



LANGUES ET GRAMMAIRES  
EN (ILE DE) FRANCE

KASSIM MOHAMED SOYIR /LLF-UP  
ANNE ZRIBI-HERTZ SFL/PARIS 8-CNRS

## LE COMORIEN (shikomori)

[quelques contrastes pertinents pour l'acquisition du Français Langue Seconde par des locuteurs du comorien]



### **LGIDF**

Le projet Langues et Grammaires en (Île-de) France propose :

- un **SITE INTERNET** (<http://lgidf.cnrs.fr/>) conçu par des linguistes, des didacticiens et des professionnels de l'Éducation nationale contenant des informations linguistiques sur diverses langues parlées en (Ile-de) France, des descriptions scientifiques des propriétés phonologiques et grammaticales, une histoire et un lexique traduits et enregistrés dans toutes les langues étudiées, des jeux linguistiques, des ressources bibliographiques pour chaque langue et des liens conduisant à d'autres sites pertinents
- des **FICHES LANGUES** qui présentent une description contrastive et les particularités spécifiques de chaque langue pour les professionnels francophones en charge de publics allophones
- des outils « **EN FRANÇAIS ET AILLEURS** » sur des thématiques du français, avec des activités pédagogiques « **REGARDONS NOS LANGUES** ».

## PRÉSENTATION GÉNÉRALE

Le comorien est une langue à syntaxe bantoue et au lexique fortement arabisé — résultat d'un ancien peuplement métissé arabo-africain. Classé dans le groupe *sabaki* avec le swahili et le kipokomo (autres langues bantoues orientales), le comorien est parlé aujourd'hui par environ un million de personnes, dont une importante diaspora, émigrée notamment en France depuis les années 1960-70. L'archipel des Comores est situé dans l'Océan Indien entre la côte Est de l'Afrique et l'île de Madagascar. Il est composé de quatre îles nommées Ngazidja (Grande-Comore), Mwali (Mohéli), Ndzواني (Anjouan) et Maoré (Mayotte). Les trois premières ont accédé à l'indépendance en 1975 en formant l'Union des Comores, tandis que Mayotte est restée française puis devenue "département français d'Outre-Mer" en 2011. Dans l'Union des Comores, le comorien (langue vernaculaire) a statut de troisième langue officielle à côté du français (langue scolaire et administrative) et de l'arabe (langue de la culture islamique). A Mayotte, le comorien (mahorais) a statut de "langue régionale" et le taux d'illettrisme est, globalement, plus élevé que dans l'Union des Comores. La langue comorienne a une coloration locale spécifique dans chacune des quatre îles mais se subdivise en deux variantes principales : occidentale et orientale (les exemples de ce document sont en shingazidja/grand comorien, variante occidentale). La politique d'alphabétisation mise en place dans l'Union des Comores à partir de 1975 a substitué une graphie comorienne en caractères latins à l'ancienne écriture arabe, bien qu'il n'existe pas encore de code orthographique officiel.

## ÉLÉMENTS DE PHONOLOGIE

Les voyelles antérieures arrondies [y] (*pu*), [ø] (*peu*), [œ] (*peur*) sont absentes en comorien, ainsi que [ə] (*petit*), et la semi-consonne [ɥ] (*lui*) : les apprenants pourront donc prononcer *devenir* [devenir], percevoir *appétit* comme un dérivé de *petit*, et ne pas distinguer oralement *lui* de *Louis* ou *buée* de *bouée*. Les voyelles nasales n'existent pas en comorien et demandent donc un apprentissage en français : *paix/pain*, *pas/paon*, *peau/pont*, ainsi que les distinctions [e]/[ɛ] (*fée-fait*, *fête*) et [o]/[ɔ] (*peau*, *paume-pomme*). En revanche, en dehors de [ɥ] déjà mentionné, les consonnes du français sont toutes disponibles en comorien. La prosodie du français ne devrait pas non plus être problématique pour les locuteurs de cette langue car une prosodie à accent final de groupe tend à s'y développer actuellement parallèlement à la disparition des tons.

