كلية الآداب والعلوم الإنسانية بصفاقس-تونس
وحدة بحث اللسانيات والنظم المعرفيّة المتّصلة بها

الصّرف بين التحويل والتصرف

تكريما للأستاذ الطيّب البكّوش

وقائع الملتقى الدوليّ الثالث في اللسانيات

صفاقس 21-22 أكتوبر 2009

إشراف: عبد الحميد عبد الواحد

تونس 2010

كلية الآداب والعلوم الإنسانيّة بصفاقس– تونس
وحدة بحث اللسانيات والنظم المعرفيّة المتّصلة بها

LSCA
Le Linguistique et les Systèmes Cognitifs Associés

الصّـــــرف
بيـــن التحويل والتحريف

تكريما للأستاذ الطيّب البكّوش

وقائع الملتقى الدوليّ الثالث في اللسانيات

صفاقس 21–22 أكتوبر 2009

إشراف : عبد الحميد عبد الواحد

تونس 2010

المحتوى

مقدمة **5**

المقدمة الخاتمة **13**

سيرة ذاتية **15**

الصيغية وموضوعها **25**
عبد الحميد دباش

الميزان الصرفي بين مصطلح المفهوم ومصطلح الوظيفة **45**
رزيق بوزغاية

جدلية الشكل والدلالة في الصيغمية العربية **69**
نواري سعودي

الأبنية المتّحدة في الأصول والمعنى وقضية أصل الاشتقاق .. **85**
محمد الصحبي البعزاوي

مبادئ التحليل الصرف – صواتي العربي القديم بين الوقائع الصوتية والسياقات الصرفية **103**
مصطفى بوعناني

أبعاد التفاعل الصرف – صواتي في الإنجازات والإدراكات اللغوية العربية: مقاربة لسانية معرفية **125**
هدى بلمكي

الجذور في العربية: دراسة مستقلة القطع **145**
مولدي اليحياوي

الوحدات الصرفية ووظائفها الدلالية في اللغة العربية **159**
صالح سليم الفاخري

الصّيغمية بين شكل البنية ودلالة الشكل **179**
الحبيب النصراوي

أثر علم الصرف في منهج ترتيب المداخل المعجميّة في القواميس العربية 205
محمد الغريبي

الكلمة ونظام الوحدات القياسية 227
مراد بن عياد

ما حظّ الفعل الماضي من البناء؟ 251
عبد الحميد عبد الواحد

"جريان الحدث" في الفعل 265
رضا الطيب الكشو

منزلة الوزن الصّرفيّ بين الوزن العروضي والوزن التصغيري 289
محمّد عبد الجبّار بوشعالة

التقابل اللغوي في تصريف الأسماء والأفعال وما يطرأ عليها من تغيّرات بين العربية والإنجليزية 305
أسماء أحمد
رشيد المومني

برنامج المحلّل الصّرفيّ الآليّ للعربيّة : الصّياغة والإشكاليّات ... 325
صالح الماجري
وبشير الورهاني

كشف وإصلاح أخطاء التّطابق في نصوص عربيّة غير مشكولة . 341
مكرم بوجلبان
شفيق علولو
لمياء هدريش بلغيث

برنامج المحلّل الصّرفيّ الآليّ للعربيّة: الصّياغة والإشكاليّات

صالح الماجري
بشير الورهاني[1]

تمهيد

يندرج مشروع صياغة محلّل صرفيّ آليّ للعربيّة في إطار مشروع المعالجة الآليّة للعربيّة. ونذكّر بأنّ هذه الغاية تتطلّب في جميع الألسن المرورَ بمراحل هي التّالية : التّحليل الصّرفيّ، فالتّحليل التّركيبيّ، فالتّحليل الدّلاليّ. وهو ما يجعل صياغة محلّل صرفيّ شموليّ أساسَ كلّ مشروع معالجة آليّة للسانٍ ما.

في هذا الإطار يعمل فريقنا المشترك على إعداد موارد وقواعد بيانات مُرَقمَنَة. إلاّ أنّ معالجتها ووسمَها آليّا يحتاج إلى محلّل صرفيّ يكون قادرا على التّعرّف إلى جميع أشكال الوحدات المعجميّة الممكنة في العربيّة (المشتقّة/ غير المشتقّة؛ البسيطة/ المركّبة؛ الخ.) ووَسْمها آليّا (قسم الكلام المنتمية إليه، مقولات الجنس والعدد، التّصريف، الجذر والوزن بالنّسبة إلى المشتقّة، الخ.).

يتوزّع هذا العرض على محاور ثلاثة: أمّا الأوّل فيذكّر سريعا بمقتضيات التّحليل الصّرفيّ الآليّ ويتعرض لبعض المحاولات في الموضوع. وأمّا الثاني فيتناول التّصوّر العامّ للمحلّل الآليّ فيعرض بنيَتَه ويقدّم عيّنات من القواعد التي تنظّمها. وأمّا الثّالث فيناقش عددا من الإشكاليّات وطرق التّعامل معها. وقد اخترنا بصفة صريحة التّركيز على الجانب اللّسانيّ من القضايا.

---

1) مخبر **المعاجم والقواميس والإعلاميّة** (LDI) باريس 13/ وحدة البحث **المعالجة الإعلاميّة للمعجم** 00/UR/0201، جامعة منّوبة/ المخبر العالمي المشترك **الألسن والتّرجمة والتّعليم** (LIA) المركز الوطني للبحث العلميّ، فرنسا.

1. مقتضيات المعالجة الآليّة وبعض المشاريع المنجزَة

نودّ، قبل الخوض في مسائل المحلّل الصّرفيّ الآليّ، أن نذكّر بالمقصود بالمعالجة الآليّة للألسن الطّبيعيّة داخل الإطار النّظريّ الذي نعتمده، وبالمستلزمات المنهجيّة لهذه الغاية التي تسعى إليها مشاريع الوصف اللّسانيّ والبرمجة الإعلاميّة في إطار فريقنا المشترك (وحدة البحث المعالجة الآليّة للمعجم/ مخبر المعاجم والقواميس والإعلاميّة).

1. 1. مقتضيات المعالجة الآليّة وموقع المحلّل الصّرفيّ

إنّ أعمال المعالجة الآليّة التي نقصدها بكلامنا ترسم لنفسها هدفا أقصى يمكن تلخيصه في محورين متكاملين هما:

- ✓ التّحليل الآليّ للأشكال وتركيبها ودلالاتها؛
- ✓ الإنتاجُ الآليّ للأشكال وتركيبُها تركيبا "سليما" من وجهة النّظر النّحويّة.

ويعني ذلك في نهاية الأمر القدرة على "فهم" الجمل وإنتاج الجمل السّليمة (المطابقة لقواعد النّحو والمقبولة دلاليّا).

ونذكّر بأنّ التّحليل الآليّ يشترط إنجازَ وصف منتظم وبرامج آليّة وفق المراحل التالية:

- التّحليل الصّرفيّ؛
- التّحليل التّركيبيّ؛
- التّحليل الدّلاليّ.

وتتطلّب صياغة المحلّل الصرفيّ، الذي لا بدّ من أن يكون شموليّا، مدوّنات تكون منطلقا للوصف وسندا لاختبار النّتائج. ويمكن أن نذكر مصادر عديدة لجمع المدوّنة وبنائها[1] :

- مؤلّفات مُرَقمَنَة؛
- مدوّنات صحفيّة،
- قواعد بيانات (مختلف الاختصاصات...)؛

---

1) يجدر أن نذكّر في هذا الصّدد بأنّ المدوّنة شيء يُبنى وفق الأهداف والمقتضيات المنهجيّة للباحث وليست شيئا خاما معطى بصفة مسبقة.

- نصوص مختلفة على الأنترنات.
- الخ.

والملاحظ طبعا توقرُ الموارد وقواعد البيانات. لكنّها موادّ خام في أغلبها لا بدّ من تبويبها ووسْمها صرفيّا بطريقة آليّة.

ويُنتظَر من المحلّل الصّرفيّ مبدئيّا:

- أن يتّسم بالشّموليّة ويأخذ في الاعتبار مختلِف خصائص النظام الصّرفيّ في العربيّة.
- أن يكون قادرا أيضا على استيعاب الأشكال التي تُستَحدَث.

1. 2. نماذج من المحلّلات الصّرفيّة

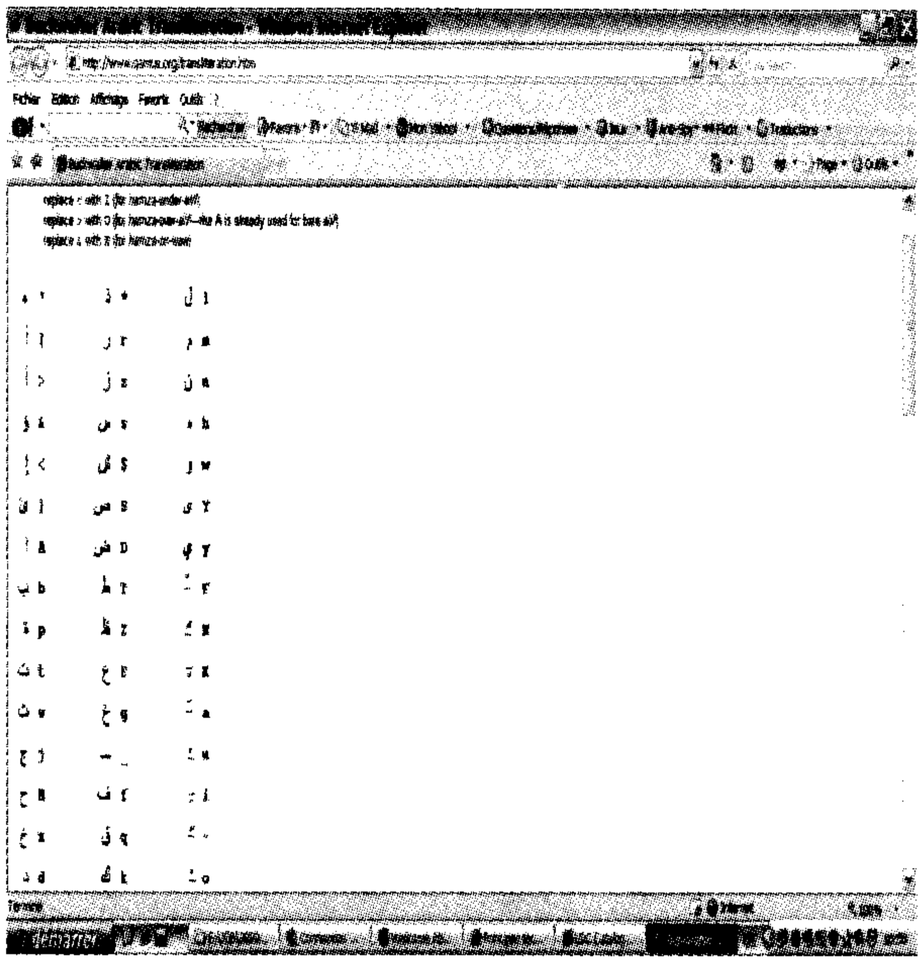
لا نعدم تجارب ومحاولات لصياغة محلّلات صرفيّة آليّة للعربيّة بعضها مازال متواصلا وبعضها انقطع منذ زمن. ونكتفي هنا بأن نذكر منها سريعا مثالين.

