

Construction d'un dictionnaire de noms composés en arabe

Faten Khalfallah Hammouda*, Kais Haddar** and Abdelhamid Abdelwahed***

* FSEG, Sfax 3018, Tunisie. Email: Khalfallah_fat@yahoo.fr

** FSS, Sfax 3018, Tunisie. Email: Kais.haddar@fss.rnu.tn

***FLSH, Sfax 3018, Tunisie. Email: Abdelhamid.abdelwahed@yahoo.fr

Résumé—La construction d'un dictionnaire de noms composés arabes peut réduire la complexité du processus de compréhension des textes traités et faciliter la traduction automatique vers d'autres langues. Dans cet article, nous allons présenter un travail permettant de construire un dictionnaire de noms composés en arabe en se référant à une plateforme d'ingénierie linguistique baptisée NooJ. Pour ce faire, nous commençons par une étude sur les applications existantes traitant les mots composés. Ensuite, nous proposons une typologie des noms composés arabes en étudiant ses ambiguïtés. Puis, nous élaborons une démarche basée sur des transducteurs exprimés dans la plateforme NooJ. Cette démarche est expérimentée sur un corpus provenant du domaine de sport. Enfin, nous clôturons par une conclusion et des perspectives.

Abstract—The construction of an Arabic compound nouns dictionary can reduce the complexity of the text understanding and facilitate the automatic translation to other languages. In this paper, we'll take steps to construct a dictionary of compound nouns in Arabic referring to linguistic platform named NooJ. For that, we start first with study about the existing applications treating the compound words. Second, we'll propose a typology of Arabic compound nouns. After that, we'll elaborate a gait based on the transducers made with the platform NooJ. This gait is experimented on a corpus belonging to the domain of sports. Finally, we'll finish with a conclusion and some perspectives.

Mots clefs—Noms composés; grammaires locales; transducteurs; NooJ

Key words—Compound nouns; local grammars; transducers; NooJ

I. INTRODUCTION

L'automatisation de l'une des applications de traitement automatique de la langue naturelle (TALN) nécessite en général une étape de compréhension du texte ou document source par le système. Cette compréhension se fait généralement par la décomposition en unités (i.e., phrases, mots).

Vu la richesse de la langue arabe, le processus de compréhension de ses textes présente d'énormes difficultés. La principale difficulté consiste à repérer des unités de sens dans un flux d'informations. Pour cela, il s'avère très utile de faire construire une base lexicale (i.e., dictionnaire) contenant l'ensemble d'unités simples (i.e., mots composés) constituant un texte traité de la langue arabe.

Le présent travail se fait dans le cadre d'un projet de coopération Tuniso-Français entre le laboratoire MIRACL, l'unité de recherche LSCA et le laboratoire LAELDI. En effet, notre objectif principal est de construire un dictionnaire de noms composés arabes à partir d'un corpus homogène (i.e., domaine de sport). Les ensembles des règles, des contraintes et des déclencheurs, que nous proposons, sont représentés à l'aide des transducteurs à états finis sur la plateforme NooJ.

Dans cet article, nous commençons par présenter quelques travaux existants sur les noms composés. Ensuite, nous décrivons la typologie des noms composés arabe. Puis, nous proposons des transducteurs permettant de mettre en œuvre notre démarche. Enfin, nous clôturons cet article par une conclusion et des perspectives.

II. ETAT DE L'ART

La notion de mots composés, en général, et de noms composés, en particulier, varie d'un auteur à un autre. Elle fait l'objet de nombreuses discussions et reste jusqu'à maintenant ouverte à la conversation. Il existe des linguistes qui ont fait des analyses poussées des noms composés sans jamais définir cette notion. Par exemple, pour Bauer dans [6], la composition est «la formation d'un nouveau lexème par l'adjonction de deux lexèmes ou plus». Certains autres auteurs voient qu'un mot composé est une suite de mots simples qui sont séparés par un séparateur. En effet dans [7], on considère que toute séquence composée doit contenir au moins un séparateur. En fait, le critère référentiel de la composition est donné par [3] «il y a composition quand deux termes identifiables pour le locuteur se conjoignent en une unité nouvelle à signifiant unique et constant».

