlerons de 'fonds bantu'), il s'agit d'un artifice de présentation, censé renvoyer à une (ou des) langue(s) bantu sur laquelle (ou lesquelles) l'influence arabe est venue s'imposer pour aboutir au shingazidja; mais, compte tenu notamment de l'état de la documentation bantu que nous utilisons (voir paragraphe suivant), on ne saurait à partir de là reconstruire la proto-langue bantu à laquelle nous faisons implicitement allusion.

En d'autres termes, le shingazidja est une langue bantu, mais qui n'existe pas indépendamment de l'influence arabe.

S'il peut nous arriver de transposer notre hypothèse dans un cadre historique et d'évoquer un état ancien du shingazidja, cela est fait par commodité et doit être pris avec réserve.

On a pu parler, à propos du swahili, de créole arabo-africain, ce qui pourrait se dire aussi du comorien ... C'est une hypothèse séduisante, mais c'est aussi, pour ce qui est du shingazidja, un débat en partie vain, ou du moins, précoce: la description et la compréhension des faits nous paraît un préalable à toute entreprise de catégorisation.

Si cette approche différentielle souligne par définition l'hétérogénéité constitutive du shingazidja, cela ne nous fera pas non plus oublier que le shingazidja, comme toute langue prise à un moment donné de son histoire, est un système, c'est-à-dire une structure organisée, relativement unifiée, cohérente, efficace; dans un article consacré à l'emprunt, le Professeur HOVIS a montré "l'identité dialectique de toute langue, [qui est à la fois] une systématique, un état structurel pour des locuteurs contemporains, et aussi le terme transitoire d'une évolution" (HOVIS 1983 : 35).

Il importe maintenant de définir ou de préciser ce que nous entendons ~~par~~ bantou et arabe, c'est-à-dire de définir les deux pôles constitutifs de la ~~langue~~ auxquels nous confronterons ses diverses manifestations.

## LES COMPOSANTS DE LA LANGUE

## 1 - LE BANTU

**Généralités**

Le bantu n'est pas une langue, mais une famille linguistique couvrant une grande partie de l'Afrique, depuis le Cameroun jusqu'à l'Est et au Sud du continent; le comorien, qui, nous l'avons dit, appartient à cette famille des langues bantu, représente ainsi l'extrême est de la diffusion bantu.

L'établissement de cette famille linguistique par des linguistes européens ne remonte guère au-delà du 19<sup>e</sup> siècle; il fait suite à l'entreprise de colonisation européenne de l'Afrique et résulte des études des langues africaines menées par des voyageurs, des religieux, des soldats, des administrateurs, des érudits: on remarqua l'existence de ressemblances constantes, d'analogies, entre de nombreuses langues africaines, géographiquement distantes; ces ressemblances seront peu à peu rassemblées et systématisées, dans une optique méthodologique inspirée des travaux antérieurs ou contemporains sur l'indo-européen ...

+ C. M. DOKE a dressé un historique de toutes les informations sur les langues bantu, déduites ou explicites - DOKE remonte ainsi au terme pharaonique de PUNT à propos duquel il écrit "*it is highly probable that Punt stands for bunt(u), the land of the people, the bantu*" (DOKE 1961 a : 1); il procède ensuite à un inventaire des premières descriptions et travaux comparatifs, jusqu'au 19<sup>e</sup> siècle (DOKE 1961 b), que prolonge une étude de COLE pour la période allant jusqu'aux années 60 (COLE 1961).

Parmi tous les chercheurs, nous mentionnerons seulement pour son apport tout particulier "le bibliothécaire allemand du Gouverneur du Cap, Wilhelm

BLEEK, dont l'ouvrage inachevé, *Comparative Grammar of South African Languages* (1862-1869) introduit dans la linguistique moderne le terme même de <<bantu>> (ALEXANDRE 1967 : 31).

+ Il est anecdotique de remarquer, incidemment, que le comorien (précisément le mohélien) fut une des premières langues bantu de l'Est à figurer dans une liste de vocabulaire: c'est la liste recueillie en 1613 par W. PAYTON, un commerçant britannique en route pour les Indes, lors d'une escale à Mohéli; cette liste d'une quinzaine de termes sera reprise plus tard par Th. HEBERT qui s'en attribuera indûment la collecte, ce qui abusera des auteurs postérieurs ... (voir, sur cette 'multiplication des sources', P. E. H. HAIR 1981 : 151); ce début précoce des études comoriennes (si l'on peut ainsi qualifier une liste de vocabulaire) fut sans suite.

D'un point de vue pratique, mais aussi méthodologique, nous nous référerons essentiellement aux travaux de Malcolm GUTHRIE, en particulier à sa critériologie, et à la somme que constituent les quatre volumes de *Comparative Bantu : an introduction to the comparative linguistics of the Bantu languages* (1967-1970), soit directement, soit à travers la glose qu'en a donnée le Professeur Pierre ALEXANDRE (dans le chapitre sur les langues bantu, in *Les langues dans le monde ancien et moderne* 1981).

#### *La critériologie bantu*

GUTHRIE a élaboré une critériologie permettant une définition objective sinon rigoureuse des langues bantu: les critères sont d'ordre morphologique et phonético-lexical; ils renvoient bien entendu aux résultats des travaux comparatifs.

Nous reprenons ici *in extenso* la présentation qu'en a fait P. ALEXANDRE (1981 : 354-355):

**A) Critères principaux.**

1) un système de genres grammaticaux, en général au moins cinq, présentant les caractéristiques suivantes:

a - les indicateurs de genres sont des préfixes, au moyen desquels les noms peuvent être assortis en classes, généralement au nombre de dix à vingt;

b - les classes sont régulièrement associées par paires opposant le singulier au pluriel de chaque genre. outre les genres à deux classes, il existe des genres à une seule classe, sans opposition singulier/pluriel, dont le préfixe peut être identique à l'un des préfixes, soit singulier, soit pluriel, d'un des genres à deux classes.

c - quand un mot a un préfixe indépendant (PI) comme indicateur de classe, tout mot qui lui est subordonné s'accorde en classe avec lui au moyen d'un préfixe de dépendance (pd).

d - il n'y a pas de corrélation des genres avec une notion sexuelle ou tout autre catégorie sémantique clairement définie.

2) une certaine proportion du vocabulaire doit pouvoir être rattachée par des règles fixes de correspondances phonétiques à un catalogue de racines communes hypothétiques.

**B) Critères subsidiaires**

3) un ensemble de radicaux (RAD) invariables à partir desquels la plupart des mots se forment par agglutination, ces radicaux présentant les traits suivants:

a - ils ont une structure -CVC-

b - en s'adjoignant un suffixe grammatical (sfv), ils constituent des bases verbales (Vb)

c - en s'adjoignant un suffixe lexical (sfn), ils constituent des thèmes nominaux (TN).

d - un élément d'extension (Ex), de structure -VC- ou -V- peut fréquemment être inséré entre le radical et le suffixe

e - les radicaux n'apparaissent sans préfixe que dans les interjections verbales (impératifs).

4) un système de voyelles équilibré, avec un /a/ médian et un nombre égal de voyelles antérieures et postérieures"

L'application de cette critériologie n'est pas sans soulever des problèmes, pour des langues se situant aux marges linguistiques du domaine - ce qui n'est pas le cas du comorien; certains ont proposé des modifications (voir les contributions de B. HEINE et de K. STALLCUP, notamment, dans l'ouvrage d'ensemble sur l'expansion bantoue, 19).

GREENBERG, en plus d'une critique de l'application de ces critères, a mis en cause la notion même de critériologie linguistique, faisant remarquer que des prédécesseurs réputés de GUTHRIE n'avaient pas cherché à en établir une et que rien de tel n'existe pour d'autres familles linguistiques, aussi peu discutées que les divers niveaux des langues indo-européennes, ou les langues sémitiques (GREENBERG 1949 : 81-82); certes, mais GUTHRIE a montré la spécificité des études génétiques portant sur les langues africaines en face des études sur les langues "classiques", conséquence de ce que HOVIS appelle "*la civilisation de l'oralité*": sans trace écrite ancienne, donc sans recul historique, on ne peut guère remonter dans le temps et reconstruire des états de langue ... Le comparatisme synchronique, c'est-à-dire la comparaison entre les langues contemporaines, permet de suppléer à l'information diachronique manquante, à condition de ne pas prendre d'une façon immédiate les 'reconstructions' ainsi opérées comme ayant valeur historique.

BENVENISTE a montré les limites de la distinction entre classement génétique et classement typologique, dans une situation qui s'applique tout à fait aux langues africaines: "*Il importe de voir que la distinction entre filiation et affinité n'est possible que dans les conditions de notre observation actuelle. Un groupement d'affinité, s'il est établi préhistoriquement, nous apparaîtra historiquement comme un indice de parenté génétique*" (BENVENISTE 1966 : 110)

Ainsi, dans cette perspective non diachronique, GUTHRIE parle-t-il de Bantu Commun (BC, *Common Bantu*) et non de proto-bantu.

C'est l'originalité de GUTHRIE par rapport à ses prédécesseurs, en particulier MEINHOFF, que de ne pas prétendre établir une reconstitution diachronique (voir ALEXANDRE, op. cit., 358)

A partir de la comparaison synchronique entre un certain nombre de langues bantu qualifiées de "test-languages", géographiquement réparties dans tout le domaine (le shingazidja n'en a pas fait partie), GUTHRIE a construit un ensemble de données de différents niveaux, qui constituent le Bantu Commun, concrétisé essentiellement dans un inventaire de quelques 2283 "séries comparatives" (CS, "*comparative series*"), représentant des reconstructions synchroniques (étymons), établies à partir de convergences sémantiques et formelles régulières entre des langues appartenant à des zones géographiques non contigues; ces formes sont données précédées d'un astérisque (GUTHRIE les désigne sous l'appellation de "*starred forms*"); les reconstructions lexicales sont munies d'indications sur les zones où elles sont attestées et le cas échéant sur leur combinatoire.

GUTHRIE avait entrepris de donner une dimension historique à certaines de ses reconstructions, qui figurent en un inventaire limité, mais cette partie est restée inachevée. Nous n'y ferons pas allusion ici.

Par ailleurs, étudiant la répartition des formes 'bantu commun' dans les différentes zones, GUTHRIE a tiré des hypothèses historiques (ou pré-historiques) sur les migrations des populations bantu: il distingue des formes à diffusion générale, et des formes dont la diffusion correspond à deux sous-ensembles géographiques du domaine, définissant ainsi un 'proto-bantu X', général, et deux 'dialectes', le 'proto-bantu A', *occidental* et le 'proto-bantu B', *oriental* (d'après ALEXANDRE, 1981., 360).

GUTHRIE distingue deux types d'étymons, qui correspondent à la structure morphologique des langues bantu - que l'on retrouve en shingazidja:

- des éléments radicaux ou lexicaux
- des éléments grammaticaux, parmi lesquels on distingue les extensions (qui sont des affixes dérivationnels), et ce que nous avons considéré comme des "actualisateurs", qui sont des éléments permettant notamment le passage des éléments lexicaux au statut d'éléments de discours (préfixes, suffixes).

Les éléments de chacune de ces deux catégories entrent dans une combinatoire: des tirets, à droite et/ou à gauche indiquent la nécessaire complémentation des éléments lexicaux: les étymons lexicaux sont toujours munis d'un tiret à gauche, indiquant qu'un préfixe est nécessaire: ainsi se présentent les "thèmes nominaux", les "radicaux verbaux" étant munis en outre d'un tiret à droite, indiquant la présence nécessaire d'un suffixe.

GUTHRIE donne souvent après les thèmes nominaux une indication sur le genre grammatical (classes en opposition de nombre) où le thème est inséré.

Nous citons les formes telles qu'elles sont données.

Parmi les radicaux, la majorité (sans doute) sont 'simplex', c'est-à-dire sans extension, alors que d'autres apparaissent munis d'une extension - sans d'ailleurs qu'une mention explicite en soit faite: il arrive que l'on trouve, dans la liste, à la suite, un radical 'simplex', et le même muni d'une extension, avec éventuellement renvoi de l'un à l'autre. Les radicaux peuvent aussi être des radicaux 'complex' (non réductibles à un radical simplex suivi d'une extension).

Ainsi, le bantu commun, construit par GUTHRIE, ne se situe-t-il pas dans une dimension historique: il ne se veut pas, du moins pas de façon immédiate, la reconstruction d'une proto-langue hypothétique ...; de ce fait, les symboles

graphiques utilisés pour en représenter les formes ne doivent pas en principe être investis d'une valeur phonétique ou phonologique, ce sont seulement des instruments de représentation, "une symbolisation arbitraire des correspondances phoniques entre les langues actuelles" (ALEXANDRE 1981 : 358); toutefois, ces symboles ont été choisis de façon à évoquer les articulations des langues réelles à partir desquelles ils ont été établis: on peut ainsi distinguer d'emblée parmi ceux-ci des consonnes orales notées \*C (\*p, \*b, \*t, \*d, \*k, \*g, \*c, \*j), des nasales \*N (\*m, \*n, \*n', \*ny), des suites nasales - consonnes (\*NC) et \*y, ainsi que des voyelles (\*i', \*i, \*e, \*a, \*o, \*u, \*u'); on verra plus loin que les consonnes orales se partagent entre 'occlusives' (\*p, \*b, \*t, \*d, \*k, \*g) et 'affriquées' (\*c, \*j), à partir de leurs 'reflets' (notre traduction de l'anglais reflex) en shingazidja, \*g étant toutefois quelque peu à l'écart.

#### Liste des symboles du B.C.:

(les symboles du bantou commun, basés sur les lettres de l'alphabet latin, ne posent pas de problème de transcription; pour les symboles non standards, nous utilisons l'indice ' suivant la lettre sur laquelle ce symbole est formé: ainsi, \*n' note la nasale vélaire \*ŋ, \*i' et \*u' les voyelles fermées \*i et \*u)

#### consonnes :

\*p \*b

\*t \*d

\*k \*g

\*c \*j

\*y

\*m \*n \*ny \*n'

\*mb \*nt \*nd \*ng \*nk (\*ny)

(\*ny: peut être considéré comme une nasale ou une séquence nasale consonne)

#### voyelles :

\*i'

\*i  
\*e  
\*a  
\*o  
\*u  
\*u'

Depuis, d'autres travaux ont vu le jour, qui proposent d'autres analyses, mettant en cause notamment le découpage en zones de GUTHRIE, qui conditionne l'établissement des séries comparatives, ainsi que ses hypothèses sur les "migrations bantu" ...; MEEUSSEN aboutit ainsi à d'autres listes étymologiques, dans une optique historique (*Bantu grammatical reconstructions*, 1967 et *Bantu lexical reconstructions*, 1969, récemment enrichies par l'équipe de TERVUREN); d'autres auteurs proposent des reconstructions régionales - notamment pour la zone orientale, DEREK-NURSE notamment (voir plus loin).

MEEUSSEN a critiqué différents points des reconstructions de GUTHRIE, notamment la différenciation entre certains symboles qui se retrouvent fréquemment dans les séries 'osculantes' (séries comparatives ne différant que sur un point - nous aurons l'occasion d'y revenir *infra*) (MEEUSSEN 1973 : 6-18); pour une appréciation et une mise en perspective des hypothèses historiques de GUTHRIE et de la problématique de constitution des séries comparatives, voir l'article détaillé de DALBY (DALBY 1975 : 481-501 & 1976 : 1-27); l'aspect historique est également abordé dans l'ouvrage collectif *L'expansion bantoue* (1977): notamment les contributions de B. HEINE déjà cité pour l'interprétation de la pré-histoire bantu et de nouvelles hypothèses sur les migrations, et d'A. COUPEZ continuant MEEUSSEN pour la reconstruction du proto-bantu, article dont Cl. GREGOIRE a souligné l'importance (GREGOIRE 1984 : 64-65).

Pour la zone orientale, nous avons signalé les recherches de DEREK-NURSE et HINNEBUSH sur les reconstructions historiques des langues *sabaki*, où figurent les parles comoriens; mais cette entreprise n'est pas achevée.

Il va de soi que notre propre travail eût bénéficié de s'appuyer sur de telles reconstructions, plus proches et plus fiables pour le shingazidja sans nul doute que le bantou commun; il est possible aussi que notre travail pourra aider à cette reconstruction.

#### *Modalités d'application au shingazidja*

Nous ne nous intéresserons donc pas, dans ce travail, aux hypothèses historiques que l'on peut déduire du bantou commun, non plus qu'aux indications des zones géographiques où la forme a été relevée, mais nous nous en tiendrons aux reconstructions de GUTHRIE, car il nous a semblé que le BC de GUTHRIE, tel qu'il se présentait, suffisait à répondre à notre objectif - essentiellement mettre en évidence ce qui est substrat vis-à-vis de ce qui est adstrat en shingazidja - et que la confrontation avec les autres reconstructions à portée générale, même si sur certains points ces dernières peuvent être plus proches du shingazidja, n'apportait aucun avantage décisif eu égard à cet objectif; par ailleurs, le travail de GUTHRIE reste unique par son ampleur et son souci méthodologique.

Rappelons que notre objectif est de faire ressortir quels sont les éléments que l'on peut rapporter au bantou dans le système phonologique du shingazidja, et plus généralement ce qui est bantou en shingazidja, et non pas d'étudier si ou dans quelle mesure les reconstructions de GUTHRIE s'avèrent pertinentes vis-à-vis du shingazidja.

En tout état de cause, le shingazidja s'inscrit aisément et sans ambiguïté dans le cadre critériologique défini par GUTHRIE: dire du shingazidja qu'il est une langue bantou signifie qu'il répond positivement à l'ensemble des critères, de GUTHRIE, principaux comme subsidiaires, ce que nous vérifierons à mesure de l'étude.

A partir de là, et de l'état de la documentation bantou, le terme de 'bantou' employé pour qualifier un aspect ou un élément quelconque du shingazidja recouvre deux acceptions:

- une acception précise implique que l'élément qualifié de bantou en shingazidja soit le 'reflet' d'un élément figurant explicitement dans le Bantou Commun de GUTHRIE:

ainsi, l'existence d'un système de classe en shingazidja est-elle une caractéristique bantou; de même, le mot *nyunba*, *maison*, est-il bantou, puisqu'il reflète le thème BC \*-nyumbá, de même sens (on constate les identités formelle et sémantique - toutefois, comme nous le verrons, les correspondances formelles ne sont pas basées sur la simple ressemblance de deux éléments, qui pourrait être fortuite mais sur un ensemble de faits convergents;

- une acception large reprend ce que nous avons dit plus haut: tout ce qui n'est pas avéré comme emprunt et qui n'a pas de raison particulière d'être suspecté de l'être, est, de façon provisoire en tout cas, réputé appartenir au substrat bantou, alors même qu'aucun étymon n'a été identifié chez GUTHRIE:

ainsi, le mot *ubabuha*, bavarder, pour lequel nous n'avons aucune indication qu'il provienne d'une langue ayant influencé le shingazidja, quoique sans étymon dans la liste de GUTHRIE (nous n'avons pas trouvé de racine \*-babuk-, ou \*-bab- avec un sens compatible ...) sera-t-il posé comme bantou.

Les éléments bantou dans cette acception le sont donc tant que l'on n'a pas trouvé qu'ils aient une source non-bantou plausible (bantou par défaut).

Pour les mêmes raisons, nous accordons primauté de principe à l'origine bantou lorsque celle-ci est en concurrence avec une étymologie 'étrangère', les deux étant selon toute vraisemblance non apparentées: ainsi *nbuzi*, chèvre, sera-t-il référé (malgré KNAPPERT 1983 : 131, cité plus bas) à l'étymon bantou \*-búdi', chèvre, plutôt qu'au persan *buz*, chèvre, etc.

Ce type de coïncidence, ou, à la rigueur, de "collision étymologique" (terme que nous reprenons à DE CALUWE DOR, 1982 : 68, au sens où "les

faisceaux des deux ou trois étymologies possibles convergent pour donner un mot unique dans la langue donnée", le mot shingazidja représentant alors à la fois l'aboutissement de l'étymon bantu et l'emprunt au persan), sans être fréquent, n'est pas unique.

Quand nous les avons relevées, nous donnons ces deux ouvertures étymologiques dans le lexique, mais la source étrangère (non bantu) sera donnée comme douteuse, et seul l'étymon bantu retenu dans les évaluations.

Nous parlerons de bantu *stricto sensu* pour les éléments avec étymon prouvé, de bantu *largo sensu* pour les éléments bantu par défaut.

Mais, même parmi les éléments shingazidja possédant un étymon, nous serons amenés à distinguer ce qui correspond à un développement phonétique régulier en shingazidja, et ce qui est un emprunt à des langues bantu voisines (swahili), expliquant ainsi un comportement irrégulier en shingazidja:

L'importance des contacts entre swahili et shingazidja ne peut être mise en doute, dès la présence des Shiraziens d'une part, et plus récemment, au XIX<sup>e</sup> siècle, d'autre part, lorsque le sultanat de Zanzibar étendit sa domination sur les Comores. De cette dernière période datent sans doute la majeure partie des nombreux emprunts lexicaux au swahili actuellement décelables, ceux-ci pouvant être tant des termes bantu - ce qui est susceptible de créer des 'doublets' entre la forme régulière et la forme venue par le swahili, synonymes ou avec une spécialisation sémantique, nous en verrons de nombreux exemples - que des termes eux-mêmes d'emprunt, provenant de langues avec lequel le swahili a été en contact (on retrouve les mêmes influences orientales que pour le shingazidja, ainsi que l'anglais, etc).

La problématique grammaticale plus ou moins explicite dans la mise en rapport du shingazidja avec le bantu commun s'inspire à la fois de la problématique grammaticale de GUTHRIE (à travers sa critériologie) et des caractéristiques morphologiques du shingazidja, qui apparaîtront plus loin:

nous ne posons pas comme nécessaire l'identité entre catégorie lexicale du BC (radical verbal / thème nominal) et formes shingazidja:

par exemple, un radical verbal BC peut être proposé comme l'étymon d'un nominal shingazidja, pourvu que la structure morphologique du nominal shingazidja rende plausible l'hypothèse que celui-ci soit issu d'une racine verbo-nominale, qu'une formation verbale apparentée soit effectivement attestée en shingazidja ou non:

ainsi, BC \*-dim-, cultiver, sera-t-il donné comme étymon tant de Ng -*lima*, cultiver, que de *ndima*, agriculture;  
de même, BC \*-bèng-, luire, sera-t-il donné comme étymon de Ng *mwendje*, lumière, alors même que nous n'avons pas relevé de verbe \*-wendja, dans le sens de luire ...

Nous admettons aussi des correspondances entre étymons 'simples' (radicaux simplex, dans la terminologie de GUTHRIE, c'est-à-dire dépourvus d'extension) et termes shingazidja pouvant s'analyser, dans une optique comparative, comme étendus, quand bien même la forme simple n'est pas attestée dans notre corpus - il s'agit alors de radicaux dits complex, voir partie morphologie - pourvu bien entendu que l'analyse ainsi effectuée se voit confortée par la sémantique (que le sens du mot shingazidja corresponde à peu près à celui de l'étymon postulé combiné avec celui de l'extension ...):

ainsi, \*-kòb-, accrocher, sera-t-il donné comme étymon de Ng -*hoboa*, décrocher, étant entendu que -*hoboa* peut s'analyser comme la combinaison d'un radical simplex donnant le verbe (non attesté) \*-hoba et de l'extension, maintenant inopérante, à valeur réversive |-u-/ -o|: -*hoboa*: |-hob-o-a|

Ces deux conditions - acceptation d'une différence de catégories grammaticales et de complexité de l'étymon - sont importantes si l'on ne veut pas restreindre considérablement le corpus retenu pour les comparaisons.

Par contre, cette problématique souligne d'emblée la difficulté d'avancer des critères stricts pour déterminer les correspondances à retenir, dès lors que les analyses ou reconstructions auxquelles nous procédons en shingazidja sont

hypothétiques, et obligent à des interprétations sémantiques où la subjectivité de l'auteur intervient.

Cela amène au point le plus problématique qui est celui des gloses: les rapprochements se font à travers l'équivalent français des gloses anglaises extrêmement brèves de GUTHRIE et du shingazidja - ce qui signifie deux traductions ...

Sur le plan des symboles, une correspondance est retenue dès lors qu'elle est vérifiée par au moins trois exemples similaires.

Les correspondances retenues sont reprises dans un index des étymons bantu, dans le vol. III, qui est accompagné d'un index des étymons douteux - qui sont des étymons non retenus mais proposés comme piste.

En tout état de cause, il importe de rester conscient que les données du B.C. de GUTHRIE ne constituent pas une langue réelle, mais un ensemble de correspondances morphologiques, lexicales et phoniques abstraites, issues de la comparaison entre diverses langues actuelles ou quasi-actuelles (la date où fut recueillie la documentation utilisée par GUTHRIE varie selon les langues), dont aucune n'actualise la totalité; cet inventaire ne représente pas autre chose qu'un état de la connaissance sur le domaine bantu, en voie, vraisemblablement, d'être remplacé (au moins pour la zone qui nous intéresse) par des travaux plus complets. C'est l'une des raisons pour lesquelles on ne saurait conclure que tout fait shingazidja ne correspondant pas à un fait répertorié par GUTHRIE n'est pas bantu ...

## LES COMPOSANTS DE LA LANGUE

## 2 - L'APPORT EXOGENE

Du fait de la situation socio-linguistique où il s'est forgé, le shingazidja a reçu un apport considérable d'emprunts à des langues "orientales" - c'est-à-dire des langues se trouvant à "l'orient" du shingazidja, qui se trouvent aussi être des langues orientales pour nous: ce sont les langues des populations shiraziennes qui se sont installées aux Comores à partir du XII<sup>e</sup> siècle, essentiellement l'arabe, sans exclure le persan, l'hindi; on peut y assimiler le portugais du fait que les contacts se sont produits pendant la période shirazienne; d'autres langues source d'emprunt sont le français et le malgache. Les emprunts identifiés s'opposent aux mots 'indigènes', réputés de bantu.

Les tableaux suivants indiquent la répartition par source

i) des articles lexicaux (calculée à partir des index étymologiques, vol. III);

ii) des mots:

*1) par article lexical*

Dans l'optique suivie, un article lexical regroupe tous les termes construits sur la même base ou lexème ou, pour les emprunts à l'arabe, issus de la même racine étymologique: un article peut, tantôt se confondre avec un seul terme, tantôt, par le jeu des processus de dérivation propres au shingazidja ou par le biais de la dérivation arabe, regrouper jusqu'à plus de dix termes.

nombre d'articles lexicaux:

- exogènes (emprunts)	844	
• emprunts orientaux:		532 + 115 = 647
dont arabes (racines arabes):	505 + 74	

persans	27 + 4	
emprunts non identifiés:	37	
• emprunts 'européens':		153 + 12 = 165
dont portugais:	9 + 4	
français:	144 + 8	
• autres emprunts:		28 + 6 = 32
dont malgaches:	10 + 1	
• emprunts divers:	18 + 5	
- 'indigènes' :	1444	
dont articles munis d'un étymon bantu:	510 + 140	
dont articles sans étymon :	794	

soit un total de 2288 articles.

× le premier chiffre exprime les emprunts 'sûrs', le second, les emprunts 'douteux'.

× Sont inclus parmi les emprunts à l'arabe tous les mots attestés dans les sources arabes consultées, quelle que soit leur origine, qu'il s'agisse de mots d'origine arabe proprement dite ou d'emprunts provenant d'autres langues orientales (persan, turc, hindi, etc) mais pour lesquels l'arabe a ou jouer le rôle de média (discussion ci-dessous).

× les emprunts 'non identifiés' sont des mots dont les caractéristiques phonétiques permettent d'affirmer qu'il s'agit d'emprunts orientaux mais pour lesquels la source n'a pas été identifiée (mais nous avons bon espoir qu'elle le soit); ils se rajoutent aux emprunts orientaux; ces termes au demeurant sont en petit nombre.

#### 11) par mot:

- emprunts identifiés :	1321
dont arabe :	1019
persan :	30
autres :	181
douteux:	91
- 'indigènes' :	2298
dont munis d'un étymon bantu :	1113
sans source :	1185

soit un total de 3619 mots

On mesure l'importance du lexique d'origine orientale en shingazidja, où l'arabe, qu'il soit la source étymologique ou seulement le média de passage, se taille la part du lion ( 505 articles et 1019 mots, représentant respectivement env. 1/4 et 1/3 des termes).

En énoncé, la proportion des emprunts varie bien entendu selon le thème abordé, le niveau de langue, la situation, les locuteurs, etc, mais elle est rarement inférieure à un tiers des mots dans tout discours.

Les emprunts touchent tous les domaines, non seulement ceux que l'on peut tenir pour culturellement marqués, comme la religion, le droit, le commerce, la navigation, la divination, la politique, les techniques, etc, mais aussi le vocabulaire courant, avec souvent, entre terme d'emprunt et terme 'bantu' (parfois même entre emprunts de différentes sources), des synonymes ou quasi-synonymes:

exemples : (en premier, le terme le plus courant):

**udiwaza** (bantu) = **usahau** (ar.) : oublier  
**umaliza** (bantu) = **utimizi** (ar.) : terminer  
**sahani** (ar.) = **shano** (bantu) : assiette  
**ranbuu** (ar.) = **shileo** (bantu) : bétel  
**fedha** (ar.) = **pesa** (pers.) : argent  
                   fedha et pesa sont des doublets, l'un et l'autre étant sans doute de même origine  
**mfunguo** (bantu) = **lakile** (fr.) : clef  
**uázimia** (ar.) = **udeside** (fr.) = **upania** (bantu) : décider  
**mfano** (bantu) = **mathalani** (ar.) : par exemple  
**sha** (bantu) = **lakini** (ar.) : mais  
 etc

Des spécialisations sémantiques entre ces termes peuvent se produire:

- **ndishe**, orphelin de mère (bantu) + **yatima** (ar.), orphelin de mère (il semble que l'arabe *yatim* signifie 'orphelin d'une façon générale')
- **-elewa**, comprendre (bantu) + **-fahamu** (ar.), surtout employé dans le sens de 'faire attention' (parfois de comprendre) (ar. *fahima*, comprendre)
- **mau**, amende infligée à celui qui a offensé un notable dans le cadre traditionnel + **lamanda** (fr.), amende

• paa la shio, 'pièce du livre', école coranique • darasa, likoli, école officielle (ar. et fr.)

Les emprunts à l'arabe se retrouvent dans toutes les catégories grammaticales, noms, verbes et fonctionnels, alors que, d'une façon générale, les emprunts d'autres provenances sont limités aux nominaux (sauf les emprunts au français, qui comportent aussi des verbes):

noms:

isimu, nom (ar.)  
 kalamu, crayon (ar.)  
 fikra, pensée (ar.)  
 likoli, école (fr.)  
 pera, goyave (port.)  
 etc

verbes:

ufikiri, penser (ar.)  
 utahani, moudre (ar.)  
 ubusu, embrasser (ar.)  
 urosi, arroser (fr.)  
 etc

fonctionnels divers (une classification détaillée sera donnée ultérieurement):

baada, après (ar.)  
 la, non (ar.)  
 ili, pour que (ar.)  
 etc

Spécificité et importance de l'apport arabe

Pour des raisons culturelles et historiques, les emprunts à l'arabe, ou pour lesquels l'arabe a servi de média de passage, dans leur grande majorité, ne se sont pas complètement assimilés aux termes indigènes et manifestent une spécificité sur les plans phonologique, morphologique et lexicologique, venant ainsi perturber des systèmes existants:

- phonologie: introduction de phonèmes spécifiques, rupture de conditionnements existants;
- morphologie: comportement grammatical particulier des verbaux;
- lexicologie: apport du système dérivationnel de la langue source

Cela explique notre hypothèse, selon laquelle on ne peut expliciter de manière satisfaisante le fonctionnement du shingazidja sans se référer à la dualité fondamentale qui le constitue (bantou vs arabe)

Aussi notre approche n'est-elle pas purement étymologique, les emprunts nous intéressant du fait de leur spécificité dans les divers plans de l'analyse, et non pas comme tels: il est clair que si les emprunts à l'arabe présentaient une intégration phonétique et morphologique totale et ne se distinguaient des termes 'indigènes' de diverses façons, il n'y aurait pas de justification particulière à une telle approche, en dehors, bien sûr, d'une étude volontairement étymologique.

Par contre, les autres langues ayant fourni des termes au shingazidja ne montrent pas de telles particularités: elles n'introduisent aucune spécificité, que ce soit sur le plan phonétique ou morphologique ou lexicologique.

C'est la raison pour laquelle nous privilégions l'influence arabe comme matrice exogène de la langue actuelle, les autres langues n'étant mentionnées, comme source d'emprunts, que par souci d'exhaustivité.

Cette situation est commune à d'autres langues de la région, qui se sont trouvées dans la même configuration socio-historique, le swahili notamment, mais les particularités phonologiques du shingazidja ainsi qu'une attitude *grosso modo* plus conservatrice ou plus fidèle des locuteurs vis-à-vis des emprunts à l'arabe, font que les conséquences y sont sans doute plus apparentes.

+ La similarité entre swahili et shingazidja - le fait que les influences ayant agi sur ces deux langues soient les mêmes - nous permettra de faire appel à l'occasion à des travaux sur le swahili (notamment étymologiques).

+ P. ALEXANDRE a fort justement rapproché swahili et anglais, sur le plan lexical ("*le swahili*) Langue bantou, donc, mais avec un lexique contenant

une proportion considérable de mots d'emprunt à l'arabe; tout comme l'anglais est une langue germanique avec une forte proportion d'emprunts romans" (ALEXANDRE 1983 : 7).

#### *L'arabe, comme source ou média de passage des emprunts orientaux*

Parmi les diverses langues orientales sources, la part de l'arabe apparaît en effet primordiale pour les raisons suivantes:

- quantitativement, l'arabe est la source immédiate largement majoritaire des emprunts orientaux;

- qualitativement, la prise en compte de l'arabe suffit à rendre compte de pratiquement toutes les spécificités des emprunts, sur les divers niveaux de l'analyse linguistique, phonologique comme morphologique. En effet,

a) une part non négligeable des mots d'origine orientale que d'aucuns considèrent comme provenant d'autres langues que l'arabe (en particulier persans, cf. KNAPPERT repris plus bas) figure dans des dictionnaires d'arabe et leur traitement phonétique en shingazidja ne permet pas - sauf en de très rares cas - de déterminer la langue de passage, et donc, d'établir une influence autre qu'arabe;

b) parmi les mots shingazidja d'origine orientale non attestés en arabe, ou que leur traitement phonétique amène à considérer comme d'origine non-arabe, la majeure partie sont des nominaux (noms), exceptionnellement des fonctionnels; en particulier, on ne trouve aucun verbe d'origine orientale qui ne soit attesté en arabe; or, il ne fait pas de doute que l'insertion d'un emprunt dans la catégorie des verbes n'est pas du même ordre que l'insertion d'un emprunt dans toute autre catégorie grammaticale - elle est la preuve d'une influence plus forte ...

Nous avons formulé cette remarque dans un article sur le comportement des verbes d'origine arabe en shingazidja (LAFON

1984 : 28); même chose en swahili selon J. KNAPPERT "(...) no Swahili verbs of Persian or Turkish origin have been found", KNAPPERT 1983 : 113).

Ceci pose à l'évidence le problème des sources documentaires, auquel s'ajoute, pour l'arabe, celui de la variété de référence - nous reviendrons sur ce dernier point en fin de chapitre.

On sait que l'arabe lui-même a emprunté de nombreux mots au persan, aux langues indiennes, au turc, etc; les dictionnaires arabes auxquels nous faisons référence (Hans WEHR, notre référence principale pour le lexique, DOZY et BELOT auxquels nous avons recours lorsque WEHR est 'défaillant') signalent d'ailleurs avec plus ou moins de précision les mots que l'arabe a empruntés et sont eux-mêmes plus ou moins ouverts aux emprunts: ainsi, de nombreux mots figurent, avec souvent l'indication d'une origine persane, dans DOZY, qui ne se trouvent ni dans WEHR, ni dans BELOT.

Cela est du reste normal, l'objectif revendiqué par DOZY, apparent dans le titre *"Supplément aux dictionnaires arabes"* étant de recenser de nombreux termes arabes ignorés par les dictionnaires précédents, souvent du fait qu'il s'agit d'emprunts; DOZY écrit *"Un autre inconvénient [des dictionnaires antérieurs] c'est que ces lexicographes étaient des puristes outrés, qui ne voulaient donner que la langue soit disant classique"* (Avertissement, p 1)

Pour ce qui nous concerne, cela signifie simplement que ces mots, quelle que puisse être leur origine première, sont, ou ont été, utilisés en arabe, et donc, que l'arabe a pu être leur vecteur en shingazidja - pour autant, bien sûr, que cela soit compatible avec leur traitement phonétique: en fait, dans le cas de mots attestés en arabe, seul le traitement phonétique peut permettre de discriminer entre emprunts via l'arabe et emprunts directs à une autre source.

Nous présentons à travers quelques exemples la problématique rencontrée.

Les étymologies arabe & persane ne peuvent être discriminées sur la base de la forme shingazidja: nous retenons par principe l'origine arabe:

En effet, nous basant sur les dictionnaires arabes de WEHR et BELOT, nous considérons comme d'origine arabe les mots shingazidja *bahati*, *chance*, *bahishi*, *pourboire*, *kitani*, *sisal*, *mهوري*, *sceau*, que KNAPPERT, dans un article cité très riche sur les emprunts perses et turcs en swahili, donne comme persans (swa. *bahati*, *bahashishi*, *katani*, *muhuri*, venant de pers. *bahs'is*, *baht*, *katān* et *muhur* respectivement [1983 : 120-1 & 129]): ces mots figurent sans mention particulière dans H. WEHR (*baht*, *baksis*, *kattān* ou *kittān*, *muhur*), BELOT attribuant aux deux premiers une origine étrangère (non précisée).

Il ne s'agit évidemment pas de contester l'origine persane de ces mots; nous n'apportons pas non plus de preuve décisive que ces mots soient entrés en shingazidja et en swahili par l'arabe - il faudrait envisager divers facteurs, dater leur entrée en arabe, cela n'est plus notre sujet.

Nous affirmons simplement que rien ne permet, sur le plan linguistique, dans de tels cas où les formes arabes et persanes sont phoniquement proches sinon identiques, de déterminer de façon indubitable la langue de passage en shingazidja: dans ce cas, aucune raison n'apparaît de postuler un emprunt au persan plutôt qu'à l'arabe ...

Notons, dans le cas de *muhur*, que WEHR le donne sous la racine arabe MHR: cela montre que ce terme est considéré comme pleinement naturalisé (cette même racine donne le mot shingazidja *mahari*, dot)

Cas où l'on dispose d'indices formels permettant de distinguer les deux langues

Il est bien entendu la langue présentant la plus grande proximité formelle que

nous retenons: deux situations:

Cas d'entrée directe à partir du persan:

le passage par l'arabe sera logiquement refusé lorsque l'emprunt comporte en shingazidja des phonèmes que l'arabe ne connaît pas - avec prudence toutefois, car des transformations peuvent se produire en shingazidja;

cette situation est restreinte à un petit nombre de situations: c'est le cas des mots orientaux comportant une bilabiale sourde [p], une vélaire sonore [g] (pour ce dernier son, avec des nuances), et, au moins dans un mot, [c]:

× [p] ([β], pv):

+ rappelons le doublet Ng *pesa*, argent, du persan *paisa* + *fedha*, id., de l'arabe *fid.d.a*;

+ Ng *pvili-pvili*, poivre : rapporté au persan *pilpil* plutôt qu'à l'arabe *filfil*, tous deux d'ailleurs de même origine: si l'arabe avait été la source du mot shingazidja, on aurait normalement \**filfili* ... (notons que la continue bilabiale pv ([β]), pour persan [p], ne s'explique pas très bien non plus dans cette position - on attendrait plutôt \**pilipili*, ce qui est d'ailleurs la forme en swahili; nous reviendrons sur cet exemple dans l'étude de phonologie dynamique.)

× [g] :

+ Ng *legeni*, moule à gâteau : rapproché de pers. *lagan* (en var. *lakan*), plutôt que de l'ar. *lakan*, de même sens.

Toutefois, Ng [g] peut provenir de la sonorisation de la vélaire arabe [q] ou de la palatale [k], dont l'opposition est neutralisée en shingazidja au profit de [k], comme dans Ng *-gabili*, variante de *-kabili*, faire face, de ar. *qabāla*, Ng *bigra*, virginité, de ar. *bikr*, vierge...

+ Ng *bange*, chanvre : ce mot sera plutôt rapporté au pers. *bang* qu'à l'arabe *banj*, car le djim ar. ج. j est généralement rendu par Ng [dj].

Toutefois, il est des cas où ar. ج. j aboutit à Ng [g] : *ngamia*, chameau, de ar. *jamal* ce qui là aussi montre qu'il faut être prudent ...

× [c] :

Ng *tshai*, thé : rapproché du persan *chai*, dont il est plus proche phoniquement, plutôt que de la forme arabe *s'äyy*.

• cas de passage avéré par l'arabe:

Dans quelques cas, à l'inverse, le phonétisme trahit un passage par l'arabe (ces cas sont pour nous difficile à identifier car dès lors que l'arabe fournit une étymologie satisfaisante, nous nous en contentons):

Ng *bwara*, pièce d'or < ar. *bāra* < pers. *para* (DOZY); si le mot était entré directement du persan en shingazidja, on devrait avoir une forme telle que \**para*, ...

D'une façon générale, simple coïncidence ou bien indice, les phonèmes introduits par les emprunts aux langues orientales et plus généralement l'influence de ces langues sur le plan phonique ne se distingue pas de celle de l'arabe. De ce fait, la prise en compte, sur un même plan, des langues orientales autres que l'arabe pour la description du shingazidja, ne paraît pas nécessitée par des arguments linguistiques décisifs.

C'est là l'argument linguistique qui permet de penser que l'élément arabe dominait parmi les populations shiraziennes.

Pour résumer, on peut dire que ce qui importe en réalité, plus que l'origine première d'un mot ou d'un phénomène, est d'une part de déterminer s'il s'agit d'un emprunt ou non (donc, puisque l'emprunt doit se prouver, d'un élément rattachable à telle source étrangère), et d'autre part de voir en quoi cela (l'origine étrangère) se manifeste dans le comportement de l'élément en shingazidja.

Or, au moins pour ce qui est des emprunts orientaux, tout ce qui leur est spécifique peut être référé à l'arabe.

Du point de vue des emprunts lexicaux, nous nous en tiendrons donc à l'origine arabe dès lors que le mot figure, dans une forme acceptable pour le shingazidja, dans un des dictionnaires arabes consultés, quelle que soit par ailleurs son origine 'première', et les étymologies persane, hindi, ou autre, ne seront retenues que pour les emprunts lexicaux aux langues orientales dont nous n'avons pas trouvé trace dans les dictionnaires consultés.

Outre la limite posée d'emblée par la référence principale à l'arabe littéraire, à travers WEHR, sur laquelle nous revenons plus bas, il est évident que ces dictionnaires, pas plus que d'autres, ne sauraient être exhaustifs, même pour la variété de langue qu'ils décrivent; il serait

évidemment préférable de pouvoir recourir à des dictionnaires monolingues arabes, mais, n'étant pas arabisant, ceux-ci ne nous sont pas accessibles. Nous qualifierons de [+ arabe] tout phénomène se rencontrant en shingazidja qui montre des présomptions d'y avoir été introduit par l'arabe (que l'arabe en soit la source première ou seulement le véhicule), et de [- arabe] tout phénomène qui ne paraît pas pouvoir être référable à l'arabe.

+ Nous n'avons pas l'ambition de KNAPPERT, qui cherche à remonter à la source première de chaque mot exogène en swahili (grec, sanscrit, turc, malais, etc) (op. cit., et conférence sur les emprunts grecs en swahili lors des Journées de Sèvres sur le swahili, mai 1983 - non publiée à notre connaissance; ces emprunts grecs s'avèrent en fait être passés, soit par l'arabe, soit par le biais de l'anglais, à l'occasion notamment de traduction de la Bible ...)

+ Signalons aussi qu'il y a lieu d'être prudent dans l'utilisation de certaines sources existantes: le dictionnaire swahili-anglais de JOHNSON (1939) mentionne un certain nombre d'étymologies, arabes mais aussi perses: voilà ce qu'en dit KNAPPERT, auquel on peut certainement faire confiance (op. cit. p 113): *"Only sparing use has been made of (...) the alleged etymologies in JOHNSON (1939), which are so often erroneous that much time can be wasted looking for a word in the dictionary that is placed under its presumed but incorrect 'original' form. Indeed, JOHNSON sometimes invented or composed Persian words..."* (op. cit. p 113).

Pour notre part, et sauf indication explicite, toutes les étymologies proposées ont été trouvées ou vérifiées dans des dictionnaires de la langue source (voir lexique pour l'inventaire des sources)

+ Les opinions sur l'existence, perçue à travers les phénomènes d'ordre linguistique, de contacts directs persan - swahili (auquel on peut assimiler ici le shingazidja) sont partagées:

Dans un article où il envisage la constitution du swahili, ALEXANDRE note que de tels contacts ne sont pas prouvés: *"certains mots d'emprunt paraissent persans, turcs ou urdu ou encore portugais d'origine, sans que l'on puisse déterminer avec certitude s'ils ont été empruntés directement ou via l'arabe"* (ALEXANDRE 1983 : 8-9).

Plus déterminée, Sh. M. ZAWAWI dans son ouvrage sur les emprunts en swahili prend pour hypothèse de départ ("*working assumption*") que l'arabe a joué dans tous les cas le rôle de média: "*that many words that have hitherto been identified as Persian, Hindi or Portuguese have not been taken directly from these languages but have, in fact, come into Swahili via Arabic, and a very old dialect of Arabic*" (ZAWAWI 1979 : 3); nous reviendrons plus bas sur un autre aspect, plus contestable selon nous, de cet ouvrage.

BALDI, dans sa thèse sur les emprunts arabes en swahili et en hacussa, n'exclut pas la possibilité de contacts directs entre urdu, persan et swahili (BALDI 1982 : 1 : 15), contacts que J. KNAPPERT admet au moins implicitement: dans l'étude très documentée que nous avons déjà citée, ce dernier écrit à propos de deux mots d'origine, l'un grecque, l'autre inconnue: "*in the case of these two words, Persian and Turkish have merely functioned as mediating languages*" (op. cit. 1983 : 114): si le perse et le turc ont pu fonctionner comme média, c'est, semble-t-il, que l'auteur suppose contact direct ...

#### L'hypothèse historique : l'arabe comme langue véhiculaire ?

Reste à envisager comment l'environnement historique et culturel du shingazidja peut expliquer le rôle prépondérant de l'arabe, alors même que, pour dominants qu'aient pu être les Arabes parmi les Shiraziens, on ne saurait exclure la présence d'autres langues.

Divers facteurs ont pu jouer:

##### *1) l'influence constante de l'arabe:*

depuis la constitution même de la langue jusqu'à nos jours, des contacts entre arabe et shingazidja se sont toujours maintenus, à un niveau ou un autre, qu'il s'agisse de relations avec des populations arabophones originaires de la côte arabe, ou d'une influence de la langue littéraire à travers l'Islam;

##### *ii) l'importance de l'arabe pour toute population musulmane:*

on sait le rôle de l'arabe, langue de la religion, pour toutes les populations musulmanes, y compris pour des communautés de langue maternelle autre que l'arabe, comme celles, originaires du golfe arabo-persique, de la côte indienne, qui ont pu se rencontrer aux Comores;

*iii) l'attitude des locuteurs Comoriens vis-à-vis de la langue arabe:*

vis-à-vis de l'arabe, les Comoriens ont une attitude bi-valente qui perdure, partagée entre une volonté de respect total de la forme, et un rejet comme étrangers de faits trop marqués comme arabes ...; cette attitude est la traduction linguistique du double souci culturel d'imitation, d'intégration dans le moule arabo-islamique et de différenciation, d'affirmation de la spécificité comorienne.

