المعالجة الآلية للوحدات الإسمية:
الأعداد المركبة في اللغة العربية نموذجاً

عزالدين غازي
جامعة القاضي عياض- مراكش-المغرب

ملخص: تندرج هذه المقالة في إطار المعالجة الآلية للمداخل المعجمية، مركزة على الوحدات الإسمية Entités Nommés وخاصة منها الأعداد المركبة، كما تأتي أيضاً في إطار بناء المعجم الآلي للغة العربية، باستغلال التقنيات المعلوماتية في سائر المستويات ومنها المعجم. وذلك بوضع قاعدة ترتكز على وصف المداخل المعجمية العددية. وبناء عليه ستتطرق مداخلتي أولا: إلى وصف وتصنيف هذه الوحدات في اطار قاعدة بيانات معجمية موسعة، وثانياً: ستعتمد تطبيق تقنيات التعرف الآلي المتاحة في منصة **نوج** (NooJ) للتطوير اللغوي والمتمثلة في خوارزميات التعرف المعتمدة على أوتومات الحالات النهائية (automates à étas finis) والمحولات (transducteurs) والرسوم (graphes)، والمنطلقة من خصوصية النحويات المحلية لهذه المتواليات وما يتميز به نظام العربية الإعرابي والصرفي والعلاقات التركيبية- الدلالية. مما سيجيب عن كيفية تقييس الكفاية الطبيعية حاسوبياً، وعن الأدوات المعلوماتية المناسبة لتمثيل ومعالجة هذه المداخل. ويأتي اختيار البيئة اللغوية المجانية المفتوحة المصدر **نوج** بناء على تمكنها الأمثل من الأدوات العقلانية ذات القدرة والكفاءة العالية في معالجة البنيات الإسمية المذكورة خاصة الأعداد المركبة كما سنرى عبر مجموعة من الرسوم الممثلة لذلك.

الكلمات المفتاحية: الأعداد المركبة- التعرف الآلي- أوتومات الحالات النهائية- المحولات- الرسوم- بيئة نوج للتطوير اللغوي.

تقديم :

تستند هذه الدراسة التي ستهدف التعرف الآلي على الأعداد المركبة، على نظرية المعجم- التركيبي التي استفادت عُمْقاً وتصوراً من نماذج فرضتها أساليب الوصف العلمي في مجالات متعددة وعلى رأسها الهندسة المعلوماتية التي استعارت اللغة العقلانية الرصينة كالتوليد والمعالجة والتحليل والأتمتة بأنواعها. حيث بينت بجلاء الخصوصية التركيبية والدلالية لهذا النوع من الأسماء، الذي يفترض بناء نماذج صورية كعمليات أولية في المعالجة النصية والمتعلقة أساساً بكيفية بناء برنامج آلي يمكننا من التعرف على البنيات اللسانية العربية المركبة والبسيطة؟ وشروط التعامل مع هذه البنيات والوحدات المركبة؟

عند كل وصف وتفسير للظواهر اللغوية المعالجة في نظام اللغة العربية، تطفح على السطح مشاكل وصعوبات تتعلق بكيفية التعرف الآلي على المتواليات([1]). لذلك

([1]) تطرح المتواليات المركبة على سبيل المثال لا الحصر إشكالية كبيرة في معالجة الغموض فيها، إذ على الرغم من وجود أنظمة آلية مثل ما ابتكره شينك (A.Schenk 1986) في اقتراحه نظاماً للترجمة الآلية لهذه البنيات يقتضي وجود بنيات مماثلة في اللغة الهدف. إلا أن هذه المنهجية لم تتوصل إلى حل لمعالجة آلية مستقلة في نظام مستقل كما

طرحت نظرية المعجم-التركيبي المبدأ العقلاني الصوري لمعالجة البنيات اللسانية التي تتسم بالغموض، وكانت الانظمة المخصصة لذلك مجهزة تماماً بهذه العُدة كالذي بنيت به منصة (NooJ)([2]) والتي استغلت في معالجة معاجم لغات عديدة، كما تساعد هذه البيئة كذلك على بناء نماذج صورية وخاصة مع تقنية أوتومات الحالات النهائية كمبدأ عقلاني لمعالجة الوحدات الاسمية المركبة بهدف صورنتها وإدخالها في خوارزميات التعرف الآلي للتحليل والتوليد والتي نسعى إليها من خلال بناء نظام آلي متمكن من معالجتها وإنتاجها كما عند المتكلم الطبيعي. وذلك بتطبيق تقنيات التعرف على نماذج من الأعداد المركبة في العربية، بداية بمرحلة الوسم المعجمي في النص المراد معالجته وكيفية التعرف على المنطقة المعتمة للبنية والمتوالية المترابطة وغيرالمترابطة مرورا بوخوارزمات التعرف الآلي وتقنيات تمثيل البيانات ثم رصد واستخراج الاعداد في نظام اللغة العربية.

1- تعريف وتصنيف الوحدات الاسمية المركبة:

اعتماداً على مقولة بايلون وفايير حول خصوصية الوجود المستقل والموحد للمفردة المركبة،واعتباراً للطبيعة الدلالية والتركيبية للمتواليات. فإن أغلب التعاريف ([3])أجمعت على"أن المفردة المركبة تعكس الصورة الموحدة"([4]) في الواقع الإحالي للغة، لكن هذا قد يحدث خلطاً لبعض الوحدات المركبة التي يمكن اختزالها في المفردة البسيطة انطلاقاً من المحتوى الدلالي للمفردة نحو: جواز (جواز السفر / جواز المرور). مع ذلك فإن هذا التعريف غير مناسب للوحدات المركبة التي تتميز، في أغلب اللغات، بكونها تربط بين الشكل والمعنى([5]). وإذا كان هذا التعريف يستند إلى المستوى الدلالي فإنه يتقيد بالمستوى التركيبي. ذلك أن ما قد يكون في لغة ما مفردة بسيطة يكون في لغة أخرى مفردة مركبة مثلا: مفردة (gare) في الفرنسية تقابلها المتلازمة المركبة في اللغتين الألمانية (bahnhof) والإنجليزية (railway station) على التوالي. هذا يعني أن ليس هناك علاقة مباشرة بين معنى الوحدة المعجمية وبين مجالها الاستبدالي سواء كان بسيطاً أو مركبا([6]ً). فالدلالة،إذن، لا يمكنها لوحدها الإحاطة بخاصية الوحدات الاسمية المركبة، بل من الضروري استدعاء المستوى التركيبي([7])الذي يؤكد البعد

فعل إيريك لا بورت (للمزيد ينظر E.Laporte,la reconnaissance des expressions figées lors de l'analyse automatique, in **langage** n°90 1988.pp 117-126.).

