

Construction d'un dictionnaire électronique des flexions verbales du kabyle à l'aide du logiciel NOOJ (Verbes primaires).

Hamid ANNOUZ

Doctorant, Lacnad - Inalco, Paris.

I)-Introduction

L'objet du traitement automatique des langues est de créer des programmes informatiques capables de traiter automatiquement la langue. Il s'agit d'une démarche qui consiste à manipuler des données linguistiques qui sont l'objet d'entrée de l'opération afin de les modifier en un objet de sortie qui est le résultat. Le traitement de ces données s'effectue par une suite d'actions appliquées par la machine à l'aide d'un programme qu'on lui installera. Nous devons donc avant tout, être capables de décrire les règles de la langue qui seront manipulées pour les transmettre à l'ordinateur qui les appliquera à son tour d'une façon automatique.

L'extraction de la représentation pragmatique d'un texte, par exemple, exige au minimum les quatre modules suivants (P. Bouillon et al, 1998, p. 15) :

Texte => Module lexical et morphologique [*représentation lexicale et morphologique du texte*] => Module syntaxique [*représentation syntaxique du texte*] => Module sémantique [*représentation sémantique du texte*] => Module pragmatique [*représentation pragmatique du texte*].

Avec l'évolution de la technologie, un grand nombre de langues se sert de TAL, et nous pourrions dire que, aujourd'hui, l'avenir d'une langue dépend de plus en plus de cette évolution. Pour cette raison, nous considérons que l'intégration du kabyle et de tous les dialectes berbères dans le domaine du TAL est indispensable.

Nous présentons dans cet exposé le résumé de ce que nous sommes en train de préparer pour le mémoire de Master 2. Il s'agit de construire un dictionnaire des flexions des verbes simples du kabyle à l'aide du logiciel NooJ. Nous voulons dire par flexion toutes les formes conjuguées d'un verbe auxquelles on a intégré la racine. Comme nous l'avons cité ci-dessus, cette tâche qui s'inscrit dans le module lexical

et morphologique est la première à réaliser. Ce choix n'est pas fortuit ; à notre connaissance, dans le domaine du TAL, aucun travail n'a été fait sur la langue kabyle.

Nous précédon's par les cinq étapes suivantes :

- La préparation des entrées dans la partie dictionnaire.
- La construction de toutes les flexions qu'un verbe peut avoir à l'aide des paradigmes d'aspects.
- Relier ces paradigmes d'aspects à une liste de verbes types qui seront notés devant les entrées du dictionnaire.
- Compiler la partie dictionnaire avec la grammaire.

II)- Aperçu sur la morphologie du verbe en berbère (kabyle)

1)-Les modalités obligatoires du verbe

-Le thème : Toute forme verbale kabyle et berbère en générale est constituée d'une racine consonantique appelée radicale, porteuse de sens et d'un schème vocalique indiquant l'aspect. A ces deux éléments qui constituent le thème s'ajoute les affixes (indices de personne) qui indiquent le sujet :

Racine + schème = thème verbal + indice de personne = il a écrit
r + *a - u* = *aru* + *y* = *yura*

-Les indices de personnes : Sont des déterminants grammaticaux qui accompagnent presque toutes les formes verbales à l'exception de participe. Les indices de personne sont identiques pour tous les thèmes et ne peuvent pas apparaître en dehors du verbe auquel elles donnent le rôle de prédicat.

	Singulier	Pluriel
1 ^{re} p.	_____ (e)y	n(e)_____
2 ^e p.m.	t(e) _____ (e)d	t(e)_____ (e)m
2 ^e p.f.	t(e) _____ (e)d	t(e)_____ (e)mt
3 ^e p.m.	i/y(e) _____	_____ (e)n
3 ^e p.f.	t(e) _____	_____ (e)nt

Remarque :

-L'indice de la troisième personne masculin singulier se réalise différemment selon les contextes¹ :

- *i-* devant un thème débutant par une consonne suivie d'une voyelle : *iwala* (il a vu).
- *ye/i-* devant un thème débutant par deux consonnes : *yelħa / ilħa* (il a marché).
- *y-* devant un thème débutant par une voyelle : *yufa* (il a trouvé).

L'aspect : Georges Mounin² définit l'aspect comme : « ...une catégorie grammaticale qui est différente des catégories du temps, du mode et de la voix, et qui manifeste le point de vue sous lequel le locuteur envisage l'action exprimée par le verbe : comme accomplie, c'est-à-dire vue dans son achèvement, son résultat, ou comme inaccomplie, vue dans sa durée, sa répétition... ».

À propos de cette dichotomie (accompli/inaccompli) S.CHAKER écrit : « De point de vue des signifiants, les deux thèmes fondamentaux sont ceux de l'aoriste et du prétérit. Les autres thèmes apparaissent tous comme secondaires par rapport à ces deux thèmes de base.. ». (1978 : 210).

Exemple :

Yečča (il a mangé), action accomplie.

(*ad*) *Yečč* (il mangera) action inaccomplie.

C'est dans cette logique aspectuelle que le berbère s'est évolué en développant des sous thèmes basés sur l'un ou sur l'autre. Actuellement dans la grammaire du berbère on parle de quatre thèmes : prétérit, prétérit négatif, aoriste et aoriste intensif.

2)-La formation des thèmes verbaux

-L'aoriste simple : La forme de l'aoriste simple correspond à l'impératif de la 2^{ème} p.sg. Elle est souvent accompagnée de la modalité préverbale *ad/a*, qui permet, selon les contextes, d'exprimer : le futur, le souhait, le potentiel, etc.

Exemple : *Ruħ* (partes) => (*ad*) *ruħey* (je partirai)

¹ NAÏT-ZERRAD K., *Mémento grammatical et orthographique de berbère, Kabyle-Chleuh-Rifain*, L'Harmattan, 2011, p. 39.

² Georges Mounin, *dictionnaire de linguistique*, éd. PUF, Paris 1974.

Le seul contexte où l'aoriste peut apparaître sans la particule préverbale est celui de l'optatif négatif : *a wer d-yuyal, drus i yexdem*, (pourvu) qu'il ne réapparaisse pas ! Il a fait tant de mal.

