

## Jeu d'étiquettes morphosyntaxiques de la langue amazighe\*

Fadoua Ataa Allah (1), Siham Boulaknadel (1) et Hamid Souifi (2)

(1) CEISIC, (2) CAL - IRCAM

يعتبر الوسم النحوي علما لتصنيف وتدليل كلمات النص أو المتن النصي إلى نوع محدد من أجزاء الكلام بناء على تعريف الكلمة وسياقها. ويعد حجر الزاوية للعديد من تطبيقات المعالجة الآلية للغات الطبيعية وخاصة التصحيح النحوي، التحليل الدلالي والترجمة الآلية. ويعتمد في ذلك على وضع وتحديد لائحة أجزاء الكلام المناسبة للخصوصيات اللغوية للغة المدروسة.

في هذا المقال، نقترح مجموعة من العلامات النحوية تضم تصنيفا مفصلا لأجزاء كلام اللغة الأمازيغية، آخذين بعين الاعتبار الخصائص الصرفية والتركيبية للغة الأمازيغية ومستنديين في بناء هذه الأقسام على نموذج "EAGLES"، الذي يساهم في الانفتاح على التطبيقات المتعددة اللغات.

*This work aims to provide the Amazigh language with a morphosyntactic tagset. In this process, morphology and syntax are considered as an inextricable asset. This tagset will assign to each meaningful unit information concerning the "shape variations of signifiers, their amalgams and their discontinuity" and information about its function in the statement. The proposed tagset is based on EAGLES guidelines in order to ensure the reuse of corpora and language's comparability in natural language processing.*

*Ce travail se veut une contribution à l'élaboration d'un Jeu d'étiquettes morphosyntaxiques de la langue amazighe. Il s'appuie sur la morphologie et la syntaxe en tant qu'un tout indissociable. Ce jeu permettra d'attribuer à chaque unité significative des informations sur les « variations de forme de signifiants, à leurs amalgames et à leur discontinuité<sup>1</sup> » et sur sa fonction dans l'énoncé. Le jeu que nous proposons ici se base sur les recommandations EAGLES, visant la réutilisation des corpus et la comparabilité entre les langues dans le domaine du traitement automatique du langage naturel.*

\* Nous tenons à exprimer nos vifs remerciements à Abdallah Boumalk et Rachid Laabdelaoui (CAL, IRCAM) pour avoir bien voulu relire et commenter ce travail.

<sup>1</sup> Martinet, A. (*Grammaire fonctionnelle du français, cf. 1.8*)

## 1. Introduction

L'étiquetage morphosyntaxique, appelé aussi *étiquetage grammatical* ou *part-of-speech tagging* (*POS tagging*, en anglais), est un préalable pour de nombreuses applications du traitement automatique des langues, particulièrement la correction grammaticale, l'analyse sémantique, l'analyse pragmatique et la traduction automatique. Toutefois, il nécessite le développement d'un jeu d'étiquettes adéquat aux spécificités linguistiques de la langue étudiée.

Cet étiquetage consiste à attribuer à toute unité significative du texte (occurrence d'un corpus), et à l'aide d'un outil informatique, un symbole correspondant à son comportement syntaxique<sup>2</sup> (nom, verbe, pronom, adverbe, préposition...), et des informations morphologiques qu'il affiche dans un contexte précis du corpus linguistique (masculin, singulier,...). Idéalement, un jeu d'étiquettes doit permettre de<sup>3</sup> :

- représenter la richesse des informations lexicales ;
- représenter l'information nécessaire à la désambiguïsation<sup>4</sup> en contexte des étiquettes morphosyntaxiques ;
- et d'encoder les informations utiles au traitement linguistique pour lequel l'étiquetage morphosyntaxique a été déployé.

Le choix de jeux d'étiquettes est particulièrement délicat. Ainsi, une panoplie de jeux d'étiquettes pour chaque langue a été développée au fil des années et au gré des projets par différents groupes. Face à ces divergences de pratique, qui forment un obstacle à l'échange et à la réutilisation des corpus, basées parfois sur des conceptions différentes de ce que doit être un jeu d'étiquettes de données, d'importants efforts d'uniformisation et d'alignement ont été déployés au sein de projets internationaux tels que EAGLES (EAG, 1996) et MULTEX (MUL, 1996).

Dans un souci de parvenir à une « standardisation » d'étiquettes morphosyntaxiques pour l'amazighe lui assurant la comparabilité entre langues, notamment dans un système multilingue, nous avons élaboré un jeu d'étiquettes en nous inspirant des recommandations fournies par le groupe EAGLES<sup>5</sup> en matière d'annotation morphosyntaxique. Ces dernières sont assez intéressantes, dans la

---

<sup>2</sup> Bien que les unités significatives (ou mots) appartiennent toutes à une ou plusieurs parties du discours, selon le contexte de leur emploi, des difficultés d'étiquetage peuvent découler de l'ambiguïté des rôles grammaticaux de certains mots dans certaines langues, où la mutation catégorielle paraît vivante et dans lesquelles la classification des catégories grammaticales s'avère difficile à établir.

