

L'amazighe dans les sciences du numérique : expérience de l'IRCAM

Fatima Agnaou¹, Khalid Ansar², Fadoua Ataa Allah³

Aicha Bouhjar⁴, Siham Boulaknadel³

(1) CRDPP, IRCAM, Rabat, Maroc

(2) CAL, IRCAM, Rabat, Maroc

(3) CEISIC, IRCAM, Rabat, Maroc

(4) CRDPP, IRCAM, Rabat, Maroc

{agnaou, ansar, ataaallah, bouhjar, boulaknadel}@ircam.ma

RESUME

Selon les études menées par l'UNESCO, les langues du monde courent, de plus en plus, le risque d'extinction. Ce qui amenuise la diversité culturelle qui fait la richesse de l'humanité. Ainsi, elle a entrepris, dans ses programmes de sauvegarde des langues en danger, des actions de préservation et de valorisation de ces langues en mettant en exergue les initiatives qui puissent les doter des technologies de support.

Avec la même vision, l'Institut Royal de la Culture Amazighe a mis en place une stratégie pour la promotion et le développement de l'amazighe. Cette stratégie s'est traduite à travers une démarche progressive en matière d'aménagement linguistique, d'implantation dans le système éducatif et de traitement de l'amazighe par le biais des sciences du numérique. Néanmoins, la voie de la concrétisation de cette stratégie est pleine de défis et d'entraves : elle nécessite, en effet, la coordination des acteurs concernés par le développement et la transmission de l'amazighe dans les domaines de la politique linguistique, la recherche scientifique et la communauté du logiciel libre.

ABSTRACT

Amazigh in the digital sciences: experience of the IRCAM

On the basis of the studies undertaken by UNESCO, it has been observed that world languages are, increasingly, exposed to the risk of extinction. This risk has influentially contributed in reducing cultural diversity which may well be viewed as one of the most prominent aspects of richness for humanity. To contend with this situation and save endangered languages, UNESCO has entertained a whole range of measures with an eye to preserving and valuing less endowed languages by carrying out a number of initiatives meant to endow these languages with the necessary technologies.

With the same line of thinking as background, the Royal Institute of Amazigh Culture has piloted and set up a strategy for the promotion and development of Amazigh. This strategy has been translated into action through steps involving the progressive planning of Amazigh, its implantation in the educational system and its introduction in the digital world. Nonetheless, an evaluation of the experience carried out by IRCAM, thus far, evinces that this strategy is beset with a whole range of obstacles and constraints, and that better results will ensue if a coordination holds between the actors interested in the development and transmission of Amazigh in the domains of linguistic policy, scientific research and language digitalization.

MOTS-CLES : amazighe, science du numérique, promotion, langue peu dotée.

KEYWORDS: Amazigh language, digital science, promotion, less resourced language.

1 Introduction

Dans un contexte global où la transformation digitale est l'un des défis majeurs dans tous les domaines (économique, social, culturel, ...), toutes les langues se doivent de pouvoir répondre aux attentes de leurs locuteurs pour pouvoir être utilisées dans les sciences du numérique. C'est dans cet environnement global que les langues peu dotées, dont la langue amazighe, sont amenées à se développer. Se pose dès lors la question de savoir comment faire de l'amazighe une langue *suffisamment dotée* pour répondre aux défis de son implantation dans la sphère publique et ce d'autant qu'elle est à présent officielle, au Maroc, aux côtés de l'arabe depuis le 1^{er} juillet 2011, soit une décennie après la création de l'Institut Royal de la Culture Amazighe (IRCAM).

En effet, le statut de la langue amazighe a changé depuis le Discours Royal d'Ajdir du 17 octobre 2001 et le Dahir (Décret Royal) qui porte création de l'IRCAM. Jusqu'alors absente de la scène publique, essentiellement orale, limitée à l'usage informel et présente sous sa forme dialectale, la langue amazighe fait désormais partie du « patrimoine culturel commun » et devient une « affaire nationale » puisqu'elle concerne tous les citoyens, soit plus de trente millions de personnes. L'identité culturelle du citoyen marocain est ainsi définie dans la pluralité de ses affluents et constitue un patrimoine à promouvoir, à valoriser et à préserver. Ce processus de patrimonialisation et plus récemment d'officialisation de la langue et de la culture amazighes induit des actions spécifiquement liées à l'aménagement du statut et du corpus de la langue. Conformément aux missions assignées à l'IRCAM, des travaux sur l'aménagement du corpus (graphie, orthographe, lexique et grammaire) ont été initiés afin de permettre à la langue de remplir les nouvelles fonctions qui lui sont à présent dévolues dans le domaine public ; prioritairement dans l'enseignement et les médias ainsi que dans les sciences du numérique.