-L'aoriste intensif: Ce thème relève de l'inaccompli, formé à partir de l'aoriste simple, il est devenu une partie intégrante du jeu des formes de base et s'oppose directement aux autres thèmes, notamment le prétérit. Il est utilisé seul pour exprimer une action qui est en train de se faire, qui se fait d'une manière habituelle ou qui se faisait d'une manière habituelle dans le passé. Exemple : *itett* (il mange habituellement / il mange actuellement / il mangeait).

-Avec les modalités préverbaux *a / la / ala*, il exprime uniquement une action qui est en train de se faire. Exemple : *a /la/ala itett* (il mange actuellement / il est en train de manger).

-Avec la modalité non réelle *ad*, l'aoriste intensif est utilisé pour exprimer une action qui se fera (éventuellement) d'une manière habituelle dans le futur. Exemples : *ad itett* (il mangera habituellement dans le futur).

-Le prétérit : Il relève de l'accompli, il exprime une action achevée dans le passé, il constitue avec l'aoriste la seule opposition aspectuelle dans le système primitif du berbère. Sur le plan morphologique, on peut distinguer deux catégories : réguliers et irréguliers. Si nous concédons que la forme de base du verbe est celle de l'aoriste, nous pourrions dire que les verbes réguliers sont ceux qui gardent cette forme au prétérit, et les verbes irréguliers sont ceux qui changent de formes.

Exemple d'un verbe régulier : *muqel* (voir) :

Aoriste : (*ad*) *muqley* (je verrai) => Prétérit : *muqley* (j'ai vu).

Exemple d'un verbe irrégulier : *afeg* (voler) :

Aoriste : (*ad*) *afgey* (je volerai) => Prétérit : *ufgey* (j'ai volé).

Aoriste : (*ad*) *nadiy* (je chercherai) => Prétérit : *nuday* (j'ai cherché).

Aoriste : (*ad*) *zrey* (je saurai) => Prétérit : *zriy* (j'ai su).

-Le prétérit négatif : Certaines formes du prétérit négatives sont identiques à celles du prétérit. Dans d'autres verbes, cette distinction est marquée par l'apparition de la voyelle /i/ dans la dernière syllabe du verbe :

Exemples : *yezdey* (j'ai habité) → (*ur*) *yezdiy* (*ara*) (il n'a pas habité).

iwala (il a vu) → (*ur*) *iwala* (*ara*) (il n'a pas vu)

Ce thème est toujours employé, selon les contextes, avec les particules de négations : *ur/u—ara, ur-- , wer--* .

3-Le système d'impératif

Les thèmes de l'impératif simple et intensif sont identiques, respectivement à l'aoriste simple et à l'aoriste intensif. Il se combine avec les indices de personnes suivants :

Indices de personne	Aoriste simple
2 ^{ème} p. sing. m/f. ___(<i>zéro</i>).	<i>azzel</i> (cours).
2 ^{ème} p. pl. masc. ___(<i>t</i>)/(<i>m</i>).	<i>azzlet/m</i> (courez, vous hommes).
2 ^{ème} p.pl. fém. ___(<i>mt</i>).	<i>azzlemt</i> (courez, vous femmes).

4-Le participe

Le participe a des caractéristiques particulières, il est : incompatible avec les indices de personne et invariable en genre et en nombre. Il existe au prétérit, à l'aoriste et à l'aoriste intensif avec une forme négative différente pour chacune des formes positives. Sa forme est construite sur la 3^e personne du masculin *y/i-----n³*:

Prétérit :.....	<i>yeččan</i> (qui a / ont mangé)
Prétérit nég. :.....	(<i>ur</i>) <i>nečči</i> (<i>ara</i>) (qui n'a/ n'ont pas mangé)
Aoriste intensif :.....	<i>itetten</i> (qui mange(nt))
Aoriste intensif nég. :.....	(<i>ur</i>) <i>ntett</i> (qui ne trouve(nt) pas)
Aoriste :.....	(<i>ara</i>) <i>yeččen</i> (qui mangera/ mangeront)

-Pour les verbes d'état, le participe du thème prétérit ne comporte que l'élément suffixé *-n* :

Meqger (il est grand)*meqgren* (étant grand).

-Avec les verbes d'état au prétérit négatif les deux formes sont possibles :

³ Comme nous l'avons souligné plus haut, l'indice de la 3^e personne singulier masculin s'écrit, *i-* devant un thème débutant par une consonne suivie d'une voyelle, *ye-/i-*, devant un thème débutant par deux consonnes et *y-* devant un thème débutant par une voyelle.

Meqquer (il est grand) (*ur*) *nmeqquer* (n'étant pas grand).
 (*ur*) *meqqren* (n'étant pas grand).

III)- Le programme NooJ

Le Logiciel NooJ⁴ est créé par Max Silberztein en 2003 afin de dépasser les limites du programme INTEX dont il est issu. Son premier objectif est de construire des dictionnaires qui regroupent l'ensemble des flexions et des dérivations des mots. Le second est d'analyser automatiquement des textes aux niveaux lexical, morphologique, syntaxique et sémantique.

D'une manière générale pour utiliser NooJ, il faut⁵ :

- Un texte (ou corpus) que l'on ouvre en début de travail. Rangé par défaut dans "*Projects*".
- Des dictionnaires rangés par défaut dans le dossier "*Lexical Analysis*".
- Des graphes syntaxiques rangés dans le dossier "*Syntactic Analysis*".



Les fichiers à construire seront rangés automatiquement dans ces dossiers à l'aide des suffixes. Exemples :

.dic => Pour les fichiers contenant des dictionnaires non compilés, rangés dans : lexical analysis.

.nod => Pour les fichiers contenant des dictionnaires compilés, rangés dans : lexical analysis.

.nom => Pour les fichiers contenant des grammaires (graphes) morphologiques, rangés dans : lexical analysis.

⁴ www.nooj4nlp.net

⁵ Ces éléments ne sont pas tous indispensable en fonction du travail que l'on veut effectuer.