En effet, le manque de définition commune concernant la composition nominale crée un véritable problème dans le domaine de recherche. Pour recenser des mots composés à partir d'un texte, il existe trois approches : l'approche linguistique (i.e., [6]), l'approche statistique (i.e., [5]) et l'approche mixte (i.e., [3]). La première est la plus proche du langage naturel, parce qu'elle donne des résultats précis mais lourds à mettre en œuvre. La seconde se base sur des probabilités et des opérations arithmétiques comme le calcul de fréquence ou de nombre d'apparitions. Et la troisième utilise les deux approches précédentes ensemble. Elle est appelée aussi hybride.

Parmi les travaux réalisés concernant les mots composés et utilisant l'approche linguistique, on peut citer le travail réalisé dans [6]. Dans ce travail, l'auteur a proposé une méthode d'extraction des mots composés à partir de textes

français, anglais et portugais. Cette méthode a été expérimentée avec le système INTEX [6], l'ancienne version de NooJ [8].

III. TYPOLOGIE DES NOMS COMPOSÉS ARABE ET AMBIGUÏTES POUR LE DOMAINE DE SPORT

Rappelons que le mot arabe peut se présenter sous différentes formes. Il peut être soit un nom, soit un verbe ou une particule. Dans cette section, nous nous intéressons aux noms composés. D'abord, nous allons découvrir les différentes formes de noms composés en nous inspirant de [1]. Puis nous allons exposer les différentes ambiguïtés qu'on peut rencontrer pour extraire les noms composés arabes. Mais, avant ceci, nous devons expliciter qu'un nom composé est toute séquence de mots simples qu'on ne peut pas les séparer par n'importe quel autre mot.

A. Les différentes formes de noms composés pour le domaine du sport

En s'inspirant des travaux de [1] et [9], nous avons pu identifier plusieurs catégories grammaticales des noms composés arabes tels que le composé annexe (مركب إضافي), le composé référentiel (مركب إسنادي), le composé substitutif (مركب بدلي), le composé descriptif (مركب نعني) et la filiation (الكينية). Dans les sous sections suivantes, nous allons nous contenter de détailler essentiellement ces cinq catégories.

A.1. La catégorie de composé annexe

Le composé annexe «المركب الإضافي» représente un syntagme nominal SN composé de deux parties. Ce syntagme est composé d'un nom indéfini suivi d'un nom défini qui le complète. La première partie représente l'annexant «المضاف» et la deuxième représente l'annexé «المضاف إليه». L'annexant «المضاف» doit être toujours simple, il n'accepte pas la composition. Il ne permet ni l'ajout du déterminant «ال» ni le duel ni le pluriel. Tandis que l'annexé «المضاف إليه» accepte plusieurs formes, Il peut être simple ou complexe et permet le rattachement de l'article déterminant «ال». Par exemple, pour le cas de «كرة القدم football», il s'agit d'un cas très simple où le mot «كرة» représente le «المضاف» et le mot «القدم» représente le «المضاف إليه» sachant que «القدم» est un complément pour le mot «كرة». Le choix de l'exemple précédent comme un cas de nom composé est dû à la raison que sa traduction automatique vers d'autres langues présente un nom simple «football».

A.2. La catégorie de composé référentiel

Le composé référentiel «المركب الإسنادي» est composé d'un attribut «المسند» et d'un corrélatif d'attribut «المسند إليه». Prenons l'exemple de «السباق الثلاثي le triathlon» qui représente un cas simple d'un nom composé puisque «السباق الثلاثي» représente un synonyme pour le nom «تريثلون». Nous pouvons citer encore le nom propre «تأبط شرًا», il représente un cas particulier où le composé référentiel est une phrase verbale dont «تأبط» est un verbe et «شرًا» est un complément.

A.3. La catégorie de composés substitutifs

C'est un composé nominal de formes "بدل substitutif" et "مبدل منه substitué". Le rôle du premier est d'expliquer le deuxième. Au niveau de cette catégorie, on peut citer l'exemple suivant «المدرّب حسن أفنديش». Le mot «المدرّب» représente le substitutif et le reste représente son substitué «حسن أفنديش».

