

Du texte aux données analysables
***Lwali n udrar* comme corpus**

Patrice POGNAN **Mohand Akli SALHI**
Sorbonne et INALCO **LEALA - UMMTO**

Agzul

Deg tezrawt-agi, nra ad d-nesken amek i yelha uđris n *Lwali n udrar* d aybalu i uslađ ama di teskla ama di tesnilest. Nettef ungal-agi d ammud, deg-s ad d-nekkes iferdisen, ama n umawal, ama d tefyar, ama d tikkisin, ad tent-id-nsewjed i usezrew.

Abstract

Starting from the text as it is edited, as a novel, we wish in this paper to take the text of *Lwali n udrar* as a linguistic and literary material to transform it into lexical, sentences and sequential data that can be analyzed in literature and in linguistics.

En partant du texte tel qu'il est édité, en tant que roman, par les Editions Tira en 2011, nous souhaitons dans ce papier prendre le texte de *Lwali n udrar* comme un matériau linguistique et littéraire pour le transformer en données lexicales, phrastiques et séquentielles. Et ce dans le but d'exposer quelques possibilités d'étude en proposant des données préparées à l'analyse. Il s'agit donc ici de présenter une étape d'une recherche globale dédiée aux traitements automatiques d'un corpus large et varié. Nous limiterons notre exposé à la présentation organisée, suivant leur nature, des données concrètes (linguistiques et littéraires) intéressantes à étudier dans le texte *Lwali n udrar*.

Les premiers traitements automatiques effectués sur ce texte sont très encourageants ; les données qu'ils fournissent nous permettent non seulement de caractériser le texte, et partant l'écriture de Belaid, mais également de décrire un état de langue et de littérature historiquement bien situé.

Le choix du texte n'est pas arbitraire. Il présente la singularité d'être à la frontière de l'oralité et de l'écriture à tous points de vue (langue, textualité, style, etc.). Aussi bien du point de vue linguistique que littéraire, *Lwali n udrar* est historiquement le premier *texte écrit*, qui présente une *longueur inédite* et une *poétique novatrice*, composé d'une manière élaborée présentant plusieurs écarts avec les textes déjà transcrits ou oraux.

Question de démarche

Après avoir récupéré le texte sous sa version word et procédé à sa correction suivant les règles de l'écriture actuelle dont nous donnons ci-après la première page :

Ssebba n tmacahut

Welleh, a Sidi, ar d tidet : sebæa u sebein n şşellaḥ i aḡ-d-yezzin dagi di læerc-agi-nney i ssney, i d-ttaddren, is i ttgallay di şşbeḥ alamma d tameddit, læemer yelli wi i yi-iḥkan tamacahutt n yiwen deg-sen, neḡ i yi-innan ula d awal yeḡ leḡyat-nsen s tidet ! Ma yehwa-yak, bnaðem, mi yeggull: aḡeq Ccix flani ! iæedda limin-nni, neḡ yeḡnet : umbeed ad d-yelhi d wayen nniḡen: ur yettnadi ara yeḡ leflani-nni d acu-t, neḡ amek iæac, neḡ amek yemmut : d Ccix kan u hadak ma kan. Awi-d ayen i d-iteddun, ma d ayen iæeddan, dḡa akken qqaren waeraben : lli fat mat!

Lḡaşun, ziy, ar yekkeb d Sidi Rebbi ar d-sley i tmacahut takka n yiwen Ccix, yerna d lwali ameqqran deg yidurar-agi-nney yakk, ar d-afey d win ara yi-d-yeḡkun fell-as, ur bniḡ yeḡ yiman-iw.

Lexmis n udrar-agi, (dḡa sdeffir udrar), ulamma d tamdint, temmed, selley kan yes-s, læemer tt-kcimey, læemer tt-ssiney yakk di ddunit-iw. Aḡas n wat tmurt-agi-nney i tt-yettsewwiqen dḡa d lexmis, ttaddren-tt-id kan akka ma yewwi-tt-id wawal.