.nog => Pour les fichiers contenant des grammaires (graphes) syntaxiques, rangés dans : syntactic analysis.

.noj => Pour les fichiers mémorisés, rangés dans : projects.

Nous verrons la fonctionnalité de certaines parties du programme en détail dans le chapitre suivant consacré à l'application du logiciel aux verbes du kabyle.

IV)-Le processus suivi

1)- Préparation des entrées du dictionnaire

Cette étape consiste à installer dans la partie dictionnaire les verbes dans leur forme simple, suivie de code : (,V+FLX =). Ce code sert à relier le verbe avec son paradigme flexionnel qui sera construit dans la partie grammaire. Le nom du dictionnaire doit être précédé par le code (#use) et suffixé par (.dic). Ce nom doit être le même que celui de la grammaire à laquelle il sera lié.

Exemple des entrées du dictionnaire :

```

#
# Special Characters: '\\', ''', '+', ', ', '#', ' ', '#', '#', '#
#
#use exemple.nof

acar, V+FLX=
addi, V+FLX=
adef, V+FLX=
aden, V+FLX=
aqen, V+FLX=
ader, V+FLX=
adu, V+FLX=
aser, V+FLX=
aslay, V+FLX=
astad, V+FLX=
asu, V+FLX=
aszag, V+FLX=
af, V+FLX=
afeg, V+FLX=
afes, V+FLX=
ay, V+FLX=

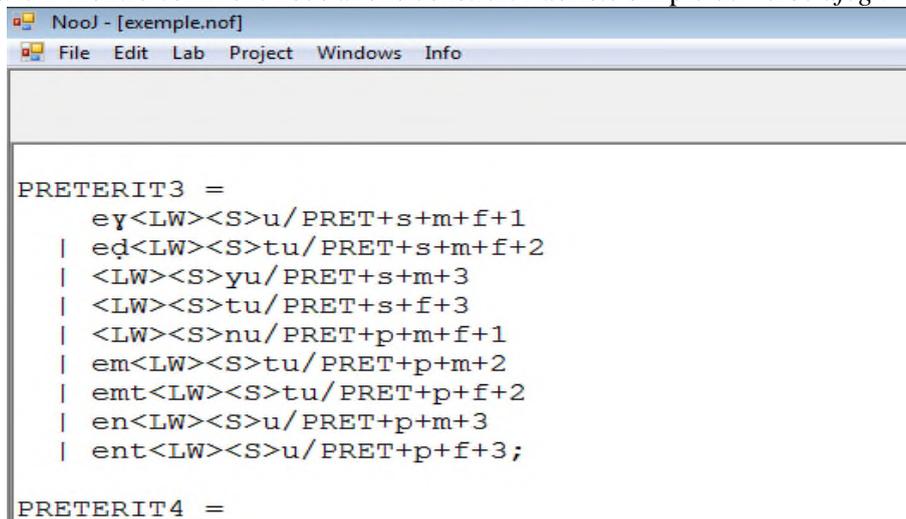
```

Figure 1

2)- Construction des paradigmes flexionnels

Cette étape se fait dans la partie grammaire. Il s'agit de construire toutes les flexions auxquelles les entrées du dictionnaire seront liées. Un paradigme flexionnel est l'ensemble des formes fléchies d'un verbe. L'obtention de ces formes se fait par la combinaison des codes et des graphes du logiciel. Dans notre travail, à la place des paradigmes verbaux nous avons choisi de construire des paradigmes d'aspects. Cette procédure nous a permis d'une part, de faire partager un nombre limité de modèles aspectuels avec un grand nombre de verbes et d'autre part, d'éviter de construire un nombre infini de modèles de verbes.

La *figure 2* montre comment nous avons construit l'aoriste simple du verbe *afeg* :



```

NooJ - [exemple.nof]
File Edit Lab Project Windows Info

PRETERIT3 =
  ey<LW><S>u/PRET+s+m+f+1
  | eq<LW><S>tu/PRET+s+m+f+2
  | <LW><S>yu/PRET+s+m+3
  | <LW><S>tu/PRET+s+f+3
  | <LW><S>nu/PRET+p+m+f+1
  | em<LW><S>tu/PRET+p+m+2
  | emt<LW><S>tu/PRET+p+f+2
  | en<LW><S>u/PRET+p+m+3
  | ent<LW><S>u/PRET+p+f+3;

PRETERIT4 =

```

Figure 2

Dans la première combinaison nous avons programmé le logiciel pour : ajouter *ey* au verbe, faire quatre sauts en arrière pour remplacer la voyelle (*a*) par la voyelle (*u*)⁶. Ces combinaisons se font dans la partie morphologie qui contient un

⁶ Le passage de *a* à *u* est une spécificité de certains verbes irréguliers.

outil qui sert à la manipulation (ajouter et /ou supprimer) de mot. Cet outil contient des options qui nous permettent de l'adapter à notre langue de travail. Pour notre langue nous avons ajouté deux options pour traduire la tension de *y* par *qq* et celle du *d* par *tt*.

Chaque paradigme aspectuel a un nom spécial et s'appliquera à tous les verbes de sa catégorie. L'**aoriste 3** s'applique à tous les verbes qui subissent les mêmes changements à l'aoriste simple comme *awed*, *ağew*, etc.