A.4. La catégorie de composés descriptifs

Le composé descriptif "المركب النعني" est composé de deux parties "descriptif "نعنت" et " décrit المنعوت ". Un adjectif "النعنت" peut se manifester sous différentes formes grammaticales, il peut être simple ou complexe. Il faut bien noter que " le descriptif "النعنت" s'accorde en genre, en nombre, en identification et en voyellation avec "le décrit المنعوت". Le composé descriptif se distingue du composé référentiel par le fait que dans une phrase le "composé descriptif المركب النعني" occupe les mêmes fonctions que ceux du nom, soit dans une phrase nominale ou verbale. Comme, il peut être composé des mêmes composants qu'un composé nominal (مركب اسمي) ou un composé prépositionnel (مركب حرفي). Au niveau de cette catégorie, on peut citer l'exemple de "الفقز من الثبات Saut sans élan".

A.5. La catégorie de filiation

La filiation "الكينية" est une marque de distinction appliquée à des personnages de premier plan pour les honorer. Elle est généralement associée à un nom de personne. Ainsi, les Arabes appellent \$ EiI- OHasani quelqu'un qui porte le nom de 'Ali, et Abo-Yosf celui qui a pour nom Ya'qūb. La Kunya est un élément composé dont le premier terme est souvent "أبو" Abu. Cette catégorie peut ne pas être prise en considération parce qu'il appartient à la catégorie des composés annexes.

B. Hiérarchie de noms composés pour le domaine du sport

Nous avons décomposé la liste des noms composés selon une hiérarchie. Cette hiérarchie est illustrée par le diagramme UML de la figure 1.

Au niveau de ce diagramme, nous explicitons qu'un nom composé peut être soit un nom de personne (joueur, arbitre, responsable, etc.), soit un nom de jeux (football, volleyball, basketball, etc.), soit un lieu (stade, complexe, centre, etc.), soit une organisation (fédération, équipe, etc.), soit un événement sportif (olympiade, etc.) ou tout autres termes sportives tels que les fonctions sportives (Président du club رئيس الجمعية), ainsi que leurs grades (Aide entraîneur مساعد المدرّب) ou les positions des joueurs (Attaquant قلب هجوم). Le nom d'un lieu (i.e., Stade 2 Mars (ملعب 2 مارس) peut être composé d'une date.

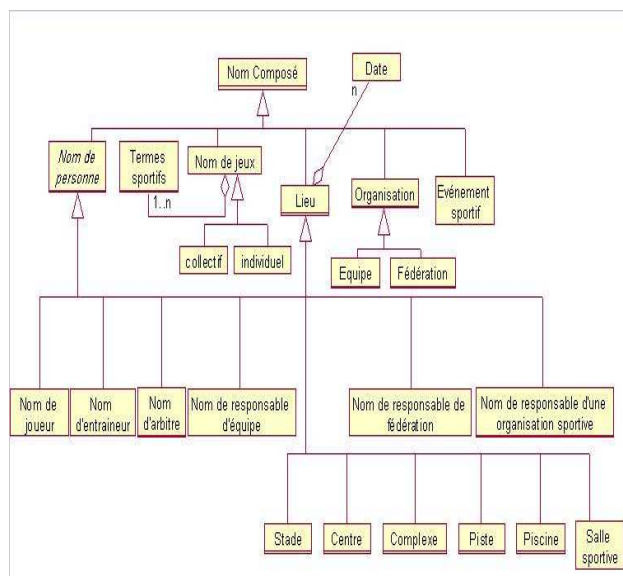


Figure 1. Diagramme de classe présentant l'hierarchie d'un nom composé arabe

L'hierarchie de la figure 1 va servir dans la construction des dictionnaires statiques dans la plateforme NooJ.

C. Ambiguïtés

Puisque les noms composés sont considérés comme référentiels, ils peuvent donc recevoir un type rendant compte de leur référent. Cette considération est le plus souvent erronée. En effet, on peut dégager plusieurs cas où la langue arabe représente des ambiguïtés de quantification (i.e., sémantiques) dans l'identification des noms composés à partir d'un corpus. Dans ce qui suit, nous allons détailler quelques cas sémantiques qui peuvent produire des problèmes lors de l'analyse sémantique.

