

# La longueur vocalique en kinyarwanda : statut linguistique et contraintes orthographiques

**Alphonse KABANO**  
Université Nationale du Rwanda\*  
Faculty of Arts, Media and Social Sciences  
Email : [akabano2@nur.ac.rw](mailto:akabano2@nur.ac.rw)  
Mobile : 250-0782198286

## **Abstract :**

*The law order of kinyarwanda (1985) prohibits the notation of long vowels, on the contrary of the practice in some other Bantu speaking countries in East-Africa. We look at the potential impact of this order in consideration of the abundance of long vowels and the linguistic constraints on its variability and predictability in this language. An approach based on the ability of vowel length recognition by the users is advocated to better assess the risk of confusion when long vowel are not marked in contrast to brief vowels. A highly efficient model is proposed for an economic and non confusing orthography. It is strongly recommended to review the kinyarwanda orthographic system so as to meet new sociolinguistic, educational, technological re-quirements and regional integration of Rwanda.*

**Key words:** kinyarwanda, vowel length, orthography.

## **1. Introduction**

Il va y avoir bientôt un quart de siècle, l'instruction ministérielle de 1985 portant fixation de l'orthographe du kinyarwanda au Rwanda stipulait, en son article 2, que : «*Le dédoublement des voyelles est prohibé, sauf dans l'orthographe des mots "saa" (temps horaire) et "yee" (oui), dans l'usage expressif de l'allongement des voyelles (yooo!) ainsi que dans l'usage scientifique*». Cette disposition reste en vigueur, et, à plusieurs reprises, l'on s'est posé des questions à propos de sa justesse. Aujourd'hui encore nous nous demandons quels sont les fondements théoriques et la rentabilité d'un tel système qui prohibe aussi en même temps, comme on le sait, la notation de la tonalité. En effet, qu'est-ce-qu'on perd et qu'est-ce-qu'on gagne en interdisant la notation de la longueur vocalique? Pourquoi, dans un texte de loi, on tient à accorder de l'importance à deux mots, dont un (*saa*) est de surcroît un emprunt, comme s'ils étaient les seuls cas problématiques dans la langue? Qu'est-ce qui auto-rise à noter les faits relatifs à l'expressivité s'il s'agit là d'éléments qui, en fait, ne font pas partie de l'information

---

\* Cet article a été rédigé lors de notre séjour au Laboratoire de Sociolinguistique (Département de linguistique, Université d'Ottawa, Canada) en 2009 en qualité de chercheur invité. Nous sommes très reconnaissant envers les responsables du Laboratoire et ceux du Département pour leur accueil.

intrinsèque et stable sur les mots? Quelle raison y a-t-il à distinguer l'orthographe courante et l'orthographe scientifique en prenant la deuxième plus 'au sérieux', pourrait-on dire? Nous entendons par rentabilité le rapport entre le coût cognitif qu'entraînerait la surcharge du texte dans lequel les voyelles longues seraient dédoublées (ou représentées par des moyens diacritiques) et la capacité de reconnaissance des mots par le lecteur lorsque le texte kinyarwanda, pour des raisons d'économie et en conformité à la loi, ne comporterait pas de dédoublement.

Même si les usagers du kinyarwanda semblent s'accommoder du système de notation dont il est question, le processus de prise de décision soulève en soi des interrogations. Le problème est d'ailleurs pertinent au Rwanda comme ailleurs. Nous pensons, par exemple, à la rencontre du SIL organisée à Dallas en 2005 sur la problématique de l'orthographe bantu. La contribution de Stegen (2005) que l'on lit dans un document électronique portait précisément sur le statut de la quantité vocalique et ses impacts sur les systèmes orthographiques dans les langues bantu de l'Afrique de l'Est (Uganda et Tanzanie). Stegen passe en revue les différentes configurations de la quantité vocalique et suggère des stratégies qui devraient mieux adéquatement guider la décision politique en matière d'orthographe. Son mérite est de proposer de procéder à des tests pour évaluer l'effet réel du phénomène de la quantité vocalique chez les locuteurs-auditeurs (scripteurs-lecteurs), de manière à intégrer dans la conception du système orthographique les contraintes relatives à la performance des usagers auxquels il est destiné. C'est dans ce genre de débat que nous relançons la question de la quantité vocalique en kinyarwanda, en particulier au moment où le Rwanda est amené à s'intégrer dans la Communauté Est-Africaine où il va certainement falloir harmoniser bon nombre de pratiques, dont l'orthographe bantu, justement.

Ce qui est en cause, sur le plan de la théorie, c'est la nature et le format des informations phonétiques et phonologiques sur les mots, telles qu'elles sont déclarées dans la mémoire des usagers de la langue, mais aussi celle des procédures utilisées pour y accéder. En effet, prohiber formellement la notation de la quantité vocalique, qui est phonologiquement parlant une dimension supra-segmentale, laisse supposer que l'information segmentale suffit à elle seule pour que le lecteur reconnaisse sans peine le mot, ou alors que les contextes dans lesquels les mots sont utilisés sont toujours suffisamment clairs pour éliminer tout risque d'ambiguïté. On reconnaîtra que la dissociation entre les plans segmental et suprasegmental est, dans le cadre de la théorie de la phonologie auto-segmentale (Goldsmith, 1979), une démarche correcte. Toutefois, l'on se demandera si la quantité vocalique en kinyarwanda est un trait sous-spécifié quelque part en structure profonde

(sur la théorie de la sous-spécification, voir Archangeli, 1988), de telle sorte qu'il soit autorisé de l'omettre dans la représentation de surface (disons-le très simplement, dans la notation orthographique) où une procédure quasi automatique permettrait de recouvrer les faits profonds.

Dans cet article, nous examinons s'il est raisonnable de prohiber la notation de la longueur vocalique en kinyarwanda. Certes, nous n'avons procédé à aucune recherche expérimentale auprès des usagers pour évaluer la lisibilité d'un texte kinyarwanda ne comportant pas de notation de la quantité; cependant, les quelques considérations que nous allons émettre pourraient servir d'hypothèses pour une investigation empirique dans des domaines variés tels que : a) le développement de la sensibilité à la quantité vocalique dans l'acquisition de la langue maternelle; b) l'apprentissage (de l'écriture et) de la lecture; c) le traitement automatique (mise au point de dictionnaires électroniques, de correcteurs orthographiques, de logiciels d'analyse et de synthèse de la parole); d) l'optimisation de la communication par SMS, etc.