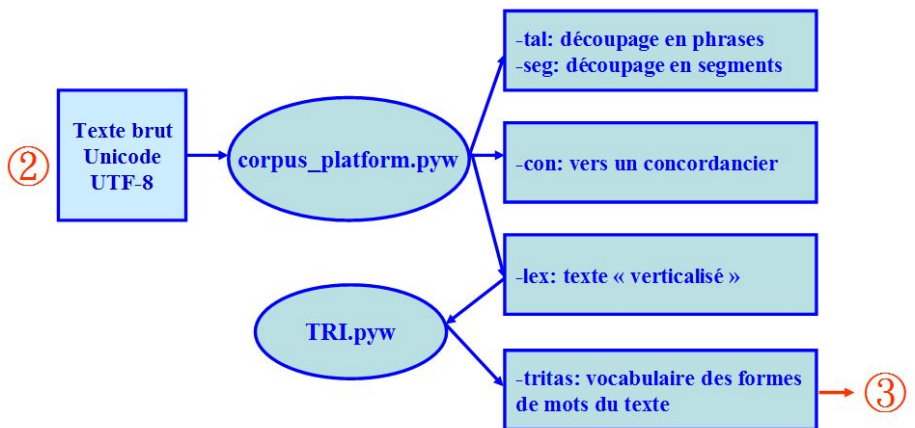
Nekk, læemer seiḡ ssebba is ara tt-issiney: tabeæ maççi ala nekk akka i d-inudan mkul tamurt, yeẓger lebḡur, yunag yeḡ mkull berr, di tmurt-is ur yessin ara di læerc-is akkin. Ula d læerc-nni... Wissen a Rebbi !

Ussan-agi iæddan, nniy-as : welleh ard awðeý ar din ziy ad zrey.

Kkrey-d deg wass n larebea şşbeý zik, ttfey abrid. Nniy-as ad nsey din axir: akken, azekka-nni ad d-kkrey meqqar imi ara yebdu ssuq. Gas yef uðar, lakin ulamma lhiy s leeql, ziy anda ulac yakk, tameddit ad ssiwðeý.

le premier élément de la chaîne de traitements est toujours le texte transformé en fichier brut (format .txt) en Unicode UTF-8 qui est devenu le standard pour le multilinguisme et son traitement informatique. Le corpus est constitué de ces 3 premiers fichiers (formats du fichier originel .pdf, .doc et .txt) et de chaque étape de traitement dûment référencée du fichier .txt.

Le premier programme (00-corpus_platform.pyw) de la chaîne de traitements est très important, puisqu'il réalise un traitement de la typographie et notamment de la ponctuation, mais aussi et surtout il produit simultanément 5 types différents de sorties suivant le schéma ci-dessous :



Les fichiers correspondants aux différentes sorties sont marqués dans leur nom par une « option » qui indique les traitements et les applications qui en résulteront : les options –tal et –seg sont orientées toutes deux vers l’analyse automatique (–tal est découpé par phrases et –seg par segments plus petits compris entre deux ponctuations,

notamment la virgule. Une 3^{ème} option –tal-L est identique à –tal à ceci près que chaque phrase est séparée par un blanc, afin de produire une impression papier qui permettent de préparer d'éventuelles corrections au sein du texte), l'option –con servira à réaliser des concordanciers et l'option –lex est l'entrée des traitements lexicographiques. Il s'agit donc d'une direction « lexique ».

A la suite de ce premier programme, les fichiers de l'option -lex, texte « verticalisé » subiront une série de traitements : tri et tassement du tri qui donnent la liste des occurrences de mots du texte en ordre alphabétique, élimination des éléments non lexicaux (signes de ponctuation, nombres), possibilité de garder au choix des unités mixtes alphanumériques (portion alphabétique munie d'un nombre, d'une unité de mesure, d'une notation, ...). Les unités lexicales qui apparaissent avec une majuscule d'une part et entièrement en minuscules d'autre part sont ramenées à une forme unique en minuscules.

L'option –lex est alors prête à subir des traitements en vue de la recherche du vocabulaire et les traitements d'analyse morphologique qui y sont liés et serviront, par ailleurs, pour l'option –tal, analyse automatique.

Voici, à titre d'illustration, une sortie de type « tal » (découpage par ponctuation majeure) pour le texte *lwali n udrar* :

Ssebba n tmacahut .

Welleh , a Sidi , ar d tidet :

sebba u sebain n şşellaḥ i ay-d-yezzin dagi di læerc-agi-nney i ssney , i d-ttadden , is i ttgallaḡ di şşbeḥ alamma d tameddit , læmer yelli wi i yi-iḥkan tamacahut n yiwēn deg-sen , neḡ i yi-innan ula d awal yeḡ leḡyat-nsen s tidet !