Cela nous amène à proposer une explication historique, qui est aussi une vision de la situation sociolinguistique d'où est sorti le shingazidja, qui postule une diglossie:

dans une situation où plusieurs communautés 'orientales' de langues et d'origine diverses étaient en présence, l'arabe, langue de la religion et du commerce, pratiquée à ce titre par l'ensemble de ces communautés, s'est imposé comme langue véhiculaire entre les différentes communautés immigrées, mais sans doute sans se substituer aux vernaculaires, et tout en intégrant du vocabulaire de ces langues; par ailleurs, ou de ce fait, l'arabe va servir de base lexicale principale dans les contacts entre ces communautés et les populations Africaines bantu, aboutissant ainsi au shingazidja, avant de disparaître lui-même avec le mélange de ces diverses communautés et leur "comorianisation" ... (voir toutefois nota ci-dessous sur une survivance de l'usage de l'arabe aux Comores).

Cela peut expliquer le nombre relativement important de mots d'origine persane présents en shingazidja, qu'ils soient entrés directement ou via l'arabe: si l'arabe a joué ce rôle de langue véhiculaire, on peut supposer que cet arabe véhiculaire était lui-même influencé par les langues avec lesquelles il était en

contact, à savoir persan, hindi, etc; mais cela n'exclut pas entièrement un contact direct entre persan, hindi, etc, et ce qui devait devenir le shingazidja, ce contact toutefois étant resté limité à certaines communautés, ou à certaines périodes.

Grâce à cette hypothèse, on peut proposer une explication pour certaines spécificités morphologiques des emprunts à l'arabe en shingazidja.

Le rôle religieux de l'arabe, son prestige, les contacts maintenus avec des populations arabophones, l'hétérogénéité linguistique elle-même, auraient aussi, dans ces conditions, empêché que l'arabe véhiculaire s'éloigne trop fortement de la langue littéraire ou classique.

Ainsi s'expliquerait que les emprunts arabes en shingazidja soient de fait proches de formes d'arabe classique ou littéraire - on n'a le plus souvent aucune difficulté, compte tenu des transformations dues au shingazidja, à retrouver les racines arabes sources.

Nota.

Signalons à ce propos qu'il semble bien que l'arabe qui est resté utilisé comme langue de communication aux Comores parmi l'élite lettrée jusqu'à ces dernières années présentait des variations non négligeables par rapport à l'arabe classique ou à d'autres variétés écrites; ces variations peuvent d'ailleurs simplement provenir d'influences comoriennes ... nous ne connaissons qu'une étude sur ce point, qui nous est malheureusement inaccessible; c'est une chronique arabe des Comores du 19<sup>e</sup> siècle, recueillie, reproduite, traduite et commentée par G. ROTTER (ROTTER 1976, référence aimablement communiquée par G. BOULINIER.)

Le problème des sources arabes : le choix de l'arabe classique

Dans notre travail, nous ferons référence quasi-exclusivement à l'arabe classique; à cela, deux ordres de raison, linguistique et pratique:

- la référence à l'arabe classique, au niveau phonologique comme lexicologique et morphologique apporte une réponse satisfaisante dans la plupart des cas (ou du moins, paraît en apporter une); nous en avons proposé une double

explication, sociologique et historique: l'arabe classique, ou coranique, a influencé et influence toujours directement le comorien, à travers la religion, l'école coranique, le droit, etc; nous avons vu par ailleurs que l'on pouvait faire l'hypothèse que le parler arabe au contact duquel le shingazidja s'est constitué n'était guère divergent d'une forme classique;

- s'il faut rechercher des influences dialectales, les parlers arabes du Yémen, et surtout de Mascate et d'Oman, paraissent géographiquement et historiquement les mieux placés pour avoir été en contact direct avec le shingazidja.

Mais l'état actuel de nos connaissances et la maigre documentation disponible sur ces parlers ne suffisent à étayer cette hypothèse.

A. S. G. JAYAKAR présente une description contrastive phonétique, morphologique et lexicologique de l'arabe d'Oman, où les niveaux de l'écrit et de l'oral ne sont pas toujours clairement distingués; sur le plan phonétique, il écrit *"With the following exceptions the sounds of the letters of the alphabet are generally the same as in standard Arabic"* (JAYAKAR 1889 : 651); les écarts indiqués ne correspondent à rien de précis en shingazidja; nous n'avons toutefois pas confronté son lexique aux données shingazidja, mais il semble que quelques mots au moins se retrouvent (selon A. LONNET, à qui nous devons de connaître cette référence).

Il est sûr qu'une étude prenant en compte les divergences dialectales de l'arabe de cette région à tous les niveaux, notamment sémantique, serait riche d'enseignement concernant la genèse du shingazidja mais aussi la reconstitution de l'histoire de cette période, mais cela dépasse largement nos compétences et sort de notre sujet.

Dans l'étude des emprunts orientaux en swahili, la question de la variété d'arabe auquel se référer s'est toujours posée avec acuité:

× les deux principaux ouvrages lexicographiques swahili disponibles, déjà relativement anciens, les dictionnaires de JOHNSON de 1939 et de SACLEUX

publié en 1939 également mais rédigé au début du siècle, contiennent des indications étymologiques:

JOHNSON fait référence à l'égyptien et à l'arabe classique *"The Egyptian dialect of Arabic appears to have had a distinct influence on the Swahili vocabulary as well as the classical Arabic of the Zanzibar Arabs (JOHNSON 1939 : viii)*; la référence à l'égyptien ne paraît pas convaincante hors quelques mots, comme nous le verrons.

Nous avons vu la prudence avec laquelle il convenait de regarder ses étymologies.

SACLEUX se rapporte au dialecte de Mascate: *"les correspondances (entre arabe et swahili) ont été établies conformément à la prononciation du dialecte de Mascate, qui est celui auquel les Swahilis ont fait le plus d'emprunts"* (op. cit., p 31); sa bibliographie ne fait état d'aucune source dans ce domaine; nous avons constaté que les étymologies arabes qu'il propose ne se retrouvent pas toujours dans les dictionnaires classiques.

× Sh. M. ZAWAWI se réfère également à l'arabe d'Oman et de Mascate, et s'appuie sur sa propre compétence de locuteur bilingue pour déterminer les emprunts du swahili (*"(...) it appears that the major donor of recent 'loan' words is the 'Umāni dialect of Arabic. (As a bilingual speaker of Swahili and the 'Umāni dialect of Arabic, I propose to use my own knowledge to identify those Swahili words which appear to have come from this source since there are few linguistic publications available.)"*) (ZAWAWI 1979 : 3); mais cela l'amène, sans autre justification que l'affirmation d'une forme selon elle ressemblante dans le parler d'Oman, hors toute vérification possible, et sans discussion, à considérer comme emprunts des termes bantu bien établis, répandus dans tout le domaine, cela n'incite pas à utiliser son corpus (nous avons discuté ce point dans un compte-rendu de cet ouvrage, LAFON 1982).

× POLOME, dans sa description du swahili, assume cette faiblesse: *"In mentioning the sources of the Arabic loans in Swahili, reference is made to a tentatively phonemicized form of Modern Written Arabic. Though historical events make it probable that the dialects of Oman (Muscat) and Mehri (Hadramaut) supplied most of the Swahili words of Arabic origin, it is practically impossible to refer to the forms of these dialects for lack of a reliable phonemic analysis of the same"* (POLOME 19867 : 198, note 144)

× BALDI, dans sa thèse déjà citée, se réfère exclusivement à l'arabe classique (avec comme référence lexicale principale le dictionnaire de H. WEHR, mais il recourt aussi à BELOT, DOZY, KAZIMIRSKY), pour des raisons pratiques: *"pour le swahili, (...) les influences sont venues directement de la péninsule arabe, donc il faudrait être familiarisé avec le sudarabique et les dialectes arabes parlés sur la côte de l'Oman jusqu'au Yémen, ou de l'autre route, de l'Égypte, souvent à travers d'autres langues de populations islamisées. (...) C'est à cause de l'ampleur des problèmes soulevés que l'on va ici limiter l'analyse des correspondances (...) avec l'arabe classique."* (BALDI 1982 : tome 1, p 14-15).

× Quant à nous, nous nous référons en premier lieu, nous l'avons dit, à l'arabe classique, à l'aide de WEHR pour le lexique et de BLACHERE & GAUDEFROY-DEMOMBYNES pour la grammaire, tout en faisant occasionnellement quelques excursions dialectales grâce à DOZY notamment, qui permettent d'expliquer certaines formes shingazidja.

Le choix de limiter l'étude différentielle à l'arabe classique procède donc d'un problème de source et d'un souci de prudence; il n'est certes pas satisfaisant, mais, comme nous l'avons déjà remarqué, soit simple coïncidence, soit confirmation de l'hypothèse sur la prédominance de l'arabe classique dans la constitution du shingazidja, il permet de résoudre la plupart des cas.

Il paraît par contre assuré qu'il n'y a pas lieu de rechercher une éventuelle influence directe du sud-arabique: aucune ressemblance frappante n'apparaît dans la confrontation du shingazidja et des listes lexicales de sud-arabique établies par JOHNSTONE (référence signalée également par A. LONNET).

#### Remarques sur les étymologies orientales et arabes

Des remarques d'ordre formel doivent être faites à propos de la recherche étymologique sur les langues orientales et particulièrement sur l'arabe, la plus importante et aussi celle qui présente le plus de difficultés méthodologiques.

Nous ne sommes au fait d'aucune de ces langues; nous sommes donc entièrement tributaire des sources et de notre capacité à les consulter: cela nous restreint, nous l'avons déjà signalé à propos de l'arabe, aux dictionnaires bilingues; par ailleurs, la recherche étymologique d'un mot shingazidja soupçonné d'être d'origine orientale suppose de le référer à toutes les séquences phoniques qui peuvent en avoir été la source jusqu'à trouver: du fait des modifications susceptibles de se produire en shingazidja par rapport aux langues orientales sources, on ne peut être certain d'avoir épuisé, pour un mot donné, les possibilités.

Ces deux raisons conjuguées font qu'il est clair que les étymologies proposées ne le sont que provisoirement, et que des spécialistes des domaines concernés seraient amenés sans nul doute à en corriger de nombreuses, ainsi qu'à en identifier sans peine parmi les cas douteux.

Pour ce qui est plus précisément de l'arabe, où notre référence est WEHR, il faut rappeler, en plus de la limitation tenant à la variété de langue choisie, que le dictionnaire n'épuise pas tous les mots formés sur une racine donnée: certains noms verbaux (*masdars*) existent mais ne sont pas donnés parce qu'ils sont de formation et de sens réguliers, etc; nous verrons les problèmes spécifiques liées au mode de citation des verbes; en outre, les dialectes présentent fréquemment des variations dans la vocalisation par rapport aux formes littéraires: cela explique des cas où l'on peut identifier de façon très vraisemblable la racine arabe d'où provient le mot shingazidja mais sans trouver le terme arabe précis; le mot arabe proposé comme étymologie (le plus rapprochant) présentera alors des divergences avec le mot shingazidja.

Enfin, nous suivons les conventions de transcription de WEHR - ce qui est très commode mais présente certains inconvénients (voir étude syllabique).

Nous signalerons dans le cours de l'étude les problèmes particuliers soulevés et les précautions à prendre à cet égard.

#### Autres langues sources d'emprunts

Le shingazidja présente aussi des emprunts à d'autres langues que les langues orientales, avec lesquelles il s'est trouvé ou se trouve toujours en contact: langues européennes (portugais, français), malgache, peut-être indonésien ou proto-malgache.

\* langues européennes: on trouve, en shingazidja, un nombre non négligeable d'emprunts lexicaux, au portugais et surtout au français.

\* portugais :

par leur nombre limité et la situation - ils n'ont pas bénéficié d'un appui culturel comparable à celui qui a accompagné les emprunts à l'arabe - les emprunts au portugais, quoique sans doute relativement anciens, n'ont jamais pu influencer profondément sur la langue; certains auteurs considèrent d'ailleurs qu'ici aussi l'arabe a joué ou a pu jouer le rôle de langue de passage (ZAWAWI, ALEXANDRE, déjà cités) (la même question se pose d'ailleurs pour quelques emprunts au français, dont le terme même, Ng *farantsa*, < fr. *france* ou ar. *faransä*).

Nous avons donc suivi la même démarche que pour les emprunts orientaux, et ne faisons figurer la forme portugaise que si une forme source acceptable n'est pas attestée dans les dictionnaires arabes consultés.

\* français :

Quant aux emprunts au français, quasi-contemporains dans leur grande majorité, ils manifestent, dans une certaine mesure, une intégration ou une adaptation seulement partielle aux canons du shingazidja; mais ils atteignent le shingazidja quand celui-ci a déjà été façonné par

l'influence arabe, et n'apportent pas d'innovations ou de transformations d'importance.

Dans ces conditions, français comme portugais ne figurent comme langues sources que dans le lexique.

Nous nous en tenons à la variété recueillie en Grande-Comore: on constate, parmi la communauté comorienne installée en France, surtout chez les jeunes, le développement d'une variété de shingazidja très influencée par le français ... mais nous doutons que cette influence se répercute de façon sensible à la Grande-Comore, hors bien entendu l'introduction d'emprunts en plus ou moins grand nombre.

Signalons, pour l'anecdote, que la recherche de l'origine première pourrait entraîner des résultats pour le moins paradoxaux: ainsi, parmi les emprunts du shingazidja au français, se trouve-t-il des termes d'origine arabe (que le français a lui-même empruntés à l'arabe, directement ou non): Ng gudro, *route bitumée* du français *goudron*, venant de *l'arabe d'Egypte qatrān*" (DAUZAT 1969).

Indonésien ou malgache: une influence indonésienne dans la constitution même de la langue est historiquement possible (l'escale sur l'archipel des populations d'origine indonésienne qui se fixèrent à Madagascar est au moins plausible); toutefois, si tel est le cas, les mots sont intégrés à la langue au point de n'être pas décelables immédiatement par leur comportement phonétique ou morphologique ... cette question est liée aux hypothèses concernant le substrat bantu à Madagascar, et la formation de la langue malgache (voir DAHL, à paraître, et SIMON, 1987); nous espérons pouvoir y revenir dans un autre travail.

Par contre, dans une époque historique, les contacts entre les Comores et Madagascar ont été patents, en particulier au XIX et XX siècles, avec la présence à Madagascar d'une forte communauté comorienne, dont une partie est rentrée récemment (1976): cela a favorisé la pénétration en shingazidja d'emprunts lexicaux au malgache: en l'état de nos connaissances, nous nous contenterons donc de relever ces emprunts, sans aller plus avant.

Des index étymologiques regroupent par langue toutes les formes citées (mots ou racines pour l'arabe); ils sont suivis d'index de cas douteux, et d'étymologies orientales inconnues.

## CORPUS

## LA LANGUE DECRITE

*Niveau de la langue décrite; phonostyle*

D'une façon générale, toute langue est constituée d'une multitude de 'styles', reflétant en partie l'appartenance sociale; le style de langue utilisé est un des facteurs permettant aux sujets de se situer les uns par rapport aux autres dans la hiérarchie sociale, et, de ce fait, la langue est investie par le locuteur d'une fonction d'identification sociale.

Pas plus aux Comores qu'ailleurs, il n'y a adéquation automatique entre appartenance sociale et niveau ou style de langue utilisé, et les différentes pratiques linguistiques auxquelles nous nous référons en terme de niveau ou de style ne sont pas des réalités hermétiques: chaque locuteur grand-comorien, comme tout locuteur de n'importe quelle langue, est à même d'adapter, dans une certaine mesure tout au moins, sa pratique à la situation précise de communication qu'il expérimente, ou, pour le dire autrement, dispose d'un éventail de styles entre lesquels il choisira, plus ou moins librement, selon la perception qu'il a de la situation où il se trouve, selon sa stratégie sociale, etc: dans cette mesure, la pratique langagière est un des lieux sans doute privilégiés où peuvent s'inscrire de façon consciente et voulue des choix ou des préférences idéologiques individuels, permettant au sujet d'affirmer son identité - étant entendu que l'intercompréhension reste toujours réalisée dans l'ensemble, pour autant, comme on dit, qu'il y ait jamais communication.

Aux Comores, le clivage essentiel, dans l'appréciation de la pratique linguistique, oppose un parler 'arabisé', à un parler qui ne l'est pas ou qui l'est moins: en schématisant, on peut définir l'un comme le parler des 'lettrés', l'autre, comme le parler des campagnes: le premier contient un nombre maximum de mots d'origine arabe 'savants', prononcés d'une façon qui se veut conforme au modèle étymologique, il est caractéristique de citadins 'lettrés'; le second contient également des mots d'origine arabe - le shingazidja comportant de toute manière des emprunts à l'arabe - mais en moins grand nombre et, surtout, mieux intégrés phonologiquement, c'est le parler des 'paysans': il existe donc une variété qui peut être qualifiée de 'soutenue' eu égard aux réalisations des phonèmes arabes, mais elle ne s'assimile pas à la langue des 'lettrés': elle se situe entre ces deux pôles théoriques, dont on doit comprendre toutefois qu'ils ne sont postulés que par commodité.

Dans une telle situation, en effet, le style de langue joue pleinement son rôle d'identification sociale, et se prête aux diverses stratégies individuelles: ainsi, dans certaines situations, des jeunes issus de milieux traditionnels, eux-mêmes plutôt 'lettrés', utiliseront-ils par réaction un parler volontairement dépouillé de tout ce qui leur apparaît marqué comme 'arabe', c'est-à-dire reflétant leur origine sociale, alors que d'autres, de milieux modestes, emploieront, dans un but d'ascension sociale, un langage truffé d'emprunts et de prononciations arabisées (parfois mal à propos: on parlera alors de 'sur- ou d'hyper-correction'); cette latitude sur le plan de la langue va de pair avec le fait que, à l'heure actuelle, l'existence et l'autonomie de la culture comorienne vis-à-vis de l'arabe ne sont pas complètement ressenties.

On peut sans doute évoquer, de ce point de vue, l'émergence progressive du 'vulgaire' français au Moyen-Age vis-à-vis du latin: pendant une longue période, tout texte français pouvait, semble-t-il, accueillir *ad libitum* des mots ou expressions latines, des formes latines ('savantes') ... on sait ce

qu'il en advint, le français actuel montrant non seulement les évolutions phonétiques normales à partir des étymons latins, mais ayant aussi accueilli moult formes introduites à diverses périodes par des grammairiens, hors de l'évolution normale ...

Ces facteurs sont à l'origine des variations nombreuses et constantes en shingazidja, vis-à-vis des emprunts à l'arabe: au niveau lexical, il existe souvent des synonymes, l'un des termes étant d'origine arabe, l'autre de fonds bantu, au plan phonique, une grande part des phonèmes spécifiques au lexique arabe sont tantôt réalisés conformément à l'étymologie, tantôt ramenés à des phonèmes présents dans le fonds bantu.

Il est intéressant, sur ce point, de se référer à la notion de "phonostyle", que nous reprenons à DRESSLER & HUGFARD. Ceux-ci, dans un article concernant la phonologie du breton sud-bigouden, distinguent quatre 'phonostyles', nommés d'après une terminologie musicale, qui s'ajoutent aux faits relevant de la distribution géographique et lexicale; la plupart des locuteurs sont à même de passer d'un phonostyle à l'autre, selon la situation, etc:

*L 'Lento', forme appartenant à un style très formel et prononcée lentement;*

*And 'Andante', forme appartenant à un style intermédiaire entre L et F;*

*A 'Allégo', forme appartenant à un style relativement négligé, intermédiaire entre And et F;*

*'Presto', forme appartenant à un style négligé de débit très rapide.*

(DRESSLER & HUGFARD 1980 : 8-10)

DRESSLER & HUGFARD proposent quatre analyses phonologiques, chacune correspondant à un phonostyle défini, avec des règles de passage de l'un à l'autre.

Cette méthodologie permet une prise en compte de la variation phonologique, et s'approche ainsi la description de la réalité perçue; il est important de noter

que DRESSLER & HUGFARD ne posent pas les différents phonostyles comme particuliers à des groupes sociaux ou des individus, mais les inscrivent dans

l'interaction sociale; il s'agit d'une grille d'analyse, chaque forme donnée étant rapportée à tel ou tel phonostyle et un même énoncé pouvant contenir des formes relevant de différents phonostyles.

Ce cadre peut s'appliquer au shingazidja pour rendre compte de la variation phonologique.

Si l'on identifie la variété la plus arabisée au phonostyle 'lento', la variété décrite nous paraît se situer entre les phonostyles 'lento' et 'andante'; elle correspond à celle pratiquée par des personnes ayant suivi, dans leur jeunesse, au moins le premier niveau de l'école coranique (*paa la shio*): cette formation leur permet de manier, avec plus ou moins de facilité selon la pratique qu'ils en ont, la graphie arabe du comorien, de lire l'arabe coranique, mais ne leur donne pas une maîtrise réelle de l'arabe proprement dit, que ce soit à l'oral ou l'écrit.

La plupart des Grand-Comoriens répondent à cette définition, du fait de la fréquentation quasi-générale du premier niveau de l'école coranique.

Cela exclut d'un côté les paysans ou les travailleurs agricoles, originaires de régions reculées de l'intérieur de la Grande-Comore, et d'un autre, les 'lettrés', qui possèdent une connaissance plus approfondie de l'arabe, acquise dans les degrés supérieurs de l'école coranique, en famille ou ailleurs (à l'étranger, etc).

Comme nous nous attachons à la description d'un phonostyle particulier, il nous a paru important que ce phonostyle reflète, non pas nécessairement l'usage statistiquement le plus fréquent, mais un usage admis, ou susceptible de l'être, par le plus grand nombre - soulevant le moins d'objections de la part des locuteurs: il nous a paru que tel était le cas du phonostyle 'soutenu' que nous

avons donc retenu; ce qui n'appartient pas à ce phonostyle est traité comme écart, ou variation - imputable à des interférences d'autres phonostyles.

Nous considérons comme des variantes phonostylistiques les réalisations concurrentes (c'est-à-dire, ressortissant d'autres phonostyles, qu'il s'agisse du phonostyle 'arabisé' ou de phonostyles peu ou pas 'arabisés' - ne retenant que peu ou pas de réalisations spécifiques aux emprunts arabes).

### *Composition du corpus*

L'enquête linguistique a été effectuée d'une part sur le terrain, en 1976 et 1977-8, principalement à Moroni et dans quelques villes côtières (Mitsamiuli, Mbeni, Ikoni), auprès de jeunes scolarisés, et d'autre part, en France grâce à l'aide amicale d'étudiants Grand-Comoriens de diverses origines (rurale comme citadine).

Le corpus provient d'enregistrements divers (contes, récits, discours politiques, recettes de cuisine, etc), de rencontres et d'enquêtes auprès de groupes sociaux spécifiques (artisans, cultivateurs, etc - trop peu nombreuses malheureusement), directes ou par l'intermédiaire de nos informateurs, sans négliger les textes écrits de tous ordres dont nous avons pu avoir connaissance (correspondances privées, tracts politiques, récits traditionnels, etc).

D'une façon générale, toutes les informations recueillies sur le terrain ont été vérifiées en France, mais nous n'avons admis aucune information nouvelle recueillie en France qui n'ait été confirmée par plusieurs sources 'indépendantes' (personnes sans liens familiaux).

Notre enquête sur le terrain remontant à la période immédiatement postérieure à l'indépendance et nos informateurs en France ayant effectué leurs années de lycée en règle générale avant l'indépendance, ce travail en principe n'a pas été

concerné par les modifications du système scolaire introduite depuis, en particulier par la diffusion de l'enseignement de l'arabe et la multiplication des bourses d'étude dans des pays arabes ...

Nos informateurs directs ont été essentiellement des jeunes scolarisés, connaissant le français, ayant fréquenté le lycée de Moroni: tous ont été, pendant quelques années au moins, en contact avec le parler de Moroni - qui est aussi, plus ou moins, celui de la radio.

Ces jeunes sont le plus souvent issus de familles bourgeoises (fonctionnaires religieux ou civils) plutôt que de travailleurs manuels, mais le cloisonnement social n'est pas tel à la Grande-Comore qu'il empêche les contacts entre personnes de milieux ou de générations différents; dans l'émigration, ces contacts sont souvent favorisés par les associations de nature diverse, l'entraide, le militantisme, toutes choses qui entraînent un relatif brassage social et régional.

#### *Norme ou forme de référence*

Dans un travail à visée lexicographique, portant sur une langue peu décrite et ne disposant pas de tradition écrite, il est clair que, à travers la variété décrite, le style, le niveau de langue et l'orthographe retenue, nous proposons, au moins implicitement, une forme de référence; toutefois, cette forme de référence ne saurait être dotée d'une valeur normative, visant, par exemple, à unifier le shingazidja sur cette base (nous espérons qu'elle n'aura pas non plus d'effet répulsif). Cela n'entre ni dans nos compétences, ni dans nos intentions.

Nous avons simplement recherché des formes médianes, reconnues et acceptées par le plus grand nombre, mais cela n'invalide aucunement des prononciations, des mots, ou des constructions à diffusion régionale (que nous mentionnons

d'ailleurs, le cas échéant); l'orthographe permet, en principe, de rester proche de la prononciation; il reste cependant que des choix, en définitive arbitraires, au niveau de la formalisation, sont nécessaires.

Nous croyons donc avoir saisi, du double point de vue du niveau et de l'état de langue, une variété 'moyenne', que l'on peut définir négativement comme n'étant ni (i) une variété régionale particulière, ni (ii) une variété sociale ('sociolecte'), qu'il s'agisse de classe d'âge ou de classe sociale à proprement parler, ni (iii) une variété datée (un état de langue particulier).

Bien entendu, lorsqu'une forme est sentie expressément comme relevant d'un registre particulier (géographique, social ou historique), nous l'indiquons.

## **CONVENTIONS DE NOTATION**

*Où l'on dit ce que signifie comment l'on écrit*

## CONVENTIONS DE NOTATION ET DIVERS

1) Shingazidja1-1) Signes*a) symboles orthographiques*

Ainsi que nous l'avons dit, nous attachons une grande importance à la perspective d'instrumentalisation de la langue.

Nous nous situons d'emblée dans une perspective orthographique: en l'absence d'orthographe officielle ou standard, nous utilisons le système orthographique proposé pour les parlers comoriens par AHMED-CHAMANGA, LAFON & SIBERTIN-BLANC, basé sur une étude phonologique et morpho-phonologique (AHMED-CHAMANGA, LAFON & SIBERTIN-BLANC, à paraître).

Tout en reprenant pour l'essentiel les conventions graphiques communément répandues dans le public comorien, ce système permet une notation quasi-phonologique; il s'éloigne du système international pour des raisons pratiques, par l'usage exclusif de caractères standards (clavier usuel), et pour éviter l'usage de symboles trop éloignés de leur réalisations, ce qui amène à utiliser des digraphes et à adopter certaines conventions spécifiques.

Visant à être un système orthographique, il s'éloigne éventuellement d'une transcription phonétique ou phonologique pour mettre en évidence certaines relations morphologiques.

L'utilisation de ce système dès le début de notre travail et dans son ensemble présente l'intérêt essentiel d'avoir, dans tous les cas, une orthographe identique, étant entendu qu'il nous paraît que le lexique doit être transcrit dans un système à vocation orthographique. Toutefois, quand les nécessités de l'analyse le demandent, cette orthographe est accompagnée par un système de

transcription utilisant les signes de l'IAI, donnés entre crochets carrés []  
(d'après BOUQUIAUX-THOMAS (1976)).

Pour faciliter la lecture, figure, sur un feuillet mobile, l'ensemble des  
graphèmes, avec leur valeur phonétique, et, pour les phonèmes 'arabes',  
l'indication de leur source.

#### correspondances avec les symboles phonétiques

##### 1) articulations simples:

p = [p]	
b = [b] (implosive), [b]	
t = [t] (dentale)	
ts = [ts]	pv = [β]
dz = [dz]	w = [w]
tr = [ʈ] (alvéolaire)	l = [l]
d = [ɖ] (alvéolaire, rétroflexe), [d]	r = [r]
tsh = [tʃ]	y = [y]
dj = [dʒ]	h = [h]
k = [k]	m = [m]
g = [g]	n = [n]
f = [f]	ny = [ɲ]
v = [v]	
s = [s]	
z = [z]	
sh = [ʃ]	
th = [θ] ([s])	
dh = [ð] ([z])	
kh = [χ] ([h])	
gh = [ɣ] ([ɦ])	

##### 2) articulations complexes:

**nc** = toutes mi-nasales ou prénasalisées, indépendamment du point  
d'articulation de la consonne - c'est la différence majeure entre notre  
système et la graphie traditionnelle; cela permet de distinguer dans  
l'écriture prénasalisée et nasale syllabique devant labiale, notée  
toujours **m**:

**np** = /<sup>~</sup>p/ # **mp** = /m-p/ (en graphie traditionnelle, une seule  
notation, **mp**);

**nb** = /<sup>~</sup>b/ # **mb** = /m-b/ (en graphie traditionnelle, une seule  
notation, **mb**) (voir en fin de paragraphe un cas particulier)

**nd** = /<sup>~</sup>d/ (en graphie traditionnelle, **nd**) ; **nf** = /<sup>~</sup>f/ (en graphie  
traditionnelle, **nf**), etc

À noter :

\* **npb** = /<sup>~</sup>b/ (& /<sup>~</sup>β/); la graphie traditionnelle n'a pas toujours  
reconnue la spécificité de cette articulation, notée tantôt **mb**,  
(comme la prénasalisée implosive et la séquence nasale syllabique  
+ implosive), tantôt **pb** (sans indice nasal); la graphie \***npv**,

quoique justifiée phonologiquement (ainsi que morphologiquement dans certains cas) paraissait trop éloignée de la réalité phonétique (explosive sonore); nous avons donc préféré **npb** qui s'apparente à la graphie déjà proposée **pb** tout en la normalisant (avec l'indication explicite de l'articulation prénasalisée).

\* **ndr** = /~r/, réalisé [~d~] : nous avons suivi ici la réalisation  
**Cw** : vélarisées : **pw** = [p<sup>w</sup>], **bw** = [b<sup>w</sup>], **dhw** = [δ<sup>w</sup>], etc

iii) hors-système:

' (apostrophe) = ['], attaque vocalique (provenant de ayn ou hamza arabe)

^ (accent circonflexe) = [ˆ], attaque vocalique et [˜], nasalisation de la voyelle, ces deux éléments, provenant de ayn arabe, étant souvent associés

**m** = [m], nasale syllabique (toujours devant consonne); notre notation, nous l'avons signalé, distingue la nasale syllabique (toujours labiale, **m**) et la nasale homorganique (toujours notée **n**), qui représentent des réalités phonologiques et morphologiques différentes.

**b) élocution délibérée**

Nous nous sommes placés ici, en principe, dans la situation dite 'd'élocution délibérée', qui est celle qui convient le mieux à un travail à visée lexicographique.

Nous avons raisonné sur des mots, parfois des monèmes, en isolation; cela suppose un découpage de l'énoncé faisant apparaître les unités lexicologiques (les mots) indépendamment de tous les faits de contraction facultatifs, très fréquents en shingazidja, même en débit normal (chute de voyelles ou de syllabe, coalescence non lexicalisée, etc): nous restituons alors en principe les formes intégrales, attestées en répétition lente, et confirmées par comparaison avec des cas où, pour des raisons d'opportunité (contexte phonétique différent), il n'y a pas contraction:

[tsamboka:djahunu] = **tsianba** (ou **tsanba**) **uka hadja hunu**, j'ai dit qu'il était venu ici:

[tsanb..] = [tsi-anb..], pr.pers. 1<sup>er</sup> sg. - racine verbale, 'dire' : [i + a] > [a]

[..anbokal = l..anb-a#uka], racine verbale, 'dire' - suffixe - conj. 'que' : [a + u] > [o]

uka, 'conj. que', se décompose à son tour en l(h)u-k-a], préf. cl. 15 - rac. verbale 'être' - suffixe; mais cela n'interfère pas ici.

[..ka:djal = l..k-a#ha-dj-a], conj. 'que' - il/elle est venu/e : [a + ha] > [a:]

[hafanyezembil = hafanya ezenbi, il a fait mal (les mauvaises [choses])

[enyumba:hangu] = enyunba ya hangu, ma maison

L'allongement du [a] final de nyunba est une indication de la présence de l'extra-préfixe de dépendance, ya.

Par contre, lorsque la forme entière supposée n'est pas attestée, même en débit lent, même si des arguments morphologiques viennent à l'appui de la reconstruction, nous n'effectuons pas de restitutions orthographiques: la contraction est alors considérée comme lexicalisée, et nous ne proposons éventuellement de décomposition que dans le cadre de l'analyse structurelle:

wontsi, eux tous

[wa-ontsil, préf. cl. 2 - thème 'tout' → [wontsil, jamais \*[waontsil]

(malaho) mema, de bonnes (maisons)

[ma-ema] →, préf. cl. 6 - thème 'maison' → [memal, jamais \*[maema]

D'autre part, dans les cas où forme contractée et forme entière sont également admises, lorsqu'il n'apparaît pas d'impératif justifiant une restitution systématique de la forme entière - en particulier dans des cas de rencontre entre voyelles entraînant ou pouvant entraîner coalescence - nous formalisons l'une ou l'autre, étant entendu que la forme de référence, celle figurant comme entrée lexicale, reste la forme développée:

faili ou feli, caractère

(embe) kaitsina .. ou ketsina, (la vache) n'a pas de ..

[ka-yi-tsi-na], négatif - préf. cl. 9 - négatif - particule de liaison

tsihisa tsili ou tsisa tsili, j'ai fini de manger

[tsisa] = [tsi-his-a], pr.pers. 1<sup>er</sup> sg. - racine verb., 'finir' - suffixe : la première syllabe de ce verbe chute très fréquemment lorsqu'il est employé comme auxiliaire; par ailleurs, les syllabes [hV], d'une façon générale, sont peu stables.

De la même façon, l'écriture reproduit certaines variantes phonostylistiques

(cf. ci-dessus, p 63); un exemple permettra de présenter le problème:

• soit (dans notre analyse) la réalisation notée kh (spécifique du lexique arabe), correspondant à ar.  $\text{خ}$  k); ce phonème, réalisé [χ], est une caractéristique du phonostyle soutenu; il est absent d'autres phonostyles, où l'on a, dans les mots où il apparaîtrait, une réalisation [h], qui, en outre, dans tous les phonostyles, la réalisation du phonème /h/, lui-même commun aux mots de fonds bantu et aux mots d'origine arabe:

ex: [χat<sup>w</sup>aril], danger dans le phonostyle soutenu, correspondant à [hat<sup>w</sup>aril] dans un phonostyle moins soutenu (ar. kat.ar.), ce que l'on peut ainsi représenter:

	phonostyle soutenu	phonostyle non soutenu
ar. k →	[χ]	
ar. h, h. BC *k →	[h]	[h]

• Inversement, l'opposition entre les deux phonèmes arabes k et q n'est pas maintenue en shingazidja dans le phonostyle considéré, les deux étant réalisés [k] (ce qui rejoint un phonème de fonds bantu); on peut toutefois rencontrer une réalisation [q] dans la langue des lettrés:

ex: [kaulul], parole dans le phonostyle soutenu, correspondant à [qaulul] chez les lettrés:

	phonostyle lettré	phonostyle soutenu
ar. q →	[q]	
ar. k ; BC *k →	[k]	[k]

La variation s'interprète alors comme le passage d'un phonostyle à un autre; ce passage peut intervenir sans que le locuteur en soit toujours conscient, selon la situation, ou même dans un énoncé suivi.

Notre transcription correspond au phonostyle retenu: cela permet une différenciation maximum dans l'écriture; nous admettons toutefois des variantes orthographiques pour refléter des variantes fréquentes de réalisation: ainsi, on écrira *khatwari* ou *hatwari*, mais toujours *kaulu*.

(ces formes sont données comme variantes dans le lexique, l'entrée étant faite à la forme caractéristique du phonostyle de référence)

## 1-2) Mode de citation des unités

### *a) mode de citation des lexèmes (éléments abstraits)*

Les gloses françaises doivent indiquer non seulement le sens des lexèmes mais aussi leur statut morphologique respectif; or, il n'existe pas en français de

catégorie grammaticale analogue, le français ne fonctionnant que partiellement à partir de racines (par ailleurs, celles-ci posent des problèmes spécifiques de définition qu'il n'est pas dans notre propos d'envisager).

Nous avons donc recours à un artifice: nous insérons les gloses de racines entre barres verticales (ce qui indique leur statut d'élément construit), précédées du terme de métalangue "*notion de ..*" suivi d'un infinitif; par convention, cela renvoie à l'ensemble des verbaux et des nominaux notionnellement reliés au mot français évoqué:

ainsi, *[notion de manger]* renvoie potentiellement non seulement à "manger, faire manger, etc", mais aussi à "nourrir, nourriture, etc", toutes notions véhiculées en shingazidja par le même lexème, *[-l-]*;

*[-pih-]*, *[notion de cuisiner]*, renvoie non seulement à "cuisiner, faire la cuisine", mais aussi potentiellement à "cuisinier, mesure pour le riz", etc.

Nous utilisons le même procédé pour les thèmes, avec un nom français ou un terme en fonction nominale, au lieu de l'infinitif, en nous efforçant de sélectionner un mot qui n'exclue aucune des potentialités sémantiques impliquées par le thème grand-comorien:

ainsi, *[notion d'existant]* renvoie tant à "homme" qu'à "chose", et glose le thème *[-ndru]*;

*[notion de maladie]* renvoie tant à "malade" qu'à "maladie", et glose *[-wadel]*, etc.

#### **b) mode de citation des constituants syntaxiques**

Bien que les constituants syntaxiques relèvent des catégories de discours, et qu'il s'agisse donc d'éléments moins abstraits que ceux relevant des catégories de langue, il n'est pas toujours possible de les citer sans avoir recours à des conventions de notation, inspirées de celles adoptées pour les monèmes, et qui correspondent aux usages en vigueur en bantouistique.

Nous rappelons que les caractères gras sont réservés aux constituants syntaxiques.

• nominaux:

i) nominaux indépendants : les noms (nominaux indépendants) sont cités tels qu'ils apparaissent dans l'énoncé, en principe au singulier; le cas échéant, le pluriel (manifesté par un changement de préfixe) est indiqué à la suite de la forme singulier par la mention du préfixe entre parenthèse suivi d'un tiret; lorsque des modifications d'ordre morpho-phonologique non systématiques (alternance consonantique, coalescence entre voyelles, etc) se produisent entre singulier et pluriel, telles que le pluriel ne puisse se déduire de la forme singulier par le simple changement du préfixe, le pluriel est donné dans son intégralité.

Alors même que l'alternance entre occlusive et continue à l'initiale lexicale est la règle dans le fonds bantu (même si elle supporte des exceptions), la notation ne se fonde pas sur ce mécanisme; l'alternance dans les nominaux sera indiquée expressément, alors que la non-alternance (donc, la présence d'une occlusive après préfixe syllabique) n'entraîne pas de mention spéciale; cette manière de procéder paraît en effet plus simple (seul le changement effectif est noté), et s'applique aussi sans difficultés aux emprunts, où l'on a vu que l'alternance n'opérait qu'exceptionnellement.

Exemples:

x pluriel obtenu par simple remplacement du préfixe:

**bwibwi** (ma-), araignée : pluriel **mabwibwi**

**apishi** (wa-), cuisinier : pluriel **wapishi**

**mbibo** (mi-), cajoutier : pluriel **mibibo**

x pluriel entraînant modification de l'initiale lexicale:

**daho** (**malaho**), maison : pluriel indiqué **malaho**

**ndime** (**nyilime**), langue : pluriel indiqué **nyilime**

x coalescence entre voyelle du préfixe et initiale radicale:

**mwana** (**wana**), enfant : pluriel indiqué **wana**

ii) nominaux dépendants : dans le discours, les adjectifs (nominaux dépendants) sont toujours accordés "en classe" à un nominal; la forme de citation neutralise cet accord, en présentant l'adjectif sans préfixe, précédé d'un tiret, comme pour le thème:

-wade, malade

-le, grand

L'adjectif peut aussi être cité dans une classe donnée (la classe est alors indiquée, soit par son numéro, soit par la présence entre parenthèse d'un nominal-type appartenant à cette classe):

**mwade**, malade (cl.1) ou (**ndru**) **mwade**, (une personne) malade

Utilisation de caractères gras ainsi que la glose distinguent alors

l'adjectif du thème:

notion de maladie!, thème : → d'où l'adjectif **-wade**, accordable dans différentes classes, ainsi que le nominal indépendant **uwade**, maladie;

notion d'éloignement!, thème : → d'où l'adjectif **-le**, grand, ainsi que le nominal indépendant **ule**, longueur, etc

Il est paru nécessaire de définir une forme verbale particulière, qui renvoie

à l'hypothèse à l'ensemble des formes verbales actualisées ou actualisables

du thème donné: cette forme est constituée de la racine, éventuellement

de l'extension, adjointe du suffixe verbal le plus fréquent et le plus

grammaticalement (-a), précédée d'un tiret qui rappelle la présence

de préfixes pour obtenir une forme réalisée; cette forme est

la "forme de citation"; on glose traditionnellement cet ensemble par

l'adjectif français:

**la**, cuisiner

**manger**

**la**, faire

La forme, construite, est toutefois identique sur le plan segmental à

la forme réelle, celle de l'impératif affirmatif de 2<sup>e</sup> pers. sg., qui est la

forme verbale minimale:

**cuisine !**

**mange !**

**fais !**

On rapproche aussi de la forme nomino-verbale (infinitif) qui consiste

en la forme minimale munie du préfixe nominal de cl. 15 (h)u-:

**la**, cuisiner; **hula**, manger; **ufanya**, faire

On distingue

**la** (notion de cuisiner! : racine verbo-nominale

**cuisine !** : impératif

**cuisiner** : infinitif

**la** cuisiner : forme de citation

### Conclusion de l'arabe

Dans les parties de ce travail, nous aurons fréquemment recours à des

exemples arabes.

Par souci d'homogénéité, et à un signe près, nous utilisons dès à présent le système de transcription du dictionnaire de Hans WEHR (1976) à partir duquel nous citons les étymologies arabes, tel qu'on peut le reconstruire (il ne figure pas de façon autonome dans le dictionnaire); la seule divergence entre la transcription de H. WEHR et la notre concerne le ayn, que, pour des raisons pratiques, nous transcrivons au moyen de l'épsilon ε (au lieu de ' ).

Les nécessités techniques nous ont en outre amené à apporter les aménagements suivants au système de transcription de WEHR:

C. = Ç (pour les emphatiques et ḥ ʕ) ; g. = ḡ (ġ)  
 ṣ = š (ar. ش)

- le tréma remplace le tiret droit sur les voyelles longues:

ā = ā̄, etc

Toutefois, pour faciliter les comparaisons phonétiques avec le shingazidja, nous avons recours, le cas échéant, aux les symboles utilisés par AL-ANI, qui sont plus proches des normes phonétiques (EL-ANI 1979 : 29). Pour faciliter la lecture, nous indiquons aussi, éventuellement, les graphes arabes originaux à côté de leur transcription (dans le lexique, les mots arabes ne figurent qu'en transcription - par contre la racine est donnée en caractères arabes).

Nous donnons donc ici, dans l'ordre de l'alphabet arabe:

- col. 1 : les lettres arabes
- col. 2 : la transcription de Hans WEHR;
- col. 3 : notre propre transcription;
- col. 4 : les symboles utilisés par EL-ANI (col. 4).

col.1 arabe	col.2 transcription (H. WEHR)	col.3 notre	col.4 EL-ANI
ب	b	b	b
ت	t	t	t
ث	t	t	θ
ج	j	j	j
ح	h	h	h

خ	k	g	g	g
د	d	f	f	f
ذ	ð	q	q	q
ر	r	k	k	k
ز	z	l	l	l
س	s	m	m	m
ش	ʃ	n	n	n
ص	s	h	h	h
ض	ʒ	w	w	w
ط	t	y	y	y
ظ	ʒ			
ع	ʕ			
ف	f			
ق	q			
ك	k			
ل	l			
م	m			
ن	n			
ه	h			
و	w			
ي	y			

ii) voyelles:

Suivant une convention très générale (qui est notamment celle utilisée par H. VEHR), les voyelles longues sont notées par un trait suscrit (ici tréma):

ä, î, ü

Conventions diverses

- \* en gras : formes orthographiques
- \* entre crochets [] : réalisations phonétiques
- \* entre barres obliques // : phonologie
- \* entre barres verticales ||: formes structurelles
- \* les tirets - - séparent des morphèmes
- \* le dièse sépare des mots #
- \* + indique une forme impossible
- \* \* indique une forme tirée du bantou commun de GUTHRIE

Par commodité, les notes sont intégrées au texte - en retrait, en corps différent (corps de notes)

## ETUDE LINGUISTIQUE

Où l'on en vient à ce dont on parle

## ETUDE LINGUISTIQUE

En shingazidja comme dans les autres langues bantu, phonologie et morphologie interfèrent constamment, au point qu'il semble difficile de décrire la distribution des différents phonèmes, leurs conditionnements, etc, en s'en tenant strictement à la phonologie - on parle alors souvent de morpho-phonologie; ainsi que le note ALEXANDRE dans sa présentation générale des langues bantu que nous avons déjà eu l'occasion de mettre à contribution,

*"il est (...) généralement impossible d'analyser le système phonique sans faire entrer en compte la: morphologie (...) plusieurs langues présentent des sous-systèmes phonologiques à conditionnement morphologique (ALEXANDRE op. cit., 363).*

Dans notre optique, cela est d'autant plus nécessaire que l'étude différentielle exige une définition précise des contextes d'apparition des phonèmes, les modifications introduites par l'apport exogène se manifestant notamment par une modification des conditionnements des consonnes.

Il nous est donc apparu préférable de commencer par la présentation morphologique de la langue (niveau monématique), qui sera précédée de l'inventaire des différentes parties du discours (niveau lexicologique); ce n'est qu'après que nous en viendrons à une présentation des phonèmes ou réalisations retenues comme distinctives (cf. p 232 )

## LA CONSTITUTION DES MOTS

Où l'on étudie comment les mots se forment  
dans les deux parties  
du vocabulaire

## PROGRAMME

Par morphologie, nous entendons l'analyse des éléments constitutifs des parties du discours (ou "constituants syntaxiques") et ce qui est nécessaire pour en dresser l'inventaire (ce à quoi nous avons référé comme à la 'constitution des mots'):

- i) l'inventaire des parties du discours correspond plus ou moins au découpage spontané en 'mots', ce qui définit le plan du 'discours';
- ii) l'analyse des parties du discours revient à l'inventaire et à la combinatoire des classes de monèmes qui en sont les éléments constitutifs, ce qui définit le plan de la 'langue'.

C'est à ce niveau que sera introduit l'aspect différentiel.

Ce plan rejoint la dichotomie saussurienne entre parole ('discours') et langue, et aussi, nous semble-t-il, la définition que le Pr. HOUIS donnait du niveau lexicologique:

*"Les rapports entre le lexique comme ensemble des unités significatives qui participent à l'énoncé et la lexicographie comme discipline qui en livre un recensement organisé, exigent que soit posée et résolue une double problématique au sein de laquelle la lexicologie trouve sa justification comme fondement de la systématique des dites unités. En pratique, cela revient à justifier les entrées du dictionnaire pour ce qui concerne la notation, la segmentation, l'identification taxinomique des unités et leur fonctionnement dans l'énoncé. (...)*

*Cette double problématique a pour champ d'élaboration, d'une part la taxinomie des unités significatives aux niveaux où elles opèrent, d'autre part la combinatoire de ces mêmes unités selon les structures requises par telle langue en particulier. Il s'agit en fin de compte de décrire la langue en fonction d'une réflexion sur cette langue, donc de l'identifier en tant qu'elle est une systématique particulière"*

## ORGANISATION DES MOTS

HOUIS poursuit en posant trois niveaux dans "la systématique d'une langue":

- le niveau monématique
- le niveau lexicologique
- le niveau énonciatif

et définit les

"constituants syntaxiques" comme les "unités minimales du niveau lexicologique, soit, les unités telles qu'elles doivent être formalisées pour assumer les fonctions syntaxiques dans les différents types d'énoncés. Il va de soi que certaines langues admettent que le constituant coïncide avec le lexème quant à leurs signifiants. Il n'en reste pas moins que cette coïncidence ne doit pas voiler la nécessité de poser deux niveaux."

