



Des traitements automatiques en lexicographie berbère Automatic processing of Berber lexicography

Patrice Pognan¹

¹Institut National des Langues et Civilisations Orientales, France, membre de PLIDAM

¹Université de Prague, République tchèque, membre de ÚFAL

patrice.pognan@gmail.com

Article information

History of the article- Historique de l'article

Received: 25/11/2021

Accepted : 13/12/2021

Published : 31/12/2021

Abstract

This article provides an overview of the steps involved in the computerization of Miloud Taïfi's "Berber-French reasoned dictionary, Moroccan vernacular" required for its set-up (structured dictionary corpus ; two indices) or for applications developed after its publication (dictionary with examples). The latter can now be viewed on line at www.linguotheque.huma-num.fr. Other works conducted with this dictionary (conjugator; dictionary look-up from French) will make use of the online description of Kabyle nominal and verbal morphology.

Keywords: Berber-French dictionary, Dictionary of Examples, Moroccan Vernacular, Computerization.

Résumé

Cet article présente un ensemble de travaux informatiques réalisés dans le cadre du « Dictionnaire raisonné berbère-français, parlars du Maroc » de Miloud Taïfi pour sa réalisation (corpus dictionnaire structuré ; deux index) ou pour des applications postérieures à sa publication (exemplier). Ce dernier est actuellement accessible en ligne sur le site www.linguotheque.huma-num.fr. D'autres travaux issus de ce dictionnaire (conjugueur ; consultation du dictionnaire par une entrée française) profiteront de l'expérience de la présentation en ligne de la morphologie nominale et verbale du kabyle.

Mots clés : dictionnaire berbère-français, exemplier, parlars du Maroc, traitements informatiques.

Agzul

Amagrad-a, d asenked n yimahilen isenselkimen yettwaxedmen deg « Usegzawal unziž tamaziyt-tafransist, tantaliwin n Lmaruk » n Milud Tayfi ama deg usewjed-ines (amud s wudem n usegzawal ; sin n yimataren) ama deg yisnasen ara yilin yer sdat (amud n yimediyaten). Amud n yimediyaten, yezmer yiwén ad t-iwali deg wadeg www.linguotheque.huma-num.fr. Imahilen nniden n usegzawal-a (aseftay n yimyagen, asexdem n usegzawal s yimsudaf n tefransist) ad lhun i usenked n tasnalɣa n yisem d umyag n teqbaylit.

Awalen-tisura : Asegzawal Tamaziyt-Tafransist, amud n yimediyaten, tantaliwin n Lmaruk, isekkiren isenselkimen



Les premiers travaux que Miloud Taïfi et moi-même avons réalisés pour la fabrication du dictionnaire berbère - français (Taïfi & Pognan, 2017) ne sont pas représentatifs de ce qu'il faut faire dans l'idéal. En effet, on construit généralement une base de données à partir de laquelle on pourra fabriquer le dictionnaire. Cela n'a pas été le cas en ce qui nous concerne, comme je vais l'exposer infra.

Par contre, au fil des années, les travaux se sont ajoutés les uns aux autres pour constituer un ensemble susceptible d'intérêt que je me permets de présenter au lecteur.

I. Les réalisations passées

Lorsque Miloud Taïfi et moi-même avons, en 2004, lors de notre deuxième action intégrée franco-marocaine, envisagé des travaux de lexicographie du berbère du Maroc, Miloud Taïfi était déjà l'auteur d'un gros dictionnaire des parlers du Maroc central vers le français (Taïfi 1992).

Il y avait donc des acquis qu'il convenait de ne pas négliger. La prise en compte de ce premier dictionnaire a été difficile du fait qu'il était manuscrit et qu'il s'est révélé impossible de recourir à la lecture optique qui donnait de très médiocres résultats où tout était à relire et corriger. Le contenu a été ressaisi sous la forme de 29 fichiers Word représentant chacun un caractère de l'alphabet.

Comme l'écrit Miloud Taïfi au début de son introduction au dictionnaire publié à l'IRCAM de Rabat en 2017 :

Ce dictionnaire est une version corrigée, augmentée et remaniée de l'ouvrage intitulé Dictionnaire tamazight-français (parlers du Maroc central) publié en 1992. Le but visé étant de reconstituer le système de la langue berbère, cette nouvelle version est suffisamment différente de la première, dans plusieurs de ses aspects, pour en proposer un autre titre qui est désormais : *Dictionnaire raisonné berbère-français. Parlers du Maroc*. Cela veut dire que la masse lexicale consignée dans cet ouvrage n'est plus, comme dans le précédent, confinée dans le dialecte tamazight qui regroupe les parlers pratiqués dans le Maroc central, mais comporte aussi, comme il est expliqué ci-après, des formes de mot et/ou des significations nouvelles attestées dans d'autres zones géolinguistiques berbérophones du Maroc.