-Pourquoi des paradigmes d'aspects :

Pour le kabyle, contrairement à d'autres langues, à la place des paradigmes verbaux nous avons choisi de construire des paradigmes d'aspects. Pour le français par exemple, il suffit de faire un modèle pour un verbe du premier groupe et l'appliquer pour tous les verbes de ce groupe. Mais si on suit cette démarche pour les verbes du kabyle on arrivera à plus de 1000 modèles de verbes pour couvrir la liste de 3500 ; pour regrouper des verbes dans un modèle, il faut qu'ils partagent les mêmes changements. Dans notre langue, même les verbes que nous concéderons d'un même groupe peuvent avoir des spécificités différentes. Exemple :

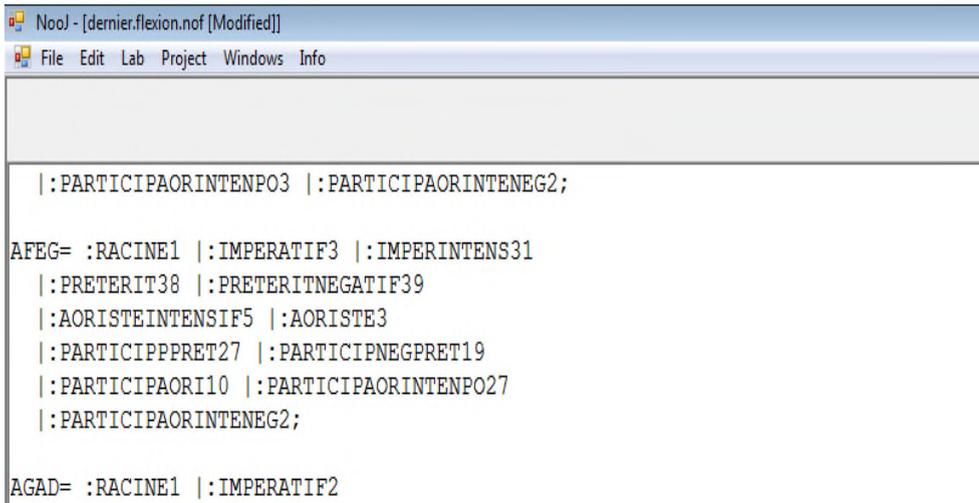
Les verbes : *afeg* (voler), *ayew* (avoir des neveux/nieces) et *awed* (arriver) sont différents pour les raisons suivantes : Premièrement, pour obtenir le prétérit et le prétérit négatif des verbes *afeg* et *ayew*, on remplace la voyelle initiale *a* par *u* par contre dans *awed*, on supprime *a* et on double la consonne *w*. Deuxièmement, le verbe *ayew* a deux formes intensives (*ttayew* / *ttayaw*).

Pour le logiciel donc, une lettre de plus ou de moins, peut exclure un verbe dans un modèle et nécessite la construction d'un autre. Pour cette raison, on a construit pour ces verbes par exemple les mêmes paradigmes aspectuels en ajoutant seulement un prétérit et un intensif de plus.

3)-Construction d'une liste de verbes types :

Il s'agit des verbes qui seront placés devant les entrées du dictionnaire, ils vont être précédés par (V+FLX=). Chacun représentera un ensemble de flexions d'une catégorie de verbes. Leur fonction est de renvoyer l'entrée du dictionnaire à son paradigme flexionnel.

La *figure 3* montre un paradigme flexionnel représenté par le verbe *AFEG* :



```

NooJ - [dernier.flexion.nof [Modified]]
File Edit Lab Project Windows Info

|:PARTICIPAORINTENPO3 |:PARTICIPAORINTENEG2;

AFEG= :RACINE1 |:IMPERATIF3 |:IMPERINTENS31
|:PRETERIT38 |:PRETERITNEGATIF39
|:AORISTEINTENSIF5 |:AORISTE3
|:PARTICIPPPRET27 |:PARTICIPNEGPRET19
|:PARTICIPAORI10 |:PARTICIPAORINTENPO27
|:PARTICIPAORINTENEG2;

AGAD= :RACINE1 |:IMPERATIF2

```

Figure 3

AFEG ne sera pas mis uniquement devant l'entrée de ce verbe *afeg* mais devant tous les verbes qui partagent les modèles : RACINE1, IMPERATIF3, IMPERINTENS31, etc. La notation des verbes types doit être en majuscule contrairement aux entrées.

La figure ci-dessous montre le dictionnaire à l'état final. Le nom du dictionnaire doit être le même que celui de la grammaire avec laquelle il sera compilé et doit être précédé par le code #use.

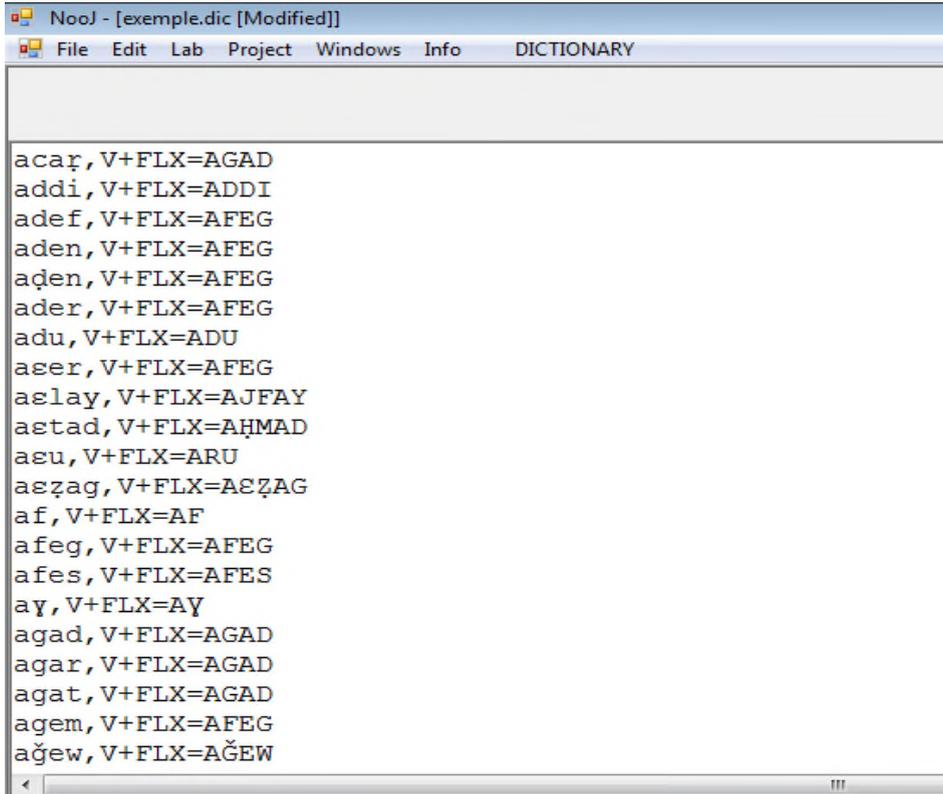


Figure 4

La première ligne veut dire : *acar*, verbe qui se fléchit selon le modèle *AGAD*.