<sup>3</sup> Nous reproduisons les mêmes objectifs visés par *Technolanguage.net* cités sur [http://www.technolanguage.net/article.php?id\\_article=296](http://www.technolanguage.net/article.php?id_article=296). Site consulté en septembre 2013.

<sup>4</sup> La désambiguïsation d'un mot consiste à faire cesser son ambiguïté en ne retenant qu'un seul de ses sens lié à une seule forme morphosyntaxique, et ce, dans le but de le rendre plus aisé à comprendre.

<sup>5</sup> <http://www.ilc.cnr.it/EAGLES96/annotate/annotate.html>. Site consulté en septembre 2013.

mesure où elles proposent un cadre, qui peut être aisément suivi, fondé sur une distinction entre *étiquettes obligatoires*, *étiquettes recommandées* et *extension particulière*.

## 2. Ambiguïté en amazighe

L'amazighe se caractérise par un comportement morphologique riche dans la formation des signifiants ; par flexion, par dérivation ou par **agglutination**. Ce qui le laisse confronté à un problème d'ambiguïté pénalisant tout processus de traitement automatique. Ainsi, l'ambiguïté se manifeste sous différentes formes selon les niveaux de traitement lexical, morphologique et syntaxique.

Dans ce contexte, nous essayons, à travers cet article, de décrire l'une des formes d'ambiguïté qui s'imposent pour le lexique amazighe et que seule la syntaxe pourrait désambiguïser comme c'est le cas<sup>6</sup> dans les exemples qui suivent :

- *Ambiguïté entre Nom et Verbe :*

Nom : ⵏⵉⵍⵉⵍⵉ [illi] (Litt. fille de moi) « ma fille »,

Verbe : ⵏⵉⵍⵉⵍⵉⵏ [ur illi] (Litt. ne pas être) « il n'existe pas ».

- *Ambiguïté entre Nom et Adjectif :*

Nom : ⵏⵉⵍⵉⵍⵉⵏ ⵏⵉⵍⵉⵍⵉⵏ<sup>8</sup> [zriγ amqgran] (Litt. voir je amqgran) « j'ai vu Amqgran »,

Adjectif : ⵏⵉⵍⵉⵍⵉⵏ ⵏⵉⵍⵉⵍⵉⵏ [argaz amqgran] (Litt. homme âgé) « l'homme âgé ».

- *Ambiguïté entre Particule modale du futur (PMF) et Particule démonstrative de proximité :*

PMF : ⵏⵉⵍⵉⵍⵉⵏ ⵏⵉⵍⵉⵍⵉⵏ [ad ffγγ] (Litt. Particule modale, sortir-1PS, « je sortirai »,

Particule démonstrative de proximité : ⵏⵉⵍⵉⵍⵉⵏ ⵏⵉⵍⵉⵍⵉⵏ [argaz ad] (Litt. homme ce) « cet homme-ci ».

- *Ambiguïté entre Particule modale de l'inaccompli (PMI) et Adverbe de lieu de proximité : ⵏⵉⵍⵉⵍⵉⵏ.*

PMI : ⵏⵉⵍⵉⵍⵉⵏ ⵏⵉⵍⵉⵍⵉⵏ [da ichtta] (Litt. PMI, 3PMS manger (inacc.)) « Il est en train de manger » ;

Adverbe de lieu : ⵏⵉⵍⵉⵍⵉⵏ ⵏⵉⵍⵉⵍⵉⵏ ⵏⵉⵍⵉⵍⵉⵏ [ssrs it da] (Litt. Poser (V. impératif), masc. sing. le, ici) « Pose-le ici ».

<sup>6</sup> Le problème de la mutation catégorielle est omni présent dans d'autres langues, c'est le cas du Français (ex. *écailler* (nom « personne qui ouvre et vend des huîtres, des fruits de mer »/verbe « dépouiller un poisson de ses écailles »), *faire* (nom « manière de faire une œuvre »/verbe « construire, fabriquer »), *gabarié*, *baiser*... ou du Vietnamien (ex. *trên* (préposition « sur »/nom « le supérieur »), *trong* (préposition « dans »/nom « l'intérieur »)... voir à ce propos Nguyen *et al.* (2003), « Une étude de cas pour l'étiquetage morphosyntaxique de textes vietnamiens », cf.3.2.

<sup>7</sup> Accompli négatif du verbe ⵏⵉⵍⵉⵍⵉⵏ *ili* « être ».

<sup>8</sup> Le nom ⵏⵉⵍⵉⵍⵉⵏ [amqgran] « Amqgran » peut être un nom propre.