C.1. La polysémie

C'est le fait d'avoir plus que deux mots qui s'écrivent de la même façon mais de sens différents. Cette situation est présente dans différents cas:

- Noms propres composés se confondent avec les noms de régions

1) ملعب فرحات حشاد *Avenue Farhat Hached*
 2) كان فرحات حشاد نقابياً *Farhat Hached était un syndicat*

Ainsi, le mot composé فرحات حشاد représente en même temps un nom de stade dans l'exemple (1) et un nom de personne dans l'exemple (2). Pour cela, le mot « ملعب » dans l'exemple (1) doit déclencher le fait que « فرحات حشاد » est un nom de stade et n'est pas une personne. Donc, il joue le rôle d'un déclencheur.

- Noms de personnes composés se confondent avec des noms d'animaux

3) مكرم الجرو هو لاعب دولي في كرة اليد التونسية *Makram Eljarou est un joueur international du handball tunisien*
 4) أعطى أبو مكرم الجرو طعاماً

Le père de Makram a donné à son petit chien à manger.

Dans l'exemple (3), le mot composé الجرو مكرم représente un nom et prénom d'une personne célèbre. Mais, elle représente un prénom suivi d'un nom d'animal (4). Le composé « هو لاعب » dans l'exemple (3) montre que « الجرو مكرم » est un nom de joueur. Donc, il représente un signe de fin de phrase et montre que ce composé est un nom de joueur.

- Noms de jeux composés se confondent avec des syntagmes nominaux

5) كرة السلة رياضة العملاقة

Le basketball est un sport des géants.

6) ضربت كرة السلة فقلبت ما بها على الأرض

Le ballon a bousculé le panier par terre

Dans l'exemple (5), le mot composé « كرة السلة » représente en même temps un nom de jeux sportif. Mais, elle représente deux mots totalement indépendants dans l'exemple (6). Le mot « رياضة » dans l'exemple (5) représente un justificatif que ce composé est un nom de jeu sportif.

C.2. La synonymie

C'est le fait d'avoir plus que deux mots qui s'écrivent de différentes façons ayant le même sens. Cette situation est présente dans différents cas. Soit l'exemple ci après.

7) تحصل نسور قرطاج على كأس إفريقيا للأمم

8) تحصل المنتخب التونسي على كأس إفريقيا للأمم

L'équipe Nationale tunisienne a obtenu la coupe d'Afrique

Les exemples (7) et (8) signifient la même chose. Le syntagme *نصور قرطاج* est synonyme au syntagme *المنتخب التونسي*.

Ainsi, on peut prendre l'exemple ci-dessous :

9) لعب أسود الترينقا مباراة قوية ضد أسود الأطلس

10) لعب الفريق الكامروني مباراة قوية ضد الفريق المغربي

L'équipe du Cameroun a joué contre l'équipe Marocaine

Les exemples (9) et (10) signifient la même chose. Le syntagme "أسود الترينقا" est synonyme au syntagme "الفريق الكامروني" et le syntagme "أسود الأطلس" est synonyme au composé "الفريق المغربي". Mais, ici, le verbe utilisé لعب appartient au champ lexical du domaine sportif.

En conclusion, toutes ces ambiguïtés nécessitent la présentation d'une solution permettant de réduire ou même résoudre les erreurs qu'ils peuvent s'entraîner. Pour cela, nous avons recours à utiliser un corpus appartenant seulement au domaine sportif, en plus que nous pouvons étudier le comportement des verbes (transitivité) et/ou utiliser les textes partiellement ou totalement voyellés.

IV. DEMARCHE PROPOSEE

Dans ce travail, nous proposons une démarche permettant de construire automatiquement un dictionnaire de noms composés arabes à partir d'un corpus en passant par différentes étapes.

- *Etape 1 : Constitution d'un corpus pour le domaine de sport*

Nous avons collecté un corpus formés de 100 textes contenant presque 20 000 mots en entrée extrait à partir d'Internet, des journaux et des magazines. Ce corpus contient tous types de mots (i.e., nom, verbe, adjectif, statut, préposition, etc.). Le fait d'avoir un très grand corpus nous permet d'avoir le maximum des cas possibles de noms composés. Notons que le corpus se repose sur des textes appartenant au domaine sport. L'analyse de ce corpus sera basée sur la plateforme NooJ (environnement linguistique) construite par Max Silberztein [6].