## ÉLÉMENTS DE GRAMMAIRE

Comme toutes les langues bantoues, le comorien est une langue à *classes nominales* : chaque nom du lexique est affecté à une classe caractérisée par une série de marques, typiquement préfixales, s'attachant au nom lui-même et/ou à tous ses dépendants (déterminants, adjectifs, pronoms de 3<sup>ème</sup> personne...). On a pu identifier pour le comorien dix-sept classes de noms (c1, c2...c6) — les classes impaires étant singulières, et les paires, plurielles. Un nom variable en nombre est affecté à une paire de classes (impaire au singulier, paire au pluriel), un nom invariable en nombre est en revanche "monoclasse" — il n'est affecté qu'à une classe, comme en français certains noms le plus souvent singuliers (*de l'eau*/ ?*des eaux*) ou pluriels ( ?*un épinard*/*des épinards*). Noter toutefois que les noms de substances "massives" ('eau', 'boue', 'sable', etc.) appartiennent couramment en comorien à la classe 6 (une classe plurielle) alors que leurs contreparties en français sont singulières. Ci-dessous un petit échantillon illustratif des classes nominales (trois noms biclasses et un nom monoclasse) :

TABLEAU 1

classe n°	Singulier	classe n°	Pluriel
1	<i>mu-ana</i> '(un) enfant'	2	<i>w-ana</i> '(des) enfants'
3	<i>m-pira</i> '(un) ballon'	4	<i>mi-pira</i> '(des) ballons'
5	<i>dzi-tso</i> '(un) oeil'	6	<i>ma-tso</i> '(des) yeux'
	-----	6	<i>ma-dji</i> '(de l')eau'

Les classes nominales bantoues sont comparables aux genres en français : l'appartenance d'un nom à une classe ou à un genre détermine la forme de ses dépendants (*un/ce petit banc vert/une/cette petite chaise verte*), et il faut mémoriser la forme et la distribution des marques d'accord sur chaque type de dépendant :

TABLEAU 2

nom	classe n°	N + dépendants	traduction française
<i>mu-ana</i> 'enfant'	1	<b>ye</b> <i>mu-ana</i> <b>m-huu</b> <b>o-i</b> C1.DF C1-enfant C1-grand C1-DM	'ce grand enfant'
<i>w-ana</i> 'enfants'	2	<b>o</b> <i>w-ana</i> <b>wa-raru</b> <b>wa-huu</b> <b>wa-nu</b> C2.DF C2-enfant C2-trois C2-grand C2-DM	'ces trois grands enfants'
<i>m-ri</i> 'arbre'	3	<b>o</b> <i>m-ri</i> <b>mu-huu</b> <b>u-nu</b> C3.DF C3-arbre C3-grand C3-DM	'ce grand arbre'
<i>mi-ri</i> 'arbres'	4	<b>ye</b> <i>mi-ri</i> <b>mi-raru</b> <b>mi-huu</b> <b>yi-nu</b> C4.DF C4-arbre C4-trois C4-grand C4-DM	'ces trois grands arbres'
<i>shi-o</i> 'livre'	7	<b>ye</b> <i>shi-o</i> <b>shi-huu</b> <b>shi-nu</b> C7.DF C7-livre C7-grand C7-DM	'ce grand livre'
<i>zi-o</i> 'livres'	8	<b>ye</b> <i>zi-o</i> <b>zi-raru</b> <b>zi-huu</b> <b>zi-nu</b> C8.DF C8-livre C8-trois C8-grand C8-DM	'ces trois grands livres'

Une différence bantou/français est que les marqueurs de classe sont généralement préfixaux en bantou alors que les marques de genre sont plutôt suffixales en français. Une autre est que la classe peut souvent

s'induire de l'observation du nom lui-même en comorien, alors que le genre n'est révélé en français que par certains suffixes (ex. *lav-**age***, masculin, *lave-**rie***, féminin — mais *accent*, masculin, *dent*, féminin) : en comorien, un nom humain commençant par *mu* appartient à la classe 1 et fait son pluriel en c2, un nom non-humain à initiale *m(u)* appartient à c3 et fait son pluriel en c4, un nom à initiale *sh/h/i* appartient à c7, etc. Il existe cependant aussi en comorien des noms dont les propriétés internes ne révèlent pas la classe (ex. *mbwa* 'chien(s)', c9 ou c10) ainsi que des noms irréguliers comme celui signifiant 'grand-mère', dont les formes singulière (*koko*) et plurielle (*makoko*) relèvent des classes 5 et 6, mais dont les dépendants s'accordent dans les classes 1 et 2 sur la base de sa sémantique (nom humain).