- *Ambiguïté entre Auxiliaire (ou particule) de prédication, coordonnant et Particule d'orientation de rapprochement :*

*Auxiliaire de prédication :  $\Lambda$   $\circ\text{X}\circ\text{X}$  [d argaz] (Litt. c'est homme) « c'est un homme »,*

*Coordonnant :  $\circ\text{X}\circ\text{X}$   $\Lambda$   $\dagger\text{O}\Theta\circ\dagger$  [argaz d trbat] (Litt. homme et fille) « l'homme et la fille »,*

*Particule d'orientation spatiale de rapprochement :  $\xi\Lambda\Lambda\circ$   $\Lambda$   $\text{g}\text{O}\Theta\circ$  [idda d urba] (Litt. 3PMS.Venir ici enfant) « l'enfant est venu (vers ici)».*

### 3. Pourquoi le modèle EAGLES ?

Le groupe *EAGLES* (*Expert Advisory Group on Language Engineering Standards*), fondé en février 1993, est une initiative de la commission européenne du programme « Ingénierie et Recherche Linguistique » (*Linguistic Research and Engineering, LRE*). Il a pour but l'élaboration de standards des ressources langagières (corpus écrit et oral, lexiques électroniques) et des moyens de structuration et d'exploitations, ainsi que des procédés d'évaluation de ressources et d'outils. Ainsi, plusieurs acteurs industriels, professionnels et universitaires à l'échelon européen ont participé à la réalisation d'une batterie de recommandations, notamment en matière d'annotation morphosyntaxique.

Les spécifications proposées par EAGLES sont le fruit d'une étroite collaboration avec le projet MULTEX (*Multilingual Text Tools and Corpora*). Elles résultent de l'observation et de l'analyse d'un ensemble de projets de corpus et de lexiques, qui ont permis de relever des traits communs aux différentes langues, et de déterminer un noyau d'informations morphosyntaxiques sur lesquelles un accord assez large peut être établi. Ce noyau se complète par des couches d'informations optionnelles, ou propres à des applications particulières.

Bien que le modèle EAGLES ne soit pas applicable tel quel à l'amazighe, il peut servir de point de départ à de nombreux acteurs pour disposer de ressources linguistiques directement exploitables ainsi que d'outils de traitement réutilisables. Le modèle a été développé d'une manière à systématiser les stratégies d'étiquetage dans un environnement aussi bien monolingue que multilingue, et par conséquent à assurer la comparabilité de l'amazighe avec d'autres langues tout en permettant une grande flexibilité.

### 4. Définition du « jeu d'étiquettes » pour l'amazighe

Pour l'amazighe, la question de la classification des catégories grammaticales (*cf.* 2) est une tâche difficile et toujours en débat, du fait que cette langue est en voie de standardisation et que ses règles sont en phase de codification. A ce jour, malgré les quelques études menées<sup>9</sup> dans ce cadre, il n'existe aucun standard reconnu pour

---

<sup>9</sup> Entre autres : (Loikkanen, 2007), (Ataa Allah et Jaa, 2009), (Outahajala *et al.*, 2010), ....

les catégories des mots. C'est dans cette perspective que nous proposons un jeu d'étiquettes pour les applications du traitement automatique de l'amazighe.

Sachant que la grammaire se définit comme l'étude systématique des éléments constitutifs d'une langue et la science qui permet de montrer la structure d'une langue et d'expliquer les règles de changement et de combinaison des mots formant un énoncé, elle se subdivise en deux parties : la morphologie, qui étudie les changements des mots, et la syntaxe qui s'intéresse à la combinaison des mots en formes plus étoffées que de simples mots. Néanmoins, la catégorie grammaticale représente l'unité dialectique de la valeur grammaticale et de la forme grammaticale, reflétée par les différentes oppositions existant dans un système syntaxique d'une langue.

Pour bien refléter toutes les relations syntaxiques possibles, généralement, il faut disposer d'un important jeu d'étiquettes. Cependant, plus le nombre des étiquettes est important plus la tâche d'annotation ou d'étiquetage est difficile. D'où la nécessité de gérer ce compromis afin de parvenir à un jeu d'étiquettes assez précis et de taille acceptable. Ainsi, en nous basant, d'une part, sur la définition du mot graphique en amazighe<sup>10</sup>, et d'autre part, sur les recommandations formulées par EAGLES, nous avons proposé deux listes d'étiquettes<sup>11</sup> spécifiques aux caractéristiques de la langue amazighe. Une assez précise, elle contient les étiquettes obligatoires qui reflètent toutes les oppositions du système syntaxique (*cf.* Tableau 1). L'autre traduit toutes les relations syntaxiques contenant les étiquettes recommandées, qui fournissent des indications plus précises telles que le nombre, le genre et l'état pour les noms et les adjectifs, le temps, le mode et la personne pour les verbes.

#### 4.1. Etiquettes obligatoires

Les étiquettes obligatoires comprennent les parties du discours d'une langue, représentées par les classes de mots ayant les mêmes propriétés sémantiques et grammaticales. Ces classes sont principalement caractérisées par le sens général catégoriel, puis par la forme grammaticale ainsi que la fonction syntaxique des mots.