([2]) بيئة مجانية مفتوحة المصدر في الشابكة قابلة للتحميل للاطلاع ينظر الموقع التالي:(2011) www.nooj4nlp.net

([3]) لقد شكل موضوع تعريف الوحدات الاسمية في اللغات الاجنبية منذ 1874 مصدر اهتمام العديد من النحاة واللسانيين نذكر منهم (م.بريال (1897) و شارل بالي (1950) و ف.برينو وشارل برينو (1949) وماروزو (1961) و م.كريفيس (1969) للمزيد انظر G.Gross,degrès de figement des noms composés,in **langages n°90** .p.59.) .

(4) G.Gross,degrès de figement des noms composés,in **langages n°90** .p.59.

([5]) نفس المرجع والصفحة.

([6]) نفس المرجع والصفحة.

([7]) أكد أندري مارتيني على أنه لا يمكن تحديد خاصية مكونات المفردة المركبة منفردة، مما يجعل أمر هذا التحديد في النظر إلى مجموع مكونات المؤلفة للاسم المركب، حيث لا يمكن أن نستبدل أحد عناصر المركب التالي (la tête de l' Afrique *)،(ارتفاع الحرارة (*علو الحرارة)، وقد ذهب جلبير (Guilbert 1975) إلى اختزال الجمل المركبة في الأفعال العماد (support) ولكن كاستون كروس ينفي ذلك لأن هذه الأفعال (المساعدة) لا تظهر في حالة التأسيم (nomalisation). ب(Corne de l' afrique

التأليفي لبعضها، ومع ذلك لا يمكنك إدراج هذا المستوى في تحليل جميع الحالات. وإذا كانت الوحدات الاسمية تلاحق جميع اللغات كما قال أ.مكاي (A.Makkai 1978)([8]) فإن جميع الوحدات المركبة ستعالج من باب كونها تتميز بدرجة ما من التحجر والتضام والتكلس، مما يجعلنا نهتدي إلى تحديد أصنافها وأنماطها من خلال مدونة تشمل جميع الوحدات الاسمية في النصوص العربية قديمها وحديثها، وهذا ما حاولنا القيام به بالتحديد من خلال بناء نموذج الوحدات الاسمية ([9]).

وللإحاطة الشاملة لابد من اعتماد مفهوم التصنيف لتسهيل عملية ترتيبها في أشكال صورية متجانسة وفي صنيفات(sub-categories) تجمعها وحدة الخصائص الموفو-تركيبية بعد إعداد قاعدة بيانات خاصة بها وفق جداول تصنيفية([10])، بتسجيل كل العناصر المركبة وتصنيفها تصنيفاً متجانساً كما تقول ليليا بيكابيا([11])،بحيث يقوم التصنيف المعتمد هنا على التآليف
المستعملة بين مختلف أجزاء الوحدات في جداول المعجم- التركيبي،مما يجعلنا أمام حالتين من التجميع المركباتي:

أ- فحص طبيعة متواليات الوحدات الإسمية التي تتطلب إجراءً تحليلياً خاصاً بها.
ب- فحص أساليب الضم والتجميع لجزء إلى آخر يماثله ويتوزع معه، وهذا ما يدعونا إلى الجمع بين نوعين من الوصف: الاهتمام أولا بالوضع المورفولوجي للعناصر التي تتألف معها الوحدات الإسمية، ثم ثانياً وصف وضعيتها الجديدة التي أنشأتها في نظام اللغة العربية([12]).

1-2- أنواع الوحدات الإسمية في اللغة العربية:

نميز في تصنيف هذه الوحدات بين ثلاث أشكال رئيسية هي حسب (Poibeau 2005([13])): أسماء الأعلام والأعداد والتواريخ التي يمكن إجمالها في الشكل التالي:

([8]) المرجع السابق، ص.61.

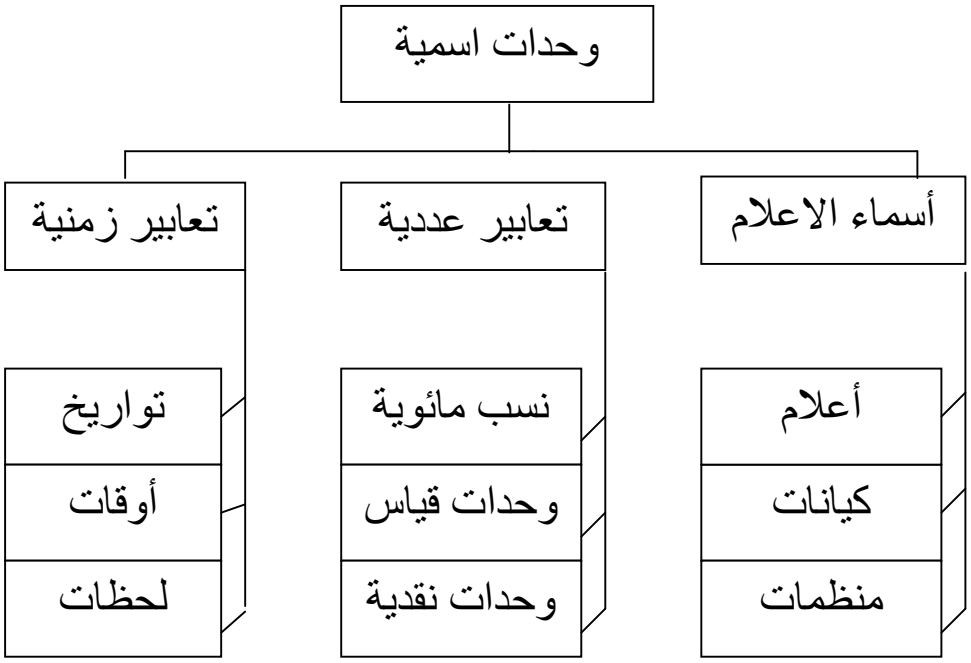
([9]) لقد أعددنا نموذجاً لقاعدة بيانات الوحدات المركبة في اللغة العربية بعد جمعها من مصادر مختلفة، بتصنيفها إما مباشرة من المعاجم وإما باستخراجها، بشكل غير مباشر، من أمهات المعاجم الحديثة المحوسبة والنصوص العربية المتداولة. هذا فضلا عن ورودها في غالب الاحيان في كتب النحو وكتب الطب والنبات والأدوية وعلوم المنطق والفلك والاصول، كما ان الألفاظ المركبة في اللغة العربية المعاصرة حاضرة بشكل كبير في لغة المعلوميات والإعلام والصحافة بأنواعها ولغة المال والأعمال والإقتصاد والسياحة والموضة... ولغة التخاطب اليومي الذي يعج بمثل هذه المفردات لكن لايتسع المقام لذكرها (للمزيد ينظر، عزالدين غازي، **بناء المعجم الآلي للغة العربية: معالجة التعبيرات المتلازمة**، أطروحة الدكتوراه، كلية الاداب عين الشق، الدار البيضاء (2009)) وسنكتفي هنا بايراد بعض الأمثلة للتوضيح في الملحق 1.