Le comorien possède un élément fonctionnel pré-nominal (glosé ici DF) qui, comme l'article défini du français, s'accorde en classe (genre) et nombre avec le nom et intervient à la fois sous interprétation "définie" (entité présupposée : ex. (1a/a')) et sous interprétation générique (espèce générale : ex. (1b/b')). Le déterminant défini comorien peut toutefois apparaître en présence du démonstratif (1c), contrairement à ce qu'on a en français (1c') :

<b>(1a)</b> <i>Ali ha ono o w-ana.</i> Ali C1.SUJ voir C2.DF C2-enfant 'Ali a vu les enfants.'	<b>(1a')</b> <i>Ali a vu <b>les</b> enfants.</i>
<b>(1b)</b> <i>Ali ye hwandza o w-ana.</i> Ali C1.SUJ aimer C2.DF C2-enfant 'Ali aime les enfants [enfants particuliers/en général]'	<b>(1b')</b> <i>Ali aime <b>les</b> enfants.</i>
<b>(1c)</b> <i>Ali ye hwandza o w-ana wa-nu.</i> Ali C1.SUJ aimer C2.DF C2-enfant C2-DM 'Ali aime ces enfants.'	<b>(1c')</b> <i>Ali aime (<b>*les</b>) <b>ces</b> enfants.</i>

Les déterminants indéfinis et partitifs du français n'ont pas de contreparties en comorien, qui utilise ici des noms nus, cf. (2a, b) — comparer *madji/ye madji* en (2b). Le numéral *-dzima*, signifiant 'un', n'intervient que pour indiquer précisément le nombre (2c) :

<b>(2a)</b> <i>Ali ha ono mu-ana / w-ana.</i> Ali C1.SUJ voir C1-enfant/ C2-enfant	<b>(2a')</b> <i>Ali a vu <b>un/des</b> enfant(s).</i>
<b>(2b)</b> <i>Ali ha no ma-dji / ye ma-dji.</i> Ali C1.SUJ boire C6-eau / C6.DF C6-eau	<b>(2b')</b> <i>Ali a bu <b>de</b> l'eau / l'eau.</i>
<b>(2c)</b> <i>Ali ha ono mu-ana m(u)-dzima.</i> Ali C1.SUJ voir C1-enfant C1-un	<b>(2c')</b> <i>Ali a vu <b>un</b> enfant (pas deux).</i>

L'organisation des constituants dans le groupe nominal est illustrée plus haut dans le tableau 2 : on constate que le numéral et le démonstratif sont postnominiaux en comorien (ils se comportent comme les adjectifs, tous postnominiaux), et que l'ordre linéaire des modificateurs postnominiaux est : Numéral-Qualificatif-Démonstratif ('...trois grands ces'). L'emploi des déterminants indéfinis et partitifs et la répartition des déterminants et adjectifs de part et d'autre du nom, en français, méritent donc une attention particulière de la part des apprenants comoriens.

L'ordre canonique des constituants dans la phrase comorienne est : Sujet-(Modal)-Verbe-Objet. Le système des temps-modes-aspects est assez complexe et très différent de celui du français. On peut retenir qu'en comorien comme en français standard, la flexion temporelle se combine avec un "marqueur de sujet" (glosé SUJ) — désinence de personne/nombre en français, affixe de personne-nombre ou de classe en comorien (personne-nombre aux personnes 1 et 2, classe à la 3ème personne). En présence d'un sujet lexical (de 3ème personne par définition), le marqueur de sujet *s'accorde* (en classe) avec le sujet (3b, c, d), mais en l'absence d'un sujet lexical (3a, b', c', d'), ce marqueur *identifie* à lui seul le sujet à la manière d'un pronom — comme le fait la désinence personnelle dans une langue comme l'espagnol (*vine ayer* 'je suis venu(e) hier'/*vino ayer* 'il/elle est venu(e) hier'). Le marqueur de sujet comorien est à cet égard comparable aux pronoms clitiques sujets du français, qui indiquent la personne, le nombre et (à la 3ème personne) le genre :