أ. محللّ بكولتر Buckwalter (ترعاه مؤسّسة Xerox[1]) الذي أتيحت لنا فرصة الاطّلاع عليه واختباره. ونبدي في شأنه الملاحظات التّالية :

•من وجهة نظر إعلاميّة، لا يعمل على الحروف العربيّة مباشرة، فهو يمرّ عبر مرحلة تقنين برموز تعتمد الخطّ اللاّتيني وفق نظام رموز وضعه الباحث نفسُه؛

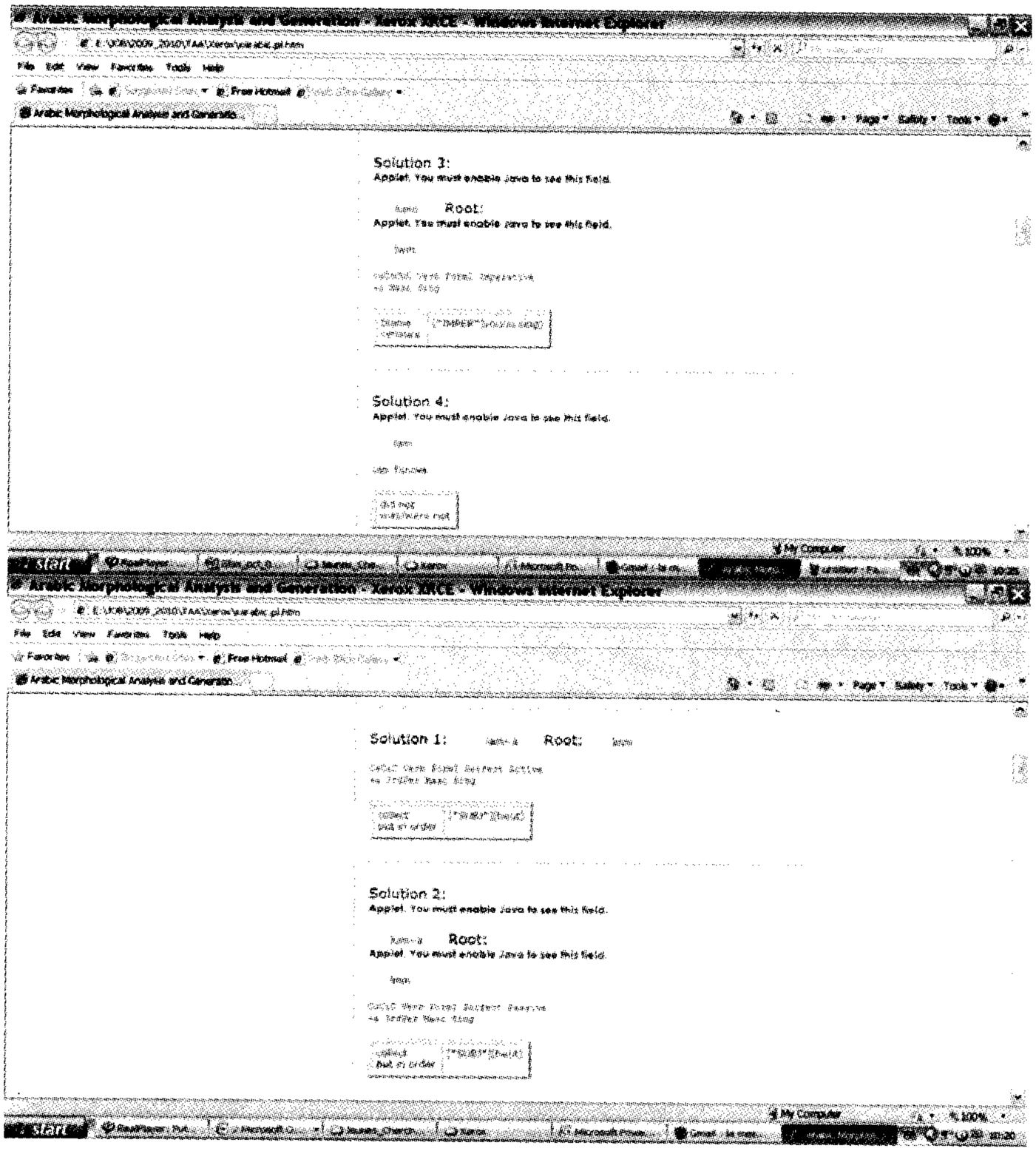
•هو موجّه لجمهور غير عربيّ اللّسان أساسا. وفي ما يلي صورة عن نظام الرّموز بالأحرف اللاّتينيّة التي يستعملها :

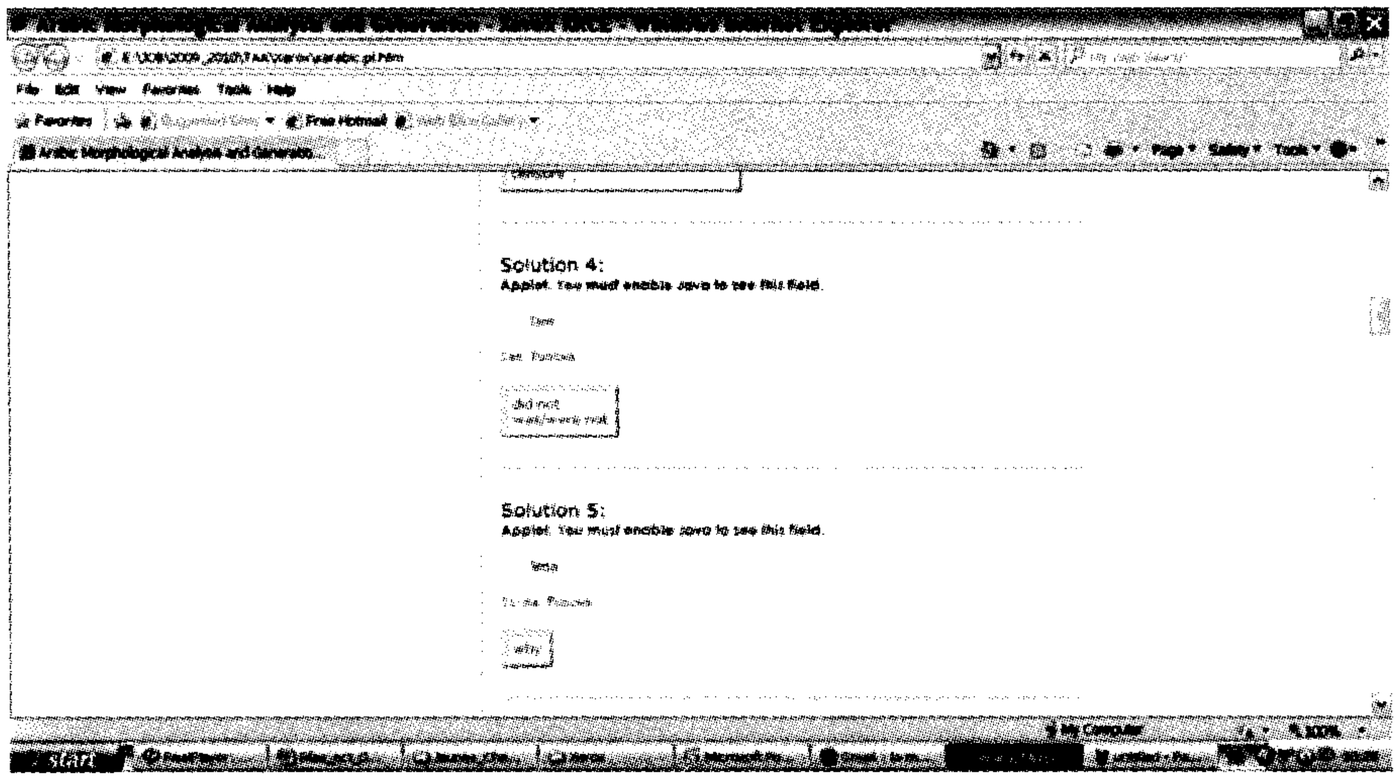
1) Xerox Research Centre Europe : http://www.arabic-morphology.com/ http://www.xrce.xerox.com/competencies/content-analysis/arabic/



• ونضيف إلى ذلك أنّ نسبة تغطيته للمعجم العربيّ تشمل الوحدات المشتقّة بصفة تكاد تكون حصريّة، فضلا عن أنّ هذا المحلّل لا يأخذ في الاعتبار الجانب الكامنَ في نظام الاشتقاق. يعني ذلك أنّه لا يمكّن مثلا من صياغة أشكال غير موجودة انطلاقا من قواعد تشكيل الوحدات في العربيّة.

• نشير أخيرا إلى أنّ هذا المحلّل يلاقي صعوبات مع الوحدات المشكولة جزئيّا. فهو يتعرّف إلى الوحدات المشكولة كلّيا ويقدّم مختلف الإمكانيّات بالنّسبة إلى الوحدات غير المشكولة، إلاّ أنّه لا يتعرّف إلى الوحدات الّتي يكون شكلها جزئيّا. وفيما يلي مثال عمّا يقترحه لتحليل الشّكل "لم" (دون حركات) :


Solution 3:
Applet. You must enable Java to see this field.
Root:
Applet. You must enable Java to see this field.
Solution 4:
Applet. You must enable Java to see this field.
Solution 1:
Root:
Solution 2:
Applet. You must enable Java to see this field.
Root:
Applet. You must enable Java to see this field.



ب. تجربة محلّل آليّ في تونس في أواخر الثّمانينات وبداية التّسعينات من القرن الماضي[1]. إذ شُرع في صياغة برنامج تصريف آليّ في إطار مشروع صياغة محلّل صرفيّ ومترجم آليّ. إلاّ أنّ الملاحَظ أنّ المشروع لم يتطوّر ولم تكن منهجيّة الصّياغة واضحة وصريحة.