Compte tenu des spécifications lexicales de chaque classe en amazighe<sup>12</sup>, le jeu d'étiquettes que nous proposons (Tableau 1) contient douze (12) catégories

---

<sup>10</sup> Selon (Ameur *et al.*, 2004 : 34; Boukhris *et al.*, 2008 : 27), un mot graphique est constitué d'une séquence de lettres, et éventuellement d'une seule lettre, délimitée par deux blancs typographiques.

<sup>11</sup> Pour une fiabilité optimale des systèmes d'étiquetage, il est à recommander d'avoir un important jeu d'étiquettes représentant toutes les relations morphosyntaxiques de la langue annotée, plus le corpus est volumineux, plus le résultat est précis. Mais cela ne va pas sans contrainte, car plus le jeu d'étiquette est important, plus la tâche d'annotation est difficile aussi.

<sup>12</sup> Ces classes se basent sur les traits morphologiques et sur les parties du discours décrites dans *Initiation à l'amazighe* (Ameur *et al.* : 2004) et la *Nouvelle grammaire de l'amazighe* (Boukhris *et al.*, 2008 : 27).

principales, à l'encontre du modèle EAGLES qui en propose 13 et dont la catégorie « Article » n'est pas supportée par l'amazighe. En outre, la catégorie « Unique » a été remplacée par la catégorie « Particule » afin de répondre aux spécificités morphosyntaxiques de l'amazighe.

N°	CATEGORIES	N°	CATEGORIES
1	Nom (N)	7	Particule (P)
2	Verbe (V)	8	Conjonction (C)
3	Adjectif (AJ)	9	Interjection (I)
4	Pronom (PR)	10	Résiduel (R)
5	Adverbe (AV)	11	Numéral (NU)
6	Préposition <sup>13</sup> (AP)	12	Ponctuation (PU)

Tableau 1 : Liste des étiquettes obligatoires

#### 4.2. Etiquettes recommandées

Dans cette section, nous envisageons de définir les étiquettes recommandées, où chaque catégorie spécifiée dans la liste obligatoire sera subdivisée à l'aide d'étiquettes plus spécifiques relevant de la catégorie morphologique par le changement de la forme du mot et de la propriété sémantique.

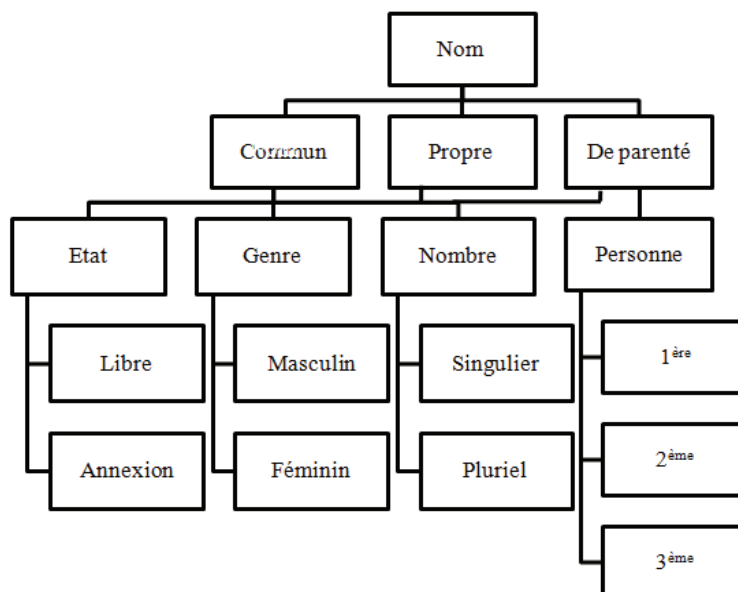
##### 1. Nom (N) :

Le nom peut être de type commun, propre ou de parenté<sup>14</sup>. Tous les types varient en genre, en nombre et en état<sup>15</sup>. En outre, les noms de parenté varient aussi en fonction de la personne.

<sup>13</sup> Le modèle EAGLES propose l'étiquette adposition (AP) qui comprend la préposition et la postposition.

<sup>14</sup> Selon les recommandations d'EAGLES, nous avons classé le nom du type numéral dans la catégorie « Numéral ».

<sup>15</sup> C'est un concept grammatical morphosyntaxique qui affecte l'initiale vocalique des noms masculins par changement de formes :  $\circ \rightarrow \text{u/wa}$ ,  $\xi \rightarrow \text{yi}$  et  $\text{u} \rightarrow \text{wu}$ . Pour les noms féminins, ils marquent l'état d'annexion par l'effacement de la voyelle ( $\circ$  [a] /  $\varepsilon$  [i]) placée après la marque préfixée  $\dagger$  [t].