Nous avons longuement discuté des questions de système linguistique, puis travaillé à l'élaboration d'une écriture en caractères latins que nous avons qualifiée de « systémique » donnant à voir la mécanique de la langue. Le nouveau dictionnaire possède près de 50% de racines en plus.

J'ai transformé ces 29 fichiers Word en fichiers bruts (extension .txt) en Unicode UTF-8, ce qui permet une programmation aisée, ici en Python. Ils ont été réunis dans un fichier unique de 2 700 000 caractères.

J'ai appliqué à ce fichier deux séries de programmes. La première contient 5 programmes dont le but est de développer les structures du dictionnaire par des indentations qui conservent les symboles typographiques utilisés par Miloud Taïfi pour distinguer différents niveaux : entrées par racines, niveau des mots, niveaux des significations, niveau des exemples, ... On obtient ainsi un corpus structuré « brut » qui permettra

ensuite toutes les transformations et tous les traitements que l'on souhaitera. La seconde série (3 programmes) calcule notamment toutes les formes flexionnelles des principales catégories lexicales : verbes, substantifs et adjectifs. C'est à partir de cette forme développée que peuvent être créées diverses applications.

Fig. 1. Corpus dictionnaire structuré

SBH	SBH1	ar., kb.							
◆ 1	sebbəh	SBH1	1.1.	VN	V	sebbəh	sebbəh	tsebbəh	ur-sebbih
	► arriver à, être à ... le matin ; rencontrer ... le matin								
	■ few, FW								
	► souhaiter le bonjour à ...								
	• la d tsebbəhn xef yiməngan, ils se disputent dès leur réveil, le matin								
	• wenna isebbəhn xef ubağus, han şşbah-nes ur izill, celui qui rencontre un singe le matin (son matin n'est pas bon), doit s'attendre à quelque malheur (croyance populaire)								
◆ 2	msebbəh	SBH1	1.2.	VN	V	msebbəh	msebbəh	temsebbəh	ur-msebbəh M
	► se souhaiter réciproq. le bonjour								
	• awd inedlibn la temsebbəhn ass n lcid, même les ennemis se souhaitent le bonjour le jour de l'Aid								
◆ 3	aşebbəh	SBH1	1.3.	VN	SUBST M	aşebbəh	uşebbəh	isebbih	yisebbih
	► première rencontre le matin ; premier acte ou événement de la journée								
	• aşebbəh n uwtul ixxa, la rencontre d'un lièvre, le matin, est de mauvais augure								
◆ 4	şşbah	SBH1	1.3.	VN	SUBST M	şşbah	şşbah		
	► matin, matinée								
	■ tıfawt, FW								
	• g şşbah, au matin, le matin								
	• seg şşbah, depuis le matin								
	• şşbah zikk, le matin de bonne heure								
	— şşbah lıxir (matin de bien), bonjour.								
	— şşbah n teslit (le matin de la mariée), petit déjeuner (généralement soupe et crêpes) offert par les parents de la mariée aux danseurs et aux invités le lendemain de la consommation du mariage.								
◆ 5	tasebhıt / tisebhıyt	SBH1	1.3.	VN	SUBST F	tasebhıt / tasebhıyt	tsebhıt / tasebhıyt	tisebhın / tisebhıyn	
	► matin, matinée								

La figure ci-dessus montre une racine verbo-nominale (VN) avec un verbe simple (1.1.), un verbe dérivé (1.2.) et des substantifs dérivés (1.3.). Les mots et racines synonymes sont indiqués par un carré, les exemples par un rond et les locutions par un tiret long. On trouvera une description détaillée du corpus et de sa construction dans quelques articles antérieurs (Pognan & Taifi 2010, 2011).

Ainsi constitué, ce corpus peut être largement exploité pour créer des applications. Les deux premières ont été publiées avec le dictionnaire. Il s'agit de deux index. Le premier assigne à chaque mot du dictionnaire sa racine. Cela permet une recherche plus facile de l'entrée en cas de doute, le dictionnaire ayant un classement par racines. Le second est beaucoup plus intéressant: il permet une consultation du dictionnaire par l'intermédiaire d'une entrée française. Les quelques témoignages spontanés que nous avons pu recueillir expriment une réelle satisfaction et témoignent de l'utilité de ce dispositif.

D'autres applications peuvent être envisagées, notamment des lexiques spécialisés, mais aussi un conjugueur automatique, un « exemplier » et la consultation du dictionnaire par l'intermédiaire d'une entrée française.