ويمكن أن نشير إلى أنّ المصرّف الذي صيغَ لا يَشكّل الأفعالَ (لا يرسم حركة العين وغيرها من الحركات الأساسيّة في التّمييز بين الصّيَغ) كما أنّه لا يأخذ في الاعتبار الشّدّة (باعتبارها علامة التّضعيف، ومن ثمّ مَكمَنُ أهمّيتها في نظام الاشتقاق وفي عمليّة التّحليل الصّرفيّ). وينتج عن ذلك عددٌ من النقائص تجعل من المحلّل برنامجا غيرَ ناجع بالنّسبة إلى العربيّة. ذلك أنّ:

- حاصل التّصريف الآليّ غير قابل للاستعمال في المحلّل لأنّها تخلط بين مختلف الأنماط وجداول التّصريف (فهو لا يميّز مثلا بين: حسِبَ وحسَبَ وحسُبَ التي تمثّل جداول تصريف مختلفة لأفعال مختلفة)؛
- لا يمكن إجراء هذا المحلّل الصّرفيّ على مدوّنات تحتوي على نصوص مشكولة كلّيا أو جزئيّا. فإذا أخذنا الشّكل كتب الذي يمكن أن يكون كَتَبَ/ كُتِبَ/ كَتْبٌ/ كُتُبٌ، الخ، فإنّ هذا المحلّل لا يسعفنا في شيء إذْ لا يفرّق بينها.

---

1) Institut Régional des Sciences Informatiques et des Télécommunications (IRSIT).

ونكتفي في هذا المجال بهاتين العيّنتين. ونحن نَلمَح بعدُ من خلالهما عددا من الشّروط المنهجيّة والاعتبارات التي يجب أن تكون حاضرة من البداية عند التّفكير في صياغة محلّل صرفيّ آليّ للعربيّة. وهو موضوع العنصر الموالي.

2. بنية المحلّل الصّرفيّ الآليّ

نودّ التّذكير سريعا في هذا السّياق بالخصائص العامّة لمعجم اللّسان العربيّ من وجهة النّظر الصّرفيّة. ويمكن من هذا المنظور التّمييز بين مستويات عديدة:

– المشتقّ (ما يمكن ردّه إلى مادّة معجميّة ووزن معلوم)؛/ غير المشتقّ (بعض الأسماء والأدوات، الخ.).

– المستعمَل/ غير المستعمَل (الذي نعتبره طاقة توليد كامنة) ونحيل هنا على كتاب **العين** للخليل بن أحمد الفراهيدي في تقليبه لحروف الجذر الواحد تقليبا رياضيّا يفحص الثّلاثيّ فالرّباعيّ فالخماسيّ ويميّز فيما يخرج عنها من مشتقّات بين المستعمَل والمُهمَل. وتلخّص هذه الفقرة من مقدّمة كتابه المنهجَ المتّبَع:

"قال الخليل : اعلم أن الكلمة الثنائيَّة تَتَصَرَّفُ على وَجْهَيْن نحو : قَدْ دَقْ شَدْ دَش والكلمة الثلاثيَّة تتصرَّفُ على ستة أوجُه وتُسمَّى مَسدُوسة وهي نحو : ضرب ضبر برض بضر رضب ربض والكلمة الرباعية تتصرَّف على أربعة وعشرين وجهاً وذلك أن حروفها وهي أربعةُ أحرف تُضْرَب في وجوه الثلاثّي الصَّحيح وهي ستَّة أوجه فتصيرَ أربعة وعشرين وَجْها يُكَتَّب مُسْتَعْمَلُها ويُلغى مُهْمَلُها وذلك نحو عبقر تقول منه عقرب عبرق عقبر عبقر عرقب عربق قعرب قبعر قبرع قرعب قربع رعقب رعبق رقعب رقبع ربقع ربعق بعقر بعرق بقعر بقرع برعق برقع. والكلمة الخماسية تتصرّف على مئة وعشرين وجها وذلك أن حروفها وهي خمسة أحرف تُضرَب في وُجُوه الرُّباعيِّ وهي أربعة وعشرون حرفاً فتَصيرُ مئة وعشرينَ وَجْهاً يُسْتَعْمَل أقلُّه ويُلغى أكثره وهي نحو : سَفَرجل سفرلج سَفجرل سجفرل سجرلف سرفجل سرجفل سلجرف سلرفج سلفرج سجفلر سرفلج سجفرل سلفجر سرجلف سجرلف سرلجف سجلفر وهكذا."[1].

ويسمح هذا التّمشّي بأن نأخذ في الاعتبار من البداية كلَّ الطّاقة التّوليديّة الكامنة في نظام الاشتقاق في العربيّة. ولذلك اعتمدناه منطلقا لقاعدة بيانات المحلّل الصّرفيّ كما سنبيّن أسفله.

---

1) الخليل: **كتاب العين**. المقدّمة، ص 50–53.

هكذا يُنتَظر من الهيكل العامّ للمحلّل أن يأخذ في الاعتبار مختلفَ هذه الجوانب وأن يزاوج فيما بينها عن طريق:

- ✓ وجوب الفصل بين المشتقّ وغير المشتقّ؛
- ✓ إمكانيّة الإضافة والتّوسّع بما يعكس الحركيّة الطّبيعيّة التي تسِم الرّصيد المعجميّ لمختلف الألسن المتداوَلة.

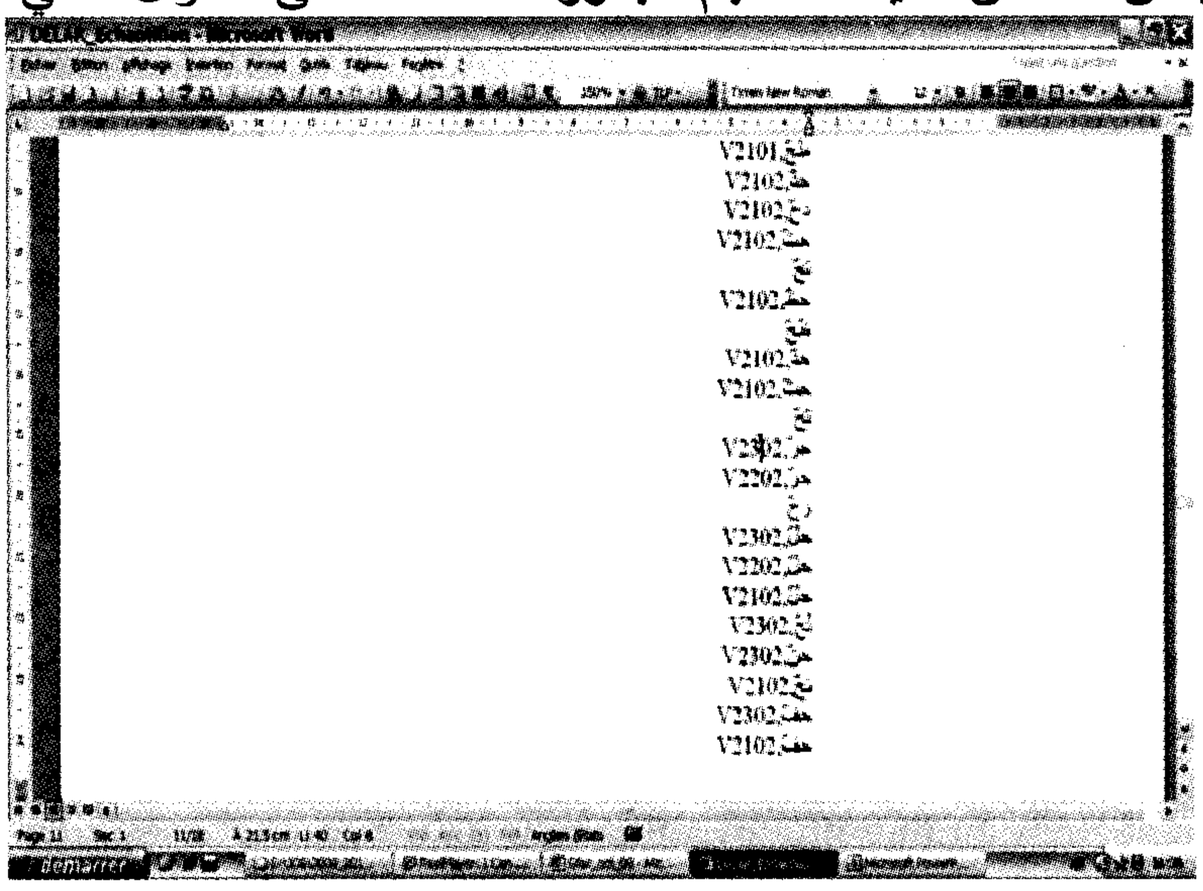
وقد اخترنا فيما يلي من عرضنا التّركيز على جانب الرّصيد القائم على الاشتقاق لما يوفّره من انتظام ولأهمّيّة القضايا التي يطرحها في آن.

أ. الوحدات المشتقّة نموذجا

تقوم صياغة بيانات المحلّل الصّرفيّ على بناء مجموعة من المعاجم الإلكترونيّة انطلاقا من عمليّة جرد أولى للجذور المتوقرة في مختلف القواميس والمعاجم ووسْمها وفق نظام من الرّموز يحمل معلومات عن نوع الجذر بحسب :

- الكمّ : عدد الحروف (ثنائيّ/ ثلاثيّ/ رباعيّ)؛
- النّوع: سالم/ صحيح/ مهموز/ مضاعف/ معتلّ (نوع حرف العلّة وموقعه)، إلخ.

يمكّن ذلك من صياغة "**معجم الجذور**" للمشتقّات على المنوال التالي:



إضافة إلى ذلك تمّت صياغة "معجم الأوزان" الذي يحتوي على مختلف الأوزان التي تتألف مع الجذور لتولّد المشتقّات وفق قواعد معيّنة. وقد استند ذلك إلى جرْد عدد من المؤلّفات التي تتناول مسائل الاشتقاق.