([10]) لقد قمنا ببناء جداول تصنيفية لقاعدة بيانات المفردات المركبة في اللغة العربية اعتماداً على تصور نظرية المعجم –التركيبي التي نتبعها في هذا البحث.وتتضمن هذه الجداول قائمة من المركبات (الأسطر العمودية) الأصناف التركيبية – المعجمية (الاسطر الفقية) تحدد العلاقة بينهما قيم الرفض والقبول (+،-) (انظر الملحق 2).

([11])عبد الرحمان يجيوي،**الأسماء المركبة في اللغة العربية، دراسة مورفولوجية –تركيبية**، رسالة دبلوم الدراسات العليا، كلية الآداب والعلوم الانسانية بفاس 1 ، سنة 1994.

([12])محمد الحناش، المعاجم الآلية للغة العربية: بناء قاعدة المعطيات، **التواصل اللساني**، المجلد. 4، ع. 1 /1992.

([13]) Mesfar Slim, **Analyse morpho-syntaxique automatique et reconnaissance des entités nommées en arabe standard**, thèse doctorale de l'université de Franche Compté, France 2008,p 150



شكل 1 يبين اصناف الوحدات الاسمية

2- إشكالية التعرف الألي على الوحدات الاسمية، صنف الأعداد المركبة :

يمكن أن نجمل حاصل الأصناف الكبرى للاعداد المركبة([14]) في نسق اللغة العربية صورياً كما في الجدول التالي (انظر الملحق 1):

- الأعداد من : صفر(0) إلى تسعة (9)
- الأعداد من: عشرة (10) إلى تسعة عشر (19)
- الأعداد من: عشرون (20) إلى تسعة وتسعون (99)
- الأعداد من: مائة وواحد (101) الى تسعة مائة وتسعة وتسعون (999)
- الأعداد من ألف وواحد (1001) إلى تسعة آلاف وتسعة مائة وتسعة وتسعون مليار (000 000 000 9 999)

([14]) أحصى جورج كروس حوالي ستة وعشرين صنفاً من أصناف المفردات المركبة في اللغة الفرنسية وكان هدفه هو تحديد خاصية التعبيرات المركبة انطلاقاً من العلاقات التي تربط بين عناصرها المؤلفة لها.(انظر المرجع السابق.صص. 67-68)

هذه الأصناف تامة فكما يقول موريس كروس (M.Gross) إن أقل درجة من التحجر والتضام لهذه الوحدات الإسمية يكفي لتدرج ضمن المتلازمات والمسكوكات.([15]) ولأن درجة التحجر هذه هي الكفيلة بتحديد الخاصية اللاتأليفية للوحدات المركبة.

أما صنف الأعداد المركبة من أحد عشر وخمسة عشر وإثنا عشر([16]) فهي تقوم بوظيفة إبلاغية لتفيد المخاطب عن المقداروالكم والقياس والترميز والترقيم. ولهذا فإن العنصرين بالعدد المركب يكونان متلازمان في دلالتهما الإحالية لأنهما يحيلان على شيء واحد في الواقع، أو على الأقل من نفس النوع والفصيلة([17]).ومن وجهة نظر الأسماء المعدودة في اللغة فقد لاحظ المشتغلون بأدبيات الجهة وجود تناظر مشابه بين المحمولات الدالة على الحالة وتلك الدالة على السيرورة و أسماء الكتلة، وقد لاحظ ليتش (1969) ([18]) أن معاني الأفعال والأسماء جميعها تتضمن العدد، لأن الأحداث والأشياء معدودة في الوجود.

هذا وتعتبر مسالة التعرف الآلي على الوحدات الإسمية في النصوص من الظواهر اللسانية المعقدة بصفة عامة، لكونها لا يمكن أن تعالج وتحلل معحمياً، أما وصفها في المعجم الآلي فيحتاج إلى بناء يدوي لأتومانات الأوضاع النهائية بناء على الأنحاء المحلية لهذه البنيات. وهذا يستوجب بناء قاعدة بيانات لجميع الأعداد المركبة في اللغة العربية وهي : من واحد إلى تسعة آلاف وتسعة مائة وتسعة وتسعون مليار (000 000 000 9 999) في معجم آلي معين. وما نبتغيه ها هنا هو تطبيق مقاربة الأنحاء الالكترونية المستمدة من النحو المحلي على شكل رسوم لاختبار أدائها المعجمي والتركيبي في النصوص. وللوهلة الأولى يمكن اختبار ذلك بالنظر في المحددات العددية ومدى تمثيلها لقيمتها العددية في الواقع الإحالي، مثل التعبير العددي التالي:

مائتان وأربع وعشرون

التي تمثلها القيمة الرقمية التالية:

(224)

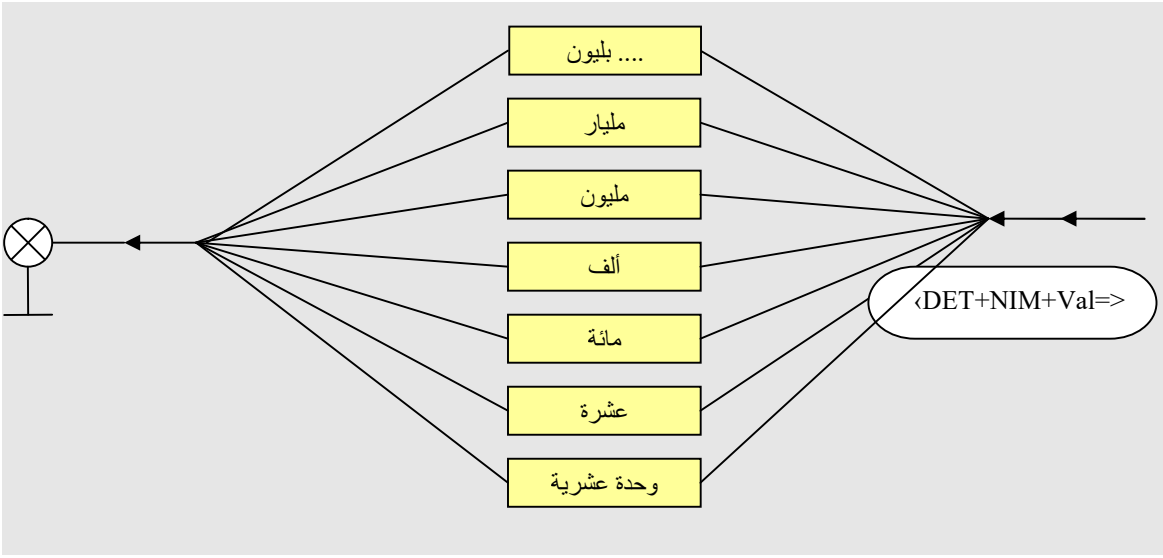
هذه المتوالية يمكن أن نجدها تتكرر في رقم ما كما في التعبير العددي التالي : خمسة آلاف وخمسة مائة وخمس(5505)التي يمكن تمثيلها حسب بيئة نوج للتظوير اللغوي بالرسم التالي:

([15]) Agata Savary, **recensement et description des mots composés-méthodes et applications** , thèse de doctorat en informatique fondamentale , soutenu le 14 décembre 2000 à l Université de Marne la Vallée , laboratoire automatique documentaite et linguistique , université Paris 7.

([16]) عالج هذا الصنف الكثير من علماء النحو القدامى منهم ابن يعيش في شرح المفصل والمبرد في المقتضب وسيبويه في الكتاب والاسترابادي في شرح الكافية في النحو .ونريد أن نبرز أهمية العدد المركب الذي أصبح مصدر ًالاهتمام كثير من علماء اللسانيات والمهندسين نظرا للوفرة الغزيرة لهذا النوع من المركبات لاسيما مع الثورة التكنولوجية .

([17]) تطرق الفلاسفة إلى مقولة العدد وفصلوا القول فيها تفصيلا وخاصة حين تحدثوا عن الطبيعة والميتافيزبقا، وعن مقولة النوع والعدد والكثرة والفصيلة نذكر هنا تمثيلا لاحصراً أرسطو وابن رشد...

([18]) عبد الطيف شوطا، تناظر الأوضاع وطبقات الأسماء في اللغة العربية، **المركبات الإسمية والحدية في اللسانيات المقارنة** ، 1999.ص.196.



شكل 2 : رسم رئيسي يمثل أنحاء محلية للتعرف على الوحدات الرقمية

3- مرحلة الوسم المعجمي للوحدات الإسمية:

لا يحصل- مهما كان الأمر- لأي برنامج آلي مُعد سلفاً للتعامل مع المعجم الآلي العربي، أن يتعرف على جميع مكونات النص المستند على عينة قليلة العدد، إلا إذا كان هذا المعجم محتوياً على جميع أنواع وأصناف الوحدات المركبة في اللغة. وقد بينت التجربة إمكانية تطبيق نفس المنهجية التي تتبع في التعرف الآلي على المفردات البسيطة أن تحدد شكلها الصوري المتمثل في متوالياتها الحرفية المحاطة ببياضين.

وقبل البدء في صياغة نماذجها، لا بد من التذكير بماهيتها الصورية، فما هي الوحدات الإسمية المركبة إذن؟

تعتبر الوحدات الإسمية المركبة كل متوالية مؤلفة على الأقل من مفردتين بسيطتين:([19])

وحدة إسمية مركبة = (إ سم بسيط 1 + أسم بسيط 2) + حرف

وقد ذهب ماكس سلبرشتاين حين عرف أن كل مفردة مركبة مقيدة، إلى اعتبار أن المتوالية هي بالمعنى التقني (منطقة مترابطة (zone connexe)) ([20]) لا يمكن النيل منها لأنها صامدة أمام أي تغييرأو استبدال حيث أنها متلازمة ومتكلسة لا يمكنك إضافة أو إدماج أي عنصر بينها، لنتأمل المثال التالي:

إني رأيت أحد عشر كوكباً

([19]) Max Silberztein,**Dictionnaires électroniques et analyse automatique du textes** Masson. 1993.p 137.

([20]) نفس المرجع، ص .144.

