

Finite State Morphology for Amazigh Language

Fatima Zahra Nejme¹, Siham Boulaknadel², and Driss Aboutajdine¹

¹ LRIT, Unité Associée au CNRST (URAC 29), Université Mohamed V-Agdal,
Rabat, Maroc

`fatimazahra.nejme@gmail.com, aboutaj@fsr.ac.ma`

² IRCAM, Avenue Allal El Fassi, Madinat Al Irfane, Rabat-Instituts, Maroc
`boulaknadel@ircam.ma`

Abstract. In the aim of safeguarding the Amazigh heritage from being threatened of disappearance, it seems opportune to equip this language of necessary means to confront the stakes of access to the domain of New Information and Communication Technologies (ICT). In this context, and in the perspective to build tools and linguistic resources for the automatic processing of Amazigh language, we have undertaken to develop a module for automatic lexical-analysis of the Amazigh which can recognize lexical units from texts. To achieve this goal, we have made in the first instance, a formalization of the Amazigh vocabulary namely: noun, verb and particles. This work began with the formalization of the two categories noun and particles by building a dictionary named "EDicAm" (Electronic Dictionary for Amazigh), in which entry is associated with linguistic information such as lexical categories and classes of semantics distribution.

Keywords: Amazigh language, Natural Language Processing, NooJ, lexical analysis, inflectional morphology.

1 Introduction

The Amazigh language in Morocco is considered as a prominent constituent of the Moroccan culture and this by its richness and originality. However it has been long discarded otherwise neglected as a source of enrichment cultural. Nevertheless, due to the creation of the Royal Institute of Amazigh Culture (IRCAM)¹, this language has been introduced in the public domain including administration, media also in the educational system in collaboration with ministries. It has enjoyed its proper coding in the Unicode Standard [9][28], an official spelling [4], appropriate standards for keyboard realization and linguistic structures that are being developed with a phased approach [5][13]. This process was initiated by the standardization, vocabularies construction [21][6][4][7], Alphabetical Arrangement [25], spelling standardization [4] and development of grammar rules [12].

However, this not sufficient for a less-resourced language [12] as the Amazigh to join the well-resourced language in information and Communication Technologies, mainly due to the lack of automatic language processing resources and tools. There-

¹ Institution responsible for the preservation of heritage and the promotion of the Moroccan Amazigh culture and its development (see <http://www.ircam.ma/>)

fore, a set of scientific and linguistic research are undertaken to remedy to the current situation. These researches are divided, on the one hand, on researches that are concentrated on optical character recognition (OCR) [8][18], and in the other hand, on those that are focused on natural language processing [20][10][15][11][16], which constitute the priority components of researches.

In this context, the present work deals with ongoing research efforts to build tools and linguistic resources for the Amazigh language. This paper focuses on the treatment of nouns and particles of Amazigh vocabulary with Finite-State programming environment. The description of nouns and particles was chosen, because nouns either particles constitute a major word category in Amazigh and play a major role in syntactic analysis. Secondly, the noun classification system in Amazigh is computationally interesting.

This paper is structured around five main sections: the first present a description of the Amazigh language particularities. The second expose the automatic Amazigh language processing, which includes an overview of programming environment, and the formalization of a set of rules. While the last section is dedicated to the conclusion and perspectives.

2 Amazigh Language Particularities

2.1 Historical Background

The Amazigh language also known as Berber or Tamazight (ⵜⴰⴷⵓⴷⴰⵢⵜ [tamaziɣt]), is belonged to the African branch of the Afro-Asiatic language family, also referred to Hamito-Semitic in the literature [19][23]. It is currently presented in a dozen countries ranging from Morocco, with 50% of the overall population² [14], to Egypt, passing through Algeria with 25%, the Tunisia, Mauritania, Libya, Niger and the Mali [17].

In Morocco, we distinguish between three major Amazigh dialects. Tarifit is spoken in northern Morocco, Tamazight in the Middle Atlas and south-eastern Morocco, and Tashelhit in south-western Morocco and the High Atlas.

Today, the current situation of the Amazigh language is at a pivotal point. It holds official status in Morocco. Its morphology as lexical standardization process is still underway. At present, it represents the model taught in must schools and used on media and official papers published in Morocco.

2.2 Amazigh Morphology

Amazigh is a morphologically rich language. It is highly inflected and also shows derivation to high degree. The morphological word classes in Amazigh are [12][3]: noun, verb and particles. In this paper, we are interested in noun and particles morphology.