Enfin, HOUIS propose le schéma suivant, qui situe la lexicographie comme l'aboutissement d'une problématique:

recensement

lexicologie systématique → lexicographie

organisé

(HOUIS 1983 : 9-11)

Notre démarche s'inscrit dans ce cadre, notre 'plan du discours' correspond au niveau lexicologique du Pr HOUIS, avec pour unités les "constituants syntaxiques", et notre "plan de la langue", au niveau monématique, avec pour unités les monèmes. Nous retiendrons particulièrement la notion de "constituant syntaxique", que nous préférerons à celle de mot, ou de partie du discours. Ce terme a en effet l'avantage de rester neutre par rapport à sa propre complexité monématique, tout en indiquant son niveau de définition.

xplan du discours : constituants syntaxiques

Au plan du discours, les unités (constituants syntaxiques), sont généralement constituées de plus d'un monème; elles proviennent de l'observation immédiate et elles sont envisagées en synchronie, indépendamment (ou presque) de tout

considérations étymologiques ou comparatives: nous en présentons un inventaire général, selon la terminologie bantu traditionnelle; bien que la différenciation étymologique ne soit pas ici pertinente, nous signalons pour chaque catégorie et sous-catégorie la présence ou non d'emprunts, ainsi que leur proportion face au vocabulaire non exogène (bantu au sens large): en effet, si les mots d'origine arabe s'insèrent dans les grandes catégories constituées et instituées par les mots du fonds bantu, et n'ont pas, de ce fait, entraîné la création de nouvelle catégorie syntaxique, il n'en est pas de même au niveau des sous-catégories qui peuvent être spécifiques, ou quasi-spécifiques, d'une provenance.

#### *Plan de la langue : monèmes*

Au plan de la langue, selon la terminologie classique d'A MARTINET, les unités (monèmes) peuvent être lexicales (lexèmes) ou grammaticales (morphèmes).

Nous verrons qu'en shingazidja les processus de formation des mots peuvent différer selon l'origine des termes: en effet, si le modèle bantu tend, par nécessité intrinsèque, à se généraliser, certains emprunts montrent des spécificités irréductibles: à côté d'un état où les emprunts sont "naturalisés" et admettent les processus de formation de mots qui caractérisent le fonds bantu - ce comportement est d'ailleurs un indice essentiel de leur naturalisation - il existe un état où les emprunts, non naturalisés, n'admettent qu'en partie ces processus; ils présentent, sur les plans morphologique et lexicologique, de nombreuses idiosyncrasies qui rendent difficile de dégager des règles...

Nous ferons appel aux langues 'sources' (bantu et arabe) pour rendre compte de ce double fonctionnement - toutefois, nous verrons que les spécificités

morphologiques des emprunts à l'état non naturalisé ne correspondent pas à des traits propres à l'arabe.

Ces deux niveaux sont interdépendants et s'inter-influencent: d'une part, les unités de langue (monèmes) sont perçues à travers l'analyse des constituants syntaxiques (c'est-à-dire des unités de discours), et d'autre part, la catégorisation de ces derniers est, au moins en partie, fonction de leurs caractéristiques morphologiques ...

Par ailleurs, la prise en compte des unités de discours permet d'établir un pont entre l'analyse spontanée en 'mots' et la description 'scientifique' (à tout le moins analytique); si cela peut avoir pour effet de rendre notre travail plus accessible à des non spécialistes, en particulier aux locuteurs non linguistes, nous aurions atteint un objectif, annexe certes, mais pour nous d'une grande importance.

A

## INVENTAIRE DES PARTIES DU DISCOURS

ou

## CONSTITUANTS SYNTAXIQUES

*(niveau lexicologique)*1) Modèle théorique

Comme nous l'avons indiqué en introduction, nous nous situons dans le modèle classique en bantouistique, proposé notamment par P. ALEXANDRE d'après GUTHRIE; le shingazidja s'insère parfaitement dans ce cadre descriptif, preuve s'il en est de son caractère bantu.

Nous reproduisons la catégorisation d'ALEXANDRE *in extenso* (tirée de sa présentation générale des langues bantu, que nous avons déjà mise à contribution dans le chapitre de Méthodologie, pour expliciter l'approche différentielle):

On peut à ce point esquisser une description schématique des parties du discours

ACCORD +		ACCORD -	
Nominaux		Verbaux	Idéophones
• préfixielle courte NI	Série préfixielle longue ND		Interjections
		Verbes	Circonstants ad-verbaux
	Adj/Pron.	Verboïdes (copules)	...
stantifs	Nomino-verbaux		

ALEXANDRE 1981 : 362)

Cette 'esquisse' définit trois oppositions hiérarchisées:

i) variables / invariables :

+ les invariables ne participent pas à l'accord de classe; ils comportent

x des constituants syntaxiques co-extensifs de monèmes (indécomposables)

x des constituants syntaxiques résultant de grammaticalisations de termes variables (verbaux, éventuellement nominaux - à ce titre, ils sont analysables), qui remplissent des fonctions spécifiques (fonctionnels):

ils sont considérés comme invariables en raison de leur caractère figé

+ les termes variables participent à l'accord de classe et se divisent en deux grandes catégories.

ii) nominaux / verbaux :

la différenciation est faite selon des "critères formels et fonctionnels, ce qui amène à admettre une catégorie intermédiaire verbo-nominale";

les nominaux à leur tour se divisent en deux catégories

iii) nominaux indépendants / nominaux dépendants :

l'opposition repose sur la longueur de la série préfixielle, et la rection:

+ les nominaux indépendants (NI) possèdent une "série courte [et] régissent l'accord de classe": cette catégorie correspond à des noms;

+ les nominaux dépendants (ND) sont munis d'une "série [préfixielle] longue, [et sont] régis": cette catégorie définit des adjectifs (dont les numéraux) et les démonstratifs.

Ces dichotomies se retrouvent exactement en shingazidja.

Nous introduisons quelques sub-divisions par rapport à ces schèmes très généraux: parmi les catégories d'invariables pour lesquelles P. ALEXANDRE donne une liste non limitative, nous distinguerons notamment des "conjonctions de subordination et "des conjonctions de coordination", selon les définitions proposées par HAZAEL-MASSIEUX (1983 : 95), les premières "introduisant une relation de dépendance [entre] des unités n'appartenant pas au même paradigme fonctionnel", les autres "réunissant par énumération des unités appartenant au

même paradigme"; dans la catégorie intermédiaire des verbo-nominaux, nous ferons place d'une part à des formes relatives (prévues explicitement par ALEXANDRE) et d'autre part aux infinitifs.

Dans la liste ci-après, nous signalons les catégories qui constituent des inventaires fermés.

Par ailleurs, nous indiquons le cas échéant l'origine étrangère des termes - les exemples étant choisis de manière à faire ressortir l'hétérogénéité morphologique éventuelle de chaque catégorie ou sous-catégorie. Nous signalons aussi les ensembles ne comprenant pas de termes d'origine étrangère.

Les exemples présentés ici sont des 'mots de la langue', donc des formes actualisées: pour les nominaux dépendants et les verbes, ce sont des formes accordées ou conjuguées mais nous indiquons également pour ces derniers la "forme de citation" (forme minimale précédée d'un tiret: voir conventions de notation, pp 75 et suiv.).

## 2) Inventaire général des classes de constituants syntaxiques

### A Termes variables

#### 1 Nominaux:

1-1 nominaux indépendants (NI) (cf. ALEXANDRE "substantifs"):

**ikonbe**, tasse

**ulime**, langue (organe)

**madji**, eau

**kalamu**, crayon ( ar.)

1-2 nominaux dépendants (ND):

+ démonstratifs (inventaire fermé; pas d'emprunt)

-**nu**, ceci, ce...ci: **owandru wanu**, ces gens (-ci)

-**la**, cela, ce...là: **emwana ola**, cet enfant-là

+ adjectifs

-**fupvi**, petit: **ndru mfupvi**, une personne petite

-**wade**, malade: **ndru mwade**, une personne malade

-hodari, courageux (ar.): **ndru rhodari**, une personne courageuse  
 -ili, deux: **wandru waili**, deux personnes

## 2 Nomino-verbaux:

2-1 infinitifs  
**ufanya**, faire  
**hula**, manger  
**uzalisa**, mettre au monde  
**uswali**, prier (ar.)

2-2 formes relatives  
**owadjao**, ceux qui viennent, la relève, les jeunes  
**eyasoma**, quelqu'un qui a étudié  
**owafikirio**, ceux qui pensent (ar.)

## 3 Verbaux:

3-1 copules, verboïdes, qui expriment un présent ou un habituel et sont en inventaire fermé; cette sous-catégorie ne contient pas d'emprunt:

+ copule (qui n'apparaît qu'avec le relatif)  
 -li-, relatif: **emndru ulio hunu**, la personne qui est ici;  
 + verboïdes (présentatif, emphatiseur):  
**ng(a)-**, présentatif (identité, localisation): **nge hunu**, il/elle est ici  
**nd-**, emphatiseur: **(nde) ndru nde yaenda**, c'est la personne qui est venue

3-2 verbes (formes conjuguées)  
 -piha, cuisiner: **hapihi**, il/elle a fait la cuisine  
 -angalia, regarder: **angalia !**, regarde!  
 -fahamu, faire attention, comprendre (ar.): **namunifahamu ndro**, comprenez-moi bien

## B Termes invariables

Les lexèmes-constituants syntaxiques de cette grande catégorie sont en petit nombre, du moins dans la partie native du vocabulaire; on y trouve des fonctionnels de diverses sortes, des mots portant en eux-mêmes leur actualisation (adverbes de temps, termes d'onomastique), les idéophones, par définition hors de tout système, etc

## 4 Conjonctions (inventaire fermé):

4-1 conjonctions de subordination  
**ukaya**, que  
**ili**, pour que (ar.)

4-2 conjonctions de coordination  
**sha**, mais  
**epvonge**, donc, alors  
**lakini**, mais (ar.)

## 5 Circonstants (inventaire limité):

A valeur temporelle.  
**lelo**, aujourd'hui  
**apvaha**, maintenant

halafu, ensuite (ar.)

6 Particules adverbiales, adnominales

ndro, bien

tsena, encore

halisi, vraiment, beaucoup (ar.)

baada, après (ar.)

7 Interjections

ewa, oui

bo (fulani), eh (un tel) !

la, non (ar.)

8 Idéophones

Ils sont rares en tant que monèmes spécifiques, mais des éléments relevant d'autres catégories peuvent être "idéophonisés":

8-1 idéophones (pas d'emprunt):

ruuu, idée d'envol, de départ rapide: pfitt !

evwiiii, idée de vitesse, de furtivité: et hop !

8-2 emplois idéophoniques:

hule, loin (démonstratif en cl. locative 17): huleee, très très loin

hata, jusqu'à (ar.): hataaa, extrêmement, énormément, en excès

Sur l'allongement vocalique expressif, voir voyelles, pp.

Quasi-nominaux

Dans cette catégorie nous rangeons des constituants qui syntaxiquement et sémantiquement s'apparentent aux nominaux, mais n'admettent pas l'accord de classe; on distingue deux sous-catégories, parallèles à celles des nominaux:

9-1 quasi-NI : cela comprend les termes d'onomastique, noms de lieu et de personne, qui sont souvent des syntagmes lexicalisés dans cet emploi mais dont les éléments restent analysables; ces termes se trouvent dans le paradigme des noms (NI) et peuvent commander des accords de classe, bien qu'ils ne prennent pas eux-mêmes de marque de classe correspondant à leur fonction nominale:

**Moroni, Moroni** (capitale de la Grande-Comore)

L'analyse de ce mot comme moro-ni permet deux interprétations basées sur l'homophonie entre mro, feu et mro, rivière, dont *moro* serait une forme ancienne, *-ni* étant un suffixe locatif ; soit 'dans le feu [du volcan]', soit 'dans la rivière'.

**Itsandra, Itsandra**, nom d'une ville de la G-C

Peut s'analyser en li-tsandrai - mais le *thème* ne paraît pas doué d'une signification particulière

**Abudu, Abudu** (nom d'homme) (ar.)

**Arilihila, "Il ne craint rien"** (nom d'homme)

Ce mot s'analyse comme lka-a-r-i#g-hilal, "il ne craint pas la difficulté"

9-2 quasi-ND : cela comprend des termes qui appartiennent au paradigme des adjectifs (nominaux dépendants), mais n'admettent pas l'accord de classe

shenda, neuf (9)

kumi, dix

sita, six (ar.)

makini, calme (ar.)

bile, bleu (fr.)

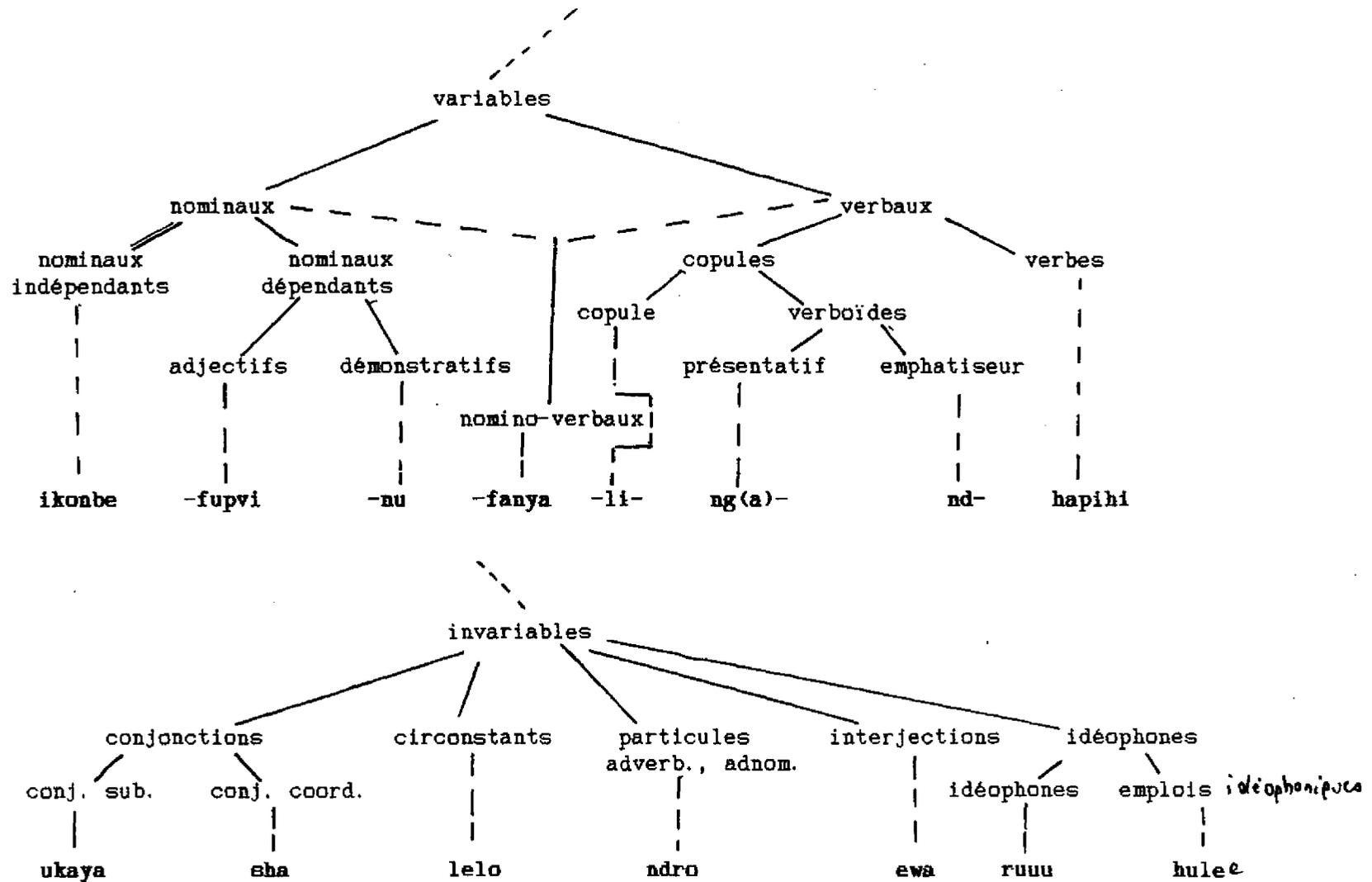
### Remarques

1) Cette typologie des constituants syntaxiques est fondée principalement sur un critère morphologique (variable/invariable); cela scinde des regroupements intuitifs (à partir des gloses françaises en tous cas), tels que adjectifs, numéraux, voire noms: la plupart des mots shingazidja correspondant à des adjectifs ou des noms français appartiennent comme on s'y attend aux catégories variables, mais, comme nous venons de le voir, quelques uns ont un comportement qui les a fait classer comme invariables. Il ne s'agit pas d'une distance due à l'écart entre le français et le comorien (le non-recouvrement strict des découpages sémantiques et syntaxiques des deux langues), mais bien d'une scission à l'intérieur de catégories sémantiques comoriennes, certains termes ayant un comportement particulier - comme on peut le voir dans la liste ci-dessus, il s'agit souvent d'emprunts.

2) Les termes non autonomes, tels que préfixes (de classe), suffixes, etc, n'étant pas considérés comme des constituants syntaxiques, ne figurent pas ici; ils constituent des inventaires fermés que nous verrons au cours de l'analyse morphologique.

**Tableau**

Le tableau suivant reprend notre inventaire, avec un exemple illustrant chaque sous-classe (traduction: se reporter à l'inventaire)



B

## ANALYSE DES PARTIES DU DISCOURS

ou

## CONSTITUANTS SYNTAXIQUES

*(niveau monématique)*

Selon la procédure annoncée, nous traitons ici des unités de langue, c'est-à-dire des monèmes, sous le double aspect de leur inventaire et de leur combinatoire.

Il faut toutefois rester conscient que les monèmes eux-mêmes, et les mécanismes de combinaison, se donnent à travers la comparaison des différents constituants, et non par l'application de catégories et de schémas définis *a priori* - même si nous situons notre travail, à travers la référence constante à la bantuistique, dans un cadre théorique donné: comme nous l'avons souligné dans l'introduction, ce cadre a été lui-même établi à partir de l'observation et de la comparaison de diverses langues bantu.

Nous proposons donc ici un modèle de représentation des processus de formation des mots en shingazidja; sa validité doit être jugée essentiellement à sa capacité (éventuelle) à rendre compte de l'ensemble des-dits constituants, tels qu'ils se rencontrent, ainsi que de la productivité et de la créativité lexicales propres au shingazidja.

Pour les raisons signalées ci-dessus, il est indispensable de prendre en compte dans cette partie le facteur étymologique: toutefois, s'agissant de morphologie,

la pregnance des mécanismes de fonctionnement bantu - qui, de par la définition même de la langue, en sont les mécanismes généraux - est très forte: cela entraîne, comme nous l'avons dit, un état où les emprunts, naturalisés, ne se distinguent morphologiquement plus de mots 'indigènes'; on peut ainsi, sur la base de leur origine et de leur comportement morphologique, opposer termes bantu et emprunts naturalisés d'une part et emprunts maintenant une spécificité d'autre part, ce qui permet de poser un critère ± natif auquel se confrontent les termes non bantu selon leur comportement grammatical.

En outre, les systémativités spécifiques que peuvent montrer les emprunts 'non naturalisés' ne sont pas directement imputables à des faits de la morphologie de leur langue source ...

Nous traiterons d'abord du modèle bantu, les exemples étant pris d'une façon générale parmi les mots non repérés comme emprunts (bantu au sens large). Nous verrons dans une seconde partie, en utilisant pour ce faire les emprunts "non-naturalisés", la manière dont les emprunts s'insèrent, en y introduisant des modifications, dans les catégories syntaxiques instituées par les termes bantu. Enfin, nous étudierons spécifiquement les processus morphologiques de naturalisation des emprunts.

Nous verrons donc successivement :

- le modèle bantu dans son application aux termes de fonds bantu
- les termes d'emprunt non naturalisés (- natif)
- les termes d'emprunt naturalisés (+ natif)

## I LE MODELE BANTU ET SON APPLICATION

## AU FONDS BANTU DE LA LANGUE

Nous présentons le modèle morphologique ou lexicologique bantu tel qu'il apparaît au travers du comparatisme, puis son application au shingazidja dans les mots de fonds bantu; cela nous amènera à préciser l'acception de certains termes de la métalangue ainsi que nos conventions de citation, avant de décrire les modalités d'application du modèle au shingazidja.

## I - 1 Préalables

1) Le modèle bantu : taxonomie et combinatoire

GUTHRIE propose une taxonomie des unités de langue (monèmes) définies à travers leur combinatoire: cette perspective est, comme nous l'avons déjà signalé, l'un des "critères subsidiaires" d'identité des langues bantu - l'on pourrait dire de 'bantuité' - (ALEXANDRE op. cit. 354); nous reprenons *in extenso* la description d'ALEXANDRE, comme nous l'avons fait pour les parties du discours (la citation suivante a déjà été donnée dans l'introduction, à propos de la critériologie bantu):

"B) Critères subsidiaires

(...)

3) [il existe] un ensemble de radicaux (RAD) invariables à partir desquels la plupart des mots se forment par agglutination, ces radicaux présentant les traits suivants:

a - ils ont une structure -CVC-

b - en s'adjoignant un suffixe grammatical (sfv), ils constituent des bases verbales (Vb)

c - en s'adjoignant un suffixe lexical (sfn), ils constituent des thèmes nominaux (TN).

d - un élément d'extension (Ex), de structure -VC- ou -V- peut fréquemment être inséré entre le radical et le suffixe

e - les radicaux n'apparaissent sans préfixe que dans les interjections verbales (impératifs)."

(ALEXANDRE op. cit. 354)

P. ALEXANDRE apporte les compléments suivants:

i) radicaux bivalents / radicaux monovalents:

"les radicaux bivalents (...) produisent à la fois des nominaux et des verbaux, et les radicaux monovalents (...) ne produisent que des nominaux; [pour ceux-ci] il ne semble pas nécessaire d'isoler le radical du thème nominal"

ii) verbal / nominal:

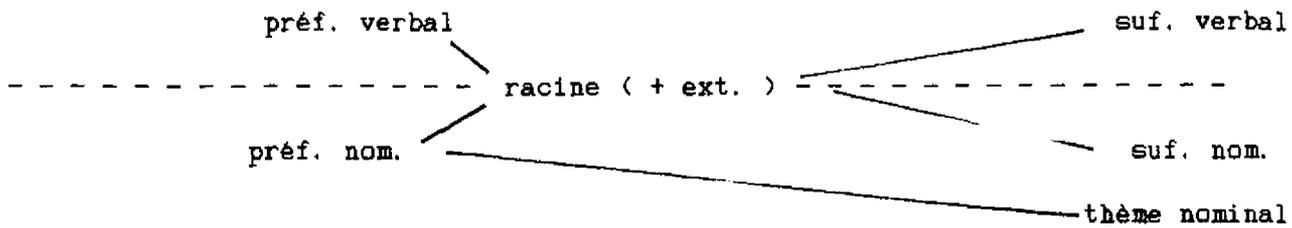
"on peut, cependant, discuter de l'opportunité de retenir le seul critère suffixiel pour opérer la distinction verbal/nominal. Il nous semble plutôt que le nominal doit être formellement appréhendé comme un ensemble (Préf.-RAD-suff. )."

iii) structure des radicaux:

"-CVC- ne représente pas la seule structure possible des radicaux"

Le schéma suivant représente la combinatoire qui découle de ces définitions:

**Verbaux**



**Nominaux**

Ce schéma représente les relations entre catégories de discours et catégories de langue, c'est-à-dire entre constituants syntaxiques et monèmes, ou encore la combinatoire des unités de langue pour aboutir aux unités de discours.

On relève une opposition entre des éléments *obligatoires* (lexèmes, préfixes, suffixes) et des éléments *facultatifs* (extensions). L'organisation de l'étude reprend cette opposition.

En exploitant la distinction opérée par ALEXANDRE entre radicaux bi-valents et radicaux mono-valents, pour lesquels l'isolation du suffixe ne paraît pas nécessaire, nous obtenons les catégories suivantes, qui rendent compte de façon tout à fait satisfaisante du shingazidja:

1/ éléments radicaux ou lexicaux : lexèmes:

1-1: racines (ou radicaux) verbo-nominales, bi-valentes, de structure majoritairement -CVC-, susceptibles d'être augmentées d'une extension (ou dérivatif); ces racines donnent des nominaux ou des verbaux, selon les types d'affixes qui leur sont adjoints;

1-2: thèmes mono-valents (non réductibles à des racines monovalentes complétées d'un suffixe); ils donnent des nominaux uniquement, à partir desquels on distinguera des thèmes proprement nominaux, qui donnent des noms, et des thèmes adjectivaux, qui donnent des adjectifs (nominaux dépendants).

## 2/ éléments grammaticaux non-radicaux : morphèmes:

2-1: affixes actualisateurs (préfixes et suffixes), qui caractérisent le mot comme verbal ou nominal; les préfixes nominaux (il faudrait dire, reprenant HOUIS, "nominants") sont dans un lien plus étroit avec l'élément radical auquel ils sont associés que les préfixes verbaux ("prédicatifs" dans la terminologie de HOUIS).

2-2: extensions ou dérivatifs, de structure -(V)C-, qui peuvent s'adjoindre aux racines verbo-nominales

Il n'est pas tenu compte ici des termes non décomposables, où il y a coïncidence stricte entre lexème et constituant syntaxique, qui constituent des invariables; ces termes ont été signalés dans l'inventaire des parties du discours.

Avant de décrire dans le détail les éléments rentrant dans ce cadre, il nous a paru utile de préciser certains aspects théoriques sur la conception des catégories grammaticales et les modes de citation.

2) Préalables théoriques et méthodologiques

Les catégories de langue se distinguent intrinsèquement des catégories de discours: alors que ces dernières sont (plus ou moins) directement observables, les premières sont constituées d'éléments abstraits, construits, reflétant une hypothèse générale sur le fonctionnement de la langue (empruntée à la bantouistique), et notre compréhension du fonctionnement du shingazidja à un moment donné.

Les notions fondamentales sont à ce niveau (le "niveau monématique" du Pr. HOUIS), comme nous l'avons vu, celles de lexèmes (racines et thèmes), et de morphèmes (affixes et extensions).

Parmi l'ensemble des morphèmes opérant en shingazidja, nous privilégions ceux qui interviennent dans la formation des mots 'lexicologiques', stables, qui sont les mots retenus dans le lexique: nous excluons ainsi en particulier les

morphèmes spécifiques des formes verbales conjuguées qui constituent des préfixes et des suffixes particuliers, dits "verbaux".

L'aspect construit des catégories de langue transparait dans leur mode de citation qui est distinct du mode de citation des constituants syntaxiques (cf. conventions de notation, pp 75).

#### a) éléments lexicaux

##### α) racine vs thème

Sur le plan lexicologique, la différence réside dans le caractère bi-valent de la racine qui s'oppose au caractère monovalent du thème: une racine donne aussi bien des verbaux que des nominaux, alors qu'un thème, par définition, ne peut donner que des nominaux.

Il suffit donc, en principe, que, sur une base lexicale donnée, soit attesté un verbal pour que l'on en infère l'existence d'une racine bi-valente.

• racine : antériorité nominale ou verbale:

Nous ne nous placeons pas dans cette polémique, qui évoque celle de la poule et de l'œuf: l'observation des constituants syntaxiques de la langue ne permet pas de poser que l'un soit antérieur à l'autre; ainsi que l'écrit HAGEGE, "*on s'enferme dans une inextricable circularité dès lors que l'on veut à toute force que le nom ait pré-existé au verbe ou inversement le verbe au nom, puisque ce n'est que l'un par rapport à l'autre qu'ils existent et se définissent.*" (HAGEGE 1984 : 17).

Par le terme même de racine verbo-nominale, - nous aurions pu indifféremment utiliser celui de racine nomino-verbale - nous postulons que la racine est neutre, indifférenciée, sans polarité préférentielle, et que de cette racine sont issus tant des verbaux que des nominaux (ou des nominaux et des verbaux).

• thème :

Il s'agit d'un élément lexical avec polarité nominale, telle que ce terme recouvre autant les noms proprement dits (nominiaux indépendants) que les adjectifs (nominiaux dépendants)

β) réserves

Deux réserves sont à prendre en considération quant à l'établissement des lexèmes:

• le corpus:

Il est par définition lacunaire; en particulier, il est toujours possible que nous n'ayons pas relevé de forme verbale pour un lexème donné alors même qu'elle existe; dans ce cas, nous présentons comme thème ce qui en réalité est racine ...

• les rapprochements:

Peut-être plus importante encore que le corpus est la question des rapprochements: face à des constituants divers présentant une convergence formelle et sémantique, la décision à prendre concernant leur rattachement éventuel à un même lexème - qui conditionne la construction d'une racine ou d'un thème - n'est pas toujours d'évidence: à côté de cas assurés, nous nous trouvons souvent confronté à des situations où le rapprochement lexicologique entre des formes nominale et verbale d'une part, entre différents verbaux ou nominiaux d'autre part, implique des choix où entre une part non réductible d'arbitraire. Cela paraît inévitable.

Nous reconnaissons et admettons cette réserve sur la signification et la valeur des rapprochements: elle n'infirmes pas les analyses et les résultats proposés mais rappelle qu'ils ne sont valables qu'autant que le sont les hypothèses sur lesquelles ils sont fondés.

Nous en donnons un exemple.

Si l'on peut affirmer sans risque que **paha**, chat (pl. **mapvaha**), relève d'un thème nominal |-pvaha| (notion de félin) et n'a rien à faire avec la racine |-pvah-| (notion de peindre) qui donne -pvaha, peindre, **mpvaho**, manière de peindre, etc, qu'en est-il de **nyumba**, maison, en face d'une racine |-unb-| (notion de créer), qui donne -unba, créer, **munbo**,



création, etc ? ou de *mlima*, colline, en face de *l-lim-l* (notion de cultiver), qui donne *-lima*, cultiver, *ndima*, agriculture, etc ?

Bien entendu, la plausibilité morphologique (à partir de l'étude des suffixes nominaux affectant les racines) doit être invoquée; en l'espèce, elle ne paraît pas déterminante: il sera vu que *-a* est bien un suffixe nominal, rare mais non exclu même dans des cas où le terme ne désigne pas un agent.

Notre attitude fait la part de notre propre subjectivité et nous ne regroupons pas *mlima* et *-lima*, non plus que *nyunba* et *-unba*.

Nous définissons donc des thèmes autonomes, *l-lima* (notion de colline) & *l-unba* (notion de maison), comme *l-pvaha* (notion de félin).

Par ailleurs, les rapprochements reposent aussi sur l'analyse morpho-phonologique, en particulier sur l'alternance consonantique, et, d'une façon plus générale, sur la distribution des consonnes en shingazidja et la répartition des reflets des symboles BC qui sera traitée ultérieurement.

Ainsi, en tout état de cause, les regroupements (ou les non-regroupements) proposés des termes 'indigènes', qui constituent notre matériau pour cette partie et qui sont donnés dans le lexique, ne le sont-ils qu'à titre provisoire.

Quelques lexèmes contredisent ces principes définitoires: ainsi on trouve certains verbes, tels qu'aucune formation nominale apparentée ne soit possible, et quelques nominaux sur lesquels peuvent se former des verbaux au moyen d'une dérivation particulière ...

Il s'agit, pour les verbes, essentiellement de verbes 'météorologiques' (*-tswa*, se coucher (pour le soleil), *-nya*, pleuvoir, etc), et, pour les nominaux, de quelques nominaux dépendants (adjectifs) exprimant la qualité (*-le*, grand, *-nene*, gros, *-radji*, large) à partir desquels, au moyen de dérivés (extensions) spéciaux, peuvent se former des verbes d'état "dé-adjectivaux" ('grandir' sur grand, 'grossir' sur gros, 'élargir' sur large).

Il s'agit en tout état de cause de processus non productifs.

Ces quelques exceptions, sémantiquement explicables, ne mettent ainsi pas en cause la généralité de l'opposition racine - thème: malgré cette particularité, nous considérons ces verbes météorologiques comme issus

de racines, et ces nominaux dépendants, comme issus de thèmes (adjectivaux).

*γ) constructions synchroniques vs reconstructions comparatives & diachroniques*

Il importe également ici de préciser un aspect de la relation que nous faisons au bantu commun dans le domaine morphologique: cela nous amène à poser la distinction de deux niveaux dans l'établissement des lexèmes, un niveau synchronique interne et un niveau comparatif à visée diachronique.

• niveau synchronique interne

Les lexèmes (racines & thèmes) sont postulés au travers de faits linguistiques shingazidja uniquement et les rapprochements lexicologiques effectués ne font pas intervenir d'autres considérations. Il nous suffit pour instituer le lexème qu'il découle de la combinatoire normale des monèmes shingazidja (voir ci-dessus la problématique des rapprochements).

Dans certains cas, la parenté ainsi déduite paraît indubitable:

Ainsi, nous rapprochons les constituants

-loa, pêcher et  
iloo, hameçon

sur la base de leur proximité formelle et sémantique, ainsi que de la régularité de leur formation (les éléments en seront vus plus loin), qui permet de postuler une racine |-lo-|, [notion de pêcher| dont ils sont tous deux issus;

de même,

-pvenya, cligner des yeux et  
npbenye, éclair

pour les mêmes raisons, qui permettent de postuler une racine |-pveny-|, [notion de briller|, dont ils sont tous deux issus;

de même,

ndru, être humain et  
hindru, chose

qui permettent de postuler le thème |-ndru|, [notion d'existant|, que viennent préciser les préfixes de classe (voir plus loin).

De même, le nominal isolé *mwendje*, lumière, que l'on analyse comme issu d'un thème *[-endje]*, sera-t-il rapporté à la racine BC *\*-bèng-*, étinceler:

cela suppose, au moins implicitement, une racine *\*-endj-*, à partir de laquelle le nominal *shingazidja* a été obtenu normalement par l'adjonction des affixes obligatoires (*\*[mu-endj-e]*; on verra que ces éléments, préfixe *[mu-]* et suffixe *[-e]*, font partie des affixes identifiés).

Mais il est clair que *mwendje*, en l'état actuel de notre documentation, reste considéré synchroniquement comme issu d'un thème nominal: nous ne tirons pas argument des considérations comparatives pour le décomposer en *shingazidja*.

Ainsi posons-nous l'autonomie sur le plan lexicologique du *shingazidja* vis-à-vis de l'approche comparative.

#### b) éléments grammaticaux

##### *affixes vs extensions*

Quoique appartenant tous deux à la même catégorie des éléments grammaticaux, extensions et affixes ne sont pas de même nature: les unes (extensions) s'adjoignent facultativement aux racines sans en modifier le statut - une racine "étendue" reste une racine - alors que les autres (affixes) sont des actualisateurs nécessaires pour permettre le passage des racines et thèmes au statut de constituants syntaxiques.

Cela amène à distinguer parmi les morphèmes, des éléments obligatoires, que sont les actualisateurs, et des éléments facultatifs, que sont les extensions.

Nous examinons maintenant les processus de formation des mots en *shingazidja* à travers les monèmes et leurs combinatoires: nous dressons les schémas généraux caractéristiques, d'une part, des termes issus de racines verbo-

nominales et d'autre part des nominaux issus de thèmes, puis nous présentons les éléments impliqués et leurs combinatoires.



- racine étendue : ext. sémantique + ext. syntaxique : -u- + -w- :  
 → verbe: -fungulwa, être détaché : |-fung-u-lw-al, rac. - ext. sém. - ext. synt. - suf. v.

**RVH** : -pih- (notion de cuisiner)

- racine primaire:  
 → verbe: -piha, faire la cuisine : |-pih-al, rac. - suf.v  
 → nom: **mpiho**, préparation culinaire, plat : |m-pih-ol, préf.n. - rac. - suf.n

**mpihadji**, cuisinier : |m-pih-adjil, préf.n. - rac. - suf.n

→ nom: **mpishi**, cuisinier : |m-pih-il, préf.n. rac. - suf.n.

- racine étendue (ext. sémantico-syntaxique) : -i-, applicatif:  
 → verbe: -pishia, cuisiner pour qqun : |-pih-i-al, rac. - ext. synt. - suf.v  
 On remarque que -i- du suffixe nominal et de l'extension entraîne une modification de la finale radicale

**Rac** : -l- (notion de manger)

- racine primaire:  
 → verbe: -la, manger : |-l-al, rac. - suf.v  
 → nom: **llo**, nourriture : |m-l-ol, préf.n. - rac. - suf.n
- racine étendue (ext. synt.) : -is-, causatif:  
 → verbe: -lisa, nourrir : |-l-is-al, rac. - ext. synt. - suf.v

**b) constituants syntaxiques issus de thèmes**

A partir des thèmes, on obtient seulement des nominaux:

**nominaux** : ————— préf.n. + thème

Exemples:

**Th** : -ndru (notion d'existant)

- **wandru**, gens, personnes : |wa-ndrul, préf.n - thème
- **hindru**, chose : |hi-ndrul, préf.n - thème

**Th** : -bibo (notion de cajou)

- **bibo**, fruit du cajoutier : |Ø'-bibol, préf.n - thème
- **mbibo**, cajoutier : |m-bibol, préf.n - thème

**Th** : -wade (notion de maladie)

- **wade**, malade (cl.1) : |mu-wade;, préf.n - thème
- **uwade**, maladie : |u-wadel, préf.n - thème

Nous allons donc voir en détails les différentes composantes des constituants syntaxiques:

1) éléments obligatoires:

- lexicaux : racines et thèmes
- grammaticaux : préfixes et suffixes;

ii) éléments facultatifs : affixes, correspondants à ce que nous avons appelé 'extensions'

## I - 3 Analyse des constituants syntaxiques

1) Eléments obligatoires1-1 Eléments lexicauxa) *Racines verbo-nominales (primaires)*

Nous nous limitons ici aux racines primaires, c'est-à-dire sans extension: nous avons vu en effet dans le schéma de formation que les racines peuvent être primaires ou étendues (munies d'une ou plusieurs extensions).

Les nombreux cas d'analyse difficile ou problématique (où la racine n'est attestée que sous une forme apparemment étendue), pour lesquels GUTHRIE a introduit la notion de racine "complexe", seront traités dans le chapitre consacré aux extensions.

Rappelons que la décomposition des racines, et partant, la reconnaissance de leur qualité (primaire, étendue ou complexe) est effectuée à l'intérieur du shingazidja *pro se*, et que le corpus de référence est constitué par le lexique à travers son organisation lexicologique.

GUTHRIE donne pour les radicaux une structure -CVC- (la structure réfère au nombre de phonèmes constitutifs de la racine); de fait, la majorité des racines verbo-nominales primaires (sans extension) de fonds bantou en shingazidja réalisent cette structure ou une structure que l'on peut considérer comme en découlant.

Toutefois cette structure n'est pas unique.

À cette structure, il est intéressant d'adjoindre le type, qui, lui, renvoie à la quantité syllabique, car tous deux conditionnent, pour les verbes qui en sont issus, la forme de certains temps.

Cette double approche permet de regrouper des structures différentes, dont la différence est souvent due à la chute d'un élément consonantique du BC. Ainsi, -CVC- représente une structure tri-phonématique et appartient au type mono-syllabique, de même que -V-, -VC- ou -CV-.

Le tableau suivant combine ces deux critères.

+ Pour des raisons de place, nous ne donnons pas les étymons éventuels; ils figurent dans le lexique

rac. primaires

struct.

struct.				infra-syllab.
mono- phon.	-C- :	-l-   notion de manger		
		-f-   notion de mourir		
		-dj-   notion de venir		
	...			
	-V- :	-u-   notion de frapper		s
				y
bi- phon.	-CV- :	-fu-   notion de laver	m	l
		-ho-   notion de tousser	o	l
		-me-   notion de pousser	n	a
	-VC- :	-end-   notion d'aller	o	b
		-on-   notion de voir		i
	-ol-   notion de pourrir		q	
	...		u	
			e	
tri- phon.	-CVC- :	-lim-   notion de cultiver		
		-pih-   notion de cuisiner		
		-kop-   notion d'emprunter		
		-vur-   notion de tirer		
	...			
				s
				y
quadri- phon.	-CVVC- :	-siuh-   notion de dormir	d	l
	-VCVC- :	-azim-   notion de prêter	i	l
	...			a
				b
				i
penta- phon.	-CVCVC- :	-tafuny-   notion de mordre		q
		-diwaz-   notion d'oublier		u
	...			e

Les racines non exogènes de notre corpus se répartissent comme suit, selon les types et les structures (nous signalons le nombre de racines munies d'un étymon):

- par structure:

- mono-phonématique (-C-, -V-): 16, dont BC 16 (voir nota ci-dessous)
- bi-phonématique (-CV-, -VC-): 38, dont BC 22
- tri-phonématique (-CVC-): 222, dont BC 109
- quadri- et penta-phonématique (-CVVC-, -VCVC-, -CVCVC-): 202 dont BC 57

- par type:

- infra-syllabique (-C-): 15 dont BC 15
- mono-syllabique (-V-, -CV-, -VC-, -CVC-): 261 dont BC 132
- bi-syllabique (-CVVC-, -CVCV-, -CVCVC-): 202, dont BC 57

soit au total 476 racines bantu dont 201 ont des étymons chez GUTHRIE.

Nota

1) D'une façon générale, les lexèmes constituent des inventaires ouverts, et le corpus ne saurait être exhaustif; l'information chiffrée donnée ici porte donc seulement sur l'importance relative des différentes structures et types de racines (si l'on admet une représentativité à notre corpus).

Toutefois, en ce qui concerne spécifiquement les racines mono-phonématiques -C- & -V-, nous pensons en avoir relevé l'ensemble des attestations: il est en effet facile d'enquêter directement sur l'existence ou non de racine de cette structure.

2) Une procédure de recherche systématique semblable est envisageable pour les autres structures, mais les possibilités croissent exponentiellement et nous ne l'avons pas fait - en outre, cela pose des problèmes méthodologiques particuliers (cette procédure a été suivie par C. SEYDOUX dans l'établissement du dictionnaire peul: communication sur le dictionnaire peul, réunion de décembre 1986 dans le cadre du Centre d'Etudes Africaines de l'INALCO).

3) Les verbes issus de racines monosyllabiques -CVC-, -CV-, -VC- ont, à certains temps (accompli affirmatif non relatif et négatif absolu) une voyelle qui est soit la reproduction de leur voyelle radicale (harmonie vocalique) soit -a (qui s'applique aux verbes d'autres structures, indépendamment de leurs voyelles internes):

- enda, aller : haende, il/elle est allé/e
- ona, voir : (wo) kwaono, ils/elles ne voient jamais
- lima, cultiver : walimi, ils/elles ont cultivé

mais :

- nika, donner : hanika, il/elle a donné
- toa, sortir (tr.) : (ye) katoa, il/elle ne sort pas (qqch)

comme les verbes 'longs' :

- eshelea, se souvenir : waeshelea, ils/elles se se sont souvenus
- maliza, terminer (tr.) : wamaliza, ils/elles sont terminés (qqch)

Les verbes issus de racines monophonémiques -C- et -V- ont à ces mêmes temps une voyelle spécifique (qui peut être -a) :

- la, manger : wali, ils/elles ont mangé
- dja, venir : (mi) tsidja, je ne viens pas

Le comportement de ces verbes peut être rapportés au bantou commun; toutefois, comme nous ne traitons pas ici de suffixes verbaux, nous ne l'analyserons pas - et nous contenterons d'indiquer dans le lexique le timbre de la voyelle suffixielle :

- enda (e), -ona (o), -nika (a), etc

On constate que la structure -CVC- est la plus fréquente; cette prédominance s'accroît si l'on envisage le type mono-syllabique -VC-, -CV- et -CVC- qui forme une unité, -VC- & -CV- étant souvent des avatars de -CVC- reflétant la structure bantou canonique \*-CVC-.

En tout état de cause, la nature même des éléments lexicaux, qui sont inanalysables, insécables, explique la prédominance des structures brèves.

#### b) Thèmes

Dans sa critériologie, GUTHRIE ne pose pas de structure type pour les thèmes, qu'il n'isole pas des racines. Nous reprenons la distinction racines - thèmes, fondamentale en shingazidja, d'ALEXANDRE qui pose la non-séparabilité d'un suffixe dans le cas des nominaux.

On trouve en shingazidja différentes structures, toujours ouvertes (à finale vocalique), ce qui est lié de la structure syllabique ouverte du shingazidja.

La structure des thèmes la plus fréquente est -CVCV, qui rappelle la forme canonique des racines (-CVC-) augmentée d'un suffixe, mais d'autres structures se rencontrent:

- CV : l-tsal, lnotion de boutl → ntsa, bout  
l-shel, lnotion de femellel → (wandru) washe, (des) femmes ('des personnes femelles')  
l-nul lémonstratif de proximitél → wandru wanu, ces personnes
- CVCV : l-garel lnotion de farine l → ugare, farine  
l-buzil lnotion de caprin l → nbuzi, chèvre  
l-radjil, lnotion de large l → (ari) radji, (une planche d'arbre) large
- VCV : l-anal, lnotion d'enfant l → mwana, enfant  
l-itsil lnotion de crudité (pour aliments) l → (muhogo) mwitsi, (du manioc) crû
- VV : l-eol lnotion de respect l → sheo, respect  
l-ill, lnotion de mauvais l → (mwaha) nbi, (une mauvaise) année
- CVV : l-shial lnotion de queue l → mshia, queue

La définition amène aussi à distinguer, selon la longueur de la série préfixielle des nominaux qui en sont issus, des thèmes donnant des nominaux indépendants (NI, noms) et des thèmes donnant des nominaux dépendants (ND, adjectifs) (thèmes adjectivaux):

\* nominaux indépendants : combinatoire préfixielle limitée

- l-shial → mshia, queue ; mishia, queues:  
ce thème apparaît dans deux formations nominales, appartenant à deux classes différentes en opposition de nombre
- l-garel → ugare, farine et nyigare, farines :  
ce thème apparaît dans deux formations nominales appartenant à deux classes différentes (en opposition de nombre)
- l-buzil → nbuzi, chèvre ; nbuzi, chèvres  
buzi, bouc , mabuzi, boucs  
ubuzi, prostitution  
ce thème apparaît dans cinq formations nominales appartenant à cinq classes différentes (dont quatre en opposition de nombre deux à deux - on remarque deux classes aux préfixes nominaux semblables (nbuzi, chèvre/s): ce sont les accords régis qui diffèrent).

nominaux dépendants : combinatoire préfixielle libre

- l-shel → mshel, washe, mishe, tshe, ntshe,, etc : .... femelle
- l-nul → wanu, unu, inu, linu,, etc : ce ... ci

l-radji → mradji, waradji, miradji, tradji,, etc : ... large  
 ces thèmes apparaissent sans restriction dans autant de formations  
 nominales que de classes différentes

Cette opposition interne est importante sur le plan sémantique: le sens des NI provient en partie d'une interaction entre préfixe et thème (voir motivation des classes, *infra*) alors qu'il n'en est pas de même pour les ND, dont le sens est indépendant du préfixe - puisqu'ils ne connaissent en principe aucune restriction dans la série des préfixes qu'ils admettent.

Nous avons déjà signalé que certains thèmes adjectivaux admettent la formation de verbaux, au moyen d'une extension spéciale.

Les constituants syntaxiques (verbes et noms) issus des racines verbo-nominales appartiennent à tous les champs sémantiques, et ne privilégient aucun rapport particulier de l'énonciateur à son propos; les racines primaires peuvent être transitives ou intransitives, admettre des sujets personnels ou être limitées à des constructions impersonnelles, etc: les constituants qui en sont issus reflètent cette ouverture.

Les nominaux issus de thèmes ont par contre certains domaines sémantiques privilégiés, qui se trouvent pour la plupart correspondre à la partie motivée des classes (voir *infra* pp. 128). Il se dégage de ce fait une très relative répartition sémantique des référents nominaux entre les deux catégories morphologiques: ainsi, des notions telles que celles référant aux créations 'naturelles' (être humain, animaux, plantes, etc) sont-elles généralement rendues par des termes issus de thèmes et non de racines.