Ma yehwa-ak , bnadem , mi yeggull :

aḥeq Ccix flani !

iædda limin-nni , neḡ yeḡnet :

umbeed ad d-yelhi d wayen nniḡen :

ur yettnadi ara yeḡ leflani-nni d acu-t , neḡ amek icac , neḡ amek yemmut :

d Ccix kan u hadak ma kan .

Awi-d ayen i d-iteddun , ma d ayen iæddan , dya akken qqaren waeraben :

Ili fat mat !

Lḡaṣun , ziy , ar yekṡeb d Sidi Rebbi ar d-sley i tmacahut akka n yiwēn Ccix , yerna d lwali ameqqran deg yidurar-agi-nney yakk , ar d-afey d win ara yi-d-yeḡkun fell-as , ur bniḡ yeḡ yiman-iw .

Quelques données

La préparation des données analysables peut donc couvrir plusieurs niveaux d'analyse, allant de l'inventaire des formes de mots à l'extraction des segments à valeur poétique et/ou stylistique, tout en passant par les segmentations en phrases en relation avec le système de ponctuation adopté dans le texte. Les données jugées importantes et qui méritent des regards attentifs sont situées dans le cadre de cet exposé au plan linguistique et littéraire.

Des données intéressant l'interrogation linguistique, nous présentons, dans un premier temps, celles en relation avec les problématiques morpho-phonologiques et prosodiques et, dans un second moment, celles liées aux relations entre les unités linguistiques où nous prendrons le cas particulier de l'affixation.

Sur l'ensemble de la masse lexicale du texte *Lwali n udrar*, une partie importante contient un segment fait du redoublement de la même consonne signifiant ainsi qu'il est phonétiquement et

phonologiquement différent du segment non redoublé. On en décompte près de 8000 occurrences dont cette image donne la dernière page de l'inventaire :

7932	tettēs
7933	tt-nessaki
7934	tt-nessaki
7935	uggadey
7936	ay-d-tewwet
7937	Daynetta
7938	akken
7939	tenniq
7940	sgellin
7941	yīsefra-nni
7942	yi-d-tewwiq
7943	nemxallaf
7944	wass
7945	ṛebbi
7946	yerra
7947	tayett-is
7948	azniq-nni
7949	yerran
7950	taddart
7951	yewwed
7952	yiseddaṛen
7953	akken
7954	yettali

Cette longue liste présente donc les formes de mots ayant ce segment redoublé. C'est un inventaire composé de manifestations concrètes de ce fait dans des contextes morphosyntaxiques divers. Le chercheur linguiste dispose ainsi de matériaux intéressants relatifs, entre autres, à l'étude de la tension consonantique ou de la gémination. Par ailleurs, en relation avec l'extraction des racines, il aura à observer, classer et analyser la nature de certaines radicales de la racine (cas d'assimilation, cas d'affaiblissement, etc.) pour déceler l'origine (ou les causes) de la tension ou de l'allongement de ce segment.

La même liste contient également des données relatives à l'étude du comportement de la voyelle neutre en relation avec le

segment redoublé tout en ayant à l'esprit la question du traitement syllabique des unités contenant ce type de segment.

L'autre type de données à présenter concerne le fait de l'affixation. L'image suivante présente une page de la liste des mots avec des préfixes :

as-anfey	V	Pref	aoriste	a-s	3.S.FM.D	
as-cetkiy	V	V	Pref	aoriste	a-s	3.S.FM.D
as-cfiy	V	Pref	aoriste	a-s	3.S.FM.D	
as-d-as	V	Pref	aoriste	a-s	3.S.FM.D	
as-d-igerru	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-ihedder	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-inesren	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-kemmley	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-kksen	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-neqey	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-qqaren	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-tawi	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-teyli	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-tezzin	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-ttawin	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-ufin	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-wwiy	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-yawi	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-yedhu	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-yedsı	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-yefhem	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-yefken	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-yehdi	V	V	Pref	aoriste	a-s	3.S.FM.D
as-d-yekk	V	V	Pref	aoriste	a-s	3.S.FM.D

Ce type d'approche est intéressant notamment dans la perspective d'élaboration automatique de la grammaire et dans la réalisation de logiciels interactifs. Outre la fréquence et l'ordre, l'opportunité sera également donnée à l'analyse de constater et d'étudier la nature des compatibilités (combinaison avec des régimes directs et/ou indirects dans le cas des verbes par exemple) des affixes avec les unités qui les portent. Par ailleurs, cette liste présente aussi des matériaux qui intéressent la conjugaison.