² It presents the Amazigh population largest in number.

Noun Characteristics

In Amazigh language, noun is a lexical unit, composed from a root and a pattern. It could appear in several forms: simple form (ⵓⵔⵗⵎⵓ [argaz] “the man”), compound form (ⵓⵔⵗⵎⵓⵙⵓⵔⵓ [buhyyuf] “the famine”), or derived one (ⵔⵓⵎⵓⵙⵓⵔⵓ [amsawaɖ] “the communication”). Whether simple, compound or derived, the noun varies in gender (masculine or feminine), number (singular or plural), and state (free or construct) [12].

1. Gender: the Amazigh noun is characterized by one of grammatical gender: masculine or feminine.
 - The masculine noun: begins with one of the initial vowels: ⵓ [a], ⵉ [i] or ⵓ [u]. However, there are some exceptions as: ⵉⵎⵎⵓ [imma] “(my) mother”.
 - The feminine noun: is marked with the circumfix +...+ [t...t]. However, there are some exceptions such as nouns which have only the initial or the final + [t] of feminine morpheme: ⵜⵓⵏⵏⵓ [tadla] “the sheaf”, ⵓⵓⵔⵓⵙⵓⵔⵓ [irmuyt] “the fatigue”.
2. Number: the noun, masculine or feminine, has a singular and plural. This latter has four forms:
 - The external plural: is formed by an alternation of the first vowel ⵓ/ⵉ [a/i] accompanied by a suffixation of ⵏ [n] or one of its variants (ⵏ [in], ⵏ [an], ⵏⵏ [ayn], ⵏⵏ [wn], ⵏⵏ [awn], ⵏⵏ [wan], ⵏⵏ [win], ⵏⵏ [tn], ⵏⵏ [yin]): ⵔⵔⵔⵓⵎⵓ [axxam] “house” -> ⵔⵔⵔⵓⵎⵓⵏ [ixxamn] “houses”, ⵜⵓⵔⵓⵔⵓⵏ [tarbat] “the girl” -> ⵜⵓⵔⵓⵔⵓⵏⵏ [tirbatin] “girls”).
 - The broken plural: involves a change of internal vowels: ⵓⵔⵓⵔⵓⵔⵓ [abaɣus] “monkey” -> ⵔⵔⵔⵓⵔⵓ [ibuɣas] “monkeys”, ⵜⵓⵔⵓⵔⵓⵔⵓ [tiɣmst] “tooth” -> ⵜⵓⵔⵓⵔⵓⵔⵓ [tiɣmas] “teeth”.
 - The mixed plural: is formed by vowels’ change accompanied, sometimes by the use of the suffixation by ⵏ [n]: ⵔⵔⵔⵓⵔⵓ [izikr] “the rope” -> ⵔⵔⵔⵓⵔⵓⵏ [izakarn] “the ropes”.
 - The plural in ⵔⵔ [id]: this kind of plural is obtained by prefixing the noun with ⵔⵔ [id]. It is applied to a set of nouns including: nouns with an initial consonant, proper nouns, parent nouns, compound nouns, numerals, as well as borrowed nouns: ⵔⵔⵔⵓⵔⵓ [butgra] “the turtle” -> ⵔⵔ ⵔⵔⵔⵓⵔⵓ [id butgra] “the turtles”.
3. State: we distinguish between two states: the free state and the construct one.
 - The free state: is unmarked. The noun is in free state if it is a single word isolated from any syntactic context, a direct object, or a complement of the predictive particle ⵏ [d].
 - The construct state: involves a variation of the initial vowel. In case of masculine nouns, it takes one of the following forms: initial vowel alternation ⵓ [a]/ⵓ [u]; adding of ⵏ [w] or adding of ⵏ [y]. For the feminine nouns, it consists to drop or maintaining of the initial vowel.

Particles

The particles are a set of Amazigh words that is not assignable to noun neither to verb. It consists of several elements, namely:

1. The pronouns

The pronoun refers to any element that could replace a noun or nominal group. It may represent a nominal group already employed or designate a person participates in the communication. The paradigm of pronouns includes: the personal pronouns; the possessive; the demonstratives; the interrogative and the indefinite ones: nk [nk] “me”.

