

DICTIONNAIRES EN LANGUES AFRICAINES: LE LOGICIEL BOSOLAN ET LA REDACTION D'ARTICLES LEXICOGRAPHIQUES EN BAMANANKAN

Issiaka BALLO

Université des Lettres et des Sciences Humaines de Bamako (ULSHB)

issiakaballo79@gmail.com

Adama Diokolo Coulibaly

Académie Malienne des Langues (AMALAN)

Résumé

Les outils de conception des répertoires lexicographiques, surtout les dictionnaires, répondent de plus en plus aux exigences de l'ère du numérique. Les bristols sont vite remplacés par les fichiers informatisés (Lehmann 2018 :255) pour la consignation provisoire des articles du dictionnaire. Vu que cela coïncide avec le moment où les langues africaines forcent leur petit bonhomme de chemin dans le concert de la lexicographie monolingue, il est essentiel de créer des outils de rédaction lexicographiques en leur nom. En réponse donc à ce constat, un logiciel de dépouillement de corpus lexicographique dénommé Bosolan vient d'être conçu et réalisé en bamanankan (bambara). Ledit logiciel est muni d'une option (Dajeboso) explicitement développée pour la rédaction d'articles lexicographiques en bamanankan. Donc, des tâches comme l'extraction des occurrences du corpus, le classement en ordre alphabétique des entrées, l'identification de l'indicatif de grammaire de l'entrée, la détermination de la rection des verbes, et la description des autres rubriques de l'article facilitent la rédaction dictionnaire en bamanankan à travers Bosolan. Le logiciel est conçu de sorte qu'il soit adaptable aux autres langues africaines. Son test grandeur nature a pu concevoir plus d'une centaine d'articles comprenant toutes les catégories lexicales qui existent en bamanankan.

Mots-clés : *bamanankan, corpus, dictionnaire, lexicographie, logiciel*

Introduction

La conception de dictionnaires peut être considérée comme une innovation dans le processus de la promotion de la plupart des langues africaines selon l'étape à laquelle ces langues se trouvent dans le processus d'aménagement linguistique. On en compte très peu, surtout les dictionnaires monolingues. Cet état de fait suscite de plus en plus une réaction de la part des chercheurs qui travaillent sur ces langues. Alors, qui parle de dictionnaire parle avant tout de lexicographe. Il se trouve que les lexicographes qui s'intéressent à la lexicographie des langues africaines commencent à se rendre compte de l'inexistence des outils appropriés de traitement lexicographique dans lesdites langues. Une chose est de faire recours aux outils lexicographiques taillés sur mesure pour les langues européennes pour faire les traitements lexicographiques des langues africaines. Mais, l'accommodation des pertinences linguistiques propres aux langues africaines demeure un grand obstacle à

ces traitements dans des outils inappropriés. A titre illustratif, le classement par ordre alphabétique selon la langue africaine respective n'est pas toujours automatique dans un outil étranger aux mécanismes internes de la langue. Les traitements relatifs aux différentes rubriques de l'article ne trouveront pas toujours d'options qui répondent valablement dans un outil qui n'est pas fait sur mesure.

Après ces constats, l'outil sur lequel le présent exposé porte a été mis sur pied pour cibler en premier lieu une langue africaine parlée au Mali : le bamanankan.

L'objectif de l'exposé est de faire part d'un progrès réalisé en matière de ressources lexicographiques numériques en langues africaines. Vu que ladite ressource est conçue pour le dépouillement lexicographique, il est important de rendre compte au public le mécanisme d'exploitation de l'outil.

Cependant, le cadre théorique exploité par le logiciel Bosolan couvre les règles lexicographiques que sont l'exploitation du corpus de dépouillement, le dictionnaire monolingue, la macrostructure du dictionnaire et la microstructure de l'article.

La question du corpus est abordée par Polguère (2008 :107) sous le terme *accès aux données linguistiques*. Il a mis un accent particulier sur l'exploration des corpus pour résoudre des problèmes langagiers (p.109). Le corpus est défini comme étant un « *ensemble limité des éléments (énoncés) sur lesquels se base l'étude d'un phénomène linguistique* » (Grand Robert électronique 2005). Alors, qui parle de corpus parle de nombre d'occurrences qui s'y trouve. Comme définition du mot occurrence, Alain (Polguère 2008 :108) formule la suivante : *On appelle occurrence d'un élément linguistique dans un corpus donnée une instance d'utilisation de cet élément dans le corpus en question* ».

Concernant la théorie sur le dictionnaire monolingue, dictionnaire cible de Bosolan, Alise (Lehmann 2018 :257) aborde la question en l'opposant au dictionnaire bilingue. Comme ces deux types sont inclus dans la notion de dictionnaire au même titre que plusieurs autres types, Polguère (2008 :226) formule la définition suivante pour le dictionnaire : « *un dictionnaire d'une langue donnée est un répertoire du lexique de cette langue qui offre une description de chaque lexie selon un patron relativement rigide* ». Cependant, un dictionnaire est l'aboutissement d'une pratique dénommée lexicographie. En ces termes, Lehmann (2018 :255) dit que la lexicographie est: « *La confection des dictionnaires : choix des unités lexicales à traiter, méthode de leur description, techniques de présentation, en vue de la publication* ».

Par ailleurs, la confection d'un dictionnaire demande d'abord une macrostructure du dictionnaire et enfin une microstructure de l'article. Selon Lehmann (2018 :267), « *la macrostructure est la structure de la nomenclature (ou ensemble des entrées), constituée et délimitée par les critères du choix des entrées* ». La macrostructure concerne aussi l'ordre

d'agencement des entrées et le choix fait par rapport au regroupement ou dégroupement des entrées homonymiques (Lehmann 2018 :277).

Pour la microstructure des articles du dictionnaire, elle est l'ensemble des rubriques programmées dans lesquelles la description plus ou moins complète du mot respectif est fournie. Lehmann (2018 :283) a regroupé la plupart des rubriques dans les mots-clés suivants : bloc-entrée, les marques, la définition, l'exemple, les relations lexicales. C'est au lexicographe de retenir les rubriques qui concernent son projet de dictionnaire. L'ensemble s'appelle l'article de dictionnaire que Polguère définit comme étant « *le bloc de texte décrivant un vocable donné* » (p.233). Au sujet de la rubrique définition, l'ordre interne des acceptions attestées chez une entrée polysémique est une fois de plus une question de choix entre 4 différents ordres : l'ordre historique, logique, l'ordre de fréquence et l'ordre linéaire (Lehmann 2018 :302).

Alors, il est bien clair que bosolan est l'outil permettant la réalisation d'articles de dictionnaire au moins en bamanankan. Mais qu'en est-il de la portée sociale de ce nouvel outil ?

La portée sociale de bosolan se mesure à l'appropriation que les particuliers, les sociétés civiles ou politiques vont en faire dans le traitement des langues africaines. Vu que la demande (Tourneux 2022 : 167) a précédé à l'offre (Bosolan), l'usage du logiciel bosolan produira de l'effet domino dans les secteurs de la politique linguistique et de l'aménagement du corpus des langues qu'il servira à traiter. On s'attend donc à observer une production massive et conséquente de dictionnaires monolingues en langues africaines, surtout en mandenkan (bamanankan, Dioula, maninkakan...). En effet, les chercheurs et les institutions en charge des langues africaines ont toujours exprimés les besoins de produire des dictionnaires en langues africaines. Donc, les effets susceptibles d'être produits par l'appropriation du logiciel bosolan sont énormes. Dans la même optique parmi les options choisies par les politiques linguistiques des pays africains, il existe la prise des langues des terroirs africains comme langue de travail dans l'administration. Ainsi, l'aménagement du corpus desdites langues se trouve être au centre de cette option. En synthèse, l'impact majeur que l'appropriation de l'outil bosolan pourra engendrer dans les sociétés africaines se sentira dans le domaine des productions lexicographiques monolingues.