Pour des raisons de simplicité, c'est l'exemplier, dit « exemplier Taifi », qui a été construit en premier. L'article (Pognan 2013a) en donne une vision précise. Cet exemplier se présente sous la forme d'une base de données Access interrogeable. Le principal inconvénient est qu'une base de données Access est dans l'immense majorité des cas une base de données

unipersonnelle. Dès le départ, l'idée de cet exemplier en ligne à l'aide de MySQL ou PostgreSQL était présente. Ce n'est qu'en 2021 que cette idée a pris forme. Elle constitue l'objet principal de cet article.

La consultation de l'exemplier apporte une autre idée: son utilisation pour illustrer une grammaire par des exemples authentiques tirés principalement de la littérature orale et surtout son exploitation comme support à la réalisation d'une méthode d'apprentissage. Cette dernière extension a été présentée il y a déjà quelques années (Pognan 2013b).

Fig. 2. L'interface Access de l'exemplier Taïfi



II. Les travaux actuels

Des applications rendues possibles par les travaux menés sur le dictionnaire, c'est certainement « l'exemplier » qui était le plus adéquat pour se lancer dans l'apprentissage de la réalisation d'un logiciel en ligne. En effet, c'est vraisemblablement la seule application qui puisse se contenter d'une table unique sur le serveur, au moins pour un usage simple.

L'exemplier est hébergé sur un serveur sécurisé (https) de Humanum, la TGIR (très grande infrastructure de recherche) consacrée aux humanités numériques.

II.1. La base de données

Pour mener à terme cette opération, il est nécessaire de savoir utiliser le gestionnaire de bases de données MySQL et de connaître la syntaxe SQL pour des opérations de chargement, de modification ou de lecture. L'application phpMyAdmin permet un accès facile à la base de données et une vision claire des différents contenus:

Fig. 3. Extrait de la table unique wp_exemplier

				ID	type	exemple	berbere	trad_litt	francais
<input type="checkbox"/>				848	expression	• a ay tru ! (litt. ô ce (que) elle a pleuré), elle a pleuré beaucoup !	a ay tru !	ô ce (que) elle a pleuré	elle a pleuré beaucoup !
<input type="checkbox"/>				849	exemple	• a ayd iswa d aman !, que d'eau il a bu !	a ayd iswa d aman !		que d'eau il a bu !
<input type="checkbox"/>				850	exemple	• a ay tğuda !, qu'elle est belle !	a ay tğuda !		qu'elle est belle !
<input type="checkbox"/>				851	exemple	• a wa awra !, (ô celui-ci, viens), ô toi, viens donc !	a wa awra !	ô celui-ci, viens	ô toi, viens donc !
<input type="checkbox"/>				852	expression	• a ta ultma ay tğid (litt. ô celle-ci ma soeur ce (que) tu es), ô toi, c'est ma soeur que tu es	a ta ultma ay tğid	ô celle-ci ma soeur ce (que) tu es	ô toi, c'est ma soeur que tu es
<input type="checkbox"/>				853	exemple	• ggullix aya nekk usar as siwlex, j'ai juré de ne plus jamais lui parler	ggullix aya nekk usar as siwlex		j'ai juré de ne plus jamais lui parler
<input type="checkbox"/>				854	exemple	• aya nekk mi tt isežra, malheur à moi à qui il a joué ce tour !	aya nekk mi tt isežra		malheur à moi à qui il a joué ce tour !
<input type="checkbox"/>				855	expression	• irdel a-s leflus (litt. il a prêté à lui de l'argent), il lui a prêté de l'argent	irdel a-s leflus	il a prêté à lui de l'argent	il lui a prêté de l'argent
<input type="checkbox"/>				856	exemple	• isifd a-wn d leflus, il vous a envoyé de l'argent	isifd a-wn d leflus		il vous a envoyé de l'argent
<input type="checkbox"/>				857	exemple	• isiwl a-m, il t'a parlé (fém)	isiwl a-m		il t'a parlé (fém)
<input type="checkbox"/>				858	exemple	• idea a-x, il nous a cités en justice	idea a-x		il nous a cités en justice

Cette table reprend la structuration de la base de données Access, à savoir une identification de chaque enregistrement (ID), le type de l'enregistrement (locution, exemple, dicton, chant, devinette, ...), l'exemple en entier tel qu'il était donné dans le dictionnaire, la partie berbère de l'exemple, la traduction littérale en français de l'exemple berbère lorsque cela est nécessaire et enfin la traduction française. Les processus automatiques de découpage des exemples sont exposés dans l'article (Pognan 2013a).