Cette contribution expose trois aspects essentiels de la problématique liée à la stratégie de l'IRCAM en matière de développement de l'amazighe, à savoir : son aménagement linguistique, son implantation dans le système éducatif et son traitement par le biais des sciences du numérique.

2 Stratégie du développement de l'amazighe

D'une manière générale, la stratégie adoptée par l'IRCAM, en matière d'aménagement linguistique, inspirée de l'expérience corse (Marcellesi, 2003), consiste en une gestion matérielle de la variation géolectale, ce qui engendre un amazighe enrichi de l'apport des différents dialectes régionaux du nord, du centre et du sud du Maroc. Concrètement, il s'agit de neutraliser ou de contenir la variation dialectale lorsque c'est possible et préférable pour une communication élargie (sur le plan phonique essentiellement). Par contre, la variation est intégrée lorsqu'elle est source d'enrichissement de la langue : la variation lexicale ou morphosyntaxique, actualisée dans la langue par des termes ou structures concurrent(e)s, est notamment prise en charge par le biais de la synonymie. Les lacunes terminologiques font quant à elles l'objet d'une création néologique commune à l'ensemble des variantes nationales ; la néologie reste cependant l'ultime recours après que toutes les ressources de la langue aient été épuisées. Autrement dit, la stratégie polynomique permet de préserver aux différentes variétés leur vitalité et, par là même, d'obtenir l'adhésion des locuteurs et de garantir la

cohésion sociale. Mais cette approche n'exclut pas une visée standardisante sur le long terme à dessein d'atténuer les variations de surface et d'outiller la langue en enrichissant son lexique. Ainsi, au fonds commun sous-jacent aux différentes variétés, vient s'ajouter un pan terminologique que les différents géolectes auront en partage. Une norme standard nationale s'impose donc pour des raisons d'intercompréhension évidentes, mais également par souci didactico-pédagogique comme c'est le cas pour toutes les langues (Ameur et Boumalk, 2004 ; Boukous, 2007 ; Agnaou, 2009).

Par ailleurs, dans une perspective d'intégration de la langue amazighe dans les sciences du numérique, l'IRCAM a développé une feuille de route (Ataa Allah et Boulaknadel, 2012 ; Ataa Allah et Boulaknadel, 2014B), structurée selon le court, le moyen et le long terme, pour le développement de ressources et outils assurant son fonctionnement à l'instar des langues conformément équipées. Elle est représentée par une chaîne partant des traitements élémentaires, passant par la constitution de briques de ressources linguistiques, et allant vers des applications génériques. Cette chaîne va de l'adaptation et l'amélioration d'outils en fonction des nouvelles technologies jusqu'au développement d'applications, en respectant les stratégies de développement inspirées des travaux de Muhirwe (Muhirwe, 2007), à savoir la standardisation, la notion d'extensibilité et de documentation ainsi que l'adoption des technologies libres.

Plus concrètement, la première action, à court terme, à laquelle l'IRCAM s'est attelé afin de permettre le passage à l'écrit de la langue amazighe, a trait au choix au système d'écriture (le tifinaghe) et des graphèmes qui devaient constituer l'alphabet.

2.1 Graphie et encodage

Fixer une norme graphique de l'amazighe passe nécessairement par le choix d'un alphabet qui doit répondre à un double objectif :

- le maintien d'un lien solidaire avec les différentes variantes de l'alphabet tifinaghe actuel, d'où la nécessité de puiser dans le fonds des graphèmes disponibles dans les différentes variantes et de considérer la création de nouveaux symboles comme un dernier recours ;
- l'adaptation du nouvel alphabet aux structures de l'amazighe standard, requérant parfois l'introduction de quelques modifications.