Toutefois, cette très relative spécialisation sémantique des thèmes nominaux n'est en aucune manière impérative: il se trouve des nominaux issus de thèmes qui désignent des réalités pour lesquelles un terme provenant d'une racine

paraîtrait concevable, de même que des nominaux bâtis sur une racine peuvent désigner des réalités écologiques:

- à côté du thème {-tsozi| (notion de larmes| qui correspond au nominal tsozi, larmes, existe la racine {-lil-| (notion de pleurer| à l'origine du verbe -lila, pleurer: un mot tel que \*malilo dans le sens de 'larmes' serait possible mais n'est pas attesté;

- si le terme nbingu, nuages, ciel, comme la plupart des autres réalités écologiques ne relève clairement que d'un thème (-bingu, (notion de nuage|), les termes synonymes upepo et nbepvo, vent, sont à rapprocher du verbe -pepea, flotter (au vent), par l'intermédiaire de la racine -ppepv-; l'on a vu, d'autre part, que nous rapprochions mare, salive, de la racine -r- (p 102), etc

De la partie irréductible de cette répartition, qui se constate dans d'autres langues où opère la dichotomie racine verbo-nominale / thème nominal, on peut peut-être tirer l'hypothèse que le système de pensée reflété par ces langues oppose des notions qui peuvent être considérées comme liées à un processus (pour les nominaux issus de racines) et d'autres, qui paraissent ne relever que de ce qu'on pourrait appeler 'essences' ou 'substances' (pour les nominaux issus de thèmes).

### 1-2) Eléments grammaticaux obligatoires

La critériologie de GUTHRIE fait état d'éléments grammaticaux, suffixes (suffixes verbaux et nominaux), et préfixes, qui s'adjoignent aux radicaux pour former les mots de la langue.

Les schémas constitutifs des parties du discours en shingazidja établis d'après l'application de ce cadre montrent, à la périphérie des lexèmes (racines et thèmes), des éléments grammaticaux affixés (préfixes et suffixes); ces éléments actualisent les lexèmes, c'est-à-dire, modifient leur statut pour permettre leur intégration dans le discours.

Ces éléments ont les caractéristiques suivantes, qui les opposent aux lexèmes (et aux extensions) et justifient leur interprétation comme morphèmes, dans la terminologie d' A. MARTINET:

- x rôle morphologique : accompagnateur et actualisateur obligés des lexèmes;
- x inventaire fermé;
- x sémantisme particulier qui se confond souvent avec leur fonction grammaticale.

Par définition, les racines bi-valentes verbo-nominales nécessitent des préfixes et des suffixes, alors que les thèmes nominaux n'admettent que des préfixes.

Les préfixes sont donc communs aux racines et aux thèmes, alors que les suffixes sont réservés aux seules racines et, dans ce cas, il existe des relations de compatibilité entre préfixe et suffixe.

Pour cette raison, nous traitons d'abord des suffixes puis des préfixes.

#### *a) Suffixes nominaux*

Ils ne possèdent que rarement des étymons chez GUTHRIE (deux formes sur six); toutefois, ces morphèmes sont prévus dans la combinatoire bantou; sans doute

les éléments dont GUTHRIE disposait ne lui ont-ils pas permis d'autres reconstructions.

Les suffixes apparaissent porteurs d'une information sémantique et grammaticale; compte-tenu de la structure syllabique ouverte du shingazidja (ce sont les éléments terminaux des constituants syntaxiques issus de racine), les suffixes sont toujours à finale vocalique - en fait ils ne sont constitués que d'une seule voyelle, celle-ci pouvant entraîner des modifications morpho-phonologiques de la finale de la racine.

Tous dressons la liste des suffixes classés selon leur forme, en indiquant les relations avec les préfixes (en regroupant éventuellement ceux en opposition de nombre) et, quand on le peut, la valeur de l'ensemble.

α) -a :

conditionnement :

suffixe très fréquent; s'emploie avec les toutes les racines, nues ou étendues (sauf passif), et se trouve combiné avec de nombreux préfixes de classe.

• avec préf. cl. 3 / 4: m- /mi- : formation très productive; nom d'action

- mfanyo**, (une) action |m-fany-ol < -fany- |notion d'être fait|
- mfanyiho**, acte (en tant qu'il a été fait) |m-fany-ih-ol < -fany- |notion d'être fait|
- munbo**, (la) création (de l'homme) |m-unb-ol < |unb-| |notion de créer| < \*-bumb-
- mfano**, exemple, cas |m-fan-ol < -fan- |notion de ressembler| < \*-pú'an-
- mo**, nourriture |m-l-ol < -l- |notion de manger| < \*-dí-
- mfungo**, noeud |m-fung-ol < -fung- |notion d'attacher| < \*-tú'ng-
- mfunguo**, clef |m-fung-u-ol < -fung- |notion d'attacher|
- mkopesheo**, prêt |m-kop-sh-e-ol < -kop- |notion d'emprunter| < \*-kóp-
- mwendo**, allure |mu-end-ol < -end- |notion d'aller| < \*-yènd-
- mwendeleo**, progrès |mu-end-ele-ol < -end- |notion d'aller|
- mfwagio**, balai |m-fwagi-ol < -fwagi- |notion de balayer| < \*-piágid-

• avec préf. cl. 5 / 6 : Ø'-, dji- / ma- : formation productive; nom d'action

- bunbo**, (la) création, (l') origine |Ø'-wunb-ol < |wunb-| |notion de créer|
- djapizo**, malédiction |dji-apiz-ol < -apiz- |notion de maudire| < \*-dáp-
- djinbo**, chanson |dji-inb-ol < -inb- |notion de chanter| < \*-yimb-
- bazo**, bouchon |Ø'-bay-z-ol < -bay- |notion de fermer|

**malipvo** (cl. 6), salaire **lma-lipv-ol** < -lipv- Inotion de payer! < \*-dip-

- avec préf. cl. 7 / 8: i-, shi- / zi- : formation productive; instrument

**iloo**, hameçon **li-lo-ol** < -lo- Inotion de pêcher!

**ikao**, groupe **li-kay-ol** < -kay- Inotion d'exister! < \*-yi'kad-

**shelo**, mesure **lshi-hel-ol** < -hel- Inotion de mesurer!

- avec préf. cl. 9 / 10: N- / N- : formation peu productive

**nkeleo**, marche, escalier **lN-he-le-ol** < -he- Inotion de monter!

**ntondoo**, morceaux **lN-tondo-ol** < -tondo- Inotion de couper en morceaux!

- avec préf. cl. 11: u- : formation peu productive et souvent équivalente à celle en cl. 3

**ufunguo**, clef (= **mfunguo**) **lu-fung-u-ol** < -fungu-

**wendo**, allure (= **mwendo**) **lu-end-ol** < -end-

Ce suffixe est très fréquent; il est presque toujours attesté avec les préfixes de cl. 3 & 4, le nominal étant alors le nom d'action du verbe.

Nous avons signalé que -o était incompatible avec l'extension passive -w- (qui entraîne souvent la vélarisation de la finale radicale); cela se conçoit sémantiquement, mais son absence dans ce cas peut aussi être le résultat d'une neutralisation due à l'analgane Cw + o + Co.

B) -a : < \*a

conditionnement :

toutes les racines; ce suffixe est le plus souvent combiné avec la classe 15 pour former des infinitifs, et avec les classes 1/2 pour désigner des agents

- avec préf. cl. 15 **lhu-l** : infinitif ; formation attestée avec toutes les racines:

**urenda**, faire **lhu-rend-al** < l-rend-l Inotion de faire! < \*-tënd-

**hula**, manger **lhu-l-al** < l-l-l Inotion de manger!

**usoma**, étudier **lhu-som-al** < l-som-l, Inotion d'étudier!

- avec préf. cl. 1 / 2 : m- / wa- : agent :

x après racine primaire : suivi d'un complément:

**mtsunga mbe**, éleveur (de boeuf) **lm-tsung-a#N-mbe** < l-tsung-l Inotion d'élever (des animaux)

**mfanya hazi**, travailleur ('faiseur de travail') **lm-fany-a#0-hazi** < l-fany-l Inotion de faire!

**mwandza sheo**, personne qui veut être respectée ('personne qui aime le respect') **lmu-andz-a#sh-eol** < l-andz-l Inotion de vouloir!

Les racines l-fany-l et l-andz-l sont particulièrement propices à cette construction, et peuvent être considérées comme des formateurs de mots composés; **mfanya-hisabu**, comptable ('faiseur de compte'); **mwandza-ndzadze**, adolescent attardé ('personne qui veut sa mère'), etc

x après passif -w- :

**rumwa**, esclave |m-rum-w-a| < |rum-| |notion de servir| < \*-túm-

• autres classes : sens divers

x après racine primaire :

**nkoa**, cl. 9 toux |N-ho-a| < |ho-|, |notion de tousser| < \*-kóód-

**nwana**, cl. 9 combat |N-wan-a| < |wan-| |notion de lutter|

**ndola**, mariage |N-lol-a| < |lol-| |notion de marier|

**avuna**, cl. 6 récolte |ma-vun-a| < |vun-| |notion de récolter| < \*-bú'n-

**rema**, (cl.3/4) champ cultivé ('défriché, coupé') |m-rem-a| < |rem-| |notion de frapper|

x après passif :

**bunwa**, cl. 5 constitution |Ø'-wunb-w-a| < |wunb-| |notion de créer|

**itrendwa**, cl. 7 action |i-Ø'-rend-w-a| < |rend-| |notion de faire|

est un suffixe sans nul doute d'origine verbale: il est employé dans des formes clairement nomino-verbales: les infinitifs (cl. 15) et dans les noms d'agent en cl. 1/2, avec une complémentation de type verbal; son étymon \*a figure parmi les marques de temps ("tense signs").

#### Rappel

Nous avons vu dans les préalables quelques exemples de nominaux à finale [a], identique au suffixe, mais que nous ne décomposons pas du fait que le rapport sémantique avec la racine ne paraissait pas suffisamment solide (alima, colline ? < -lim- |notion de cultiver|, nyunba, maison ? < -wunb-, |notion de créer|

#### Y) -adji :

conditionnement :

après racine primaire, combiné avec les préfixes des classes 1/2 et 11

• avec préf. cl.1 / 2: m- / wa- : valeur d'agent

**mfumadji**, chasseur |m-fum-adjil| < -fum- |notion de chasser|

**msonyadji**, couturier |m-sony-adjil| < -sony- |notion de coudre|

**mpihadji**, cuisinier |m-pih-adjil| < -pih- |notion de cuisiner|

**alimadji**, paysan |m-lim-adjil| < -lim- |notion de cultiver|

Cette forme peut être en variante avec celle formée à l'aide du suffixe -i (ci-dessous); alimizi = alimadji; msonyi = msonyadji; mpishi = mpihadji, etc

• avec préf. cl. 11 u- :

**unonyadji**, oppression |u-nyony-a-djil| < -nyony- |notion d'opprimer|

Comme le suffixe -a, toutes les racines primaires admettent ce suffixe; par rapport à -a, -adji se caractérise par le blocage de toute complémentation de type verbal.

δ) -i: < BC \*i' ;

conditionnement :

après racine primaire ou étendue; par ailleurs, ce suffixe entraîne une modification de la finale radicale (en synchronie ou en reconstruction), qui aboutit à une fricative pré-palatale:

h + i → shi (< \*k + \*i')

l + i → zi (< \*d + \*i')

w + i → zi (< \*b + \*i')

Dans quelques cas, h + i → si ; w + i → vi, b + i → dzi

Ce conditionnement est semblable à celui d'une extension notée -i-, à valeur causative, venant d'un étymon identique (voir plus loin, pp

Comme le précédent, ce suffixe se combine essentiellement avec les préfixes de cl. 1/2 et 11; il n'apparaît pas productif actuellement.

• avec préf. cl.1 / 2 : m- /wa- : agent

**mlozi**, pêcheur |m-lo(w)-i| < |lo(w)-| |notion de pêcher| < \*-dób-  
**mlezi**, éducateur |m-lel-i| < |lel-| |notion d'enseigner| < \*-déd-  
**mpishi**, cuisinier |m-pih-i| < |pih-| |notion de cuisiner| < \*-yi'pik-  
**mwanzishi**, écrivain |mu-andzih-i-| < |andzih-| |notion d'écrire| < \*-yándik-  
**mrumishi**, serviteur |m-rum-ih-i-| < |rum-| |notion d'utiliser| < \*-túm-  
**mtsesi**, facétieux |m-tseh-i| < |tseh-| |notion de rire| < \*-cèk-  
**mwidzi**, voleur |mu-ib-i| < |ib-| |notion de voler| < \*-yi'b-  
**mlewi**, ivrogne |m-lew-i| < |lew-| |notion de boire|  
**msonyi**, tailleur |m-sony-i| < |sony-| |notion de coudre|  
**mtsungi**, nageur |m-tsung-i| < |tsung-| |notion de nager|

• avec préf. cl.11 u- : activité de ..

**ulozi**, pêche |u-low-i|  
**ulezi**, éducation |u-lel-i|  
**urumishi**, utilité |u-rum-ih-i|  
**widzi**, vol |u-ib-i|  
**umani**, querelle |u-man-i| < |man-| |notion de se disputer| < \*-yúman-

• avec autres préf.: divers

**fundzi**, fundi, maître |Ø'-fund-i| < |fund-| |notion d'enseigner| < \*-tú'nd-

Les deux formes sont en variante - nous y reviendrons (étude des sons, p ); fundzi est conforme au conditionnement normal du suffixe.

Ce suffixe est le seul suffixe spécifiquement nominal pour lequel nous disposons d'un étymon (\*i', "nominal word-building suffix of agent").

ε) -e :

conditionnement :

exclu après extension causative, ce suffixe se combine avec diverses classes, mais il est peu productif; quand le terme désigne un humain, il exprime un agent non actif

• humain :

**mtrume** (cl.3/4), prophète, 'envoyé' |m-Ø<sup>h</sup>-rum-| < |rum-| (notion d'envoyer)|

**irewe** (cl.7/8, diminutif), infirme |i-r-ew-e| < |r-| (notion de jeter)| < \*-tá-

Formé sur passif : |rew-|, 'être apporté'; irewe est glossé comme 'celui qui est envoyé [par Dieu]' (cf. pp 102)

• autres :

**ntongole**, pointe |N-tongol-e| < |tongo-| (notion de tailler)| < \*-còng-  
**npbenye**, éclair |N-pveny-e| < |pveny-| (notion de filer (étoile))| < \*-pèni'-

**nkeme**, cri |N-kem-e| < |shem-|, (notion de crier)| < \*-kém-

Noter l'irrégularité de l'initiale de nkeme (la palatalisation de \*k ne se produit pas); ce mot est apparenté à -shemeza, crier, de formation régulière - ce qui permet de poser la racine |shem-|.

On relève deux formations que l'on peut interpréter comme archaïques :

**mzade**, femme qui vient d'accoucher |m-zad-e| < |zad-| (notion de mettre au monde)| < \*-bí'ád-

**mzalendro**, personne bien née |m-zal-e#ndro| < |zal-|, (notion de mettre au monde)|

|zad-|, |zal-| (et |zay-|) ; racine à plusieurs formes; la forme |zad-| n'apparaît que dans mzade; si l'on suppose une évolution dans le sens de l'affaiblissement des consonnes depuis la langue commune (\*d > Ng d + l + y + Ø), cela montre que le suffixe -e est ancien - les deux mots étant restés insensibles à l'évolution, chacun figeant un état de langue.

(Une autre hypothèse étymologique est possible, qu'il s'agisse d'une trace du suffixe verbal BC \*-i'de lui-même archaïque, donné par GUTHRIE comme "aspect of completion"; on aurait

\*-bíád- + -i'de + mzade; le maintien de l'occlusive dans cette position serait alors normal devant voyelle fermée étymologique, nous y revindrons p )

Remarque

Un certain nombre de cas où ce suffixe -e serait détectable demandent interprétation; par exemple:

**-zee**, vieux (adj.) : faut-il, comme SACLEUX en fait l'hypothèse, le rapprocher de la racine |zay-|, (notion de mettre au monde) (ci-dessus) ?

Ce mot ferait alors apparaître un suffixe -e : \*|zay-e| (on suppose alors une harmonie vocalique régressive pour aboutir à la forme réelle);

**mkatre**, pain : faut-il rapprocher de la racine |katr-|, (notion de couper) :

\*|m-katr-e|, 'ce qui est coupé' ? (\*-kant-, couper; \*-kaaté, pain)

Nous n'avons pas retenu ici les apparentements mais ils restent plausibles.

τ) -u

conditionnement :

avec racines primaires ou étendues; par ailleurs, de même que le suffixe -i, -u entraîne une modification de la finale radicale (en reconstruction), pour aboutir à une fricative:

\*k + \*u → fu ;

\*d + \*u → vu

en remplaçant les symboles BC par leurs différents reflets en Ng :

h, Ø + u → fu ;

l, y, Ø + u → vu

Ce suffixe ne forme que des adjectifs (nominaux dépendants), qui s'accordent en classe avec le nominal indépendant qui le régit - ils admettent donc (en principe) tous les préfixes de classes.

A partir de ces adjectifs, la classe 11 forme des noms de qualité.

• adjectifs :

x [-vu]

-**ledjevu**, faible < -ledjew- (notion de s'affaiblir) < \*-dègid-  
 -**polevu**, tranquille < -pol- (notion d'être calme) < \*-pód-  
 -**lemevu**, fatigué < l-lemew-ul < -leme- (notion d'être fatigué)  
 -**onvu**, maigre < -hond- (notion de maigrir) < \*-yónd-  
 -**kinaivu**, auto-suffisant < -kinay- (notion d'avoir à satiété)

x [-fu]

-**babufu**, bavard l-babuh-ul < (notion de bavarder)  
 -**nyongofu**, droit l-nyongoh-ul < -nyongo(h)- (notion de déplier)  
 -**trukufu**, noble l-trukuh-ul < -truku(h)- (notion d'élever) .

noms : cl. 11 : u- :

Tous les adjectifs ci-dessus peuvent être substantivés en classe 11 :

**unyongofu**, droiture l-nyongoh-ul  
**utrukufu**, grandeur l-trukuh-ul  
 etc

on trouve aussi quelques substantifs alors que l'adjectif n'est pas attesté:  
**uzorofu**, lourdeur, difficulté l-uzoroh-ul < -zoroh- (notion d'alourdir)

Remarque

On trouve quelques nominaux à finale /u/, incontestablement en relation avec des formes verbales: on hésite pourtant à les faire dériver du même lexème (racine) du fait qu'on ne retrouve ni la forme du suffixe précédent (avec modification de la finale radicale), ni sa caractéristique

adjectivale; en outre, le timbre de la finale peut trouver une autre explication:

**mfu**, mort ? < -f- (notion de mourir) → -fa, mourir  
 (\*-kú'-, mourir ; \*-kú' mort )

**msihu**, bâton pointu qui sert à dégager la noix de coco de son écorce  
 ? < -sih- (notion de tenir) → -sihua, enlever  
 l'écorce de la noix de coco (réversif)

-**huu**, grand ? < -hu- (notion de grandir) → -hua, grandir  
 (\*-kúd-, grandir ; \*kúdù, grand ) (cas signalé p 102)

et, sur le plan synchronique uniquement

**mdjuhu**, descendant ? < -djuh- (notion de se lever) → -djuha, se lever  
 (\*-júkúdù, petit-enfant )

Dans ces mots, différentes raisons militent en faveur d'une interprétation lexicale (et non grammaticale) de la voyelle finale: le timbre /u/ peut en effet s'expliquer par:

i) harmonie de la finale avec la voyelle radicale: -**huu**, **mdjuhu**,

ii) voyelle lexicale ou relevant d'un morphème non suffixiel:

**msihu** (où /u/ peut être la voyelle du réversif);

-**huu**, **mfu** : -fa a pour voyelle spécifique à l'accompli /u/ (ci-dessus p 110)

(pour mémoire) iii) étymologie:

-**huu** < \*-kúdù;

**mfu** < \*-kú' mort

-**djuhu** < \*-júkúdù

Dans ces cas, nous ne considérons donc pas cette finale /u/ comme un suffixe et ne rapprochons pas ces mots à travers un lexème qui leur serait commun - nous partons de sème)

Par contre, dans le cas précis des adjectifs en -u, nous admettons systématiquement la décomposition en racine, nous fondant sur la valeur particulière de ce suffixe et sur la généralité de cette dérivation:

\* le plus souvent, cette analyse est confirmée par l'existence effective de verbaux apparentés, la racine apparaissant souvent complexe:

-**babufu**, bavard ⇨ -**babuha**, bavarder → |**-babuh-**|

-**ledjevu**, faible ⇨ -**ledjea**, s'affaiblir → |**-ledjew-**|

-**kinaivu**, suffisant ⇨ -**kinaya**, suffire → |**-kinay-**|

\* nous n'avons relevé que trois adjectifs en -u (sur 14 dans le corpus) sans verbe relié attesté; dans ce cas, nous posons une racine hypothétique synchronique notée +:

-**sinbifu**, opprimé, → +|**-sinbih-**|

-**renbavu**, mou → +|**-renbay-**|

-**ivivu**, envieux → +|**-ivih-**|

On note dans ce dernier cas l'irrégularité de la forme [vu], alors qu'on attendrait [fu]; il peut s'agir d'une assimilation de sonorité à la consonne radicale précédente.

**Récapitulation**

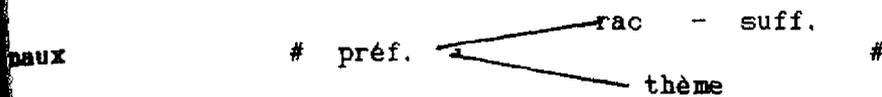
Le tableau ci-dessous présente les différents suffixes, en liaison avec les préfixes de classe en compagnie desquels ils apparaissent (nous n'avons pas fait figurer ici les extensions; les classes sont indiquées par la forme la plus courante de leur préfixe & leur numéro).

préf. cl.	rac	suffixes	valeur
m-/mi- (cl. 3/4) Ø <sup>*</sup> -/ma- (cl. 5/6) i-/zi- (cl. 7/8) N-/N- (cl. 9/10)	- ... -	o	nom d'action
m-/wa- (cl. 1/2)	- ... -	a (+ complément) adji	agent
u- (cl. 11)	- ... -	i, adji	action
divers	- ... -	e	agent passif <sup>1</sup>
...	- ... -	u	adjectifs
hu- (cl. 15)	- ... -	a	infinitif
m-/mi- (cl. 3/4) Ø <sup>*</sup> -/ma- (cl. 5/6) i-/zi- N-/N- (cl. 9/10)	- ... -	a	divers

<sup>1</sup> quand il s'agit d'humains

) *préfixes nominaux*

Ainsi que nous l'avons vu, les préfixes sont communs à tous les nominaux, qu'ils soient dérivés de racines ou de thèmes:



La présence d'un préfixe d'un certain type (préfixe nominal) est d'ailleurs l'un des facteurs principaux d'identification des nominaux comme tels.

Les préfixes définissent ce qu'il est convenu d'appeler 'le système de classes nominales', où rentrent tous les nominaux, et dont l'existence est une des conditions de 'bantuité' d'une langue. Les préfixes caractérisent la classe à laquelle appartient le nominal, ils ne marquent pas seulement le nominal mais aussi tous les mots qu'il régit; cet ensemble de morphèmes constitue ce qu'on appelle le schème d'accord, caractéristique et définitoire de chaque classe; en effet, alors que le préfixe nominal peut se retrouver identique dans différentes classes (ex. cl. 1 & 3, 9 & 10), le schème d'accord sera, lui, toujours différent.

Sur le plan morphologique, chaque 'classe' se caractérise donc par :

- i) le (ou les) préfixe(s) des nominaux indépendants qui lui sont propres,
- ii) son 'schème d'accord' (l'ensemble des marques d'accord régies, qui se porte sur tous les mots 'variables' dans la dépendance d'un nom, y compris les adjectifs ou nominaux dépendants)
- iii) éventuellement une opposition de nombre (voir point suivant).

Il est coutume, depuis les premières études sur les langues bantu, de présenter les classes nominales dans un ordre précis, et de leur attribuer un numéro qui fait référence à l'ensemble des classes inventoriées dans le domaine bantu: le schéma de GUTHRIE, par exemple, prévoit 19 classes possibles, dont chaque langue actualise qu'un certain nombre; dans une étude particulière, la numérotation est donc relative à celle de l'ensemble, ce qui explique qu'elle ne soit pas continue.

Nous suivons cette habitude, qui fournit un cadre commode où s'insèrent les faits grand-comoriens et facilite grandement les études comparatives; cela explique que notre numérotation saute de 11 à 15, les classes intermédiaires, 12, 13, 14 n'étant pas actualisées en shingazidja.

Le principe de numérotation des classes fait que, jusqu'à la classe 11 du moins, les classes impaires sont des classes singulier, les classes paires, pluriel; en principe, le pluriel d'une classe est formé dans la classe immédiatement supérieure.

La plupart des 'classes' se regroupent, par deux ou trois, sur la base de l'opposition de nombre: on a coutume de parler alors de 'genre', mais nous n'utiliserons généralement pas cette terminologie, source, nous semble-t-il de confusion, en raison de la divergence avec l'acception usuelle du terme de genre et parce que toutes les classes où se forme le pluriel d'une autre ne se réduisent pas à ce rôle de contre-partie.

Avant d'étudier chaque classe à travers la combinatoire du préfixe (dans ses différents allomorphes) avec les lexèmes, nous présentons, sous forme de tableau, l'inventaire des préfixes des nominaux indépendants complétés par les préfixes d'accord, puis envisageons la question de la motivation des classes.

#### *α) Inventaire des préfixes de classe*

La première liste donne, pour chaque classe, la forme structurelle du préfixe des nominaux indépendants, le ou les allomorphes principaux et l'étymon chez GUTHRIE.

La seconde liste donne les principaux accords de classe, en fonction des catégories grammaticales des éléments.

#### *1) Liste des préfixes des nominaux indépendants*

GUTHRIE distingue des classes dont le préfixe a plusieurs allomorphes conditionnés ("prefix clusters") et des classes où les allomorphes du

préfixe sont en distribution lexicale (non conditionnée) ("composite class"); pour ces dernières, il définit une forme principale et des formes secondaires, numérotées Xa, Xb, etc (où X = numéro de la classe). (GUTHRIE 1956 : 80 & 87)

Cette catégorisation ne nous est pas apparue opératoire en shingazidja: si les préfixes possèdent en général des allomorphes (où la distinction conditionnée vs non conditionnée n'est d'ailleurs pas toujours claire), nous nous écartons ici de la lettre guthrienne et ne constituons pas des classes Xa, etc; les différents allomorphes de chaque préfixe sont listés sous leur classe: ainsi, nyi- ne constituent pas une classe 10a mais sont donnés après le préfixe N- de cl 10 (nous en verrons les raisons plus bas).

1	lmu-1 : m-, mw-	< *mu
2	lwa-1 : wa-	< *ba
3	lmu-1 : m-, mw-	< *mu
4	lmi-1 : mi-	< *mi
5	lØ¹-1 :	< *yi¹
	ldzi-1 : dzi-	< *di¹
	ldji-1 : dji-	( < *gi)¹
6	lma-1 : ma-	< *ma
7	lshi-1 : shi-	
	lhi-1 : hi-	< *ki
	li-1 : i-	
	lki-1 : ki-	
8	lzi-1 : zi-	< *bi¹
9	lN-1 : n(C)	< *ny²
10	lN-1 : n(C)	< *ny
	lnyi-1 : nyi-	
11	lu-1 : u-	< *du & *bu
15	lhu-1 : hu-, u-	< *ku
16	lpvo-1 : pvo-	< *pa
17	lho-1 : ho-	< *ku
18	lmo-1 : mo-	< *mu

1 \*gi' + Ng dji- ; cl. 21, voir plus bas, p

2 Nous nous basons exclusivement sur les formes données par GUTHRIE dans sa liste des étymons BC (nous revenons plus loin sur ce point, p 14)

Les formes réalisées des préfixes sont le plus souvent identiques à leurs formes structurelles; rappelons

cl. 1 & 3 : m-, nasale syllabique

cl. 5 : |Ø'-| : ce préfixe rend l'initiale lexicale occlusive; il n'a pas de support segmental isolable;

cl. 9 & 10 : nC représente une prénasalisée, amalgame entre le préfixe |N-| et l'initiale lexicale

2) Liste des préfixes des nominaux dépendants (adjectifs)

Les préfixes des nominaux dépendants sont identiques à ceux des nominaux indépendants, sauf aux classes reprises ci-dessous:

7 |shi-| : shi-

li-| : i-

Ces deux allomorphes sont souvent en variante pour un même lexème

9 |N-| : nC

lnyi-| : nyi-

Allomorphes qui peuvent être en variante

11 |mu-| : mu-

cl. 7 : l'allomorphe hi- n'est pas attesté;

cl. 11 : les nominaux dépendants prennent ici l'accord de la classe 3

β) Liste des accords de classe

Nous donnons ici les principaux accords de classe, autres que les accords nominaux: déictique (pré-préfixe ou augment), préfixe verbal sujet (à l'accompli non relatif), infixe objet, accord de démonstratifs.

Les éléments donnés ici ont été exposés dans LAFON 1985

		démonstratif				
préf. NI	déictique (pré-préf.)'	préfixe verbal sujet	inf. objet	proximité -nu	rappel -o	
1	m(u)-	(y)e	ha-	-m-	oyi <sup>2</sup>	uwo
2	wa-	(w)o	wa-	-wa-	wanu	wao
3	m(u)-	(w)o	u-	-m-	unu	uwo
4	mi-	(y)e	i-	-i-	inu	iyo
5	Ø'-	(y)ele	li-	-li-	linu	lilo
		le				
6	ma-	(y)e	ya-	-ya-	yanu	yayo
7	shi-	(e)she (y)e	shi-	-shi-	shinu	isho
8	zi-	(e)ze	zi-	-zi-	zinu	izo
9	N-	(y)e	i-	-i-	inu	iyo
10	N-	(y)eze	zi-	-zi-	zinu	izo
		ze				
11	u-	(w)o	u-	-m-	unu	uwo
15	hu-	(h)o	hu-		unu	uwo
16	pvo-	(pva-)	pva-		pvanu	yapvo
17	ho-		hu-		hunu	yiho
18	mo-		mu-		pvanu	yimo

1 pré-préfixes : ont une valeur de déterminant; ils possèdent des formes à initiale vocalique ou continues, plus ou moins conditionnée par le contexte (forme à initiale consonantique après pause, forme à initiale vocalique, venant souvent s'amalgamer avec la voyelle finale du mot précédent, en l'absence de pause); en outre, aux cl. 5, 7, 8 & 10 il existe une forme longue et une forme brève: en cl. 7, la forme est en relation avec celle du préfixe (préf. hi-, i- → pré-préf. she; préf. shi- → pré-préf. e :

eshehitswa, la tête leshe-hi-tswa  
 esheitranda, le lit leshe-i-tranda  
 eshiba, l'os le-shi-ba

2 : oyi, forme non régulière du démonstratif -nu en cl. 1

Exemples:

donnons quelques exemples d'énoncés montrant ces différents éléments  
 ordés :

- cl. 1 : **emwana uwo hafariki**, cet enfant est mort  
 le-mu-ana#uw-o#ha-farikil, pré-préf. cl. 1 - préf. cl. 1 - thème # préf.  
 accord cl. 1 - dém. # préf. verbal sujet cl. 1 - rac.
- cl. 2 : **owandru wanu ngwao wawade**, ces gens sont malades  
 lo-wa-ndru#wu-nu#ng-wa-o#wu-wadei, pré-préf. cl. 2 - préf. cl. 2 - thème #  
 préf. cl. 2 - dém. # présentatif - préf. cl. 2 - dém. # préf. cl. 2 - thème
- cl. 4 : **emiri inu idanguha**, ces arbres sont tombés  
 le-miri#i-nu#i-danguh-a, pré-préf. cl. 4 - préf. cl. 4 - thème # préf.  
 accord cl. 4 - dém. # préf. verbal sujet cl. 4 - rac. - suff. verbal
- cl. 5 : **lebwe linu linyongoha**, cette pierre est allongée  
 lle-Ø-we#li-nu#li-nyongo-h-a, pré-préf. cl. 5 - préf. cl. 5 - thème # préf.  
 accord cl. 5 - dém. # préf. verbal sujet cl. 5 - rac. - ext. - suff.v.
- cl. 6 & 16 : **ngapvo malaho mendji**, il y a beaucoup de maisons  
 ng-apv-o#ma-laho#ma-indjii, présentatif - préf. cl. 16 - dém. # préf. cl. 6  
 - thème # préf. ND cl. 6 - thème
- cl. 9 : **enuni iindjia ho-nyunbani**, l'oiseau est entré dans la maison  
 le-N-nuni#i-indji-a#ho-N-nyunba-nii, pré-préf. cl. 9 - préf. cl. 9 - thème #  
 préf. verbal sujet cl. 9 - rac. - suff.v. # pré-préf. cl. 17 - préf. cl. 9 -  
 thème - suff. locatif
- cl. 10 : **zenuni zindjia ho-nyunbani**, ces oiseaux sont entrés dans les maisons  
 lze-N-nuni#zi-indji-a# ho-N-nyunba-nii, pré-préf. cl. 10 - préf. cl. 10 -  
 thème # préf. verbal sujet cl. 10 - rac. - suff.v. # pré-préf. cl. 17 - préf.  
 cl. 10 - thème - suff. locatif

γ) 'valeur' des préfixes : motivation des classes

La motivation sémantique des classes, c'est-à-dire leur adéquation éventuelle à des catégories sémantiques, a été l'objet de moult débats, les uns voyant ou voulant voir des corrélations, d'autres, le refusant.

Dans sa définition d'ensemble du système de classe, GUTHRIE refuse toute corrélation, tout en reconnaissant que l'appartenance de classe est "importante" pour les nominaux indépendants:

"as it is in one type of word only, the independent nominal, that the class is an integral characteristic of the word, it is in respect to these nominals that the assorting into classes is important." (GUTHRIE, op. cit. 79)

Cela ne fait pas justice cependant du fait que l'on constate, dans les termes de fonds bantou du grand-comorien comme dans les autres langues bantou, une indéniable spécialisation sémantique, pour certaines classes du moins: ainsi que l'écrit ALEXANDRE

"il est très fréquent de relever, dans un genre donné, une proportion statistiquement significative de substantifs relevant effectivement d'une catégorie sémantique donnée". (ALEXANDRE 1981 : 357)

Cette caractéristique, qui n'a de sens bien entendu que pour les nominaux indépendants, se manifeste tant avec ceux issus de thèmes qu'avec ceux issus de racines (pour ces derniers, en liaison avec le suffixe): il paraît donc légitime d'imputer aux préfixes la responsabilité de cette valeur sémantique.

Cela est particulièrement clair lorsqu'un même lexème donne plusieurs nominaux:

-lo- | notion de pêcher | < \*-dób-

- avec préf. 1 / 2 'animé':  
mlozi (wa-), pêcheur |m-/wa-low-i|
- avec cl. 7 / 8 'instrument, objet':  
iloo (zi-), hameçon |i-/zi-low-o|
- avec préf. cl. 11 'abstrait, activité':  
→ ulozi, pêche |u-low-i|
- avec préf. cl. 15 : infinitif:  
uloa, pêcher |hu-low-a|

-fany- | notion de faire |

- avec préf. 1 / 2 'animé':  
mfanya-hazi (wa-), travailleur |m-/wa-fany-a#0-hazi|
- avec préf. 3 / 4 'action':  
mfanyo (mi-), acte |m-/mi-fany-o|
- avec préf. 15 infinitif:  
ufanya, faire |hu-fany-a|

-ndru | notion d'existant | < \*-ntù

- avec préf. 1 / 2 'animé':  
mndru (wandru), personne (gens) |mu-/wa-ndru|
- avec préf. 5 'augmentatif':  
djindru (ma-), géant |dji-ndru| / |ma-dji-ndru|
- avec cl. 7 / 8 'objet':  
hindru (zindru), chose/s |hi-/zi-ndru|

-buzi | notion de caprin | < \*-búdi\*

- avec préf. cl. 9 / 10 : 'animaux':  
nbuzi (-), chèvre |N-buzi|
- avec préf. cl. 5 / 6, : 'augmentatif':  
buzi (ma-), grosse/s chèvre/s |0°-buzi|

Le tableau suivant recense les principales valeurs sémantiques que l'on peut

attribuer aux classes, et donc aux préfixes:

-l / |wa-l : préf. cl. 1 / 2 en opposition de nombre :

être humain

lmu- / lmi- : préf. cl. 3 / 4 en opposition de nombre :

arbres, plantes, parties du corps; action ('nom verbal')

HAIG DER-HOUSSIKIAN a suggéré le trait "sentient" pour rendre compte de ces notions ("*biological entities which are neither human nor animal; plant life in general*" (HAIG DER-HOUSSIKIAN 1974 : 17); ce trait correspondrait peut-être à 'matière vivante non animée' ? En shingazidja, il est remarquable que les plantes sont très généralement dans ces deux classes (une seule exception, le bananier, trindi, en cl. 5)

lØ'- / lØj- / lma- : préf. cl. 5 / 6 en opposition de nombre :

augmentatif ; animaux, objets divers; action

lma- préf. cl. 6 : liquide

lN- / lN'- : préf. cl. 9 / 10 en opposition de nombre :

animaux, objets divers

lshi- / lhi- / li- / lzi- préf. cl. 7 / 8 en opposition de nombre :

parties du corps; objet, ustensile, instrument; diminutif

lu- préf. cl. 11 :

nom abstrait, activité, unité (un terme pris dans un tout)

lhu- préf. cl. 15 : infinitifs

lpvo- préf. cl. 16 : lieu / moment où l'on est

lho- préf. cl. 17 : lieu où l'on va classes dites locatives

lmo- préf. cl. 18 : à l'intérieur de

Cet aspect du système de classes nominales ne doit toutefois pas suggérer que la combinatoire entre préfixe et lexème soit libre au sens où tout préfixe pourrait s'adjoindre à n'importe quel lexème pour former un nom, dès lors qu'apparaît une compatibilité sémantique:

x dans le cas des thèmes, le plus souvent, un seul préfixe est attesté (ou deux en opposition de nombre);

x dans le cas des racines, comme nous l'avons vu, la latitude combinatoire est plus grande.

Cette relative motivation sémantique des classes joue un grand rôle, comme on le verra, dans l'intégration des nominaux d'origine étrangère.

#### δ) système de classe

Nous présentons, pour chaque classe nominale, un aperçu de son contenu sémantique éventuel, son ou ses préfixes pour les NI et les ND, la forme reconstruite du BC correspondante et des exemples issus de racines et de thèmes.

Nous reviendrons *in fine* sur le phénomène dit d'alternance consonantique, qui se produit au contact du préfixe de classe et de l'initiale lexicale, ainsi que sur les accords adjectivaux.

Les exemples suivent l'ordre des classes; les classes 'pluriel' reprennent, quand c'est possible, les nominaux donnés au singulier; les adjectifs sont donnés en syntagmes, c'est-à-dire avec un nom auquel ils sont accordés; nous indiquons les lexèmes (racines ou thèmes) dont les nominaux sont issus.

Quoique cette partie traite du seul fonds bantu, nous l'avons pris ici dans son acception large, celle de tous les termes non empruntés: nous admettons donc ici les termes sans étymon, pourvu qu'il n'apparaisse aucune raison de les suspecter d'être d'origine étrangère. De ce fait, nous n'indiquons pas systématiquement les étymons éventuels des lexèmes.

cl.1 : NI & ND : m-, mw-, (mu-) < \*mu-

\* préfixe :

mu- exceptionnellement (un accord de ND)

mw- devant voyelle

m- devant toute consonne (y compris /m/) et exceptionnellement devant voyelle; se syllabise devant toute consonne sauf /w/ (nasale syllabique)

contenu sémantique:

ce qui désigne les humains en tant que créatures animées, agissantes, notamment les noms d'agent formés sur les racines avec suffixe -a, -adii ou -i.

exemples:

\* mu- : un seul mot

(andru) mudu, (une personne) noire (et non \*mdu)

Cette forme s'explique sans doute par son ancienneté, à double titre; l'étymon du lexème est \*yí'dù, vraisemblablement accentué sur la première syllabe, et mu- est sans aucun doute la forme originelle

du préfixe; l'allomorphe *mu-* ici employé résulte d'une influence de la structure di-syllabique étymologique, et de la place de l'accent sur la pénultième; on peut imaginer une évolution d'une forme *\*mu-* à *\*mwidu*, puis *'mudu*.

S'y ajoute actuellement le souci d'éviter une confusion éventuelle avec *[mdu]*, pied (*andu*)

x *mw-* :

*mwana mtiti*, un petit enfant *[mu-ana]* < *-ana* (notion d'enfant); *-ti* (notion de petitesse)

*mwongozi*, responsable, guide < *-ongoz-* (notion de diriger)

*mzungu mwema*, un européen gentil *[mu-zungu#mu-ema]* < *-zungu* (notion d'européen); *-ema* (notion de bon)

x *m-* :

*mfanya-hazi*, travailleur *[mu-fany-a#Ø-hazi]* < *-fany-* (notion de faire); *-hazi* (notion de travail)

*mlimadji*, paysan *[mu-lim-adji]* < *-lim-* (notion de cultiver)

*mwezi mdzima*, un seul mendiant *[mu-mey-i#mu-dzima]* < *-mey-* (notion de mendier); *-dzima* (notion d'unité)

*ndru*, un homme *[mu-ndru]* < *-ndru* (notion d'existant)

*Mungu*, Dieu *[mu-(u)ngu]* < *-(u)ngu* (notion de divinité)

Thème difficile à définir dans la mesure où ce mot est réalisé par *[nungu]* ou *[m-gu]*, et où il ne connaît pas d'autres attestations (notamment, pas de pluriel)

classe 2 : NI & ND : *wa-*, *w-* < *\*ba*

α préfixe :

*wa-* devant consonne

*w-* ou *wa-* devant voyelle, en répartition lexicale (voir voyelles, supra)

β contenu :

pluriel de la classe 1

γ exemples:

*wandru wadu*, des personnes noires *[wa-ndru#wa-du]*

*wana watiti*, de petits enfants *[wa-ana#wa-titi]*

*waongozi, wongozi*, des responsables *[wa-ongozi]*

*wazungu wema*, de gentils européens *[wa-zungu#wa-ema]*

*wafanya-hazi*, travailleurs *[wa-fany-a#Ø-hazi]*

*walimadji*, paysans *[wa-lim-adji]*

*wamezi waili*, deux mendiants *[wa-mey-i#wa-ili]* < *-ili* (notion de double)

classe 3 : NI & ND : *m-*, *mw-*, *m-* < *\*mu*

α préfixe : identique à celui de la cl. 1 :

*mu-* dans un accord de ND

*mw-* devant voyelle

*m-* syllabique devant toute consonne (sauf /w/); exceptionnel devant voyelle

contenu sémantique:

- nom d'action (sur racine nomino-verbale, avec suffixe -o)
- nom de plantes, d'arbres (presque systématiquement)
- nom de parties du corps (cf. cl. 7)
- divers

exemples:

x **mu-** : un seul cas

(**ri**) **mudu**, (un arbre) noir |(**mu-ri**)#**mu-du**!

Même remarque que ci-dessus

x **mw-** :

**wendeleo**, progrès |**mu-end-ele-o**! < -endele- |notion d'avancer! (intensif de -end-)

**waha mwema**, une bonne année |**mu-aha#mu-ema**! < -aha |notion d'année!

**mbuu mweu**, un baobab blanc |**mu-buu#mu-eu**! < -buu |notion de baobab!; -eu |notion de blancheur!

x **m-** :

**mfanyo**, action |**mu-fany-o**! < -fany- |notion de faire!

**mpango mhuu**, un programme important |**mu-pang-o#mu-huu**! < -pvang- |notion d'organiser!; -huu |notion de grandeur!

**mlalo mdziro**, un hébergement difficile |**mu-lal-o#mu-dziro**! < -lal- |notion de dormir!; -dziro |notion de lourdeur!

**mwunbo**, création |**mu-wunb-o**! < -wunb- |notion de créer!

Analgane ici entre la voyelle du préfixe et la continue initiale du lexème; le résultat serait le même si la racine était à initiale vocalique (les raisons pour considérer que la racine possède un [w] initiale est le nominal de cl. 5 **bunbo**, création, où la consonne alterne avec w.) (voir alternance plus loin, et surtout étude des sons, sur la réflexion des occlusives étymologiques initiales, p )

**ri mnene**, un gros arbre |**mu-ri#mu-nene**! < -ri |notion d'arbre!; -nene |notion de grosseur!

**o**, main |**mu-hono**! < -hono |notion de main!

La classe 3 se distingue de la cl. 1 par les accords (autres qu'adjectivaux)

qu'elle commande ainsi que par la formation du pluriel:

**enzungu woyi hahimi**, cet européen est debout : cl. 1

↔ **omri unu uhimi**, cet arbre est debout : cl. 3

cl. 1 → pluriel cl. 2 ↔ cl. 3 → pl. cl. 4

Les raisons nous paraissent suffisantes, compte-tenu de notre définition des classes qui ne se limite pas au préfixe, pour instituer des classes différentes, contrairement à de nombreux auteurs qui rassemblent sous un même numéro ces classes sur la base de leur analogie préfixielle [et compliquent ainsi l'analyse, sans, nous semble-t-il, qu'il en découle des avantages, par ex. HAIG & HOUSSEKIAN, déjà cité, et LUMWANU MCMLXX (1970)]

Les noms de la classe 11 commandent les accords de la classe 3, notamment pour les adjectifs

classe 4 : NI & ND : mi-, m- < \*mi

α préfixe :

m- devant voyelles

mi- devant consonnes

β contenu:

pluriel de 3 (quand il est possible)

γ exemples:

**miri midu**, des arbres noirs |mi-ri#mi-du|

**maha mema**, de bonnes années |mi-aha#mi-ema|

**mibuu meu**, des baobabs blancs |mi-buu#mi-eu|

**mifanyo**, actions |mi-fany-o|

**mipango mihuu**, de grands programmes |mi-pang-o#mi-huu|

**miri minene**, de gros arbres |mi-ri#mi-nene|

**mihono**, des mains |mi-hono|

classe 5 : cette classe connaît différents préfixes dont la distribution est au moins partiellement lexicale.

α préfixes:

dj- (dz-) devant voyelles

Ø-, dzi-, dji- devant consonnes (répartition lexicale)

(le rapport entre les allomorphes, qui peuvent être en variation pour un même lexème, est traité ci-dessous)

β contenu sémantique:

• nom d'action ou du résultat de l'action exprimée par une racine (avec suffixe -o; cf. cl. 3/4);

• augmentatif

• noms de profession, de métier, termes de parenté (lorsqu'ils ne sont pas en cl. 1/2)

• divers

β exemples:

\* **dzi-**, (dz-) < \*di' : forme vestigielle: quelques NI provenant de thèmes; un seul ND

**dzinyo**, une dent |dzi-nyo| < -nyo |notion de dent| (< \*yí'nò)

**dzitso dzidu**, un oeil noir |dzi-tso#dzi-du| < -tso |notion d'oeil| (< \*yí'cò)

Les thèmes des NI sont identifiés par leur pluriel (manyo, matso); dzi- est bien un préfixe; on remarque dans ces deux cas la chute de la première syllabe étymologique.  
Pour dzidu, cette forme ajoute un argument à l'hypothèse d'une lexicalisation ancienne (voir ci-dessus, p 131)

(dzina, ci-dessous)

\* dji-, dj- : NI & ND provenant de lexèmes généralement à initiale vocalique ou prénasalisée; pour les ND, également lorsque l'initiale lexicale n'alterne pas; pour les NI, contrairement au fonctionnement normal des préfixes, dji- se maintient au pluriel en cl. 6 (ci-dessous) :

djapizo, malédiction |dj-apiz-ol| < -apiz- |notion de maudire|

djindru djema, un bon géant |dji-ndru#dji-ema| < -ndru |notion d'existant|; -ema |notion de bon|

djitretrezi, bavard |dji-tretr-e-i| < -tretr- |notion de bavarder|

djilemevu, fatigué |dji-leme-ul| < -lem- |notion d'être fatigué

En variante rare demevu, tretrezi, avec initiale alternante; point développé ci-dessous.