Dans la section 2, nous présentons les faits de surface relatifs à la manifestation de la quantité vocalique en kinyarwanda. Dans la section 3, nous proposons une analyse des faits sous-jacents qui essaie de capter la représentation permanente de la quantité vocalique chez les locuteurs. Enfin, dans la section 4, nous discutons les implications orthographiques et nous suggérons des voies de révision de l'orthographe courante en kinyarwanda.

## **2. Les faits de surface relatifs à la quantité vocalique**

### **2.1. La distribution des voyelles brèves et longues**

Commençons par un décompte très simple. Selon les estimations établies par Coupez (1980), les voyelles brèves (a, e, i, o, u) couvrent 75% de l'ensemble des usages vocaliques en kinyarwanda, au moment où leurs contreparties longues représentent 25%. On gardera en tête que la voyelle constitue le sommet de la syllabe en cette langue. Dans un texte kinyarwanda, une voyelle sur quatre est donc réalisée longue. L'asymétrie est manifeste entre voyelles brèves et voyelles longues, mais, si on essaie d'imaginer le nombre total de syllabes qu'il y aurait dans un texte, on comprendra que les voyelles longues sont tout de même massivement présentes. Le décompte est sans doute à vérifier. Or, lorsqu'on sait que le kinyarwanda comporte d'importantes combinaisons consonantiques (Walli-Sagey, 1986), on réalise à quel point le dédoublement d'une voyelle sur quatre dans la notation orthographique viendrait alourdir davantage le texte écrit par la multiplication de signes, et exigerait du coup un décodage visuel ardu (par exemple : *gushwáanyuura* 'déchirer'). Les choses seraient d'autant plus délicates si la tonalité devait aussi en même temps être notée; par

exemple, au lieu de noter *guhora* tout court, on aurait *guhóra*, ‘refroidir’, *guhóra*, ‘venger’ et *guhóra* ‘accuser, condamner pour’. L’opposition *umwĩshywa* ‘neveu, nièce’ vs. *umwĩshywa* ‘espèce d’herbe’ montre qu’avec une voyelle longue, le ton haut peut affecter soit la première, soit la deuxième voyelle; le dédoublement permettrait ainsi tout au moins de contraster le ton montant et le ton descendant (le ton haut double existerait aussi en kinyarwanda).

## 2.2. Le caractère phonémique

Pourquoi s’attacherait-on à visualiser ce genre de distinctions dans l’orthographe? C’est que la quantité vocalique a fondamentalement une valeur oppositive en kinyarwanda (ainsi : *guhaga* ‘pomper’ vs. *guhaga* ‘être rassasié’), et personne ne peut nier, de ce point de vue, que la notation de cette opposition dans l’orthographe est théoriquement de nature à optimiser la reconnaissance des mots lorsque, évidemment, le contexte n’apporte pas d’indices pour la désambiguïsation. Dans l’énoncé : *yariye kuva kare none dore amaze guhaga* ‘Il mange depuis un certain temps, et voilà qu’il vient d’être rassasié’, *guhaga* se lit, s’interprète sans aucun problème comme *guhaga*, étant donné le contexte. Mais alors, on ne saurait naïvement compter sur le rôle du contexte pour décider l’orthographe des mots, pas plus qu’on ne recommanderait de noter la quantité uniquement quand le contexte n’aide pas à la désambiguïsation, démarche qui ne saurait être réellement pratique.

Si nous nous en tenons pour l’instant au plan du lexique, nous nous demanderons si la portée oppositive de la quantité vocalique est une donnée quantitativement significative pour que cela vaille la peine de la marquer dans l’orthographe. Umubyeyi (2000) ne dénombre qu’une vingtaine de paires minimales basées sur la quantité en kinyarwanda, mais cela uniquement au niveau des racines. Dans une communication personnelle, Bizimana (2008) qui a assidûment travaillé sur le dictionnaire kinyarwanda depuis quelques décennies, nous a fourni quelques trente-cinq paires minimales du genre. On pourrait penser que cela ne représente rien au sein de l’ensemble de la langue. En dehors des items lexicaux de base, le système est cependant de loin plus productif : par exemple, le jeu des suffixes qui peuvent s’opposer quant à la quantité contribue à introduire beaucoup plus d’oppositions lexicales (par exemple : *guhũngura* ‘(faire) apparaître, montrer’ vs *guhũngura* ‘(faire) cesser le travail des champs’), qui sont, il n’y a pas de doute, directement listées dans le lexique mental. Le fait de prohiber la notation de la quantité vocalique prive donc le lecteur d’une information capitale, difficile à quantifier en ce qui nous concerne ici.