Pour répondre à cet objectif, il est tenu compte de quatre principes : l'historicité, la simplicité, l'univocité du signe et l'économie. Dans une seconde étape, d'autres paramètres ont été pris en compte dans le choix des caractères. Il s'agit de la fréquence des graphèmes dans les différentes variantes du libyque - tifinaghe, de leur simplicité au niveau de l'écriture manuelle (facilité psychomotrice), de l'esthétique des symboles et de la cohérence d'ensemble du système d'écriture proposé (Bouhjar, 2004). Après une étude approfondie du système phonético-phonologique des différentes variantes de l'amazighe présentes au Maroc, la sélection de 33 graphèmes a finalement été opérée. L'adoption officielle de la graphie tifinaghe en tant que système d'écriture pour la langue amazighe au Maroc a été annoncée le 10 février 2003.

Parallèlement, le tifinaghe a franchi une étape technologique initiée par le codage, un processus qui a connu deux phases contrastées :

- Une première phase transitoire a consisté en une adaptation de la norme ISO 8859-1 pour coder les caractères tifinaghes en codage ANSI afin de répondre à l'urgence de l'introduction de l'amazighe dans le système éducatif marocain en 2003. Cependant, la portée de ce codage privé est limitée et la gestion des textes comportant plusieurs systèmes

démarche globale de l'IRCAM s'appuie sur un certain nombre de principes et critères hiérarchisés et suit les principales étapes de tout travail terminologique. Ainsi, trois principes fondamentaux ont dicté la démarche à adopter lors du travail terminologique : la pertinence, la motivation (transparence des propositions dans la mesure où seul un haut degré de lisibilité garantit un haut degré d'appropriation des termes) et le respect de la morphologie et de la syntaxe de l'amazighe. Ces principes et critères ont permis de produire les lexiques sectoriels suivants : le lexique grammatical (Ameur *et al.*, 2009 ; Ameur *et al.*, 2011), le lexique des médias (Ameur *et al.*, 2009 ; Ameur *et al.*, 2013) et le lexique administratif (Ameur *et al.*, 2015).

Parallèlement, l'IRCAM a entrepris un projet de construction de bases de données lexicales et terminologiques numériques. La base de données terminologiques contient des entrées terminologiques relevant de plusieurs thématiques (El Azrak et El Hamdaoui, 2011). Cette base de données a été exploitée dans la réalisation de l'application mobile « LEXAM » (Frain *et al.*, 2014). Par ailleurs, une base de données lexicales, qui regroupe le lexique usuel appartenant aux différentes variantes de l'amazighe, a été élaborée. Cette dernière fait l'objet d'une exploitation Web et libre d'accès afin de répondre aux besoins des acteurs travaillant dans les domaines de la linguistique, l'enseignement, la traduction et la communication.

En vue de développer la compétence lexicale chez les apprenants de l'amazighe, l'IRCAM a réalisé cinq ressources numériques, en l'occurrence $\text{ⵜⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ}$ [Tamawalt inu tawlafant] « Mon vocabulaire illustré » (Aagnaou et Karoum, 2009), $\text{ⵜⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ}$ [Tamawalt n imzzyann] « Dictionnaire pour enfants » (Ataa Allah, 2011), $\text{ⵜⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ}$ [Tinml n tmazight] « Ecole amazighe » (Zenkouar *et al.*, 2007), $\text{ⵜⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ}$ [Izwiln s tmazight]¹ « Chiffres en amazighe » et $\text{ⵜⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ}$ [Awal inu amzwaru]² « Mes premiers mots ». Ces ressources ont été conçues dans le but d'outiller l'apprenant de l'amazighe du vocabulaire de base en langue qui soit en harmonie avec le processus de l'habilitation de la langue et avec les thèmes fixés dans les programmes scolaires.

D'autres contenus numériques ont été conçus pour le développement de la compétence communicative. Il s'agit de deux ressources, à savoir Imudar iramyarn d inamyarn « Animaux sauvages et domestiques » $\text{ⵜⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ ⵜⵉⵔⵉⵎⴰⵎⴰⵏⵜ}$ [IgDaD d ibukha] « Oiseaux et insectes » (Zenkouar *et al.*, 2008). Elles visent, en plus du développement du vocabulaire de la faune, la maîtrise de la compréhension et la production orale chez les apprenants.