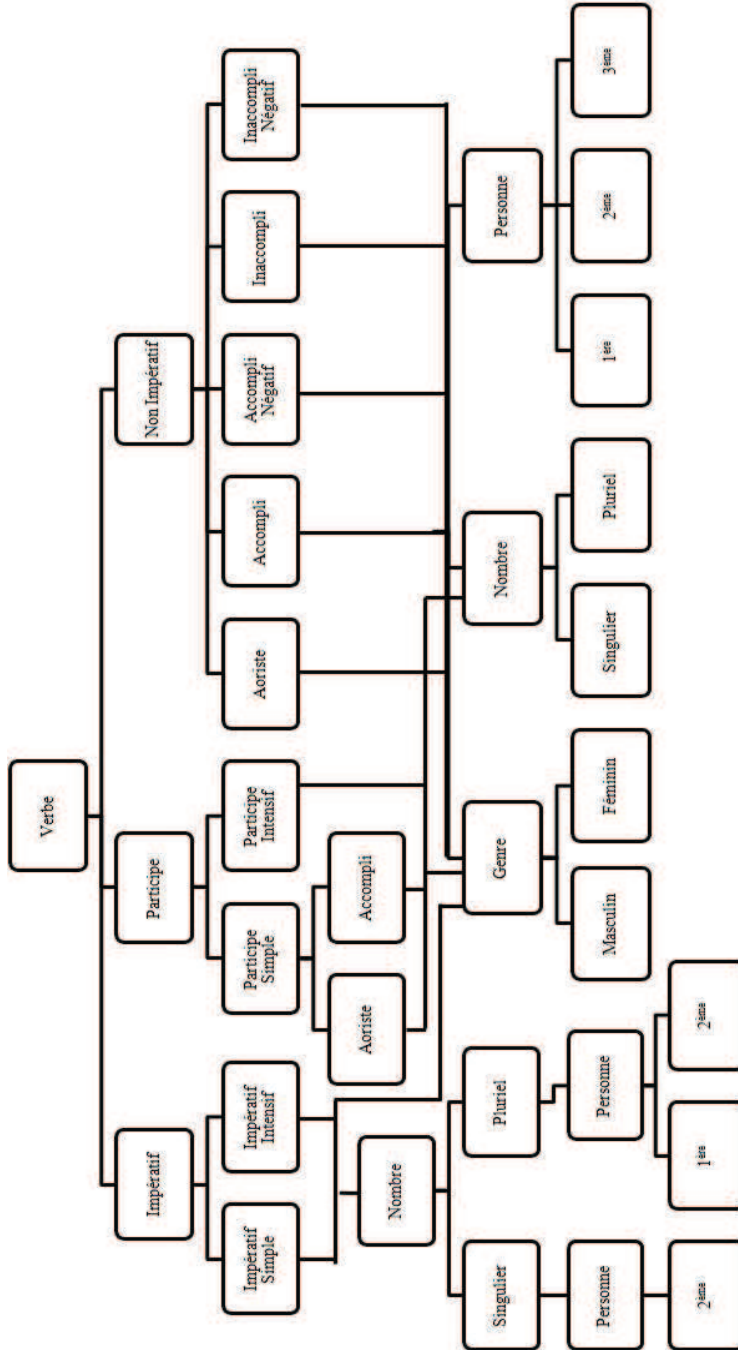
## 2. Verbe (V) :

Le verbe se combine avec les modalités aspectuelles des quatre thèmes : l'aoriste, l'accompli, l'inaccompli et l'accompli négatif. Néanmoins, afin de couvrir toutes les variantes amazighes du Maroc, nous introduisons dans le schéma du verbe, en plus des quatre thèmes verbaux communs, la forme de l'inaccompli négatif présente dans les parlers du Rif et de Figuig.

Dans sa conjugaison, le verbe reçoit des désinences verbales qui sont de trois types : (1) non impératives, (2) impératives et (3) participiales.

Lorsque les deux formes simples<sup>16</sup>, de l'impératif et du participe, se combinent avec les modalités de l'inaccompli (gémination d'une consonne radicale ou la préfixation de : ++ *tt*), elles sont appelées successivement impératif intensif et participe intensif.

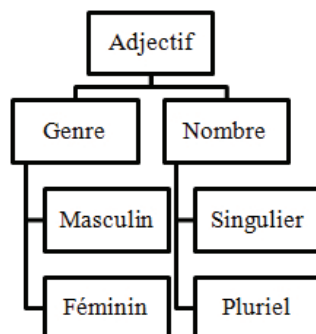
<sup>16</sup> Le participe simple est obtenu par la combinaison des désinences de la forme participiale avec l'aoriste ou l'accompli (Manuel de conjugaison, p. 14).





### 3. Adjectif (AJ) :

L'adjectif est un nom de qualité qui suit directement le nom qualifié et il s'accorde avec lui en genre et en nombre. Il se distingue du complément de nom par sa position syntaxique<sup>17</sup>



### 4. Pronom (PR) :

On distingue les pronoms autonomes et les pronoms affixes. Le premier type qui concerne ce travail comprend en plus des pronoms personnels autonomes, les démonstratifs, les interrogatifs<sup>18</sup>, les possessifs et les pronoms régimes. Tandis que le deuxième type comporte les pronoms affixes aux présentatifs, aux quantifieurs et à la préposition.