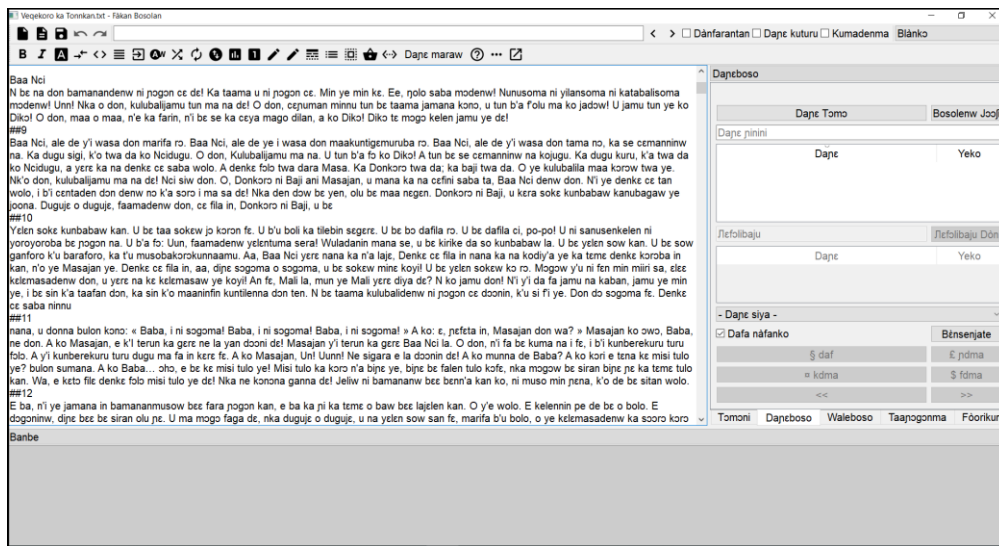
Sur un tout autre plan, le présent exposé relate le cheminement systématique que le logiciel Bosolan offre au lexicographe dans le cadre de la confection de son dictionnaire. L'exposé montre combien Bosolan peut faire le travail et combien cette innovation est une solution fiable pour la lexicographie au moins en bamanankan.

1. Point sur le logiciel Bosolan

IL est question d'un nouvel outil numérique de traitement lexicographique dans notre exposé. Qu'est-ce que c'est ? La question est globalement

répondue dans la citation suivante. La citation expose un préambule assez complet sur le logiciel. Elle est tirée d'un article ayant délivré une ébauche sur le logiciel (Ballo 2022:91).

Le programme Bosolan (...) est un logiciel de traitement lexicographique que nous avons spécifiquement mis au point (Ballo, 2016) pour le bamanankan. Il est tout aussi adaptable aux langues africaines en plus de l'anglais et le français. Le logiciel est hors ligne. Il doit être installé sur des ordinateurs contenant Windows. Son interface d'utilisation est en bamanankan. Le menu y est assez fourni. Il comporte plus d'une trentaine d'options. Sa fonction principale est le traitement lexicographique simple et complexe. De ce fait, Bosolan est muni de plusieurs rubriques. Parmi les plus remarquables, il y a Tòmɔni (extraction simple ou composite des occurrences d'un corpus), Dapɛboso (rédaction de dictionnaire), Waleboso (traitement des verbes d'un corpus), Sigannadapɛ (filtre pour l'édition semi-automatique), Taapɔgɔnma (traitement des unités corrélatives du corpus).



Capture d'écran de la page d'accueil du logiciel Bosolan avec un affichage de corpus.

2. Bosolan : comment marche-t-il ?

Comme on le dit souvent, pour travailler avec un outil, il te faut se le procurer au préalable. Un outil informatique est destiné à être installé sur un gadget qui le reconnaît : ordinateur, tablette, téléphone, iPhone... Donc, comme indiqué dans la citation ci-haut, Bosolan est à installer sur l'ordinateur. Une fois installé sur le gadget, le double-clic d'ouverture habituelle lance l'ouverture d'une page assez réduite qui vient s'afficher au beau milieu de l'écran. Cette page invite l'utilisateur à renseigner le formulaire Tɔgɔ (Nom). L'utilisateur renseigne ledit formulaire avec son

Nom à l'intégral ou en abrégé. A la suite, il clique sur le bouton Co (Ok) pour ainsi voir la page des menus s'ouvrir.

En ce qui concerne le menu, Bosolan possède 5 barres déjà bien fournies en options en plus de la page d'affichage du contenu du fichier actif. Les 3 premières barres sont horizontalement superposées en haut de l'écran sous lesquels la page d'affichage occupe la majeure partie de l'écran. La 4^{ème} barre, toujours horizontale, occupe, en bas de l'écran et sous la page d'affichage, un espace un peu plus large que les 3 premières réunies. La 5^{ème} et dernière barre de la présente version de Bosolan occupe verticalement la partie droite de l'écran dans les limites de la page d'affichage. Cela sous-entend que ladite barre occupe par défaut un quart de l'espace de la page d'affichage. Chaque barre possède des indications de tâches spécifiques ou d'information. On énumère ces indications de la manière suivante.

A la toute première barre, la plus haute des trois superposées en haut de l'écran, correspond dans l'angle gauche de l'écran l'icône et l'enseigne « Fàkan Bosolan » pour signifier l'appartenance de l'outil à l'entreprise qui l'a réalisé. Les autres indications sur cette barre se suivent dans l'angle droit de l'écran. Ces indications arrangées de gauche à droite sont respectivement l'icône de réduction sur la barre des tâches de la page affichée, celle de réduire la page à la moyenne et enfin celle d'agrandir la page à tout l'écran.

A la deuxième barre superposée correspondent plusieurs options sur toute la ligne. Pour ne citer que quelques-uns, ces options commencent à gauche par trois icônes principales : kura (nouvelle page), Dayelen (Ouvrir), Lasago (Enregistrer). L'option *Dayelen* porte toute son importance dans le repérage des fichiers à afficher dans la page d'affichage (infra sous-titre 3).

La troisième barre, sous les deux autres dans leur superposition à trois, est la bande sur laquelle les options d'édition et de mise en forme du corpus se juxtaposent de gauche à droite. La barre abrite également d'autres options qui exécutent des tâches spécifiques.

La page vierge qui s'affiche par défaut sous la barre d'édition fait office d'espace d'affichage du texte à traiter. Sous cet espace se trouve la bande *Banbe* (Console). Elle sert d'espace d'affichage pour les tâches ordinaires exécutées lors du traitement du corpus.

La 5^{ème} barre, qui est une bande verticale à droite de l'écran, comporte des sous rubriques de traitement du corpus lexicographique. Ces sous rubriques dont chacune exécute une tâche quelconque dans le dépouillement lexicographique sont au nombre de 5 : tòmòni (extraction des unités), dapɛbosò (rédaction lexicographique), walebosò (analyse des verbes), taapɔgɔnma (unité corrélatrice), fòorikun (identification des limites des syntagmes dans l'énoncé). Lorsqu'on lance chaque sous rubrique de la bande, ses options de traitement s'affichent sur lesquelles un clic permet de lancer la tâche respective.

Par ailleurs, de toutes les options du logiciel Bosolan, celle sur laquelle porte le présent exposé est la rédaction lexicographique (Dajɛboso). La section suivante nous édifie davantage sur cette option.

3. Comment Bosolan procède-t-il à l'extraction des occurrences du corpus

Bosolan applique l'extraction des occurrences ciblées sur un corpus (infra sous-titre 7.2). De ce fait, il faut dire que le logiciel est doté d'un corpus de base. Le contenu de son corpus est extensible au fur et à mesure que des textes supplémentaires se rendent disponibles. Le clic sur l'option *Dayɛɛn* permet de parcourir l'emplacement des fichiers dans l'ordinateur pour atteindre le dossier système « bosoli_daga ». Ce dossier est fait sur mesure pour ne contenir que les fichiers textes composant le corpus.