2.3 Ressources grammaticales

On peut penser que les faits morphosyntaxiques soient moins sujets à la variation dans la mesure où il s'agit du niveau le plus stable. Or, une observation approfondie de certains faits morphosyntaxiques permet de relever que ce n'est pas le cas. En effet, certains phénomènes tels que l'état construit, la morphologie verbale et les déterminants, entre autres, résistent au processus d'unification. Disposant de peu de recherches et d'ouvrages de référence élaborés dans une perspective comparative et normative, l'IRCAM, conscient de telles contraintes, a concentré ses efforts sur la publication d'ouvrages de référence à visée standardisante sur le plan morphosyntaxique à partir d'une approche comparative inter-dialectale (Boukhris *et al.*, 2008 ; Laabdelaoui *et al.*, 2012). A partir de ces premiers ouvrages, il s'agissait également de doter la

¹ <https://play.google.com/store/apps/details?id=ma.ircam.chiffres&hl=fr>

² <https://play.google.com/store/apps/details?id=com.ircam.vocabetlettres&hl=fr>

langue amazighe des premiers outils de base de grammaire en TAL. Le premier outil élaboré concerne l'étiqueteur morphosyntaxique qui a pour objectif la catégorisation grammaticale des unités lexicales d'un corpus de textes. Dans cette optique, il y a eu lieu d'élaborer un premier corpus numérique de référence à visée exhaustive (Boulaknadel et Ataa Allah, 2012 ; Boulaknadel et Ataa Allah, 2013). Ce corpus est constitué de textes bruts, représentatif des variantes de l'amazighe marocain et des différents genres (conte, poésie et articles de presse). Bien que le projet a progressé, il faut néanmoins noter qu'il a fallu surmonter un certain nombre de difficultés liées essentiellement à l'homogénéisation de la graphie et de l'orthographe du corpus. Par ailleurs, un jeu d'étiquettes, inspiré du modèle EAGLES (EAG, 1996), a été proposé (Ataa Allah *et al.*, 2014) afin d'assurer l'exploitation dans différentes applications du TAL, notamment dans le contexte multilingue. Le second outil concerne le conjugueur qui prend appui sur le *Manuel de conjugaison* (Laabdelouai *et al.*, 2012). Il est conçu selon une méthode à base de règles fondée sur une extension, par les arbres discriminatoires, de l'approche à deux niveaux (Koskenniemi, 1984). Il permet l'accès à distance à la conjugaison des verbes de la langue amazighe (Ataa Allah et Boulaknadel, 2014A).

D'autres outils sont présentement à l'étude mais, vu la pluralité des plateformes existantes et l'absence d'études en traitement automatique de l'amazighe, ils requièrent en amont des recherches approfondies menées en parallèle. Ces outils concernent l'analyse morphologique, les entités nommées et la traduction automatique.

Les travaux de recherche sur la morphologie de la langue amazighe ont été initiés par deux études (Nejme *et al.*, 2016 ; Ataa Allah, 2014), reposant respectivement sur les systèmes Nooj (Silberstein, 2007) et Xerox (Beesley et Karttunena, 2003). Ces études se basent sur le principe des transducteurs bidirectionnels utilisés à la fois pour l'analyse et la génération. Du point de vue de la génération, le plus haut niveau représente une description abstraite des règles morphotactiques qui définissent les conditions d'ordre et de combinaison entre les différents morphèmes. Ce niveau est associé à un niveau plus concret, dans lequel les règles morpho-phonologiques s'appliquent et lient chaque lexie à ses formes de surface. Les règles modélisées par des transducteurs séparés à chaque niveau sont combinées avec un lexique qui contient les formes radicales de l'ensemble des lexies. Concernant l'analyse, le transducteur traite les formes entrantes pour reconnaître les segmentations éventuelles en identifiant le radical et les différents affixes qui lui sont collés.

Parallèlement à ces travaux, des études d'analyses syntactico-sémantiques ont été menées. Elles consistent à réaliser un extracteur automatique des entités nommées amazighe (Talha *et al.*, 2015 ; Boulaknadel *et al.*, 2014) en exploitant le principe des transducteurs à états finis pour définir les contextes d'apparition des unités à extraire.