Au niveau de l'analyse littéraire, nous nous contenterons d'évoquer, entre autres données, les marques situant les interpellations

que le narrateur (ou les narrateurs) adresse à son récepteur (narrataire). Interpellations qu'on peut interpréter comme des indications permettant le contact avec le lecteur du texte. Par ailleurs, ces mêmes interpellations participent grandement à la structuration de la narration en permettant des digressions, des commentaires, des positionnements du narrateur par rapport à l'histoire qu'il narre, etc.

L'une des ce nombreuses interpellations, l'expression « *ma yehwa-ak* » est des plus importante d'abord par le nombre de fois qu'elle a été utilisée et puis par la définition du contrat de réception qu'elle propose aussi bien pour le narrataire que pour le lecteur. En voici un pointage détaillé de cette expression située dans ces contextes d'utilisation :

Nekk, ussan-nni, ad tettafed yakk leğyub-iw ččuren d lekwayed ketben , ama dya d isefra , ama d timucuha , ama ula n wayen i yi-d-yeqqar uqerru-w. Cwiṭ, **mayehwa-ak**, s rreyba-nni seiṭ yer teqbaylit : akken kan bdiy heffdey ad tt-ketbey qqarey-tt. Day netta , anda ufiy ney sliy i usefru, leqqdey d llqed, am win ara yettarranyer tkufit. zriy deg wannect-nni ulac lxezna am tlawin .

[...]

— Ma tessned lehyatula n yiwen di lawleyya is i tettgallaḍ ? zriy : **ma yehwa-ak**, ula d nekk, cwiṭ dya seg wakken i k-walay ... ur tettxalafed ara wiyid, i byiy ad ak-ħkuṭ yef Ccix Ḥmed u Eli-agi-nney. Yerna, atas atas, tidet i Rebbi, imi la teqqareḍ tkettbeḍ s teqbaylit , tketbeḍ isefrat tawin fell-as ... ih , isefra-s ... Ladya, mi ara as-qqaren : tnefeed-ay s lbaraka-k, teefud-ay ma d ayen nexdem ... Ih ! lḥaşun ... tezriḍ ? ...

[...]

— Mi ara mlilen sin ur nemyussan ara d lmaḍi, qqimen i lhedra, zran ad mfaragen deg wass-nni, ad iniy yettili gar-asen nnšib n tidet. Nekk, lḥaşun ad ak-iniy : kra n win ara iheddren, - yerna

ad d-yelfu ula d win ara ikettben-, almi di ddhen-is ad tettwayellet tmeslayt-nni-ines. Win ara ak-yinin iheddeṛ kan, ney iketteb kan i yiman-is, ihi tili dya ulayyer yakk yessembiwel imi-s ney afus-is. Daynetta, ad ak-skiddbey ma nniy-ak ... yur-k anda tetteawadeḡ ayagi ara k-iniy yef Ccix, yur-k ad t-teḥkuḡ , yur-k dihin, yur-k dagi. Ala! d lekdeb, ma nniy-ak-t. Lxattima, ihi ead, imi... tettwaktab teqbaylit, i d-yenna d dhen-iw, ahat... ahat... Yerna lḥaşun, niy nekk ula d isem-iw ulay yer t-tissined . Awi-d kan ma yella kra... uyellet ! Sakin, kečč a mmi-s n tmurt, deḡber aqerru-k ! Teggra-d deg wayen ara ak-d-ḥkuḡ, ma d ayen yella nney... zeggdey , ney sneyseḡ... tfehmed ur ay-yecrik wara n ssuq : **ma yehwa-ak** yakk, ḡeel-it d albeed n tmucuha-nni n walef lila u lila, di syin truḡed yer tmurt-nwen, nekk s axxam-iw... Niy akka ?

[...]