Le contenu du fichier choisi dans *bosoli_daga* s'affiche sur la page d'affichage du logiciel. Une fois le texte sur cette page, la sous rubrique Dajɛboso de la barre verticale est indiquée pour continuer le processus d'extraction des occurrences ciblées du texte affiché.

L'opérateur clique sur Dajɛboso pour découvrir ses différentes options de traitement. Parmi ses options, le bouton « Dajɛ Tɔmɔ » constitue la première étape de l'extraction des occurrences du texte. Un clic sur le bouton génère la liste de chaque occurrence unique du texte. La liste s'affiche dans une page réduite sous l'option *Dajɛ Tɔmɔ* avec la marge « Dajɛ ninini » intercalée entre elles. Le logiciel classe par défaut ses occurrences répertoriées en ordre alphabétique croissant. Il s'agit bien de l'ordre alphabétique bamanan.

L'espace d'affichage de la liste des occurrences est en deux colonnes : la colonne des unités extraites avec comme entête « Dajɛ » à gauche et le nombre de la fréquence d'occurrence « Yeko » devant chaque unité dans la colonne droite. Un clic sur l'entête de chaque colonne reclasse le contenu de la colonne respective. Pour la colonne Dajɛ, l'ordre change du croissant au décroissant (A-Z / Z-A) tandis que pour la colonne Yeko, l'ordre change de l'unité ayant le moins de fréquence d'occurrence à l'unité ayant le plus de fréquence.

4. Comment Bosolan traite-t-il l'article de dictionnaire

Comme il a été souligné dans la section précédente, la mise à disposition de la liste des occurrences est l'étape qui jette les bases de la rédaction d'articles dans Bosolan. A la suite, vient alors la question de structuration de l'article, macrostructure et microstructure à la fois, ainsi que celle de sa description. Il s'agit ici de penser à fournir suffisamment de renseignement langagier sur les unités retenues. Et l'ensemble des informations fournies sur l'unité un est tout appelé article de dictionnaire. Un article de dictionnaire est donc constitué des composantes appelées les rubriques. Le nombre de ces rubriques varie selon l'objectif, le public

cible du dictionnaire ou encore selon le type de dictionnaire à confectionner. Les rubriques obligatoires présentes dans la plupart des articles dictionnaires sont l'entrée, la partie du discours, la phrase définitoire et l'exemple définitoire (Lehman 2018 :257, 283). A ceux-ci s'ajoutent plusieurs autres tel la synonymie, l'antonymie, les citations, l'étymologie, la phonie, la graphie, les marques d'usage. Cependant, de toutes ces rubriques, quelles sont celles retenues dans un article de dictionnaire traité dans Bosolan ?

En guise de réponse à cette question, nous disons que les rubriques suivantes sont celles qui structurent un article du logiciel Bosolan : entrée, indicatif de grammaire, phrase définitoire, exemple définitoire, synonyme, antonyme. Chaque article est assorti d'informations statistiques. A commencer par la rubrique entrée, les sections suivantes donne des précisions sur le mode de traitement de chaque rubrique de l'article dictionnaire dans Bosolan.

4.1 L'entrée

Tout d'abord, il faut choisir parmi les occurrences inventoriées du corpus les unités qui deviendront des entrées de ton dictionnaire. L'idéal serait de considérer l'ensemble des occurrences comme entrée du dictionnaire. Mais hélas, tous les mots d'un corpus ne sont pas toujours à attester dans un répertoire lexicographique comme le dictionnaire. On peut soutenir cet argument en disant qu'il existe des mots de moindre importance dans un corpus aux côtés des mots de grande importance. En plus, Alise (Lehmann 2018 :267) évoque la même question de la façon suivante : « *Aucun dictionnaire de langue ne peut ni ne doit recueillir tous les mots observables dans les discours. Les nomenclatures sont nécessairement limitées par les dimensions physiques des dictionnaires, même les plus grands* ». Donc, le choix s'impose au lexicographe. Partant de cela, Bosolan donne la latitude au lexicographe de faire le choix de l'unité à traiter comme entrée. On peut alors choisir les entrées dans la liste soit à tous azimuts selon un mérite plus ou moins subjectif, soit par ordre alphabétique ou encore en se basant sur le critère de la fréquence des occurrences d'où on se réfère aux nombres de la colonne Yeko.

Dans ce contexte, le choix le plus objectif des entrées du dictionnaire est celui de la fréquence des unités (Lehmann 2008 :26). La fréquence d'emploi d'une unité dans un corpus est très souvent synonyme de son importance dans le lexique de la langue. C'est pour dire que l'unité «kàla (baton)» se limitant seulement à 8 occurrences dans un corpus fournira moins d'acception dans le traitement qu'un corpus comptant plus de 100 occurrences de la même unité. Il est préférable de se fixer un intervalle dans les fréquences. Par exemple, les unités possédant une fréquence supérieure ou égale à 10 peuvent être les cibles.

Une fois le choix dégagé, l'opérateur défile la liste des occurrences répertoriées dans Bosolan. Il cherche ainsi les unités faisant office du nombre de fréquence retenue. A chaque unité de ce genre, il procède à

un clic sur l'unité. Aussitôt, la statistique sommaire correspondant à l'unité cliquée s'affiche en haut de la bande verticale juste au-dessus de *Daje Təmə*. Le choix de l'entrée prend fin à cette étape. Il reste à l'opérateur de décliner l'identité grammaticale de l'entrée. Cela faisant, le sous-titre suivant explique comment peut-on choisir l'indicatif de grammaire de l'entrée dans Bosolan.

4.2 Les indicatifs de grammaire

On reste toujours dans la barre verticale et dans la sous rubrique *Daje təmə*. Nous y sommes pour renseigner la rubrique « indicatif de grammaire ». A titre de rappel, l'indicatif de grammaire comprend la catégorie lexicale (partie du discours) et la catégorie grammaticale (genre et nombre...) de l'entrée (Boutin-Quesnel 1978:52). La catégorie grammaticale n'étant pas un phénomène pertinent dans la métalangue bamanan, le logiciel n'en fait pas cas. Par contre, les livrets de grammaire monolingue du bamanankan énumèrent plus ou moins une vingtaine de catégorie lexicale parmi lesquelles certaines se recoupent parce que certains grammairiens ont été plus dissociant que d'autres dans les métalangues. Le rôle du concepteur du logiciel Bosolan n'étant pas d'assurer une police quelconque sur le nombre de la métalangue de la langue, il n'a fait que recenser les catégories lexicales disponibles dans les différents livres de grammaire du bamanankan. Par conséquent, il a dressé une liste exhaustive pour que chacun y trouve son compte. La liste des catégories lexicales est intégrée au logiciel sous forme de fichier système. L'option « *Daje siya* (Nature du mot) » est le bouton qui fait office de point d'accès à la liste.

4.2.1. L'option *Daje siya*

Elle est une liste déroulante qui s'affiche lorsqu'on la clique. Elle se trouve aussi dans la barre verticale parmi les différentes options de la sous rubrique *Daje bosə*. Lorsque la liste se déroule, le lexicographe, connaissant bien les catégories lexicales de la langue, choisit la catégorie à laquelle répond l'entrée respective. Cependant, de toutes les catégories lexicales du bamanankan, le verbe possède un aspect à prendre en compte dans la rédaction lexicographique : la rection (Tourneux 2022:171). D'autres appellent ledit aspect la valence (Vydrine 2020 :91). Il s'agit ici de faire mention de la transitivité du verbe. Le logiciel Bosolan en a tenu compte. De ce fait, dès que l'opérateur choisit le verbe comme catégorie d'une entrée quelconque, l'espace « *Dafa nəfanko* » sous « *Daje siya* », auparavant désactivé, s'active aussitôt. C'est l'espace dans lequel il faut cliquer sur la rection respective de l'occurrence.