II.2. L'interface en ligne et le fichier correspondant

Pour pouvoir exploiter en ligne les informations de l'exemplier, il est nécessaire d'avoir une page permettant l'interrogation. Cette page présente l'alphabet utilisé dans l'écriture systémique qui sous-tend le dictionnaire:

Fig. 4. Présentation de l'alphabet latin utilisé

Berbère: alphabet utilisé dans l'exemplier:

A a	B b	D d	Ḍ ḍ	F f	G g	Ġ ġ	H h	Ḥ ḥ	X x
I i	Ž ž	K k	L l	M m	N n	Q q	R r	S s	Ş ş
Š š	T t	Ṭ ṭ	U u	W w	Y y	Z z	Ẓ ẓ	ε ε	

La page d'interface présente une possibilité de recherche par le français et par le berbère (les mots doivent être entrés dans l'alphabet du dictionnaire). La recherche par le français se fait dans les deux champs « traduction littérale » (trad_litt) et « français » et pour le berbère à partir du champ « berbère ». Ce qui sera renvoyé à l'utilisateur sera, pour chaque enregistrement adéquat, le champ « exemple » précédé du champ « type ».

L'interface utilisateur se présente ainsi:

Fig. 5. Interface utilisateur

31 MARS 2021 PAR PATRICE POGNAN
EXEMPLIER "TAÏFI"

Consultation de l'exemplier "Taïfi"

Choisissez l'un des deux modes de saisie ci-dessous

Entrez un mot français

Entrez un mot berbère (tamazight du Maroc central)

udem

Valider Vider les données

Valider Vider les données

Même si l'on peut construire une partie de l'interface utilisateur de manière intuitive dans un éditeur pour le concepteur, dit « visuel », c'est dans tous les cas un fichier HTML implémenté par l'intermédiaire d'un « éditeur de code » qui produit l'interface. Dans le fichier HTML, un appel est fait à un programme PHP `exemplier-FRA.php` pour le français et au programme `exemplier-BER.php` pour le berbère.

II.3. Les fichiers d'exploitation de la base de données

Les deux fichiers écrits en PHP (`exemplier-FRA.php` et `exemplier-BER.php`) font appel à une feuille de style d'un logiciel supplémentaire « CSS » qui permet une mise en page et une typographie riches pour améliorer les performances du HTML. Ces fichiers ont pour première tâche de récupérer par l'intermédiaire d'une variable (`$fra` ou `$ber`) le mot recherché par l'utilisateur.

Pour le français, la variable `$fra` est recherché dans les deux champs traduction littérale et français de la base de données. Si aucune occurrence n'est trouvée, un message approprié est renvoyé à l'utilisateur. Dans le cas contraire, c'est la liste de tous les enregistrements dans lesquels le mot recherché a été trouvé qui est renvoyée à l'utilisateur à l'aide d'une fonction d'impression qui colorie en bleu la première (ou l'unique) occurrence du mot recherché.

Le programme pour le berbère possède la même structure, mais est plus intéressant. A partir de la variable `$ber`, le programme calcule plusieurs formes candidates pour la recherche. C'est en premier lieu la forme entrée par l'utilisateur, puis, si cette forme contient un schwa (écrit « e »), on produit une forme sans schwa. La forme annexée de chacune de ces deux formes est ajoutée à la liste. Sont ensuite calculées d'éventuelles formes féminines (s'il y a lieu) libres et annexées. Le calcul des pluriels est très complexe et est seulement en cours d'étude. Il pourrait y avoir une alternative plus productive: la consultation des unités lexicales issues du dictionnaire et rassemblées dans une autre table. Mais je préfère largement un logiciel qui calcule à un logiciel qui ne fait que consulter des tables.

II.4. Les résultats présentés à l'utilisateur

Fig. 6. Extraits de résultat pour le mot « udem »