V) - Compilation et résultat

Avant d'avoir toutes les formes fléchies de tous les verbes, ce dictionnaire doit être associé (compilé) à sa grammaire en suivant les instructions suivantes :

- Aller dans la rubrique : *Lab* > *dictionary* pour ouvrir l'outil de compilation.
- Cliquer sur *set* pour sortir le fichier qui contenant le dictionnaire à compiler.

- Revenir à l'outil de compilation et cliquer sur *compile*.
- Enfin en aura le dictionnaire (exemple-flx.dic) contenant toutes les formes fléchies.

La *figure 5* montre les flexions du verbe *acar* :

```

Dictionary contains 185090 entries

acar, acar, V+FLX=AGAD+IMP+s+m+f
acaret, acar, V+FLX=AGAD+IMP+p+m
acarem, acar, V+FLX=AGAD+IMP+p+m
acarem, acar, V+FLX=AGAD+IMP+p+m
acarem, acar, V+FLX=AGAD+IMP+p+f
ttacarem, acar, V+FLX=AGAD+IMPERINTENS+p+f
ttacarem, acar, V+FLX=AGAD+IMPERINTENS+p+m
ttacaret, acar, V+FLX=AGAD+IMPERINTENS+p+m
ttacar, acar, V+FLX=AGAD+IMPERINTENS+s+m+f
ucarent, acar, V+FLX=AGAD+PRET+p+f+3
ucaren, acar, V+FLX=AGAD+PRET+p+m+3
tucarem, acar, V+FLX=AGAD+PRET+p+f+2
tucarem, acar, V+FLX=AGAD+PRET+p+m+2
tucare, acar, V+FLX=AGAD+PRET+s+m+f+2
yucar, acar, V+FLX=AGAD+PRET+s+m+3
tucar, acar, V+FLX=AGAD+PRET+s+f+3
nucar, acar, V+FLX=AGAD+PRET+p+m+f+1
ucarey, acar, V+FLX=AGAD+PRET+s+m+f+1
ucarent, acar, V+FLX=AGAD+PRETNEG+p+f+3
ucaren, acar, V+FLX=AGAD+PRETNEG+p+m+3
tucarem, acar, V+FLX=AGAD+PRETNEG+p+f+2
tucarem, acar, V+FLX=AGAD+PRETNEG+p+m+2
tucare, acar, V+FLX=AGAD+PRETNEG+s+m+f+2
yucar, acar, V+FLX=AGAD+PRETNEG+s+m+3
tucar, acar, V+FLX=AGAD+PRETNEG+s+f+3
nucar, acar, V+FLX=AGAD+PRETNEG+p+m+f+1
ucarey, acar, V+FLX=AGAD+PRETNEG+s+m+f+1
ucarent, acar, V+FLX=AGAD+PRETNEG+p+f+3
ucaren, acar, V+FLX=AGAD+PRETNEG+p+m+3
tucarem, acar, V+FLX=AGAD+PRETNEG+p+f+2
tucarem, acar, V+FLX=AGAD+PRETNEG+p+m+2
tucare, acar, V+FLX=AGAD+PRETNEG+s+m+f+2
yucar, acar, V+FLX=AGAD+PRETNEG+s+m+3
tucar, acar, V+FLX=AGAD+PRETNEG+s+f+3
nucar, acar, V+FLX=AGAD+PRETNEG+p+m+f+1
ucarey, acar, V+FLX=AGAD+PRETNEG+s+m+f+1
ttacarent, acar, V+FLX=AGAD+INTEN+p+f+3
ttacaren, acar, V+FLX=AGAD+INTEN+p+m+3
tettacarem, acar, V+FLX=AGAD+INTEN+p+f+2
tettacarem, acar, V+FLX=AGAD+INTEN+p+m+2
tettacare, acar, V+FLX=AGAD+INTEN+s+m+f+2
yettacar, acar, V+FLX=AGAD+INTEN+s+m+3
tettacar, acar, V+FLX=AGAD+INTEN+s+f+3

```

Figure 5

Dans la deuxième ligne on peut lire : *acaret* est issu du verbe *acar*, qui se fléchi selon le modèle *AGAD* et il est à la 2^{ème} p. m. f. de l'impératif simple.

-Application de dictionnaire à un texte

1) Evaluation

En appliquant le dictionnaire sur un corpus, nous constatons deux genres de lacunes : le silence et le bruit.

1- Le silence (des formes verbales non signalées) : Nous relevons dans le texte les formes suivantes : *isselhu, sliley, issefhem, issewjed, issefqad, yettwareba, ssasayen, tettwaxdem, tennekmal, yesdukkel*. Ces formes sont toutes des dérivées, donc elles ne sont pas prises en compte dans notre travail consacré uniquement aux formes simples.

2- Le bruit (des formes non verbales peuvent être signalées) : Certains adverbes, affixes du verbe, prépositions et d'autres formes peuvent être considérés comme des verbes par le logiciel à cause de leurs formes identiques à celles du verbe (homographes).

Nous relevons dans le texte les cas suivants :

-Le nom propre *Rebbi* « dieu » est confondu avec l'impératif intensif de verbe *rbi* « surfaire », à la deuxième personne du singulier.

-Le complément, affixe de verbe *-ak* deuxième personne du singulier, est confondu avec la forme simple de verbe *ak* « pousser des soupirs ».

-La conjonction *am* « comme » est confondue avec la forme simple de verbe simple *am* « commencer, entreprendre », à la deuxième personne de l'impératif simple.

-Le pronom interrogatif *ayen* « ce que » est confondu avec la forme de l'aoriste simple de verbe *ay* « crier », à la troisième personne masculin pluriel (*ad*) *ayen* « ils crieront ».

-Le complément affixe de verbe, troisième personne singulier *-as* et son allomorphe *-yas* sont confondus avec : la forme simple de verbe *as* « venir » et avec la forme de l'aoriste simple à la troisième personne masculin (*ad*) *yas* « il viendra ».