4.2.2. Comment Bosolan détermine-t-il la rection des verbes

L'espace « *Dafa nəfanko* » dédié au traitement de la rection du verbe est muni des différentes options sur la transitivité du verbe bamanan. La

rection du verbe est surtout plus complexe en bamanankan qu'en français ou en anglais. C'est pourquoi on peut énumérer jusqu'à 4 différentes rections dans cette langue : intransitif (dafalen : daf), antétransitif (ɲɛdafama : ɲdma), post-transitif (kɔdafama : kdma), circontransitif (fàñfiladafama : fdma). Un même verbe peut en avoir autant que les différentes acceptions qu'il couvre. Raison pour laquelle, un verbe comme bɔ (sortir) peut se retrouver avec toutes les rections selon ses différents emplois dans le corpus:

- *Kà bɔ : Ne bɔra (je suis sorti) daf*
- *Kà a bɔ : Ne ye wari bɔ (j'ai dépensé de l'argent) ɲdma*
- *Kà bɔ a fɛ : A bɔra a fa fɛ kojùgu (il ressemble trop à son père) kdma*
- *Kà a bɔ a ma : A y'a ka dùguta a bɔ a ma (il a annulé son voyage) fdma*

Les phénomènes de verbe antétransitif, post-transitif ou encore circontransitif sont propres à la métalangue du bamanankan. Ce sont donc des terminologies forgées en français pour rendre les concepts respectif wale ɲɛdafama, wale kɔdafama, wale fàñfiladafama. Les mêmes phénomènes entraînent avec eux des terminologies comme antécomplément (ɲɛdafa), postcomplément (kɔdafa), circoncomplément (fàñfiladafa) (Ballo 2020 :468).

Par ailleurs, on disait à la fin de la section 4.2.1 que l'espace dédié au traitement de la rection du verbe s'active lorsqu'on a choisi le verbe comme catégorie d'une entrée. Cette activation invite l'opérateur à cliquer sur la rection correspondant au verbe sous traitement. Mais avant cela, dès le choix de la catégorie lexicale, les flèches de navigation dans le texte étant activées en même temps que l'espace de la rection, l'opérateur doit cliquer d'abord sur la flèche d'avancement en avant.

4.2.2.1. Les flèches de navigation dans le corpus

Le verbe est choisi comme catégorie lexicale de l'entrée. L'étape suivante est la recherche de la première occurrence de l'entrée dans le corpus pour pouvoir déterminer sa rection dans son contexte énonciatif. A propos, de même qu'on dit que le mot n'a de sens que dans une phrase, on ne saurait se situer sur la rection du verbe tant qu'il n'est pas dans une phrase. Le clic sur la flèche met le logiciel à la recherche de la première occurrence de l'entrée respective dans le corpus actif. L'occurrence s'affiche sous la couleur de surbrillance grise comme contraste dans son lieu d'occurrence dans la phrase en particulier et dans le corpus en général.

4.2.2.2. Le choix de la rection

La première occurrence de l'entrée étant indiquée dans le texte, l'opérateur peut facilement déterminer la rection que doit porter l'entrée dans cet environnement réel. Par conséquent, il clique sur la rection respective dans l'espace de la rection. Le logiciel insère aussitôt les parenthèses dans le texte, derrière l'occurrence en question. Les

parenthèses contiennent le symbole attribué à la réaction choisie muni à gauche et à droite d'une barre de division. Une fois à ce niveau, les options des réactions se désactivent pour laisser activer le champ à traiter dans l'étape suivante. Il s'agit du traitement de la rubrique « phrase définitoire ».

4.3. La phrase définitoire

Nous faisons allusion à la rubrique définition dans l'article. Le champ dédié à cette tâche dans la page Даґебосо du logiciel se trouve juste sous le champ d'affichage de la liste des occurrences extraites du corpus. L'opérateur est guidé vers le champ par l'enseigne « Ләғәлибәжу (définition sommaire) » qui disparaît dès qu'on entame le remplissage du champ.

Le lexicographe, ayant l'entrée sous les yeux dans la phrase du corpus, devine une définition plus ou moins provisoire avec laquelle il remplit la case « Ләғәлибәжу ». Il valide sa définition en cliquant sur le bouton « Ләғәлибәжу Дөн (Valider la définition) » qui s'active pendant le remplissage du champ de la définition. Ce bouton suit horizontalement le champ « Ләғәлибәжу ». Le bouton, une fois cliqué, laisse apparaître une boîte à dialogue sur l'écran au milieu du texte. C'est la boîte à dialogue à renseigner dans l'étape prochaine : le synonyme et l'antonyme. Mais avant cela, qu'en est-il du traitement de la rubrique « exemple définitoire »? Un élément de réponse sera donné dans la section 4.4.

Par ailleurs, l'exposé ne fait pas ressortir la question du type de définition lexicographique à consigner dans le champ Ләғәлибәжу. Pour raison, tout lexicographe est suffisamment muni de bagage intellectuel pour faire la part entre les types de définitions et par la suite opter pour un d'entre eux. Les types de définitions sont entre autres la définition par inclusion, la définition métonymique, la définition morphosémantique, la définition synonymique et antonymique, et la définition métalinguistique (Lehmann 2018 :298).

4.4. L'exemple définitoire

L'exemple définitoire est un énoncé sciemment fourni dans l'article lexicographique avec la mention de l'entrée tel qu'elle se trouve en vedette. C'est pour dire qu'on n'y remplace pas l'entrée par un mot de ses autres formes possibles. Cependant, l'exemple définitoire peut soit être forgé de toute pièce soit peut provenir d'une source publiée : le corpus d'extraction de l'entrée ou un autre texte édité.

Chacune des deux sources comporte ses propres avantages et inconvénients à la fois. De ce fait, lorsque l'exemple est une construction délibérée de l'opérateur, un exemple rédigé ad hoc selon Gaudin (2000 :84), son inconvénient réside dans le fait que la crédibilité de la phrase est à revoir par une tierce pour approuver sa cohérence avec le sens défini ou encore pour vérifier sa grammaticalité (orthographe, syntaxe). Pour un exemple tiré d'un document, surtout si c'est automatiquement tiré du corpus lui-même, un des inconvénients est la longueur de la phrase : souvent jugée soit trop laconique si l'extraction

automatique tombe sur une phrase trop courte soit trop encombrante à cause du manque de toute économie du point de vue espace qu'elle occupe dans l'article.

Dans Bosolan, l'exemple définitoire est la rubrique dont le contenu suit directement celui de la phrase définitoire dans la présentation finale du résultat du traitement de l'article (infra sous-titre 7.2). Ce contenu est suivi par celui des synonyme et antonyme au cas où l'entrée en a fourni. Pourtant, lors du traitement, l'étape de l'exemple définitoire est occultée d'où l'opérateur assiste plutôt à la succession directe de l'étape des synonyme et antonyme à celle de la phrase définitoire. La raison de cette dissimulation de l'étape est que Bosolan extrait automatiquement les exemples définitoires du corpus. Il fait la capture automatique de la phrase dans laquelle baigne la première occurrence de l'entrée dans le corpus. Donc, les exemples définitoires d'un article traité à l'aide de Bosolan sont des phrase-exemples (Gaudin 2000 :84) parce qu'ils proviennent des sources publiées. L'étape de traitement de l'exemple définitoire étant occultée lors de l'opération, l'étape de la rubrique des synonymes et antonyme prend le relai.

4.5. Le synonyme et l'antonyme

Lorsque la boîte à dialogue s'affiche sur l'écran suite à la validation de la définition, on peut y voir deux champs à remplir avec leur enseigne respectif : $\text{K}\alpha\text{r}\alpha\text{r}\alpha\text{g}\alpha\text{v}\alpha\text{m}\alpha$ (synonyme) et $\text{S}\iota\alpha$ (antonyme). L'opérateur, le lexicographe en question, sait si son entrée possède ou non un synonyme ou antonyme. Il remplit les champs s'il connaît un synonyme ou antonyme. Dans le cas contraire, il passe en cliquant sur le bouton de validation « $\text{O}\kappa$ (Ok) » en bas de la boîte. A ce stade, le logiciel vient de faire le tour complet des rubriques prises en compte dans la rédaction d'un seul sens de l'entrée active.