Exemplier Taïfi

Il contient 13641 exemples et locutions

udem

- locution....— **udem** n ubağus (visage de singe), laid, laideron.
- locution....— **udem** n wakal (litt. visage de la terre), ras-de-sol.
- locution....— **udem** n ulemisir (litt. visage de la peau de mouton ...), visage jaunâtre, maladif.
- locution....— **udem** n řebbi (visage de Dieu) ; g wudem n řebbi, pour l'amour de Dieu ; gratuitement.
- exemple.....● isbenter-as usafar nna da iteswa **udem**, le remède qu'il prend lui a enflé le visage
- expression...● ifka **udem** i lewcer (litt. il a donné la face à un endroit escarpé), il s'est dirigé vers un endroit escarpé, fig. : il s'est mis dans une situation difficile
- exemple.....● tewser ay da ax itsekruššum **udem**, c'est la vieille qui nous ride le visage
- expression...● ikkes **udem** i terbat (litt. il a ôté le visage à la fille), il a défloré la fille
- exemple.....● isemsex **udem** ġer midden, il s'est avili aux yeux de tous
- exemple.....● iřewđ **udem**-nes s tafukt, il a le visage brûlé par le soleil
- exemple.....● itkur **udem**-nes s ukeffis, son visage est souillé de suie
- exemple.....● sity d **udem**-nek ad raeax is yarid, approche ton visage pour que je voie s'il est propre
- exemple.....● amterreħ teđher g **udem**-nes, quand qn. est heureux (mène une existence douce) cela se voit sur son visage
- locution....— **udm** amazdar, revers (d'une chose).
- exemple.....● **udm** n řřbel, (visage de réussite), personne de bon augure, porte-bonheur
- expression...● **udm** n lexřart (visage de perte), personne de mauvais augure, porte-malheur
- expression...● uħeqđ **udm** n řebbi (formule de serment), par la face divine
- expression...● qeddemx-ak nn **udm** n řebbi (je mets devant toi le visage de Dieu), je t'en supplie, je t'en conjure
- expression...● yanney-as **udm** (litt. il a vu à lui le visage), il a pris en considération sa demande (parce qu'il le respecte)
- exemple.....● **udm** n wakal, ras de sol, surface de la terre
- exemple.....● mer ġur-s **udm**, ur itadža ibba-s g tmar, s'il avait de l'amour-propre, il n'aurait pas laissé son père dans la misère

Les extraits ci-dessus montrent que le programme a renvoyé les formes « udem » et « wudem » ainsi que les formes « udm » et « wudm » affixées (uniquement suffixes pour les substantifs, préfixes et suffixes pour les verbes) ou non, marquées en bleu. L'exemplier peut être consulté à <https://linguothèque.huma-num.fr/exemplier-taifi/>

III. Les prolongements à donner

La mise en ligne de l'exemplier et l'apprentissage, au moins élémentaire, des multiples logiciels nécessaires permet d'envisager plusieurs applications dérivées du dictionnaire Taïfi (III.1. ci-dessous) et de la coopération avec les collègues de l'Université de Tizi-Ouzou (III.2.).

La toute première application découle directement de l'exemplier en ligne : l'utilisation des résultats pour illustrer une grammaire en construction ou soutenir la réalisation d'une méthode d'apprentissage. En ce qui concerne la tamazight du Maroc central, il n'existe guère que la méthode de Bisson

qui est déjà ancienne (Bisson 1940). Ce prolongement avait débuté avec la base de données en Access et il est activement repris actuellement. Les relevés sont classés par thèmes susceptibles de constituer des leçons : temps, lieu, habitat, artisanat, animaux domestiques, animaux sauvages, ... Ce sera, bien évidemment, un projet collectif.

III.1. au dictionnaire raisonné berbère – français

En dehors de ce qui découle directement de l'exemplier, il y a au moins deux applications privilégiées issues du dictionnaire, un conjugueur et l'utilisation du dictionnaire par l'intermédiaire du français.

III.1.1. Elaboration d'un conjugueur automatique en ligne

De la même manière que l'extraction des exemples, j'ai procédé également à l'extraction de certaines catégories lexicales, notamment le verbe. C'est un ensemble de 5000 verbes qui a été constitué dont je donne ci-dessous un extrait :

Fig. 7. Extrait de la liste des verbes pour la réalisation d'un conjugueur en ligne

şaḥa	ŞḤ2	1.1.	VN	V	şaḥa	şaḥa	tşaḥa	ur-şaḥa	
şleb	ŞLB1	1.1.	VN	V	şleb	şleb	teşlab	ur-işlib	
tuşleb	ŞLB1	1.2.	VN	V	tuşleb	tuşleb	ttuşlab	ur-tuşlib	Tu
şelleb	ŞLB2	1.1.	VN	V	şelleb	şelleb	tşellab	ur-şellib	
şleḥ	ŞLḤ1	1.1.	VN	V	şleḥ	şleḥ	teşlaḥ	ur-şliḥ	
şleḥ	ŞLḤ2	1.1.	VN	V	şleḥ	şleḥ	teşlaḥ	ur-şliḥ	
şalḥ	ŞLḤ2	1.1.	VN	V	şalḥ	şalḥ	tşalaḥ	ur-şalḥ	
tuşleḥ	ŞLḤ2	1.2.	VN	V	tuşleḥ	tuşleḥ	ttuşlaḥ	ur-tuşliḥ	Tu
mşalaḥ	ŞLḤ2	1.2.	VN	V	mşalaḥ	mşalaḥ	temşalaḥ	ur-mşalaḥ	M
semşalaḥ	ŞLḤ2	1.2.	VN	V	semşalaḥ	semşalaḥ	tsemşalaḥ	ur-semşalaḥ	
Sm									
şleḥ	ŞLḤ3	1.1.	VN	V	şleḥ	şleḥ	teşlaḥ	ur-şliḥ	

Les verbes simples ou dérivés présentent tous les 4 formes fondamentales (aoriste, prétérit, aoriste intensif, prétérit négatif) nécessaires à la production automatique de toutes les formes. Les procédures informatiques afférentes sont en grande partie prêtes ayant été réalisées pour un projet sur le kabyle (cf. III.2.).