-L'adverbe *tura* « maintenant » est confondu avec la forme du verbe *aru* « écrire » à la troisième personne du singulier féminin « elle a écrit ».

-Le nom féminin *lmut* « la mort », est confondue avec la forme de l'impératif simple pluriel masculin du verbe *lmu* « mariner ».

-Le nom masculin pluriel *medden* « les gens », est confondue, avec la forme du verbe *medd* « passer, donner de la main à la main », de la deuxième personne du pluriel (*ad*) *medden* « ils passeront/donneront de la main à la main ».

-Le pronom possessif de la troisième personne pluriel du singulier *-nsen* est confondu avec la forme de verbe *ens* « passer la nuit », à la troisième personne de pluriel masculin de l'aoriste simple.

-Le pronom personnel indépendant *nekk* « moi » est confondu avec la forme du verbe *ekk* « venir de, provenir » à la première personne du pluriel de l'aoriste simple.

-La conjonction de coordination alternative *ney* « ou, ou bien » est confondu avec la forme de verbe *ney* « tuer », à l'impératif simple singulier.

-Le pronom indéfini *nniden* « autre » est confondu avec la forme de verbe *nned* « enrouler, s'enrouler », à la troisième personne du pluriel féminin de prétérit négatif (*ur*) *nniden (ara)* « ils ne sont pas enroulés ».

-Le démonstratif *--nni* « ce, cette, ces, le, la, les... » est confondu avec trois formes de verbe *ni* « être enfilée (aiguille) » : celle de la première personne pluriel masculin de prétérit négatif, celle de la première personne pluriel de l'aoriste simple et celle de prétérit négatif.

-Le connecteur *ahat* « peut-être » est confondu avec la forme de verbe *aha* « se mettre, se hâter, faire vite », à la première personne de pluriel masculin de l'impératif simple.

-La préposition *ger* « entre, parmi » est confondue avec la forme de verbe *ger* « mettre, introduire », à la première personne de l'impératif simple.

-La préposition *yer* « vers, à » est confondu avec la forme de verbe *yer* « étudier », à la première personne de l'impératif simple.

2) Levée des ambiguïtés :

2.1) La racine : avant d'appliquer le dictionnaire au texte, nous avons écarté la racine en ajoutant devant chaque modèle le code # pour les classer parmi les commentaires non pris en compte par le logiciel. Cela nous a permis d'écarter certaines prépositions comme *n* (de, appartenant à) ou *s* (avec, au moyen de, par) qui peuvent être considérées comme racines de certains verbes comme *ini* (dire) ou *usu* (tousser).

2.2) Les homographes : pour les affixes des prépositions et tous les autres homographes nous envisageons de construire des dictionnaires d'affixes, de

prépositions, de noms, d'adverbes, etc., avec des graphes syntaxiques dont on indiquera les propriétés de chaque élément et les différentes possibilités qui permettront au logiciel de ne reconnaître que le verbe.

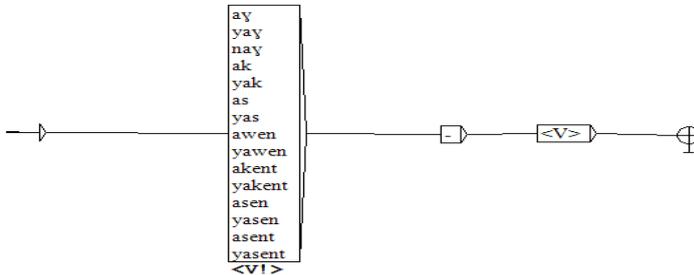
➤ **Exemples de désambiguïssations :**

Nous avons, ci-dessous, essayé de construire trois graphes afin d'éliminer les affixes de verbes et de prépositions qui constituent des cas d'homographies :

1- Affixes de verbes homographes de formes verbales :

Graphe (1) pour éliminer des homographes précédant un verbe :

r syntactic grammar.



Ce graphe traduit la structure suivante : si une de ces formes est préfixée à un verbe par l'intermédiaire d'un tiret, elle sera écartée de la catégorie des verbes.

- Exemples d'occurrences écartées dans le texte :

[Concordance for Text beleid at ali.not]			
Edit Lab Project Windows Info TEXT CONCORDANCE			
Display:	5 characters before, and	5 after.	Display: <input checked="" type="checkbox"/> Matches <input checked="" type="checkbox"/> Outputs
Before	Seq.	After	
k-yehwa, s wayen i	yak-yehwan/<V!>	; sakin ayen yellan a t	
atan nmektiy-d : sgellin, mi	yawen-udrey/<V!>	leqmer, ur ufiy iyer ara	
a mmi, ilaq-ay ad	ak-nezweg/<V!>	, ad ak-d-nawi tamettut	
d-yezzi useggas-agi ad	as-nanef/<V!>	i Seedi : atan lecyal-is	
yer tmurt n win akkenni	yas-yefkan/<V!>	lherz i Seedi. Abrid iqlul	
lhekma s tjennit, hed ur	as-yezmir/<V!>	. Walakin, si lgiha d-yerran	
yer tmezzyant, tis sebaa, ad	as-tiniq/<V!>	diyen : tif-itent yakk ! Dya	
-yas : _Kker a tedduq ad	ak-mley/<V!>	tin ara yi-txedmed ass	
awal : 11 tsellek-it akken a	yas-tenna/<V!>	, ney ur t-tsellek, w	
-t almi yerwa asnuqel, almi	yas-tekkes/<V!>	ssekran, tenna-yas : _Akkagi ay	
yettwaxdem xir n wakken i	yak-yecred/<V!>	baba. Atan yer-k kra	
-agi i baba akken i	yak-yenna/<V!>	; lameena, a k-weşşiy yef	
iyi-d-ttadred nekkini, anda	yas-teqqareq/<V!>	d nekk i yak-yeqdan	
yas-teqqareq d nekk i	yak-yeqdan/<V!>	tagi. Lukan ad slen netta	
slen netta d-yemma, ad	ay-ččen/<V!>	, ččen ula d tamurt iyef	
udem-is. Teud ulac wi	yasen-izemren/<V!>	, ata yusa-d win ara	
a twaliq, ixef-is... Ad	ak-teyli/<V!>	tcacit ; yerna, i wakken ur	
tamacahut : elaxater, şşbeñ-nni, mi	yas-nnan/<V!>	i Façma : Awi-yas aewin	
wass-a, taxatemi-iv ur	as-tezmir/<V!>	ara. Yella wayen i yi	
yeddelmen Seedi imi yuzi ad	as-vay/<V!>	awal. ara s-yinin : d	

635/635

-Résultat du graphe (1) :

Cette image nous montre un échantillon des formes désambiguïsées comme :

-*yak* : homographe de l'aoriste simple du verbe *ak* «pousser des soupirs » dans sa forme de la troisième personne du singulier.