Pourtant, le dépouillement n'a pas encore exploré l'ensemble des occurrences de l'entrée active dans le corpus. Dès que la rubrique « $\text{K}\alpha\text{r}\alpha\text{r}\alpha\text{g}\alpha\text{v}\alpha\text{m}\alpha$ et $\text{S}\iota\alpha$ » est validée, le logiciel part automatiquement à la recherche de l'occurrence suivante de l'entrée dans le corpus au lieu d'attendre à ce que l'opérateur le fasse manuellement en cliquant toujours sur la flèche d'avancement. C'est pour dire que le logiciel en a fini avec le premier sens (acception) de l'entrée sous traitement. L'exploration des autres occurrences pourra nous dire si l'entrée a d'autres sens en plus de ce qui est traité.

4.6. Le traitement de la polysémie dans Bosolan

Bosolan parcourt le corpus à la découverte de l'occurrence suivant la précédente traitée. La couleur de surlignage grise invite toujours l'opérateur à se fixer sur le sort lexicographique de l'occurrence dans son environnement d'emploi réel. Il doit répéter les mêmes étapes de traitement des rubriques faites chez l'occurrence précédente et dans le même ordre dicté par le logiciel excepté l'étape de la rubrique de l'indicatif de grammaire. Cette dernière, une fois choisie dès le traitement de la

première occurrence d'une entrée, n'est plus à refaire jusqu'à terminer la boucle du traitement des occurrences de l'entrée respective. Ainsi, la première étape de traitement à laquelle l'opérateur pense de l'occurrence active est sa rubrique définition pour les entrées de toute catégorie lexicale sauf la catégorie verbe. Pour les entrées de catégorie verbe, il commence plutôt les étapes à partir du choix de la recton avant de continuer pour l'étape de la rubrique définition.

Ce n'est qu'arrivé à l'étape de la rubrique définition pour chaque catégorie que le lexicographe découvre la question de polysémie. A ce stade, deux cas de figure s'offrent à l'opérateur : soit il trouve que l'occurrence actuelle porte un sens identique au sens ou à l'un des sens déjà traités de l'entrée, soit il trouve que son sens est distinct de l'autre ou d'un autre sens traité. La comparaison des sens est rendue facile à ce titre grâce au champ de la console des définitions.

4.6.1. Les occurrences à sens identique dans le corpus

A chaque fois que l'opérateur arrive à une nouvelle occurrence de l'entrée, il doit procéder à un exercice de comparaison entre le ou les sens existant dans la console des définitions et le sens qu'il découvre chez l'occurrence active. Si le sens est identique à un sens figurant parmi les sens déjà définis dans la console, il lui suffit de faire un double clic sur la définition correspondante dans la console. Aussitôt, la statistique de la fréquence de ladite définition augmente dans la colonne Yeko de la console. Plus le nombre d'occurrence ayant le même sens augmente, plus ledit sens est attesté comme acception de ladite entrée. La fréquence est un critère de confirmation et de crédibilité d'un sens quelconque de l'entrée. C'est pour dire que plus un sens est attesté chez les occurrences d'une entrée dans le corpus, plus ledit sens est crédible. Cependant, les mots étant super polysémiques dans le lexique de la langue, l'opérateur ne tarde jamais à tomber sur un ou des sens distincts du sens des occurrences qu'il a déjà traitées. Alors, qu'en est-il des occurrences qui se démarquent des autres du point de vue sens ?

4.6.2. Les occurrences à sens distinct dans le corpus

La question d'occurrences à sens distinct est abordée par Mel'cuk (1995 :160, 184, 186) sous les appellations *regroupement des lexies, division ou regroupement des acceptions*. Ce dernier le traite avec la lexie « vendre » comme exemple d'explication. Par contre, il en est question ici sous le nom d'occurrences à sens distinct parce que le l'opérateur découvre des distinctions de sens entre les différentes occurrences d'une même entrée pendant son parcours.

L'opérateur continue de parcourir les différentes occurrences donc. Il tombe sur une occurrence suivante. Mentalement, il trouve que son sens est différent de celui des occurrences déjà traitées. C'est dire qu'il fait face à une occurrence à laquelle il ne peut pas appliquer une définition précédemment donnée à l'entrée. C'est la preuve que ladite entrée est polysémique. L'opérateur formule alors dans Лэфolibaju une définition

fidèle au sens de l'occurrence en vigueur. Cette définition, une fois validée, devient une de plus consignée à propos de l'entrée dans la console des définitions. Ainsi, dans la console des définitions, on trouve tout un étalage de définitions distinctes sur l'entrée au fur et à mesure que la découverte d'acception nouvelle continue tout le long du corpus. Ces consignations sont constamment classées dans l'ordre de fréquence de chaque définition. Jusqu'à ce que le dépouillement atteigne le bout de l'ensemble des occurrences de l'entrée active, il est toujours probable de tomber sur de nouvelles acceptions de l'entrée. Une fois le tour bouclé dans le dépouillement d'une entrée, l'opérateur choisit l'entrée suivante toujours parmi celles ciblées dans la liste de base au-dessus du champ *Λεϕolibaju*. Une fois encore, le processus de dépouillement est répété de la rubrique de l'indicatif de grammaire jusqu'aux synonyme et antonyme. Par ailleurs, quelques explications additives sont à signaler sur quelques corollaires du dépouillement lexicographique d'une langue comme le bamanankan. La section suivante y apporte des précisions.

5. Corollaires du dépouillement lexicographique en bamanankan

Le bamanankan a des particularités assez récurrentes qui sont souvent à l'origine de certains traitements subsidiaires pendant le dépouillement d'une entrée. Ces particularités sont entre autres la récurrence des unités homographes « da (nom : bouche), da (nom : nombre), da (nom : oseille), da (verbe : coucher, ...), da (verbe : créer) » dans le corpus, et la notation des tons.

En ce qui concerne le problème d'homographie, il est d'ordre lexicologique et se traite sous le principe des relations lexicales. L'homonymie qui se subdivise en homophonie et homographie dans l'orthographe de certaines langues ne comprend que l'homographie en bamanankan vu que l'orthographe de cette dernière a opté pour le principe : à tous les homonymes homophones, la même orthographe.

Le lexique général du bamanankan est très fécond en matière d'unités homographes (Ballo 2020a :140). Surtout, le phénomène de déverbal vient encore épaissir la situation lorsqu'on sait que chaque verbe est potentiellement un substantif en bamanankan. Donc, un infinitif devenu déverbal par conversion reste toujours avec sa forme de l'infinitif. Le dépouillement lexicographique fait les frais de ces ressemblances morphologiques des occurrences. Dès qu'on lance le traitement d'une entrée homographique comme « da » dans Bosolan, on s'aperçoit rapidement que ce type d'entrée possède des occurrences qui ne s'apparentent pas aux comportements de l'entrée sous traitement. Ces occurrences particulières diffèrent des occurrences réelles de l'entrée soit de par leur catégorie, soit de par leur sens synchronique (Polguère 2008 :60) ou encore de par leur comportement distributionnel (Lehman 2018 :101).

Lorsque l'opérateur du dépouillement découvre une occurrence de ce genre, il le saute tout simplement à l'aide des flèches d'avancement. Il part donc à l'occurrence suivante. Le traitement d'une telle occurrence ne

doit en aucun cas être mélangé au traitement en vigueur de l'entrée. Son traitement ou le traitement de toute autre occurrence de la sorte fera l'objet d'un autre tour de dépouillement de l'entrée.