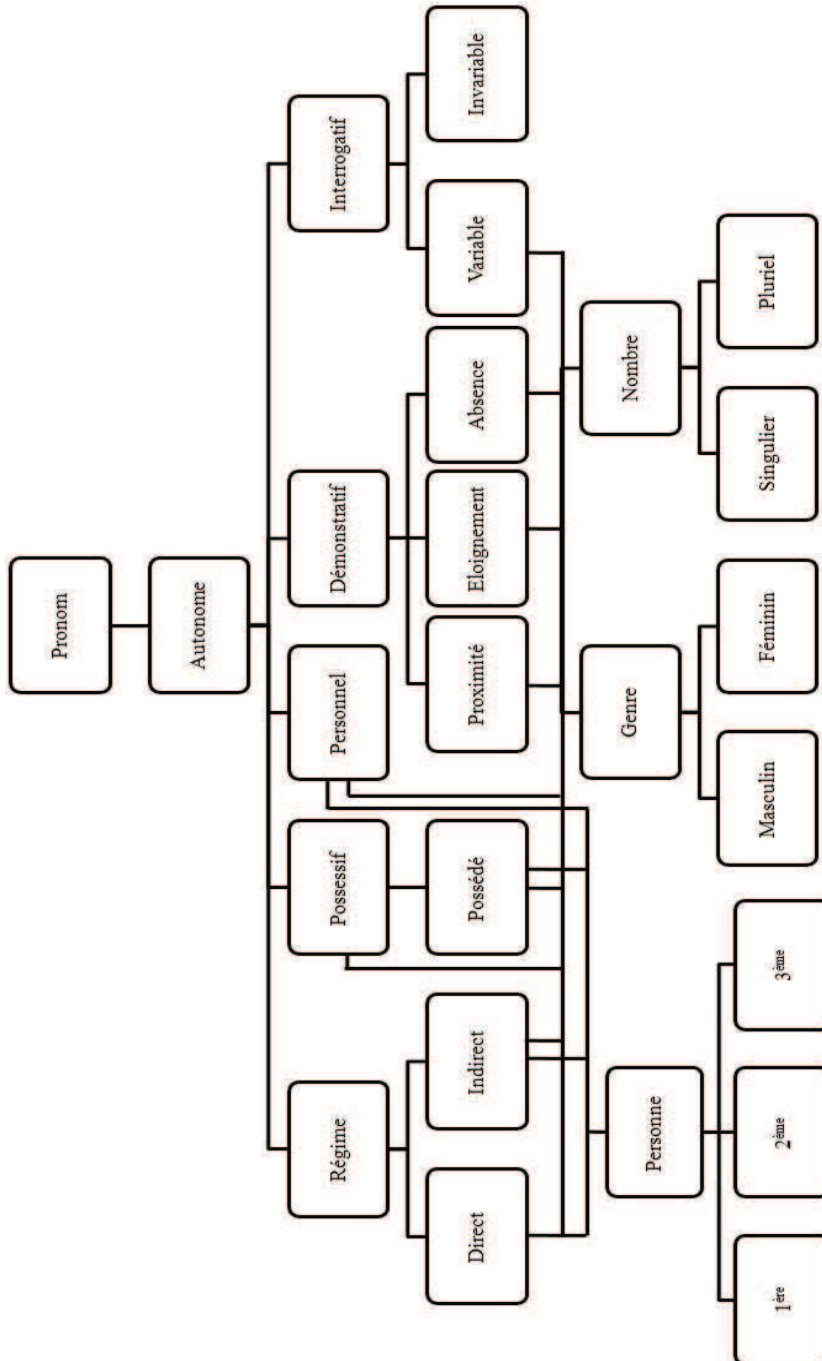
Partons du principe que toute unité lexicale est délimitée par deux blancs typographiques, les présentatifs, les quantifieurs et la préposition forment ainsi un seul mot graphique avec leur complément pronominal, ce qui explique l'exclusion des pronoms affixes du jeu des étiquettes recommandées.

Concernant les pronoms régimes (complément d'objet direct/indirect), beaucoup d'amazighisants les considèrent comme étant des « affixes ». Or, en nous basant, d'une part, sur le rôle syntaxique de chacun des deux pronoms (leur substitution à un nom ou à un syntagme prépositionnel, leur ordre dans l'énoncé verbal...), et d'autre part sur la tâche principale de l'étiquetage morphosyntaxique, nous les considérons comme « pronoms autonomes unifonctionnels » par opposition aux «

<sup>17</sup> Le complément du nom affiche l'état d'annexion qui résulte de la présence de la préposition I [n] « de » généralement supprimée pour les locuteurs dans leur langage quotidien. Il se distingue de l'adjectif par sa structure syntaxique étant donné que ce dernier est toujours juxtaposé au nom qu'il détermine, sans le biais d'aucun fonctionnel.