- *Etape 2 : Construction des dictionnaires utiles*

Nous avons construit, en utilisant NooJ et selon l'hierarchie élaborée, 20 dictionnaires de noms composés arabes. Chaque dictionnaire va contenir une catégorie de noms composés (nom de joueur, nom d'arbitre, nom de stade, nom d'organisation, etc.) qui seront sélectionnés par nécessité dans la phase d'analyse.

- *Etape 3: Collection d'une liste de déclencheurs*

Nous constituons la liste de déclencheurs qui permet de déclencher la présence d'un nom composé qui vient juste après, appartenant au domaine du sport tel que le cas de (Stade ملعب), vu que la présence de « ملعب » nécessite la présence de son nom juste après. Pour cela elle est considérée comme déclencheur.

- *Etape 4: Collection de mots représentant le début de noms composés*

Nous ajoutons une autre liste de mots représentant le début de noms composés appartenant au domaine du sport tel que le cas de (Course سباق) pouvant être présent dans les différents cas suivants (Equitation سباق الخيل, Cyclisme سباق الدراجات).

- *Etape 5: Collection de mots représentant la fin de noms composés*

Nous ajoutons une liste contenant les mots représentant la fin de noms composés du domaine du sport tel que le cas de (gauche أيسر) dans les cas suivant (Elier gauche مدافع أيسر, جناح أيسر).

- *Etape 6: Génération des dictionnaires de flexions*

Nous générons les différentes formes fléchies du lexique dans des dictionnaires de flexions pour traiter toutes les formes de chaque nom dépendant du domaine du sport.

- *Etape 7: Construction des transducteurs*

Nous passons, au niveau de cette étape, à élaborer les transducteurs NooJ à partir d'une liste de grammaires simples qui va expliciter toutes les règles de composition nominale en se basant sur la typologie étudiée des noms composés arabes. Cet ensemble de transducteurs (voir VI) joue le rôle d'analyseurs. Parmi eux, on trouve ceux qui présentent des analyseurs lexicaux et d'autres présentant des analyseurs syntaxiques. Cette étape est illustrée dans la figure 2.

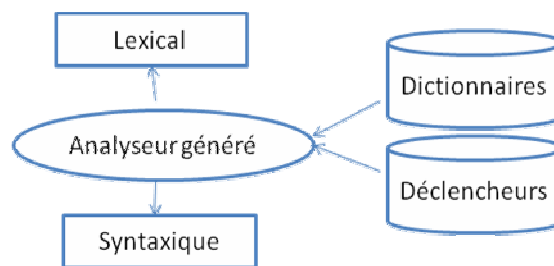


Figure 2. Processus de génération d'analyseurs

Enfin, nous passons notre corpus, sous format NooJ, déjà construit sur ces transducteurs pour avoir la liste finale des noms composés qui va construire à son tour le dictionnaire en question.

La grammaire établie se présente comme suit:

NOM_COMP → nom indef SN | SN | nom_def
NOM_PR → nom_def FIN_MOT | DEC SN

SN → nom_def (adj_def)*

NOM_PR → DEC1 prénom | DEC prénom | prénom

DEC → ملعب | كرة | قاعة | نادي | إستاد | مركز

DEC1 → المدرب | الرئيس | الحكم

FIN_MOT → الارتقاء | الأولمبية | الإرسال

Cette grammaire va servir à faciliter la construction des grammaires (transducteurs) dans la plateforme linguistique baptisée NooJ puisque nous allons étudier toutes les règles possibles pour la composition nominale.

La grammaire ci après permet de présenter un nom de filiation arabe.

NOM_FIL → DEB_MOT mot | (prénom)*

DEB_MOT → ابو | ابن | عبد | أبو | بن

V. EXPERIMENTATION ET EVALUATION

Après avoir collecté le corpus et généré les analyseurs à partir des dictionnaires construits, il fallait passer à l'expérimentation. Cette étape s'avère la plus importante dans notre démarche. Elle permet de mesurer la fiabilité du travail établi.

Le graphe NooJ ci-dessous de la figure 3 représente une description d'un transducteur qui permet d'extraire les noms composés appartenant à la catégorie de filiation en utilisant la grammaire précédente.