2. The prepositions

The preposition is a closed paradigm that combines simple and complex forms that express various semantic values. For the first form, it consists of several formats: I [n], xi [i], O [s], X [g], Ax [di], Xx [zg], Xh [xh], XO [gr], al [al]/ ar [ar], Ola [bla], vr [vr], dar [dar], ag [agd], d [d]. For the second one, it composed of two or three prepositions that one of which can be used adverbially: I [izdar n] “below”.

3. The adverb

The adverbs are the elements that change the meaning of a verb. Generally, they are classified according to their semantics. Thus we distinguish adverbs of place, time, quality and manner: da [da] “here” (adverb of place).

Further these elements, there are also: aspectual, orientation and negative particles; subordinates and conjunctions.

Generally, the particles are known as uninflected words. However in Amazigh, some of these particles are inflected, such as the possessive and demonstrative pronouns: wad [wad] “this one” \rightarrow wid [wid] “these ones”.

3 Automatic Amazigh Language Processing: Development and Evaluation

To develop a module for automatic lexical-analysis of the Amazigh, we followed the following steps: (1) construction of electronic dictionary, (2) formalization of the two categories noun and particles and (3) evaluation.

To achieve this goal, we use finite state machines integrated in the linguistic development platform NooJ.

3.1 NooJ Platform

NooJ³, released in 2002 by Max Silberstein [22], is a linguistic development platform that provides a set of tools and methodologies for formalizing and developing a set of Natural Language Processing (NLP) applications. It presents a package of finite state tools that integrates a broad spectrum of computational technology from finite state automata to augmented/recursive transition networks. Thus, it presents a complete platform for formalizing various types of textual phenomena (orthography, lexical and productive morphology, local, structural and transformational syntax). For each of these formalization levels, NooJ propose a methodology, one or more formalisms,

³ See <http://www.nooj4nlp.net/> for information of NooJ.

tools, software development and a corresponding parser that can be used to test each piece of the linguistic formalization on a large corpora.

Currently, users of NooJ form a large community including linguists, teachers and computer specialists in NLP, who share the same scientific and technical goals, to develop a large-coverage of language resources in fifteen languages (Arabic, Armenian, Bulgarian, Catalan, Chinese, English, French, Hebrew, Hungarian, Italian, Polish, Portuguese, Spanish, Vietnamese and Belarusian).

Given these advantages, we have undertaken to adopt NooJ for formalization, description and analysis of Amazigh language. We begin our work by the formalization of the Amazigh language vocabulary. This formalization is described and stored into inflectional grammars, and can recognize all the corresponding inflected forms. To test these grammars, we built an electronic dictionary which the lexical entries are attached to a set of linguistic information automatically generated using inflectional grammars which will be used for lexical analysis of texts.

3.2 Building of Dictionary

As part of developing the analysis module for the Amazigh language, we elaborate our dictionary «EDicAM» (Electronic Dictionary for Amazigh) for simple nouns (ex. ⵓⵛⵔⵉⵎ [axxam]) and inflected particles (ex. ⵉⵎⵉⵏ [winu] “mine” (masc.) -> ⵜⵉⵎⵉⵏ [tinu] “mine” (fem.)).

Each entry into Amazigh dictionary generally presents following details: the lemmas, Lexical category, Semantic feature, French translation and the inflexional paradigm.

Our dictionary contains, currently, 4480 entries of simple nouns as singular masculine and 44 particles. These particles include personal pronouns, demonstrative and possessive ones. Each entry in our dictionary is associated with an inflectional class allowing to generate all the corresponding inflected forms (feminine, plural and constructed state).

3.3 Formalization

Given that any linguistic analysis must go through a first step of lexical analysis, which consists in testing membership of each word of the text to the lexicon of the language, we began our work by formalization of the Amazigh vocabulary. This study is focus on the formalization of the tow sets: nouns and particles, using finite state machines integrated to linguistic development platform NooJ.

Formalization of Morphological Rules

This study presents the formalization of the noun and particles categories in the NooJ platform. For this, a set of rules has been defined allowing to generate from each entry, its inflectional information: gender, number and state for the nouns.

The formalization is based on use of some generic predefined commands such as:

- <LW> move at the beginning of lemma,
- <RW> move at the end of lemma,

- <S> delete current character,
- delete last character,
- <L> go left,
- <R> go right.