<b>(3a)</b> <i>tsi -on-o...</i> 1SG.SUJ-voir-SFX 'j'ai vu...'	
<b>(3b)</b> <i>ye mu-ana ha -on-o...</i> C1.DF C1-enfant C1.SUJ -voir-SFX Lit. 'l'enfant il a vu...'	<b>(3b')</b> <i>ha -on-o...</i> C1.SUJ -voir-SFX 'il (= enfant) a vu...'
<b>(3c)</b> <i>le paha li -on-o...</i> C5.DF C5-chat C5.SUJ -voir-SFX Lit. 'le chat il a vu...'	<b>(3c')</b> <i>li -on-o...</i> C5.SUJ -voir-SFX 'il (= chat) a vu...'
<b>(3d)</b> <i>o w-ana wa -on-o...</i> C2.DF C2-enfant C2.SUJ -voir-SFX Lit. 'les enfants ils ont vu...'	<b>(3d')</b> <i>wa -on-o</i> C2.SUJ -voir-SFX 'ils (= enfants) ont vu...'

Bien que le pronom sujet se combine couramment avec un sujet lexical en français parlé informel (ex. *Jean, il a dit oui*), ils s'excluent mutuellement en français écrit formel. Il en va différemment en comorien, où les phrases comme (3b, c, d) sont parfaitement standard. Ce contraste mérite peut-être d'être souligné généralement aux bantouphones apprenant le français.

Le marqueur de sujet comorien est suffixé au marqueur d'aspect Inaccompli, mais à l'Accompli il est préfixé au radical du verbe. Ci-dessous un échantillon du verbe *huona* 'voir' aux aspects Inaccompli et Accompli :

NB-PERS/CLASSE	INACCOMPLI	TRADUCTION F	ACCOMPLI	TRADUCTION F
<b>1SG</b>	<i>nga-m on-o</i> INAC-1SG.SUJ voir-SG	'je vois'	<i>tsi -on -o</i> 1SG.SUJ-voir-SFX	'j'ai vu'
<b>2SG</b>	<i>ng-o on-o</i> INAC-2SG.SUJ voir-SG	'tu vois'	<i>hu -on -o</i> 2SG.SUJ -voir-SFX	'tu as vu'
<b>3SG/c1</b>	<i>ng-u on-o</i> INAC-C1.SUJ voir-SG	'Il (= enfant) voit'	<i>ha -on -o</i> C1.SUJ -voir-SFX	'il (enfant) a vu'
<b>1PL</b>	<i>nga-ri on-ao</i> INAC-1PL.SUJ voir-PL	'Nous voyons'	<i>ri -on -o</i> 1PL.SUJ -voir-SFX	'nous avons vu'
<b>2PL</b>	<i>nga-m on-ao</i> INAC-2PL.SUJ voir-PL	'Vous voyez'	<i>m -on -o</i> 2PL.SUJ -voir-SFX	'vous avez vu'
<b>3PL/c2</b>	<i>ng-wa on-ao</i> INAC-C2.SUJ voir-PL	'Ils (= enfants) voient'	<i>wa -on -o</i> C2.SUJ -voir-SFX	'ils (enfants) ont vu'

Les compléments suivent canoniquement le verbe en comorien. Pour les verbes à double complément comme 'donner', les deux compléments sont construits directement (sans préposition) dans l'ordre : Bénéficiaire - Thème (4b) — comme c'est le cas dans beaucoup de langues, par exemple l'anglais : *John gave Mary the book*, lit. 'Jean a donné Marie le livre'. Si un seul complément est pronominalisé, le pronom est antéposé au verbe (4c, d). Mais si deux compléments directs sont pronominalisés en même temps, c'est le pronom Bénéficiaire qui est placé à gauche du verbe (4e) :