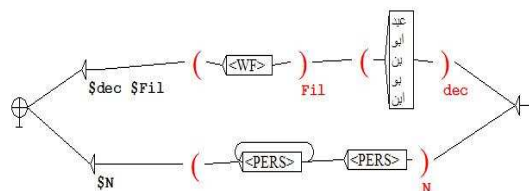


Figure 3. Contraintes morphologiques

Dans la figure 3, nous avons présenté un transducteur qui permet d'extraire les noms de personnes composés ainsi que les filiations telles que « عبد الله » و « bin mftaah بن مفتاح ». Les sorties de ce transducteur peuvent être des noms de joueurs, d'entraîneurs, d'arbitres ou d'autres responsables. Aussi, le transducteur ci-dessous permet de recenser les noms de personnes (joueurs, arbitres, etc.).

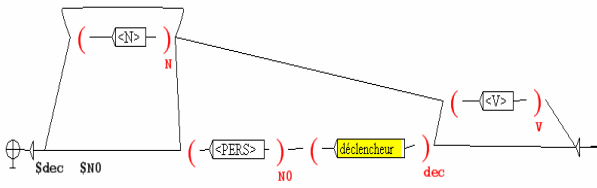


Figure 4. Extraction des noms propres

Dans la figure 4, la case colorée en jaune nommée déclencheur contient l'ensemble des déclencheurs qui peuvent être trouvés avant les noms propres. Comme son nom l'indique, le déclencheur contient l'ensemble de mots qui permettent de déclencher une catégorie bien définie à l'avance (ici les noms propres).

Le transducteur ci après représente l'extraction des noms de stades en les accordant l'annotation <N+lieu> puisqu'ils représentent un endroit.

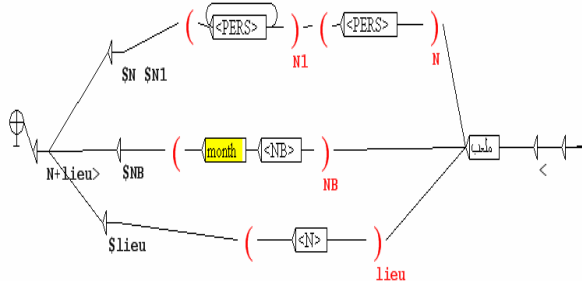


Figure 5. Extraction des noms de stades

La figure 5 permet d'extraire les noms de lieu (stades), il peut par exemple recenser les composés « الطيب المهيري » ou « ملعب 2 مارس » s'ils sont précédés par le mot « ملعب ».

Les critères d'évaluation utilisés sont le rappel, la précision et la F-mesure. Ces derniers peuvent être mesurés par les règles ci-après [5].

$$R = \frac{\text{Nombre de réponses pertinentes du système}}{\text{Nombre de réponses idéales}} \quad (1)$$

$$P = \frac{\text{Nombre de réponses pertinentes du système}}{\text{Nombre de réponses fournies par le système}} \quad (2)$$

$$F\text{-mesure} = \frac{P}{R} \quad (3)$$

En appliquant ces règles sur la liste corpus collecté, nous aboutissons les résultats ci- après.

Nombre de corpus	1 (domaine sport)
Nombre de textes	100 textes
Nombre de mots	20 000
R = 11 000/15 100	0.72 (72%)
P = 11 000/ 16 000	0.68 (68%)
F-Mesure	0.94 (94%)

Pour analyser le corpus établi en se référant à la plateforme NooJ, il suffit de cocher les dictionnaires en question au niveau des préférences et de cliquer sur l'option « linguistic analysis » du menu corpus. Ceci est explicité par la figure 6.

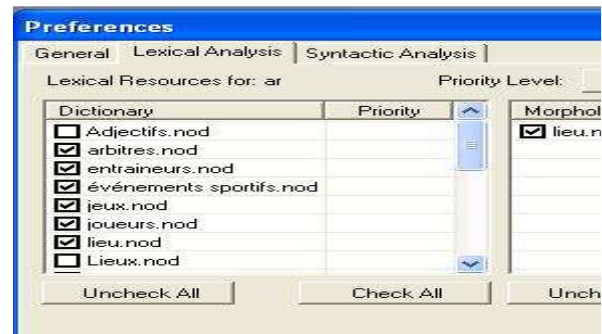


Figure 6. Choix de dictionnaires à utiliser

D'après la figure 6, Nous pouvons distinguer que les dictionnaires contenant les adjectifs et les lieux ne sont pas appelés au niveau de l'analyse.