III.1.2. Consultation en ligne du dictionnaire raisonné berbère - français par l'intermédiaire du français

Il s'agit d'une application majeure dont la réalisation a débuté. Elle doit permettre la consultation en ligne du « dictionnaire raisonné berbère - français. Parlers du Maroc » par l'intermédiaire du français.

C'est l'index 2 que j'ai construit en annexe du dictionnaire qui permettra cette réalisation. On y trouve sous une entrée française, plus ou

moins complexe, les racines berbères correspondant aux différentes acceptions du mot français.

La circulation au sein de l'entrée française pour arriver aux racines berbères nécessite la construction de 4 tables. Les deux premières sont montées sur le serveur. Un long travail manuel de gestion des doublons (130) est en cours pour assurer la circulation entre ces deux tables. Les deux dernières tables seront montées lorsque ce point sera résolu.

Ensuite, d'autres tables devront contenir le corpus dictionnaire suivant les différents types d'enregistrement (cf. I. Les réalisations passées). Chacune des racines berbères concernées devra retourner l'ensemble des informations la concernant. Les procédures informatiques nécessaires sont en grande partie également élaborées grâce au projet sur le kabyle (III.2.).

Fig. 8. Extrait de l'index français - racines berbères

abreuver (du bétail à un point d'eau): ĠB4	accablé: LF3, NBR1
abreuvoir: SRŽ2, ŠRŽ3	~ de malheurs: ĤDŠ1
abri: DRG1, DRY3, NTL1, SFL2, SNFY1	~ de soucis: ĠF1
~ à bétail: NWL2	~ par les fortes chaleurs de l'été: GLF3
~ pour se protéger: ŽY1	accablement: DĚQ1, ŠHM1
~ sous roche: FR10	accabler: DĚQ1, ĎR5, NBR1, NĠ2, QHR1, ŠHM1
se mettre à l'~: DRY3	~ de problèmes: ŠF3
tenir à l'~ des regards: ĤŽB1	accalmie: NFŽ2
abricot: MŠ8	accéder (aux désirs): NĚM1
abricotier: MŠ8	accélérer: DMR1, MR2, Z3
abrité: SDRY1	accentuer (s'~): NHM1
abriter (s'~): DRY3, NTL1, SDRY1, SFL2	

III.2. traitements automatiques du kabyle et apprentissage de la langue

Un projet réalisé en collaboration avec mes collègues de Tizi-Ouzou cherche à mettre en place un environnement d'apprentissage du kabyle.

Il est basé sur notre programme d'analyse de la morphologie du kabyle (Pognan, Sadi, Fitas & al. 2018) que nous complétons actuellement. Ce programme est appliqué sur notre corpus textuel de manière thématique. Les mots relevés par le programme permettent de construire une base de données détaillée.

Cette base de données a une structuration:

Fig. 9. Tables d'une base de données du kabyle



qui permettra aussi d'élaborer la structure qui enregistrera le dictionnaire de Miloud Taïfi pour la construction d'une consultation en ligne. On distingue dans la figure 9 ci-dessus une entrée (en tant que mot), suivie de catégories lexicales ayant chacune des significations pouvant être illustrées d'exemples. A ces diverses acceptions d'un mot peuvent être liés des antonymes et synonymes ainsi que des informations morphologiques détaillées pour les verbes, les substantifs et les adjectifs.

Je donne ci-dessous des illustrations des résultats renvoyés à l'utilisateur. En premier lieu, la construction d'informations concernant un nominal « asuf » pouvant être substantif ou adjectif suivant la signification (les informations morphologiques flexionnelles peuvent être par un simple clic présentées ou cachées suivant les besoins):

Fig. 10. Sortie concernant le nominal « asuf »

asuf (ⵓⵙⵓⴼ)

√: SF, dérivation nominale

1. adjectif

☞ unique

▼ Flexion détaillée

Masculin	Etat libre	Etat annexé	Féminin	Etat libre	Etat annexé
Singulier	asuf	wasuf	Singulier	tasuft	tasuft
Pluriel	isufen	yisufen	Pluriel	tisufin	tsufin

2. substantif, masc.

☞ Ling. singulier (antonyme de pluriel)

- Tinfelyiwin (asuf : tanfelya) = métamorphoses : Tinfelyiwin (singulier : tanfelya) = métamorphoses.