En outre, dans la perspective de réaliser un système de traduction automatique de la langue amazighe, deux études ont été entamées. La première porte sur l'alignement de corpus parallèles pour l'amazighe (Miftah *et al.*, 2017). Tandis que la deuxième (Taghbalout *et al.*, 2017) concerne l'intégration de la langue amazighe dans le projet UNL (Universal Networking Language) (Boguslavsky *et al.*, 2005). Ce projet tend à favoriser l'éclosion du multilinguisme dans la société d'information et de permettre à toute personne d'accéder à l'information sur Internet par le biais de sa langue native sans limitations des barrières linguistiques. Il permet de coder une représentation du sens d'un texte pour servir de pivot inter-langue dans les systèmes UNL de la traduction automatique.

3 Points forts et limites

Dans le cadre de sa stratégie relative à la promotion de la langue amazighe, l'IRCAM accorde une grande importance à l'aménagement de la langue et à son intégration dans les sciences du numérique. Ainsi, des efforts ont été fournis en vue de doter la langue amazighe de fondements essentiels et de ressources indispensables pour son développement. Désormais, l'amazighe est passée de la situation de langue vernaculaire orale avec des transcriptions différentes assez disparates à celle d'une langue dotée d'une écriture normée et encodée selon les standards du plan multilingue de base ISO-Unicode (norme 10646). Elle dispose à présent d'une orthographe stabilisée, d'une grammaire, de ressources et d'outils numériques.

En dépit de ces efforts pour développer une stratégie, qui favorise l'intégration de l'amazighe dans les sciences du numérique, celle-ci est confrontée à différentes contraintes.

Ainsi, la contrainte d'ordre linguistique est à mettre en relation avec le système d'écriture de l'amazighe qui pose les difficultés majeures suivantes :

- Variation dialectale : le processus de standardisation étant encore tout récent, les ressources disponibles nécessitent d'être redressées afin de répondre aux normes orthographiques préconisées par l'IRCAM. Cette étape reste incontournable pour pouvoir exploiter de manière optimale ces ressources dans les sciences du numérique.
- Pratiques scripturaires : l'usage du tifinaghe n'étant pas généralisé, le spécialiste du traitement automatique de l'amazighe se retrouve face à une multitude de formes d'écriture et de modes de transcription. Ce fait limite considérablement la portée de l'action vu l'ampleur des tâches préliminaires à effectuer avant toute exploitation numérique.
- Disponibilité de ressources : le nombre restreint de ressources enfreint fortement le développement numérique de l'amazighe pour des raisons évidentes de fiabilité des résultats. De même, sur le plan qualitatif, il faut relever que la nature des ressources disponibles ne couvre généralement que le domaine littéraire ; les médias émergent tout doucement. Ce constat met en péril la représentativité des corpus exploités.
- Morphologie amazighe : certaines catégories grammaticales n'étant pas encore stabilisées dans la mesure où elles nécessitent une étude plus poussée, la réalisation d'outils de traitement s'en trouve entravée.
- Sur le plan éducatif, les limites relevées concernent principalement les points suivants :
 - o Le niveau de compétences des enseignants exige une mise à niveau constante par la formation à l'utilisation technique des outils élaborés.
 - o La non généralisation de l'équipement, de salles de cours, en ordinateurs et Internet.

4 Conclusion

En définitive, il est à noter que la langue amazighe a relevé le défi de son introduction dans le monde du numérique puisqu'à présent une panoplie d'outils existent et sont mis à la disposition des utilisateurs. Cependant, beaucoup reste encore à réaliser et à entreprendre compte tenu des contraintes soulevées pour suffisamment doter la langue amazighe. Dans cette perspective, des travaux de recherche sont en cours et d'autres sont envisagés.

Références

ANDRIES P. (2008). UNICODE 5.0 EN PRATIQUE, CODAGE DES CARACTERES ET INTERNATIONALISATION DES LOGICIELS ET DES DOCUMENTS. DUNOD, FRANCE.

AGNAOU F., AFOULKI, M. (2006). ⵏⴰ ⵎⵓⵎⴰ ⵜⴰⵎⴰⵣⵉⵖⵜ (AD NLMD TIFINAGHE). PRODUCTIONS DE L'INSTITUT ROYAL DE LA CULTURE AMAZIGHE. RABAT.