The nouns

To formalize the inflexional rules of the noun, we have created, through graphs integrated in the linguistic development platform NooJ, an inflectional analysis system describing flexing models in standard Amazigh of Morocco.

1. Gender

To formalize the gender we built this rule that generate from a masculine entry its feminine correspondent.

Table 1. Example of gender rule

The rules in NooJ	Explanation	Examples
<LW>+<RW>+	This rule adds the morpheme "+" [t] at the beginning and at the end of the noun.	$\xi\ominus\mathfrak{H}\xi$ [isli] "married"(masc.) -> $+\xi\ominus\mathfrak{H}\xi+$ [tislit] "married" (fem.).

2. Number

For the Amazigh plural, we have many plural forms which are generally unpredictable due to Amazigh complex morphology. We searched formal rules to unify the calculation of these plural forms. To achieve this goal, we have relied on the new grammar of Amazigh [12], on the works of Boukhris [13] and those of Oulhaj [24]. According to these works and to an heuristic study of the nouns in the Taifi dictionary [27] and those of Amazigh language vocabulary [5], we have raised, actually, 303 classes which 97 classes is for the external plural, 99 for the broken plural, 104 for the mixed plural and 3 classes for the plural in $\xi\wedge$ [id]. Each class corresponds to a scheme and contains a set of nouns, for example for the external plural, if the noun begins and ends with a 'o...o' [a...a], the plural forms will be built by a vowel alternation accompanied by suffixation of \mathfrak{H} [tn]. Thereafter, we provided some examples of rules for each of these plural types.

- The external plural

Table 2. Plural form for the masculine nouns beginning and ends with o [a]

The rules in NooJ	Explanation	Examples
<LW> ξ <S><RW> \mathfrak{H}	The initial vowel is transformed into ξ and the suffix \mathfrak{H} [tn] is add at the end of the noun.	$\circ\ominus\xi\circ$ [asira] "desk" -> $\xi\ominus\xi\circ\mathfrak{H}$ [isiratn] "desks".

- The broken plural

Table 3. Plural for the nouns of VCn form

The rules in Nooj	Explanation	Examples
<LW>ξ<S><RW><L>◦	The rule changes the initial vowel into ξ [i] and include ◦ [a] before the final consonant.	ⵎⵣⵣⵓⵎ [azgzl] “abbreviation” -> ξⵎⵣⵣⵓⵎ [izgzal] “abbreviations”

- The mixed plural

Table 4. Example of plural forms for the masculine nouns

The rules in Nooj	Explanation	Examples
<LW>ξ<S><RW>ⵢ<L2>ξ	The rule change the initial vowel into ξ [i], include the vowel ξ [i] before the last consonant and add a suffix ⵢ [n] at the end of the noun.	ⵎⵓⵔⵓⵎ [ahudr] “the fact of lean” -> ξⵎⵓⵔⵓⵎⵢ [ihudirn].

- The plural in ξⵢ [id]

Table 5. Example of plural in ξⵢ [id]

The rules in Nooj	Explanation	Examples
<LW>ξⵢ " "	The rule adds ξⵢ [id] before the noun.	ⵉⵔⵓⵎ [butgra] “tortoise” -> ξⵢ ⵉⵔⵓⵎ [id butgra] “tortoises”.

3. State

Table 6. Example of constructed state

The rules in Nooj	Explanation	Examples
<LW><R>ⵔ	The rule deletes the initial vowel and adds ⵔ [u] at the beginning of the noun.	ⵎⵓⵔⵓⵎ [afiras] “pear” -> ⵔⵎⵓⵔⵓⵎ [ufiras].

The particles

Similarly to the noun, we formalized the inflectional particles, such as: demonstrative pronouns, possessive pronouns etc., with a collection of graphs and subgraphs presenting the inflectional grammars.

In the following, we will cite an example of possessive pronouns. The possessive pronouns are formed by the combination of determination supports with affixes pronouns of noun complement.

1. Gender

To formalize the gender of particles, we built five morphological rules that generates from a masculine entry its feminine correspondent.

Table 7. Example of gender graph of particles

The rules in Nooj	Explanation	Examples
<LW><S>+	This rule removes the first consonant and adds the discontinuous morpheme + [t] at the beginning of the particle.	ⲓⲛⲓⲛ [winu] “mine” (masc.) -> +ⲛⲓⲛ [tinu] “mine” (fem.).