ولابدّ من الإشارة في هذا السّياق إلى عنصر ضروريّ في هذا الهيكل هو مجموعة القواعد المولّدة للوحدات المشتقّة انطلاقا من معجم الجذور ومعجم الأوزان، وهو عمل تقنيّ بحت يُنجَز في ضوء المعلومات اللّسانيّة وقواعد الاشتقاق المعروفة التي تتحكّم في النّظام. وتمكّن هذه العناصر الثّلاثة برنامجَ التّحليل الصّرفيّ الآليّ من العمل في الاتّجاهين. فهو ينتج/ يتعرّف إلى:

✓ مختلف الأشكال البسيطة للأفعال المصرّفة (مختلف الأزمنة والضّمائر) وذلك بواسطة مكوّن من مكوّنات المحلّل هو برنامج **تصريف الأفعال**؛

✓ مختلف الأسماء والصّفات البسيطة وتصرّفها جنسا وعددا بواسطة مكوّن آخر هو **برنامج تصريف الأسماء والصّفات**؛

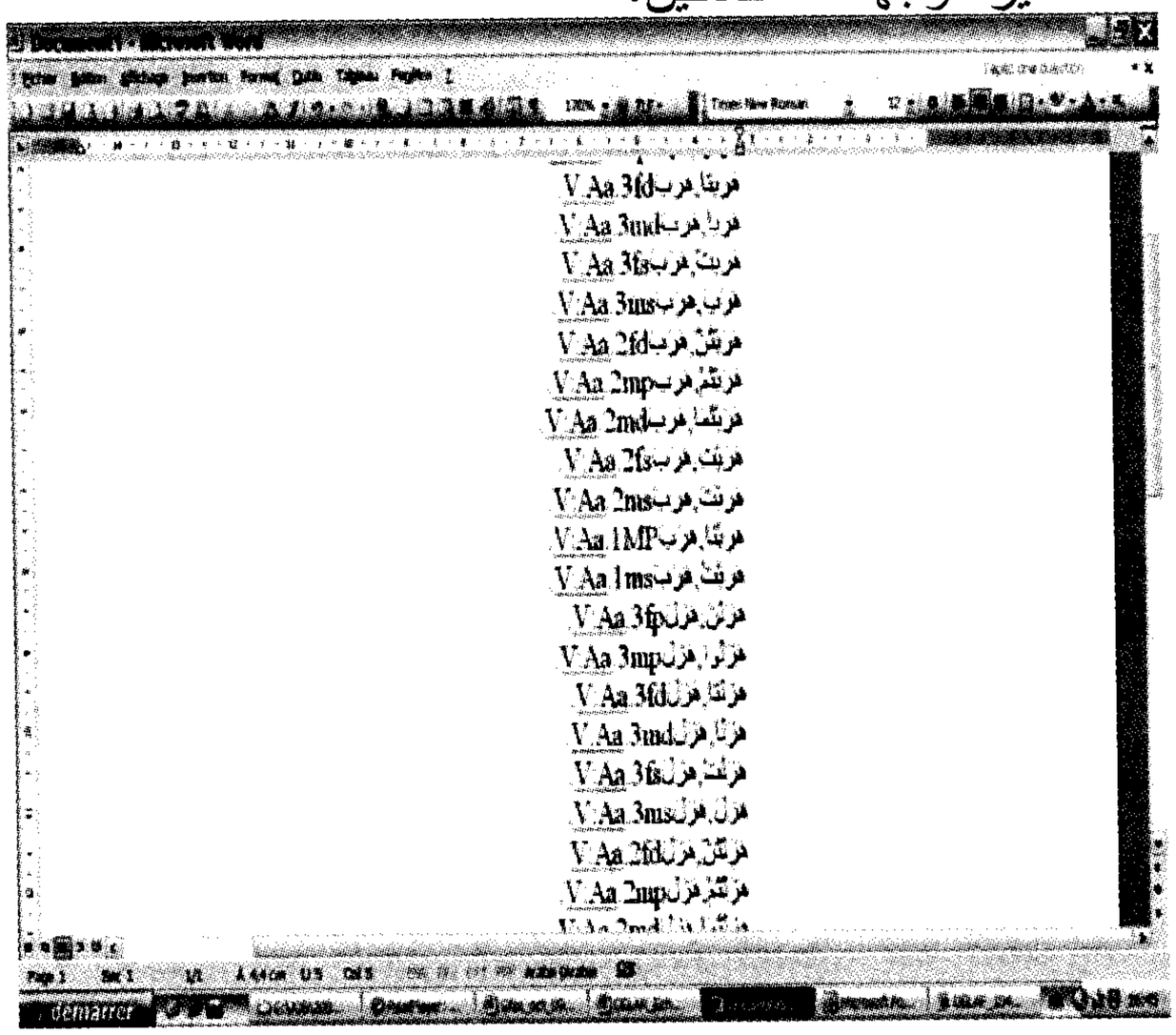
هكذا تكون للبرنامج ميزة أساسيّة هي الاقتصاد في الحجم وربح الوقت عند المعالجة والتّنفيذ. ذلك أنّ المصرّفين المذكورين لا يخزّنان الأشكال المشتقّة، بل يعتمدان قواعد الصّياغة المخزّنَة في معجم القواعد المشار إليه أعلاه، وهي بالطّبع أقلّ عددا من المشتقّات. وتبوَّب القواعد إعلاميّا وفق الهيكل التّالي:

- معجم قواعد اشتقاق الأفعال (مجرّدة ومزيدة بمختف أصنافها)؛
- معجم قواعد اشتقاق الأسماء (المصادر بأنواعها)؛
- معجم قواعد اشتقاق الصّفات (اسم الفاعل، اسم المفعول، الصّفة المشبّهة، إلخ.).
- معجم قواعد الرّبط بين العناصر المترابطة اشتقاقيّا (أفعال وصفات وأسماء تشترك في الجذر: ضَرَب، ضارب، مضروب، ضرّاب، ضرْب، ضربة، إلخ.).

ويكون الإنتاج أو التّحليل فوريّا عند طلب المستعمل الذي يتّصل بالبرنامج عن طريق واجهة معدّة للغرض. وفيما يلي فكرة عن برنامج تصريف الأفعال.

ب. برنامج تصريف الأفعال نموذجا

يتّسم هذا البرنامج بدرجة عالية من الانتظام بفضل نظاميّة تصريف الأفعال في العربيّة، ولاسيّما المزيد منها. وهي نظاميّة تغطّي القسم الأكبرَ من المعجم الفعلي للعربيّة حتّى بالنّسبة إلى الثّلاثيّ المجرّد الذي نلحظ فيه ظواهر عديدة من عدم الانتظام والخروج عن القياس. ونكتفي بأن نشير هنا إلى أنّ هذه النّظاميّة الطّاغية تترجمها بوضوح حركةُ العين التي توزّع الأفعال على أصناف وفق أنماط معلومة في الماضي وفي المضارع[1]. وفيما يلي نموذج عن نتيجة التّصريف التي يقدّمها البرنامج في نسخة غير موجّهة للمستعملين:

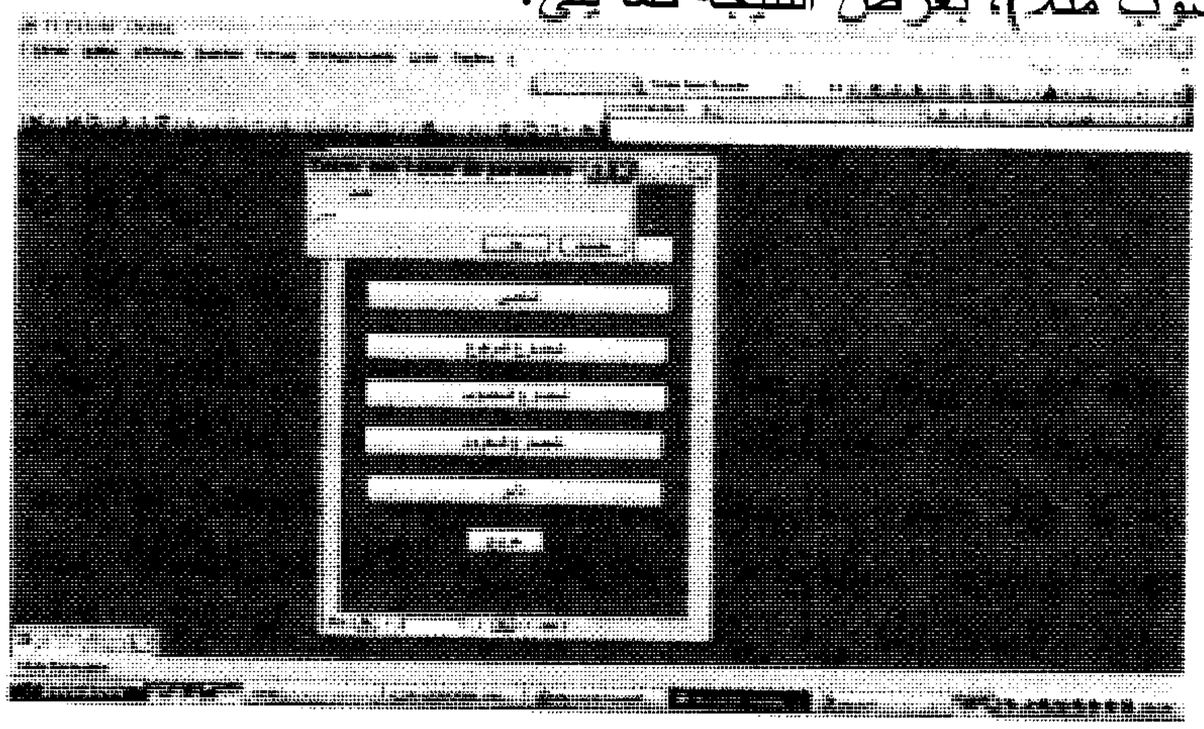


حيث تحمل الرّموز المرافقة للتّصريف مجموع المعلومات الصّرفيّة المتعلّقة بنوع الجذر وبحركة عين الفعل في الماضي وفي المضارع والشّخص إلخ. ومعلوم أنّ هذه المعطيات على الشّكل الذي قُدِّمت عليه هنا ليست موجَّهَة للمستعملين كما ذكرنا. ولذلك يحتوي برنامج المحلّل على مكوّن يترجم هذه المعلومات في شكل قابل للاستغلال بالنّسبة إلى المستعمل "العاديّ" عن طريق واجهة التّحاور مع البرنامج. فيمكن أن

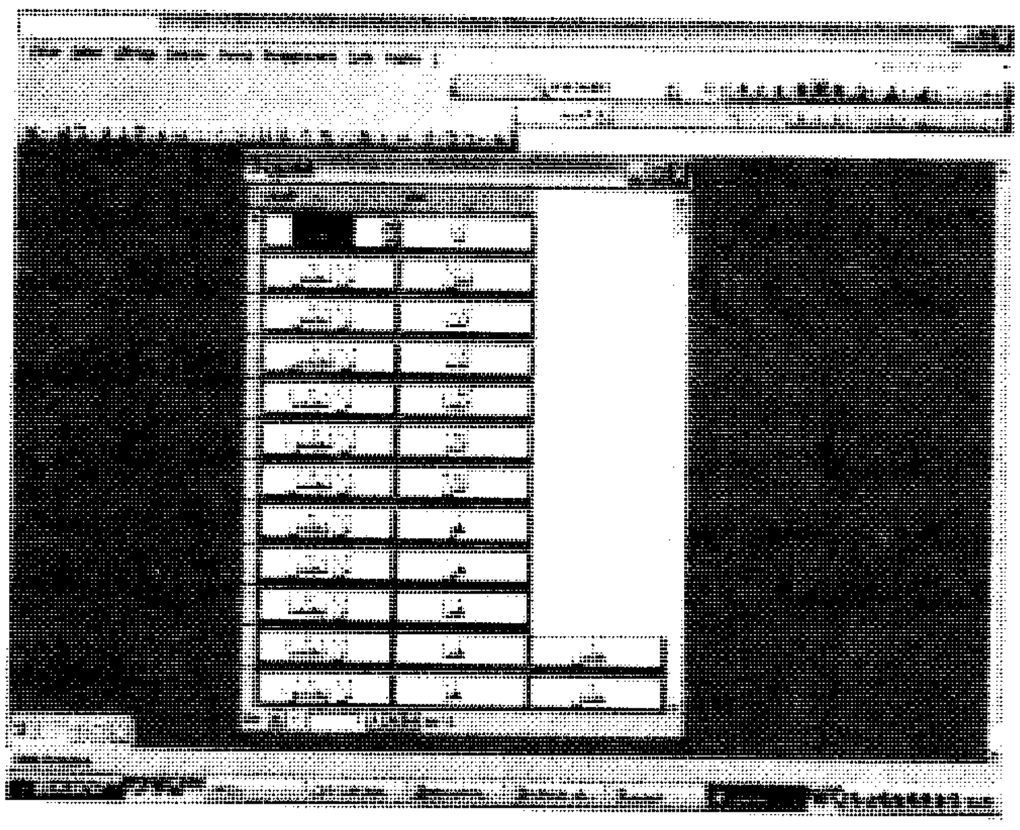
---

1) انظر في هذه المسائل: ط. البكّوش (1992).