اذ لا يمكنك قول: *إني رأيت أحد X (كوكبا) عشر

إضافة إلى أوتومات الأوضاع النهائية اعتمد سيمون أولمان دالة القيمة (V) لتحليل معطيات الأعداد المركبة بتحديد تركيبها الصحيح في الخوارزم، كما في الشكل التالي:

$$V: \ X \longrightarrow V(X)$$

فلكل الأعداد أبجدية تامة (*terminal alphabet*) هي $x \in T$ ، بحيث أن قيمة $V(x)$ هي أصناف الأعداد التي تنتمي لنسق العدد من واحد إلى مائة نحو:

v (واحد) =1 ، v (ثمانية) =8 ، v (سبعة عشر) = 17

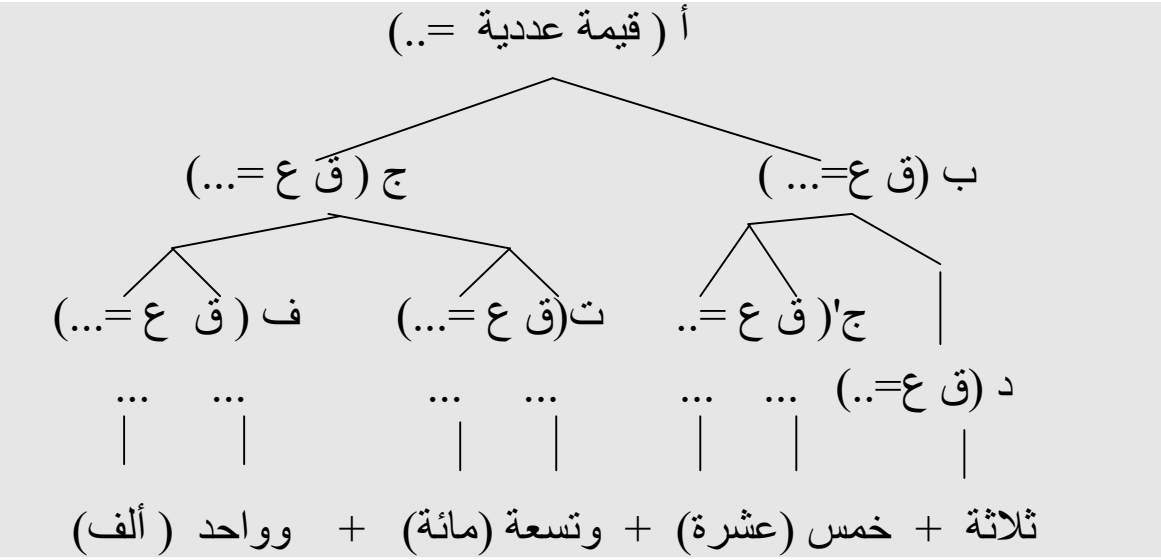
v (مائة)= 100 ، v (صفر) =0 ...

أما القيمة: $v(n\ q)$ فتحدد عملية الضرب والزيادة كما يلي :

$$V(x\ y) = v(x) * v(y) \ , \ y \in \{\gamma,\delta\}$$

$$V(x\ y) = v(x) + v(y) \ \ y \ , \ \in \{\gamma,\delta\}$$

يمكن تطبيق العمليات بهذه المعادلات اعتماداً على القيم المحددة سلفاً، كما يمكن تمثيل هذه القيم بالرسوم وأوتوماتات الأوضاع النهائية الخاصة بذلك تطبيقاً على نموذج من نسق العدد في اللغة العربية وهو: ازداد محمد سنة " **ثلاثة وخمسين وتسعمائة و ألف**" كما يلي :



شكل 3 : يبين شجرة تمثيل القيم الرقمية للأعداد المركبة الناتجة عن العمليات الحسابية.

أما فيما يتعلق بإعراب الأعداد المركبة فتخضع إلى التغيرات التصريفية كالإفراد والتثنية والجمع والتذكير والتأنيث.وتتلخص معالجتها هنا بالانطلاق من العوامل

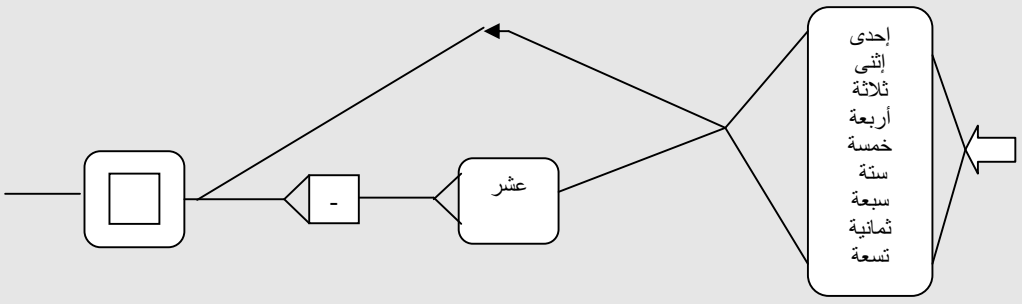
الإعرابية التي تظهر في بداية آخر متوالية للكلمة المؤلفة للمفردة المركبة،كوحدة إسمية، إذ يقوم البرنامج بعد ذلك بإعطاء رمز لهذه المتوالية، مع العلم أن نظام العدد في اللغة العربية ذو شكل موحد مما يساعدنا على تحديد العنصر الذي يلحقه الإعراب واستبداله بالأشكال التي تلائم العمليات التصريفية للعدد المركب وفق ترميز محدد ودقيق.(21) مثل (الأعداد العشرية):

1- ثلاثة عشر، خمسة عشر، سبعة عشر
2- خمسة وعشرون، ستة وعشرون
3- أربعة وخمسون، واحد وثمانون
4- مائة وواحد، مائة وخمسة، مائة وخمسة وستون

يتمكن النظام الالي نوج من إتاحة الإمكانية الكبيرة في تحليل نسق العدد في اللغة العربية بتطبيق أوتومات الحالات النهائية بهدف اختبار قبول أو عدم قبول المفردة المركبة في نظام الكفاية المعجمية، اعتماداً على حسابات قيم العدد التي تعتبر في حد ذاتها قيماً لأنحاء محلية.

وتجدرالإشارة كذلك، في هذا الإطار،أن موريس كروس (2002)(22) ذهب الى تطبيق نفس مقاربة معالجة نسق العدد على الوحدات الإسمية الأخرى كالتواريخ والأزمنة والنسب المائوية .