En plus de l'asymétrie numérique entre les voyelles brèves (75%) et les voyelles longues (25%), nous avons affaire à une autre asymétrie, d'ordre qualitatif, du fait que toutes les voyelles (brèves et longues) ne sont pas impliquées dans des oppositions pertinentes (ainsi, on aura *kurema* mais pas *kureema*, 'créer'; *kureeba* mais pas *kureba*, 'voir, regarder', etc.). Avec de telles 'lacunes' lexicales, cela signifie que si l'on se passe de la notation de la quantité vocalique sur un certain nombre de mots (combien au juste?), cela ne devrait pas poser de problème de confusion. Encore que la quantité vocalique (brève ou longue) fait inaliénablement partie des informations de base sur les mots et que rien ne permet définitivement de le prédire. Autrement dit, le lecteur doit au préalable connaître le mot pour pouvoir lui assigner une quantité, et donc le lire correctement, ce qui est en soi un paradoxe. Dans la liste des mots suivants, qui ne sont pas du tout impliqués dans des oppositions phonologiques dans le système : *kubora* 'pourrir', *gusoma* 'lire', *gusuka* 'verser', *kurega* 'tendre', *gusena* 'polir', *guheka* 'porter sur le dos', qu'est-ce qui pourrait indiquer au lecteur que seuls les trois premiers portent des voyelles brèves, si ce n'est l'information déjà mémorisée? À juste titre, Bizimana (1983) parle de la quantité fortuite. Il est bien clair que l'assignation de la quantité à bien des mots que l'on rencontre pour la première fois en lecture est problématique si ce paramètre n'est pas noté. Nous ouvrons la porte à la recherche de patrons ('patterns') phonétiques et/ou phonologiques qui seraient en œuvre dans la langue et qui commanderaient l'assignation de la quantité par défaut. Le domaine de l'intégration des emprunts venant de langues sans quantité vocalique (par exemple le swahili) serait ici intéressant : pourquoi les mots empruntés au swahili ont-ils tantôt une voyelle longue tantôt une voyelle brève : *ibaati* 'tôle', *ishaati* 'chemise', mais *ingumí* 'poing', *iposhó* 'ration'? Est-ce par une sorte d'analogie à des mots déjà existant en kinyarwanda? Ou ceci est-il ultimement dicté par des patrons réguliers en kinyarwanda, dont le locuteur ordinaire aurait une connaissance intuitive suffisamment précise, auquel cas cela ne vaudrait simplement pas la peine de surcharger l'orthographe par la notation de la quantité vocalique?

### 2.3. La variabilité

Le caractère pertinent de la quantité vocalique en kinyarwanda et l'importance qu'il y a à la répercuter dans l'orthographe sont fortement restreints par un aspect intéressant, à savoir la variation, qu'elle soit idiosyncratique ou dialectale, synchronique ou diachronique. Nous avons en effet des doublets lexicaux qui prouvent que la forme phonologique des mots, quant à la quantité, n'est pas définitivement spécifiée et fixe. En principe, pour les doublets, que la quantité soit notée ou pas, cela ne devrait

pas constituer une entrave majeure. Seulement, en consultant la liste des variations lexicales établie par Kimenyi dans un document électronique (2010), nous sommes tout de même surpris par le fait qu'il n'en dénombre que deux cas : *inzára / inzára* 'ongle', *igipóropóoro / igipóropóoro* 'objet métallique vieux et inutilisable'. Nous avons l'impression qu'on pourrait trouver d'autres cas (par exemple : *indegé / indegé* 'avion', remarquons la redistribution du ton haut), quoiqu'il nous soit quasi impossible d'en donner une estimation quantitative exacte. Ce qui est sûr, par contre, c'est qu'en dehors du champ strictement lexical, la variation dans la réalisation de la quantité est un phénomène qui ne surprend pas (par exemple : *kwíiyanga / kwíiyanga* 'se détester soi-même'; mais *kwáanga* et non *kwánga* 'détester'). Au niveau syntaxique, la variation dans la quantité est aussi un fait notablement caractéristique : ainsi, par exemple *avuga até* 'il parle comment?', va se réaliser phonétiquement, selon les régions, cela est cependant à vérifier, par une voyelle brève ou longue à la frontière entre les deux mots. Encore une fois, l'étendue de la variation dans ce genre de contexte est à regarder d'un peu plus près.

#### 2.4. Le conditionnement

Pour ce qui est du conditionnement, nous enregistrons en kinyarwanda, comme dans la plupart des langues bantu, des cas importants de voyelles longues prévisibles. L'on comprend aisément que, dans ce genre de contextes, l'information sur la quantité de la voyelle se déduit, se devine, se calcule à partir du contexte, et que, par conséquent, il serait logiquement redondant, superflu, de la marquer dans l'orthographe. Par exemple, les mots isolés ne comportent jamais de voyelles longues en position finale ni en position initiale (sans attaque consonantique). L'on a donc pas besoin d'en informer le lecteur. Par contre, les voyelles sont longues en contextes d'allongement vocalique compensatoire (Umubyeyi, 2000), c'est-à-dire devant la combinaison nasale moraique + consonne (*kureemba* 'être gravement malade'), après les consonnes palatalisées (*kubyúuka* 'se lever') et les consonnes vélarisées (*kubwúura* 'dire quelque chose à quelqu'un'). La formulation de généralisations se heurte à quelques problèmes de transparence (sur la notion de transparence des règles en phonologie, voir Kiparsky, 1973).

(1) Certains contextes normalement déclencheurs de la longueur vocalique semblent désactivés. Ainsi, nous dirons *kubyúna* 'danser' et non *kubyúna*; *imbyíro* 'suie' et non *imbyúro*, *impyísí* 'hyène' et non *impyúsí*. Les reconstructions historiques (par Guthrie, Meeussen) montrent que les proto-formes respectives *\*-bín-*, *\*-bído*, *\*pítí* ne laissent pas voir de contextes qui conditionnent l'allongement vocalique. La structure con-sonne + semi-

voyelle n'est qu'un réflexe direct des consonnes /b/ et /p/ devant la voyelle \*/i/ qui en est venue à être responsable de leur palatalisation.

(2) Il est des cas d'allongement vocalique après des consonnes simples qui dérivent diachroniquement de contextes d'allongement compensatoire canonique. Ainsi, *haasí* 'par terre', *kunuuuka* 'dégager une mauvaise odeur', et le radical *-óose* 'tout(e/s)' viennent respectivement de \**páncí*, \**-nunk-* et \**-óncé*. L'on voit que la nasale s'est amuie (et n'a pas simplement migré, nous y reviendrons), ce qui aurait eu pour effet d'allonger la voyelle pour des raisons de sauvegarde du nombre de positions segmentales dans la syllabe. Une explication alternative est qu'on retrouve le contexte nasale moraique + consonne qui aurait aussi déclenché l'allongement vocalique avant d'être occulté par l'histoire. La voyelle longue n'est donc pas fortuite.