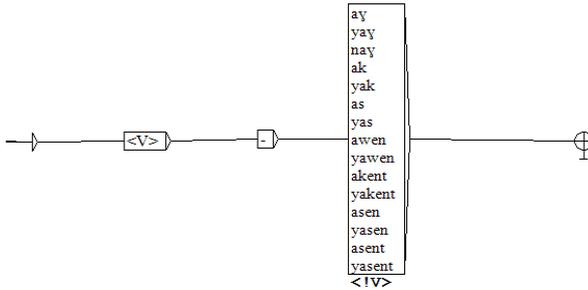
-*yas* : homographe du verbe *as* «venir » dans sa forme de la troisième personne du singulier.

-*ay* : homographe de la forme de l'impératif simple du verbe *ay* « acheter ».

- *Graphes syntaxiques (2) pour éliminer des homographes suivant un verbe*⁷

⁷ Ce graphe concerne les pronoms affixes suivant un verbe.

grammar.



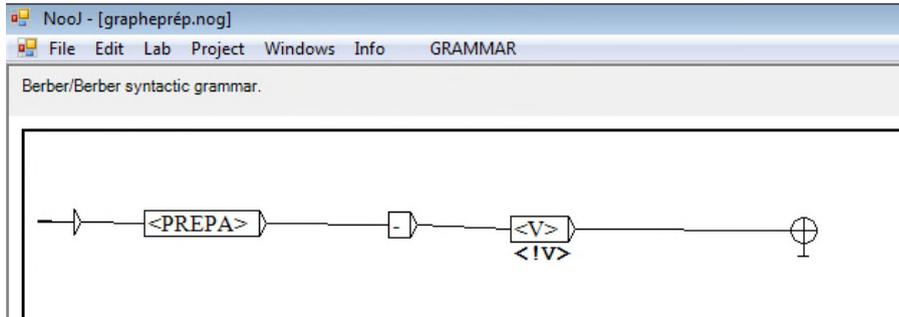
- Exemples d'occurrences écartées dans le texte :

Before	Seq.	After
tmurt anect-ila-tt. Timesejwt,	içher-ak/<IV>	ljal_ ulac win ikesben am
axxam tameddit, yettef mmi-s,	yenna-yas/<IV>	: _Tura a mmi, keçç aql
ak-nuqem tameyra, ncalleh. Yernu	yefka-yas/<IV>	yakk tisura n yexxamen, n
dimi t-issefhem akken iwulem,	yenna-yas/<IV>	Seedi mmi-s : 2 A k
ula d nekk aql-i	kksey-ak/<IV>	aybel : ur tesseid ara tħebbred
ama d ccyel nniçen yakk,	yessa-yasen/<IV>	ifellahen, açal ! Netta, d acu
ala tigi, berka-t. Seedi,	yerna-yas/<IV>	Rebbi tin ur d-ssasayen
-ssasayen la idrimen la ttrebga :	yefka-yas/<IV>	Rebbi şşifa d zzyen d
-t yid-s, yer dinna,	yenna-yas/<IV>	: _Akken a tferrgeç ! Iferreğ. Azekka
-yuyen ? A yuy ljal, medden	tħebbiren-as/<IV>	, ziy-nni netta ar yefka
ayen iwumi tezmer texdem-it,	tenna-yas/<IV>	-t ; lameena, ur tezmir i
deg wawal : win iyef keçtrent	yettes-aset/<IV>	: Ihi eni, ula d netta
ihusa-t-id umexluq-nni,	yenna-yas/<IV>	: _A gma, semmeħ-yi ma
wayen iss ara k-eiwney. _	Yerra-yas/<IV>	-d Seedi : _Ayen iyi-iblan
Amexluq-nni ihres-it-id,	yenna-yas/<IV>	: _Amen s win k-ixelqen
ul nniçen yeskaki-t-id :	yenna-yas/<IV>	: D acu txeşreç ? Argaz-agi
yella kra n lqedra yessa :	ħku-yas/<IV>	d acu k-yuyen, d
acu k-iruhej. Yettef Seedi	yeħku-yas/<IV>	-tt-id akken tella. Akken
akken tella. Akken dimi ifukk,	yenna-yas/<IV>	-d winna : _D aya i
Ulac i vsehlen am aya :	sey-as/<IV>	ddwa ara k-fkey tura

1397/1397

2- Affixes de prépositions, homographes de formes verbales :

- Graphe (3) pour éliminer des affixes de prépositions, homographes de formes verbales :



Remarque : <PREPA> est le nom d'un dictionnaire de prépositions construit dans la partie dictionnaire.