D ayen i si, nekk s yiman-iw, akken ara ak-iban, amek d-wwiy s lexbar, ilaq-iyi... d tamacahut weqbet tamacahut... Deg yimrabden-agi nniḡen iwessawen di taddart, at Ssi Belqasem, yella, leewam-agi ieeddan, yiwen deg-sen, Ssi Ccrif, nemxalaḡ nekk yid-s, **ma yehwa-ak** nemḡabab am watmaten. Ladya, neszedda nekk yid-s kra n teswiḡet akka di Fṛansa, nemyexdam lxir, yerna, lḥaşun, yufa-yi di kra n temsalt akken n ddiq, almi, taneggarut, asmi d-nuḡal nemlal di tmurt ad iniy iḡeel-iyi lxir n leemum-agi-ines. Tabee d leemum-is-agi yakk i d wi yebḡa i d ssebba-s almi ass-agi ideḡ nella mazal-it kan netta di Fṛansa, ur zriy yemmut ney yedder...

[...]

Uḡalen ead ihi, **ma yehwa-ak**, byan ad t-eezlen di taddart : nnan-as : imi zik-nni d Bu Leyṡuṡ kan neemmed-as i wudem n Reḡbi, ma d tura, atan yuḡal d ameeguz : yettagi ad yefk leetab i yiman-is... eni d tuckerḡa ayen la yettaker ?...

[...]

— Belqasem ! Ur yuklal ara " Sidi " : yessexsar kan deg wawal n Rebbi. Dya, **mayehwa-ak**, lketra, fell-as i d-usiy ad k-weşşiy : yur-k anda d-yettruḥu yer-k, ma ur d-iruh ara yagi ? Ad ak-yesqizzeb, ad ak-d-yawi yak timucuha-nni-ines i si yettkellix medden! Tabee, akken i yessuffey i ubrid n Rebbi at wadda almi ifuk deg-sen ula d leeqel, byan ad t-rren d ccix. Walakin, ihun Rebbi, mazal Imumnin : at ufella-yagi ur qbilen ara . **Ma yehwa-ak** ihi, niy tura wagar-asen kan, ala lḥağ Emer kan i ixussen deg-sen cwiṭ tamusni ! Lakin, ulac uyilif : bna dem ad inadi kan lehna d leşlah n tmurt. Win yessnen Rebbi s tidet ur t-iyelleb ara ccitan. Yerna, muqel kan : nekkni netṭebbir, Rebbi yetṭebbir. Imcumen eezmen nniqal ad ay-ceeelen times i taddart, tusid-d kečč tessexsid-tt ! Ihi tura a Ccix, ad yecbeḥ lḡamee yis-k, ad yeemer ... Ahat ass-agi dya ara tebdud ?...

[...]

Ladya, **ma yehwa-ak**, tidet i Rebbi, tamezwarut yakk di temsal is ara t-yefhem bna dem d lwali, mačči cwiṭ kan : dya , d taduklinni yesdukkel taddart ass-nni kan yef i tt-id-yekcem . Ala tin kan i tettun kra n wid d-iḥekkun fell-as .

[...]

— Ad qebley, **ma yehwa-ak**, asmi d-yessehder yelli-k elaxater, mi ara yesked bna dem, d ayen isehlen am wulac, imi thedder si zik, mačči d tagugamt. Wamma taqsit n txatemt, yef wakka d-teḥkid... almi... s acu yellan din... almi... Lḥaşun, tufid-iyi-d aql-i wehmey !...

[...]

Deg wass-nni almi d asmi iyab, ur yuyal ara yembawel di lḥara ney di tdekkant ney di tqubbet akkin taddart n Tizi n Tfilkut s yiman-is, yerna ikcem-itt kan yiwen n wass si leemer-is s lekmal: asmi yemmut lḥağğ Emer at Rabeḥ, iruh d netta i ibedden i tentelt-is. Di syin akkin, almi d asmi i t-rren ula d netta

deg użekka-nni i yas-heggan di tqubbet asmi i tt-bnan, ur yeṭaxer leemer syagi. Tabee yuṭal yekkes ṭṭmee n tuffya akka ney akka dya si zik, segmi kan i yas-tebna Lla Faṭima taqubbet-nni. Tfehmed a di syin d asawen ur t-id-isaḥ ad as-yexḍu .Yerna, **ma yehwa-ak** , mkulci d tannumi a bna dem .

[...]