(3) Il est d'autres cas d'allongement vocalique dont les données historiques indiquent qu'ils résultent d'un processus de transformation apparenté à l'allongement compensatoire. Dans les exemples *izúko* 'foyer', *gushaaka* 'chercher' et *isuumo* 'chute' dérivés respectivement des radicaux \**gíko*, \**-caka* et \**-cum-*, l'allongement de la voyelle résulte probablement du passage de l'occlusive ou de l'affriquée à la fricative. Les mesures effectuées par Ruzindana (1991) montrent qu'une voyelle qui suit une fricative en position pénultième est plus longue que celle qui suit une occlusive dans la même position. Il y a donc lieu de poser que le contexte de la fricative est dans une certaine mesure déclencheur de la longueur vocalique. Des mesures instrumentales précises s'imposent cependant à ce sujet pour établir si la longueur vient réellement jouer ici un rôle compensatoire dans un contexte où la fricative est articulée de manière raccourcie. Comme pour les deux cas précédents, le conditionnement de la quantité vocalique, s'il était vérifié sur une base beaucoup plus consistante, reste résiduellement enfouie au fin fond de l'histoire de la langue et n'est pas accessible à la conscience des locuteurs. Voilà des cas qu'il est difficile d'envisager aujourd'hui comme participant du conditionnement.

### 3. Les faits sous-jacents

#### 3.1. La théorie

Pour voir plus clair dans les faits de surface, il faut nécessairement une assise théorique. En voici quelques aspects saillants.

(1) Il faut distinguer deux types de représentations : phonétique et phonologique. La phonétique s'occupe des sons du langage dans leur aspect physique, alors que la phonologie examine leur fonctionnement dans le cadre d'un système linguistique particulier. La phonologie requiert un certain degré d'abstraction et débouche sur des représentations qui peuvent se distancer de la manifestation physique immédiate des éléments. Pour

paraphraser Carr (1993, p.30), nous dirons que la représentation phonétique montre les sons que nous produisons, alors que la représentation phonologique, abstraite, est concernée par les sons que nous percevons. Les unités phonologiques sont d'ordre de la perception, et les règles phonologiques expriment les généralisations qui permettent de décoder les signaux phonétiques en termes de ces unités.

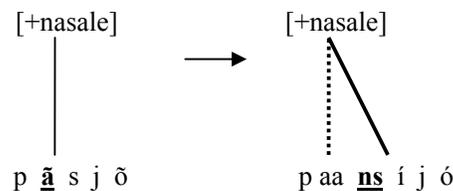
Prenons un exemple simple en français : tout le monde sait que le nominal *décision* dérive du verbe *décider* auquel on adjoint le suffixe *-ion*. Si personne ne peut être tenté de croire que ce nominal vient de *\*déciser*, c'est qu'à un certain niveau de représentation ou de connaissance mentale, *décision* a la forme *décid-ion*. Cela dit, le problème crucial est de pouvoir démontrer que cette analyse reflète une quelconque réalité psychologique. De *décid-ion* et *décision*, quelle forme est enregistrée dans la mémoire au titre de la connaissance permanente du locuteur-auditeur? Donnons aussi l'exemple de la quantité vocalique en kinyarwanda. Lorsqu'on parle de voyelle(s) longue(s), on ramène en réalité toute une série de variations dans la réalisation à une certaine image acoustique invariante. Dans une recherche basée sur des mesures acoustiques, Myers (2005) a établi qu'en kinyarwanda la voyelle longue compensatoire est en réalité moins longue en avant de la complexe nasale moraique + consonne qu'après la complexe consonne + semi-voyelle. Pourtant, en phonologie, on maintient que nous avons affaire à la même chose, en l'occurrence, une voyelle longue que l'auditeur percevra vraisemblablement comme telle.

(2) Ceci nous amène à parler de la structure sous-jacente ou profonde et de la structure de surface, la deuxième dérivant de la première. Le modèle s'inspire de la grammaire générative. En phonologie générative, l'idée est que les sons physiques que nous entendons et transcrivons sont la partie émergente de l'iceberg; en profondeur se retrouvent des formes abstraites de représentation qui permettent la perception des sons et une batterie hautement organisée de règles. Faisant partie de la grammaire générative, la phonologie générative est une approche qui se donne pour objectif de caractériser la connaissance inconsciente que les locuteurs-auditeurs ont de leur langue (Carr, 1993, p.98). De la sorte, nous dirons que la phonétique appartient à la structure de surface tandis que la phonologie est du ressort de la structure sous-jacente.

(3) Revenons à la théorie de la sous-spécification (Archangeli, 1988). Une des possibilités, lorsqu'on reconstitue les formes sous-jacentes en phonologie, est que les propriétés phonétiques d'un morphème ou d'un mot sont différentes de la représentation en structure profonde. Il n'y a pas de correspondance obligée entre l'input (la forme sous-jacente) et l'output (la forme de surface) lors du 'mapping' de celui-là à celui-ci, et la contrainte

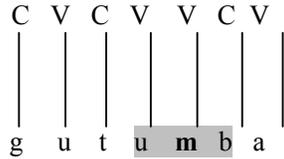
‘Maximiser l’input’ (‘Maximize the input’), pour nous exprimer en termes optimalistes, reste violable. Ainsi, un trait ne sera pas spécifié en structure sous-jacente si sa valeur est prédictible et générée au cours de la dérivation suite à l’application de règles de la grammaire. Sur la base de mesures neuronales, il est connu que la tendance générale du cerveau est d’éviter des représentations prédictibles (Wheeldon & Waksler, 2004). Pour tout dire, la représentation phonologique d’un mot, écrit Carr (1993, p.29), est sa forme une fois dépouillée des caractéristiques phonétiques prédictibles et gouvernées par des règles.