يكتفي البرنامج، عند طلب التّصريف فحسب (الفعل **جلس** في المضارع المنصوب مثلا)، بعرض النتيجة كما يلي:



ويُعرَض ذلك بشكل من الأشكال في الواجهة الأماميّة للبرنامج، وفق نظام من الوظائف التّفاعليّة بين المستعمل والبرنامج:



ولئن كان برنامج تصريف الأفعال لا يطرح صعوبات كبرى بفضل خاصّيّة الانتظام التي تميّزه، فإنّ قضايا كثيرة تُطرَح، لسانيّا وإعلاميّا بالنّسبة إلى بقيّة مكوّنات المحلّل الصّرفيّ. وسنعرض لأهمّها في العنصر الموالي.

3. عدد من الإشكاليّات

نتطرّق في هذا العنصر الأخير إلى عدد من القضايا المتعلّقة بالمحلّل الصّرفيّ. ونركّز فيها أساسا على جوانب الوصف اللّسانيّ والعلاقة بالمرحلتين الأخريين من المعالجة الآليّة، أي المحلّل التّركيبيّ والمحلّل الدّلاليّ.

أ. الأشكال غير القياسيّة

ذكرنا أعلاه أنّ النّظام الصّرفيّ في العربيّة على درجة كبيرة من النّظاميّة. إلاّ أنّه من الضّروريّ حلّ مشاكل الظّواهر غير القياسيّة. ومن أهمّها صياغة المصادر المتّصلة بالثّلاثيّ المجرّد. ويبدو أنّه لا مفرّ من جرد هذه الأشكال انطلاقا من القواميس والمعاجم ووسمها يدويّا في مرحلة أولى. ولا بدّ من أعمال وصف لسانيّ في هذا السّياق تتناول هذا الجانب وتتثبّت من نسبة تغطية بعض الإشارات التي نجدها عند النّحاة في هذا المجال. من ذلك اختصاص وزن ما من مصادر الثّلاثيّ المجرّد بدلالة معيّنة (فِعالة: مهَن، فُعال: مرض، فَعَلان: حركة واضطراب).

ونشير في هذا الصّدد إلى أنّنا درسنا عيّنة من مداخل **المعجم الوسيط** (من الهمزة إلى الطّاء) الواردة على وزن **فِعالة** والتي يُفتَرَض أنّها تدلّ على حرفة/ صناعة. وقد لاحظنا أنّه من مجموع 105 تدلّ 55 فقط على حِرفة. وبناء على ذلك تتأكّد في نظرنا ضرورة الدّراسة الميدانيّة المنتظمة لهذه الظّواهر لاسيّما إذا ما وضعنا في الاعتبار أهمّيّة هذه المعلومات بالنّسبة إلى صياغة المحلّل الدّلاليّ في مراحل لاحقة.

ب. معالجة الأشكال المتلاصقة

معلوم أنّ العربيّة لسان مُدمِج كثيرا ما يؤلّف بين الوحدات المعجميّة في الكتابة على صورة أشكال متلاصقة لا يفصل بينها بياض. وقد يبلغ ذلك حدّ ورود جمل كاملة في شكل واحد متلاصق على غرار الجملة الاستفهاميّة :

أرَأيْتَهُمَا؟

حيث يدمج هذا الشّكلُ الفعل والفاعل والمفعول به إضافة إلى همزة الاستفهام الدّاخلة على الجملة. لا بدّ إذن من أن يأخذ المحلّل التّركيبيّ في الاعتبار ظاهرة الأشكال المتلاصقة وأن يعتمد في ذلك على المحلّل الصّرفيّ في جانب كبير.

ويجدر أن نشير في هذا الصّدد إلى وجود عدد من المشاريع والمحاولات لصياغة محلّلات تركيبيّة للجمل في العربيّة. إلاّ أنّنا نكتفي

هنا بالإشارة إلى أنّ أيّ برنامج تحليل صرفيّ لا بدّ له- في رأينا- من أن يستند إلى برنامج تحليل نحويّ من جهة، ولا بدّ له من ناحية ثانية من أن يعتمد أرضيّة نظريّة تعرّف مفهوم الجملة باعتبارها وحدة التّحليل الدّنيا وتضبط العلاقات بين مكوّناتها[1].

ج. الأشكال المركَّبة

نقصد بذلك ظواهرَ التّركيبيّة المتكلّسة بصفة عامّة ومختلف الأنماط التي تتوزّع عليها :

- أسماء (مفكّ براغ، هزّة أرضيّة، غارة جويّة، الخ)؛
- صفات (بهيّ الطّلعة، ثقيل الظّلّ، الخ.)؛
- أفعال (أشعل فتيل الأزمة، وقع في شرّ أعماله، الخ.)؛
- حروف/ أدوات (بصدد، بهدف، في حلّ من كذا، الخ.).

ولا شكّ أنّ هذه الظّواهر لا بدّ أن تكون موضوعَ أعمال وصف لسانيّ تتناول بالدّرس مدوّنات استعمال لغاية ضبط القائمات وصياغة مقاييس التّصنيف. وينتج عن هذه الأعمال معاجم هي قائمات تضمّ مختلف أصناف الأشكال المركّبة وخصائصها. وهي قائمات يستعملها المحلّل الصّرفيّ للتّعرّف إلى هذه الوحدات في المدوّنات ووسمها وعدم الخلط بينها وبين الوحدات الحرّة.

د. مسألة الحركات

نذكّر بأنّ العديد من مشاريع التّحليل الصّرفيّ المقترَحة لا تأخذ في الاعتبار رسم الحركات ونؤكّد أنّه لا قيمة لمحلّل صرفيّ لا يأخذ في الاعتبار هذا المعطى، لا سيّما وأنّ:

- القسم الأكبر من النّصوص المتوفّرة (رقميّة وغيرها) لا نجد فيه رسما للحركات؛
- الغالب على المؤلّفات الأدبيّة مثلا حالة وسطى هي ما يسمّى بـــ"الشّكل الوظيفي".

---

1) يعتمد فريقنا مثلا تعريف الجملة عند هاريس (1964، 1976) على النّحو الذي تواصل به في إطار أعمال فريق المعجم- النّحو ومخبر المعاجم والقواميس والإعلاميّة. ومن أهمّ المنطلقات النّظريّة التي نعتمدها: الجملة وحدة دنيا للوصف والتّحليل؛ تحدّد الجمل بما هي عمليّة إسناد تربط علاقة بين مسند ومعمولاته؛ تُرَدّ الوحدات المعجميّة إلى إحدى وظائف ثلاث: المسند، المعمول، المحيّن. ونحيل في هذا الصّدد إلى مجمل الأعمال المنجزة في إطار مخبر المعاجم والقواميس والإعلاميّة.

من الضّروريّ إذن أن يأخذ المحلّل في الاعتبار جميع هذه الحالات ويدخل في الحساب مختلف الإمكانيّات.

كيف نتعامل مع النّصوص غير المشكولة؟

– في مرحلة أولى لا يمكن للمحلّل الصّرفيّ إلاّ أن يعدّد مختلِف الإمكانات التي يمكن أن ترجع إليها وحدة غير مشكولة كما رأينا أعلاه (العنصر 1. بالنّسبة إلى "كتب" مثلا)؛

– ثمّ : التّعويل على المحلّل التّركيبيّ الذي "يضيّق" الاحتمالات على أساس الجوار المباشر للوحدة وباعتماد معطيات مختلفة (إحصائيّة، تخطيط معمولات، التواردات المشتركة، إلخ.).

هـ. العلاقات بين مختلف الأصناف الاشتقاقيّة (الاشتقاق الدّلاليّ)

إنّ البنية الهندسيّة المحكمة (عموديّا وأفقيّا) التي يظهرها نظام الاشتقاق في العربيّة قد لا يطابقها الاستعمال. ولئن كان النّظر في آليّة الاشتقاق من منظور قضايا الاشتقاق الدّلاليّ ليس من مشمولات برنامج التّحليل الصّرفيّ، فإنّنا نعتقد أنّها قضايا لا بدّ من أخذها في الاعتبار من البداية. ذلك أنّه من منظور الاستعمال يمكن أن نجد مشتقّات تعود إلى:

- أشكال مشتقّة من أشكال أخرى هي بدورها مشتقّة أو غير مشتقّة :

أقام أربعينيّة عليّ => أربعون(ين)

مأتم، مأدُبة، إلخ.

- أشكال متّصلة اشتقاقيّا لكنّها في قطيعة تركيبيّة–دلاليّة من حيث البنية الإسناديّة التي تنتظم فيها. فإذا انطلقنا من المادّة **(ع، م، ل)** أمكن أن نصوغ المشتقّات التّالية وأن نقول إنّ بينها علاقات اشتقاقيّة:

– أفعال : عمِل

– صفات : عامِل، عَمِيل.

– أسماء : عَمَل، عمليّة

أمّا إذا تفحّصنا هذه المشتقّات نفسها في الاستعمال فإنّنا نلاحظ وجود حالات مختلفة من حيث العلاقات التّركيبيّة–الدّلاليّة:

– فعل :

عمِل سامي في هذه الشّركة

– صفتان

(1) سامي عامل في هذه الشّركة

(2) سامي عَميل (لـــ + متمّم)

– اسمان

(3) قام سامي بعمَله

(4) أجرى سامي عمليّة (+ جراحيّة/ سبر آراء...)

ويمكن أن نبدي من هذا المنظور الملاحظات التّالية:

✓ الجملة القائمة على الفعل "عَمِل" والجملة (1) القائمة على الصفة "عامِل" والجملة (3) القائمة على الاسم "عمَل" متكافئة = مجال العمل (الشّغل، الحرفة...)؛

✓ نسجّل في (2) "عميل" و(4) "عمليّة" انتقالا من مجال إلى آخر بحيث لا يمكن الحديث عن علاقات تكافؤ.