\* Ø'- (+ occlusive) < \*yi' : NI & certains ND

panbo, ce qui enveloppe |Ø'-panb-ol| < -pvanb- |notion d'envelopper|

dima, agriculture |Ø'-lim-a| < -lim- |notion de cultiver|

buzi dume, bouc (caprin mâle) |Ø'-buzi#Ø'-ume| < -buzi |notion de caprin; -ume |notion de mâle|

koko, grand-mère |Ø'-hoko| < -hoko |notion de grand-mère|

paha tshe, chatte |Ø'-pvaha#Ø'-she| < -pvaha |notion de chat; -she |notion de femelle|

bwe titi, une petite pierre |Ø'-we#Ø'-titi| < -we |notion de pierre; -titi |notion de petitesse|

daho kuu, une grande maison |Ø'-laho#Ø'-huu| < -laho |notion de maison; -huu |notion de grand|

tranga, voile |Ø'-rangal| < -ranga |notion de voile|

djuai, oeuf |Ø'-djuai| < -djuai |notion d'oeuf|

L'initiale du mot est identique à l'allo-morphe dji- de cette classe; mais rien n'indique qu'il y ait eu confusion; pluriel madjuai, étymon \*-jáyí

dzina, nom |Ø'-dzinal| < -dzina |notion de nom| (< \*-yí'nà, \*-gí'nà)

dzina ; comme dans le cas précédent, la syllabe initiale n'est pas préfixielle - ce qui apparaît dans la formation du pluriel (madzina); le thème est bien l-dzinal, il n'y a aucun argument synchronique pour isoler un thème \*-na.

Si l'on retient pour étymon \*-yí'nà, par comparaison avec les deux cas étymologiquement similaires dzinyo < \*-yí'nò et dzitso < \*-yí'cò, où dzi- est préfixiel (pl. manyo, matso), cela fait supposer une lexicalisation du préfixe; cela suppose donc une histoire du mot différente; cela peut aussi être un argument pour préférer ici l'étymon 'occulant' \*-gí'nà (voir aussi p ).

siwa dene, une grande île |Ø'-siwa#Ø'-nene| < -siwa |notion d'île|

dene ; le thème est bien l-nene|, ce qui apparaît par tous les autres accords de classes (mnene, cl. 1 & 3, ci-dessus, etc); ce cas est discuté plus bas

L'utilisation des allomorphes dzi- ou dji- apparaît purement lexicale (et non conditionnée): en fait, dzi- est résiduel et ne se trouve plus que dans quelques mots, alors que dji- est vivant.

On remarquera que le préfixe que nous avons symbolisé par  $\emptyset^*$  entraîne, dans le cas de lexèmes à initiale continue (pv, w, tr, l, h) la modification de la consonne initiale, qui de continue devient occlusive: cela est dû à la règle d'alternance consonantique, qui s'explique en diachronie par la présence d'une voyelle fermée présente dans l'étymon du préfixe ... (voir chapitre suivant, étude des sons, pp ).

D'un point de vue synchronique, la caractéristique fondamentale de cette classe, qui découle de ce mécanisme morpho-phonologique n'est pas un préfixe à proprement parler, mais plutôt une 'marque' ou un 'indice' préfixiel, lié au conditionnement morpho-phonologique de l'initiale lexicale :

x quand l'initiale lexicale n'est ni continue, ni nasale (ni prénasalisée), et préférentiellement occlusive, le préfixe est  $\emptyset^*$ .

x lorsque l'initiale lexicale est une continue (pv, w, tr, l, h), la nasale /n/ ou une prénasalisée, la langue recourt alors à un préfixe syllabique, dji- (ou dzi-).

Cela est particulièrement apparent pour les adjectifs:

i) si l'initiale lexicale est vocalique, nasale, ou prénasalisée, ils seront construits avec l'allomorphe syllabique dji-, ou induiront des reconstructions du thème:

ainsi des formes

**djeu, djema**, à partir des thèmes -eu, (notion de blancheur), -ema, (notion de bonté) : |dji-eu|, |dji-ema|.

**djinkudu**, rouge, à partir de -nkudu, (notion de rougeur) : |dji-nkudu|, et, avec substitution de l'initiale lexicale (reconstruction d'un thème),

**dene, dono** (en variante **djinono**), à partir de -nene, (notion de grand), -nono, (notion de bonne santé) : | $\emptyset^*$ -nene|, | $\emptyset^*$ -nono|.

Nous parlons dans ce dernier cas de reconstruction de thème plutôt que d'alternance car n et d ne sont pas en relation d'alternance dans le restant de la langue; par ailleurs, les thèmes -nene et -nono sont bien établis à travers les autres accords de classe (cl. 1 & 3, **nnene, nnono**; cl. 4 **minene, minono**, etc); enfin, ces formes à initiale occlusive servent de base, on le verra, aux accords de cl. 9/10 avec préf. **IN-**.

ii) si l'initiale est occlusive, l'allomorphe syllabique et le préfixe  $\emptyset^*$  peuvent être en variante:

**djitretrezi** ou **tretrezi**, bavard |dji-tretr-e-i| ou |Ø'-tretr-e-i| < -tretr-  
 notion de bavarder|

**djilemevu** ou **demevu**, fatigué |dji-leme-u| ou |Ø'-leme-u| < -leme-  
 notion d'être fatigué|

**djiwade** ou **bwade**, malade |dji-wade| ou |Ø'-wade| < -wade, notion de  
 maladie|

Dans certains cas, cette variation peut être significative: ainsi, la forme **bwade** sera utilisée en cl. 5 avec des animés plutôt que **djiwade**:

**lekoko ngilo bwade**, la grand-mère est malade (plutôt que **lekoko ngilo djiwade**, qui serait très péjoratif, la personne étant en quelque sorte réifiée (un autre accord, dit notionnel, est aussi possible, qui fera accorder l'adjectif en cl. 1 des animés: **lekoko nge bwade**; c'est sans doute la forme la moins connotée)

Le comportement des adjectifs en cl. 9/10 est souvent parallèle à celui de cl. 5 décrit ici.

Nota

Tous les adjectifs ne permettent pas cette polymorphie: ainsi, |l-huul|, notion de grandeur|, |l-itsil|, notion de maturité|, etc, s'accordent toujours avec initiale occlusive: **kuu**, **bitsi**.

Comme nous avons signalé, parmi les valeurs de cette classe, une valeur augmentative: cela signifie que, lorsqu'un nominal est normalement dans une autre classe que la 5, il sera souvent possible de le placer (par changement de préfixe ou modification de l'initiale) en classe 5 pour lui conférer valeur augmentative:

ainsi, **nbuzi**, chèvre, est normalement en cl. 9: en classe 5 **buzi** signifie grosse chèvre, et s'emploie surtout pour désigner le bouc **buzi dume**);

**ndru**, personne, est en cl. 1; en cl. 5, **djindru** signifie géant.

Les adjectifs de cl. 5 font leur pluriel en classe 6 mais il est remarquable que les termes gardent alors la marque de cl. 5 (initiale occlusive ou préf. -):

**mbuzi maume**, des boucs |ma-Ø'-buzi#ma-ume|

**madjindru**, des géants |ma-dji-ndru|

Ce mécanisme est exclu pour les nominaux normalement en cl. 5; mais, lorsqu'ils ont une initiale alternante (occlusive en cl. 5, continue en cl. 6), un augmentatif peut être formé en cl. 6 par maintien de l'occlusive initiale (au lieu de la continue):

**daho**, maison → **malaho**, des maisons

→ **madaho**, de grandes maisons |ma-Ø'-laho|

**paha**, chat → **mapvaha**, des chats

→ **mapaha**, de grands chats |ma-Ø'-pvaha|

De l'existence de cette valeur augmentative on peut inférer l'hypothèse suivante, qui explique aussi la co-occurrence des deux allomorphes syllabiques *dzi-* et *dji-* du préfixe de cette classe

*dzi-* est sans doute le reflet normal de BC \*di', préf. cl. 5; *dji-* correspond lui au préfixe de la classe 21 augmentative \*gi (source bibliographique, ci-dessous p 261); comme par ailleurs le phonème /dz/ a été concurrencé par /dj/ introduit à la suite de l'influence arabe, il a pu se produire une confusion entre /dj/ et /dz/, qui a entraîné une confusion des classes \*21 et \*5, ce qui a entraîné une valeur augmentative dans la classe 5. /dz/ ne reste alors plus que dans des mots anciens, ce que montrent les quelques formes où il est attesté (l'adjectif -du, noir, dont l'ancienneté est confirmée par les accords de cl. 1 & 3, les NI où le préfixe *dzi-* est devenu partie du lexème *dzinyo*, *dzitso*)

classe 6 : NI & ND : ma-, m- < \*ma

α préfixe :

m- ou ma- devant voyelles

ma- devant consonnes

β contenu :

- pluriel des classes 5 et 11 (voir aussi 10)
- nom de masse, non-dénombrable (en particulier liquides)
- pluriel généralisateur pour nominaux de toutes classes (pluralisateur)

γ exemples :

\* pl. cl. 5 :

**manyo**, dents |ma-nyo|

**matso madu**, des yeux noirs |ma-tso#ma-du|

**madjapizo**, malédictions |ma-dji-apiz-o|

Noter ici le maintien du préf. de cl. 5 au pluriel

**mahoko matretrezi**, des grand-mères bavardes |ma-hoko#ma-tretre-i|

**mahoko malemevu**, des grand-mères fatiguées |ma-hoko#ma-lene-u|

**mapvanbo**, enveloppes |ma-pvanb-o|

**mapvaha mashe**, chattes |ma-pvaha#ma-she|

**mawe matiti**, petites pierres |ma-we#ma-titi|

**malaho mahuu**, de grandes maisons |ma-laho#ma-huu|

**maranga**, voiles |ma-ranga|

**madzina**, noms |ma-dzina|

**masiwa manene**, de grandes îles |ma-siwa#ma-nene|

\* pl. des augmentatifs : rétention du préf. de cl. 5 (sur ce point, voir p 261)

**madjindru mema**, de bons géants |ma-dji-ndru#ma-ema|

**mabuzi maume**, boucs |ma-Ø'-buzi#ma-ume|

**madaho**, de grandes maisons

x cl. 11 : (voir cl. 11) :

**mani**, herbes [ma-ani]

**mawade**, maladies [ma-wade] (= **nyiwade**, et **nbwade**, 10)

La formation du pluriel de mots de cl. 11 en cl. 6 est souvent en variante avec une formation en cl. 10 (souvent à préf. nyi-

x pluralisateur :

**maya**, ensemble des doigts [ma-aya] (à côté de **shaya** / **zaya**, doigt/s, cl. 7/8)

x indénombrable :

**mare**, salive [ma-r-e] < -r- [notion de jeter]

**madji**, eau [ma-dji] < -dji [notion d'eau]

**madziwa**, lait [ma-dziwa] < -dziwa [notion de lait]

**maendeleo**, progrès [ma-end-ele-o] (cf. **wendeleo**, cl. 3)

Les noms de liquide (**mare**, **madji**, **madziwa**) ne sont attestés qu'en classe 6; ils sont hors de l'opposition de nombre (indénombrable).

Nous avons signalé que les augmentatifs gardaient le préfixe de cl. 5 : **mabuzi**, **madjindru**, **madaho**, **mapvaha**.

classe 7 : différents préfixes à répartition lexicale, sans doute dérivés d'une même forme < \*ki

préfixes :

sh-, shi- devant voyelles

hi-, i-, shi- devant consonnes (répartition lexicale)

contenu :

- instrument (souvent sur racine avec suffixe -o)
- partie du corps (plus petite que dans 3/4)
- diminutif (dont infirmité)
- noms d'animaux
- manière (shi-, ki-)
- divers

exemples:

1) hi- : quelques NI issus de thème seulement

**hitswa**, tête [hi-tswa] < -tswa [notion de tête]

**hindru**, chose [hi-ndru] < -ndru [notion d'existant]

ii) shi- : NI & ND; obligatoire avec thèmes à initiale vocalique et avec le démonstratif:

**shipvaa shihuu** [shi-pvaa#shi-huu] (ou **ihuu**, **hi-huu**), une grande prison < -pvaa [notion de prison]; -huu

shema shinu, cet albinos |shi-ema#shi-nu| < -ema |notion d'albinos|; -nu |démonstratif de proximité|  
 shaya shema, un bon doigt |shi-aya#shi-ema| < -aya |notion de doigt|  
 shiuno, hanche |shi-uno| < -uno |notion de hanche|  
 shingazidja, grand-comorien (langue) |shi-ngazidja| < -ngazidja |notion de Grande-Comore|  
 (hindru) shindrwadjina |(hi-ndru)#shi-ndrwadjina| (ou indrwadjina |i-ndrwadjina|), une autre (chose) < -ndrwadjina, |notion d'altérité|

iii) i- : NI & ND (sauf à initiale vocalique)

iloo, hameçon |i-lo-o| < -lo- |notion de pêcher|  
 ilevu, menton |i-levu| < -levu |notion de menton|  
 itranda ihuu |i-tranda#i-huu| (ou shihuu), lit < -tranda |notion de lit|;  
 irewe, faible, débile |i-r-ew-e| < -rew- |notion d'être apporté| (passif de -r- |notion de jeter, mettre| (cf. pp )

iv) ki-, shi- (variante libre): deux valeurs, selon le nominal introduit :

\* devant adjectif d'ethnie, cela signifie la langue:

(luga ya) shingazidja, la langue grand-comorienne |shi-ngazidja|  
 kiarabu, l'arabe |ki-arabu|

\* devant un nominal qui peut être déjà muni d'un préfixe de classe, pour former une locution adverbiale, 'd'une façon ... , à la manière de ...'; shi- et ki- sont alors en fonction d'extra-préfixe indépendant' (EPI, ALEXANDRE):

(mwendeso wa) ki-mdji, l'organisation de la vie du village |ki-m-dji|  
 (namna ya) ki-aswilia, (une manière) traditionnelle |ki-Ø-aswilia|  
 (nyile za) shi-zungu, (nyile za) shi-mzungu, une perruque ('cheveux à l'européenne') |shi-(m)-zungu| (les uns disent l'un, les autres, etc, - peut-être les anciens disent-ils davantage shimzungu, et les jeunes, shizungu

Bien que les deux formes soient dans ces deux emplois en variante libre, on peut peut-être déceler une tendance à la spécialisation, shi- pour la langue, ki- pour les autres emplois.

Ces différentes formes du préfixe viennent du même étymon, \*ki et cumulent les reflets possibles de \*k, comme nous le verrons:

\* passage à la continue (→ hi-) : reflet analogue à celui des consonnes initiales de préfixes BC (voir p )

\* palatalisation devant voyelle d'avant (→ shi-) : reflet 'normal' du préfixe

\* chute de h (→ i-): phénomène fréquent

\* maintien de l'occlusive (ki-) : reflet minoritaire de \*k

Il est intéressant de mettre en regard la manière dont elles reflètent l'étymon avec leur fréquence et leur emploi:

shi- : forme la plus fréquente; outre son rôle de préfixe nominal, il est le préfixe d'accord unique pour les démonstratifs et tous les accords non nominaux (pré-préfixes, affixes verbaux, etc);

hi- : forme rare; n'apparaît que dans un petit nombre de noms issus de thèmes, à l'exclusion des adjectifs et de tout élément régi; c'est un préfixe mort, qui n'est plus productif; on peut imaginer qu'il est à l'origine de i-.

i- : forme assez courante, productive; apparaît dans les nominaux, issus de racines ou de thèmes, et en accord adjectival (en variante avec shi-).

ki- : emploi spécialisé, en doublet avec shi-; la présence de cet allomorphe du préfixe de cl. 7 en shingazidja est sans doute due à une influence récente du swahili (où ki- est la forme principale) et ce d'autant que ki- n'entraîne pas de modification des mots eux-mêmes, puisqu'il leur reste extérieur, venant se préposer à un préfixe déjà en place.

Notons que le doublet ki-/shi- confirme la généralité d'emploi de shi-.

classe 8 : NI & ND : zi- < \*bi\*

préfixe:

z-, zi- devant voyelles

zi- devant consonnes

contenu:

pluriel de la classe 7 (lorsqu'il est possible)

exemples:

zitswa, têtes |zi-tswa|

zindru, choses |zi-ndru|

zipvaa zihuu, de grandes prisons |zi-pvaa#zi-huu|

zema zema, de gentils albinos |zi-ema#zi-ema|

aya, doigts |zi-aya| voir aussi **maya**, cl. 6)

ziuno zinu, ces hanches |zi-uno#zi-nu|

zindru zindrwadjina, des autres choses |zi-ndru#zi-ndrwadjina|

ziloo, hameçons |zi-lo-o|

zilevu, mentons |zi-levu|

zitrandu zihuu, grands lits |zi-trandu#zi-huu|

zirewe, débiles, infirmes |zi-r-ew-e|

zi- correspond aux trois préfixes de la classe 7 qui admettent un pluriel, hi-, hi- et i-, ce qui est une autre confirmation de leur origine commune.

classe 9 : NI & ND : N-, ny- < \*ny

Rappelons que \*ny est la forme reconstruite par GUTHRIE dans l'inventaire des étymons; toutefois, la majorité des langues bantu dont le shingazidja ont pour préfixe principal sinon unique en cl. 9 (et 10) une nasale homorganique de la consonne initiale lexicale

(notée ici N-), s'amalgamant avec celle-ci pour constituer une prénasalisée.

Nous nous en tenons strictement, ici comme ailleurs, aux représentations de GUTHRIE; pour ce qui est du shingazidja, le reflet attendu *nyi-* à partir de \*ny étant un allomorphe du préfixe de classe, cela intègre \*ny comme forme reconstruite plausible.

α préfixe :

*nyi-* devant voyelles et dans certains accords adjectivaux

*N-* : devant consonnes, nasale homorganique amalgamée à l'initiale lexicale (devant nasales /n, ny/, il y a fusion entre le préfixe et l'initiale lexicale)

β contenu sémantique :

difficile à définir: tout est susceptible de rentrer dans cette classe:

- nominalisations de racines (suff. *-a* et *-a*)
- noms d'animaux (cf. cl. 7)
- phénomènes naturels
- divers (cf. cl. 5)

γ exemples:

i) *nyi-* : NI & ND à initiale vocalique; ND où une forme /nC/ est impossible: quand l'initiale lexicale est déjà une prénasalisée ou quand la règle d'alternance n'a pas été appliquée :

*nyunba nyindrwadjina*, une autre maison |*nyi-unba#nyi-ndrwadjina*| < *-unba* |notion de maison|; *-ndrwadjina* |notion d'altérité|

*nyungu nyirahafu*, une marmite propre |*nyi-ungu#nyi-rahafu*| < *-ungu*, |notion de marmite|; *-rahafu*, |notion de laver|

En variante pour l'adjectif, une forme où la règle d'alternance s'applique, *ndrahafu*

(*ngozi*) *nyinkudu*, (peau) rouge |(N-*gozi*)#*nyi-nkudu*| < *-gozi* |notion d'écorce, de peau|; *-nkudu* |notion de rougeur|

(*nbuzi*) *nyonvu*, une chèvre maigre |(N-*buzi*)#*nyi-onvu*| < *-buzi* |notion de chèvre|; *-hond-* |notion de maigrir|

Les raisons générales de ne pas considérer /ny/ comme une prénasalisée \*iN-yi sont développées dans la présentation des sons (voir p. ...).

Sur le plan morphologique, cette analyse qui fait apparaître des thèmes à initiale vocalique est justifiée par la comparaison avec les termes apparentés en classe augmentative 5:

*nyunba* # *djunba*, palais |*dji-unba*| ; on isole un thème |*-unba*|, plutôt que \**-nyunba*;

*nyungu* # *djungu*, grosse marmite; de même, on isole un thème |*-ungu*|. Le BC offre dans chaque cas deux étymons osculants;

\**yũmbá* # \**nyũmbá*; \**yũngú* # \**nyũngú*.

Un cas où une telle analyse n'est pas justifiée est donnée ci-dessous (préf. N- devant nasales),

*nyitretrezi*, bavard |*nyi-tretr-e-i*| < |*-tretr-* |notion de bavarder|

*nyilemevu*, fatigué |*nyi-leme-u*| < |*-leme-* |notion d'être fatigué|

nyilemevu est en variante avec une forme plus rare où la règle d'alternance s'applique, ndemevu.

ii) N- :

x N- : → /nC/ (N-C) : le préfixe nasal s'amalgame à la consonne initiale du lexème pour former une prénasalisée; le cas échéant, la règle d'alternance s'applique

**nkuo**, grattoir |N-hu-ol| < -hu- |notion de gratter|

**nkoa**, toux |N-ho-a| < -ho- |notion de tousser|

**ndola nkuu**, grand mariage |N-lol-a#N-huu| < -lol- |notion de se marier|; -huu, |notion de grande taille|

**npbenye**, éclair |N-pveny-e| < -pveny- |notion de briller|

**ndrongoo**, affaire |N-rongo-ol| < -rongo- |notion de parler|

**nbuzi ndjeu**, une chèvre blanche |N-buzi#N-dji-eu| < -buzi |notion de caprin|; -eu, |notion de blancheur|

**nbwa ndjema**, un gentil chien |N-bwa#N-dji-ema| < -bwa |notion de chien|; -ema, |notion de bonté|

ndjema, ndjeu : l'accord est formé à partir de la forme de classe 5, dont le préfixe (dji-) est intégré au radical; la forme du thème est attestée par d'autres accords (mwema, wevu, cl. 1 & 3, etc; voir aussi ci-dessous).

**ngozi**, peau épaisse |N-gozi| < -gozi |notion d'écorce, de peau|

**ntanbi**, vermicelles |N-tanbi| < -tanbi |notion de vermicelles|

x N- : → Ø devant nasales /n, ny/: en cas d'initiale lexicale nasale, le préfixe fusionne avec l'initiale; quelques NI & ND:

**nyama**, viande |N-nyama| < -nyama, |notion de viande|

**nyunyi** ou **nuni**, oiseau |N-nyunyi|, |N-nuni| < -nyunyi, -nuni |notion d'oiseau|

Ici, aucun argument ne vient à l'appui d'une décomposition \*iny-ama|, ou \*iny-unii|, ce qui distingue ces cas de ceux tels que nyungu et nyunba; le BC propose toutefois similairement deux étymons, \*y-àná & \*ny-àná, \*y-ùni' et \*ny-ùni'.

(ngozi) **nene**, (peau) épaisse |N-gozi#N-nene| < -nene |notion de grosseur|

En variante rare; ndene, issu d'une reconstruction de thème à partir de l'accord de la cl. 5 (voir ci-dessous); le thème est isolé par comparaison avec d'autres accords (cl. 1 & 3 mnene, etc) .

La forme de base du préfixe de la classe 9 est |N-i|, entraînant une prénasalisée initiale par amalgame avec la consonne initiale du lexème; ces prénasalisées rentrent dans le cadre de l'alternance consonantique, dont les autres termes sont les continues et les occlusives.

Le préfixe nyi- est utilisé quand une prénasalisée est impossible:

- devant lexèmes à initiale lexicale prénasalisée (mais pas devant nasales)

- devant lexèmes à initiale vocalique

- devant lexèmes à initiale consonantique non alternante (occlusive sourde, continue)

L'accord des nominaux dépendants (adjectifs) est intéressant:

i) cas les plus fréquents (qui sont les cas réguliers):

x formation d'une prénasalisée : une prénasalisée se développe par amalgame du préfixe et de l'initiale lexicale consonantique - il peut s'agir d'une consonne rentrant dans le cadre de l'alternance ou d'une consonne qui reste extérieure à ce processus (fricatives, affriquées):

-huu (notion de grande taille) → **nkuu**, grand  
 -fupvi (notion de petitesse) → **nfupvi**, petit

x utilisation de l'allomorphe **nyi-** : avec lexème à initiale vocalique ou prénasalisée (lexicale):

-hond- (notion de maigrir) → **nyonvu**, maigre  
 -ndrwadjina (notion d'altérité) → **nyindrwapdjina**, autre

ii) cas particuliers: des adjectifs ne respectent pas les règles suivantes: deux tendances s'opposent:

x *prénasalisation* :

certaines adjectifs, très fréquents, à initiale lexicale vocalique ou nasale, ne sont pas accordés avec le préfixe **nyi-**, il y a reconstruction, à partir de la forme de cl. 5, d'un lexème à initiale consonantique occlusive, ce qui permet la formation d'une prénasalisée:

**ndjema**, bon (N-dji-ema) < **djema**, accord cl. 5 de -ema (notion de bonté)  
**ndjeu**, blanc (N-dji-eu) < **djeu**, accord cl. 5 de -eu (notion de blancheur)

et, de façon plus surprenante:

**ndene**, gros < **dene**, accord cl. 5 du thème -nene attesté par tous les autres accords de classe : ce qu'on peut représenter par (N-Ø'-nene)

En variante, un accord à préf. assimilé à l'initiale, **nene**

**ndonno**, en bonne santé < **donno**, accord cl. 5 du thème -nonno : **ndonno** peut être représenté par (N-Ø'-nonno)

En variante, un accord avec préf. syllabique, **nyinono**

Cela montre la force de la prénasalisation comme indice de cette classe.

x *maintien de l'initiale lexicale inchangée ('constance de l'initiale')* par

*l'emploi de **nyi-*** :

certaines adjectifs, dont l'initiale lexicale s'identifie à l'un des termes de l'alternance, gardent ou peuvent garder (nombreuses variations) leur initiale inchangée grâce à la forme syllabique **nyi-**:

**nyitretrezi**, bavard < -tretr- (notion de bavarder)  
**nyilemevu**, fatigué (= **demevu**) < -leme- (notion d'être fatigué)  
**nyirahafu**, propre (= **ndrahafu** < -rahar- (notion d'être propre)  
**nyinono**, sain (= **ndonno**, ci-dessus)

classe 10 : NI & ND : N-, nyi- < \*ny

préfixe :

identique à la classe précédente: ces classes se distinguent au niveau des accords régis autres que nominaux

contenu

- pluriel de la classe 9
- pluriel de la classe 11 (voir aussi 6)

exemples:

\* pl. de cl. 9 :

nkuo, grattoirs;  
ndola nkuu, grands mariages;  
etc:

\* pl. de cl. 11 :

ntsihu, journées;  
ndime ou nyilime, langues;  
ndongo ou nyilongo, argiles  
nyidjuzi, connaissances;  
nyizoevu, habitudes;  
nyilime ou ndime, langues;  
nyilongo ou ndongo, argiles;  
nyigangi, charmes  
nyigunba, pièges à poisson

Dans le cas des pluriel de mots de cl. 11, nyi- apparaît ici régulièrement devant continues.

On a noté la variation fréquente dans la formation du pluriel entre nyi- et N-:

L'utilisation de nyi- dénote souvent une influence ou une origine campagnarde, ou du moins est perçue comme telle. Au niveau syntaxique, on constate généralement une allitération entre les deux formes du préfixe, dans le cas

d'accords adjectivaux:

nyilime nyirahafu, des langues propres ou ndime ndrahafu, plutôt que  
nyilime ndrahafu ou ndime nyirahafu, admis toutefois.

On regroupe parfois les nominaux indépendants admettant le préfixe nyi- comme marque du pluriel (ce sont en général des NI de cl. 11 au sing.) dans une sous-classe ainsi définie, 10a; il nous a paru plus simple de n'en rien faire, en raison des nombreuses variations constatées et du fait que nyi- est aussi attesté comme préf. d'accord adjectival de cl. 9.

Nous considérons donc N- et nyi- comme des allomorphes du préfixe des classes 9 & 10.

classe 11 : NI seulement : u-, w- < \*du, \*bu

α préfixe :

w- devant voyelles

u- devant consonnes

β contenu sémantique:

• nom d'abstraction, de qualité - en particulier sur racine avec suff. -i et -u et avec thème adjectival

• terme d'unité (parmi un grand nombre)

β exemples:

ulezi, éducation lu-lel-i | < -lel- | notion d'éduquer |

udjuzi, connaissance lu-djuw-i | < -djuw- | notion de savoir |

uzoevu, habitude lu-zoev-u | < -zwey- | notion d'être habitué |

unyongofu, fait d'être allongé lu-nyongoh-u | < -nyongoh- | notion d'être droit |

weu, blancheur lu-eu | < -eu

ulime, langue (organe) lu-lime | < -lime | notion de langue |

usiku, journée lu-sihu | < -sihu | notion de jour |

ulongo, argile lu-longo | < -longo | notion d'argile |

ugangi, sorcellerie lu-gangi | < -gangi | notion de sorcellerie |

ugunba, piège à oiseau lu-gunba | < -gunba | notion de piège à oiseau |

wani, herbe lu-ani | < -ani | notion d'herbe |

Tous les accords gouvernés par les NI de cette classe, y compris ceux des ND, rejoignent ceux de la classe 3 (m-, mw-):

ulezi mwama, une bonne éducation;

usiku mdziro, une journée difficile

Il arrive que l'on rencontre pour les ND un accord analogique en u-: [usiku udziro]; ce fait est assez rare.

Le pluriel est formé en classe 10 (souvent avec préf. nyi-) et/ou 6; il arrive que le préfixe de classe 11 se maintienne:

uso, visage → nyiuso, nyiso, visages

On ne sait pas trop dans ce cas comment fixer le thème: |-so| ou |-uso| ? (le BC donne \*-ci'6)

Cette classe regroupe en shingazidja les classes 11 \*du et 14 \*bu du BC: il y a des confusions des reflets des préfixes, \*du et \*bu pouvant aboutir tous deux à Ng u-. Cette convergence explique les deux valeurs de cette classe, et la formation composite du pluriel: la classe \*11 du BC étant celle des unités et \*14, celle des noms abstraits, \*11 formant le pluriel en \*10 (ou \*10a, selon l'analyse faite), \*14 en \*6.

classe 15 : NI issus de racines seulement (infinitifs) : hu-, u-, hw-, w- < \*ku  
 a préfixe :

hw- ou w- devant voyelles

hu- ou u- devant consonnes

La présence d'une aspiration est liée à la structure de la racine

Ce préfixe, dont la forme de base est sans nul doute *ihu-*, subit de nombreuses transformations: sur le plan phonétique, /h/ est instable; en outre, devant voyelles, /u/ se dévocalise souvent; de plus, du fait de la nature du constituant (verbe), il y a souvent, à travers une resyllabisation, intégration de tout ou partie du préfixe dans le radical verbal, ce qui entraîne une ré-introduction du préfixe, etc:

-andzih-, l'otion d'écrire! : ( BC \*-yàndik-)

*tsiandziha*, j'ai écrit: *tsi-andzih-a*, [préf. verbal sujet - rac. - suff. l  
 Mais on entend aussi bien [tsiandziha], [tsiwandziha], [tsihandziha],  
 que [tsihwandziha]; or, le préfixe classe 15 n'apparaît pas dans cette  
 forme verbale, ce que montre la comparaison avec des verbes à initiale  
 consonantique: *tsifanya*, j'ai fait *tsi-fany-a*!

La racine est restructurée comme -wandzih-, -handzih-, hwandzih-, avec  
 intégration au lexème d'une partie du préfixe de classe 15, d'où les  
 formes mentionnées.

Cette restructuration apère aussi en classe 15 où se produit une ré-  
 introduction du préfixe: [uwandziha], [huwandziha], [hubandziha], pour  
*hwandziha*, écrire *ihu-andzih-a*!

De même, -ona, voir ( \*-bón- ) : *tsiono*, j'ai vu *tsi-on-o*, [tsiwono],  
 [tsihwono], [tsihono]; à l'infinitif, *hwona*, [hwona], [huwona], [hona],  
 [uhwona], etc

La consonne initiale ((w, h, hw)) qui peut apparaître en initiale  
 lexicale est donc à la fois

- i) un élément d'appui pour une recomposition syllabique;
- ii) un élément d'origine morphologique (préf. cl. 15 \*ku + Ng (h)u- +  
 /hw/, /w/, /h/
- iii) éventuellement le reflet de la consonne initiale \*b & \*d + Ng w,  
 \*y + Ng h (voir p )

β contenu :

cette classe est celle des infinitifs, qui sont à cheval sur les catégories  
 verbale et nominale: au niveau syntaxique, ils assument la fonction de  
 nominaux indépendants sans en posséder toutes les qualités, leur schème  
 d'accord étant incomplet (pas de préfixe de ND notamment); en revanche, ils  
 conservent certaines prérogatives verbales, comme la possibilité de  
 complémentation.

Il est vraisemblable que cette classe soit de même origine que la  
 classe locative 17 (ci-dessous); certaines formes accordées se  
 rejoignent (pré-préfixe cl. 15 & préf. cl. 17, ho):

hohwenda hahe, sa venue lho-hu-end-a#hahel, pré-préf, ou préf, cl, 17  
 - préf, cl, 15 - racine - suff.  
 Le pré-préfixe de cl, 15 est identique au préf, de cl, 17 qui peut  
 jouer le rôle d'un extra-préfixe-indépendant (EPI); la séquence [ho]  
 peut donc être interprétée comme l'un ou l'autre

β exemples:

i) racines 'mono-phonématiques' (selon notre catégorisation, voir p. 108):

x -C- (racines 'infra-syllabiques') : hu-, (u-):

**hula**, manger < -l- notion de manger!  
**hufa**, mourir lhu-l-a! < -l- notion de mourir!  
**hudja**, venir lhu-dj-a! < -dj- notion de venir!  
**hupva**, donner lhu-pv-a! < -pv- notion de donner!  
**upva**, brûler lu-pv-a! < -pv- notion de brûler!

La forme u- du préfixe d'infinitif semble ressortir ici de la  
 différenciation d'homonymes (hupva ≠ upva)

x -V-, -w- (racines 'mono-syllabiques') : u- :

**uua**, frapper lu-u-a! < -u- notion de frapper!  
**uwa**, tomber lu-w-a! < -w- notion de tomber!

ii) racines di-phonématiques à initiale vocalique (-V...-): hu-, hw-, h- ou  
u-, w-, selon l'initiale, la fréquence du mot, etc :

**hwona**, **uona** (uwona), voir lhu-on-a! < -on- notion de voir!  
**hwenda**, **henda**, aller lhu-end-a! < -end- notion d'aller!  
**hwanguha**, **wanguha**, **hanguha**, grandir lhu-anguh-a! < -anguh- notion de  
 grandir!  
**huambia**, **hunbia**, **hwunbia**, chanter une berceuse lhu-unbi-a! < -unbi- notion  
 de chanter une berceuse!

iii) racines diphonématiques à initiale /h/: u- (parfois hu-):

**uheu** (**huhea**), monter lu-he-a! < -he- notion de s'élever!  
**uhou** (**huhou**), tousser lu-ho-a! < -ho- notion de tousser!  
**uhundra** (**huhundra**), trouver lu-hundr-a! < -hundr- notion de trouver!

iv) autres racines (-CV...-, où C#h) : u-, hu- en variation libre :

**ufanya**, **hufanya**, faire lhu-fany-a! < -fany- notion de faire!  
**(h)usoma**, étudier lhu-som-a! < -som- notion d'étudier!  
**(h)ubua**, ouvrir lhu-bu-a! < -bu- notion d'ouvrir! (réversif de -bay-)  
**(h)ulima**, cultiver lhu-lim-a! < -lim- notion de cultiver!  
**(h)udzodza**, trouver lhu-dzodz-a! < -dzodz- notion de trouver!  
**(h)ukura**, être rassasié lhu-kur-a! < -kur- notion de rassasier!

Classes 16, 17, 18: extra-préfixe-indépendant (EPI) & ND : classes locatives

préfixe :

	EPI	ND	
cl. 16 :	pvo-	pva-	< *pa
cl. 17 :	ho-, o-	hu-, hw-	< *ku
cl. 18 :	mo-	?	< *mu

contenu :

Ces classes ne sont pas semblables aux autres classes: elles ne possèdent aucun nominal en propre, mais sont seulement des classes d'accord, où sont intégrés des nominaux appartenant normalement à d'autres classes; le préfixe des NI s'identifie ici à un extra-préfixe indépendant, ce qui explique l'absence de forme propre de préfixe de nominal indépendant dans ces classes.

Les nominaux qui y sont intégrés sont souvent munis d'un suffixe locatif; l'ensemble constitué de l'EPI et du nominal avec suffixe locatif constitue un syntagme locatif.

On retrouve ici un fonctionnement analogue à celui du préfixe de cl. 7 ki-. Le rôle de ces classes est essentiellement syntaxique, puisqu'elles permettent l'intégration des syntagmes locatifs dans la phrase, en régissant les accords.

Les trois classes locatives définissent des localisations spatiales et temporelles différentes (en liaison avec les suffixes locatifs); on dénote des traces d'une opposition de nombre entre 16 et 18.

cl. 16 : localisation spatiale avec singulier; localisation temporelle, pendant, dans

cl. 17 : localisation spatiale précise; lieu où l'on va

cl. 18 : localisation spatiale avec pluriel; intérieur de quelque chose (rare)

exemples :

\* cl. 16 : pvo-

pvo-ndjini, dans le village < -dji, notion de village!

pvo-djioni, dans l'après-midi < -djio, notion d'après-midi!

pva-dzima, ensemble < -dzima notion d'unité!

Cette forme s'explique comme un accord adjectival figé de cl. 16

\* cl. 17 ho-, o-

ho-dahoni, à la maison < -daho notion de maison!

ho-dahoni yiho, dans cette maison < -o, démonstratif de rappel

ho-harnwa enyunba, dans la maison < -hari notion de lieu intérieur!

La ressemblance, sans doute non fortuite, de cet élément avec le préfixe de la cl. 15 a été notée.

x cl. 18 : mo-

**mo-midjini**, dans les villages < -dji, notion de village!

**mo-midjini yimo**, dans ces villages

Nous n'avons pas relevé d'attestations pour cette classe de l'accord adjectival autre que démonstratif.

Outre des locutions faisant intervenir une autre classe locative (cl. 16, ci-dessous), la cl. 18 paraît attestée presque exclusivement avec des nominaux pluriels; cela n'est pas sans rappeler des parlers ruraux de Zanzibar où les classes locatives 16 & 18 sont normalement en relation de singulier à pluriel (nous sommes redevable à l'amabilité de Gérard PHILIPPSON pour cette information).

Un seul nominal appartient à l'une de ces classes en propre:

**pvahanu**, endroit, lieu : cl. 16

On peut imaginer que ce nominal résulte de la lexicalisation de l'accord de classe 16, suivi du démonstratif (-nu), -ha- pouvant s'analyser comme un formatif qui se retrouve dans divers éléments, dont les possessifs ...

A ces classes (surtout 16) peuvent s'accorder des termes invariables tels que **apvaha**, maintenant, **lelo**, demain, etc:

**apvaha pvanu**, au point où nous en sommes,

**lelo pvanu**, aujourd'hui même, etc.

Un accord cl. 9 est aussi possible; **apvaha yinu**, **lelo yinu**.

En outre, certains termes accordés à ces classes (démonstratifs, copules, etc) peuvent s'employer isolément avec des significations spécifiques:

x cl. 16 :

**pvanu**, alors, donc [pva-nu] ; **ngapvo**, il y a, [ng-apv-o] ; **epvonge**, alors, [e-pvo-ngel], etc

x cl. 17 :

**hunu**, ici [hu-nu] ; **hule**, loin, là-bas [hu-le] ; **ngiho**, il y a [ng-ih-o], etc

La locution (**ho-**)**harmwa**, dans, qui introduit des NI sans suffixe locatif, est intéressante: elle est à cheval sur deux classes locatives: elle se décompose en **ho** + **hari** + **mwa**, où l'on retrouve

+ l'EPI de la cl. 17 ho-, qui peut être omis;

+ le terme signifiant l'intérieur, **hari**

+ la marque d'accord de la cl. 18 qui introduit le NI, **mwa**

**ho-harmwa enyumba**, dans la maison [ho-hari-mwa#e-N-nyumba]

[hoharmwa], [oharmwa], [harmwa] : dans le cas où le préfixe cl. 17 n'est pas réalisé [ho], il se reporte sur l'aspiration du [h] initial de [hari] qui est plus forte.

**Recapitulation**

a) *Préfixes de classe*

Le tableau reprend les principaux préfixes des nominaux indépendants et des nominaux dépendants, en distinguant selon que l'initiale du lexème est consonne ou voyelle.

Les accords des nominaux dépendants, notamment en classes 5 & 9/10 qui présentent de nombreuses irrégularités, sont repris indépendamment.

Pour les classes locatives, figurent les extra-préfixes indépendants

	Nominaux indépendants		Nominaux dépendants	
	#C-	#V-	#C-	#V-
1	m-	mw-	m-	mw-, m-
2	wa-	(w-)	wa-	w-
3	m-	mw-	m-	mw-, m-
4	mi-	m-	mi-	m-
5	(dzi-) dji- Ø'-	dj-	(dzi-) dji- Ø'-	dj-
6	ma-	m-	ma-	m-
7	(hi-) i-	i-	i-	
8	shi- zi-	sh- z-	shi- zi-	sh- z-
9	N- nyi-	ny-	N- nyi-	ny-
10	N- nyi-	ny-	N- nyi-	ny-
11	u-	w-	m-	mw-, m-
15	hu-, u-	hw-, w-, h-		
16	pvo-		pva-	
17	ho-		hu-	
18	mo-			

*β Principaux regroupements de nombre*

Le tableau suivant représente les oppositions de nombre entre classes, pour les termes de fonds bantu.

singulier		pluriel
1	—————	2
3	—————	4
5	—————	6
7	—————	8
9	—————	10
11	—————	

classes isolées:

6 - liquides

11 - abstraction

15 - infinitif

16, 17, 18 - locatifs

*γ) alternance consonantique et nominaux dépendants*

Nous avons signalé, pour différentes classes, un phénomène d'alternance consonantique, qui fait varier l'initiale lexicale appartenant aux continues (pv, w ou Ø, r, l, h, sh) avec des occlusives (p, b, tr, d, k, tsh) après certains préfixes.

Ce mécanisme n'est pas systématique.

La tendance est la suivante:

- i) après préfixe syllabique (position intervocalique) : continues;
- ii) après préfixe |Ø'-| de cl. 5 (position initiale absolue) : occlusives;
- iii) après préfixe nasal |N-| : amalgame entre le préfixe et l'initiale lexicale, aboutissant à une prénasalisée (occlusive sonore)

Ainsi, un même lexème aura, pour initiale, tantôt une continue, tantôt une occlusive, tantôt une prénasalisée, selon la nature du préfixe de classe; cela joue à travers l'ensemble du système de classe, tous les préfixes étant

concernés, mais ce conditionnement est plus ou moins contraignant; il apparaît plus nettement dans les formes des adjectifs, par définition accordés dans toutes les classes, ainsi qu'entre les classes 5 à préf.  $\emptyset$ -l et 9 à préf.  $\text{N-l}$ , et les classe 6 à préf.  $\text{lma-l}$  (pluriel de 5), cl. 11 à préf.  $\text{lu-l}$  (sing. de 10), et cl. 15 à préf.  $\text{luh-l}$  (infinitif verbal).

Nous reviendrons dans la partie suivante sur la cause possible de ce qui apparaît comme une tendance à une alternance consonantique; nous donnons des exemples d'adjectifs puis de nominaux, après avoir rappelé les consonnes en alternance.

pv / p / npb  
 w ou  $\emptyset$  / b / nb  
 r / tr / ndr  
 l, y / d / nd  
 h / k / nk  
 sh / tsh / ntsh (pour les adjectifs seulement)

1) adjectifs

Le tableau suivant donne les formes accordées d'un certain nombre d'adjectifs, illustrant chacune des séries alternantes. Certains d'entre eux admettent aussi en classes 5 et 9/10 les allomorphes syllabiques des préfixes de classes, l'initiale lexicale restant alors continue (voir remarques ci-dessous).

les classes 1 & 3, ainsi que 9 & 10, aux accords identiques, sont regroupées; ne figurent ni la classe 11, dont les NI commandent des accords de cl.3, ni les classes 15 et 16, 17, 18, qui n'ont pas ou exceptionnellement d'accords adjectivaux).

Liste des adjectifs (la forme de citation est à initiale continue)

- itsi, crû
- ii, mauvais
- wade, malade
- pvungufu, minable
- radji, large
- ruvu, vide
- renbavu, mou
- le, long
- huu, grand
- havu, sec
- hara, peureux
- shashi, quelque
- she, femelle

	cl.1 & 3	cl.2	cl.4	cl.5	cl.6	cl.7 <sup>1</sup>	cl.8	cl.9
g, w / b	mwitsi mwii mwade	waitsi wai wawade	miitsi mii miwade	bitsi bi bwade	maitsi mai mawade	shiitsi shii shiwade	ziitsi zii ziwade	nbitsi nbi nbwade
pv / p	mpvungufu	wapvungufu	nipvungufu	pungufu <sup>2</sup>	mapvungufu	shipvungufu	zipvungufu	nbungufu
tr / r	mradji mruvu mrenbavu	waradji waruvu warenbavu	miradji miruvu mirenbavu	tradji truvu trenbavu	maradji maruvu marenbavu	shiradji shiruvu shirenbavu	ziradji ziruvu zirenbavu	ndradji ndruvu ndrenbavu
l / d	mle	wale	mile	de	male	shile	zile	nde
h / k	nhuu nhavu nhara	wahuu wahavu wahara	mihuu mihavu mihara	kuu kavu kara	nahuu nahavu nahara	shihuu shihavu shihara	zihuu zihavu zihara	nkuu nkavu nkara
sh / tsh	nshashi nshe	washashi washe	nishashi (nishe)	tshashi tshe	nashashi nashe	shishashi shishe	zishashi zishe	ntshashi ntshe

1 cl.7: un accord avec i- est aussi possible, mais il ne modifierait en rien l'initiale de l'adjectif

2 en variante, bungufu

(1) Nominiaux indépendants

Nous donnons des exemples en nous limitant à des cas évidents:

pv / p / npb :

- paha, chat + **mapvaha**, chats (cl. 5 / 6)
- upvea, balayer + **peleo**, balai (cl. 5)
- upvenya, cligner des yeux + **npbenye**, éclair (cl. 9)

w, Ø / b / nb :

- bwe, pierre + **mawe**, pierres (cl. 5 / 6)
- hwundjilia, rassembler + **bundjilio**, rassemblement (cl. 5)
- hwunda, construire + **bundo**, construction (cl. 5)

On remarque l'occlusive bilabiale en classe 5 qui alterne avec Ø dans les verbes, qui permet de postuler une racine *-wund-*, *-wundjili-* (avec continue w) (voir reflets de \*b, p )

- uwana, lutter + **nbwana**, combat (cl. 9)

r / tr / ndr :

- travu, joue + **maravu**, joues (cl. 5 / 6)
- urongoa, parler + **ndrongoo**, parole (cl. 9)

l, y / d / nd :

- daho, maison + **malaho**, maisons (cl. 5 / 6)
- ulima, cultiver + **ndima**, agriculture (cl. 9)

h / k / nk :

- koko, grand-mère + **mahoko**, grand-mères (cl. 5 / 6)
- uhca, tousser + **nkoo**, toux (cl. 9)

Remarques sur les adjectifs

1) Sans aller contre cette distribution, certains adjectifs admettent aussi les préfixes syllabiques *dji-* et *nyi-* des classes 5 et 9/10 (avec parfois une implication sémantique, cf. p 137) (nous ne reprenons ici que les formes différentes quand les adjectifs figurent dans le tableau précédent):

	cl.5	cl.9/10
minable	<b>djipvungufu</b>	<b>nyipvungufu</b>
	( <b>djibungufu</b> )	( <b>nyibungufu</b> )
malade:	<b>djiwade</b>	<b>nyiwade</b>
vide	<b>djiruvu</b>	<b>nyiruvu</b>
peureux		<b>nyihara</b>
ainsi que :		
fatigué	<b>djilemevu,</b>	<b>nyilemevu</b>
	<b>demevu</b>	<b>ndemevu</b>

etc

2) On trouve aussi des adjectifs (généralement issus de racine, avec suffixe *-vu*) dont l'initiale appartient à l'un des inventaires en question qui ne

manifestent pas l'alternance et acceptent toujours les préfixes syllabiques des classes 5 et 9/10 (nous ne donnons que les accords typiques: cl.1, 2, 5, 6, 9)

	cl.1	cl.2	cl.5	cl.6	cl.9
exigeant:	mpviavu	wapviavu	djipviavu	mapviavu	nyipviavu, npviavu
vert	mrutrutru	warutrutru	djirutrutru	marutrutru	nyirutrutru, ndrutrutru
calme	mrulivu	watrulivu	djitrulivu	matrulivu	nyitrulivu
malhonnête	mbabaifu	wababaifu	djibabaifu	mababaifu	nyibabaifu

Malgré des idiosyncrasies nombreuses, il semble que le comportement des adjectifs soit lié, au moins en partie, à la structure morphologique des termes (les adjectifs 'primaires' sont issus de thèmes adjectivaux, les adjectifs 'non primaires', de racine, par l'adjonction de suffixe à valeur adjectivale)

- adjectifs à initiale continue ou occlusive:

× initiale alternante : ND primaire

× initiale alternante ou constante (fluctuation) : ND non primaire

- adjectifs à initiale extérieure à ce cadre (ni continue, ni occlusive alternante) :

× initiale affriquée ou fricative : préfixes 5 & 9 non syllabiques (Ø-, H-)

× initiale prénasalisée : préfixes 5 & 9 syllabiques (dji-, nyi-)

× initiale voyelle ou nasale : cas particulier

Comme les adjectifs, d'une façon générale, ont un comportement difficilement prévisible, les accords caractéristiques de chacun (c'est-à-dire ceux des classes 1, 2, 5, 6, & 9-10) sont donnés dans le lexique.