▼ Flexion détaillée

Masculin	Etat libre	Etat annexé
Singulier	asuf	wasuf
Pluriel	isufen	yisufen

Voici une sortie concernant un verbe. Je donne d'abord la forme avec la morphologie flexionnelle cachée (que je donnerai séparément pour des raisons de taille) :

Fig. 11. Partie principale d'une sortie verbale

bdu (ⵔⵓⵎ)

√: BD, dérivation verbo-nominale

verbe

☞ commencer

- Ad bduy ibi : Alors je commence.

- Adfel ittaggad agu, agu ittaggad adu, adu ittaggad ageffur, seg wasmi tebda dduint, (trad litt: depuis le jour quand a commencé le monde / depuis que le monde a commencé (existe)) : La neige redoute le brouillard, le brouillard redoute le vent, le vent redoute la pluie, c'est comme cela depuis que le monde est monde.

► Flexion détaillée

Sous l'entrée est indiquée la racine et le type de dérivation (verbo-nominale ou nominale), puis la catégorie lexicale, la signification et enfin les

exemples. La morphologie flexionnelle peut être ensuite activée si on le souhaite.

La partie flexionnelle est elle-même divisée en 3 parties, la présentation des formes de base, puis la conjugaison des temps et enfin l'impératif et les participes.

Les formes de base

Fig. 12. Les formes verbales principales

▼ Flexion détaillée

Formes verbales (FV) principales

impératif	bdu
-----------	-----

FV1		FV2	inaccompli	FV3	accompli	FV4	accompli nég.
aoriste	bdu	aoriste intens.	bddu	prétérit	bdi / bda	prétérit nég.	bdi

Ce sont les formes à partir desquelles sont calculées toutes les autres formes verbales. On possède là l'appareil formel qui servira à la construction du conjugeur en ligne du berbère du Maroc central.

Les conjugaisons des aoristes et prétérits

Fig. 13. La conjugaison des aoristes et des prétérits

Conjugaison

aoriste	singulier	pluriel	aoriste intens.	singulier	pluriel
1. pers. m.&f.	bdu-y	ne-bdu	1. pers. m.&f.	beddu-y	n-beddu
2. pers. masc.	te-bdu-d	te-bdu-m	2. pers. masc.	t-beddu-d	t-beddu-m
2. pers. fém.		te-bdu-m-t	2. pers. fém.		t-beddu-m-t
3. pers. masc.	i-bdu / ye-bdu	bdu-n	3. pers. masc.	i-beddu	beddu-n
3. pers. fém.	te-bdu	bdu-n-t	3. pers. fém.	t-beddu	beddu-n-t

prétérit	singulier	pluriel	prétérit nég.	singulier	pluriel
1. pers. m.&f.	bdi-y	ne-bda	1. pers. m.&f.	ur bdi-y	ur ne-bdi
2. pers. masc.	te-bdi-d	te-bda-m	2. pers. masc.	ur te-bdi-d	ur te-bdi-m
2. pers. fém.		te-bda-m-t	2. pers. fém.		ur te-bdi-m-t
3. pers. masc.	i-bda / ye-bda	bda-n	3. pers. masc.	ur i-bdi / ur ye-bdi	ur bdi-n
3. pers. fém.	te-bda	bda-n-t	3. pers. fém.	ur te-bdi	ur bdi-n-t

L'impératif et les participes

Fig. 14. L'impératif et les participes

impératif	aoriste	aoriste intens.
2. pers. sing.	bdu	beddu
2. pers. pl. M	bdu-t	beddu-t
2. pers. pl. F	bdu-mt	beddu-mt

participes	positif	négatif
aoriste	i-bdu-n / ye-bdu-n	
aoriste intens.	i-beddu-n	ur n-beddu
prétérit	i-bda-n / ye-bda-n	ur ne-bdi

La conclusion

Sera brève: un très gros travail informatique a été réalisé pour la construction et la publication du dictionnaire raisonné berbère - français, mais qui n'était pas visible.

J'ai décidé de mettre en ligne, à la disposition de tous, un certain nombre d'applications. Ce travail est à peine débuté. Il n'est réellement visible que par la présence en ligne de l'exemplier Taïfi. D'autres applications viendront dans les mois qui viennent. Ce seront essentiellement la consultation du dictionnaire raisonné par l'intermédiaire du français et la création d'un environnement d'apprentissage du kabyle.

J'avoue que ces travaux participent à mon apprentissage personnel du berbère. Mon espoir le plus vif est qu'ils puissent, à terme, servir à beaucoup d'autres.