✓ ظاهرة الاختصاص بمجال معيّن هي من أهمّ الظّواهر المتحكّمة في البنية العامّة لنظام الاشتقاق في الاستعمال.

هكذا إذن نرى أنّه لا بدّ من دراسة الاستعمال بشكل منتظم استقصائيّ، ولا بدّ لصياغة برنامج التّحليل الصّرفيّ أن تأخذ في الاعتبار هذه القضايا المركزيّة من البداية، وأن يُنظر إلى مجموع المحلّلات (الصّرفيّ، التّركيبيّ، الدّلاليّ) على أساس التّكامل لا القطيعة.

4. خاتمة

نودّ أن نذكّر أخيرا بالنّقاط التّالية:

– لا بدّ للمحلّل الصّرفيّ من صفة الشّموليّة بالنّسبة إلى الوحدات المستعمَلة، وبالنّسبة إلى ما يُعتبَر طاقة إنتاج كامنة في معجم اللّسان.

– من الضّروريّ أن يستند هذا العمل إلى أعمال وصف لسانيّ منتظم وأن يضع في الاعتبار التّكامل بين مختلف مراحل المعالجة الآليّة للعربيّة؛

– تستند مثل هذه المشاريع إلى أعمال كثيرة تطبيقيّة على مدوّنات استعمال فعليّة. ومن ثمّ الحاجة الأكيدة إلى الوصف المنتظم والاستقصائيّ؛

– نشير أخيرا إلى أنّ العمل على المحلّل الصّرفيّ وفق المنهج الذي رسمناه قد بلغ أشواطا متقدّمة بفضل الجهود المتضافرة بين اللّسانيّين ومختصّي الإعلاميّة في مخبر المعاجم والقواميس والإعلاميّة الذي سيضع عن قريب نسخة تجريبيّة على ذمّة المستعملين على موقعه الإلكترونيّ.

وعلى هذا الأساس نؤكّد ضرورة إنجاز مثل هذه الأعمال في إطار فرق بحث نظرا إلى النّفس الطّويل الذي تتّسم به والجهد الجماعيّ الذي تتطلّبه.

قائمة المصادر والمراجع

عصفور (ابن)، **الممتع في التّصريف.**

الزّمخشري، **المفصّل.**

الخليل بن أحمد الفراهيدي، **كتاب العين.**

المعجم الوسيط

البكّوش (الطيب)، **التّصريف العربي من خلال علم الأصوات الحديث**، تونس، 1992.

فارس (ابن)، **مقاييس اللّغة.**

منظور (ابن)، **لسان العرب.**

الورهاني (بشير)، **الأفعال النّاقلة في العربيّة المعاصرة: بحث في الخصائص التّركيبيّة والدّلاليّة.** كلّيّة الآداب والعلوم الإنسانيّة بسوسة، 2008.

ALOULOU C. (2003), « Analyse syntaxique de l'arabe : le système MASPAR », in *Actes de RECITAL 2003*, Batz-sur-Mer, France. pp 867- 877.

COUTOIS B., 1990, « Un système de dictionnaires électroniques pour les mots simples du français », *Langue Française* 87, Paris : Larousse.

Haddad A., Ben Ghezala H., Ghenima M., « Conception d'un catégoriseur morphologique fondé sur le principe d'Eric Brill dans un contexte Multi-Agents ». 26th conference on Lexis and Grammar, Bonifacio, 2-6 October 2007.

MEJRI S., 1997, *Le figement lexical.* Université de Tunis.

Mouelhi Z. 2008, « *AraSeg*: un segmenteur semi-automatique des textes arabes », JADT 9es Journées internationales d'Analyse statistique des Données Textuelles.

OUERHANI B. & AMRI H., à paraître, « Actualisation des racines arabes et génération automatique des unités lexicales », 2ème colloque des doctorants et jeunes chercheurs, *L'actualisation prédicative et argumentale*, Paris, 19-22 décembre.

TAHIR Y., NOUREDDINE Ch. et MOSTAPHA H., « Modélisation à objets d'une base de données morphologiques sur la langue arabe », JEP-TALN 2004, Traitement automatique de l'arabe. Fès, 2O avril 2004.

سرعة لكن ذلك يكون على حساب دقة النتائج. ويمكننا إلغاء مرحلة اختيار كيفية الإصلاح أي جعله آليا، إذا طوّرنا قدرات مورف 2 وذلك بجعله يرتّب آليا الاحتمال المناسب لكل كلمة في الجملة وحلّ حالات اللبس (les cas d'ambiguïté).

2-2 المرحلة الثانية : التحليل التركيبي لكشف الأخطاء

تنقسم هذه المرحلة إلى ثلاث مراحل أخرى وهي تعمل بطريقة متسلسلة كالآتي: مرحلة تقسيم الجملة إلى مركّبات أوّليّة (syntagmes initiaux) ومرحلة تجميع المركّبات الأوّليّة إلى مركّبات ممتدّة (syntagmes étendus) حسب نوع الخطأ المراد إصلاحه وأخيرا مرحلة كشف الأخطاء.

2-2-1 مرحلة تقسيم الجملة إلى مركّبات أوّليّة

تهتمّ هذه المرحلة بتقطيع الجملة إلى عدّة مركّبات أوّليّة اعتمادا على مجموعة من الكلمات الأدوات (mots outils) التي نسمّيها الحدود التركيبيّة (frontières syntagmatiques) ونذكر منها: حروف الجرّ، حروف العطف، أدوات النفي، أدوات الاستفهام، أدوات النصب، أدوات الجزم، أدوات الاستثناء، ظروف المكان، ظروف الزمان، كان وأخواتها، إنّ وأخواتها، كاد وأخواتها، أفعال الظّن، أفعال الرّجاء، أفعال المقاربة والشّروع، الأسماء الموصولة، أسماء الإشارة، الضّمائر المتّصلة، الضّمائر المنفصلة، إلخ...

وتعتمد عمليّة التقطيع على مجموعة من القواعد بحيث تتحصّل في النّهاية على مجموعة من المركّبات الأوّليّة جاهزة للتجميع.

2-2-2 مرحلة تجميع المركّبات الأوّليّة إلى مركّبات ممتدّة

في هذه المرحلة يتمّ تجميع المركّبات الأوّلية التي توجد بينها علاقة تطابق، إلى مركّبات ممتدّة حسب نوع الأخطاء المراد إصلاحها. ويقوم نظام ديكورا 2 بثلاث عمليّات تجميع اعتمادا على مجموعة من القواعد استخلصناها من معاجم قواعد اللغة العربيّة (الدحداح، 1989) (شمس الدين، 2000) :

- التجميع الأوّل: وهو بناء المركّبات الممتدّة لكشف أخطاء الإعراب.
- التجميع الثاني: وهو بناء المركبات الممتدة لكشف أخطاء التعريف.
- التجميع الثالث: وهو بناء المركبات الممتدة لكشف أخطاء في الجنس والعدد والضمير وصفة العاقل.

2-2-3 مرحلة كشف الأخطاء

بعد بناء المركبات الممتدّة، يقوم نظام ديكورا 2 عن طريق برنامج كاشف الأخطاء بعمليّة التدقيق ويكون ذلك استنادا إلى الخصائص الصرفية للكلمات التي تكوّن المركّب الممتد، واعتمادا على تطبيق قواعد اللغة العربية بالنسبة إلى تجميعين الأول والثاني (ما هي الخصائص؟ وكيف يجب أن تكون؟) وعملية توحيد الخصائص (unification) بالنسبة إلى التجميع الثالث (يجب أن تكون خصائص الكلمات متساوية).

وفي حالة وجود خلل في إحدى خطوط التطابق (أي الخصائص)، بمعنى أن الخاصّيّة الموجودة مختلفة عمّا يجب أن تكون عليها حسب القاعدة المطبقة وذلك بالنسبة إلى مركبين الممتدين الأول والثاني أو وجود اختلاف في الخصائص بالنسبة إلى المركّب الممتد الثالث، يقوم برنامج كاشف الأخطاء بتمريرها إلى برنامج مصلح الأخطاء. أما في حالة عدم وجود خلل، يقع محو المركب الممتدّ نهائيّا لخلوّه من الأخطاء.

2-3 المرحلة الثالثة: التّحليل المتعدّد المعايير لإصلاح الأخطاء

تنجز هذه المرحلة اعتمادا على برنامج مصلح الأخطاء الذي يقوم بتوليد الجملة الصحيحة المناسبة (احتمال واحد) بالنسبة إلى أخطاء الإعراب وتوليد جميع الاحتمالات الممكنة بالنسبة إلى بقيّة الأخطاء. عند توليد احتمال، يقوم البرنامج باستبدال بعض الخصائص الصرفيّة للكلمات التي تخلّ بقواعد التطابق ويمكن لهذه الكلمة أن تكون اسما أو صفة أو فعلا.

في حالة استبدال الخصائص الصرفيّة لاسم أو صفة، يقوم برنامج مصلح الأخطاء بإرسال الاسم أو الصفة في صيغة المفرد إلى برنامج مصيغ الكلمات صحبة الخصائص التي يجب أن تكون عليها أي الجنس والعدد المناسبان. أما في حالة استبدال الخصائص الصرفيّة لفعل ما، يقوم برنامج مصلح الأخطاء بإرسال أصل الفعل إلى برنامج مصرّف الأفعال بمعيّة الخصائص التي يجب أن يكون عليها، أي حالة الفعل (معلوم أو مجهول) والضمير والزّمان المناسبين. وسنتوسّع لاحقا في وصف هذين البرنامجين المهمّين في عمليات توليد الاحتمالات.

ويجدر بنا أن نذكر أنه في حالة وجود العديد من الاحتمالات المتنافسة، يقوم برنامج مصلح الأخطاء بالتوازي مع عمليات إنتاج الاحتمالات باحتساب وزن كل احتمال وذلك قصد التفريق بين مختلف الإحتمالات وترتيبها ترتيبا تنازليا ثمّ عرضها في لائحة احتمالات الإصلاح لمساعدة مستعمل البرنامج على اختيار الاحتمال الأكثر صوابا.

ويتمّ احتساب وزن أيّ تغيير يطرأ على إحدى الكلمات التي تكوّن الجملة المراد إصلاحها، وذلك حسب ثلاثة معايير (تمّ اقتراح هذه المعايير في نطاق أطروحة الدكتوراه التي قدّمتها د. لمياء هدريش بلغيث سنة 1999) (BELGUITH,1999) :

• المعيار الأوّل وهو المعيار التردّدي: (critère de fréquence) يقترح الاحتمال الذي يغيّر أقل ما يمكن من الخصائص الصرفية للكلمات بمعنى أنّه يحافظ على الخصائص الأكثر تردّدا. ويكون عدد الخصائص هو العدد المسند لهذا المعيار. وهو معيار للتعظيم بمعنى أن قيمته سوف تضرب في العدد 1 (critère à maximiser).