2. Number

The plural of particles can be formed either by a suffixation ⲓ [n] or by a vowel alternation / final consonant sometimes accompanied by a suffixation ⲓ [n]. The following table shows the graph corresponding to the second case.

Table 8. Example of a graph for the particles plural

The rules in Nooj	Explanation	Examples
<RW>ⲓ	The final consonant is transformed into ⲓ [u] and the suffix ⲓ [winnun] “yours”. [n] is added at the end of the particle.	ⲓⲛⲓⲛⲓ [winnk] “yours” -> ⲓⲛⲓⲛⲓⲓ [winnun] “yours”.

For the rest of non-inflectional particles, we construct a set of morphological grammars that allow these particles to be known on the texts. In the following, we will cite the example of adverb.

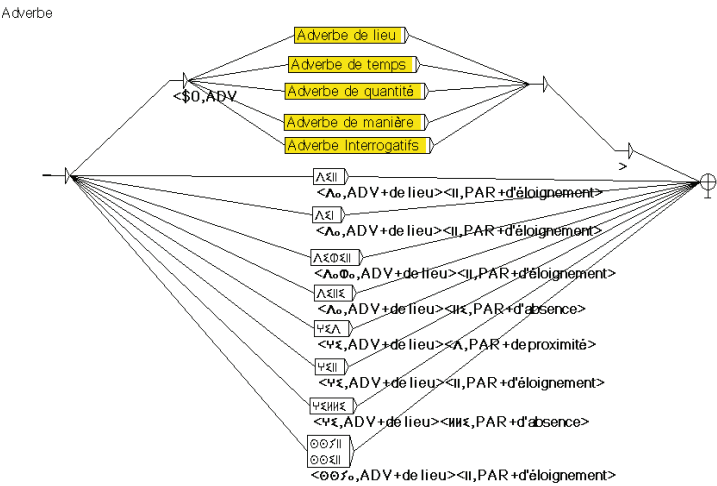


Fig. 1. Example for a formalization of Adverb

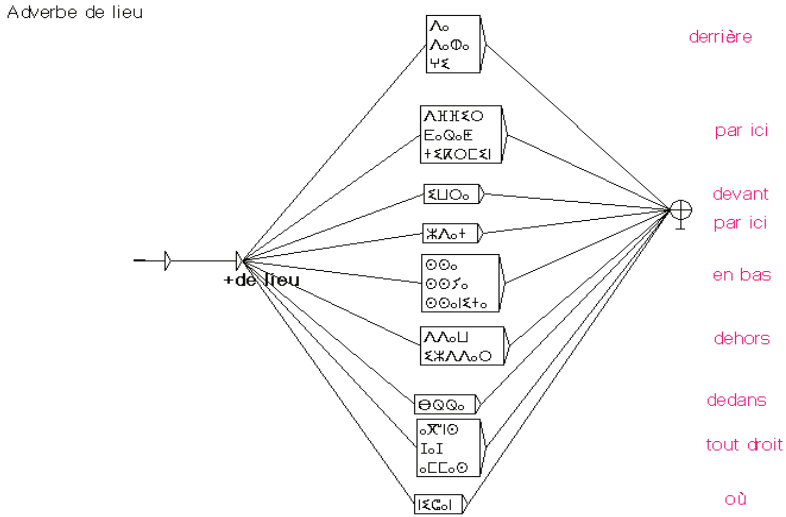


Fig. 2. Formalization of adverb of place

For the complex particles, we construct a set of local grammars which we cite an example:

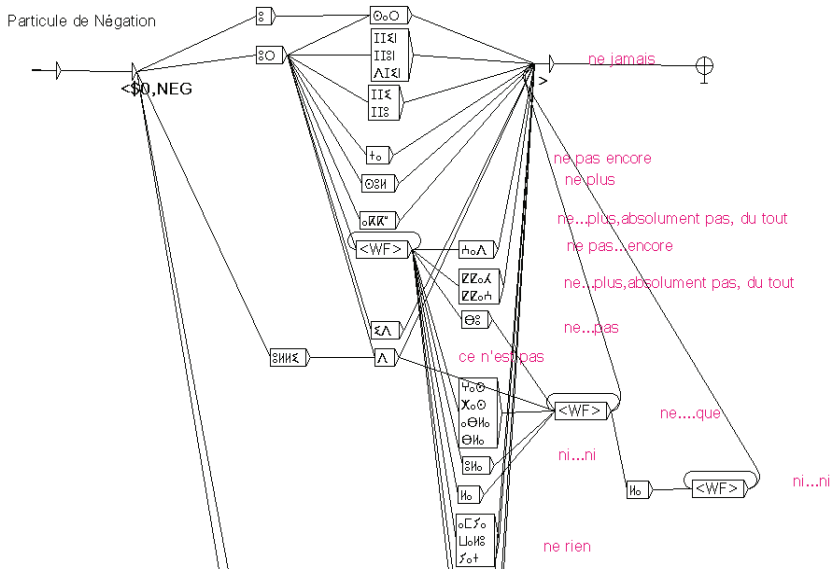


Fig. 3. Formalization of complex particles

3.4 Evaluation

To evaluate our resources, we extract a list of 1000 nouns based on lexicon school [1] and 119 particles from the new grammar of Amazigh [12].

We present in the table below, an example of the result of experiments obtained:

Table 9. Example of result of our experiments

Results	Number of recognized forms		Number of unrecognized forms	
	Number	%	Number	%
Noun	982	98,2%	18	1,8%
Particles	109	91,59%	10	8,4%

The results allow us to review, correct and complete all our resources.

As indicated in the table of evaluation, the applications of our resources have resulted a lexical coverage about 97%. The unrecognized word forms (from nouns and particles) include:

- 14 entries: corresponds to the foreign nouns that are not part of our electronic dictionary;
- 10 entries: corresponds to the foreign particles that are not part of our morphological grammar;
- 4 entries: corresponds to the nouns inflected with a rule that does not correspond to its exact form, for example if we take the feminine noun as follows: "ⵜⵉⵍⵉⵎⵉⵔⵜ" [Timelwit - soft], normally the general form of inflected feminine nouns is as follows:
 - If the noun is in the form 'ⵜ...ⵔ' (v: vowel), the vowel 'o' is transformed into 'ⵉ' [i] and a suffix of 'ⵉ' [in] is applied.
 - If the noun is in the form 'ⵜ...ⵔ' (c: consonant), the vowel 'o' is transformed into 'ⵉ' [i], the last 'ⵔ' is deleted, and a suffix 'ⵉ' [in] is applied.

Therefore, the noun inflexion in the vocabulary is ⵜⵉⵍⵉⵎⵉⵔⵜ. These types of nouns form the exceptions in Amazigh.

4 Conclusion and Future Works

The main objective of this work was to develop a module for automatic lexical-analysis of the Amazigh which can recognize lexical units from texts. To achieve this goal, we have made in the first instance, a formalization of the morpho-syntactic categories noun and particles. This work demonstrates the application of finite state approach in the analysis of Amazigh vocabulary.

Based on the results obtained, the preferred method seems to be appropriate and encouraging. However, other rules must be added to improve the result.

For future work we plane to formalize the rest of Amazigh word categories. The aim is to develop a morphological analyser of Amazigh that will be an input for many other applications such as learning systems and machine translation.