Éléments facultatifs : extensions

GUTHRIE donne comme 'extensions' des éléments de structure -VC- ou -C- de caractère facultatif, qui peuvent être insérés entre le radical et les suffixes; ce sont en shingazidja des morphèmes de dérivation, qui se post-posent immédiatement à la racine et en modifient le sens ou la construction, sans en changer le statut: une racine "étendue" conserve son ambivalence verbale et nominale, ainsi que son statut d'élément de langue: elle doit recevoir des affixes actualisateurs (préfixe, suffixe) pour devenir élément de discours (constituant syntaxique).

Toutes les extensions possèdent un étymon dans la liste de GUTHRIE.

Les extensions d'une part sont appréhendées à travers la mise en évidence de racines étendues en face de racines primaires, ce qui suppose la constitution de familles lexicologiques, à partir de la comparaison entre différents mots, phoniquement et sémantiquement proches, actualisant les racines, et d'autre part elles sont identifiées à travers leur forme et leur valeur.

Pour ce faire, deux critères interviennent:

**un critère formel:**

chaque extension doit être repérée dans ses différents avatars : dans leur conditionnement et les modifications morpho-phonologiques qu'elle peut entraîner au contact avec les racines;

**un critère sémantique:**

1) *concernant la racine:*

en principe, tous les constituants syntaxiques issus d'une même racine primaire sont reliés sémantiquement; il peut arriver toutefois que la relation entre tel dérivé et la racine ne soit pas évidente;

ii) *concernant l'extension:*

la valeur ou les valeurs de chaque extension sont (relativement) constantes et se retrouvent dans tous les cas où l'extension est actualisée. Toutefois, ce critère ne saurait être appliqué de façon trop rigoureuse.

Contrairement à GUTHRIE qui estimait que, si une même forme possédait deux valeurs distinctes selon les critères qu'il mettait en avant (essentiellement des fonctions syntaxiques, voir plus bas), il fallait en conclure à l'existence de deux extensions homophones, nous proposons une définition avant tout formelle des extensions (GUTHRIE 1962, cité ci-dessous): il nous semble plus efficace, tant sur le plan stratégique (économie de la description) que théorique, d'admettre, avec Th. C. SCHADEBERG, qu'une seule et unique extension, définie formellement, peut avoir plus d'une valeur, et donc de reconnaître une unité morphologique en dépit d'une diversité "lexicale" éventuelle; cela permet d'éviter une catégorisation trop arbitraire: à propos de l'extension -am- en swahili, Th. C. SCHADEBERG écrit "(...) j'accepte les particularités de tous les procédés dérivatifs: je considère -and-am- et -lal-am- comme étant des éléments lexicaux propres, mais je maintiens, d'autre part, qu'ils comportent tous deux le même morphème -am- (...). En effet, l'établissement de la valeur significative d'une extension verbale dépend en grande partie de l'intuition du chercheur. (Th. C. SCHADEBERG 1982 : 56).

De ce fait, nous ne proposons pas d'intitulé général pour chaque extension, que nous définissons avant tout formellement; nous rappellerons toutefois pour chacune la terminologie traditionnelle (où les intitulés ont été choisis à partir de la valeur la plus fréquente de l'extension, ou celle qui a été mise en lumière en premier dans le cours du développement des études bantu ...)

Dans les cas les plus favorables, l'extension est mise en évidence à travers l'existence d'une racine primaire et de racines étendues sémantiquement reliées, qui permettent d'isoler la valeur proto-typique de l'extension; dans d'autres cas, il peut arriver que la racine primaire qui correspondrait ne soit pas attestée ou que la relation sémantique entre la racine primaire et la racine supposée étendue ne soit incontestable, ou encore que la valeur théorique de l'extension ne se retrouve pas dans la forme supposée étendue... : la décision de rapprochement doit alors s'appuyer sur diverses considérations, formelles et

antiques, toujours internes à la langue. Rappelons que nous ne faisons pas intervenir le comparatisme dans la constitution des familles lexicologiques sur la base de la distinction opérée entre racines synchroniques et racines diachroniques, *supra*, pp 101).

Dans les cas où l'identification, à partir de nos relevés, de racines étendues nous a paru demander une extrapolation trop large rentrent d'office dans les racines dites 'complexes', pour lesquelles nous ne faisons pas d'analyse (voir fin de paragraphe).

Cela nous paraît le seul moyen de limiter l'arbitraire.

Nous présentons les différentes classes ou catégories d'extensions, définies selon un critère sémantico-fonctionnel.

Nous verrons que les extensions peuvent se combiner de façon claire, la racine résultante cumulant les valeurs des différentes extensions recensées; dans certains cas, à partir de l'allomorphe choisi de telle ou telle extension, nous sommes amenés à supposer une formation composite (la base étant une racine déjà connue mais sans que cette première extension transparaisse sémantiquement ni soit isolable), la racine primaire pouvant être ou ne pas être attestée.

Comme il a déjà été signalé, il arrive exceptionnellement que des extensions soient post-posées à des thèmes adjectivaux, le résultat en étant l'obtention de verbaux: cet aspect est traité *in fine*.

### Typologie des extensions

Sur la base de leur valeur et de leur comportement syntaxique, on distingue plusieurs classes d'extensions, qui sont compatibles, dans l'ordre indiqué par le schéma de formation (p 105):

		valeur	productivité	nécessité
ext.	1	sémantique	-	-
ext.	2	sém-synt	+	-
ext.	3	syntaxique (diathèse)	+	-

Ces trois critères ne permettent pas en réalité une catégorisation nette:

- valeur : la valeur définit trois classes, sur la base des modifications apportées par la présence de l'extension, qui peuvent être d'ordre sémantique pur ou syntaxique pur, ou relever des deux à la fois; nous référons aux catégories d'extensions par leur valeur;
- productivité : le critère de productivité permet d'isoler les extensions purement sémantiques dont la production est bloquée, qui n'apparaissent que dans un nombre limité d'exemples;
- nécessité : le critère de nécessité montre qu'aucune extension n'est obligatoire - ce qui est conforme à l'interprétation des extensions comme éléments facultativement adjoints aux racines.

Notre définition de la valeur des extensions rejoint en partie la classification syntaxique de GUTHRIE, qui distingue i) des extensions impliquant un participant supplémentaire ("*extensions involving an extra object*"), ii) des extensions impliquant un participant de moins ("*extensions involving one object less*"), et iii) des extensions impliquant le même nombre de participants ("*extensions involving the same number of objects*"). Les premières correspondent à notre catégorie d'extensions sémantico-syntaxiques, les secondes, plus ou moins aux extensions dites de diathèse, les troisièmes, aux extensions sémantiques pures (GUTHRIE 1962 : 95-97).

Nous dégageons, pour chaque extension, regroupée dans les catégories énumérées, sa forme et son conditionnement (allomorphes, modifications morpho-phonologiques au contact de la racine), mis en regard de son étymon chez GUTHRIE, sa ou ses valeurs, son domaine; nous envisageons ensuite brièvement la combinatoire des extensions entre elles.

Comme les extensions sont souvent définies syntaxiquement, on ne s'étonnera pas que les exemples fassent le plus souvent intervenir des constituants syntaxiques verbaux où la valeur de l'extension est mieux mise en valeur.

1) *Les extensions sémantiques*

Elles ne sont pas ou plus productives; elles ne sont sans doute pas ressenties, intuitivement, par les locuteurs, comme des éléments isolables, et nous n'avons relevé aucun cas où elles soient combinables entre elles.

Elles n'affectent que le fonds bantu.

C'est nous en tenons ici aux cas où forme primaire et forme étendue peuvent être mises en relation; du fait de leur très faible rendement, certaines extensions isolées ici doivent leur prise en compte en fait à la perspective comparatiste de cette étude et/ou à l'existence de racines allongées ou complexes les évoquant (voir plus loin, pp 130).

Dans la mesure où elles ne sont pas productives, leur évaluation chiffrée est significative - elle est donnée ici - même si l'on ne peut exclure, bien entendu, des cas non relevés.

Elles modifient le sémantisme de la racine et correspondent à ce que l'on entend généralement par dérivation; elles ne modifient pas le rapport de la racine à l'énoncé.

1) *extension -u-, -o- < \*-ud-*

Les réalisations dépendent de la voyelle radicale:

/u/ après voyelles extrêmes /i, u, a/

/o/ après voyelles intermédiaires /i, e/

Nombre total d'attestations: 11

leurs

Cette extension possède deux valeurs reconnues :

- réversif (contraire de la racine simple) : 7 (dont 5 sûres)
- intensification de la racine simple : 4

SACLEUX avait appelé cette forme "extractif", pour rendre compte de ces deux valeurs; POLOME, pour le swahili, note que "the essential function of the [o(1)] suffix is to convey the meaning opposite to that commonly attached to the root; it is therefore usually called reversive or conversive (...). In some cases, however, the semantic content of the basic verb is intensified, instead of being reversed" (POLOME 1967 : 90)

A côté de ces deux valeurs reconnues, il existe des exemples d'interprétation sémantiquement plus douteuse, où la définition du rapport sémantique n'est pas évident - quoique sa réalité ne nous paraisse guère pouvoir être mise en doute.

Exemples :

Nous donnons la forme simple en regard de la forme étendue, ainsi que les étymons - il arrive qu'il y ait un étymon pour chacun des deux niveaux de racine

x valeur réversive: 4 + 1

*-tú'ng-	>	-funga, attacher	⇔	-fungua, détacher	<	*-tú'ngud-
*-dùng-	>	-unga, lier	⇔	-ungua, couper		
		-baya, fermer	⇔	-bua, ouvrir		
		-wanba, attrapper	⇔	-wanbua, détacher		

A ces cas bien établis, on peut ajouter, sur la base de l'analyse synchronique, un cas toutefois plus discutable:

*-pá-	>	-pva, donner	⇔	-pua, ôter	<	*-yi'pud-
-------	---	--------------	---	------------	---	-----------

On remarque l'alternance initiale et le non rapport des séries comparatives BC entre elles; rappelons que les rapprochements sont justifiés à l'intérieur du shingazidja uniquement - en tout état de cause, le BC montre qu'il s'agit dans doute d'une convergence accidentelle en shingazidja entre deux racines originellement différentes

x valeur intensive: 4

*-càk-	>	-tsaha, chercher	⇔	-tsahua, choisir	<	*-càkud-
*-kám-	>	-hama, traire	⇔	-hamua, têter		
*-càmb-	>	-tsanba, se laver le cul	⇔	-tsanbua, rincer		

x valeur évoquant le réversif: 2

*-dàg-	>	-laha, dire adieu	⇔	-lahua, insulter		
		-siha, tenir	⇔	-sihua, enlever la bourre de coco		

Hormis le cas de -pva ⇔ -pua, où il n'y a pas lieu - et pour cause - de supposer une parenté dans la langue commune - le fait que des étymons soient attestés simultanément pour la racine primaire et pour la forme étendue montre que cette extension est ancienne; de fait, il est possible qu'elle n'ait jamais été opératoire en shingazidja, mais que les formes étendues proviennent directement d'un état de langue antérieur - cela expliquerait bien entendu qu'elle ne se soit jamais appliquée aux emprunts.

2) extension *-an-* < \**an*

Extension très peu attestée: 2 cas seulement

Valeur:

Dans la mesure où la valeur est identifiable à partir d'une telle faiblesse d'attestations, indique un état statif

Exemples:

-dòb- > -loa, être mouillé    ☉ -lowama, être trempé  
 La présence d'une consonne /w/ dans la forme étendue est liée à des raisons accentuelles

-dù'b- > -vuba, pêcher    ☉ -vubama, s'évanouir

3) extension *-any-* < \**an + #i*

Cette extension apparaît composite, consistant dans une extension *-an-* suivie d'une extension *-i*, ces deux morphes étant chacun identique à une extension sémantico-syntaxique; nous la considérons ici comme une unité, car, malgré l'analogie formelle avec des extensions identifiées, les cas ci-dessous ne permettent pas de décomposition.

Nombre d'attestations : 6

Valeur :

donne une idée de réunion, de mise en commun, d'où son appellation d'associatif

GUTHRIE intitule ce qui correspond à la première partie de cette extension "corporate"

Exemples:

-biha, frapper    ☉ -bihanya, mélanger  
 -bwabwa, être mou    ☉ -bwabwanyiha, être plein d'eau, mou  
 Ici avec extension *-ih-* ; pas de forme \*-bwabwanya relevée

-huntsa, aider    ☉ -huntsanya, réunir

-pát- > -para, obtenir    ☉ -paranya, mélanger

-càng- > -tsanga, cotiser    ☉ -tsanganya, réunir < \**-càngani-*

-dùng- > -unga, joindre    ☉ -unganya, rattacher

4) extension *-ar-*, *-ats-* < \**at-*

Extension peu attestée: 2 cas seulement (comme le statif)

Elle se présente sous deux formes, correspondant à deux reflètes possibles de \**t* (en intervocalique, comme c'est le cas ici, nous verrons que le reflet régulier est une continue, *r*).

Valeur:

Dans la mesure où la valeur est identifiable, cette extension indique un contact, comme son intitulé traditionnel de contactif le laisse entendre.

Exemples:

\*-dí'ng- > -zinga, entourer @ -zingara, attrapper  
 \*-tú'ng- > -funga, attacher @ -fungatsa, remorquer

On voit que les extensions purement sémantiques sont en fait très peu attestées (11 cas pour -u-, le plus répandu, sur 21 au total de cette catégorie).

Outre un souci d'exhaustivité, elles ont cependant leur place dans notre présentation du fait, comme il a été dit, de la perspective analytique et historique dans laquelle nous nous sommes situés; une étude purement synchronique et ne faisant pas place au comparatisme ne les reconnaîtrait sans doute pas (ou seulement la première) pour des éléments indépendants.

#### *b) extensions sémantico-syntaxiques*

Catégorie d'extensions facultatives et productives, au contraire de la catégorie précédente, elles peuvent s'adjoindre à toute racine, primaire ou déjà étendue (avec extension sémantique), pourvu que le sens le permette; elles affectent sans restriction particulière les mots d'origine étrangère, arabe notamment, mais avec des spécificités formelles; nous traiterons à part de leur application aux emprunts mais donnerons d'ors et déjà quelques exemples pour montrer l'unité de fonctionnement de la langue.

Ces extensions interviennent sur la valence des racines, l'augmentant d'un ou plusieurs participants: elles affectent les relations du constituant syntaxique issu de la racine avec les autres participants de l'énoncé; dans le cas de constituants verbaux (verbes), elles entraînent souvent, ou permettent, la

présence d'un participant supplémentaire au procès, en modifiant leur complémentation.

On retrouve la définition de GUTHRIE "*extensions involving an extra object*".

au niveau des gloses françaises, cela peut impliquer l'emploi d'un autre terme.

Nous donnons ici, le cas échéant, des exemples en énoncé.

-1) extension -i-, -e-, -li-, -le- < \*id

Les réalisations dépendent

\* de la voyelle radicale pour ce qui est du timbre de la voyelle de l'extension:

/i/ après voyelles extrêmes /i, u, a/ :

-lima, cultiver → -limia, cultiver pour

-hula, acheter → -hulia, acheter pour

-anba, dire → -anbia, dire à

/e/ après voyelles intermédiaires /o, u/

-rema, frapper → -remae, frapper pour

-ona, voir → -onea, mépriser

\* de la finale radicale pour la forme de l'extension:

finale continue /h/ : palatalisation → /sh/ :

-piha, cuisiner → -pishia, cuisiner pour

La palatalisation est conforme au reflet de \*k devant voyelle d'avant (cf. étude des sons p ).

finale continue /w, y/ : chute de la consonne radicale → extension

-lV- :

-pua, enlever → -pulia, enlever à

-zaya, mettre au monde → -zalia, mettre au monde (en précisant le lieu)

La forme -lV- est donnée en synchronie; il y a en réalité intégration de la consonne finale du lexème pour constituer un allomorphe de l'extension; en effet, la consonne finale étymologique de ces radicaux est \*d, dont les reflets réguliers en intervocalique sont (l, y, Ø); l'extension applicative est alors construite sur cet état de la racine;

-pua, enlever < \*y-i'pud-

+ -pulia : \*-pui-i- + -pu-li-

-fungua, détacher (réversif) < \*tù'ngud

+ -fungulia : \*-fungul-i- + -fungu-li-

-zaya, enfanter < \*bi'ád-

+ -zalia : \*-zal-i- + -za-li-

Comme ce processus opère pareillement, en l'absence de toute consonne finale étymologique, avec des racines 'bantu' qui ne possèdent pas d'étymon chez GUTHRIE, ainsi qu'avec des termes d'emprunt, nous considérons qu'il y a eu déport de la consonne étymologiquement radicale dans l'extension;

emprunt:

-busu, embrasser (ar.)

+ -busulia : -busu-li-

d'où \*-pul-i- + -pu-li- ; \*-fungul-i- + -fungu-li- ; \*-zal-i- + -za-li-  
Ce même phénomène se produit aussi avec d'autres extensions à initiale  
vocalique.

Par ailleurs, si la consonne finale étymologique de l'extension ( < #id )  
n'est plus perceptible au niveau phonique, certaines formes composées  
( 'causatif', voir plus bas ) nous amèneront à la ré-introduire au niveau  
de l'analyse structurelle.

Valeurs :

Cette extension sert à introduire un autre participant au procès; on  
distingue essentiellement deux valeurs, dans l'une, il s'agit du  
destinataire de l'action (bénéficiaire, d'où le nom de bénéfactif donné  
parfois à cette extension - à comprendre dans une acception large,  
incluant celui au nom duquel ou pour lequel on fait l'action, par  
substitution), dans l'autre, le lieu (localisation) du procès.

La présence de l'extension et du complément qu'elle permet s'impliquent  
mutuellement.

Exemples :

\* valeur attributive ou bénéfactive

rihulu ndrovi, nous avons acheté des bananes

rihulja owatrotro ndrovi, nous avons acheté des bananes aux enfants  
(destinataires)

On remarque le changement de timbre du suffixe, harmonie vocalique dans  
le premier cas (verbe issu de racine -CVC- monosyllabique), -a dans le  
second (verbe étendu, issu d'une racine -CVCV-) (voir p 110)

haanba uka ..., il/elle a dit que ...

hanianbja uka ..., il/elle m'a dit que ...

\* valeur locative

hazaya, elle a accouché

hazalja ho-dahoni, elle a accouché chez elle

hafu, il/elle est mort/e

hafja ho-shanbani, il/elle est mort/e au champ

Mots d'emprunt : même mécanisme; la voyelle est presque exclusivement /i/  
(puisque conditionnée par les voyelles radicales, ici en principe extrêmes, voir  
pp )

-djiri, se passer :

adjali idjiri, un accident s'est produit

-djirja : adjali inidjiria, un accident m'est arrivé

## b-2) extension -ili-, -ele- < #idid

Les réalisations de cette extension, qui se présente comme une reduplication de  
la précédente, suivent le même conditionnement que ce dernier.

Valeur :

On distingue une valeur proprement intensive: répétition, prolongement ou intensification de l'action, et une valeur instrumentale (ce qui sert à ...)

Exemples :

x valeur intensive

-enda, aller → -endelea, continuer

-hima, être debout → -himilia, faire face

x valeur instrumentale

-piha, cuisiner → -pishilia, servir à cuisiner (d'une marmite, etc)

Mots d'emprunt : très rare; un seul cas relevé

-tabani, moudre → -tahanilia, servir à moudre (d'une meule)

3) extension -(i)z-, -(e)z-, -(i)s-, -(e)s- < \*i'

-(i)sh-, -(e)sh- < \*iki'

Compte-tenu de la valeur la plus fréquente de ces deux séries d'allomorphes, ils sont regroupés dans la littérature dans une extension unique appelée 'causatif'.

Le conditionnement des réalisations fait intervenir plusieurs facteurs; il explique s'il l'on admet une perspective comparative et diachronique, et que certaines formes où apparaît le causatif sont composites.

On peut estimer, comme la présentation le suggère, qu'il y a en fait deux causatifs (ou deux constructions du causatif), l'un, issu directement de BC \*i', l'autre, issu de BC \*iki', ce dernier s'appliquant presque exclusivement aux emprunts.

La première forme (\*i') a induit des modifications morpho-phonologiques affectant la consonne finale des lexèmes (des alternances), comme il arrive souvent au contact d'une voyelle fermée étymologique (voir partie suivante); ici, cela a consisté en une 'fricativisation' de cette finale, qui devient, le plus souvent, sifflante; de ce processus, on peut alléguer, que dans une étape ultérieure du développement de la langue, un nouvel allomorphe de cette extension, comprenant les consonnes fricatives résultantes sourde ou sonore /s/ & /z/ a été extrait par comparaison des modifications produites, et, sans doute aussi, analogie et convergence avec le reflet du second causatif (\*iki').

Quant à la seconde forme (\*iki'), il faut remarquer que l'allomorphe shingazidja lui correspond tout à fait du point de vue de la réflexion des

symboles BC en shingazidja, puisqu'il montre une palatalisation comme reflet d'une séquence \*k + voyelle d'avant, ce qui constant (voir pp. ) Reste à déterminer pour quelle raison les emprunts se sont vus appliqués un allomorphe spécifique. Un comportement morphologique particulier aux emprunts en tous cas ne saurait surprendre : nous en verrons des exemples dans le chapitre suivant.

Cette division en deux allomorphes d'origine étymologique BC distincte s'expliquerait mieux si les deux reconstructions de causatifs données par GUTHRIE présentaient une différence sémantique qui justifiait leur répartition actuelle.

Une autre possibilité serait que le causatif des termes d'emprunt soit composite; il s'agirait alors d'une association d'une extension de forme -ih- (reconstruite comme \*ik, à quoi GUTHRIE prête différentes valeurs) et du causatif \*i' : la séquence \*ik + \*i' aboutit aussi normalement à Ng -ish- que \*iki'. Dans la mesure toutefois où nous n'avons pas d'indication claire qu'une telle composition se soit produite, et où la mention chez GUTHRIE de la série \*iki' ne rend pas nécessaire cette hypothèse, nous nous en tenons à cette dernière (voir cependant ci-dessous).

G. PHILIPSON (communication personnelle) suggère, pour expliquer la présence et le domaine de cette forme -ish-, un emprunt du shingazidja au swahili; en swahili, le causatif est toujours formé avec -ish- dans les emprunts, les termes bantu connaissant soit une alternance du même ordre que celle du shingazidja ( <\*i' > soit -iz- (même problématique comparative d'ailleurs, le causatif swahili pouvant originer dans une forme unique \*iki' ou dans la composition \*ik + \*i'); le swahili ayant été langue de prestige et de pouvoir aux Comores au XIX<sup>e</sup> siècle, moment où beaucoup d'emprunts à l'arabe sont entrés en swahili qui, dans l'hypothèse en question, a servi d'intermédiaire vers le shingazidja, on peut imaginer que

i) des termes d'origine arabe ayant été empruntés en grand nombre au swahili par le shingazidja déjà munis d'un causatif, ce causatif ait été abstrait et généralisé en shingazidja à tous les mots d'origine arabe, ii) ou que, moins vraisemblablement nous semble-t-il, le shingazidja ait emprunté directement le causatif swahili pour les termes d'emprunt.

Cette hypothèse d'une influence swahili ne nous convainc toutefois pas entièrement, il nous semble que dans ce cas, la distribution en shingazidja des formes avec -ish- dans les emprunts ne devrait pas être aussi générale qu'on le constate, et qu'il devrait y avoir beaucoup plus de doublets avec des formes 'régulières' que les quelques cas recensés.

Nous examinons à présent les réalisations.

+ -(i)z-, -(e)z-, -(i)s-, -(e)s- < \*i' : cette extension entraîne un certain nombre d'alternance en position finale de radical.

Cela s'applique à des racines primaires (ou complexes) et à des constructions composites, où c'est le choix de l'allomorphe (sourde ou sonore) qui permet de déterminer la forme sur laquelle le causatif est constitué.

- fana, ressembler →
- \*-fan-ih- → -fanisa, comparer
- ona, voir →
- \*-on-eh- → -onesa, montrer
- fanya, faire →
- \*-fany-ih- → -fanyisa, faire faire
- hima, être debout →
- \*-him-ih- → -himisa, maintenir
- homa, tarder →
- \*-hom-eh- → -homesa, retarder (qqun)
- anda, commencer →
- \*-and-ih- → -andisa, commencer
- enda, aller →
- \*-end-eh- → -endesa, conduire
- zaya, accoucher →
- \*-zal-ih- → -zalisa, accoucher (tr.)

causatifs avec /z/ : < finale y ou l < \*d : on suppose une forme intermédiaire avec une finale y ou l , qui peut être celle de l'extension donnée comme -i-/-e- correspondant à \*id (et que l'on note alors par -il-):

- lima, cultiver →
- \*-lim-il- → -limiza, faire cultiver | -lim-il-i-|
- linda, attendre →
- \*-lind-il- → -lindiza, faire attendre | -lind-il-i-|
- vura, tirer →
- \*-vur-il-i- → -vuriza, lancer | -vur-il-i-|
- yela, laver :
- \*-yel-el- → -eleza, expliquer | -yel-el-i-|

La racine primaire peut ne pas être attestée :

- angamia, disparaître : \*-angam-il- →
- angamiza, détruire | -angam-il-i-|
- djipvia, satisfaire : \*-djipv-il-
- -djipviza, faire plaisir à | -djipv-il-i-|

Ainsi s'expliquent des 'doublets' tels que :

- ola, pourrir < \*-bòd- →
  - ozà, faire pourrir (par ex, un fruit avarié qui en contamine d'autre) ; | -ol-i-|
  - olesa, faire pourrir (volontairement) : \*| -ol-eh-| → | -ol-eh-i-|
  - ruma, envoyer < \*-tóm- →
  - rumiza, appeler : \*-rum-il- → -rum-il-i-
  - rumisa, organiser : \*-rum-ih- → -rum-ih-i-
  - para, obtenir < \*-pát- →
  - patsa, prêter : -par-i-
  - parisa, se procurer : \*par-ih- → | -par-ih-i-|
  - pvoa, se refroidir < \*-pód- →
  - poza, refroidir (tr.) : -pol-i-
  - pvolesa, refroidir volontairement (tr.) : \*-pol-eh- → | -pol-eh-i-|
- On note ici l'alternance initiale pv / p
- lala, coucher →
  - laza, coucher qqun : -lal-i-
  - laliza, enterrer un objet sacré avec un cadavre : \*| -lal-il-| → | -lal-ih-i-|

Dans quelques cas, les sens sont identiques : les deux formes composites se rejoignent :

- tsonga, provoquer →
  - tsongeza, -tsongesa, provoquer :
    - tsongeza : \*<sup>-</sup>-tsong-el- → -tsong-el-i-
    - tsongesa : \*<sup>-</sup>-tsong-eh- → -tsong-eh-i-
- djua, savoir < \*<sup>-</sup>-jú<sup>-</sup>b- →
  - djuliza, -djua, faire savoir, annoncer :
  - djuliza : \*<sup>-</sup>-djul-il- → -dju-li-li-
  - djuza : -djul-i-

Sous l'angle comparatif, on note ici la réfection de cette racine par rapport à son étymon: \*b et \*d ont en finale radicale les mêmes reflètes (ð, y ou w, selon l'environnement vocalique, voir p. 1); l est par ailleurs un reflet possible de \*d en intervocalique, appliqué ici par analogie avec les racines dont la consonne finale reflète \*d (type \*<sup>-</sup>pód- + -pvoa, -poza)

\*-ish- / -esh- < \*ik

Cet allomorphe est presque exclusivement consacré aux mots d'emprunt à l'arabe; on le trouve toutefois dans quelques racines bantu.

La voyelle est conditionnée par la voyelle radicale - dans les emprunts, on a presque toujours /i/, puisque ceux-ci n'ont qu'exceptionnellement des voyelles intermédiaires.

- taabisha, faire de la peine < taabu, peine
- karibisha, rapprocher < karibu, près de
- badilisha, changer < -badili, changer

Quelques emprunts forment le causatif avec -is- ou -iz- ( < -i- ) (on doit supposer ici aussi des formes intermédiaires pour expliquer la répartition des allomorphes):

- wadjiri, prendre en location → -wadjirisa, donner en location :
  - wadjir-ih-i-
- daina, toujours + -dainiza, continuer : -daim-il-i- (également, -daimisha, continuer (= -dumisha))
- swali, prier + -swaliza, faire prier : -swal-il-i- (également -swalisha, même sens)

Pas plus dans ce cas qu'ailleurs, nous n'avons d'explication à proposer aux formes 'composites' - il semble que ce soit là un fait de langue.

\*sh- se rencontre dans le fonds bantu dans quelques racines :

- kopa, emprunter → -kopesha, prêter
- vuka, (se) sauver → -vukisha, sauver qqun

Ces mots présentent une irrégularité phonétique (occlusives intervocaliques), qui nous font supposer une influence swahili; on peut admettre que la forme du causatif montre aussi une telle influence.

Il faut noter que la forme du causatif est peu prévisible; d'une façon générale, il semble que la forme de l'extension sentie actuellement dans la langue est une sifflante (< /z, s/ ), soit la palatale /sh/.

A l'appui de cette interprétation, les doubles causatifs, où une forme avec -iz- / -is- / -ish- est construite sur une racine déjà causative (alternance):

- dunga, suivre → -dundza, imiter  
→ -dundzanisa, aligner
- lingana, égaler → -lindzana, comparer  
→ -lindzanisha, comparer
- eshelea, se souvenir : -eshelel-  
→ -esheleziza faire se souvenir
- vura, tirer → -vuriza, lancer  
→ -vuriziza, (se) faire lancer

De ce fait, la référence à une construction causative dans le lexique fera appel à l'un de ces symboles (plutôt qu'à la forme étymologique -i-).

Valeurs :

Le rôle essentiel de cette extension est d'introduire un participant supplémentaire au procès: la fonction diffère selon la construction normale du verbe:

- verbes déjà transitifs : le causatif permet d'exprimer celui à qui l'on "fait faire", c'est une valeur proprement causative;
- verbes non transitifs : le causatif les rend transitifs

On a vu, dans les exemples donnés pour expliciter les réalisations, que les traductions oscillaient entre ces deux valeurs.

Dans les emprunts, on retrouve, outre ces deux valeurs, une fonction de renforcement de la transitivité du verbe.

Exemples :

- \* valeur causative (sur transitif) :

ngamlo nyama, je mange de la viande  
ngamhuliso nyama, je te fais manger de la viande  
haono enyunba, il/elle a vu la maison  
haonega Mariama enyunba, il/elle a montré à M. la maison  
hafahamu, il/elle a compris  
hawafahamisha, il/elle leur a expliqué

- \* valeur transitive (sur intransitif) :

emarkabu iroro, le bateau a coulé  
warotsa emarkabu, ils ont coulé le bateau  
eshehitswa shiwehe, 'la tête est penchée', il/elle a la tête penchée  
hawege eshehitswa, il/elle a penché la tête  
halola, il s'est marié (d'un homme)  
hawaloza, elle les a mariés (d'une entremetteuse)

mots d'emprunt :

ehazi idumu, le travail continue  
wadumisha ehazi hata ..., ils ont prolongé le travail jusqu'à ...

x valeur de renforcement (emprunt) :

tsibadili nia, j'ai changé d'idée  
 tsibadilisha emaesha, j'ai modifié [les conditions] de vie

b-4) extension -an- (-lan-) < *lan*

Conditionnement : l'allomorphe -lan- s'emploie après racine à finale vocalique (même problématique que pour la première extension de cette catégorie):

-ona, voir → -onana, se voir  
 -rongoa, parler → -rongolana, se parler

Valeur :

Cette extension implique un sujet multiple (et pluriel), qui fait fonction d'agent et de patient: l'action est faite par A sur B et B sur A, A & B étant impliqués dans le sujet, d'où son appellation de réciproque.

Le réciproque exclut tout objet grammatical; il est incompatible avec l'extension passive -iw-; le sujet, le plus souvent, appartient souvent aux animés.

Exemples :

haono Mariama, il/elle a vu M.

waonana, ils/elles se sont vus/es

monano, réunion

halola, il s'est marié

walolana, ils se sont mariés [ensemble] (terminologie moderniste, qui sous-entend qu'il s'agit d'un mariage résultant du choix des époux)

nots d'emprunt :

bahusudi edjirani, il envie son voisin

wahusudiana, ils s'envient [les uns les autres]

c) Extensions syntaxiques (de diathèse)

a) valeur:

ces extensions affectent essentiellement le sens de la relation du procès aux participants en modifiant leur fonction dans l'énoncé: c'est la raison pour laquelle nous parlons d'extensions de diathèse (ou de voix).

Comme pour les extensions sémantico-syntaxiques, c'est à travers les actualisations verbales que la valeur des extensions de diathèse est le mieux mise en lumière.

Elles peuvent affecter fonds bantu comme emprunts.

Elles sont au nombre de deux dans notre analyse, passif et moyen-neutre.

c-1) extension passive : -w-, -lw-, -iw- < #u, #ibu

Conditionnement :

Après consonne finale radicale, le passif -w- s'amalgame à cette consonne pour former une vélarisée (d'où la série complexe des vélarisées).

Après voyelle, l'allomorphe -lw- est utilisé :

-anba, dire → -anbwa, être dit

-rongoa, parler → -rongolwa, 'être parlé'

Ici comme pour l'applicatif, la forme -lw- est donnée synchroniquement, à partir notamment du comportement de certains emprunts ; il est clair qu'il s'agit d'une reconstruction à partir de l'intégration de la consonne finale radicale étymologique :

-zaya, accoucher < \*bi'ád-

+ -zalwa, être mis au monde, naître ; -zal-w- → -za-lw-

-toa, sortir (tr.) < \*tód-

+ -tolwa, être choisi ; -tol-w- → -to-lw-, sur le modèle de

-abudu, adorer (ar.) →

+ -abudulwa, être adoré -abudu-lw-

C'est -iw- qui est utilisé dans les mots d'emprunt, quoique -lw- soit aussi attesté, dans certains cas, après finale vocalique (ci-dessus) :

-husudi, envier → -husudiwa, être envié (ar.)

-ruhusu, permettre → -ruhusiwa, être autorisé (ar.)

-abudulwa, ci-dessus

Valeur :

Celle-ci est celle d'un passif: le patient devient le sujet grammatical, l'agent peut être éventuellement présent, dans ce cas, il est introduit par un fonctionnel *ad hoc*, ni; les verbes à double objet admettent deux constructions passives.

Cette extension est souvent utilisée pour rendre un agent diffus ou 'masqué', correspondant aux tournures avec 'on' du français. Elle est fréquemment co-occurente d'un applicatif.

Exemples :

Ali hapusua efalasika, A. a cassé la bouteille

efalasika ipasulwa (ni Ali), la bouteille a été cassée (par A.)

haanba uka ..., il/elle a dit que ...

haanbya Mohamadi, 'il est appelé M.', on l'appelle M., il s'appelle M.

hanianbia uka ..., il/elle m'a dit que .. (applic.)

tsianbiya uka ..., 'j'ai été dit que ...', on m'a dit que .. (avec applicatif)

hanika wandru mapesa, il/elle a donné de l'argent aux gens (verbe à double objet)

(owandru) wanikiya mapesa, 'les gens ont été donné de l'argent', les gens ont reçu de l'argent (avec applicatif)

emapesa yawanikiya, de l'argent leur a été donné (avec applicatif)

Nota

La présence de l'extension passive entraîne, comme nous l'avons signalé, la neutralisation des distinctions entre les suffixes verbaux ; *-o* (inaccompli) est toujours ramené à *-a*, suffixe verbal non marqué à valeur générale, *-e* (subjonctif) peut être ramené à *-a* :

suf. *-o* :

actif : Mohamadi ngololo Aziza, M, épouse A.  
passif (théorique) : Aziza ngololwa ni Mohamadi, A, est épousée par M, (et non \*ngololwo)

suf. *-e* :

actif : (lazimu) Mohamadi yalole Aziza, (il faut) que M, épouse A,  
passif : (lazimu) Aziza yalolwa/yalolwe ni M,, (il faut) que A, soit épousée par M, (les variations paraissent ici individuelles)

2) extension *-ih-*, *-eh-*, *-h-* < \*ik

Conditionnement : *-Vh-* après consonne, *-h-* après voyelle; comme on l'a déjà vu dans la plupart des extensions, le timbre de la voyelle de l'extension est conditionnée par celle du radical :

*-zima*, éteindre → *-zimiha* : *-zim-ih-*

*-pvera*, plier → *-pvereha* : *-pver-eh-*

*-pasua*, casser → *-pasuha* : *-pasu-h-*

Dans ce dernier cas, on peut aussi analyser comme \**-pas-uh-*, faisant intervenir une extension *-uh-* (< \*uk)

Cette extension ne s'applique pas aux racines primaires à finale /h/; on relève au contraire des constructions composites, dont la forme intermédiaire montre la finale reconstruite comme /h/ (< *-ih-*) : dans ce cas, cette consonne se palatalise devant l'extension syntaxique *-ih-* :

*-ruma*, envoyer → \**-rumiha* → *-rumishiha*, servir : *-rum-ih-ih-*

On est amené à poser une forme intermédiaire \**-rumiha* ; *-rum-ih-*, dont nous avons vu d'autres exemples dans la formation du causatif; l'extension *-ih-* n'est pas celle dont nous parlons ici; leur co-occurrence possible est un des éléments à l'appui de cette différentiation); le verbe résultant *-rumishiha* fait donc apparaître une succession de deux extensions *-ih-* homophones : *-rum-ih-ih-*

Avec les emprunts, on a en général seulement *-ih-* :

*-taabiha*, avoir de la peine < *taabu*, peine (ar.)

*-dhulumiha*, être exploité < *-dhulumu*, exploiter (ar.)

leur :

Cette extension, souvent appelée moyen-neutre, correspond aux deux valeurs attribuées par GUTHRIE à l'une des extensions notées \*ik : neutre et potentiel (*neuter*, *potential*).

Le moyen-neutre s'apparente, par le sens de la relation (et la construction) au passif, en ce sens que le patient est normalement en position sujet; mais ici, l'agent ne peut être exprimé. Le moyen-neutre

définit une sorte de circularité de l'action, où l'intervention d'un agent particulier n'a pas de place.

Selon le contexte et le sens du procès, cette extension permet de rendre soit la voie 'moyenne' ou pronominale du français, soit une valeur potentielle, correspondant aux suffixes -able/-ible des adjectifs déverbatifs français.

Exemples :

x valeur 'moyenne'

Ali havundzu enyungu, A. a cassé la marmite  
 enyungu ivundziha, la marmite (s')est cassée  
 wagubanya (emaele), ils/elles ont remué (le riz)  
 wagubanyiha, ils/elles sont mélangés, en désordre  
 mfanyo, une action  
 mfanyiho, une action faite (envisagée dans sa réalisation)

x valeur potentielle :

tsili maele, j'ai mangé du riz  
 emaele yaliha, le riz est mangeable

#### Combinaison d'extensions

Extensions sémantiques, sémantico-syntaxiques et syntaxiques peuvent être concomitantes, jusqu'à trois degrés, avec différents ordres possibles selon la place du réciproque; la seule incompatibilité réside entre réciproque et passif, qui s'excluent mutuellement (nous ne tenons compte ici que des extensions identifiées et non des formes composites).

L'ordre est indiqué par les schémas suivants, qui reprennent les deux positions du réciproque:

i) réciproque en tête des extensions sémantico-syntaxiques (rares)

	extensions		
	1	2	3
	- sém. - u	sém.-synt. i, i, s	- synt. - w
Rac	- ats - am any	- an Z, Z, i	- h

ii) réciproque en dernière position

	u	i, i, Z	w	
Rac	- ats - am any	- Z, Z, i	- ih	- an -

Chaque forme vaut pour l'ensemble de ses allomorphes; nous rappelons les appellations courantes:

- u- : -u-, -o- (réversif)
- ar- : -ar-, -ats- (contactif)
- am- : -am- (statif)
- any- : -any- (associatif)
- an- : -an-, -lan- (réciproque)
- i- : -i-, -e-, -li-, -le- (applicatif)
- i-i- : -ili-, -ele- (intensif, double applicatif)
- Z- : -iz-, -is-, -ez-, -es-, -ish-, -esh- (causatif)
- Z-Z- : -iziz- (double causatif)
- w- : -iw-, -ew-, -lw- (passif)
- ih- : -ih-, -eh-, -h- (moyen-neutre)

Les formes superposées sont mutuellement exclusives.  
Toutes les extensions sont facultatives; les flèches indiquent l'incompatibilité entre passif et réciproque

a) avec extension sémantique

1) extension sémantique + extension sémantico-syntaxique :

\* réversif + applicatif + réciproque :

fung- |notion d'attacher|

> -fung-u- |notion de détacher| (+ rév.)

> -fungu-li- |notion de détacher pour| (+ appli.)

> -funguli-an- |notion de se détacher les uns des autres| (+ réc.)

\* contactif + causatif :

-nam- |notion de coller|

> -nam-ats-z- |notion de coller| (+ cont. + caus.)

\* contactif + applicatif + réciproque :

-zing- |notion de tourner|

> -zing-ar- |notion de tenir| (cont.)

> -zingar-i- |notion de tenir (qqun)| (+ appli.)

> -zingari-an- |notion de se tenir l'un l'autre| (+ réc.)

rizingariana, on se tenait les uns les autres [par la main, etc]

ii) extension sémantique + extension syntaxique :

x réversif + passif :

- fung- (notion d'attacher)
- > -fung-u- (notion de détacher) (+ rév.)
- > -fungu-lw- (notion d'être détaché) (+ passif)

x réversif + moyen-neutre :

- fung- (notion d'attacher)
- > -fung-u- (notion de détacher) (+ rév.)
- > -fungu-h- (notion d'être détachable) (+ m-n)

iii) extension sémantique + extension sémantico-syntaxique + extension syntaxique :

x réversif + applicatif + passif :

- fung- (notion d'attacher)
- > -fung-u- (notion de détacher) (+ rév.)
- fungu-li- (notion de détacher pour) (+ appli.)
- > -funguli-w- (notion d'être détaché pour ..) (+ passif)

β) extension sémantico-syntaxique + extension sémantico-syntaxique :

x applicatif + causatif : voir causatif sur formes composites

x causatif + applicatif :

- l- (notion de manger)
- > -l-is- (notion de nourrir (qqn)) (+ caus.)
- > -lis-i- (notion de nourrir (qqn pour qqn)) (+ appli.)

**halisi mwana hahe**, il/elle a nourri son enfant (même ambiguïté qu'en français sur le 'propriétaire' de l'enfant)

**hamlisia mwana**, il/elle lui a nourri l'enfant, il/elle a nourri son enfant à sa place

x applicatif + causatif + réciproque :

- ishi- (notion d'entendre)
- > -ishi-li- (notion d'écouter) (+ appli.)
- > -ishili-z- (notion d'écouter qqn) (+ caus.)
- > -ishiliz-an- (notion de bien s'entendre) (+ réc.)

**waishilizana**, ils/elles s'entendaient bien

x causatif + réciproque :

- pvir- (notion de passer (intr.))
- > -pvir-z- (notion de passer (tr.)) (+ caus.)
- > -pvits-an- (notion de se croiser) (+ réc.)

**ripvitsana**, nous nous sommes croisés

-rongo- (notion de parler)

- > -rongo-z- (notion de parler à qqn)
- > -rongoz-an- (notion de se parler)

**warongozana**, ils/elles ont bavardé

γ) extension sémantico-syntaxique + extension syntaxique :

x applicatif + passif :

-amb- notion de dire!

> -amb-i- notion de dire à! (+ appli.)

> -ambi-w- notion d'être dit ["on (me,...) dit .."]! (+ passif)

tsianbiwa uka ..., on m'a dit que ...

applicatif + moyen-neutre :

-he- notion de monter!

> -he-le- notion de monter au-dessus de ..! (+ appli.)

> -hele-h- notion d'être monté sur! (+ m-n)

tsihuheleha usiu, je suis en colère contre toi ("je t'ai élevé le visage")

causatif + passif :

-sih- notion de descendre (intr.)!

> -sih-z- notion de descendre (tr.)! (+ caus.)

> -sis-iw- notion d'être descendu! (+ passif)

emizigo zisisiwa, on a descendu les valises

causatif + moyen-neutre :

m- n-neutre + réciproque :

-fan- notion de ressembler!

> \*-fan-ih- notion d'être ressemblant!

-fan-ih-an- notion de se ressembler!

wafanibana, ils/elles se ressemblent

Remarques : racines complexes ou étendues ?

Dans son article de 1962, GUTHRIE définit les racines complexes comme suit :  
*"any longer radicals which cannot be broken down into a simplex radical and an extension are termed 'complex', although the shapes and complex and extended radicals may be quite comparable."* (GUTHRIE op. cit. 93).

De fait, pour chaque extension identifiée, il existe de nombreux constituants, dont la forme et la valeur peuvent laisser supposer qu'il s'agit de formes étendues alors que les racines primaires correspondantes ne sont pas attestées; il existe aussi des cas où les deux formes sont attestées, mais sans que l'on puisse établir de rapport sémantique, même lâche, entre elles.

Dans cet esprit, en cas de non attestation de la racine primaire, nous admettons une décomposition lorsque au moins deux constituants syntaxiques apparaissent reliées lexicologiquement par des processus réguliers à travers une racine hypothétique, qui serait leur élément commun, ces constituants étant soit deux verbaux, soit un verbal et un nominal.

La racine hypothétique, construite, le plus souvent complexe, apparaîtra alors dans le lexique sous forme d'un verbal précédé d'un +, qui signale son statut hypothétique, en maigre - puisque ce n'est pas, par définition, une forme énoncée, et sans glose. Elle introduira aux formes étendues et éventuellement aux nominaux (ex. ci-dessous).

Il importe aussi de remarquer que l'absence de racines primaires tout comme la non reconnaissance de rapport sémantique ou fonctionnel, peuvent provenir seulement des lacunes du corpus ... et qu'à mesure de la progression des recherches, le nombre tend à en diminuer. Il faut donc se garder de jugements définitifs en ce domaine.