AGNAOU F. (2009). « CURRICULA ET MANUELS SCOLAIRES : POUR QUEL AMENAGEMENT LINGUISTIQUE DE L'AMAZIGHE MAROCAIN ? », ASINAG N°3, RABAT, PUBLICATIONS DE L'IRCAM, PP. 109-126.

AGNAOU F., KARROUM, M. (2009). ⵜⴰⵎⴰⵣⵉⵖⵜ ⵏ ⵜⴰⵎⴰⵣⵉⵖⵜ (TAMAWALT INU TAWLAFANT). PRODUCTIONS DE L'IRCAM, RABAT.

AMEUR M., BOUMALK A. (2004). STANDARDISATION DE L'AMAZIGHE : ACTES DU SEMINAIRE ORGANISE PAR LE CENTRE DE L'AMENAGEMENT LINGUISTIQUE, RABAT, 8-9 DECEMBRE 2003. PUBLICATIONS DE L'IRCAM.

AMEUR M., BOUMALK A., IAZZI E.M, SOUIFI H., ANSAR K. (2009). VOCABULAIRE DES MEDIAS, FRANÇAIS – AMAZIGHE – ANGLAIS - ARABE. PUBLICATIONS DE L'IRCAM. IMPRIMERIE EL MAARIF AL JADIDA. RABAT.

AMEUR M., BOUHJAR A., BOUMALK A., EL AZRAK N., LAABDELAOUI R. (2009). VOCABULAIRE GRAMMATICAL. PUBLICATIONS DE L'IRCAM. IMPRIMERIE EL MAARIF AL JADIDA. RABAT.

AMEUR M., BOUHJAR A., BOUMALK A., EL AZRAK N., LAABDELAOUI R. (2011). VOCABULAIRE GRAMMATICAL DE L'AMAZIGHE : APPLICATION PHRASEOLOGIQUE. PUBLICATIONS DE L'IRCAM. IMPRIMERIE EL MAARIF AL JADIDA. RABAT.

AMEUR M., ANSAR K., BOUHJAR A., EL AZRAK N. (2013). TERMINOLOGIE AMAZIGHE DE L'AUDIOVISUEL. PUBLICATIONS DE L'IRCAM. IMPRIMERIE EL MAARIF AL JADIDA. RABAT.

AMEUR M., ANSAR K., BOUHJAR A., EL AZRAK N. (2015). TERMINOLOGIE ADMINISTRATIVE. PUBLICATIONS DE L'IRCAM. IMPRIMERIE EL MAARIF AL JADIDA. RABAT.

ATAA ALLAH F. (2011). CONCEPTION D'UN DICTIONNAIRE IMAGIER SONORE EN LIGNE DE LA LANGUE AMAZIGHE. ACTES DE THE 8TH MULTIDISCIPLINARY SYMPOSIUM ON DESIGN AND EVALUATION OF DIGITAL CONTENT FOR EDUCATION (SPDECE 2011). CIUDAD REAL, ESPAGNE, 15-17 JUIN 2011. 158-164.

ATAA ALLAH F., BOULAKNADEL S. (2012). TOWARD COMPUTATIONAL PROCESSING OF LESS RESOURCED LANGUAGES: PRIMARILY EXPERIMENTS FOR MOROCCAN AMAZIGH LANGUAGE. THEORY AND APPLICATIONS FOR ADVANCED TEXT MINING. RIJEKA: INTech. NOVEMBRE 2012.

ATAA ALLAH F. (2014). FINITE-STATE TRANSDUCER FOR AMAZIGH VERBAL MORPHOLOGY. LITERARY & LINGUISTIC COMPUTING. OXFORD UNIVERSITY PRESS, DOI:10.1093/LLC/FQU045.

ATAA ALLAH F., BOULAKNADEL S. (2014A). AMAZIGH VERB CONJUGATOR. ACTES DE THE 9TH INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION (LREC 2014). REYKJAVIK. ICELAND. 26-31 MAI 2014. 1051-1055.

ATAA ALLAH F., BOULAKNADEL S. (2014). LA PROMOTION DE L'AMAZIGHE A LA LUMIERE DES TECHNOLOGIES DE L'INFORMATION ET DE COMMUNICATION. ASINAG N°9, PUBLICATIONS DE L'IRCAM, PP. 33-48.