References

1. Agnaou, F., Bouzandag, A., El Baghdadi, M., El Gholb, H., Khalafi, A., Ouqua, K., Sghir, M.: *Lexique scolaire*. IRCAM, Rabat (2011)
2. Ameur M., Boumalk A.(DIR): Standardisation de l'amazighe, Actes du séminaire organisé par le Centre de l'Aménagement Linguistique à Rabat, Décembre 8-9 (2003); Publication de l'Institut Royal de la Culture Amazighe, Série: Colloques et seminars (2004a)
3. Ameur, M., Bouhjar, A., Boukhris, F., Boukouss, A., Boumalk, A., Elmedlaoui, M., Iazzi, E., Souifi, H.: *Initiation à la langue amazighe*. IRCAM, Rabat (2004b)
4. Ameur, M., Bouhjar, A., Boukhris, F., Boukouss, A., Boumalk, A., Elmedlaoui, M., Iazzi, E.: *Graphie et orthographe de l'amazighe*. IRCAM, Rabat (2006a)
5. Ameur, M., Bouhjar, A., Boukhris, F., Elmedlaoui, M., Iazzi, E.: *Vocabulaire de la langue amazighe (Français-Amazighe)*. série: Lexiques (1). IRCAM, Rabat (2006b)
6. Ameur, M., Bouhjar, A., Boumalk, A., El Azrak, N., Laabdelouai, R.: *Vocabulaire des médias (Français-Amazighe-Anglais-Arabe)*. série: Lexiques (3). IRCAM, Rabat (2009a)
7. Ameur, M., Bouhjar, A., Boumalk, A., El Azrak, N., Laabdelouai, R.: *Vocabulaire grammatical*. série: Lexiques (5). IRCAM, Rabat (2009b)
8. Amrouch, M., Rachidi, A., El Yassa, M., Mammass, D.: *Handwritten Amazigh Character Recognition Based On Hidden Markov Models*. International Journal on Graphics, Vision and Image Processing 10(5), 11–18 (2010)
9. Andries, P.: *Unicode 5.0 en pratique, Codage des caractères et internationalisation des logiciels et des documents*. Dunod, France, Collection InfoPro (2008)
10. Ataa Allah, F., Jaa, H.: *Etiquetage morphosyntaxique: Outil d'assistance dédié à la langue amazighe*. In: *Proceedings of the 1er Symposium International sur le Traitement Automatique de la Culture Amazighe*, Agadir, Morocco, pp. 110–119 (2009)
11. Ataa Allah, F., Boulaknadel, S.: *Online Amazigh Concordancer*. In: *Proceedings of International Symposium on Image Video Communications and Mobile Networks*, Rabat, Maroc (2010)
12. Berment, V.: *Méthodes pour informatiser des langues et des groupes de langues peu dotées*. Thèse de doctorat de l'Université J. Fourier - Grenoble I, France (2004)
13. Boukhris, F., Boumalk, A., Elmoujahid, E., Souifi, H.: *La nouvelle grammaire de l'amazighe*. IRCAM, Rabat (2008)
14. Boukous, A.: *Société, langues et cultures au Maroc: Enjeux symboliques*, Casablanca, Najah El Jadida (1995)
15. Boulaknadel, S.: *Amazigh ConCorde: an appropriate concordance for Amazigh*. In: *Proceedings of the 1er Symposium International sur le Traitement Automatique de la Culture Amazighe*, Agadir, Morocco, pp. 176–182 (2009)
16. Boulaknadel, S., Ataa Allah, F.: *Building a standard Amazigh corpus*. In: *Proceedings of the International Conference on Intelligent Human Computer Interaction*, Prague, Tchech (2011)
17. Chaker, S.: *Le berbère*, Actes des langues de France, 215–227 (2003)
18. Es Saady, Y., Rachidi, A., El Yassa, M., Mammass, D.: *Printed Amazigh Character Recognition by a Syntactic Approach using Finite Automata*. International Journal on Graphics, Vision and Image Processing 10(2), 1–8 (2010)

19. Greenberg, J.: *The Languages of Africa*. The Hague (1966)
20. Iazzi, E., Outahajala, M.: Amazigh Data Base. In: *Proceedings of HLT & NLP Workshop within the Arabic World: Arabic Language and Local Languages Processing Status Updates and Prospects*, Marrakech, Morocco, pp. 36–39 (2008)
21. Kamel, S.: *Lexique Amazighe de géologie*. IRCAM, Rabat (2006)
22. Silberztein, M.: An Alternative Approach to Tagging. In: Kedad, Z., Lammari, N., Métais, E., Meziane, F., Rezgui, Y. (eds.) *NLDB 2007*. LNCS, vol. 4592, pp. 1–11. Springer, Heidelberg (2007)
23. Ouakrim, O.: *Fonética y fonología del Bereber*, Survey at the University of Autònoma de Barcelona (1995)
24. Oulhaj, L.: *Grammaire du Tamazight*. Imprimerie Najah El Jadida (2000)
25. Outahajala, M.: Les normes de tri, Du clavier et Unicode. *La typographie entre les domaines de l'art et de l'informatique*, Rabat, Morocco, pp. 223–237 (2007)
26. Outahajala, M., Zekouar, L., Rosso, P., Martí, M.A.: Tagging Amazigh with An-CoraPipe. In: *Proceeding of the Workshop on Language Resources and Human Language Technology for Semitic Languages*, Valletta, Malta, pp. 52–56 (2010)
27. Taifi, M.: *Le lexique berbère (parlers du Maroc central)* (1988)
28. Zenkouar, L.: Normes des technologies de l'information pour l'ancrage de l'écriture amazighe. *Etudes et Documents Berbères* 27, 159–172 (2008)