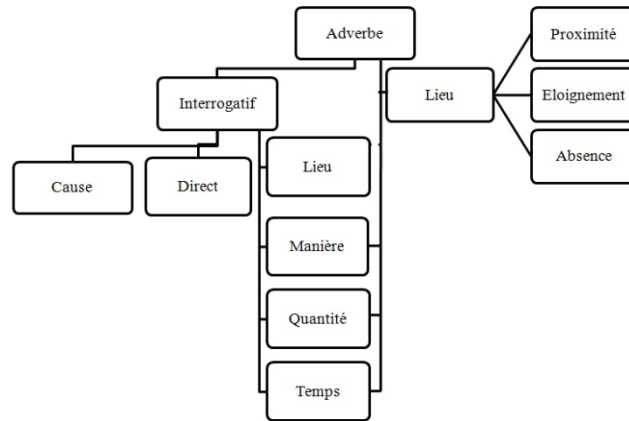
<sup>18</sup> En plus des pronoms interrogatifs associés aux pronoms déictiques (proximité, éloignement, absence) variables en genre et en nombre, il existe des pronoms interrogatifs invariables qui portent sur une partie de la phrase (ⵉⵎ [ma], ⵓ [u], ...). Voir (*La nouvelle grammaire de l'amazighe*).

pronoms autonomes plurifonctionnels », connus sous le nom « pronoms personnels autonomes ».



## Jeu d'étiquettes morphosyntaxiques de la langue amazighe

Généralement, l'adverbe est un mot invariable, sauf exception liée à des emplois régionaux. Il exprime le lieu, le temps, la qualité et la manière.

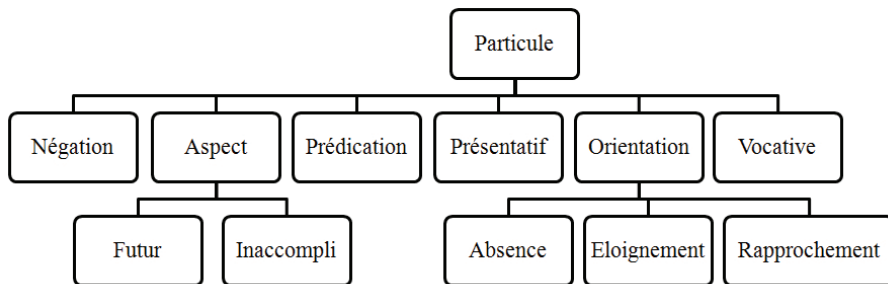


### 6. Préposition (AP) :

La préposition relève de la catégorie générale des mots de relation. Elle exprime des valeurs sémantiques diverses, notamment, la localisation spatio-temporelle, l'instrument, la direction, la possession, l'appartenance et l'accompagnement.

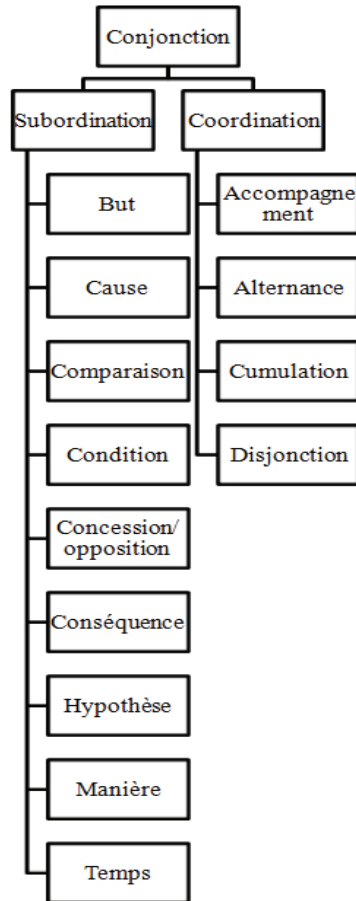
### 7. Particule (P) :

Les particules sont un ensemble de mots assez courts qui jouent le rôle d'indicateurs grammaticaux au sein d'une phrase.



8. Conjonction (C) :

La conjonction peut marquer la subordination ou la coordination.



9. Interjection (I) :

Il est question ici d'un mot invariable, autonome et n'ayant pas de fonction grammaticale. Il exprime une sensation, une émotion, un agacement, un étonnement ...