Nous donnons quelques exemples de racines complexes dont la décomposition est opérée à partir de deux attestations au moins de constituants syntaxiques reliés, et de racines complexes non décomposées. Les décompositions peuvent

appuyer sur les formations composites (causatif, moyen-neutre), qui permettent de reconstituer des formes intermédiaires.

racines postulées (au moins deux attestations de constituants syntaxiques liés)

1) à travers la comparaison de deux constituants verbaux étendus

-demea: < \*-dèm-

issu de la comparaison entre

-lemewa, être fatigué : passif

-lemeza, fatiguer (qqun) : causatif

Il s'agit ici d'une racine complexe évoquant l'applicatif -e-, ce que montre aussi la formation du causatif

-lekea:

issu de la comparaison entre

-lekelea, s'installer : applicatif

-lekeza, installer (qqun) : applicatif

Il s'agit ici d'une racine complexe évoquant l'applicatif -e-

-angama:

issu de la comparaison entre

-angamia, (se) détruire : applicatif

-angamiza, détruire : causatif

Il s'agit ici d'une racine complexe évoquant le statif -an-

douteux :

-pvaha<sup>3</sup>: < \*-pák-

issu de la comparaison entre

-pvahiza, emmener (qqun) : causatif

-pashia, embarquer : applic.

On note l'alternance initiale pv/p < \*-p et l'absence inattendue et inexplicable de palatalisation de h devant la voyelle de l'extension

11) à travers la comparaison d'un nominal et d'un verbal étendu

-pep- :

issu de la comparaison entre

npbepvo, vent < \*-pépò

-pepea, flotter au vent

Rappelons que npb correspond à IN-pvI

racines complexes non décomposables (dans l'état de nos connaissances)

racines complexes se présentent pour chaque extension: les plus fréquentes

constituées de racines de type bi-syllabique à finale vocalique (-CVCV-,

-V-), où la voyelle 'finale' évoque l'applicatif ou le réversif.

passons en revue les différents types de racines complexes, selon les extensions qu'elles évoquent.

1) racines complexes évoquant des extensions sémantiques:

• réversif -u-, -o- :

-fudua, piler

-hunua, désherber < \*-kún-

L'étymologie suppose que la forme est originellement étendue

-pvahua, servir un plat < \*-pákud-

L'extension (si telle doit être l'interprétation de la séquence \*-ud-) est manifeste dans l'étymon

-rongua, parler

-hobua, décrocher < \*-kòb-

L'étymologie suppose que la forme est originellement étendue

• associatif -any- :

-danganya, tromper

-dabwanya, troubler

• statif -am- :

-angamia détruire (ici, applicatif, voir ci-dessus)

Signalons enfin un cas d'extension non retenu dans l'inventaire ci-dessus, par défaut d'attestations, mais dont la forme est donnée en BC :

neutre -uh- < \*uk

-lalua, dissoudre < \*-dád-, briser

L'étymologie suppose que l'extension a affecté une forme simple maintenant non attestée dans ce sens (\*-lala)

Nous avons signalé quelques autres mots où apparaît une séquence [uh] que l'on pourrait rapporter à cette extension: -danguha, tomber ; -pasua, briser

ii) racines complexes évoquant des extensions sémantico-syntaxiques:

• applicatif -i-, -e-, etc :

-erea, apporter (à)

-epvea, éviter < \*-yèpid-

L'extension (si telle doit être l'interprétation de la séquence \*-id-) est manifeste dans l'étymon

-hutsia, porter un pagne < \*-kút-, envelopper

-walalia, guetter

• causatif -z-, -s-, -is-, -ez-, etc:

-pvurisa, se dépouiller

-tsihitsa, tresser

-fikiza, reboucher

• réciproque -an- :

-fumania, surprendre

-anahana, se souvenir

iii) racines complexes évoquant des extensions syntaxiques:

• passif -iw-, -ew- :

-shelewa, se réveiller

• -ih- des formes composites:

-andziha, écrire < \*-yándik-

-aniha, étendre < \*-yánik-

Dans quelques cas, l'absence de tout rapport sémantique ou fonctionnel entre deux formes, dont l'une pourrait être, sur le seul plan formel, la base de l'autre interdit la mise en relation:

- tsaha, chercher ↔ -tsahanya, pétrir
- unba, créer ↔ -unbia, chanter une berceuse

#### 4) Extensions dé-adjectivales

En principe, les extensions portent sur des racines qu'elles transforment en racines étendues, le résultat en restant un élément de même statut (racine), qui ne préjuge pas de l'appartenance catégorielle des constituants qui en sont issus (verbaux ou nominaux): c'est le fonctionnement que nous venons de voir.

Toutefois, on relève deux types d'extensions portant sur des thèmes adjectivaux qui eux entraînent un changement de catégorie grammaticale: le résultat permet l'obtention de verbes; nous n'avons pas relevé en effet que les formes obtenues s'assimilent aux racines bi-valentes, elles ne paraissent être actualisées que comme verbaux.

Ces extensions ne sont pas productives; les thèmes adjectivaux qui présentent ce fonctionnement ont tous un étymon chez GUTHRIE, de même que les morphèmes d'extensions opérant.

En raison du peu d'attestations de ce phénomène, l'ensemble des cas relevés figure ici.

#### a) extension "neutre" : -pv- , -p- < \*p

Le terme de "neutre" traduit l'appellation *neutral* de GUTHRIE, qui signale aussi que cette extension s'applique sur un thème.  
(pv est le reflet normal de \*p dans cette position (intervocalique), p étant le "reflet minoritaire", voir p )

Les deux allomorphes sont en distribution lexicale.

On trouve dans tous les cas chez GUTHRIE l'étymon de l'adjectif comme du verbe.

- x -pv- :  
-angu, facile < \*-yàngu  
→ -angupva, devenir facile < \*-yàngup-
- nene, gros < \*-néné  
→ -nenepva, grossir < \*-nénép-
- x -p- :  
-le, grand < \*-dè

→ -lepa, grandir < \*-dèp-

On ne voit guère comment expliquer l'irrégularité de cette forme (occlusive en intervocalique); sans doute s'agit-il d'une formation ancienne, avant le passage occlusive → continue, entrée ainsi dans la langue, alors que les autres formes, régulières, seraient plus tardives et auraient sans doute été formées en shingazidja même; cela supposerait que l'extension a joué dans un état du shingazidja une autre hypothèse serait l'emprunt à une langue qui n'a pas connu cette évolution.

Mais rien ne permet de comprendre pourquoi l'évolution de mots semblables, par la forme et la nature, est différente.

b) : extension -h- < \*ik

Cette extension s'applique au moins à un thème adjectival, avec la même valeur d'ailleurs que la précédente:

-radji, large < \*-tádi  
→ -radjiha, s'élargir

Ici, point d'étymon à la forme verbale.

La valeur du moyen-neutre est tout à fait adaptée à l'emploi qui en est fait ici.

## II LES EMPRUNTS

ans le fonds bantu, nous avons vu que les constituants syntaxiques analysaient en monèmes, lexicaux (racines et thèmes), et grammaticaux obligatoires, actualisateurs affixés, facultatifs, extensions).

ous avons vu que, pour les termes invariables, il y avait coïncidence entre catégorie de langue et catégorie de discours, ces constituants n'étant pas analysables.

ous examinons maintenant la manière dont les emprunts d'origine orientale, essentiellement arabe, s'intègrent morphologiquement au grand-comorien et quelle influence leur présence a pu avoir sur les systèmes morphologique et lexicologique.

on constate d'emblée une distinction entre deux états d'intégration, qui recoupe les deux plans de l'analyse morphologique:

- \* un état de non-naturalisation, ou d'intégration minimale (ou intégration passive) où les emprunts s'intègrent directement aux catégories de discours ou constituants syntaxiques,
- \* un état de naturalisation ou d'intégration achevée (ou intégration active), où les emprunts admettent les mécanismes de formation bantu: on peut alors reconstruire des lexèmes (racines et/ou thèmes), sur le modèle des lexèmes de fonds bantu.

ous définissons un critère ( $\pm$  natif) pour qualifier chacun de ces états.

es deux classes qui ressortent de cette distinction sont évidemment des clartés: à côté de fonctionnements nettement marqués qui situent les termes d'un côté ou de l'autre, on relève un fonctionnement que l'on peut considérer

comme intermédiaire, et l'on pourra, dans une certaine perspective, admettre une catégorie de transition.

Sur le plan lexicologique, parce qu'ils ne constituent pas des familles dérivationnelles sur le mode bantu - puisque non analysables en affixes et lexèmes - les termes d'emprunt non naturalisés ne sont pas liés entre eux par la combinatoire lexicologique bantu; cela ne signifie pas qu'ils soient pour autant nécessairement isolés du point de vue lexicologique: d'une part, à leur niveau, ils peuvent être apparentés avec d'autres emprunts, naturalisés ou non, et d'autre part, ils peuvent être la base lexicale (nous parlerons de support lexical) d'emprunts naturalisés, qui admettent, eux, les processus bantu.

Nous envisageons successivement l'aspect morphologique, dans les emprunts non naturalisés et naturalisés, et puis l'aspect lexicologique, c'est-à-dire les réseaux de relations lexicologiques développés par les termes d'emprunt entre eux.

#### Nota

L'intégration, même minimale ou passive, suppose une intégration sur le plan phonique, notamment syllabique: les emprunts se distinguent ainsi des "citations", fréquentes également en shingazidja, qui sont des éléments étrangers rapportés et qui ne rentrent pas dans notre problématique.

## I Aspect morphologique

Il est clair, dans le processus d'emprunt du shingazidja aux langues bantoues, qu'ont pénétré des mots, c'est-à-dire des unités lexicales et sémantiques déjà formées, et non des éléments abstraits, lexématiques ou morphématiques.

## I-1 Emprunts intégrés sans modification aux catégories de discours

Nous traitons ici de ces unités qui s'insèrent directement comme constituants syntaxiques, variables (noms ou verbes) ou invariables, sans subir de changement.

On constatera, sans qu'il faille selon nous en faire un principe, la conservation très générale des catégories grammaticales de la langue source quand elles correspondent à celles du shingazidja, ce qui est le cas des grandes divisions des parties du discours); les termes variables ne sont pas analysables en monèmes, et n'admettent pas les différents processus qui affectent les constituants syntaxiques bantu.

Sur le plan morphologique, les phénomènes varient en fonction de la catégorie syntaxique 'réceptrice'; il y a lieu de discriminer en premier lieu entre catégories 'variables' et catégories 'invariables', puis à l'intérieur des catégories 'variables', entre verbes et noms.

1-1) Emprunts non naturalisés intégrés dans les catégories variables

Sur le plan morphologique, la spécificité essentielle des emprunts entrés directement dans les catégories des verbaux et des nominaux est le fait qu'ils n'admettent pas les morphèmes actualisateurs nécessaires aux constituants syntaxiques 'bantu', suffixes ni préfixes nominaux:

- ces nominaux et ces verbaux n'admettent pas les suffixes vocaliques (ou les voyelles suffixielles) propres aux termes bantu issus de racines;
- ces nominaux n'admettent pas les préfixes de classe que l'on trouve dans les nominaux bantu.

Ce sont ces deux 'refus' d'acceptation des morphèmes et mécanismes bantu, au profit du maintien de la forme originelle, qui font parler d'intégration passive.

Nominaux et verbaux non intégrés ont donc en commun le caractère lexical de leur voyelle finale:

- x elle est étymologique pour les nominaux dont la source est à finale vocalique
- x elle est conditionnée, de façon systématique pour les verbes, et pour les nominaux dont la source est à finale consonantique.

Nous donnons brièvement le conditionnement de cette voyelle finale, puis traitons de l'intégration passive des termes dans chacune de ces deux catégories (verbaux puis nominaux).

**a) voyelle finale**

Nous signalons brièvement les facteurs intervenant dans la détermination du timbre de la voyelle finale (l'on se reportera à l'étude des voyelles pour davantage d'exemples).

x voyelle conditionnée

Le conditionnement qui est presque absolument strict pour les verbes ne s'applique pas aux nominaux à voyelle finale étymologique.

• après /r/ et /s/, harmonie vocalique sur la voyelle radicale quand celle-ci est /u/ ou /i/ :

- fasiri, expliquer < ar. *fassara*
- hisi, se rendre compte < ar. *h.assa*
- hairi, grâce divine < ar. *kair*

• après labiales /b, m, f/ et parfois après /d/ : /u/ :

- djibu, répondre < ar. *'ajäba*
- fahamu, comprendre < ar. *fahima*
- adabu, respect < ar. *adab*

• après autres phonèmes : /i/ :

- rudi, revenir < ar. *radda*
- tahani, moudre < ar. *t.ah.ana*
- asubuhi, matin < ar. *as.-s.ubh.*

voyelle étymologique :

Dans le cas où la source est à finale vocalique, les nominaux d'emprunt non naturalisés conservent leur voyelle étymologique :

- ada, coutume < ar. *äda*
- nabii, prophète < ar. *nabi'y*

Il y a ici blocage des suffixes; la présence d'une voyelle finale étant nécessaire du fait de la structure syllabique du shingazidja, une voyelle épenthétique est suppléée le cas échéant pour les nominaux, systématiquement pour les verbaux, sur la base d'un conditionnement phonétique propre au shingazidja.

Nous examinons maintenant l'intégration des verbes, puis des noms.

## b) intégration des verbes

### a) description morphologique

Du point de vue étymologique, rappelons que les verbes d'origine orientale en shingazidja sont tous attestés en arabe (cf. introduction p 47).

(sur la voyelle interne des verbes d'origine arabe, voir voyelles, p 274 ; nous la supposons ici lexicale)

Dans le fonds bantu, les formes verbales sont caractérisées par un changement de suffixe, à valeur grammaticale; ce changement exprime le mode (suffixe -e, subjonctif vs -a, indicatif ou non marqué): pour les verbes d'origine arabe, il

il y a blocage, leur voyelle finale ne se modifiant pas - ce qui n'empêche évidemment pas l'emploi de ces verbes à ces formes (nous donnons un exemple de verbe de fonds bantu pour faire ressortir la spécificité):

bantu:

**ufanya**, faire (hu-fany-a) ; **wafanye** !, qu'ils fassent ! (wa-fany-e)

arabe :

**ufahamu**, faire attention (hu-faham-u) ; **wafahamu** !, qu'ils fassent attention ! (wa-faham-u)

Les conséquences de ce blocage sont d'ordre syntaxique, et échappent quelque peu à notre propos: il y a perte d'une information syntaxique, suppléée par le contexte (on peut sans doute mettre cela en rapport avec le développement de l'emploi, toutefois facultatif, de fonctionnels tels que des conjonctions de subordination, souvent elles mêmes des emprunts ...; les langues bantu en effet utilisent peu de fonctionnels.

Par contre, ces verbes admettent sans restriction tous les éléments préfixés, porteurs de marques de catégories verbales, dont le préfixe infinitif de cl. 15, (catégorie nomino-verbale): c'est le seul exemple, on l'avouera particulier, où des emprunts 'non naturalisés' admettent un préfixe de classe; nous dirons que cela est dû au caractère ambigu de cette classe.

#### *β) analyse ou interprétation structurelle*

Les verbes bantu étaient analysés en préf. - racine - suffixe, comment analyser les verbes d'emprunt de cette catégorie, ou, en d'autres termes, comment interpréter l'absence de suffixe ?

Comme il semblerait tout à fait artificiel de postuler un suffixe Ø, on pourrait peut-être envisager dans une perspective dynamique de postuler une série verbale correspondant aux thèmes du fonds bantu, qui, en principe, sont réservés à la catégorie nominale.

### ② Intégration des nominaux

#### a) description

La composition finale, qui est celle du suffixe pour les nominaux bantu issus de racines est occupée par une voyelle lexicale: cela ne suffirait pas toutefois pour les nominaux à poser une intégration passive, spécifique des emprunts, car cela rejoint la description des nominaux bantu issus de thèmes.

La caractéristique supplémentaire et décisive est que les emprunts non naturalisés n'admettent pas non plus de préfixe de classe.

Cependant, dans le discours, ces nominaux régissent les accords des classes 9/10; aussi, malgré l'absence de préfixe, ces termes sont intégrés dans ces classes (rappelons que le préfixe nominal normal de ces deux classes dans le fonds bantu est {N-} (nasale homorganique)).

Les noms d'emprunt placés dans ces deux classes sont très nombreux; cela tient au large éventail de valeurs sémantiques de ces classes, y compris une valeur non déterminée (divers).

Nous donnons quelques exemples de nominaux d'emprunt de sens divers, régissant les accords de classe 9/10; leur initiale (initiale lexicale) peut être n'importe quelle réalisation correspondant à un phonème de la langue source, y compris des réalisations caractéristiques, pour le fonds bantu, de telle ou telle autre classe: nous suivons le classement que cela suggère:

avec initiale identique à un allomorphe de préfixe

cl. 7 {hi-, shi-, i-} :

**hila**, difficulté < ar. *h.ā'īl*

**istiklale**, indépendance < ar. *istiqlāl*

**shari**, le bien < ar. *s'arr*

cl. 4 {mi-}

**miladi**, début du calendrier < ar. *milād*

cl. 5 {dji-}

**djahazi**, boutre < ar. *jihāz*

cl. 6 {ma-}

**maiti**, cadavre < ar. *mayyit*

**mali**, richesse < ar. *māl*

cl. 9/10 [N-]

**nafaka**, chance < ar. *nafaqa*

x avec initiale occlusive (cf. préf. cl. 5 [Ø'-]) :

**bahari**, mer < ar. *bah.r*

**taabu**, difficulté < ar. *ta'ab*

**dini**, religion < ar. *din*

**kanuni**, article de loi < ar. *qānūn*

x avec initiale continue

**wilaya**, 'préfecture' < ar. *wilāya*

**ladha**, saveur < ar. *ladha*

x autre initiale :

**adabu**, respect < ar. *adab*

**fikra**, pensée < ar. *fikra*

**sitara**, secret < ar. *sitāra*

Nous verrons que cette intégration minimale en classe 9/10 est souvent en variation pour la même source lexicale avec un traitement où l'emprunt est naturalisé.

*β) analyse et interprétation structurelle : une catégorie de transition ?*

Nous avons caractérisé ce type d'intégration, passive, par l'absence de tout préfixe de classe, mais une autre interprétation est possible, qui permettrait une intégration plus satisfaisante dans les catégories 'bantu': en effet, dans la mesure où ces termes commandent les accords de classe, il paraît plus justifié de les intégrer parmi les thèmes en posant un préfixe [Ø-] (bien entendu sans effet sur la consonne initiale des termes):

**hila**, [Ø-hila] ; **sitara**, [Ø-sitara] comme **mwana**, enfant [mu-ana], etc

Cette catégorie apparaîtrait alors comme intermédiaire entre les emprunts non naturalisés et les emprunts intégrés par assimilation à des thèmes, que nous allons voir dans la suite.

### 1-2) Emprunts non naturalisés intégrés dans les catégories invariables

Sous l'appellation d'invariables, nous avons référé, dans le fonds bantu, à des constituants de fonctions diverses, dont le point commun est de ne pas relever

de l'accord de classe, et de ne pouvoir se décomposer en monèmes, reliés par la combinatoire bantu: il s'agit de cas où constituant syntaxique et catégorie de langue (monème) coïncident.

Il se trouve que des termes d'emprunt s'insèrent dans toutes les sous-catégories afférentes à ce type de constituants: l'intégration des emprunts y est en fait facilitée par le fonctionnement même de cette catégorie, puisqu'il n'y a ici affixation d'aucun élément.

On relève l'adaptation phonétique, les termes étant le cas échéant munis d'une voyelle finale conditionnée, selon les principes exposés pour les termes variables.

L'inventaire des invariables d'origine arabe montre (nous reprenons les exemples cités dans l'inventaire des parties du discours)

i) des conjonctions :

*ili*, pour que < ar. *li*  
*lakini*, mais < ar. *lakin*

ii) des circonstants :

*halafu*, après < ar. *kalf*  
*halisi*, vraiment < ar. *kälis*.

iii) des interjections :

*la*, non < ar. *lä*

iv) des quasi-nominaux :

*sita*, six < ar. *sitta*  
*makini*, calme < ar. *makin*

v) des termes 'idéophonisés'

*hataa*, extrêmement < ar. *h.attä*

Ces termes ne prennent ni ne commandent, dans leur emploi usuel, l'accord de classe.

Cela n'est pas inattendu pour des fonctionnels divers, la classe des "quasi-nominaux" mérite explication: elle contient en effet des termes qui appartiennent aux paradigmes de nominaux dépendants sans en avoir le

comportement morphologique; ce comportement est à mettre en parallèle avec ce que nous venons de voir pour les nominaux indépendants.

Soient les énoncés suivants :

mots d'origine arabe :

\* invariable :

ye **makini** / wao **makini**, il/s est/sont calme/s

\* variable :

ye **masikini** / wao **wamasikini**, il/s est/sont pauvre/s

mot de fonds bantu :

ye **wade** / wao **wawade** il/s est/sont malade/s

On voit, dans ces exemples, l'accord de classe en oeuvre pour **-wade** et **-masikini**; cet accord n'opère pas avec **makini**, qui reste invariable.

Dans la catégorie des adjectifs ou nominaux dépendants d'origine arabe, le partage entre termes relevant de l'intégration passive et termes naturalisés (admettant les préfixes de classe) ne paraît pas suivre une logique particulière; on relève en outre des fluctuations pour un même terme.

## I-2 Emprunts naturalisés

Les emprunts apparaissent aussi dans un état naturalisé, où ils sont intégrés dans les processus morphologiques bantu (intégration active): on peut alors retrouver les différents monèmes définis dans l'étude des catégories de langue; on distingue deux classes de nominaux, les uns semblables aux nominaux bantu issus de thèmes, décomposables en préfixe et lexème, les autres semblables aux nominaux bantu issus de racines verbo-nominales, décomposables en préfixe, racine, extension et suffixe, ainsi que des verbes, toujours décomposables en (préfixe), racine, extension et suffixe.

Nous examinons les emprunts assimilés à des thèmes puis ceux assimilés à des racines.

1) Intégration par assimilation à des thèmes

On se rappelle que les thèmes de fonds bantu sont caractérisés par le fait que leur suffixe n'est pas isolable et que, par l'adjonction de préfixes de classe, ils donnent des nominaux et des nominaux uniquement, un même thème pouvant donner différents nominaux selon les préfixes adjoints; nous avons dit toutefois qu'il n'y avait pas ici une grande latitude de combinaison, les cas apparaissant lexicalisés.

Les emprunts arabes que nous assimilons à des nominaux issus de thèmes montrent ce même comportement: ils se présentent comme décomposables en préfixe et élément lexical inanalysable porteur d'information, tel que cet élément accepte en général au moins deux préfixes en opposition de nombre, parfois davantage; nous assimilons cet élément à un thème.

Deux aspects sont à envisager : les modalités d'intégration morphologique et les relations lexicologiques qui en découlent.

### 1-1) Intégration morphologique

Il est intéressant d'examiner de quelle manière les préfixes (c'est-à-dire l'insertion dans une classe nominale) sont ici choisis: deux facteurs interviennent:

- la valeur ou motivation sémantique des préfixes, telle que définie dans le fonds bantu;
- la ressemblance de l'initiale du mot arabe avec un préfixe ou une marque de classe donnée (analogie phonétique)

Ces deux facteurs peuvent parfois se conjuguer et se renforcer l'un l'autre, ou, dans le cas des classes moins précises sémantiquement, ne pas être isolable l'un de l'autre.

On constate aussi des fluctuations, un même mot d'emprunt pouvant être naturalisé - l'initiale sera alors interprétée dans un cadre préfixiel - ou rester non naturalisé - le mot étant intégré en cl. 9 sans préfixe, rejoignant la situation précédemment décrite.

L'intégration des emprunts dans le système de classes a eu pour effet de modifier le contenu sémantique de certaines classes, avec un effet d'appel ou de renforcement: les emprunts nouveaux viendront naturellement s'intégrer dans une classe où des emprunts de même type ou de même valeur se trouvent déjà, le premier lot d'emprunts ayant contribué à donner une nouvelle motivation à cette classe ...

Nous rappelons les préfixes, avec leurs valeurs sémantiques et les relations de nombre entre les classes, dans le fonds bantu (cf. p 129-130):

cl. 1 / 2 : |mu-| / |wa-| : être humain

cl. 3 / 4 : |mu-| / |mi-| : arbres, plantes, parties du corps, action

- cl. 5 / 6 : Ø<sup>h</sup>-l ou ldji-l / lma-l : augmentatif; animaux ; action ; divers
- cl. 6 : lma-l : liquide;
- cl. 7 / 8 : lshi-l, lhi-l, li-l / lzi-l : parties du corps; objet; ustensile;  
diminutif
- cl. 9 / 10 : lN-l / lN-l : divers
- cl. 11 : lu-l : abstraction; activité; unité

Nous traitons d'une part de l'affectation de préfixes sur des bases sémantiques 'pures' (par adjonction d'un préfixe à la forme arabe), puis de l'intégration par analogie phonétique (identification de l'initiale à un préfixe), puis des cas de conjonction des deux facteurs, parmi lesquels les cas où l'on note une convergence entre modèle bantou et arabe; enfin nous signalons les cas où il y a des fluctuations entre deux classes, qui peut aller de pair avec une hésitation concernant le traitement morphologique (l'emprunt pouvant être non naturalisé ou naturalisé).

#### a) Affectation de préfixes sur des bases sémantiques

Ce facteur ressort dans les cas où les valeurs sémantiques des classes sont fortement définies, notamment les classes 1/2, référant à des humains, la classe 5/6 dans sa valeur augmentative, 7/8, diminutifs, et plus nettement encore quand le même emprunt lexical se retrouve dans d'autres classes.

La reconstruction de thèmes ne fait pas ici problème, dès lors en particulier que plusieurs classes (singulier et pluriel) sont possibles (voir plus loin).

- cl. 1/ 2 : termes référant à des humains: préf. m-/wa-

ar. *bazzāz*, marchand de tissu → **mbazazi** (wa-), commerçant

ar. *fud.ūlī*, curieux → **mfudhuli** (wa-), sans-gêne

ar. *kaid*, ruse → **mkaidi** (wa-), têtue

ar. *s'ait.ān*, Satan → **mshetwani**, diable

ar. *hindi*, indien → **muhindi** (wa-), indien

ar. *ʿarab*, arabe → **mwarabu** (waarabu), arabe

ar. *sawāh.il*, côte → **mswahili** (wa-), swahili

ar. *bābā*, père → **mbaba** (wa-), père

Ce mot peut aussi provenir d'un étymon bantou, \*-bāábá, père; n'oublions pas toutefois sa quasi-universalité

ii) cl. 3 / 4 : action, instrument, arbre : préf. m-/mi-

- ar. *siwāk*, baguette pour se nettoyer les dents → *msiwaki*, brosse à dent
- ar. *safar*, voyage → *msafara* (mi-), voyage
- ar. *wirāṭa* → *mwaratha* (cl. 3), héritage
- ar. → *mdarasini* (mi-), cannellier
- ar. *limūn* → *mlimu* (mi-), citronnier
- ar. *sākka*, percer ou *s`qqa*, fendre → *mshakiki* (mi-), brochette

iii) cl. 5 / 6 : augmentatif : préf. Ø'- :

- ar. *ibriq*, pichet → *birika* (ma-), bouilloire
- ar. *s`irka*, association → *shirka* (ma-), compagnie

Une intégration par analogie phonétique aurait placé ces mots plutôt en cl. 7 (cf. *ibiriti*, *shiwara*, ci-dessous).

La comparaison des sens, entre les sources arabes et les termes shingazidja, tendrait plutôt à indiquer un facteur sémantique.

La valeur augmentative se perçoit notamment dans les cas de double intégration:

- ar. *s`ari'a* →
- x affectation à base sémantique:
- sharia*, cl. 5, la loi (du Prophète)
- x analogie phonétique:
- sharia* (*mazaria*), cl. 7/8, loi/s

(autres exemples plus bas)

mais il faut comprendre que presque tous les mots d'emprunt à initiale occlusive, 'non-naturalisés' (cl. 9/10 sans marque préfixielle), peuvent alternativement être insérés dans la classe 5/6 avec une valeur augmentative, ce qui ne pose pas de problème: morphologique car cette classe classe a un préfixe Ø'- également:

- ar. *bah.r* →
- bahari* (-), mer, océan : accords 9/10, 'normaux'
- bahari* (ma-) : accords 5/6 :
- ebahari inu*, cette mer, cl. 9
- lebahari linu*, cette grande mer, cette mauvaise mer, cl. 5
- zebahari*, les mers, cl. 10
- emabahari*, toutes les mers, cl. 6

cl. 7 / 8 : diminutifs

ar. *sukkar*, sucre → *isukari* (zi-), bananes (sp.)

ar. *dār*, maison → *ilari* (zi-), mezzanine

On note l'intégration morphophonologique de l'initiale lexicale occlusive, qui passe à la continue après préfixe syllabique

ar. *naenas*, menthe → *inana* (zi-), menthe

ar. *kūfiya*, calotte → *ikofia* (zi-), dans l'expression *ma-ikofia*, jeune qui n'a pas accompli le grand mariage (jeune à la petite calotte)

La valeur diminutive est souvent confirmée, comme pour les augmentatifs ci-dessus, par l'existence, à partir de la même source arabe, d'un autre nominal *shingazidja*, éventuellement non naturalisé :

ar. *sukkar* →

*sukari* (-), sucre : 9/10

*isukari* (zi-), banane : 7/8

cl. 9 / 10 : divers

ar. *dārasini*, cannelle → *ndarasini* (-), cannelle

ar. *limūn*, citron → *ndimu* (-), citron

On remarque la transformation de l'initiale étymologique continue qui rentre dans le cadre des consonnes alternantes

ar. *zumūr*, instrument à vent → *ndzumari*, *ndumari* (-), hautbois

Nous reviendrons sur la variante /ndz/ ← /nd/ ([<sup>h</sup>d]) dans l'étude des sons

ar. *tasbih*, chapelet → *ntasbihi* (-), chapelet

ar. *jamal*, chameau → *ngamia* (-), chameau

Cette classe connaît de nombreux cas de variation sans implication sémantique, entre naturalisation (préf. |N-| marquant le nominal) et non naturalisation (préf. Ø-):

ar. *fā'ida* →

*faida* (-), profit : intégration passive, sans préf.

*nfaida* (-), profit : intégration active

cl. 11 : abstraction, activité : préf. u- :

ar. *bazzāz*, marchand de tissu → *ubazazi*, commerce

ar. *s'arab*, divertissement → *usherehi*, commémoration

ar. *tājir*, commerçant → *utadjiri*, richesse

ar. *tamaddun*, civilisation → *utamaduni*, civilisation

Intégration par analogie phonétique

Les cas les plus fréquents d'intégration active montrent une intégration par analogie phonétique, c'est-à-dire par l'identification de l'initiale étymologique à une marque de classe, qui peut être un préfixe exprimé (l'ensemble des



préfixes du modèle bantu sauf |Ø'-| de cl. 5), ou, dans le cas du préfixe 'occlusivisant' de cl. 5, |Ø'-|, une initiale lexicale occlusive.

Cette intégration apparaît nettement dans les oppositions de nombre, lorsque le pluriel ou le singulier sont formés par substitution du préfixe: cela justifie alors pleinement la reconstruction de thème.

Il arrive aussi qu'il n'y ait pas opposition de nombre, quand le terme est placé dans une classe à dominante singulier ou dans une classe à dominante pluriel, selon son sens.

Ici peuvent se ranger des cas de convergence (accidentelle) entre arabe et bantu, quand la syllabe initiale du mot arabe est un préfixe, qui se trouve correspondre au préfixe bantu.

Dans le cas de mots formés par analogie phonétique avec des préfixes de classes normalement pluriel, remarquons qu'il peut s'agir de pluriels arabes.

cl. 1/ 2 : humain : pas d'exemples ici, il y a toujours intégration sémantique également (voir ci-dessous)

cl. 3 : |mu-|

ar. *imtihān*, examen → **mtihani**, examen

ar. *mudda*, période → **muda** (mi-), période

ar. *mu'jiza*, miracle → **muujiuza**, merveille

*mu'jiza* : le préfixe arabe *mu* est interprété comme un préfixe bantu.

Notons l'allomorphe |mu-| du préf. de cl. 3, rare dans le fonds bantu (attesté seulement pour l'accord adjectival *mu*, noir)

ar. *mazh.*, plaisanterie → **mzaha** (mi-), faux-pas

Ces cas pourraient sans doute être aussi considérés comme des cas de conjonction du facteur analogique et du facteur sémantique.

cl. 4 : |mi-|

x avec reconstruction d'un singulier cl. 3 :

ar. *miqas.s.*, ciseaux → **miyasi**, cl. 4 → **mkasi** (mi-), ciseaux

x sans reconstruction d'un singulier :

ar. *mi'ād*, promesse → **miadi**, promesse(s)

Ce mot est accordé en classe 4, qui est une classe normalement pluriel; en fait, on ne peut déterminer si le sens est pluriel ou singulier, le mot pouvant référer en français tant à une promesse unitaire qu'à un ensemble d'engagements.

cl. 5 : différents préfixes :

préf. |Ø'-| devant occlusive (transformation en occlusive de l'initiale lexicale lorsque celle-ci est une continue membre des alternances consonantiques voir p 135-136)

Les nominaux à initiale occlusive sont souvent intégrés dans cette classe (ils commandent les accords de cette classe et forment leur pluriel en cl. 6); il n'y a pas en général alternance au pluriel, l'initiale reste le plus souvent inchangée:

- ar. *baḡil*, avare → *baḡili* (ma-), avare
- ar. *ballās.*, jarre → *balasi* (ma-), jarre
- ar. *barr*, terre ferme → *bara* (ma-), terre
- ar. *bāra*, pièce → *bwara* (ma-), pièce
- pers. *darīcha*, fenêtre → *dirīsha* (ma-), fenêtre

On trouve ici des cas d'intégration morphophonologique remarquable, lorsque l'initiale lexicale rentre dans l'alternance consonantique:

- ar. *dār*, maison → *dari* (malari), étage
- pers. *dām*, filet → *dema* (malema), filet
- ar. *lif*, fibre de coco → *difu* (malifu), enveloppe de la noix de coco

Intégrés dans cette classe, on trouve aussi des nominaux à initiale non occlusive (ce qui se rencontre aussi dans le fonds bantu):

- ar. *fuḡarā'*, pauvres → *fukara* (ma-), pauvre
- ar. *s.adaqa*, aumône → *swadaka* (ma-), aumône
- ar. *danb*, péché → *dhanbi* (ma-), péché
- ar. *d.amāna*, responsabilité → *dhwamana* (ma-), responsable

préf. |dji-| :

lorsque le nominal d'emprunt a pour initiale dji ou dj, phonème fréquent dans les emprunts à l'arabe (2 réalisé [j]), il y a le plus souvent hésitation entre intégration analogie phonétique à la classe 5 et intégration sans préfixe en cl. 9 /10 . Même dans les cas d'intégration par analogie phonétique où nous n'avons pas relevé d'autre traitement co-occurent, l'initiale n'est jamais assimilée à un préfixe au point de s'effacer au pluriel; toutefois, ici aussi, cela rejoint un fonctionnement présent dans le fonds bantu, avec les cas tels que *djindru* (madjindru), géant, ou *djinbo* (madjinbo), chanson, où le préfixe singulier est intégré au lexème:

ar. *jarrāfā*, filet → *djarifa* (ma-), filet  
 ar. *jirān*, voisins → *djirani* (ma-), voisin

cl. 6 : préf. *ma-*

La classe 6 est une classe pluriel le plus souvent.

Comme *ma* est un préfixe arabe, il peut y avoir convergence entre formation arabe et formation bantu et reconstruction d'un singulier en classe 5, ce qui montre bien l'intégration par analogie phonétique;

\* avec reconstruction d'un singulier :

ar. DHB > ar. *madhab*, école juridique → *madhibabi* → *dhibabi* (ma-), école juridique

ar. (M)L'K > ar. *malā'ika*, anges → *daika* (*malaika*), ange

On remarque l'adaptation morphophonologique remarquable; à partir du pluriel arabe *malā'ika*, le shingazidja a reconstruit un nominal à alternance initiale d/l

\* sans 'singulier' reconstruit :

ar. *mas.lah.a*, affaire → *masulaha*, arrangement

ar. *mawā'iz.*, exhortations → *mawaidha*, conseil

Ce dernier mot est issu d'un pluriel arabe

ar. εYS' > ar. *ma'is'a*, vie → Ng *maesha*, vie

ar. εS.Y > ar. *maes.iya*, désobéissance → Ng *maaswi*, blasphème

ar. εNY > ar. *maenan* → *maana*, raison

En dépit du préfixe pluriel, le sens de ces mots est plutôt singulier.

Il arrive que *ma-* perde sa valeur de préfixe, le mot passant alors en cl. 9/10 et formant un pluriel en cl. 10:

*maana*, raison : cl. 6

*maana* (-), raison : cl. 9/10

yo nde *maana*, c'est la raison ... : cl. 6

zo nde *maana*, ce sont les raisons ... : cl. 10

cl. 7 :

préf. |i-|

Un cas très remarquable :

ar. *kibrīt*, allumettes → *ibiriti* (zi-), allumettes

On note le passage de ar. *ki* à Ng *i-*, qui peut faire supposer une médiation du swahili *kibiriti*

préf. |shi-| :

nombreux termes ne forment pas de pluriel :

ar. *s'abahān*, cuivre → *shaba*, cuivre

ar. *s'adaha*, rendre confus → *shadaha*, trouble

avec pluriel en cl. 8 :

ar. *s'ūri*, demander conseil → *shiwara* (zi-), confidence

nombreux cas intégrés en cl. 5/6 'remotivée', souvent avec formation d'un pluriel avec double préfixe; voir ci-dessous)

cl. 8 : préf. |zi-|

le seul exemple dans le corpus, qui ne forme pas de singulier

ar. *zaman*, période → *zama*, époque

A noter, la chute de la consonne désinentielle arabe

cl. 9 & 10 :

le préfixe bantu le plus fréquent est la nasale homorganique qui prénasalise les consonnes initiales lexicales et qui peut avoir pour allomorphe *nyi-*; cette situation étant inconnue de l'arabe, il ne peut y avoir intégration par analogie phonétique; le seul cas où une analogie peut se produire est celui où le thème commence par une nasale arabe, /n/ ou /m/ :

ar. *nabiy*, prophète → *nabi* (-), prophète

ar. *nadr*, vœu → *nadhiri* (-), vœu

ar. *nafaqa*, dépense → *nafaka* (-), générosité

ar. *maenan*, → *maana* (-), raison

Il y a ici confusion entre naturalisation et non naturalisation: ces formes peuvent s'interpréter dans l'un et l'autre cadre, puisque la non naturalisation, caractérisée par l'absence de préfixe et l'intégration en classe 9/10 donnerait ici le même résultat: d'où les deux analyses structurelles pour un mot tel que *nafaka* :

Ø-nafaka! ou |Ø-nafaka! : intégration passive

|N-nafaka! : intégration active

cl. 11 : |u-|

ar. *wardiya*, rosaire → *uradi* (nyi-), chapelet

- ar. *was.iya*, testament → *waswia* (*nyaswia*), dernière parole d'un mourant  
 ar. *waqt*, temps → *wakati* (*nyakati*, *makati*), temps

On remarque les deux pluriels, en variante; cela correspond aux deux classes (10 et 6) avec lesquelles la classe 11 est en opposition de nombre

### c) Conjonction des facteurs sémantique et analogique

Dans certains cas, jouent à la fois le facteur sémantique et le facteur d'analogie phonétique; cela se vérifie pour certaines classes, à la motivation sémantique forte; on retrouve des cas de convergence (accidentelle) entre arabe et bantu:

x cl. 1 *lmu-l*, humain, correspondant au préfixe arabe de formation de nom verbal *mu*:

- ar. *ʿLM* > ar. *muʿallim*, enseignant → *mwalimu* (*walimu*), devin  
 ar. *DeW* > ar. *muddaʿi*, requérant → *mdai* (*wa-*), personne exigeante

La syllabe initiale arabe *mu*, correspondant au préfixe arabe *mu*, est interprétée comme le préfixe bantu *lmu-l* (*mw-* devant voyelles, *m-* devant consonne); cela se voit dans les pluriels shingazidja *walimu*, *wadai*, qui permettent de dégager les thèmes *l-(a)limu*, *l-dai*; les termes réfèrent aux humains (cette situation s'oppose aux cas, type *mwadhini*, où l'assimilation de l'initiale n'est pas, ou seulement partiellement, faite, ci-dessous, p 25)

cl. 11 : *lu-l*, abstraction (analogie avec l'initiale lexicale arabe)

- ar. *wud.ū'*, propreté → *udhu*, propreté  
 ar. *ʿumr*, vie → *umuri*, vie  
 ar. *wafāh*, décès → *wafati*, *ufati* (*nyifati*), mort  
 ar. *ʿuḍr*, excuse → *udhuri*, excuse

La syllabe initiale arabe est assimilée au préfixe; cette assimilation va de pair avec le fait que le ayn étymologique éventuel n'est pas représenté; on peut sans doute admettre que ces mots réfèrent à des abstractions.

### d) remotivation des classes

La 're-motivation' des classes renvoie au fait que la présence massive d'emprunts avec un sémantisme semblable dans une classe donnée entraîne

intégration systématique de tout nouvel emprunt ayant cette valeur dans cette classe; cela se produit surtout pour la classe 5/6.

Un grand nombre de termes référant à des activités humaines se trouvent parmi les emprunts intégrés dans cette classe (avec toutefois des hésitations concernant les accords régis): ces classes sont l'objet d'une re-motivation: tous les termes désignant des humains qui ne sont pas par nature en cl. 1/2 ont vocation à se trouver en 5/6, quelque soit leur initiale, même si celle-ci est analogue à un préfixe d'une autre classe. Toutefois, il convient de distinguer ici entre l'appartenance que l'on peut dire lexicale (préfixielle), et l'appartenance syntaxique: si ces nominaux appartiennent à la classe 5/6 sur le plan de l'apparence et de la formation du pluriel ( $\emptyset$ /ma-), les accords qu'ils régissent sont soit conformes à leur appartenance préfixielle (cl. 5/6), soit notionnels (cl. 1/2), les variations étant généralement significatives: l'accord en 1/2 est l'accord le plus fréquent, celui qui n'est pas connoté, que l'on peut considérer comme normal, alors que l'accord en 5/6 rejoignant les valeurs sémantiques de ces deux classes, de connotation (ainsi, l'accord en 5/6, surtout au singulier en 5, est en principe péjoratif - exemples ci-dessous).

\* initiale identique à préf. *!mu-* cl. 1 (et correspondant d'ailleurs au préfixe arabe *mu*):

ar. 'DN > ar. *mu'addin*, muezzin → **mwadhini** (ma-), muezzin

ar. FTY > ar. *muftin*, muphti → **mufti** (ma-), muphti

La formation du pluriel et les accords régis montrent la différence de traitement d'avec les cas où l'initiale étymologique (ici préfixielle) est assimilée à un préfixe bantu, type *mwalimu*:

*mwalimu* (walimu) : *!mu/wa-alimu* : assimilation de l'initiale

*mwadhini* (manwadhini) : *!Ø'/ma-mwadhini* : non-assimilation

*mwalimu* (wa-) commande uniquement les accords 1/2, *mwadhini* (ma-), les accords 1/2 notionnels ou 5/6 'grammaticaux' (connotés)

\* initiale identique à préf. *!i-* cl. 7 :

ar. *islām*, l'islam → **isilamu** (ma-), musulman

\* initiale identique à préf. *!shi-* cl. 7 :

ar. *s'abāb*, jeune → **shababi** (ma-), jeune

ar. *s'ahīd*, témoin → **shahidi** (mazahidi), témoins

- ar. s'arīf, prsonne honorable → sharīfu (**mazarīfu**), 'sharīf'  
 ar. s'ūrī, demander conseil → shawīri (**ma-**), conseil  
 ar. s'aik, personne âgée → shekhu (**mashekhu, mazekhu**), 'cheikh'

On note les doubles préfixes au pluriel (**maz..**), constitués du préfixe lma-1 de classe 6 suivi du préfixe de cl. 8; ce comportement montre que dans de tels mots, malgré l'accord en cl. 5 ou 1, /sh/ est senti comme porteur de la valeur singulier qui est celle du préfixe de classe 7 auquel il est analogue: il ne peut donc pas se maintenir au pluriel.

On peut aussi en déduire que l'accord de ces mots en classe 5/6 est récent (quoique général d'après nos relevés), et qu'il s'est substitué à un accord analogique 7/8; cela serait un argument dans le sens de la re-motivation de la classe 5/6.

× initiale identique à préf. lu-1 cl. 11 :

- ar. waliy, saint → wali (**ma-**), saint homme  
 ar. wat.aniya, nationalisme → watwani (**ma-**), patriote

× initiale identique à préf. lN-1 cl. 9 (devant nasales) :

- ar. nākudā, capitaine → nahuda (**ma-**), capitaine

En énoncé, ces nominaux peuvent commander les accords de 5/6 (grammaticaux et connotés) ou de 1/2 (notionnels):

- emwadhini oyi**, ce muezzin : accord cl. 1 normal  
**lemwadhini linu**, ce mauvais muezzin : accord cl. 5 péjoratif  
**omashababi waili wanu**, ces deux jeunes : accord cl. 2 normal  
**emashababi maili yanu**, ces deux jeunes vauriens : accord cl. 6 péjoratif  
 A vrai dire, le choix des accords dépend aussi des parties du discours concernées (augment ou pré-préfixes, démonstratifs, préfixes verbaux, etc)

**Remarque : fluctuations significatives ou non significatives**

Des fluctuations, significatives ou non, se manifestent fréquemment entre intégration 'minimale' (sans préfixe ou préf. lØ-1 cl. 9) et intégration analogique, le plus souvent en cl. 5/6, lorsque l'initiale lexicale est occlusive. Dans de nombreux cas, ces fluctuations, entre intégration minimale et intégration en cl. 5/6, sont significatives, l'intégration en cl. 5/6 ayant valeur augmentative; mais il n'est pas possible de prévoir pour quel mot la variation est significative: en fait, les locuteurs réagissent par rapport à la norme ou à l'habitude pour chaque mot, c'est-à-dire que, si un mot est inséré

normalement en cl. 9/10, sa mise en cl. 5/6 sera significative, le mot pouvant être aussi indifféremment dans l'une ou l'autre classe (mais un mot normalement en cl. 5/6 ne pourra vraisemblablement pas être mis aussi en 9/10 sans préfixe). En outre, il est évident qu'interviennent ici des variations individuelles ou régionales.

L'insertion en 9/10 sans préfixe apparaît alors comme un premier stade d'intégration.

Dans le lexique, nous signalons ce qui nous est apparu comme l'appartenance ou les appartenances normales de classe, et, le cas échéant, lorsqu'elle nous est apparue fréquente ou intéressante pour une quelconque raison, telle ou telle annotation apportée par un changement de classe (bien entendu, on ne saurait induire de l'absence de mention dans le lexique l'inexistence ou l'impossibilité de l'intégration en telle classe qui ne figure pas).

Nous donnons ici quelques exemples de fluctuation non significative, et renvoyons ci-dessus pour les fluctuations significatives (rappelons que, outre la formation du pluriel, la classe d'intégration est prouvée par les accords de genre en énoncé).

Les fluctuations non significatives (en tête, la forme qui nous est apparue la plus courante; l'appartenance de classe est donnée par la forme plurielle):

- ar. *ibriq*, pichet → **birika** (ma-), bouilloire cl. 5/6, ou **birika** (-), bouilloire cl. 9/10
- ar. *madrassa*, école → **madarasa** (-), école cl. 9/10 ou **madarasa** (cl. 6), école
- ar. *jihāz*, boudre → **djahazi** (ma-), boudre cl. 9/10 ou **djahazi** (-), boudre
- ar. *jawāb*, réponse → **djawabu** (ma-), réponse ou **djawabu** (-), réponse
- ar. *maenan* → **maana** (-), sens ou **maana** (cl. 6), sens (voir aussi supra p. 206)
- ar. *jais* "armée" → **djeshi** (ma-), armée ou **djeshi** (-), armée
- ar. *fā'ida*, profit → **faida** (-), profit ou **nfaida** (-), profit

Dans ce dernier cas, ainsi que nous l'avons signalé, la classe d'accord est la même, seule la façon de l'atteindre diffère: l'un des mots est un emprunt non naturalisé (sans préfixe adjoint), l'autre a un préfixe exprimé.

x fluctuations significatives (rappel)

bahari (-), mer vs bahari (ma-), 'grande' mer  
(voir ci-dessus pour des exemples en contexte)

### 1-2) Des thèmes, des thèmes

Nous avons dit que les modalités d'intégration morphologique décrites ci-dessus définissaient un fonctionnement analogue à celui des thèmes bantou; ceci se prouve par la comparaison entre divers nominaux montrant une même base lexicale mais différant par le préfixe, qui contribue alors au sens de chaque terme, le cas minimum étant l'expression du nombre; souvent est impliquée l'intégration à base sémantique, particulièrement les classes connotatives (diminutif, augmentatif) et la classe 11 (noms abstraits).