• المعيار الثاني وهو المعيار الرّقني: (critère typographique) يهتم بالأخطاء اللفظيّة (erreurs lexicales) النّاتجة عن نقص أو زيادة حرف من الحروف أو استبدال حرف بآخر أو قلب مواقع حرفين متتاليين، وينتج عن هذا الخطأ كلمة صحيحة لكنّها تخرق قاعدة من قواعد التطابق. ويسند لكل نوع من الأخطاء الرّقنيّة عددا بين 0 و1. ويكون هذا المعيـــار للتقليل بمعنى أن قيمته سوف تضرب في لعدد (-1) (critère à minimiser).

• المعيار الثالث وهو المعيار الصرفي: (critère morphologique) يشجّع على تغيير الخصائص الصرفيّة للكلمة الخاطئة دون استبدالها بكلمة أخرى لا تقاسمها نفس الجذر. نأخذ على سبيل المثال كلمة "ولد" في صيغة المذكّر وكلمة "بنت" في صيغة المؤنّث. ويسند لهذا المعيار العدد 1 إذا تغيّر أصل الكلمة و0 في الحالة الأخرى. ويعتبر هذا المعيار للتقليل أيضا.

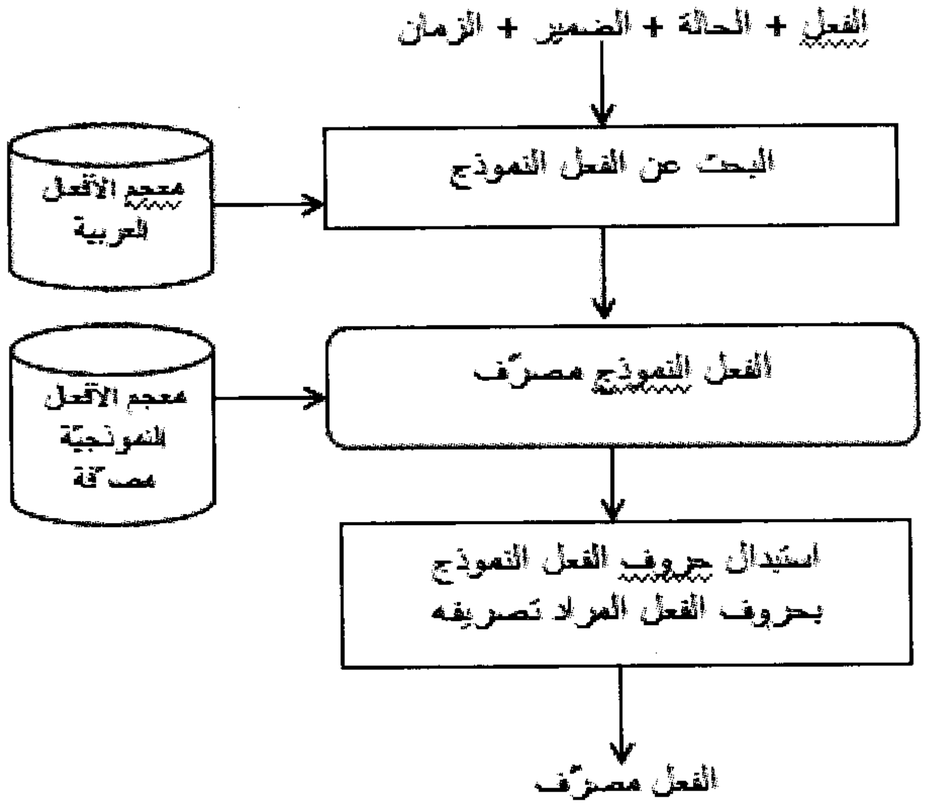
وفي النهاية يحتسب الوزن الجملي لكل احتمال وترتّب هذه الأخيرة تنازليّا في قائمة الاحتمالات المقترحة للمستعمل.

3- برنامج مصرّف الأفعال

يعتبر هذا البرنامج بالغ الأهمّيّة لا يستطيع نظام ديكورا 2 الاستغناء عن خدماته. فهو الذي تُوكل إليه مهمّة تصريف الأفعال ويستعمل من طرف برنامج مصلح الأخطاء الذي يقوم بإرسال الفعل صحبة الحالة التي يجب أن يكون عليها والضمير والزمان. فيتسلّمها هذا البرنامج ويقوم بتصريف الفعل ومن ثمّ إرجاعه مصرّفا.

وقد قمنا بعمليّة مسح في معاجم تصريف الأفعال العربية (الدّحداح، 1999) (ديشي، 1999) (وهي تتناول الأفعال العربيّة مشكولة) وشرعنا في

عمليات مقارنة بين أنماط تصريف الأفعال النموذجية فوجدنا أن كثيرا منها تتشابه إذا أزلنا عنها الشكل. وبهذه الخطوة حصرنا مختلف أنماط الفعل العربي غير المشكول ووجدنا 216 فعلا نموذجا تجسّد نمط تصريف 14000 فعل عربي، كما قمنا بإدخالها جميعا في قاعدتي البيانات المستخدمة لتحقيق صفة التكامل والشمول في هذا البرنامج. ويبيّن الرّسم 2 كيفية تصريف الفعل.



الرّسم 2 : مراحل تصريف فعل

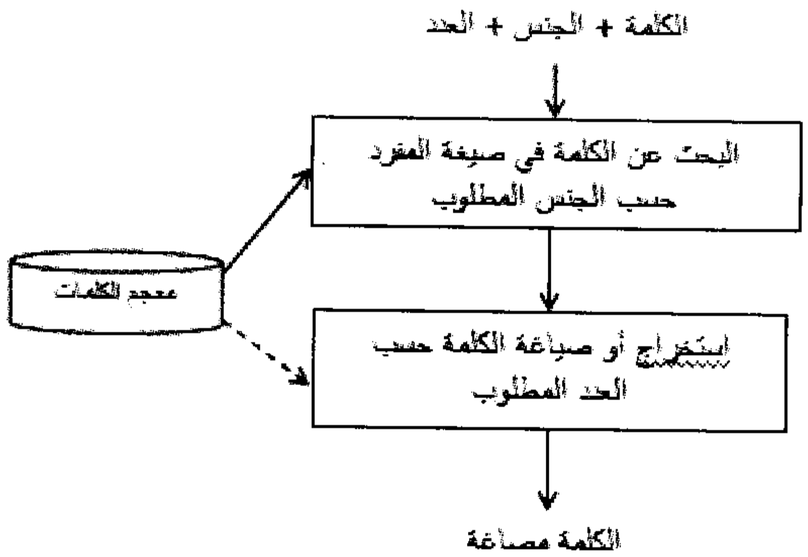
وتتمّ عملية تصريف فعل بالبحث عن فعله "النموذج" في قاعدة معجم الأفعال العربية ومن ثمّ إيجاد الفعل النموذج مصرّفا في قاعدة معجم الأفعال النموذجيّة مصرّفة، ويتمّ بعد ذلك استبدال حروف الفعل النموذج بحروف الفعل المراد تصريفه.

إنّ برنامج مصرّف الأفعال يعدّ إنتاجا كاملا يمكن إدماجه في أنظمة أخرى. وقد تمّ انجازه في مخبر MIRACL وطمحنا من خلاله إلى تحقيق هدفين رئيسيين هما: أولا، مساعدة ديكورا 2 على إنتاج اقتراحات الإصلاح. ثانيا، مدّ هذا البرنامج لكلّ الباحثين الذين يريدون مصرّفا للأفعال يساعدهم لربح الوقت والاستفادة منه.

4- برنامج مصيغ الكلمات

يعتبر هذا البرنامج مهمّا أيضا، فهو الذي توكل إليه مهمة صياغة الأسماء والصفات، ويستعمل من طرف برنامج مصلح الأخطاء الذي يقوم بإرسال الكلمة صحبة الجنس والعدد المناسبين ثمّ يتم بعد ذلك صياغتها.

وصياغة الكلمة تكون سهلة في حالتي المثنى والجمع السّالم ولكنها ليست بالأمر الهيّن في حالة جمع التكسير، فاللغة العربية ليس فيها قواعد ثابتة وتعتمد على السّماع أو ما-قالته العرب. ولاجتياز هذه المسألة الصعبة، قمنا بتخزين الكلمات في صيغة المفرد صحبة صيغتها في الجمع وصيغتها في المؤنث (أو المذكر) إذا كانت الكلمة في صيغة المذكّر ومرادفتها في صيغة المؤنّث ليستا مشتقتين من نفس الجذر. ويبيّن الرّسم 3 كيفية صياغة كلمة.



الرّسم 3: مراحل صياغة كلمة

وتتمّ عملية صياغة الكلمة بالبحث عن الكلمة المناسبة حسب الجنس المطلوب وذلك في قاعدة معجم الكلمات ومن ثمّ النظر إلى العدد المطلوب. إذا كان العدد في المفرد، تؤخذ الكلمة كما هي. وإذا كان العدد في المثنى تكون صياغتها سهلة حيث أنّه يكفي إضافة الحرفين "ان" إلى الكلمة. أمّا إذا كان العدد في الجمع، يقع الرجوع إلى المعجم لاستخراج الكلمة في صيغة الجمع. عندما تكتمل صياغة الكلمة يقع إرجاعها إلى برنامج مصلح الأخطاء الذي يتسلّم الكلمة مرفوعة ثمّ يجرّها أو ينصبها حسب حالة إعرابها في الجملة.

5- مثال لتشغيل نظام ديكورا 2

أنجز ديكورا 1 سنة 1999 في نسخته الأولى بلغة البرمجة (BELGUITH, 1999) ويعالج هذا النظام الأخطاء في الجنس والعدد والضّمير

فقط، ثم وقع تطويره وذلك بإضافة أخطاء التعريف والإعراب وصفة العاقل. وتمّتإعادة برمجته بلغة "جافا" Java. ويتم تبادل المعلومات بين مختلف البرامج المدمجة أو تسجيل البيانات بواسطة ملفّات "إكس أم أل" XML. وتعتبر لغة البرمجة جافا واستعمال ملفات "إكس أم أل" من آخر التقنيات التي توصّل إليها العلم في عالم البرمجة.