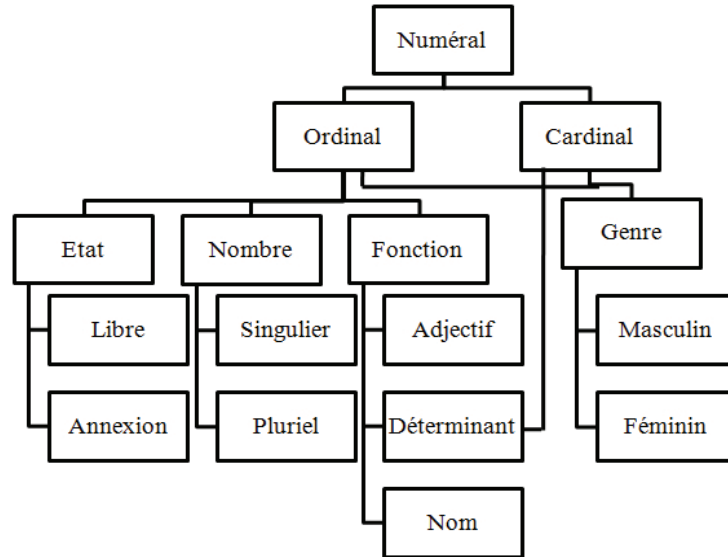
10. Résiduel (R) :

La valeur résiduelle est affectée aux termes qui ne rentrent pas dans les catégories usuelles. Par exemple: les mots étrangers, les formules mathématiques et les symboles.

Même si ces termes ne font pas partie du lexique de la langue traitée, ils se produisent assez fréquemment. Par conséquent, ils ont besoin d'être étiquetés.

11. Numéral (NU) :

La classe du numéral se compose de deux sous catégories : cardinal et ordinal. Cette classe peut se combiner avec les trois modalités démonstratives (déterminants) : de proximité, d'éloignement et d'absence.



12. Ponctuation (PU).

## 5. Conclusion

Dans la perspective de doter la langue amazighe d'outils nécessaires au traitement automatique, l'élaboration d'un jeu d'étiquettes est une étape préalable dans le processus d'automatisation. Ainsi, nous avons proposé deux jeux d'étiquettes grammaticales, dont le premier correspond à des étiquettes obligatoires et le deuxième à des étiquettes recommandées, où nous avons essayé de prendre en considération les particularités et la richesse grammaticales de l'ensemble des variantes amazighes marocaines.

## Références bibliographiques

Ameur, M. et al. (2004), *Initiation à la langue amazighe*, Publications de l'Institut Royal de la Culture Amazighe, Série : Manuels-N°1, Rabat.

Ataa Allah, F., Jaa, H. (2009), « Etiquetage morphosyntaxique : Outil d'assistance dédié à la langue amazighe ». Actes du 1er Symposium International sur le Traitement Automatique de la Culture Amazighe. 12-13 décembre 2009, Agadir, Maroc, p. 110-119.

Boukhris, F. et al. (2008), *La nouvelle grammaire de l'amazighe*, Publications de l'Institut Royal de la Culture Amazighe, Série : Manuels-N°2, Imprimerie El Maârif Al Jadida, Rabat.

EAG (1996). EAGLES, Recommendation for the Morphosyntactic Annotation of Corpora. EAGLES Document EAG-TCWG-MAC/R.  
<http://www.ilc.cnr.it/EAGLES96/home.html> [15.1.2007].

Laabdelaoui, R. et al. (2012), *Manuel de conjugaison amazighe*, Publications de l'Institut Royal de la Culture Amazighe, Série : Manuels-N°5, Rabat.

Loikkanen, S. (2007), « Étiquetage morpho-syntaxique de textes kabyles ». Actes de la conférence en Traitement Automatique des Langages Naturelles. 5-8 juin 2007, Toulouse, France. p. 193-202.

Martinet, A. (1979), *Grammaire fonctionnelle du français*, Paris, Didier-Crédif, XII + 276 p.

MUL (1996). Multilingual Text Tools and Corpora. <http://www.lpl.univ-aix.fr/projects/multext>.

Nguyen, T. M. H. et al. (2003), « Une étude de cas pour l'étiquetage morpho-syntaxique de textes vietnamiens », Actes de la conférence Traitement Automatique du Langage Naturel, 11-14 juin 2003, Batz-sur-Mer, France.

Outahajala, M. et al. (2010), « Tagging Amazigh with AnCoraPipe », Actes de l'atelier Language Resources and Human Language Technologies for Semitic Languages, 17 mai 2010, Valette, Malta. p. 52-56.

Paroubek P., Rajman M. (2000), « Étiquetage morpho-syntaxique », In J.-M. Pierrel, Ed., *Ingénierie des langues*, Paris : HERMES Science Europe, p. 131-150.

Valli A. et al. (1999), « Étiquetage grammatical des corpus de parole : problèmes et perspectives », *Revue française de linguistique appliquée*, Vol. 4, No. 2, p. 113-133.

### **Sitographie :**

[http://www.technolangue.net/article.php3?id\\_article=296](http://www.technolangue.net/article.php3?id_article=296).

<http://www.ircam.ma/doc/publica/nouvel-gram-amazigh.pdf>.

<http://www.ircam.ma/doc/publica/initiation-langue-amazighe-1.pdf>.

<http://www.ilc.cnr.it/EAGLES96/annotate/annotate.html>.

[http://www.lexilogos.com/vietnamien\\_dictionnaire.htm](http://www.lexilogos.com/vietnamien_dictionnaire.htm).