- Exemples d'occurrences éliminées dans le texte :

Before	Seq.	After
i yas-yefkan teswa, tecfa	fell-as/<IV>	. terfed allen-is yer tyeryert
tlernmast n lwali ur bnin	fell-asen/<IV>	. Akken dimi éčan, fučken imensi
-ay-d. Nekni, la tettvaliq	deg-ney/<IV>	di sebaa yid-ney d
-is. Asmi ten-tessen, tferreq	gar-asen/<IV>	, tuyal tettlaei-ten, thedder-as
d-tewweđ kan, a tali	fell-as/<IV>	, a tettlaei syün, a thedder
la sen-thedder ; almi ieedda	fell-asen/<IV>	alkden lan, d wamek i
Tebda cwit tasa-nsent tergagi	fell-as/<IV>	, yerna uread ti-zin d
sidi-s, ar d leib	fell-as/<IV>	. Terr-as-d, s leeqel
-s asmi i yay-łkan	fell-as/<IV>	. Ney tugadem ad tuy indunen
tula d nettat, ar ieedda	fell-as/<IV>	wayen-nni ; walakin, nettat yerra
axir-as : ur teddi ara	fell-as/<IV>	temlel. 68 Yettef-itt akkenni Seedi
A tamurt idelawen, Atna wulawen	Fell-am/<IV>	ryan d ssejra !... Akken dimi
ur yejli w'ara nagad	fell-as/<IV>	. Atmaten-is di sebaa yid
ladya i uštati-nni testatay	fell-asen/<IV>	ass kamel ; anda yella lexbar
wass-en d asawen ar	fell-asen/<IV>	mkul ass d tamejra, elaxađer
tmurt-nsent ajas. Teyli-d	fell-asen/<IV>	tmeddit, yejli-d lneyreb, yejli
yiwet tejra d taslayant. Ulint	fell-as/<IV>	almi d taqacuct, qimint dınna
-itt, teyli-d yer lqaea	gar-asen/<IV>	. Di dđđiqqa tuyal d ticrihin
irebbi-s. Almi d-teyli	fell-as/<IV>	tmeddit, tessawed-d yer 78 yiwet
tekcom yar-sen Fadma, rebhen	fell-as/<IV>	. am akken tewwi-d sseed

Résultat du graphe (3) : cette image nous montre l'élimination de quelques affixes homographes de formes verbales à savoir :

-*as* : homographe du verbe *as* «venir» dans sa forme de l'impératif simple, singulier.

-*asen* : homographe de l'aoriste simple du verbe *as* «venir» dans sa forme de la troisième personne du pluriel masculin.

-*asent* : homographe de l'aoriste simple du verbe *as* «venir» dans sa forme de la troisième personne du pluriel féminin.

-*am* : homographe de l'impératif simple, singulier du verbe *am* «commencer / entreprendre».

-*ney* : homographe de l'impératif simple, singulier du verbe *ney* «tuer».

VI)- Conclusion et perspectives :

On a confirmé à travers ce travail que le logiciel NooJ peut être adapté aux spécificités morphologiques du kabyle. Néanmoins, si l'orthographe n'est pas fixée de manière définitive, l'intégration du kabyle dans le TAL sera une tâche difficile. Dans notre cas, même si nous avons eu recours aux recommandations de l'INALCO, nous étions obligé de faire un travail plus détaillé et précis.

Nous avons essayé d'intégrer le maximum de formes verbales que nous pourrions trouver dans un texte à traiter. Exemples : (*acar/ččar/, bded/ bedd, etc.*) ou, dans le cas de l'intensif, (*keččem/ttekcām, ttruḥ / ttruḥu* ou *deyyel / tdeyyil / ddeyyil / ttedyal, etc.*). Ce que nous avons fait jusqu'à présent peut permettre au logiciel de reconnaître les différentes formes fléchies d'un verbe primaire. Néanmoins, les verbes dérivés sont aussi fréquents, la seconde tâche donc, sera l'intégration de cette catégorie. Dans d'autres perspectives, comme le montre le texte, nous devons, passer au traitement syntaxique de tous les éléments qui permettront au logiciel d'écarter toute forme d'ambiguïté.

-Codes et abréviations utilisés :

<LW> => Pour revenir tout en arrière du mot.

<RW> => Pour aller tout devant.

<L>	=> Pour aller de droite à gauche du mot.
<R>	=> Pour aller de gauche à droite du mot.
	=> Pour effacer un caractère on commençant de droite à gauche
<S>	=> Pour effacer un caractère on commençant de gauche à droite.
<E>	=> Pour garder le mot tel qu'il est.
<D>	=> Pour doubler le caractère.
" "	=> Pour laisser un vide.
/	=> Pour passer de la forme fléchie à l'aspect et à la personne qui lui correspond.
	=> Pour passer d'une forme à une autre à l'intérieur d'un même paradigme.
;	=> Fin de la flexion d'un paradigme.
+	=> Pour passer d'une indication à une autre (féminin singulier => f+s).
IMP	=> Pour désigner l'impératif.
IMPERINTENS	=> Impératif intensif.
PRET	=> Prétérit.
PRETNEG	=> Prétérit négatif.
INTEN	=> Aoriste intensif.
AOR	=> Aoriste.
PARTICIPPPRET	=> Participe positif du prétérit.
PARTICIPNEGPRET	=> Participe Négatif du prétérit.
PARTICIPAOR	=> Participe de l'aoriste simple.
PARTICIPAORINTENPO	=> Participe de l'aoriste intensif positif.
PARTICIPAORINTENEG	=> Participe de l'aoriste intensif négatif.
<!V>	=> L'unité désignée n'est pas un verbe.

Bibliographie :

- B. PIERRETTE & *al.*, *Traitement automatique des langues naturelles*, Duculot, Belgique, 1998.
- C. FUCHS, *Linguistique et traitement automatique des langues*, Hachette, France, 1993.
- J.-M. DALLET, *Dictionnaire Kabyle-Français, Parler des at Manguellat Algérie*, SELAF, Paris, 1982.
- J.-M. DALLET, *II Dictionnaire Français-Kabyle*, SELAF, Paris, 1985.
- K. NAÏT-ZERRAD, *Manuel de conjugaison kabyle*, L'Harmattan, Paris, 1994.
- K. NAÏT-ZERRAD, *Grammaire moderne du kabyle, tajerrumt n teqbaylit*, Karthala, Paris, 2001.
- K. NAÏT-ZERRAD, *Mémento grammatical et orthographique de berbère, Kabyle-Chleuh-Rifain*, L'Harmattan, 2011.
- S. CHAKER, *un parler berbère d'Algérie (Kabylie), syntaxe*, Thèse de troisième cycle, université de Provence 1978.

Ressources électroniques :

<http://www.NooJ4nlp.net>