ATAA ALLAH F., BOULAKNADEL S., SOUIFI H. (2014B). « JEU D'ETIQUETTES MORPHOSYNTAXIQUES DE LA LANGUE AMAZIGHE », ASINAG N°9, PUBLICATIONS DE L'IRCAM, PP. 171-184.

BEESLEY K. R., KARTTUNENAUURI L. (2003). FINITE STATE MORPHOLOGY. CSLI PUBLICATIONS, STANFORD. CA.

BOGUSLAVSKY I., CARDEÑOSA J., GALLARDO G., IRAOLA L. (2005). THE UNL INITIATIVE. AN OVERVIEW. COMPUTATIONAL LINGUISTICS AND INTELLIGENT TEXT PROCESSING. 6TH INTERNATIONAL CONFERENCE, CICLING 2005. MEXICO CITY. MEXIQUE. FEVRIER 13-19.

BOUHJAR A. (2004). LE SYSTEME GRAPHIQUE TIFINAGHE-IRCAM. ACTES DU SEMINAIRE ORGANISE PAR LE CENTRE DE L'AMENAGEMENT LINGUISTIQUE, STANDARDISATION DE L'AMAZIGHE. PUBLICATIONS DE L'IRCAM. IMPRIMERIE EL MAARIF AL JADIDA. RABAT.

BOUKHRIS F., BOUMALK A., ELMOUJAHID E., SOUIFI H. (2008). LA NOUVELLE GRAMMAIRE DE L'AMAZIGHE. PUBLICATIONS DE L'IRCAM. IMPRIMERIE EL MAARIF AL JADIDA. RABAT.

BOUKOUS A. (2007). L'ENSEIGNEMENT DE L'AMAZIGHE (BERBERE) AU MAROC : ASPECTS SOCIOLINGUISTIQUES. REVUE DE L'UNIVERSITE DE MONCTON. NUMERO HORS SERIE. 81-89. DOI : 10.7202/017709AR

BOULAKNADEL S., ATAA ALLAH F. (2012). INITIATIVE POUR LE DEVELOPPEMENT D'UN CORPUS DE LA LANGUE AMAZIGHE », ACTES DE LA 5EME CONFERENCE INTERNATIONALE SUR LES TECHNOLOGIES D'INFORMATION ET DE COMMUNICATION POUR L'AMAZIGHE (TICAM 2012). RABAT. MAROC. 26-27 NOVEMBRE 2012.

BOULAKNADEL S., ATAA ALLAH F. (2013). BUILDING A STANDARD AMAZIGH CORPUS. ADVANCES IN INTELLIGENT SYSTEMS AND COMPUTING: PROCEEDING OF THE INTERNATIONAL CONFERENCE ON INTELLIGENT HUMAN COMPUTER INTERACTION (IHCI 2011). PRAGUE. CZECH REPUBLIC. AUGUST 29-31, 2011. 179: 91-98. SPRINGER BERLIN HEIDELBERG. ISBN: 978-3-642-31602-9.

BOULAKNADEL S., TALHA M., ABOUTAJDINE D. (2014). AMAZIGHE NAMED ENTITY RECOGNITION USING A RULE BASED APPROACH. ACTES DE LA 11TH ACS/IEEE INTERNATIONAL CONFERENCE ON COMPUTER SYSTEMS AND APPLICATIONS (AICCSA'2014). DOHA. QATAR. 10-12 NOVEMBRE. 478-484.

EAGLES. (1996). RECOMMENDATION FOR THE MORPHOSYNTACTIC ANNOTATION OF CORPORA. EAGLES DOCUMENT EAG-TCWG-MAC/R. [HTTP://WWW/ILC.CNR.IT/EAGLES96/HOME.HTML](http://www/ilc.cnr.it/EAGLES96/home.html).

EL AZRAK N., ELHAMDAOUI A. (2011). REFERENTIEL DE LA TERMINOLOGIE AMAZIGHE : OUTIL D'AIDE A L'AMENAGEMENT LINGUISTIQUE. ACTES DU 4^{EME} ATELIER INTERNATIONAL SUR L'AMAZIGHE ET LES

NOUVELLES TECHNOLOGIES DE L'INFORMATION ET DE COMMUNICATION, SOUS LE THEME : LES RESSOURCES LANGAGIERES: CONSTRUCTION ET EXPLOITATION (NTIC 2011). RABAT. MAROC. 24-25 FEVRIER 2011.