Revenons, à titre d’exemple, sur le cas de la voyelle longue en avant de la complexe nasale moraique + consonne pour expliciter dans quel sens, en structure sous-jacente, elle est analysée comme fondamentalement brève. En phonologie auto-segmentale, on conçoit deux rangées parallèles (‘tiers’) occupées respectivement par des auto-segments à la rangée supérieure et des segments à la rangée inférieure, les deux rangées étant associées par des lignes, et la contrainte majeure étant que celles-ci ne peuvent pas se croiser. Ce modèle phonologique multilinéaire permet des processus connus tels que la propagation (‘spreading’) tonale, la migration de la nasale, l’harmonie consonantique ou vocalique, etc. où l’on voit que des traits phonologiques localisés sur la rangée supérieure, sont indépendants et peuvent justement se propager au-delà des consonnes ou des voyelles individuelles. Par exemple, il est intéressant de remarquer que, du fait que le kinyarwanda n’a pas de voyelles nasales, le trait [+nasale] qui affecte les voyelles dans les mots empruntés au français ne sera pas perdu, mais sera transféré et segmentalement récupéré sur la syllabe contigue pour former une pré-nasale, comme dans [pãsjõ] ‘pension’, qui donne [paansíjó], la voyelle étant alors dénasalisée (pointillées):



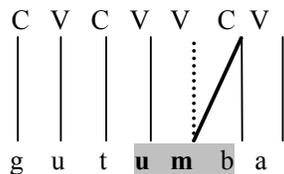
Dans cette perspective, une des approches de l’allongement compensatoire se sert de la structure syllabique canonique couplée aux unités temporelles auto-segmentales C(onsonne) et V(oyelle) (on peut utiliser aussi des positions X) pour montrer que la position V occupée par la nasale est sauvegardée lorsque celle-ci migre.

Voici alors comment on va représenter le mot *gutuumba* ‘gonfler’ en kinyarwanda

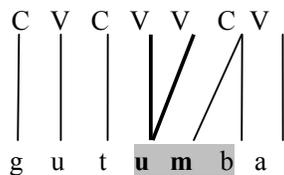


À partir de cette structure sous-jacente où manifestement nous n’avons aucune voyelle longue, les contraintes de syllabification à la base de la réalisation de la structure de surface vont essentiellement faire appel à deux règles : la dissociation et la propagation. D’abord, la nasale moraique (en caractère gras) se détache de l’association avec l’élément V qui la surplombe et migre, pour ainsi dire, pour aller former corps avec la consonne /b/, ceci étant une propriété toute particulière de la nasale, puis la voyelle précédente se propage, s’allonge pour couvrir la position V ainsi laissée vide, la compensation visant ainsi à sauvegarder numériquement la structure syllabique sous-jacente. V, qui est à l’origine associé à la nasale, en vient maintenant à s’associer avec la voyelle /u/ précédente, qui acquiert ipso facto le trait [+long]. Un argument irréfutable en faveur d’une telle analyse est qu’en reduplication partielle à gauche, la copie de la syllabe *tu* est à voyelle brève : *gututuumba* et non *gutuutuumba* ‘gonfler à répétition’ puisque le contexte normalement déclencheur de l’allongement vocalique (la nasale à droite) n’est plus là. Voici ci-dessous la représentation des deux règles :

(1) dissociation en (pointillées) et ‘migration’ de la nasale (de V à C)



(2) allongement vocalique compensatoire



### 3.2. Le statut de la quantité vocalique

Dans sa typologie des langues bantu, Hubbard (1995) distingue quatre grands groupes:

(1) Les langues à voyelles brèves sans aucune forme d'allongement ; toutes les syllabes sont donc monomoraïques, comme en kilega (Meeussen, 1974) : *kumana* 'appeler' > /ku-ma-na/ ; *mwana* 'enfant' > /mwa-na/.

(2) Les langues qui attestent une longueur vocalique, soit canonique, en d'autres termes, sous-jacente (l'allongement affecte une seule syllabe qui compte deux moras), soit transmoraique (la longueur affecte particulièrement la voyelle qui précède la séquence nasale + consonne; elle affecte à la fois la syllabe dans laquelle figure la voyelle allongée et celle dans laquelle figure la nasale qui migre, chacune d'elles étant partiellement affectée, voir notre analyse auto-segmentale précédente). L'exemple du luganda (Hubbard, 1995) est illustratif : *omuntu* 'personne' > /o-mun-tu/ > o-muu-ntu.

(3) Les langues à voyelles pseudo-longues ; il ne s'agit pas de longueur vocalique à proprement parler étant donné que, dans ces langues, deux voyelles successives (ho-morganiques ou hétéroorganiques) dans la syllabe s'articulent en séquence sous forme de diphtongue. La syllabe n'est donc pas bimoraïque mais est plutôt faite d'une séquence vocalique bisyllabique puisque les deux voyelles sont rattachées à deux syllabes différentes, chacune d'elles constituant une syllabe entière à part. Le swahili en est un exemple : *maua* 'fleurs' > ma-u-a.

(4) Les langues qui combinent deux caractéristiques ou plus parmi (1), (2) et (3). On notera qu'une langue est susceptible d'évoluer d'un type à l'autre. Par exemple, le swahili qui ne connaît pas de longueur vocalique est tout de même indéniablement issu du proto-bantu où ce trait est déclaré.

Hyman (2003, pp.48-49), repris par Stegen (2005) et par Schroeder (2008), montre que la longueur vocalique tire son origine de cinq contextes dans les langues bantu :

(1) La voyelle longue peut être sous-jacente (sans être nécessairement membre de l'opposition voyelle brève / voyelle longue, comme en kinyarwanda).

(2) Elle peut être issue de la concaténation de deux voyelles (par-dessus la frontière entre les morphèmes ou suite à l'effacement d'une consonne intervocalique comme en tonga, voir Maddieson, 2003, p.38).

(3) Elle peut naître de la semi-vocalisation suivie de la procédure d'allongement compensatoire.

(4) Elle peut être enclenchée en avant d'une nasale moraique + une consonne.

(5) Elle peut être réalisée en position pénultième dans la plupart des langues bantu de l'Est et du Sud qui ne connaissent plus le contraste voyelle brève / voyelle longue (selon Anyanwu 2001, l'allongement vocalique à la pénultième n'existe pas au niveau du mot mais au niveau du syntagme, caractérise certains types de phrases, et n'a pas d'effet sur le système prosodique des langues concernées).