Nous pouvons donc ouvrir un texte du corpus pour l'appliquer. Après avoir choisi la liste des dictionnaires à utiliser, il suffit d'effectuer une analyse linguistique par l'outil de NooJ « linguistic analysis ». Puis, effectuer une analyse lexicale ou syntaxique en se référant aux transducteurs élaborés. La figure 7 montre l'exemple du texte choisi qui a pris le format NooJ (.not).



Figure 7. Texte extrait du corpus établis pour l'évaluation

La figure 7 représente un texte extrait du corpus élaboré. Le texte précédent résulte la liste des noms composés dans la figure 8.

Text	After	Seq.	Before
	ملعب 20 مارس ملعب 2 مارس ملعب	<N+lieu> راس	ملعب 20 مارس
	ملعب 20 مارس <N+lieu> 20 مارس	<N+lieu> 20 مارس	ملعب 20 مارس
	ملعب فرحات حشاد ملعب 31 مارس	<N+lieu> 20 مارس	ملعب 20 مارس
	ملعب فرحات حشاد <N+lieu> 31 مارس مدرسة الند	<N+lieu> 20 مارس	ملعب 20 مارس

Figure 8. Affichage de noms composés extraits

Dans la figure 8, nous trouvons les résultats obtenus après le passage du texte à travers les transducteurs sachant qu'il faut cocher la case Outputs. En conclusion,

nous pouvons dire que l'extraction des noms composés à partir d'un corpus s'avère une tâche très difficile puisque les règles utilisées peuvent engendrer des ambiguïtés qui ne sont pas totalement résolues.

Notons bien que cette application peut classer le domaine du texte, s'il appartient au domaine sportif ou non, après le résultat obtenu.

VI. CONCLUSION ET PERSPECTIVES

Dans cet article, nous avons présenté la typologie des noms composés en étudiant ses ambiguïtés dans le domaine de sport. Puis, nous avons décrit la démarche proposée pour la construction d'un dictionnaire de noms composés arabes basée sur des transducteurs exprimés dans NOOJ.

Comme perspectives, nous sommes entrain de compléter l'ensemble de transducteurs nécessaires à la construction de dictionnaire de noms composés et de résoudre certaines ambiguïtés sémantiques.

REFERENCES

- [1] A. Addahdah: معجم قواعد اللغة العربية في جداول و لوحات , librairie Liban, Bairout (1992)
- [2] F. Khalfallah, K. Haddar, "Construction d'un dictionnaire de noms composés en arabe," *GEI'08, 17 -19 Mars, Sousse, Tunisie*.
- [3] E. Benveniste: Fondements syntaxiques de la composition nominale. et Formes nouvelles de la composition nominale. Problèmes de linguistique générale, 2, Gallimard, Paris (1974) 145-176
- [4] Mesfar, S.: "Standard Arabic formalization and linguistic platform for its analysis", *the challenge of arabic for nlp/mt, laseldi, Franche-Comté University, France*
- [5] T. Poibeau : Evaluation des systèmes d'extraction d'informations : une expérience sur le français, langues (spécial sur l'évaluation), vol (2) (1999)110-118.
- [6] A. Savary : Recensement et description des mots composés-méthodes et applications, thèse de doctorat en Informatique Fondamentale, Paris (2000) pp.10.
- [7] M. Silberztein : Le dictionnaire électronique des mots composés, dans : Dictionnaires électronique du français, Editions Blandine Courtois et Max Silberztein, langue française, n° 87, Paris : Larousse, LADL (1990) 71-84
- [8] M. Silberztein, et al.: "NooJ, un outil TAL pour l'enseignement des langues". Article mis en ligne (2005) 123-134.
- [9] S. Kabbabi, A. Bennour, S. Missaoui et M. Chaouch, *Manuel scolaire de langue de 8^{ème} année de base*, Ed 2006, pp 31-92