Cependant, Bosolan n'est pas encore muni de mécanisme automatique prenant en compte cet état de fait. Donc, pour traiter toute occurrence de ce type, l'opérateur doit faire preuve de bonne astuce. Il peut donc passagèrement modifier l'orthographe de l'entrée dans les limites de ses occurrences qui ne sont pas encore traitées. Cela permet de différencier la forme des occurrences à traiter (les occurrences ayant subi le saut parce qu'elles ne doivent pas bénéficier d'un traitement unitaire que celles traitées) de la forme de celles déjà traitées lors du tour de dépouillement précédent. Il peut alors accéder facilement aux occurrences cibles rendues distinctes par la modification de leur orthographe.

Par ailleurs, le problème de notation de tons n'étant pas encore totalement tranché dans le cercle des décideurs, Bosolan ne prend en compte que la notation du ton de l'indéfini qui tombe uniquement sur les substantifs. Les deux tons lexicaux, hauts et bas qui introduisent aux apprenants dans les centres d'alphabétisation, sont laissés pour compte dans la version présente de Bosolan. Un lexicographe souhaitant différencier ces mots de tons différents dans son répertoire peut le faire de son propre chef soit en marquant l'accent grave sur ses entrées à ton bas ou en marquant le ton contraire.

Pour la notation du ton de l'indéfini sur les substantifs employés avec ce ton, Bosolan fournit une rubrique facultative à ce sujet. Pour cela, lorsque la catégorie *togo* (substantif) est retenue comme catégorie lexicale d'une entrée sous traitement, la rubrique « *Yal'i b'a fe kà faranfasibali tàamasijε Sigi wa ?* (Voulez-vous marquer le ton de l'indéfini) » s'affiche sur l'écran dans une boîte à dialogue. Cette instruction s'affiche juste à la suite de la validation de la rubrique synonyme et antonyme. Cela invite l'opérateur à se décider du marquage du ton si l'occurrence en question est prononcée à l'indéfini. La marque incorporée au logiciel pour cette fin est l'apostrophe que le logiciel appose à l'occurrence, si son choix est de le marquer, juste avant sa lettre initiale. Le fait d'instruire au logiciel de ne marquer que les substantifs à ton indéfini à côté de ceux à ton défini est une décision venant du concepteur du logiciel. C'est quand même une décision fondée sur un fait objectif : les substantifs se prononçant avec le ton de l'indéfini sont moins nombreux dans un corpus que ceux qui se prononcent avec le ton du défini. Donc, une règle grammaticale qui donne moins de tâches aux usagers de la langue est beaucoup plus salutaire que son contraire. Le choix de l'apostrophe comme marque de notation est aussi arbitraire de la part du concepteur.

6. Les limites du logiciel bosolan

Le logiciel Bosolan est limité dans son dépouillement de corpus lexicographique sur deux plans majeurs : la prise en compte intégrale de tout ce qui peut être rubricé de l'article dictionnaire et son adaptabilité aux exigences lexicographiques de la plupart des langues africaines.

Sur le plan de la diversité des rubriques d'un article de dictionnaire, l'idéal serait de donner la latitude à l'opérateur de générer lui-même autant qu'il en a besoin des rubriques de la microstructure de son article. La version actuelle du logiciel ne prenant pas en compte ce cas de figure, cela réduit le champ de manœuvre du lexicographe. Rien qu'en ne se référant qu'au grand ensemble bloc-entrées de Lehmann (2018 :283), on s'aperçoit aisément qu'on peut y compter jusqu'à 4 bonnes rubriques qui ne sont pas chacune représentées dans l'article de Bosolan : graphie, prononciation, la catégorie et l'étymologie. Pour cela, le logiciel doit proposer un mécanisme conséquent d'ajout de rubrique. Un bouton « Ajout rubrique » serait la bienvenue dans ce sens.

Sur le même sujet, la démarche d'article composite entretenue par Doucouré (2021) dans la confection de son dictionnaire monolingue *Bamanankan dapegafe* n'a pas de répondant dans Bosolan. La démarche consiste à gloser sous une vedette une sous-entrée morphologiquement semblable. Cette démarche a consisté chez Doucouré à mettre dans le même article le déverbal et son verbe (cf. article du verbe *sɔrɔ* « trouver » et son déverbal *sɔrɔ* p.571). La littérature sur la question parle aussi de superarticle (Polguère 2008 :234) ou encore regroupement d'entrées (Lehmann 2018 :277).

Sur le plan de la prise en compte du traitement lexicographique d'autres langues africaines, les interfaces dans les langues concernées sont à concevoir. Il faudra également chercher à mieux saisir les pertinences et récurrences métalinguistiques, grammaticales et surtout syntaxiques qui sont le propre de ces langues.

En plus, la question du traitement unitaire ou homonymique des entrées (Lehmann 2018 :100, 106) et même des sens restent à satisfaire davantage. Le mécanisme actuel de traitement de ce phénomène est toujours manuel. On peut l'améliorer afin qu'il soit possiblement automatique ou systématique. En fin de mot, la présentation finale de la rédaction lexicographique chez Bosolan se fait en colonne unique contrairement aux pratiques actuelles qui présentent en colonne multiple (infra sous-stire 7.2). Mais, cela n'est point un obstacle pour appliquer la multiple colonne sur le texte. Le lexicographe peut faire recours au module Word pour faire la mise en page de son texte au nombre de colonne qu'il souhaite.

7. Quel présentation Bosolan fait-il pour les résultats de la rédaction

A part le Trésaurus d'Etienne Robert en 1531, répertoire présenté en colonne unique (Gaudin 2000 :20), la présentation des résultats de travaux lexicographiques en français se voit en colonne multiple à l'impression : double, triple ou encore quadruple colonne.

La présente section de l'exposé traite les aspects suivants sur comment Bosolan présente les résultats de la rédaction en ces termes :

- *les traces des traitements effectués;*
- *le bilan des rédactions faites à l'aide de Bosolan ;*
- *la microstructure brute de l'article.*

7.1. Les traces des traitements effectués

Bosolan présente le résultat de sa rédaction en trois différentes traces en occurrence une trace laissée dans le corpus, une trace laissée dans la console des définitions et une trace enregistrée dans un fichier système. La première trace laissée par le logiciel est insérée juste après l'occurrence traitée entre parenthèses et barres de division (/). Les informations du traitement déjà contenu dans cette trace sont les mentions de la catégorie de l'entrée, le symbole de sa rection (uniquement pour les verbes), la définition et enfin deux crochets contenant le nom de l'auteur et la date du traitement. En voici un exemple : *girin (/w Èkà teliya wajibiya a kan [IB | 2022-09-22]/)*.

Concernant la deuxième trace, elle s'affiche dans le champ de la console des définitions juste en bas du champ « Λεφολιβαжу ». C'est un champ destiné à lister les différentes définitions des acceptions attestées chez une entrée. Le champ est en double colonne : colonne des définitions énumérées (Dajε) et colonne de la fréquence de chaque définition (Yeko). La troisième trace est en principe enregistrée dans un fichier système appelé « Dajε bosolenw màra ». L'opérateur n'en a pas besoin pendant qu'il est dans le vif du dépouillement. Le logiciel n'enregistre cette trace qu'après que l'opérateur ait cliqué sur le bouton enregistrer (Lasago) de Bosolan pour sauvegarder le fichier du corpus traité. A la suite, il peut y jeter un coup d'œil s'il le souhaite en cliquant le bouton correspondant « Dajε bosolenw màra » parmi les icônes de la barre des éditions du logiciel. L'article s'affiche alors sur la console « Bânbe » en bas de l'écran. Le contenu affiché du fichier système *Dajε bosolenw màra* est synonyme du nombre d'articles traités jusque alors.

7.2. le bilan des rédactions faites à l'aide de bosolan

Bosolan a déjà traité plus de 100 entrées dans un corpus. Le contenu dudit corpus est varié mais il n'est pas encore assez représentatif quant à la couverture d'un grand nombre de domaine. Les domaines représentés dans le corpus se limitent seulement à la littérature et à la religion. Pire encore, sa représentation en quantité de textes reste aussi à améliorer vu que les différents textes qui le composent ne dépassent pas 9 pour un nombre d'occurrences qui ne s'élève qu'à 300 000 mots environ.