لنأخذ على سبيل المثال الجملة الخاطئة "*الولد صغير تلعب ويضحك*". وتبين النّافذة 1 نتائج كشف الأخطاء وهي على التوالي:

الخطأ الأوّل: خطأ في التّعريف بين كلمة «ولد" وكلمة «صغير"

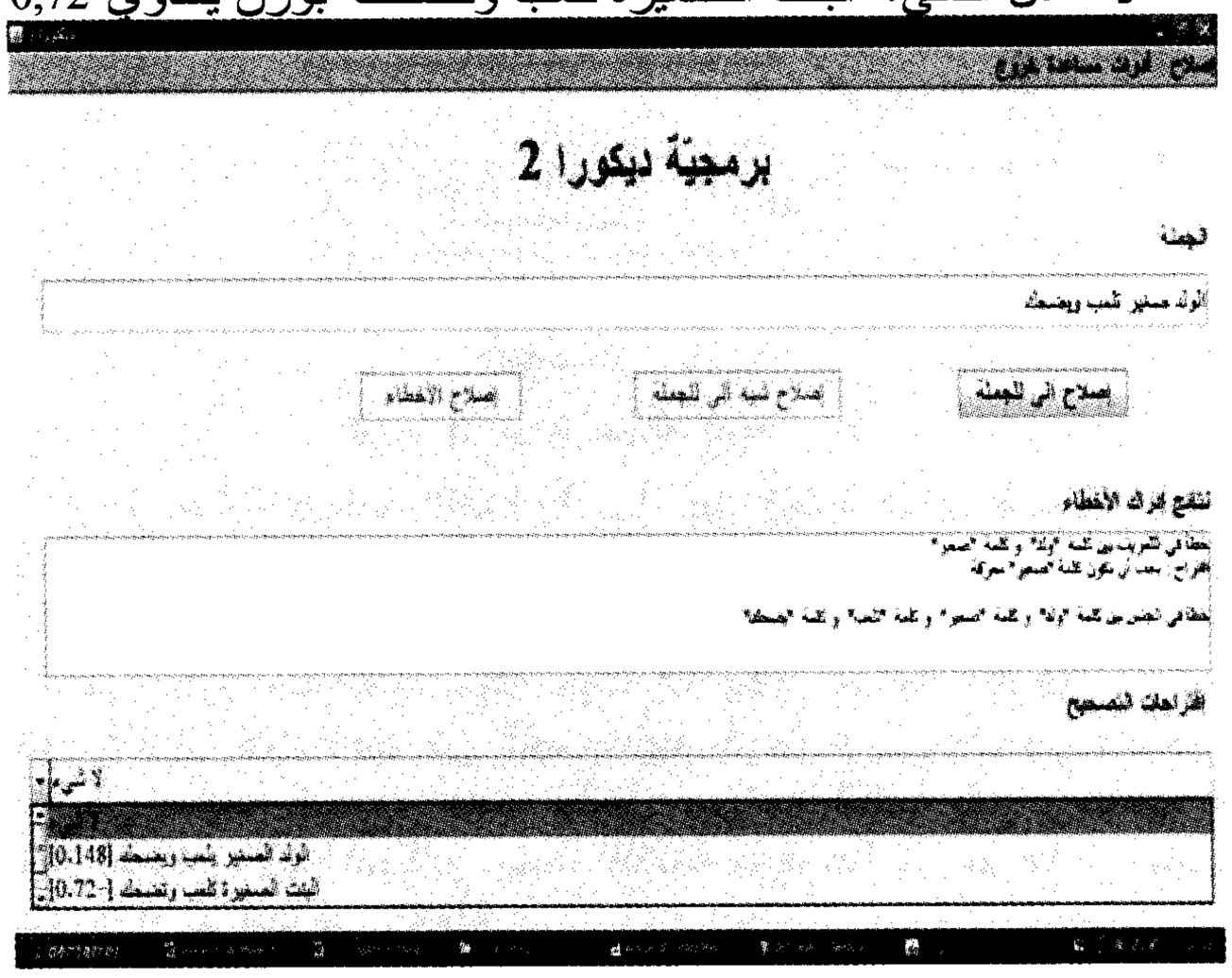
اقتراح : يجب أن تكون كلمة "صغير" معرّفة

الخطأ الثّاني : خطأ في الجنس بين كلمة "ولد" وكلمة "صغير" وكلمة "تلعب" وكلمة "يضحك"

كما تبين هذه النّافذة احتمالات إصلاح الجملة مرتّبة حسب أوزانها وهي:

الاحتمال الأول: "الولد الصغير يلعب ويضحك" بوزن يساوي 0,148

الاحتمال الثّاني: "البنت الصغيرة تلعب وتضحك" بوزن يساوي 0,72 –



النّافذة1:عرض لبرنامج ديكورا2

ونلفت الانتباه أنّه قد اخترنا في هذا المثال طريقة الإصلاح الآلي، أمّا إذا أردنا أن نختار بأنفسنا الخصائص الصرفيّة لكل كلمة فيتحتّم علينا اختيار الإصلاح شبه الآلي، حيث يعرض علينا ديكورا 2 نتائج المحلل الصرفي مورف 2 كما هو مبيّن في النّافذة 2. ويتمّ اختيار الخصائص بطريقة يدوية من طرف المستعمل.

رقم	المحتوى	النوع	أصل الكلمة	الجنس	العدد	الضمير	عاقل	الصيغة	الزمان	معرّف
1	ولد	اسم : اسم	ولد	مذكر	مفرد		عاقل			معرف
2	صغير	صفة_مشبهة...	صغير	مذكر	مفرد		عاقل/غير عاقل			غير معرف
3	تلعب	فعل : غير ناسخ	لعب	مؤنث	مفرد	الغائب		معلوم	مضارع	
3	تلعب	فعل : غير ناسخ	لعب	مؤنث	مفرد	الغائب		مجهول	مضارع	
3	تلعب	فعل : غير ناسخ	لعب	مذكر	مفرد	المخاطب		معلوم	مضارع	
4	و	حرف عطف								
4	يضحك	فعل : غير ناسخ	ضحك	مذكر	مفرد	الغائب		معلوم	مضارع	
4	و	حرف عطف								
4	يضحك	فعل : غير ناسخ	ضحك	مذكر	مفرد	الغائب		مجهول	مضارع	

النّافذة2:عرض لنتائج مورف2

6- تقييم نظام ديكورا 2

لتقييم نظام ديكورا 2، قمنا باختباره على حوالي 300 جملة مستخرجة من مواضيع إنشاء لتلامذة التعليم الأساسي والثانوي، ومن مطالب بعض المواطنين لدى البلديات ولدى شركة اتصالات تونس وسعينا إلى تنويع المستوى الثقافي للأشخاص.

وقد قسّمنا الأخطاء إلى ثلاثة أقسام:

- الأخطاء الحقيقيّة المعالجة ببرنامج ديكورا 2. ونرمز إلى عددها بحرف ق
- الأخطاء الحقيقيّة التي لم يستطع ديكورا 2 إدراكها. ونرمز إلى عددها بحرف "ر"
- الأخطاء غير الحقيقيّة المضافة من قبل ديكورا 2. ونرمز إلى عددها بحرف "ب"

ويكون الخطأ حقيقيّا عندما يعرفه أخصائي في اللغة العربية.

وتحتسب معايير التقييم (وهي الشّموليّة RAPPEL والتدقيق PRECISION) على النحو التالي :

الشّموليّة = ق / (ق + ر)

التدقيق = ق/ (ق + ب)

وقد توصلنا إلى النتائج التالية :

ق = 18 ؛ ر = 15 وب = 3

وبالتالي تكون المعايير: الشّموليّة = 54,5% والتدقيق = 85,7%

ويمكن القول بأن النّتائج المحققة مشجّعة. وتفسّر بعض النقائص بوجود حالات تتجاوز مستوى التحليل الصرفي إلى مستوى التحليل التركيبي.

7- خاتمة

قدمنا في هذه الورقة برنامج ديكورا 2 الذي أنجز في إطار عمل ماجستير. وقد توصلنا إلى نتائج مشجّعة. ونطمح في الأخير إلى إدخال المزيد من التّحسينات على ديكورا 2 حتّى يتمكّن من إصلاح نصوص كاملة وليس جملا معزولة ولا يكون ذلك ممكنا إلا بتحسين نظام مورف 2 الجاري إنجازه حتّى الآن في مخبر MIRACL وإدماج برنامج يقسّم بطريقة ناجعة النص إلى جمل معالجا أغلب حالات اللبس ومشاكل أخرى متعدّدة في تحليل اللغة العربية.

كلمة شكر : نتقدّم بالشكر الجزيل إلى الأستاذ عبد الحميد عبد الواحد الذي أتاح لنا ــ نحن الحاسوبيين ــ فرصة التعرّف على الباحثين اللّسانيين وتبادل الأفكار والمعلومات بيننا. ونعترف إليه بالجميل لما قدّمه لنا من مساعدات لإنجاز هذه الورقة.

قائمة المصادر والمراجع

❖ (الدحداح، 1989) أنطوان الدحداح، **معجم قواعد اللغة العربيّة في جداول ولوحات**، مكتبة لبنان، الطبة الرّابعة، بيروت 1989.

❖ (الدحداح، 1999) أنطوان الدحداح، **معجم تصريف الأفعال العربيّة**، مكتبة لبنان، الطبة الرّابعة، بيروت 1999.

❖ (ديشي، 1999) يوسف ديشي وسامي عمّار، **الشّامل في تصريف الأفعال العربيّة**، هاتييه (Hatier)، باريس، أكتوبر 1999.

❖ (شمس الدين، 2000) شمس الدين، **مرجع الطلاب في الإعراب**، دار الكتب العلميّة، الطبعة الأولى، بيروت 2000.

❖ (BELGUITH, 1999) BELGUITH HADRICH Lamia, « Traitement des erreurs d'accord de l'arabe basé sur une analyse syntagmatique étendue pour la vérification et une analyse multicritère pour la correction »,Thèse de doctorat en informatique, Faculté des sciences de Tunis, février 1999.

❖ (BELGUITH, 2004) BELGUITH HADRICH Lamia, BEN HAMADOU Abdelmajid, «Traitement des erreurs d'accord : une analyse syntagmatique pour la détection et une analyse multicritère pour la correction ». RSTI-RIA, volume 18 – N°5-6/2004, pages 679 à 707.

❖ (BOUJELBEN, 2005) BOUJELBEN Makram, BELGUITH HADRICH Lamia et ALOULOU Chafik, « Détection des erreurs d'accord dans des textes arabes non voyellés », Cinquièmes journées scientifiques des jeunes chercheurs en Génie Electrique et Informatique : GEI'2005, Sousse, Tunisie, 2005.

❖ (BOUJELBEN, 2008) BOUJELBEN Makram, ALOULOU Chafik et BELGUITH HADRICH Lamia, « Toward a robust detection/correction system for the agreement errors in non-voweled Arabic texts», The International Arab Conference on Information Technology, ACIT'2008, December 16th- 18th 2008, Hammamet, Tunisia.

❖ (CHAABEN, 2004) CHAABEN Nouha, « Conception et réalisation d'un analyseur morphologique pour l'arabe non voyellé », Mémoire de mastère en informatique, Ecole Nationale d'Ingénieurs de Sfax, Juillet 2004.