FRAIN J., ATAA ALLAH F., AIT OUGUENGAY Y. (2014). LEXIQUE AMAZIGHE POUR MOBILE, ACTES DE LA 6^{ÈME} CONFERENCE INTERNATIONALE SUR LES TECHNOLOGIES D'INFORMATION ET DE COMMUNICATION POUR L'AMAZIGHE (TICAM 2014). RABAT. MAROC. 24-25 NOVEMBRE 2014.

KOSKENNIEMI K. (1984). TWO-LEVEL MORPHOLOGY: A GENERAL COMPUTATIONAL MODEL FOR WORD-FORM RECOGNITION AND PRODUCTION, THÈSE DE DOCTORAT, UNIVERSITÉ DE HELSINKI, FINLANDE.

LAABDELAOUI R., BOUMALK A., IAZZI E.M, SOUIFI H., ANSAR K. (2012). MANUEL DE CONJUGAISON AMAZIGHE. PUBLICATIONS DE L'IRCAM. IMPRIMERIE EL MAARIF AL JADIDA. RABAT.

MARCELLES J. B. (2003). SOCIOLINGUISTIQUE. EPISTEMOLOGIE, LANGUES REGIONALES, POLYNOMIE. PARIS, L'HARMATTAN.

MIFTAH N., ATAA ALLAH F., TAGHBALOUT I. (2017). SENTENCE-ALIGNED PARALLEL CORPUS AMAZIGH-ENGLISH. ACTES DE L'INTERNATIONAL CONFERENCE ON INFORMATION AND COMMUNICATION SYSTEMS. IRBID. JORDANIE. 4-6 AVRIL 2017.

MUHIRWE, J. (2007), TOWARDS HUMAN LANGUAGE TECHNOLOGIES FOR UNDER RESOURCED LANGUAGES. COMPUTING AND ICT RESEARCH. (ÉD) JOSEPH KIZZA ET AL. KAMPALA.

NEJME F., BOULAKNADEL S., ABOUTAJDINE D. (2016). AMAMORPH: FINITE STATE MORPHOLOGICAL ANALYZER FOR AMAZIGHE. JOURNAL OF COMPUTING AND INFORMATION TECHNOLOGY. 91-110.

SMITH, F. (2004). UNDERSTANDING READING: A PSYCHOLINGUISTIC ANALYSIS OF READING AND LEARNING TO READ. MAHWAH. NEW JERSEY. LAWRENCE ERLBAUM ASSOCIATES PUBLISHERS.

SILBERZTEIN M. (2007). AN ALTERNATIVE APPROACH TO TAGGING. NLDB 2007: 1-11

TALHA M., BOULAKNADEL S., ABOUTAJDINE D. (2015). DEVELOPMENT OF AMAZIGHE NAMED ENTITY RECOGNITION SYSTEM USING HYBRID APPROACH. ACTES DE LA 16TH INTERNATIONAL CONFERENCE ON INTELLIGENT TEXT PROCESSING AND COMPUTATIONAL LINGUISTICS. CAIRE. EGYPT. 14-20 AVRIL 2015.

TAGHBALOUT I., ATAA ALLAH F., EL MARRAKI M. (2017). TOWARDS UNL BASED MACHINE TRANSLATION FOR MOROCCAN AMAZIGH LANGUAGE. INTERNATIONAL JOURNAL OF COMPUTATIONAL SCIENCE AND ENGINEERING.

ZENKOUAR L., AIT OUGUENGAY Y. (2007). ⵜⴰⵎⴰⵣⵉⵖⵜ ⵜⴰⵏⵎⴰⵣⵉⵖⵜ (TINML N TMAZIGHT). PRODUCTIONS DE DE L'INSTITUT ROYAL DE LA CULTURE AMAZIGHE, RABAT.

ZENKOUAR L., AIT OUGUENGAY Y. (2008). ⵏⵏⵉⵎⵓⵏ ⵏⵏⵉⵎⵓⵏ ⵏⵏⵉⵎⵓⵏ (ANIMAUX SAUVAGES ET DOMESTIQUES). PRODUCTIONS DE DE L'INSTITUT ROYAL DE LA CULTURE AMAZIGHE. RABAT.