<b>(4a)</b> <i>Tsi -on-o y-e shi-o.</i> 1SG.SUJ voir-VV C7-DF C7-livre	<b>(4a')</b> <i>J'ai vu le livre.</i>
<b>(4b)</b> <i>Tsi nik-a Ali y-e shi-o.</i> 1SG.SUJ donner-SG Ali C7-DF C7-livre	<b>(4b')</b> <i>J'ai donné le livre à Ali.</i>
<b>(4c)</b> <i>Tsi shi nik-a (Ali).</i> 1SG.SUJ C7.OBJ donner-SFX (Ali)	<b>(4c')</b> <i>[le livre] Je l'ai donné (à Ali).</i> <b>(4c'')</b> <i>*J'ai donné (à) Ali lui/le.</i>
<b>(4d)</b> <i>Tsi m nik-a y-e shi-o.</i> 1SG.SUJ C1.OBJ donner-SFX C7-DF C7-livre	<b>(4d')</b> <i>[Ali] Je lui ai donné le livre.</i> <b>(4d'')</b> <i>*J'ai donné lui le livre.</i>
<b>(4e)</b> <i>Tsi m nik-a sho.</i> 1SG.SUJ C1-OBJ donner-SG C7-OBJ	<b>(4e')</b> <i>[Ali, le livre] Je le lui ai donné.</i> <b>(4e'')</b> <i>*Je lui ai donné lui/le.</i>

Les règles de placement des pronoms compléments sont donc globalement très différentes en comorien et en français.

Les questions OUI/NON sont signalées en comorien par l'intonation (5b), optionnellement complétée par l'insertion de la particule interrogative *ye* à l'initiale de la phrase (5c), que l'on peut rapprocher du marqueur interrogatif *est-ce que* [ɛsk] en français. L'inversion du verbe et du pronom sujet, attestée dans les questions oui/non en français écrit formel (5b'), n'a pas d'équivalent en comorien :

<b>(5a)</b> <i>Ali ye hwandza o w-ana.</i> Ali C1.SUJ aimer C2.DF C2-enfant 'Ali aime les enfants.'	<b>(5b)</b> <i>Ali ye hwandza o w-ana ?</i> Ali C1.SUJ aimer C2.DF C2-enfant 'Ali aime les enfants ?' <b>(5b')</b> <i>Ali aime-t-il les enfants ?</i>
	<b>(5c)</b> <i>Ye Ali ye hwandza o w-ana ?</i> Q Ali C1.SUJ aimer C2.DF C2-enfant 'Est-ce qu'Ali aime les enfants ?'

La particule interrogative *ye* intervient aussi en comorien à l'initiale des questions partielles, mais le constituant questionné reste dans sa position canonique (6a), sans se déplacer à l'initiale comme il doit le faire en français normé :

<b>(6a)</b> <i>Ye Ali ye hwandza ndo ?</i> Q Ali C1.SUJ aimer quoi	<b>(6a')</b> <i>Qui est-ce que Ali aime ?</i> <b>(6a'')</b> <i>Qui Ali aime-t-il ?</i> <b>(6a''')</b> <i>*Est-ce qu'Ali aime qui ?</i>
<b>(6b)</b> <i>Ye Ali ye hwandza hindri ?</i> Q Ali C1.SUJ aimer quoi	<b>(6b')</b> <i>Qu' est-ce que Ali aime ?</i> <b>(6b'')</b> <i>*Est-ce qu'Ali aime quoi ?</i>

### ÉLÉMENTS BIBLIOGRAPHIQUES

< <http://lgidf.cnrs.fr/sites/lgidf.cnrs.fr/files/images/biblio.comorien.vf.pdf> >

### GLOSSAIRE

c1, c2... = classe nominale n°. ; DF = déterminant défini ; DM = (adjectif) démonstratif ; INACC = (aspect) inaccompli ; OBJ = pronom objet ; PL = pluriel ; Q = particule interrogative ; SG = singulier ; SFX = suffixe (verbal) ; SUJ = marqueur de sujet ; 1PL/1SG, 2PL/2SG = 1ère, 2ème personne (singulier ou pluriel)

REFERENCE halshs-[01493383](https://halshs.archives-ouvertes.fr/halshs-01493383)  
2017

Logo LGIDF : Stanca SOARE Illustration  
empruntée au site : [majik.over-blog.com](http://majik.over-blog.com)