Bien entendu, il n'y a pas ici liberté combinatoire totale, pas plus que dans le fonds bantou, mais il est possible que la latitude soit en réalité plus grande pour les emprunts; en effet, dans la mesure où il s'agit d'emprunts dont l'intégration en classe est susceptible de fluctuations, la liaison préfixe-thème est sans doute moins forte.

La reconstruction des thèmes est opérée par soustraction du préfixe.

Toutefois, nous considérons qu'il y a bien 'thèmes', par systématisation et généralisation, même lorsque la formation nominale reste isolée (pas d'autres formations apparentées), dès lors qu'un nominal est sujet d'une intégration morphologique, qu'il soit muni d'un préfixe (intégration sémantique), ou que son initiale étymologique soit assimilée à un préfixe (intégration analogique, y compris cl. 5 à préfixe  $\emptyset$ -1).

Nous donnons quelques exemples, reprenant des mots donnés ci-dessus.

traduction des étymologies arabes, qui figure ci-dessus, n'est pas reprise.

avec deux formations nominales à préfixe (+ natif), dépassant l'opposition nombre :

ar. *bazzāz* →

- mbazazi*, commerçant : cl. 1
- wabazazi*, commerçants : cl. 2
- ubazazi*, commerce : cl. 11

D'où la construction du thème [-bazazi], (notion de commerce)

ar. *kaid* →

- mkaidi*, têtu : cl. 1
- wakaidi*, têtus : cl. 2
- ukaidi*, obstination : cl. 11

D'où la construction du thème [-kaidi], (notion d'obstination)

ar. *mu'alim* →

- mwalimu*, maître : cl. 1
- walimu*, maîtres : cl. 2
- uwalimu*, maîtrise : cl. 11

D'où la construction du thème [-(<w>alimu)], (notion de maîtrise)

ar. *dār* →

- dari*, étage : cl. 5
- malari*, étages : cl. 6
- ilari*, mezzanine : cl. 7
- zilari*, mezzanines : cl. 8

D'où la construction du thème [-lari], (notion d'étage), à alternance initiale

ar. *limūn* →

- mlimu*, citronnier : cl. 3
- milimu*, citronniers : cl. 4
- ndimu*, citron : cl. 9
- ndimu*, citrons : cl. 10

D'où la construction du thème [-limu], (notion de citron), à alternance initiale

ar. *dāras.īni* →

- ndarasini*, cannelle : cl. 9
- ndarasini*, cannelle : cl. 10
- mdarasini*, cannellier : cl. 3
- midarasini*, cannelliers : cl. 4

D'où la construction du thème [-darasini], (notion de cannelle)

ar. *tājir* →

- tadjiri*, personne riche : cl. 5 & 1
- matadjiri*, personnes riches : cl. 6 & 2
- utadjiri*, richesse : cl. 11

D'où la construction du thème [-tadjiri], (notion de richesse)

ar. *baḥil* →

- bahili*, avare : cl. 5 & 1
- mabahili*, avares : cl. 6 & 2
- ubahili*, avarice : cl. 11

D'où la construction du thème [-bahili], (notion d'avarice)

ar. *ṣ'abāb* →

- shababi*, jeune : cl. 5 & 1
- mashababi*, jeunes : cl. 6 & 2
- ushababi*, jeunesse : cl. 11

D'où la construction du thème [-shababil, (notion de jeunesse)

ar. waliy →

walii, saint homme : cl. 5 & 1

walii, saints hommes : cl. 6 & 2

uwalii, sainteté (rigueur morale) : cl. 11

D'où la construction du thème [-waliil, (notion de sainteté (rigueur morale)

ii) avec un seul nominal +natif, dans deux classes en opposition de nombre :

ar.

msiwaki, brosse à dent : cl. 3

misiwaki, brosses à dent : cl. 4

D'où la construction du thème [-siwakil, (notion de brosse à dent)

ar. →

mshakiki, brochette : cl. 3

mishakiki, brochettes : cl. 4

D'où la construction du thème [-shakikil, (notion de brochette)

ar. →

shirka, compagnie : cl. 5

mashirka, compagnies : cl. 6

D'où la construction du thème [-shirka], (notion de compagnie)

pers. daricha

darisha, fenêtre : cl. 5

madarisha, fenêtres : cl. 6

D'où la construction du thème [-dirishal, (notion de fenêtre)

ar. jarrāfā →

djarifa, filet : cl. 5

madjarifa, filets : cl. 6

D'où la construction du thème [-djarifa], (notion de filet)

ar. s'ūrī →

shiwara, confiance : cl. 7

ziwara, confidences : cl. 8

D'où la construction du thème [-waral, (notion de confiance)

ar. jamal →

ngamia, chameau : cl. 9

ngamia, chameaux : cl. 10

D'où la construction du thème [-gamial, (notion de chameau)

ar. wardiya →

uradi, chapelet : cl. 11

nyiradi, chapelet : cl. 10

D'où la construction du thème [-uradi], (notion de chapelet)

Remarques : double préfixe :

Dans les cas de double préfixe (pl. cl. 6 de mots à initiale singulier analogue au préf. de cl. 7 /sh/, qui entraîne la forme du préfixe de cl. 8 /z/), on est amené à postuler deux thèmes:

ar. s'aik →

shekhu, 'cheikh' : cl. 5 & 1

mashekhu ou mazekhu, 'cheikhs' : cl. 6 & 2

On postule deux thèmes, [-shekhu], d'une part, pour shekhu et mashekhu, et [-zekhu], à partir de mazekhu

## II LES EMPRUNTS

Dans le fonds bantu, nous avons vu que les constituants syntaxiques analysaient en monèmes, lexicaux (racines et thèmes), et grammaticaux (obligatoires, actualisateurs affixés, facultatifs, extensions).

Nous avons vu que, pour les termes invariables, il y avait coïncidence entre catégorie de langue et catégorie de discours, ces constituants n'étant pas analysables.

Nous examinons maintenant la manière dont les emprunts d'origine orientale, essentiellement arabe, s'intègrent morphologiquement au grand-comorien et quelle influence leur présence a pu avoir sur les systèmes morphologique et lexicologique.

On constate d'emblée une distinction entre deux états d'intégration, qui recoupe les deux plans de l'analyse morphologique:

- \* un état de non-naturalisation, ou d'intégration minimale (ou intégration passive) où les emprunts s'intègrent directement aux catégories de discours ou constituants syntaxiques,
- \* un état de naturalisation ou d'intégration achevée (ou intégration active), où les emprunts admettent les mécanismes de formation bantu: on peut alors reconstruire des lexèmes (racines et/ou thèmes), sur le modèle des lexèmes de fonds bantu.

Nous définissons un critère ( $\pm$  natif) pour qualifier chacun de ces états.

Les deux classes qui ressortent de cette distinction sont évidemment des polarités: à côté de fonctionnements nettement marqués qui situent les termes d'un côté ou de l'autre, on relève un fonctionnement que l'on peut considérer

ar. s'ahid →

shahidi, témoin : cl. 5 & 1

mazahidi, témoins : cl. 6 & 2

Deux thèmes : |-shahidi| et |-zahidi|, notion de témoignage|

(1) avec un nominal n'apparaissant que dans une seule classe + natif: la reconstruction est plus hypothétique

ar. ma'is'a →

maesha, vie : cl. 6

D'où la construction du thème |-esha|, notion de vie|

ar. mi'ad, →

miadi, promesse(s) : cl. 4

D'où la construction du thème |-adi|, notion de promesse|

ar. mawā'iz. →

mawaidha, conseils : cl. 6

D'où la construction du thème |-waidha|, notion de promesse|

ar. s'adaha →

shadaha, trouble : cl. 7

D'où la construction du thème |-adaha|, notion de trouble|

ar. 'umr →

umuri, vie : cl. 11

D'où la construction du thème |-muri|, notion de vie|

Remarque : cas des emprunts non naturalisés de même base lexicale

Il arrive qu'une forme intégrée, permettant la reconstruction d'un thème, co-existe avec l'emprunt non intégré (nominal -natif) :

ar. sukkar →

sukari, sucre : cl. 9 et sukari, sucre : cl. 10

isukari, banane (sp.) : cl. 7

isukari, bananes (sp.) : cl. 8

Le thème est reconstruit sur la base des nominaux intégrés : |-sukari|, notion de sucre|

ar. kufiya →

kofia, calotte : cl. 9 et kofia, calottes : cl. 10

ikofia, petite calotte (ci-dessus) : cl. 7

ikofia, petites calottes (ci-dessus) : cl. 8

Le thème est reconstruit sur la base des nominaux intégrés : |-kofia|, notion de calotte|

Si l'on se situe dans la perspective où cette intégration minimale est vue comme une intégration avec préfixe |Ø-|, la même analyse structurelle peut s'appliquer, ce qui permet de faire apparaître les thèmes:

sukari : |Ø-sukari|

kofia : |Ø-kofia|

etc

## 2) Intégration par assimilation à des racines

Dans le fonds bantu, les racines sont caractérisées par leur bi-valence verbo-nominale: au moyen de différents morphèmes obligatoires (actualisateurs), elles donnent naissance tant à des nominaux qu'à des verbaux; elles ont aussi la propriété de pouvoir recevoir des morphèmes de dérivation appelés extensions, qui constituent avec elles des racines étendues aux mêmes propriétés morphologiques que les racines simples ou primaires.

Or, ces extensions, du moins certaines d'entre elles, peuvent aussi s'appliquer aux emprunts arabes, de façon à constituer un élément lexical au fonctionnement analogue à celui d'une racine bantu, c'est-à-dire bi-valent verbo-nominal, apte à recevoir les divers actualisateurs 'bantu', en particulier les suffixes, dont nous avons vu qu'ils étaient incompatibles avec les emprunts non naturalisés.

Il s'agit donc d'un cas patent d'intégration morphologique, puisque, du seul point de vue morphologique, on ne peut pas distinguer ici les termes indigènes réputés bantu de ceux d'origine étrangère.

Rappelons que les extensions ont toutes des étymons bantu.

On peut ainsi, à partir des formes étendues, en procédant par suppression d'éléments, reconstruire l'équivalent de racines. Ces 'racines' (une séquence identique à ce qui tient lieu de racine) se retrouvent, le plus souvent, dans des termes d'emprunt soit non naturalisés, soit naturalisés (nominaux); nous considérons ces termes, qui n'ont pas le statut de lexème, puisque ce sont des éléments de discours, comme les 'supports lexicaux' de ces formes étendues.

soient les constituants syntaxiques

-badili, changer < ar. *badala*

-badilisha, changer (tr.)

mbadilishadji (wa-), réformateur

Le rapport à la fois formel et sémantique entre ces trois constituants n'est pas douteux.

On identifie dans les deux derniers des affixes bantu : **-a** (verbal) & **m/wa-adj** (agent), alors que le premier est un verbe non naturalisé typique (finale **-i** conditionnée), inanalysable.

Si l'on décompose les deux termes analysables, on fait apparaître une séquence **'-badilish-'**:

**-badilisha** : **-badilish-a**  
**mbadilishadji (wa-)** : **m/wa-badilish-adj**

Si l'on admet de retrouver dans cette séquence l'allomorphe **-ish-** de l'extension causative, on obtient la séquence **-badil-**:

**-badilish-** < **-badil-ish-**

Cette séquence **-badil-** correspond à la partie proprement étymologique du verbe non naturalisé **-badili**, dont la voyelle finale est conditionnée.

A partir de là, on pose que **-badili** est bien le support lexical des formes étendues causatives **-badilisha**, **mbadilishadji**, etc.

**Nota**

Dans cette présentation, ainsi que dans les exemples d'une façon générale, le suffixe **-a** des verbes est mis pour les deux suffixes **-a** et **-e**, à fonction syntaxique; si le verbe admet le suffixe **-a** à la forme de citation, cela signifie qu'il admet tant bien le suffixe **-e**, c'est-à-dire, qu'il s'intègre dans les processus réguliers.

Les extensions concernées par ce processus sont les deux classes d'extensions définies par leur valeur totalement ou partiellement syntaxique et/ou leur caractère productif: sont exclues les extensions purement sémantiques, non productives: on trouvera donc ici

- les extensions sémantico-syntaxiques (applicatif, intensif, causatif, réciproque; toutefois l'intensif est excessivement rare);
- les extensions syntaxiques ou de diathèse, passif et moyen-neutre.

Comme dans le fonds bantu, ces extensions peuvent se combiner et l'on trouvera des formations à plus d'un degré d'extensions (parfois la combinaison est obligatoire).

Les extensions sont employées dans leur valeur normale, telle qu'elle a été vue dans l'étude des mécanismes bantu (dans le cas de l'applicatif, voir toutefois remarque ci-dessous). Rappelons que les extensions sémantico-syntaxiques ont

pour objet essentiellement de permettre une complémentation supplémentaire, alors que les extensions syntaxiques indiquent l'orientation du procès. Conséquence de ces valeurs, les termes issus de ces racines étendues sont, ici, aussi, en majorité mais sans exclusive, des verbes.

Les extensions purement sémantiques (et non productives) sont exclues des emprunts; cela est lié bien sûr à leur caractère non productif, mais aussi à leur valeur même: on peut admettre qu'il soit plus difficile d'intervenir sur le sémantisme d'emprunts que sur leur valence; si le shingazidja a besoin du 'contraire' d'un terme emprunté (pour faire allusion à l'extension sémantique la plus répandue, le réversif), la langue pourra recourir éventuellement à l'emprunt du terme exprimant ce 'contraire', plutôt qu'à sa fabrication au moyen d'un procédé dérivationnel.

Tous les termes d'emprunt dont nous avons déjà vu l'intégration morphologique, quel que soit leur niveau d'intégration et leur catégorie grammaticale (invariables & variables, verbes ou noms), sont concernés ici et peuvent servir de support lexical - la seule restriction concerne la compatibilité entre le sens ou la valeur de l'extension et celle du lexème d'emprunt; toutefois, il est de fait que le support lexical est fréquemment un verbe.

Par analogie, on extrait des "racines" même dans le cas où le support lexical n'est pas attesté isolément (sur le modèle des racines allongées bantu, décomposables sous certaines conditions).

Ce fonctionnement entraîne la constitution de familles dérivationnelles, où les termes sont liés entre eux, d'une part à travers les processus combinatoires bantu pour les termes issus des formes étendues, et d'autre part, de fait, par le rapport formel et sémantique à leurs supports lexicaux.

Nous rappelons les différentes extensions concernées, ainsi que leurs étymons, et examinons la manière dont elles s'appliquent aux emprunts; ni la catégorie

grammaticale du terme jouant le rôle de support lexical à la forme étendue, ni la forme (le timbre de la voyelle finale lexicale), ne paraissent intervenir dans les modalités 'techniques' d'application de l'extension.

Nous distinguons les cas où le support lexical existe, qui sont les plus fréquents, des cas où le support lexical est absent et où l'on est amené à la reconstruction analogique d'une racine, qui ne peut s'appuyer sur un terme attesté.

#### Extensions (rappel)

Nous ne donnons ici que les allomorphes attestés dans les racines tirées des emprunts; rappelons que les verbes d'emprunt ont presque toujours l'allomorphe -ish- du causatif, voir p 157 et suiv.)

Sémantico-syntaxiques :

applicatif : -i-, (-li-) < \*id

intensif : -ili- < \*idid

causatif : -ish-, (-z-, -s-) < \*ik, \*i'c

réciproque : -an- < \*an

Syntaxiques :

passif : -iw-, -w- < \*u, \*ibu

moyen-neutre : -ih- < \*ik

Le conditionnement de la voyelle des extensions explique qu'on ne trouve, sauf cas exceptionnels, parmi les allomorphes des extensions, que les voyelles qui s'identifient aux voyelles arabes, /i/ et /a/, à l'exclusion notamment de /e/.

#### Existence d'un support lexical attesté

Le support lexical est un terme d'emprunt, qui peut être non naturalisé (verbe, nom sans préfixe de classe, invariable), ou naturalisé (nom pleinement intégré dans le système de classe): dans tous les cas, la voyelle finale est lexicale, soit étymologique (/a, i, u/), soit conditionnée phoniquement /i, u/.

La post-position d'une extension permet la reconstruction d'une racine, généralement par chute de la voyelle finale lexicale (on a alors le plus souvent une structure A finale consonantique) devant la voyelle de l'extension,

la nature de la voyelle lexicale n'intervenant pas. Il arrive aussi, rarement, que la voyelle finale lexicale se maintienne, et que l'extension soit de forme -1....

Il peut se trouver deux supports lexicaux, un verbal et un nominal, également plausibles.

A la racine étendue, viennent s'affixer les différents actualisateurs qui caractérisent les constituants syntaxiques issus de racines: préfixes et suffixes, dont la combinatoire donne tant des verbes que des nominaux (noms et adjectifs).

Comme pour les racines étendues de fonds bantu, les constituants syntaxiques résultants sont le plus souvent verbaux, du fait de la valeur notamment syntaxique des extensions.

La présentation est ordonnée selon la complexité croissante (un degré, deux degrés), selon les extensions, et selon la catégorie grammaticale du support lexical.

#### *a-1) formations à un degré d'extension*

##### *1) applic. -i- (-li-) :*

x support lexical verbe :

-kusudi, avoir pour but < ar. *qas.ada*

→ \*-kusud- + -i- → l-kusudi-l, notion d'avoir pour but

→ -kusudia, avoir pour but

kusudio (ma-), but

-dhibití, prendre en main < ar. *d.abat.a*

→ \*-dhibit- + -i- → l-dhibití-l, notion de prendre en main

→ -dhibitia, prendre en main

udhibitifu (cl. 11), preuve

-baki, rester < ar. *baqiya*

→ \*-bak- + -i- → l-baki-l, notion de rester

→ -bakia, rester

-abudu, adorer < ar. *ʿabada*

→ \*-abud- + -i- → l-abudi-l, notion d'adorer

→ -abudia, adorer

-arifu, faire savoir < ar. *ʿarrafa*

→ \*-arif- + -i- → l-arifi-l, notion de faire savoir

→ -arifia, faire savoir

support lexical nominal :

**azima** (-), projet < ar. *cazima*

→ \*-azim- + -i- → |-azimi-|, [notion de faire des plans]

→ -azimia, projeter

**azimio** (ma-), projet

**hotuba** (-), discours < ar. *ḥut.ba* et -hotubu, discourir < ar. *ḥat.aba*

→ \*-hotub- + -i- → |-hotubi-|, [notion de faire un discours]

→ -hotubia, discourir

Parfois, la voyelle finale est maintenue, et l'on trouve l'allomorphe -li- de l'applicatif:

**-djibu**, répondre < ar. *'ajāba*

→ \*-djibu- + -li- → |-djibuli-|, [notion de répondre à/pour]

→ -djibulia, répondre à ou pour qqun

L'applicatif est caractérisé par une voyelle de même timbre que la voyelle finale, conditionnée (après consonnes autres que labiales) : cela a des conséquences sur son emploi: souvent, la finale -i des verbes d'emprunt non naturalisés est assimilée à un applicatif, sous la pression du modèle dominant: cela permet d'intégrer ces verbes dans les fonctionnements normaux (suffixe); cela est facilité par la valeur souvent peu ressentie de l'applicatif; cela implique la synonymie entre forme non naturalisée et forme naturalisée, dans des cas tels que -kusudi et -kusudia, etc, l'applicatif devenant une 'coquille vide' (un simple appendice formel ayant perdu toute valeur sémantique); l'on trouvera d'ailleurs, fréquemment sinon systématiquement, une racine étendue applicative dans les combinaisons d'extensions et dans le cas où le support lexical n'est pas attesté.

ii) intensif : -ili-

Le seul exemple dans le corpus, à partir d'un verbe à voyelle -i :

**tahani**, moudre

→ \*-tahan- + -ili- → |-tahanili-|, [notion de moudre]

→ -tahanilia, moudre (se dit la meule)

iii) causatif : -ish- (-iz-, -is-) :

support lexical verbe :

**badili**, changer < ar. *badala*

→ \*-badil- + -ish- → |-badilish-|, [notion de changer]

→ -badilisha, changer (tr.)

**mbadilishadji** (wa-), réformateur

**baki**, rester < *passim*

→ \*-bak- + -ish- → |-bakish-|, [notion de conserver]

→ -bakisha, conserver

**dhwahiri**, apparaître < ar. *z.ahara*

→ \*-dhwahir- + -ish- → |-dhwahirish-|, [notion d'annoncer]

→ -dhwahirisha, annoncer

-**dumu**, durer < ar. *dāma*

→ \*-dum- + -ish- → |-dumish-|, notion de durer|

→ -**dumisha**, durer

x support lexical nominal :

**furaha** (-), joie < ar. *farah.* & -*furahi*, être content < ar. *farih.a*

→ \*-furah- + -ish- → |-furahish-|, notion de faire plaisir|

→ -**furahisha**, faire plaisir

**khatwari** (-), danger < ar.

→ \*-khatwar- + -ish- → |-khatwarish-|, notion de prévenir|

→ -**khatwarisha**, prévenir

**alama** (-), signe < ar. *alāma*

→ \*-alam- + -ish- → |-alamish-|, notion d'enseigner|

→ -**alimisha**, enseigner

**djimla** (-), somme < ar. *jumla*

→ \*-djiml- + -is- → |-djimlis-|, notion de totaliser|

→ -**djimlisa**, totaliser

**mdjimliso** (mi-), somme

**hakika** (-), vérité < ar. *h.aqīqa*

→ \*-hakik- + -ish- → |-hakikish-|, notion de prouver|

→ -**hakikisha**, prouver

x support lexical invariable :

**karibu**, près < ar. *qarīb*

→ \*-karib- + -ish/is- → |-karibish-| ou |-karibis-|, notion de recevoir|

→ -**karibisha**, -**karibisa**, recevoir

**daima**, toujours < ar. *dā'iman*

→ \*-daim- + -ish/iz- → |-daimish-|, |-daimiz-|, notion de durer|

→ -**daimisha**, -**daimiza**, durer

Même sens que -**dumisha**, ci-dessus

La répartition des allomorphes du causatif est liée à l'étymologie: les emprunts se construisent en règle générale avec -ish-, qui leur est spécifique; mais on trouve aussi, rarement et de façon lexicalisée (non prévisible), les allomorphes -iz- et -is- avec certains termes d'emprunt (exemples ci-dessus).

#### iv) réciproque -an- :

Le réciproque est toujours employé après applicatif (voir plus bas, combinaison)

#### v) passif : -iw-, -lw- :

x support lexical verbe :

-**shitaki**, porter plainte

→ \*-shitaki- + -lw- → |-shitakilw-|, notion d'être attaqué en justice|

→ -**shitakilwa**, être attaqué en justice

-**swifu**, parler de qqun

→ \*-swif- + -iw- → |-swifiw-|, notion d'avoir entendu parler de|

→ -**swifiwa**, avoir entendu parler de

-**dhulumu**, voler < ar. *z.alama*

→ \*-dhulum- + -iw- → |-dhulumiw-|, notion d'être exploité|

→ -**dhulumiwa**, être exploité

**laumu**, blâmer

→ \*-laumu- + -lw- → |-laumulw-|, notion d'être blâmé

→ -laumulwa, être blâmé

support lexical nominal :

**lakini** (-), prière des morts & -lakini, lire la prière des morts

→ \*-lakin- + -iw- → |-lakiniw-|, notion de se faire sermonner

→ -lakiniwa, se faire sermonner

La répartition des formes -iw- et -lw- ne paraît pas pertinente; ces allomorphes peuvent éventuellement être en variante.

(vi) *moyen-neutre* : -ih- :

support lexical verbe :

**badili**, changer < *passim*

→ \*-badil- + -ih- → |-badilih-|, notion d'être changé

→ -badiliha, être changé

→ badiliho (**ma-**), changement

**bariki**, bénir < ar. *bāraka*

→ \*-barik- + -ih- → |-barih-|, notion d'être prospère

→ -barikiha, être prospère

**dhulumu**, voler < *passim*

→ \*-dhulum- + -ih- → |-dhulumih-|, notion d'être exploité

→ -dhulumiha, être exploité

support lexical nominal :

**shuhuli** (**ma-**), cérémonie

→ \*-shuhul- + -ih- → |-shuhulih-|, notion d'être affairé

→ -shuhuliha, être affairé

**sta'arabu** (cl. 11), culture

→ \*-sta'arab- + -ih- → |-sta'arabih-|, notion d'être cultivé

→ -staarabiha, être civilisé

**taabu** (-), difficulté

→ \*-taab- + -ih- → |-taabih-|, notion d'avoir des problèmes

→ -taabiha, avoir des problèmes

**djimla** (-), somme < *passim*

→ \*-djiml- + -ih- → |-djimlih-|, notion d'être totalisé

→ -djimliha, être totalisé

**tarehe** (-), date

→ \*-tareh- + -ih- → |-tareh-|, notion de faire date

→ -tareha, faire date

On remarque la confusion entre la syllabe finale de la forme reconstruite \*-tareh- et l'extension

## 2) formations à deux degrés d'extensions

Les extensions peuvent se combiner, selon les mêmes principes de combinatoire

que pour le fonds bantou; nous en voyons quelques exemples.

Au moins une extension, le réciproque, apparaît toujours en combinaison, après applicatif.

*i) combinaison facultative*

-dhwaifu, pauvre < ar. *d.a'if*

→ \*-dhwaif- + -ish- → |-dhwaifish-|, notion d'appauvrir qqun|

-dhwaifisha, appauvrir

→ |-dhwaifish-| + -ih- → |-|dhwaifishih-|, notion de s'appauvrir|

→ -dhwaifishiba, s'appauvrir

*ii) combinaison obligatoire : réciproque*

Le réciproque est toujours associé à l'applicatif (du moins dans notre analyse où nous ramenons systématiquement la voyelle /i/, présente dans la forme étendue, à l'applicatif)

-fariki, partir < ar. *fāraqa*

→ \*-farik- + -i- + -an- → |-farikian-|, notion de se séparer|

→ -farikiana, se séparer

hitlafi, divergence < ar. *iktilāf*

→ \*-hitlaf- + -i- + -an- → |-hitlafian-|, notion de se séparer|

→ -hitlafiana, se séparer de

-husudi, envier < ar. *h.asada*

→ \*-husud- + -i- + -an- → |-husudian-|, notion de se jalouser|

→ -husudiana, se jalouser

miadi (cl. 4), promesses ou -wadi, arriver

→ \*-ad- + -i- + -an- → |-adian-|, notion de s'engager|

→ -adiana, s'engager

*b) absence de support lexical attesté*

Dans de nombreux cas, la forme étendue est attestée en l'absence de tout support lexical isolable en shingazidja (c'est-à-dire, en l'absence d'un emprunt, non naturalisé ou naturalisé, auquel la forme étendue puisse être rapportée); par analogie avec les décompositions précédentes, on est amené à postuler systématiquement une forme 'primaire' découlant d'un support lexical hypothétique, même si basée sur une seule attestation: en effet, contrairement au fonds bantu, il n'est pas loisible ici de supposer l'existence de racines complexes (dans le fonds bantu, nous ne décomposons une racine 'allongée' que si deux éléments concordants au moins le permettent, par ex., deux formes

endues); pour les emprunts, on aura ainsi des reconstructions basées, par exemple, sur le seul applicatif.

Nous donnons un exemple du processus de construction du support lexical, avant de présenter des exemples interprétés:

soient les constituants

-*hudhuria*, assister, -*hudhurisha*, inviter < ar. *h.ad.ara*, et  
*mhudhurio* (mi-), rassemblement, < ar. *h.ud.ür*

Il est clair que ces formations sont des constructions applicatives et causative à partir d'une forme \*-*hudhur-*; mais nous n'avons pas relevé de support lexical *ad hoc*, qui serait \*-*hudhuru*, ou \**hudhuru*; la reconstruction de la racine étendue se passe alors du support lexical:

\*-*hudhur-* + -i- → |-*hudhuri-*|, notion d'assister à |  
 → -*hudhuria* & *mhudhurio* (mi-)  
 -*hudhurisha*

Dans le lexique, nous reconstruisons un support lexical plausible, plutôt verbal puisque les verbes sont plus souvent à l'origine des formes étendues que les autres catégories du discours); c'est ce support qui est référé à l'étymologie arabe, puisque les dérivations par extension sont par définition bantu; ce support est noté selon la procédure suivie pour les racines bantu reconstruites, en maigre, sans glose, et précédé du signe +: dans l'exemple, *mhudhuru* < ar. *h.ud.ür*.

Cet support permet de montrer la construction des formes étendues; il peut demander l'entrée lexicale.

Dans le cas pris en exemple, on ne se laissera pas troubler par l'existence d'un emprunt non naturalisé référé à la même racine arabe (mais à un autre terme issu de cette racine):

*hadhwara* (-), cérémonie < ar. *h.ad.ara*

En effet, ce mot ne peut fournir de support lexical satisfaisant aux formes étendues: *quid*, alors, du changement vocalique ?

La relation lexicologique évidente entre ces termes est en réalité médiatisée par l'arabe; ce point fait l'objet du paragraphe suivant (aspect lexicologique).

Encore une fois, nous sommes très conscient que l'insuffisance du corpus est seule responsable, vraisemblablement, de la plupart des cas de ce type; il nous semble que le traitement que nous en proposons pallie pour une

part cette faiblesse, par la reconstruction des formes manquantes, dont on peut supposer qu'un approfondissement de l'enquête vérifiera l'hypothèse.

1) *co-existence de plusieurs formes étendues*

- +ghafili (non attesté) < ar. *g.afala*  
 → +ghafil- + -i- → |ghafil-i|, (notion d'énerver)  
 → **ughafilifu** (cl. 11), énervement  
 +ghafil- + -ish- → |ghafilish-i|, (notion d'énerver qqun):  
 → **ghafilisha**, énerver qqun  
 → +ghafil- + -ih- → |ghafilih-i|, (notion d'être énervé)  
 → **ghafilaha**, être énervé  
 +mkini < ar.  
 → +mkin- + -ih- → |mkinih-i|, (notion de confirmer)  
 → **mkiniha**, confirmer  
 → +mkin- + -ish- + -ih- → |mkinishih-i|, (notion d'être possible)  
 → **mkinishiha**, être possible

ii) *une seule forme étendue*

x applicatif :

- +gharimu < ar. *g.arima*  
 → +gharim- + -i- → |gharimi-i|, (notion de coûter)  
 → **gharimia**, coûter  
 Le nominal non naturalisé (sans préf.) *gharama* (-), dépense, quoique relevant de la même racine étymologique, ne peut servir de support lexical en raison du changement vocalique  
 +djamali < ar. *jamala*  
 → +djamal- + -i- → |djamali-i|, (notion d'être honoré)  
 → **mdjamalifu** (wa-), notable  
 Même chose en ce qui concerne les autres avatars de cette même racine arabe en shingazidja, *djimala*, somme, et ses propres dérivés (-djimalisa, totaliser, etc)  
 +dhumani < ar. *d.amina*  
 → +dhuman- + -i- → |dhumani-i|, (notion de servir à)  
 → **dhumania**, servir à & **dhumano** (ma-), but  
 Le rapport avec *dwanana*, responsable, ne peut être établi à travers la dérivation shingazidja

x causatif :

- +djadidi < ar. *jaddada*  
 → +djadid- + -ish- → |djadidish-i|, (notion de renforcer)  
 → **djadidisha**, renforcer  
 +rekebu < ar. *raqaba*  
 → +rekeb- + -esh- → |rekebish-i| (notion de corriger)  
 → **rekebesha**, corriger  
**rekebesho** (mi-), réforme  
 On remarque la voyelle /e/ exceptionnelle; dans la mesure où elle est lexicale, l'extension la reprend

x passif :

- +dhwamiri < ar. *d.amir*  
 → +dhwamir- + -iw- → |dhwamiriw-i|, (notion d'être recherché)  
 → **dhwamiriwa**, être recherché

moyen-neutre :

baini < ar. *bā'in*, clair

→ *+-bain-* + *-ih-* → *|-bainih-|*, inotion d'apparaître

→ *-bainyiha*, apparaître

Le rapport direct à *baina*, parmi, de ar. *baina*, entre, < BYN, ne peut pas être posé en raison de la divergence sémantique: la racine arabe BYN est polysémique; il est préférable de reconstruire un support hypothétique issu de ar. *bā'in*,

### conséquences lexicologiques

Les conséquences lexicologiques sont ici de deux ordres: d'une part, les termes issus des éléments assimilés à des racines étendues par les procédés bantu d'affixation sont reliés entre eux à travers ces procédés; d'autre part, il se crée une relation hiérarchisée entre le support lexical (quand il est attesté) et les termes issus de la racine.

Plus reprenons l'exemple déjà examiné:

*-badilisha*, changer et

*abadilishadji* (*wa-*), réformateur :

Ces deux termes sont construits par affixation à partir de la racine causative *badilish-*; ils sont reliés entre eux par ce biais.

Mais ces termes sont aussi reliés au support lexical *-badili*, changer, par le biais d'une relation spécifique, hiérarchisée (l'un donne les autres).

Ainsi, les familles lexicologiques constituées par les termes d'emprunts peuvent-elles présenter des similarités avec celles constituées par les termes du fonds bantu.

Mais les emprunts peuvent aussi offrir un autre type de relation lexicologique; ce point est développé maintenant.

## II Aspect lexicologique

Nous avons exposé les processus d'intégration morphologique des termes d'emprunt en shingazidja; la présence en grand nombre d'emprunts a, on l'a vu, contribué à modifier quelque peu le système de classes nominales, et l'appréhension de la langue dans la mesure où existent des termes qui ne relèvent pas des catégories monématiques bantu.

La présence des emprunts orientaux a aussi des incidences sur l'organisation lexicologique de la langue: l'organisation lexicologique selon le modèle bantu est donnée par la combinatoire monématique, il est donc attendu que les emprunts qui n'y rentrent pas modifient aussi cet aspect des choses.

Ainsi, les relations lexicologiques des termes d'emprunt entre eux se situent à plusieurs niveaux, selon leur modalité d'intégration:

1) *emprunts naturalisés : modèle bantu*

Comme les mots de fonds bantu, les emprunts, naturalisés, sont ou peuvent être reliés entre eux par les processus morphologiques bantu, qui les sous-tendent: ainsi, l'emprunt donnant naissance à un lexème (thème ou racine bi-valente), différents constituants pourront être reliés entre eux par modification des affixes actualisateurs.

x dans le cas des termes assimilés à des thèmes nominaux, on pourra avoir plusieurs nominaux reliés:

**mbazazi** (wa-), commerçant &  
**ubazazi**, (cl. 11), commerce

x dans le cas des termes assimilés à des racines nomino-verbales étendues, on pourra avoir des nominaux et des verbaux:

-**badilisha**, changer (tr)  
**mbadilishadji** (wa-), réformateur  
-**badiliha**, être changé  
**badiliho** (ma-), changement

*(ii) racines & support lexical*

Dans le cas précis des racines, nous avons vu que celles-ci étaient le plus souvent construites sur un support lexical, qui peut être un terme non naturalisé ou un thème; il y a évidemment relation, qui ne passe pas de fait par les canons bantu, entre ce support et les termes qui sont construits à partir de cet élément. Cette relation est établie à partir de l'identité formelle entre le support et les formations étendues (nous avons parlé de relation hiérarchisée).

-badili, changer  
 → -badilisha, changer (tr)

*iii) la filière étymologique*

Mais il existe aussi des relations entre termes d'emprunt à l'arabe en shingazidja (à tout niveau d'intégration) qui ne peuvent s'expliquer ni par des processus shingazidja, ni par une construction hiérarchisée: ils ne peuvent être établis qu'à travers la langue source.

Ainsi, des relations entre

-fikiri, penser et fikra (-), pensée  
 adabu (-), respect et adabii (ma-), personne respectueuse  
 mwarabu (wa-), arabe (personne) et -staarabiha, être cultivé  
 etc

Cette situation, totalement spécifique des emprunts, nécessite quelques développements.

*Rapprochements à base étymologique*

Des termes d'emprunt de catégories grammaticales et de niveaux d'intégration indifférents peuvent être rapprochés sur la base de leur proximité formelle et sémantique en shingazidja.

Ces rapprochements, dans un nombre non négligeable de cas, sont indiscutablement ressentis par les locuteurs, et ne sauraient être négligés dans une analyse lexicologique et morphologique de la langue.

La parenté lexicologique entre les termes ne peut être établie à l'intérieur du shingazidja, sans référence au système de dérivation de l'arabe: en effet, le shingazidja a emprunté, nous l'avons dit, des mots de l'arabe, mais ces mots eux-mêmes, à l'intérieur de l'arabe, s'analysent et peuvent se ramener à telle ou telle racine arabe, en fonction des processus de dérivation afférents; la parenté ne peut s'établir que par la médiation de la langue source, l'arabe.

Ces faits ne sauraient surprendre des locuteurs français, habitués à considérer (sans doute) comme lexicologiquement reliés *île* et *insulaire*, *doigt* (dwal) et *digital*, etc: or, ces regroupements ne peuvent se fonder que, sinon sur des processus synchroniques dérivationnels français

Nous donnons quelques exemples de termes shingazidja de catégories grammaticales différentes, liés par une commune origine (racine étymologique arabe), dont il paraît peu discutable que le rapprochement soit admis et ressenti par les locuteurs; la racine arabe est en majuscule:

Nous ne citons ici que la racine arabe; on se reportera au lexique pour des indications étymologiques détaillées

x des verbes et des noms

- fikiri, penser & fikra, pensée < ar. FKR
- staadjabu, être étonné & adjabu, merveille < ar. eJB
- daraka, responsabilité & -diriki, s'occuper de < ar. DRK
- dhulumu, exploiter & dhwalimu, exploiteur < ar. Z.LM
- hotubu, discourir & hotuba (-), discours < ar. KT.B

x des nominaux entre eux

- adabu, respect & adabii, personne respectueuse < ar. 'DB
- huba (-), désir & mahaba, amour < ar. H.BB
- baridi (-), froid et burda (-), froid < ar. BRD

x des verbes et des invariables

- kurubia, approcher & karibu, près de < ar. QRF
- dumu, durer & daima, toujours < ar. DWM
- etc

De tels cas sont légion.

Parmi les différents procédés de dérivation de l'arabe, certains, du moins perçus à travers les emprunts faits par le shingazidja, sont particulièrement

quents: par exemple les modifications de schème qui distinguent les 'noms verbaux' (*mas.dar*) et les verbes, ou les formations dérivées arabes à préfixe (VIè, VIIè et Xè formes).

Nous n'avons ni la prétention ni l'objectif d'exposer ici les mécanismes de la dérivation arabe: ceux-ci ne sont considérés qu'à travers ce qui en apparaît en shingazidja - ainsi n'est-il pas utile pour notre propos de distinguer entre les différents *mas.dar* non plus qu'entre les différentes formes dérivées arabes à préfixe

Comme le fait remarquer BLACHERE à propos des *mas.dar* (noms verbaux) arabes passés dans des langues que l'arabe a influencées (où l'auteur, significativement, n'indique pas de langues africaines), ces termes sont déroutants à s'en tenir à la langue elle-même (*"Le nom verbal (...) ce vocabulaire arabe abstrait, religieux et philosophique, a passé avec l'Islam dans le persan, dans le malais, dans le turc. Introduit dans des langues d'un système tout différent de celui des langues sémitiques, il y produit un étrange effet et déroute le lecteur qui ignore la grammaire arabe"* (BLACHERE & GAUDEFROY-DEMONBYNES, 1975 : 78).

\* *mas.dar* + verbe

*badala*, équivalent & *-badili*, changer < ar. BDL  
*baraka*, pouvoir charismatique & *-bariki*, favoriser < ar. BRK  
*daraka*, responsabilité & *-daraki* ou *-diriki*, s'occuper de < ar. DRK  
*dhwamana*, responsable & *-dhwamini*, prendre en charge < ar. D.MM  
*farka*, séparation & *-fariki*, partir < ar. FR  
*amani* (-), sécurité et *-imani*, croire < ar. 'MM  
*badala* (-), équivalent et *-badili*, changer < ar. BDL  
*ibada*, précepte religieux et *-abudu*, croire < ar. εBD  
*difai* (-), défense et *-dafii*, défendre < ar. DFe  
 etc

\* formes simples et dérivées arabes

*adjabu* (-), merveille et *-staadjabu*, être étonné < ar. εJB  
*mwarabu*, arabe (personne) et *-staarabiha*, être cultivé < ar. εRB  
*-fasiri*, expliquer et *-tafasiri*, expliquer < ar. FSR  
*-arifu*, faire savoir et *taarifa* (-), message < ar. εRF  
 etc

Mais, pour fréquents que soient les termes shingazidja résultant des processus arabes dérivationnels évoqués ci-dessus, les processus en question n'en sont pas devenus indigènes, ils n'ont pas été repris par la langue comme tels, ils

restent étrangers à son fonctionnement morphologique, ils ne s'appliquent pas: on ne pourra pas fabriquer un nom à partir d'un verbe d'origine arabe en shingazidja, ou un verbe à partir d'un nom, par modification du schème, ni fabriquer des dérivés en *ta* ou *sta*, etc; il s'agit, comme l'a écrit ALEXANDRE, "*de processus de dérivation morts*"; les termes n'existeront en shingazidja que s'ils ont été empruntés individuellement chacun à l'arabe.

Les cas, déjà signalés d'analogie entre préfixes arabes et préfixes bantu sont peu nombreux par rapport aux cas non motivés en shingazidja - ils ne sont pas en position, en tous cas, pour modifier l'aspect fondamentalement immotivé des relations étymologiques dont nous traitons ici.

**mwadhini** (ma-), muezzin et -adhini, appeler à la prière < ar. 'DM  
**mdai** (wa-), requérant et -dai, exiger < ar. DeW  
**maaswi**, blasphème et -aswi, blasphémer < ar. eS.Y  
 etc

Cela pose le problème de la constitution des familles lexicologiques shingazidja, c'est-à-dire des rapprochements à effectuer dans un lexique: à côté des exemples comme ceux cités ci-dessus, choisis pour leurs convergences multiples et à ce titre non seulement acceptés mais aussi effectués spontanément (nous semble-t-il) par tous les locuteurs, l'apparementement entre la majorité des termes issus de la même racine arabe ne sera pas reconnu, ou l'attitude variera d'un locuteur à l'autre, selon sa propre réflexion sur sa langue, sa connaissance de l'arabe, l'intégration des termes en shingazidja, etc; ainsi que l'écrit ALEXANDRE, parlant des emprunts du swahili, "*les différentes dérivations d'un même thème arabe ne sont généralement pas ressenties comme morphologiquement apparentées*" (ALEXANDRE 1984 : 9).

Nous donnons quelques exemples où le sentiment de parenté paraît éminemment variable:

**mwalinu** (wa-), devin et **mtaalamu wa-**, expert < ar. eLM  
**alama** (-) signe et **elmu** (-), sortilège < ar. eLM  
 et aussi ces termes entre eux, **mwalinu** et **alama**,  
**dawa** (-), procédure judiciaire & **dua** (-), bénédiction < ar. DeW  
**halafu**, après & -hitlafiana, se disputer < ar. KLF

-dhuru, nuire & madhwara, tort < ar. D.RR  
 dhulma (-), fraude et dhwalimu (ma-), exploiteur < Z.LM  
 etc

conçoit qu'il paraisse difficile de baser la décision de rapprochement sur  
 seule intuition des locuteurs - en tous cas en l'état (inexistant à notre  
 connaissance) des travaux lexicographiques oeuvre de locuteurs grand-comoriens  
 mêmes; de ce fait, nous sommes amené à la solution étymologique: les  
 familles lexicales sont constituées par référence à l'étymologie, et seront  
 regroupés tous les termes shingazidja d'origine arabe dont la source  
 est référée, dans la lexicographie arabe, à une même racine.

cela provoque des regroupements, justifiés par l'étymologie, mais non intuitifs,  
 éventuellement même contraires à l'intuition des locuteurs, mais paraît le seul  
 moyen d'éviter trop d'arbitraire; cela bien entendu nous rend tributaire de nos  
 références en lexicographie arabe (en effet, en arabe comme en d'autres  
 langues, tous les regroupements lexicologiques ne sont pas uniment acceptés;  
 certaines traditions s'opposent; rappelons que nous nous abritons ici derrière  
 autorité lexicographique que représente le dictionnaire de WEHR, et que  
 écarter des rapprochements effectués dans la lexicographie arabe serait hors  
 propos).

ainsi, on regroupera, sur la base de la racine arabe commune:

ada, grand mariage & udi, luth < ar. eAD  
 dawa (-), procédure judiciaire, dua (-), prière & -dai, exiger < ar. DaW  
 dari (malari), étage, diwara (ma-), roue & mudiria (-), moudiria < ar. DWR  
 dhahabu (-), or & dhihabi (ma-), rite < ar. DHB  
 -dhibiti, prendre en main & madhbuti, efficace < ar. D.BT.  
 etc

Comme exemple de regroupements discutés en arabe, citons la racine unique  
 (chez WEHR) DWR qui donne *dār*, maison (→ Ng dari) et *dawwār*, en rotation  
 (→ Ng diwara), ou QS.S., qui donne *qis.s.s*, histoire (→ Ng kiswa, histoire)  
 et *miqas.s.*, ciseaux (→ Ng mkasi), etc

(informations dues à l'amabilité d'Antoine LONNET)

Cette solution a aussi l'avantage de faciliter l'insertion dans la langue d'emprunts non relevés ici, notamment quand il s'agit de termes issus d'une racine arabe déjà attestée: il ne faut pas sous-estimer la compétence en arabe de certains au moins des locuteurs grand-comoriens: ceux-ci, en cas de non-attestation d'un terme en shingazidja, seront souvent à même de retourner vers les sources arabes et d'emprunter alors le terme souhaité; il ne faut pas sous-estimer non plus le fait que l'arabe serve de réserve lexicale au shingazidja, et que, dans les circonstances actuelles où son emploi officiel aux Comores est encouragé, où les relations culturelles et humaines avec les pays arabes se développent, il paraît vraisemblable que le shingazidja acceptera de plus en plus de mots arabes, éventuellement avec une forme et un sens 'redressés' (plus proches de l'original).

\* \*

\*

En conséquence, la constitution des familles de mots d'origine arabe fait appel à plusieurs éléments:

- x des dérivations bantu pour les emprunts naturalisés;
- x le rapport au support lexical éventuel pour les racines étendues construites
- x les relations étymologiques à travers l'arabe

conçoit ainsi que les familles lexicologiques arabes soient souvent plus riches en mots que les familles bantu.

Il paraît souhaitable qu'une étape ultérieure de la réflexion sur le shingazidja soit d'opérer un partage dans les regroupements étymologiques, sur la base d'une évaluation de ces regroupements dans le cadre du shingazidja: il s'agirait de dissocier les termes shingazidja qui, quoiqu'issus d'une même racine étymologique, ne présentent pas en shingazidja de relation formelle et sémantique ressentie, de ceux dont la relation est vivante dans la langue; inversement, il serait tout à fait légitime de rassembler des termes issus de racines distinctes en arabe, mais dont les avatars se télescopent en shingazidja, sur les plans formels (soit neutralisation d'oppositions phonologiques, soit racines arabes homophones, soit accidents phonétiques), et sémantique (que nous n'envisageons pas ici); comme nous l'avons dit, cela ne peut se faire sans l'intervention active de locuteurs.