On remarque que dans certaines langues, comme le kimatuumbi (Odden, 1996), on peut avoir des oppositions significatives dans des contextes de prénasalisation, de labialisation et de palatalisation. On note aussi que d'un point de vue acoustique, la réalisation de la voyelle longue peut connaître des variations qui reflètent son origine et son statut. Nous avons vu le cas du kinyarwanda (Myers, 2003) où la prénasalisation des consonnes enclenche une longueur physiquement moins longue que la semi-vocalisation. On observe dans des langues comme le sukuma par exemple (Maddieson, 2003, p.37) que la voyelle longue conditionnée par le contexte phonologique est réalisée en quelque sorte à mi-chemin entre la voyelles brève et la voyelle longue sous-jacentes. La sensibilité des locuteurs-auditeurs est donc bien réelle.

Hyman montre, d'autre part, que la voyelle longue est abrégée ou neutralisée dans des contextes précis, ce qui est un cas de conditionnement : a) elle est bloquée au début et à la fin du mot ou du syntagme, comme nous l'avons indiqué pour le kinyarwanda ; b) elle est limitée et réalisée seulement en position pénultième ou antépénultième, comme en mwiini ; c) elle est abrégée dans des syllabes fermées comme, par exemple, en luganda en avant d'une consonne géminée. Ces contraintes sont importantes puisqu'elles permettent de mieux circonscrire le domaine d'application et de prévisibilité de la quantité vocalique.

#### **4. Les choix orthographiques**

##### **4.1. Les principes**

Le concept de développement de l'orthographe d'une langue réfère à un processus qui doit être basé sur un certain nombre de principes fondamentaux que Schroeder (2008) expose avec clarté. Nous les dégagerons rapidement :

(1) L'orthographe est la propriété de la communauté, et les usagers sont à associer à toutes les étapes, de sa conception à son implantation afin de s'assurer pleinement de leur réceptivité et d'encourager leur compréhension des décisions.

(2) L'analyse linguistique est centrale, et il faut rester en interaction avec les linguistes pour se rendre compte des défis qu'ils rencontrent dans leur travail et pour s'informer du savoir-faire qui découle de leurs résultats.

- (3) La perception et les préférences des locuteurs natifs doivent jouer un rôle significatif dans la prise des décisions.
- (4) L'usage effectif de l'orthographe est à la fois le but, la source de feedback et un précieux outil d'évaluation.
- (5) La performance des usagers dans la lecture et l'écriture (testée par rapport à la facilité et à la rapidité de décodage) est à considérer tout le long du processus.
- (6) Dynamique, l'orthographe doit faire l'objet de constantes révisions qui reflètent le contexte politique, éducationnel et social, qui est lui aussi changeant.

#### **4.2 Les choix**

En ce qui concerne la représentation graphique de la longueur vocalique dans les langues bantu, le constat est que les pratiques varient considérablement d'une langue à une autre (ou d'un pays à l'autre), et ceci peut être dû en partie à la différence de configuration de ce paramètre au sein des différents systèmes linguistiques, mais aussi, certainement, au manque d'un cadre de concertation. Partant de l'origine de la longueur vocalique, Stegen (2005) et Schroeder (2008) émettent des recommandations qui se résument en trois points essentiels :

- (1) Dans les langues (ou les contextes?) où la quantité vocalique est phonémique, on adoptera le dédoublement de la lettre pour représenter la voyelle longue par opposition à la lettre unique qui représente la voyelle brève.
- (2) La voyelle longue issue de la concaténation de voyelles (transmorphémiques ou par dessus la frontière entre les mots) tend ou devrait tendre à être notée par une double lettre également.
- (3) La voyelle longue compensatoire sera notée par une lettre unique si elle est perçue en affinité avec la voyelle brève (sous-jacente) ou par une double lettre si sa perception est en affinité avec la voyelle longue (en particulier dans les langues où les contextes d'allongement compensatoire sont compatibles avec la valeur contrastive); ceci n'est malheureusement pas encore une pratique que l'on pourrait qualifier de standard dans l'orthographe bantu.

On ne peut juger de l'efficacité d'un tel système que sur la base de chaque langue individuellement, après une recherche expérimentale sérieuse et en tenant compte de l'expérience et de la performance des usagers. La question qui resurgit est de savoir sous quel format l'information phonologique est précisément encodée dans le cerveau des locuteurs-auditeurs et avec quel degré de transparence leur connaissance est exécutée

pour produire des formes phonétiques de surface. Le problème est délicat; on notera par exemple il y a un peu d'hésitation aux points (2) et (3).

L'implantation de la notation de la longueur vocalique requiert de la patience pour que les usagers en arrivent à se familiariser. Pour Schroeder (2008), les gens en général n'ont pas d'objection à ce que la longueur vocalique soit représentée dès qu'ils réalisent que la quantité joue un rôle contrastif important dans leur langue. Au début, l'usage de la double lettre va sembler rébarbatif même pour les gens avec une culture avancée et va affecter la capacité d'épeler dans la langue maternelle. C'est pourquoi les règles d'épellation gagneront à être enseignées et pratiquées systématiquement aussitôt que l'orthographe est adoptée. Avec le temps, les lecteurs et scripteurs se rendront compte de l'efficacité de la notation de la quantité vocalique dans l'accès lexical des mots.

#### **4.2. Un modèle pour l'orthographe kinyarwanda**

S'agissant de la notation de la longueur vocalique en kinyarwanda, les faits que nous avons présentés ne nous permettent pas de dire de manière définitive ce qu'il convient de faire pour rationaliser le développement de l'orthographe. Voici cependant quelques idées générales :

- (1) Vingt-quatre ans se sont écoulés depuis le lancement de la dernière loi sur l'orthographe du kinyarwanda; il est très urgent de revisiter cette loi au regard d'un certain nombre d'impératifs socio-culturels, politiques, éducationnels, en l'occurrence: a) le contact accru du kinyarwanda avec d'autres langues (y a-t-il, par exemple, des interférences chez les sujets bilingues kinyarwanda-swahili puisqu'on sait que cette dernière langue n'a pas de quantité vocalique?), b) la nécessité d'harmonisation de l'orthographe dans la Communauté Est-Africaine (par exemple, quelle est la pratique pour le kirundi voisin? sait-on que l'orthographe standard du ganda qui date de 1947 prescrit de ne pas noter la voyelle longue par une double lettre uniquement dans les contextes de début et de fin de mot, sauf, comme en kinyarwanda, pour le mot *yee* 'oui'? (<http://www.uganda.com/luganda.htm>), etc.), et c) les perspectives de traitement automatique du kinyarwanda, etc.
- (2) Toute prise de décision en matière d'orthographe doit se laisser guider par l'analyse linguistique du phénomène d'allongement vocalique au point de vue de son origine, de son statut et de sa portée significative.
- (3) Un travail d'évaluation de la rentabilité du système orthographique tel qu'il est pratiqué s'impose, et, pour ce faire, des tests psycholinguistiques sont nécessaires pour mieux faire ressortir la performance des usagers.
- (4) La mise au point d'une prochaine version de la loi sur l'orthographe doit inclure des stratégies fermes d'implantation et de suivi.