Bibliographie

- Azdoud, Driss, 1997, *Lexique commun des Ait Hadiddou du Haut Atlas*. Thèse de doctorat d'état. Faculté des Lettres et des Sciences humaines. El Jadida.
- Bentolila, Fernand, 1981, *Grammaire fonctionnelle d'un parler berbère. Aït Seghrouchen d'Oum Jeniba*, Paris, Sela.
- Bisson, Paul, 1940, *Leçons de berbère tamazight : dialecte des Aït Nahir (Aït Nâaman)*, Rabat, éditions Mouncho.
- Boumalk, Abdallah et Bounfour, Abdallah, 2001, *Vocabulaire usuel de tachelhit, tachelhit-français*, Rabat, éd. Centre Tarik Ibn Ziyad.
- Chaker, Salem, 1995, *Linguistique berbère. Etudes de syntaxe et de diachronie*, Louvain, Peeters.
- Naït-Zerrad, Kamal, 1995, *Tajeřrumt n tmaziyt tamirant (taqbaylit)*, Alger ENAG.
- Naït-Zerrad, Kamal, 1998, *Dictionnaire des racines berbères (formes attestées)*, Louvain, Peeters.
- Naït-Zerrad, Kamal, 2004, *Linguistique berbère et applications*, Paris L'Harmattan.
- Oussikoum: Dictionnaire tamazight-français (parler des Ayt Wirra). Thèse de doctorat d'état. Faculté des Lettres et des Sciences Humaines. Beni-Mellal, 1995.
- Pognan, Patrice, 2013a, « L'exemplier Taïfi, en hommage au maître et à l'ami », in *Etudes et recherches en linguistique et littérature berbères*, eds Samira Moukrim et Hachem Jarmouni, Fès, 2015, pp. 11-24.
- Pognan, Patrice, 20117, « Un nouvel horizon pour le dictionnaire raisonné berbère-français - parlers du Maroc: l'élaboration d'une méthode d'apprentissage », *Auréoles berbères, Mélanges offerts à Michel Peyron*, Tetouan, éd. Abdelaziz Allati, Rüdiger Köppe Verlag, Berber Studies, vol. 47, pp. 137-148.
- Pognan, Patrice.: 2021, « De Chaker à Naït-Zerrad : la description grammaticale support des TAQ (traitements automatiques du kabyle) », in *Les études berbères à l'ère de l'institutionnalisation de tamazight*. Mélanges en l'honneur de Salem Chaker et Abdellah Bounfour. Merolla, Caubet, Naït-Zerrad et Cassuto eds. L'Harmattan, Paris, 2021, pp 237-248.
- Pognan, Patrice, Taïfi, Miloud, 2010, « Traitements informatiques du dictionnaire raisonné berbère – français, élaboration de « l'exemplier Taïfi », *Journées d'étude JEC « Linguistique(s) de corpus : de la constitution à l'exploitation des corpus*, Paris.
- Pognan, Patrice, Taïfi, Miloud, 2011, « Un dictionnaire en tant que corpus : traitements informatiques du dictionnaire raisonné berbère – français de Miloud Taïfi. Rabat », *4ème atelier international sur l'amazighe et les TIC*. Rabat : IRCAM.
- Pognan, Patrice, Sadi, Nabila, Fitas Rachida, Salhi ohand.-Akli, Achour, Ramdane, 2018, « Analyse morphologique automatique du kabyle. Traitement de l'affixation. Présentations linguistique et algorithmique », Bayreuth, *10ème Bayreuth-Frankfurt-Leidener Kolloquium zur Berberologie*.

- Rahho, Rachida, 2005, *Dictionnaire berbère-français. Parler des Beni-Iznassen*. Thèse de doctorat. Faculté des Lettres et des Sciences Humaines, Dhar El Mehraz, Fès.
- Serhoual, Mohamed, 2002, *Lexicologie, lexicographie et sémantique berbères : 1) Lexicologie amazighe 2) Dictionnaire tarifit - français*. Thèse de doctorat d'état. Faculté des Lettres et des Sciences Humaines, Tétouan.
- Taïfi, Miloud, 1992, *Dictionnaire tamazight - français. Parlers du Maroc central*, Paris, L'Harmattan-Awal.
- Taïfi, Miloud, 2006, « Si les berbérophones ne s'entendent pas oralement, qu'ils s'écrivent ! Pour une écriture grammaticale du berbère », In *Standardisation de l'amazighe*. Publications de l'IRCAM, Rabat, pp. 30-43.
- Taïfi, Miloud, 2017, *Dictionnaire raisonné berbère - français. Parlers du Maroc* (1223p), Pognan: index "Formes de mots berbères - racines" (166p) et index "Significations françaises - racines berbères" (319p), Rabat, IRCAM.