Mi t-twalid , ad terwuḍ. Atan yer-k, netta meqqar ad t-twalid, ad ak-id-iwali ; ad as-thedreḍ, ad ak-d-yehder. Ula d aqerru-s d yifassen-is ad ten-tettnaleḍ, ad ten-tessuduneḍ **ma yehwa-ak** meyya tikwal deg wass ; yerna ladya ladya, netta d argaz, d argaz kan am kečč d nekk, d læbd : daynetta, imi itett itess , yettlus ibrenyas d txutam n ddheb, yessibnin i seksu d wudi , itess asebsi, daynetta eni teuddeḍ ur ak-id-yettwali ara mi ara tæeddiḍ akken yer tsenduqt ad tegreḍ yer-s lweeda-k, ney ead, ma iḥemmel-ik Rebbi, yehda-t-id, yettef-itt netta s yiman-is s afus-is ?

[...]

Yerna, lḥaşun, **ma yehwa-ak**,- niy ad nini kan tidet- mi ara d-yemmekti bna dem argigi n tgecrar d temsal-nni yakk n " ṭaf ! ṭaf! ", a gma, nekk, lḥaşun, yeḥ leḥsab n tmusni-w, ur walay ara Bu Leyṭuṭ meskin ... ad yezmer ... Tabee, yerna nenna-t-id yagi, " icudd-it " uqelmun n ddheb yer Tizi n Tfilkut yiwet n tikkelt ; tikkelt tis snat, taneggarut, tcudd-it yer tqubbett-is n tmurt s lekmal akk, ladya, d Lla Faṭima. Almi d ass aneggaru yella netta " yexḍa ccyeḥ". Tfehmed, a mkullec, almi ula, yerna diyen ladya, almi ula d ayen yellan deg yiman-is netta weḥd-s, n terwiḥt-is, ar t-yexḍa d lmaḍi almi d asmi yemmut, imi d nettat i d-yeggran s deffir-s.

[...]

Lemmer d lebyi, imi nseeu Tidadacin, ad asent-nesu ula d imeslayen ilaqen. Lakin lleh yaleb... yerna, **ma yehwa-ak**, ad

neḥmed Ṛebbi, ad t-ncekkeṛ skud teqqim-ay-d Ibaṛaka n yizmwawen am bab n umkan-a!... Dya , tabee nekk la d-yettadren lekwafer d tmeslayt-nsen, eni bdiy la sxerwidex ... Yerna, a sidi, aql-i s leeqel-iw : aḥeqq amkan-agi ar kteṛ n xemsetṭac n ssna aya mačči swiy tiqit n ccrab!...

L'étude de cette expression et ses implications sur la narration, sur la relation narrateur/narrataire, sur le type de propos qu'elle introduit, sur sa valeur stylistique, etc. permettra une meilleure lecture du texte *Lwali n udrar*.

En conclusion, nous souhaitons insister sur la qualité (entre autres linguistique et littéraire) de ce texte. La méthodologie propre aux traitements automatiques des langues apporte à l'analyse, qu'il soit linguistique, critique littéraire ou autres, des données organisées et fiables permettant des approches précises.

Références bibliographiques

- Ameziane Amar, (Dir.), 2013 : *Les Cahiers de Belaid At-Ali. Regards sur une œuvre pionnière*, Bejaia, Tira Editions.
- At Eli Beleid, 2011, *Lwali n udrar*, Tira Editions, Bejaia.
- Pognan Patrice, 2012, « Extraction automatique du syntagme nominal prépositionnel berbère ». In *Etudes berbères VI – Essais sur la syntaxe et autres articles, 6. Bayreuth-Frankfurt-Leidener Kolloquium zur Berberologie – Bayreuth 2010*, eds. D. Ibrizimov, R. Volken, H. Stroomer. RüdigerKöppeVerlag.
- Pognan Patrice, Salhi Mohand Akli et Taïfi Miloud, 2012, « Du corpus aux applications, une chaîne de préparation automatique des textes berbères ». Frankfurt am Main: *7ème Bayreuth-Frankfurt-Leidener Kolloquium zur Berberologie*.
- Pognan Patrice et Salhi Mohand Akli, à paraître, « L'univers des contes kabyles (indexé) par ses personnages », *Interstices* n° 14 (Casablanca, Maroc).

Pognan Patrice et Salhi Mohand Akli 2013, « Pour une indexation raisonnée des corpus littéraires kabyles », *La linguistique de corpus : recueil, annotation, exploitation et diffusion, Iles d Imesli* n° 05. <http://revue.ummto.dz/index.php/idi/article/view/295/215>