Les 200 entrées déjà traitées comme échantillons couvrent la plupart des parties du discours répertoriées dans la grammaire du bamanankan. La catégorie du verbe reste quand même la dominante de l'échantillon vu que le verbe présente plus de caractéristiques lexicales que les autres catégories.

Une équipe de 4 membres a fait ce travail de test grandeur nature du logiciel. L'équipe a transformé les résultats de traitement des 200 articles en une base de données lexicographiques. L'équipe continue avec le traitement pour en faire un dictionnaire monolingue bamanan en fin de compte.

Par ailleurs, il est de coutume de faire la présentation finale des articles traités en multiple colonne comme cela se voit dans le dictionnaire tout fini. Cependant, la version actuelle de Bosolan ne génère pas directement la présentation en double colonne. En lieu et place, Bosolan génère un fichier contenant les articles traités en une seule colonne. Ce fichier, appelé *Daje bosolenw màrà*, est accessible sur la barre des éditions avec son icône caractéristique. Le clic de ce dernier affiche son contenu en prose dans la console *Bànbe* (supra sous-titre 2).

Pour avoir une présentation en double colonne, l'opérateur peut copier le contenu de *Bànbe* pour le coller dans un fichier Word. Une fois dans Word, le style journalistique est possible. Voici un échantillonnage de 5 articles finalement mis en double colonne dans Word. Il est suivi des détails sur la microstructure brute des articles

bari w ~ G 1 JK 1 Kb 1 NT 4

1. fdma k'a bùgò n'a ye Δ A ye ñεε bari sokε la, sokε girinna ka Taa Kala Bukeyi kòrɔba ka sokε kama, sokε fila kalansola ñogòn na. (G p.53 | y.3 | KB | 2021-01-05)
2. fdma mògò: k'i fàrikolo mèrε a la Δ U tora k'u panpan, k'u bari a la. ≈ nòrɔ a la ≠ waaro (NT p.81 | y.2 | KB | 2021-01-15)
3. fdma k'a turu a la Δ Tama y'o jε, ka taa o bari kogo la gaan. ≈ genge ≠ bò (JK p.38 | y.1 | KB | 2021-01-10)
4. kdma k'a datugu n'a ye Δ Mògò minnu ye jugumanw kε , juguman kelen sara y'a ñogonna juguman kelen ye, dɔgoyama kòni b'olu datugu, tangalibaga si t'olu boolo ka bɔ Ala ma, a bε kε i n'a fɔ su dibifin mɔnimɔni dɔ tigelen bε ka bari olu ñɛdaw la, olu de ye tasuma donbaaw ye, u bε dumε a kòrɔ; 28. (Kb p.430 | y.1 | KB | 2021-01-29)

dugura tg ~ G 4

- Jirisun dòn kungo la min bε Kùnbaya kà yεgen fo a den m̀nen bu binkεnεnama ǹonɔma bε Dun kà a kolo la s̀ den cogo la samijε fε. Δ Sisan sa, ne ye fεn min ye, ne b'o jira e la dugura kòrɔ yan bi. (G p.52 | y.4 | IB | 2022-06-27)

fɔ w ~ NB 14

- ñdma Kà kùnnafoni bɔ i da la Δ Gasi tε, juguman tε, A' min mana fɔ , Ne tulo b'o la. (NB p.7 | y.14 | IB | 2022-06-26)

jira w ~ G 9

1. fdma k'a Bɔ a kàla ma n'a keli ye a ñε 'kòrɔ walima n'a mìsali b̀lili ye a tigi kan. Δ Sisan sa, ne ye fεn min ye, ne b'o jira e la dugura kòrɔ yan bi. (G p.52 | y.5 | IB | 2022-06-27)
2. fdma K'a Ɓini a k'a Ye a ñε na n'i ka kofoli ye Δ N b'a fε i ka bilakoronin bila n ñε, ka Taa Naa tɔgò saba ka yòrɔ jira n na, n ka jigin olu kan, o ka fisa aw kanjiginni ye. (G p.88 | y.4 | IB | 2022-06-27)

kasi w ~ AK 2 NT 20 G 3 JK 5 Kb 11

1. daf kà ñeji bə Δ Denniba sinna ka kasi koyi! ≠ yεε (NT p.84 | y.38 | YB | 2021-02-23)
2. ñdma kà dùsutəərə blà a la Δ U ye anw (Ala) dusu kasi minke, an y'an ta sara u la, an y'u bæε lajelen to ji la. ≠ yεε, niyəndiya (Kb p.992 | y.2 | YB | 2021-04-16)
3. kdma kà jò a ka fili la Δ A ko : « Se i nən na, i ka dukεε k'i wasa, wa i ka kasi i ka fili la. (AK p.180 | y.1 | YB | 2020-12-11)

yaala w ~ NT 3

- ñdma ka bə i nə na ka taa fan dantigεbali də la Δ A taalen bæ də yaala, marifa b'a bolo jago la. (NT p.123 | y.3 | IB | 2022-07-2)

7.3. la microstructure brute de l'article

La microstructure d'un article de Bosolan a été présentée tout le long de l'exposé de façon diffuse. Néanmoins, nous en revenons ici de manière plus condensée relatant la structure exacte des articles issus du dépouillement.

Dans la présentation à double colonne, section précédente, on voit bien que les articles sont en ordre alphabétique. L'article qui vient le premier est celui dont l'entrée est le mot *bari*. Sur toute cette première ligne, les informations se suivent ainsi qu'il suit : l'entrée (*bari*), la catégorie lexicale de l'entrée (*w* = verbe), le symbole tilde (~) pour marquer la séparation, une série de lettres abrégées G JK Kb NT dont chacune est suivi d'un nombre. Ces abréviations tiennent pour le code des documents structurant le corpus et qui sont dépouillés lors du traitement de l'article respectif. Nous en voyons 4 pour cet article tandis qu'il figure 9 documents au total dans le corpus. Les 4 codes désignent uniquement les documents dans lesquels figurent des occurrences de l'entrée. Les chiffres tiennent pour le nombre d'occurrence que l'entrée respective possède dans les documents correspondants. Comme exemple, NT4 veut dire que dans le document NT, le dépouillement a vu 4 occurrences de l'entrée *bari*.

Le reste de l'article est constitué des lignes numérotées de 1 à 4. Ces numéros indiquent le nombre de sens décrit sur l'article respectif. Plus on dispose de numéros dans un article, plus la disjonction des sens de l'entrée est grande.

A chaque verbe, les numéros sont suivis par la suite d'informations comprenant la rection du verbe (*fdma/kdma/daf/ñdma*) au cas où l'option est cochée, l'énoncé définitoire, l'exemple définitoire introduit par le symbole incrément (Δ), les symboles *presque égal à* et *pas égal à* (\approx , \neq) pour introduire respectivement le synonyme et l'antonyme de l'entrée selon le sens respectif au cas où ils en ont été fournis. La série d'informations qui vient à la finale est comprise entre parenthèse. Elle est compartimentée en 4 champs successifs à l'aide de la barre verticale (|) dont chaque compartiment comprend : le code source et page d'extraction de l'exemple définitoire, la fréquence du sens respectif dans le corpus (*y...*), le nom de l'auteur puis la date de consignation du sens correspondant.

Les séries d'informations comprises dans les articles des autres catégories lexicales sont identiques à celles du verbe excepté l'absence de la rection chez les autres. Pour les substantifs, la marque de l'apostrophe est notée sur l'occurrence de l'entrée dans l'exemple définitoire juste avant la lettre initiale du mot lorsque l'opérateur en a fait le choix.

Cependant, qu'en est-il de l'ordre d'agencement des acceptions dans les articles de Bosolan ? Il s'agit bien des articles ayant une entrée polysémique.