أما عن حزمة أوتومات الحالات النهائية المطبق على نظام العدد في اللغة العربية ، فيمكننا أن نختبر النماذج التالية والتي تتعلق بتوليد الأعداد من إحدى عشر(11) إلى تسعة عشر (19) كما يلي:

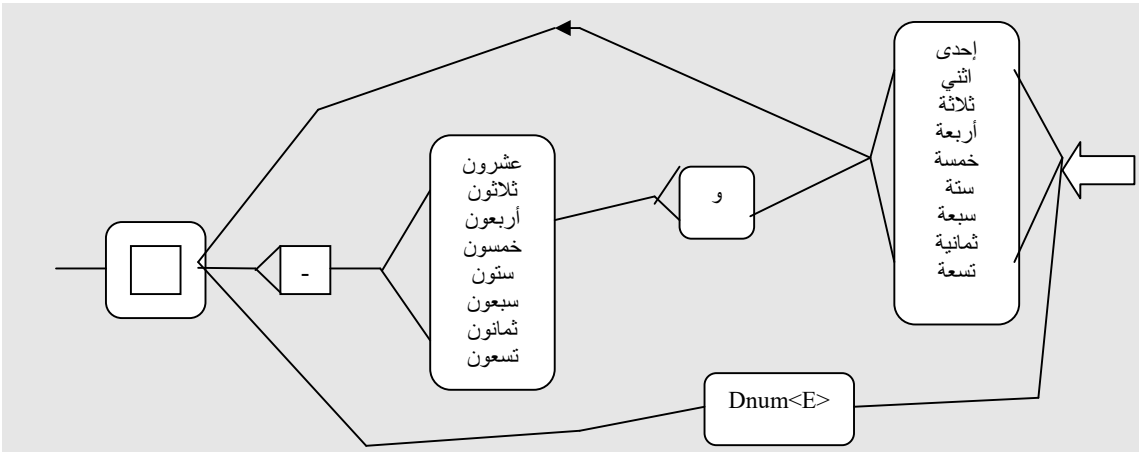


شكل 4:رسم يمثل توليد العدد المركب في العربية من إحدى عشر إلى تسعة عشر

نفس الأمر ينطبق على نسق توليد العدد من إحدى وعشرين إلى تسعة وتسعين كما هوالأمر مع الأتومات التالي:

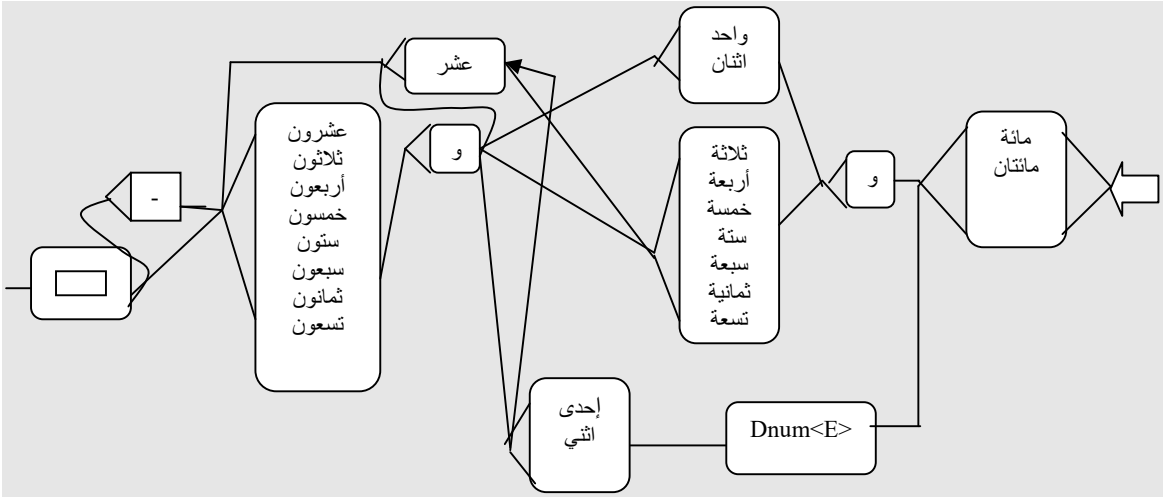
([21]) نفس المرجع، ص. 15.

([22]) M.Gross, les déterminants numéraux, un exemple: les dates horaires **in Langages** n°145, mars2002.



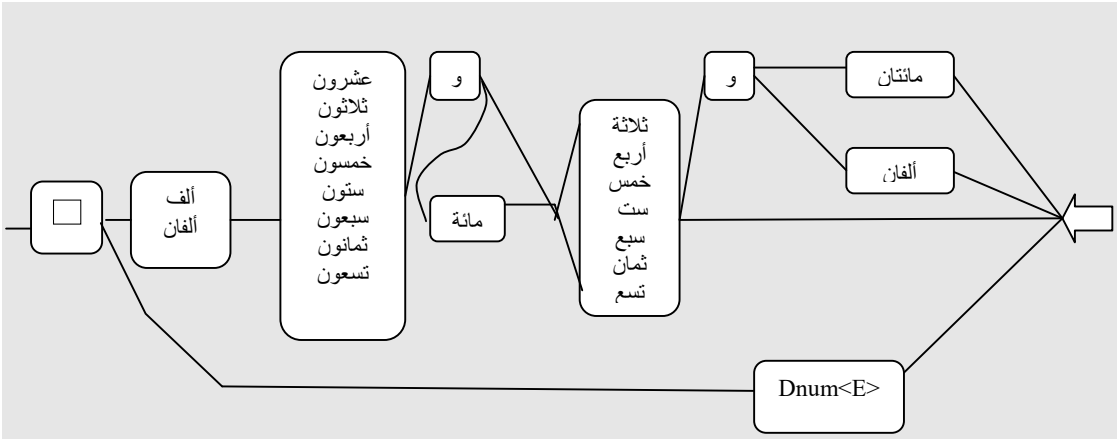
شكل 5 :رسم يمثل العدد المركب في العربية من إحدى وعشرين إلى تسعة وتسعين

ونفس الاسلوب مع توليد العدد من مائة وواحد إلى مائة و تسعة وتسعين :



شكل 6 :رسم يمثل توليد العدد المركب في العربية من مائة وواحد إلى مائة و تسعة وتسعين

وأخيرا توليد العدد المركب في العربية من مائة وواحد إلى ما فوق الألف كما في الشكل التالي:



شكل 7 : رسم يمثل توليد العدد المركب في العربية من مائة وواحد إلى ما فوق الألف

4-القيمة العددية المحددة وأمن لبس الوحدة الإسمية في النص:

اعتمادا على الأدوات العقلانية المذكورة والموظفة في منصة نوج المفتوحة المصدر وكذا على مجموع الظواهر الموفو- تركيبية التي تتيحها الأنحاء المحلية في نظام اللغة العربية حيث تتميز طريقة اقتران علامات العدد بالإسم في الوحدة العددية بالعلاقات الإعرابية لكن بتأثير عاملي معكوس (*anti-accord*) ([23])، يظهر الفرق بين العدد كإسم والاسم كإسم، إذ أن المراد هو تحديد الكمية بماهيتها وليس تمثيلا بألفاظها إجابة عن السؤال (بكم) حول حقيقة الشيء. وهكذا يعرب العدد ويشكل في الأنحاء المحلية التالية:

- في حالة الافراد والتثنية فإن القيمةالعددية في العربية تصبح علامة على الاسم وفي حالات خاصة ومطردة .

- فيما بين ثلاثة (3) وعشرة (10) فإن القيمة العددية تحمل علامة الجنس المعكوسة لتلك التي مع الاسم الذي يتبعه، إذا كان مؤنثا يصبح مذكرا والعكس صحيح نحو: العدد خمسة (5) في المثالين التاليين:

خمسة كتب و خمس سيارات

- عندما تساوي القيمة العددية إحدى عشر (11) أو إثنى عشر (12) تصبح الوحدات العشرية العددية مركبة متضامة ومتكلسة تماماً، وفي هذه الحالة يكون إعرابها مناسبا جنسا وعددا (منصوب بالفتحة الظاهرة في آخره) مع الاسم الذي يليه نحو :

أحد عشر كتاباً و إثنتا عشرة سيارة

- أما عندما تكون القيمة العددية من ثلاثة عشر(13) إلى تسعة عشر(19)، فإن العنصر العددي الاول ذو قيمة ما بين ثلاثة (3) وتسعة (9) يحمل الجنس المختلف عن الاسم الذي يليه، أما العنصر الثاني من الوحدة العددية فيحمل الجنس نفسه الذي يمثله الاسم الذي يليه نحو:

خمسة عشر كتاباً و خمس عشرة سيارة

([23]) سليم مصفار، ص.166 (2008).

-وإذا كانت القيمة العددية تساوي عشرين (20) ...إلى غاية تسعين (90) يبقى الجنس والعدد متماثلان لا يتغيران نحو:

ستون كتابا وستون سيارة

-أما إذا كانت القيمة العددية تساوي أربعة وعشرون (24) وخمسة وثلاثون (35) وستة وسبعون (76) فإن الاتفاق يكون جزئيا في جنس الاسم الذي يليه، في حين أن العناصر الأولى للوحدة العددية ما بين القيمتين ثلاثة (3) وتسعة (9) تتفق مع جنس الاسم الذي يتبعه، أما العنصر العددي الثاني الحاملة للقيم عشرون (20) وثلاثون (30) وتسعون (90) لا تتغير نحو:

خمسة وعشرون كتابا وخمس وعشرون سيارة

- غير أن القيم العددية من مائة (100) وألف (1000) ومليون (1000000000) لا تلحقها تغييرات إعرابية نحو:

مائة كتاب و ألف سيارة و مليون درهم

يتضح مما سبق أن القيم العددية غير نوع الاسم لا تثنى ولا تجمع تظرا للطبيعة الدلالية المختلفة عن طبيعة الاسم العادي الذي ينوب عن العدد في كل هذه التصريفات لأن وظيفتها الصورية تمثل عدد الاشياء الجامدة والمتحركة في الوجود فيكون الوقع على النوع لا على العدد وفي نظام اللغة العربية ما يؤكد ذلك حينما نريد أن نميز بين رجل ورجال وشجر وأشجار في الأمثلة التالية([24]):

سبعة رجال و خمسة أشجار

التي يتغيرصرفها مع القيم التالية:

مائتان وأربعة وعشرون رجلا و أربع مائة وخمس وستون شجرة .

كما في التمثيل التالي:

<DET+NUM+Val=7 > سبعة</DET>

<DET+NUM+Val=224 > مائتان واربع وعشرون رجلا </DET>

خلاصات واستنتاجات:

كانت بغيتنا وصف نظام التعرف على الوحدات الإسمية بعدما خصصنا هذه المقال لمعالجة الأعداد كوحدات إسمية مستقلة بواسطة المحلل الصرفي وقواعد التعرف الآلي الموجودة في بيئة نوج وخاصة الأنحاء المحلية الإلكترونية المستحدثة لهذا الغرض، لأنه يستطيع أن يتعرف على الأشكال المعالجة ويصنفها بما في ذلك تلك التي تتصف بالغموض في النص. ومن أجل بلوغ النسبة الكبرى للتعرف على جميع الأشكال الملتبسة بات استكمالا للقواعد المعجمية، إعداد قاعدة للوحدات الإسمية الأخرى كالتواريخ ووحدات القياس والنسب والتعبيرعن الزمن وأسماء الكيانات وغيرها في اللغة العربية أولا، وثانياً اعتماد المعالج الآلي نوج(NooJ) المُعَد بمجموعة من أتوماتات الأوضاع

([24]) سليم مصفار، ص.163 (2008).

النهائية وبالرسوم الممثلة لتلك الأنحاء، ولتحسين أداء البرمجي يجب توفير غطاء لغوي أوسع لهذه القاعدة في البيئة المستعملة للتحليل والمعالجة.

إن معالجة المتواليات العددية في نظام اللغة العربية، بالأساليب والأدوات العقلانية المعلوماتية هو بهدف إعداد معجم إلكتروني موسع للغة العربية، إذ وضحت التجربة أنها لا تختلف في شيء عن باقي اللغات فيما يتعلق بالبنيات اللسانية المتصفة بالثبات و اللاتأليف. ومع أن الميزات والخصوصيات المتوصل إليها وصفياً وتصنيفاً وتعرفاً آلياً، كفيل بتيسير عمل مصممي المعجم ومهندسيه، بغية تمكين المستعمل العربي للحاسوب من ضبط آليات النظام اللساني العربي في جميع مستوياته الصورية والتطبيقية.