Examinons d'un peu plus près l'article 2 de l'instruction ministérielle qui proscrit le dédoublement des voyelles. Nous donnerons quelques orientations :

(1) Il s'avère bénéfique et économique de trouver une façon de faire qui s'applique pareillement à l'orthographe courante et à l'orthographe scientifique; le dédoublement du système est incongru, et il est des domaines que l'on aurait du mal à catégoriser (par exemple les manuels de grammaire, les dictionnaires, l'instruction ministérielle fixant l'orthographe, etc. sont-ils ou non 'scientifiques'?).

(2) Les cas d'allongement vocalique pour des raisons de simple expressivité ne devraient pas être notés dans l'orthographe puisqu'ils ne font partie en aucune façon de l'information phonologique de base sur les mots, mais aussi vu leur grande variabilité; on pourrait tout au plus en prendre compte dans des situations 'marginales', comme dans les bandes dessinées, par exemple.

(3) Étant donné la grande récurrence de voyelles longues en kinyarwanda (25% des usages vocaliques), il est encombrant de vouloir tout noter; on opérera donc un choix, le plus judicieux possible, en fonction de la valeur oppositive de la quantité et en tenant compte d'autres facteurs qui influent sur l'accès lexical, tels que la fréquence des items (un travail patient pourrait amener à identifier les mots, y compris les emprunts, ou les contextes qui sont le plus fréquemment sujets à l'ambiguïté), l'interaction avec la tonalité, etc.

(4) Les cas d'allongement vocalique compensatoire doivent être traités comme des voyelles (sous-jacentes) brèves dans la mesure où l'instruction phonétique relative à l'allongement est contextuellement aisément recouvrable.

(5) Les cas de variation dans l'assignation de la quantité vocalique ne posent pas de problème particulier et ne nécessitent pas d'être notés.

(6) La notation de l'allongement vocalique se fera à l'aide d'une double lettre.

## 5. Conclusion

La question concernant l'orthographe se situe au cœur des efforts à déployer dans le cadre de la promotion des langues bantou et doit, de ce fait, être intégrée dans la politique culturelle générale et, en particulier, dans la préoccupation de la recherche scientifique. Si nous avons décidé de nous interroger sur les raisons et l'impact de la non-orthographe de la longueur vocalique en kinyarwanda, nous étions bien conscient que le problème était déjà connu des linguistes et des décideurs politiques. Mais ce qui est nouveau dans l'approche que nous préconisons, c'est que les locuteurs-auditeurs (scripteurs-lecteurs) sont des acteurs qui doivent être associés dans

la conception, l'implantation et le suivi de la pratique orthographique. Dans le modèle que nous proposons, les usagers de l'orthographe doivent être testés, consultés quant à leurs préférences et, une fois l'orthographe décidée, entraînés à une meilleure maîtrise du système.

La quantité vocalique est un trait phonétique et phonologique important en kinyarwanda en raison de la centralité de la voyelle au sein de la syllabe dans cette langue; et l'on sait qu'il se peut fort bien que, du point de vue de la perception des sons du langage, la syllabe soit l'unité de base, comme nombre de travaux psycholinguistiques l'ont suggéré. La présence massive des voyelles longues (25%) dans un texte kinyarwanda, leur ambivalence vu qu'elles sont à la fois phonémiques et conditionnées par le contexte, et le fait qu'elles sont soumises à la variation tant en synchronie qu'en diachronie révèlent la complexité du système vocalique. L'on ne sait pas encore au juste la nature exacte des informations phonologiques que possède le locuteur-auditeur en structure sous-jacente et l'ampleur des règles de dérivation en activité lors de la réalisation des énoncés en structure de surface. Or, si l'on accepte que l'orthographe ne peut pas être conçue uniquement en fonction de la connaissance phonologique implicite des usagers, connaissance qui peut être tout aussi abstraite qu'opaque, du moins l'on doit le plus possible s'en approcher. Nous avons essayé de discuter la configuration de la longueur vocalique en kinyarwanda pour lancer un appel à la reconsidération de l'instruction ministérielle de 1985 qui prohibe inconditionnellement le dédoublement des lettres comme mode de sa notation. Le travail à faire, à notre avis, peut se résumer à la recherche et à l'évaluation systématique des contextes où la quantité vocalique prête à confusion en kinyarwanda afin de modéliser l'orthographe en minimisant le coût et en maximisant le rendement .

À part les aspects proprement techniques de la démarche linguistique que nous entrevoyons par rapport à la longueur vocalique, il est important de penser à élargir les horizons, notamment : a) en explorant l'ensemble des éléments de l'orthographe du kinyarwanda (pensons par exemple au système tonal que nous avons déjà évoqué en interaction avec la quantité vocalique); b) en créant une synergie dans le traitement de l'orthographe des langues bantu en vue de répondre à la nécessité d'harmonisation des pratiques suivant les langues et les pays ou alors les sous-régions, et enfin c) en mariant les faits de langue aux nouvelles technologies dédiées au traitement de l'information.