En lexicographie, à titre de rappel, on peut énumérer jusqu'à 4 principaux ordres dans lesquels on agence les différents sens que couvre un mot. Il s'agit de l'ordre chronologique ou historique, l'ordre logique, l'ordre de fréquence et l'ordre linéaire (Lehmann 2018 :126, 304).

De ces différents ordres d'agencement des acceptions, celui retenu par Bosolan pour les polysèmes est l'ordre de fréquence (Lehmann 2018 :304). Donc, la numérotation des sens dans l'article va du sens le plus fréquent au sens le moins fréquent pour l'ensemble des occurrences de l'entrée dans le corpus. L'information qui indique la fréquence est celle qui se trouve dans le deuxième compartiment de la dernière série d'information sur le sens : le compartiment *yeko* (nombre d'occurrence) abrégé en *y.* avec le nombre respectif accolé. La succession des sens se voit avec les numéros introduisant les définitions respectives (voir article *bari* ou *kasi*).

Par contre, une entrée chez qui le traitement n'a attesté que seul sens, une entrée monosémique, la question de l'ordre d'agencement n'y est plus. La définition de ce seul sens est donc introduite par le symbole du disque noir (●) comme on peut le constater dans l'article sur *dugura*.

Sur tout autre plan, toutes les informations consignées dans l'article ne sont pas à garder à tout prix dans le dictionnaire à l'impression. A part les rubriques qui délivrent une description sur l'entrée, la plupart du reste est facultative voir même embarrassante dans le répertoire publié.

C'est pourquoi le lexicographe peut revenir sur les articles en appliquant une série d'édition de plus. Cette dernière édition débarrasse les articles des rubriques accessoires comme les informations comprises dans la dernière parenthèse ainsi que la série d'informations qui succède au tilde sur la ligne de l'entrée.

Cette dernière édition se fait dans une autre rubrique fonctionnelle de Bosolan : la rubrique *Sɛbɛnca* (édition). Une fois cliquée, cette option affiche une boîte à dialogue contenant les articles traités dans la marge gauche de la boîte et les options d'édition dans l'autre marge. Lorsqu'on opte pour une édition donnée, on choisit la tâche correspondante à appliquer sur les articles. On choisit soit la dissimulation des informations accessoire inutiles à l'impression ou changer l'ordre des rubriques dans l'article. Une fois le choix fait, on coche dans la liste les rubriques sur lesquelles tu appliques la tâche respective.

8. Bosolan : processus de réalisation

L'idée de création d'un outil de dépouillement lexicographique pour le bamanankan en particulier et les langues africaines en général était en

gestation dans la tête du concepteur depuis le premier trimestre de 2013. A l'époque, le concepteur venait tout juste de soutenir son mémoire de master de recherche en fin de l'année précédente. Les recherches effectuées et la rédaction de son mémoire intitulé « *The structure of the bamanankan lexicon* », une mémoire à dominante « lexicométrique », ont vite réveillé en lui le souci de créer un programme informatique. Dès lors, il élaborait en mai 2013 la première ébauche de cahier des charges de l'outil futur à mettre sur pied. L'outil, une fois mis sur pied, devrait être capable de résoudre quelques problèmes que son concepteur avait rencontrés dans le domaine de la statistique lexicale du bamanankan. Dans le cahier des charges élaboré à l'époque, on peut citer les points suivants :

- *Retrouver les éléments d'une liste dans le corpus et les remplacer respectivement*
- *Chercher l'élément quelconque dans le texte (mot, expression...). A chaque occurrence de l'élément, voir à gauche ou droite de l'élément si son compagnon [insérer compagnon ou liste de compagnons] l'accompagne dans la phrase et ensuite copier la phrase pour aller coller dans un fichier de destination. Son compagnon peut être intercalé à 0 élément / à 1 élément / à 2 éléments / de lui. Apposer alors en indice le chiffre du nombre d'élément entre les compagnons.*

Ces deux points tirés du cahier des charges sont révélateurs en ce sens qu'ils demeurent des options importantes du logiciel Bosolan dont la première version test a finalement vu le jour en octobre 2016. L'option Fàlenɲogɔnna (remplacement) et la rubrique Taapɔgɔnma (Unités corrélatives) sont respectivement les aboutissements du point 1 et 2 cités ci-haut.

Donc, de cahier des charges en cahier des charges, Bosolan a vu les options de son menu augmentées de presque dix à plus de 30 dans la version actuelle.

A chaque conception de cahier des charges, le concepteur l'envoie chez un prestataire qui s'y connaît en programmation informatique. Ce dernier traduit les concepts structurés en programme informatique. Le concepteur reçoit de la part du programmeur, un prototype qu'il vérifie les fonctionnalités.

Conclusion

Du début jusqu'à la fin de notre exposé sur la nouvelle ressource numérique *Bosolan*, le processus de traitement d'un article de dictionnaire est développé de façon linéaire et systématique. Il s'agit de découvrir un nouvel outil fait sur mesure pour la lexicographie du bamanankan et en bamanankan. Les différentes incapacités que les lexicographes rencontrent lors de la rédaction de dictionnaires en bamanankan à l'aide d'outils inappropriés sont enfin terminées. Il leur suffit de s'approprier le logiciel Bosolan pour la prise en charge aisée des aspects linguistique propres au bamanankan. Le plus simple de ces aspects irrésolus dans les

outils inadaptés mais heureusement intégrés à Bosolan est le classement en ordre alphabétique des entrées et le classement des différentes acceptions du polysème en ordre de fréquence. Un des plus complexes des aspects pris en compte par Bosolan est la rection des verbes. En fin de mot, il faut dire que Bosolan est la solution pour la rédaction lexicographique du bamanankan avec deux atouts : il est numérique et il résout les casse-têtes des problèmes propres à la langue bamanan.

Bibliographie

- Ballo, Issiaka**, 2022, *l'exploitation de l'environnement lexical dans le traitement automatique du ka possessif en bamanankan* in Cahiers du Cédimes, N° 2, pp.89-102
- Ballo, Issiaka**, 2020a, *Reconnaissance automatique des 3 ma : la méthode de l'ordre d'occurrence des unités bamanan*, in Collection FLE/FLA n°2, volume 1, Kenyata University, pp.139-153.
- Ballo, Issiaka, Coulibaly, Adama Diokolo**, 2020b, *Reconnaissance automatique du ka infinitif dans un corpus bamanan : la méthode de l'environnement lexical* in Germivoire n°13, volume ½, Université Félix Houphouët-Boigny, Abidjan-Cocody, pp. 467-478.
- Dukure, Mamadu, Baalo, Isiyaka**, 2021, *Bamanankan dajegafe*, Bamako, EDIS.
- Dumestre, Gérard**, 2003, *Grammaire fondamentale du bambara*, Paris, Karthala.
- Dumestre, Gérard**, 2011, *Dictionnaire bambara-français*, Paris, Karthala.
- Le Grand Robert de la langue française, version électronique**, 2005, SEJER.
- Lehmann, Alise et Martin-Berthet, Françoise**, 2008, *Introduction à la lexicologie : sémantique et morphologie*, Paris, Armand Colin, 3è éd.
- Lehmann, Alise, Martin-Berthet, Françoise**, 2018, *Lexicologie : sémantique, morphologie, lexicographie*, Paris, Armand Colin, 5è éd.
- Mel'cuk, Igor A., Clas, André, Polguère, Alain**, 1995, *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve, Duculot.
- Polguère, Alain**, 2008, *Lexicologie et sémantique lexicale : Notions fondamentales*, Montréal, PUM, 2è éd.
- Tourneux, Henry**, 2022, *De la nécessité de se doter d'outils lexicographiques adéquats pour le développement*, in Editions des archives contemporaines, Paris, pp. 167-181.
- Vydrin, Valentin**, 2020, *Vers une lexicographie mandingue sur la base de grands corpus annotés* in Mandenkan, N° 63, pp. 89-110.