ونجمل شتات ما ذكرناه في النتائج التالية :

- تعتبر الوحدات العددية حاصل حسابي بين المستوى الصوري والمستوى التركيبي والاحالي،وهو المحدد للعلاقات المادية بين المتواليات اللسانية والمستوى التمثيلي الكامن في الكفاية المعجمية للأفراد.
- تتميز الوحدات العددية ذات قيمة رقمية طبيعية يمكن تمثيلها في المتوالية اللغوية.
- ينشأ الغموض من أول عملية للتعرف الآلي المطرد وهو من أهم المشاكل المطروحة على المتواليات العددية المركبة، مما يستدعي بناء برمجيات أكثر تطوراً و ملائمة لنظام اللغة الطبيعي. ولن يتأتى ذلك، إلا ببناء قاعدة بيانات شاملة لها يتضمن جميع الأصناف وليس بالاعتماد فقط على أمثلة قليلة العدد.
- تجاوز محنة فقدان كل استقلال معجمي ودلالي عند إدماج عنصر في الوحدة الإسمية، ببناء أنحاء محلية جزئية لا تختلف عن النظام اللغوي العام، مما سيساعد برامج التعرف الآلي القائم على أوتوماتات الأوضاع النهائية بوصف أدق لعناصر الوحدة الإسمية، لأنه تركيب خاص في ذاته لا يخرج إطلاقاً عن نظام الكفاية المعجمية.

وقد تبين من خلال نماذج لأوتومات الأوضاع النهائية التي قاربنا بها بعض العينات القليلة العدد للعدد في اللغة العربية،وأردت بها أن نمثل لعملها الخوارزمي (لأن ذلك يحتاج إلى بحث مستفيض لا يتسع له المقام الآن) توضيحاً لمدى قوة و إنجازية وجودة واتساع والاستغلال المنسق لبيئة **نوج** التطويرية لما لها من أهمية:

- في التنمية اللغوية بإتاحتها بناء واختبار وإبقاء التعريفات الشكلية بغطاء أوسع للغة، في شكل معاجم حاسوبية أو أنحاء إلكترونية متممة لبعضها البعض، نظراً لطبيعة عناصر اللغة المشكلة للمادة المعجمية.
- وفي تطوير التطبيقات لمعالجة اللغة صرفياً وتركيبياً وتحليل المدونات بالنسبة للمختصين في التوثيق والتطبيقات الحاسوبية.

وتتضمن هذه البيئة المجانية المبنية بأدوات عقلانية وهندسية، أدوات للمعالجة الآلية للغة مثل نظام إدارة البيانات ولغة البرمجة مفتوحة المصدر، كما تمنح إمكانية معالجة المدونات المعجمية وطرق البحث والاختبار مثل الصيغ التركيبية والدلالية.وهذا ما يتطلب صورنة اللغة العربية ووضع منصة جيدة للتحليل.

ومن جانب آخر، يعتبر **نوج** محللا آلياً لمنظومة المعاجم وخاصة عند توحيد المداخل المعجمية"super-lemmas" من جهة، ومن جهة أخرى، عند تدبير منظومة المعاجم جنبا إلى جنب مع جداول المعجم- التركيبي .وقد أصبح بإمكان المستخدم ولوج معجمه سواء عن طريق الجداول أو عن طريق القوائم،أو حتى ولوج الخاصيات

الصرفية والتركيبية في الآن نفسه. وهذا جدول يلخص المعطيات الإحصائية للغة العربية في نظام نوج كما يلي ([25]):

المعجم الآلي	عدد الوحدات المعجمية	عدد الأشكال المعربة
الأفعال	10500	2734122 شكل منها 1290795 فعلا، و1443327 مصادر وغيرها..
الأسماء	15000	280267 شكل
الصفات	4600	163866 شكل
أسماء الأعلام	12400	
الكيانات	5300	

المراجع والمصادر:

[1] سليم م،زايد محمد، دراسة في المعاجم الفرنسية الحاسوبية، **أشغال الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية، مدينة الملك عبد العزيز للعلوم والتقنية** ، الرياض- أبريل 2008.

[2]غازي عزا لدين، **بناء المعجم الآلي للغة العربية: معالجة التعبيرات المتلازمة** أطروحة الدكتوراه، كلية الآداب عين الشق ،الدار البيضاء 2009.

[3] الحناش محمد، برنامج لساني– حاسوبي للتعرف الآلي على التعابير المسكوكة في اللغة العربية، **مجلة التواصل اللساني** م.3 سنة 1996.

[4] الحناش محمد،النحو التأليفي(مدخل نظري)،**مجلة دراسات أدبية ولسانية**.(1985) صص35-83.

[5] الحناش محمد، التعبيرات المسكوكة في اللغة العربية، تقديم نظري ومنهجي ،**منشورات جامعة عجمان** (2002).

[6] الحناش محمد، المعجم الآلي للغة العربية، بناء قاعدة معطيات ،مجلة **التواصل اللساني**، م.2 ع.1 سنة 1992.

[7] E. Laporte, la reconnaissance des expressions figées lors de l 'analyse automatique, in **langage** n°90, 1988.pp. 117-126.

[8] Labelle. J, lexique - grammaires comparées, formes verbales figées en français de Québec. in **langage** 90.1988.Larousse.

[9] Max Silberztein **Dictionnaires électroniques et analyse automatique du textes** Masson. 1993.p 137.

[10] Martin Simon Ulmann, l' analyse des mots composés Allemands, in **Notes techniques** n°94/8. LATL, Université de Genève p.1-16.

[11] Merella. Conenna., les expressions figées en Français et en Italien: problèmes lexico-syntaxiques de traduction. in **langage** n°90.88.

[12] Mesfar Slim, **Analyse morpho-syntaxique automatique et reconnaissance des entités nommées en arabe standard**, thèse doctorale de l'université de Franche Compté, France 2008.

[13] www.nooj-association.org (2018).

([25]) نفس المرجع أعلاه.