## Références

- Anyanwu, R.J., 2001. On the manifestation of stress in African languages. *Typology of african posodic systems workshop*. Bielefeld University, 18-20 May.
- Archangeli, D., 1988. Aspects of underspecification theory. *Phonology*, 5, pp.183-207.
- Bickmore, L.S., 1995. Accounting for compensatory lengthening in CV and moraic frameworks. In J. Durand & F. Katamba, ed. *Frontiers of phonology, atoms, structures, derivation*. London & New York : Longman, pp.119-148
- Bizimana, S., 1983. La quantité vocalique en rwanda. In F. Jouannet, ed. *Le kinyar-wanda, langue bantu du Rwanda, études linguistiques*. Paris : GERLA – SELAF, pp.33-53.
- Carr, P., 1993. *Phonology*. London : Macmillan Press Ltd.
- Clements, G. N., 1982. *Compensatory lengthening an independent mechanism of phonological change*. Bloomington : Indiana University Linguistics Club.
- Clements, G. N., 1986. *Compensatory lengthening and consonant gemination in luganda*. In L. Wetzels & E. Sezer, ed., *Studies in compensatory lengthening*, Dordrecht : Foris Publications, pp.37-57.
- Coupez, A., 1980. *Abrégé de grammaire rwanda*. Butare : Institut National de la Recherche Scientifique.
- Coupez, A., 1980. Trace de dix voyelles en proto-bantu. *Annales du Musée Royal de l'Afrique Centrale*. Tervuren : Musée Royal de l'Afrique Centrale, p.66-73.
- Davitskaya, D., 2002. *Compensatory lengthening : phonetics, phonology, diachrony*. London & New York : Routledge.
- Gerla, 1985. *Notation orthographique du kinyarwanda*. Actes de colloque (3-5 juin 85). Ruhengeri : Université Nationale du Rwanda.
- Goldsmith, J.A., 1979. *Autosegmental phonology*. New York : Garland.
- Hayes, B., 1989. Compensatory lengthening in moraic phonology. *Linguistic inquiry*, 20(2), pp.253-306.
- Halde, J.I, 1990. Compensatory lengthening in Friulian. *Probus*, 2, pp.31-46.
- Hubbard, K., 1995. Transmorphemic compensatory lengthening in Chichewa. *Papers of the 31<sup>st</sup> Regional Meeting of the Chicago Linguistic Society*, 1, pp.193-204.
- Hyman, L.M., 2003. Segmental phonology. In D. Nurse & G. Phillipson, ed. *The ba-ntu languages*. London & New York: Routledge, pp.42-58.

- Hyman L.M. & Katamba F., 1990. Final vowel shortening in luganda. *Studies in African Linguistics*, 21(1), pp.1-59.
- Goldsmith J.A., 1990. *Autosegmental & metrical phonology*. Cambridge : Basil Blackwell.
- Jouannet, F., 1985. *Prosodologie et phonologie non linéaire*, Paris : ACCT.
- Kimenyi, A. *Lexical variation in kinyarwanda : grammaticization at work*. Disponible sur : <http://www.kimenyi.com/lexical-variation.php>. [Consulté le 27 juillet 2010].
- Kimenyi, A., 1983, Faut-il adopter une orthographe phonétique en kinyarwanda. *Education Sciences et Culture*, 5, pp.35-40.
- Kiparsky P, 1973. *Abstractness, opacity and global rules*. Bloomington : Indiana University Linguistic Club. Repris in O. Fujimura, ed. *Three dimensions of linguistic theory*, Tokyo: TEC, pp.57-86.
- Maddieson, I., 2003, The Sound of the bantu languages. In D. Nurse & G. Philippson, ed. *The bantu languages*. London & New York: Routledge, pp.15-41.
- Meeussen, A. E., 1967. Bantu grammatical reconstructions. *Africana Linguistica* III. Tervuren : Musée Royal de l'Afrique Centrale, pp.81-121.
- MEEUSSEN, A. E., 1969. *Bantu lexical reconstructions*, promanuscripto.
- MINEDUC, 1978. *Commission sur l'orthographe officielle du Kinyarwanda, Synthèse des résolutions*. Kigali : DPS.
- Myers, S., 2005. Vowel duration and neutralization of long vowel length contrasts in kinyarwanda. *Journal of phonetics*, 33(4), pp.427-446.
- Myers, S., 2003. F0 in Kinyarwanda. *Phonetica*, 60(2), pp.71-97.
- Myers, S., 1995. The phonological word in Shona. In J. Durand & F. Katamba, *Frontiers of phonology, atoms, structures, derivation*. New York : Longman, pp.69-92.
- Odden, R., 1996. *The phonology and morphology of Kimatuumbi*. Oxford: Clarendon
- Roca, I., 1994. *Generative phonology*. London & New York: Routledge.
- Rosenthal, S., 1997, The distribution of prevocalic vowels. *Natural language and linguistic theory*, 15(1), p.139-180.
- Ruzindana, M., 1991. *Interlanguage study of vowel duration in the advanced kinyarwanda speakers of English*. Thèse de doctorat. University of Reading.
- Schroeder, Leila. 2010. *Bantu orthography manual*. rev. ed.. Dallas: SIL International e-Books, 9. Disponible sur <http://www.sil.org/silepubs/abstract.asp?id=52716>. [Consulté le 27 juillet 2010]
- Stegen, O., 2005, Vowel length in Eastern bantu orthographies, assignment for *SIL Bantu orthography meeting*. Dallas: SIL International, 7-12 May.

- Troubetzkoy, N. S., 1967. *Principes de phonologie*. (Traduit de l'allemand). Paris: PUF.
- Umubyeyi, M.J., 2002. *L'allongement vocalique compensatoire en kinyarwanda : étude structurale*. Mémoire de licence. Butare : Université Nationale du Rwanda.
- Walli-Sagey, E., 1986. On the representation of complex segments and their formation in kinyarwanda. In L. Wetzels & E. Sezer, ed. *Studies in compensatory lengthening*, Dordrecht : Foris Publications, p.251-295.
- Wheeldon, L. & Waksler, R., 2004. Phonological underspecification and mapping mechanisms in speech recognition of lexicon. *Brain and language*, 90(1-3), pp. 401-412.
- <http://www.buganda.com/luganda.htm> [Consulté le 27 juillet 2010]