

قواعد بيانات القرآن الكريم كأساس للمعجم الآلي الموسع للغة العربية

أ.د. محمد زكي خضر
الجامعة الأردنية

الخلاصة

لا بد للمعجم الآلي للغة العربية أن يحتوي مفردات جامعة ونصوصاً رصينة . وليس هناك ما يحقق هاتين الخاصيتين أفضل من نصوص القرآن الكريم . هذا بالإضافة إلى أن الكلمات والتعبير الأكثر استعمالاً في اللغة العربية واردة في القرآن الكريم في غالبيتها العظمى .

يحاول البحث أن يبين أسلوب تكوين قواعد البيانات الخاصة بالقرآن الكريم بدءاً من إعداد النص القرآني من كافة جوانبه مستندياً في ذلك إلى أسلوب كتابة المصحف بالرسم العثماني وإلى أسلوب كتابة الخط العربي الحديث والضرورات الطباعية الحديثة وإلى قواعد الكتابة المعروفة ومنها قواعد كتابة الهمزة وواو الجماعة والألف المقصورة والتاء المربوطة ، إضافة إلى الشواذ ، مع إعطاء فكرة عن المعالجات الآلية اللازمة لكي يدخل النص إلى قواعد البيانات . لذلك فإن قاعدة البيانات الأولى هي قاعدة بيانات الكتابة العربية.

بعد إعداد النص القرآني بشكل آيات متسلسلة يجب تحديد بدايات الجمل القرآنية ونهاياتها. قاعدة البيانات الثانية اللازمة هي قاعدة البيانات الصرفية بما فيها من فصل للواصق السابقة واللاحقة والأفعال المشتقة منها الكلمات وأوزانها الصرفية والجذور التي تعود لها الكلمات .

أما قاعدة البيانات الثالثة فهي قاعدة البيانات النحوية والتي تستكمل تقسيم المصحف إلى جمل مستقلة وفي إعراب الكلمات وعلامات الإعراب والتشكيل ومحل أشباه الجمل والجمل من الإعراب وإلى التذكير والتأنيث والإفراد والتنثنية والجمع وغير ذلك .

قاعدة البيانات الرابعة هي قاعدة بيانات المعاني والدلالة بدءاً من الجذور ثم الأفعال مع تجميع للأفعال ذات المعنى المتقارب أو المتضاد والمعاني المختلفة للكلمات مع تحديد كل استعمال ، وفيما إن كان الفعل لازماً أو متعدياً لمفعول واحد أو لأكثر من مفعول أو متعدياً بحروف الجر ومعاني كل فعل عند تعديه بحرف الجر ومعاني العبارات المكونة من أكثر من كلمة .

قاعدة البيانات الخامسة هي قاعدة بيانات الصوت والنطق . ونظراً لأن العناية بالقرآن الكريم في ترتيله وتجويده كانت عظيمة لا توجد في لغة أخرى غير اللغة العربية ، لذا فإن معالجة قواعد الوقف والتجويد في القرآن الكريم بشكل صوتي يمكن أن يكون نموذجاً للدراسات اللغوية الصوتية للغة العربية . وتجدر الإشارة إلى توفر ترتيل وتجويد للقرآن بأصوات مقرئين من شتى أصقاع العالم العربي والإسلامي يمكن أن يغني قاعدة البيانات هذه ، بالإضافة إلى تضمينها مختلف القراءات القرآنية .

يمكن تكوين قاعدة بيانات سادسة للخط العربي بمختلف خطوطه على مر التاريخ بما فيه النماذج الفريدة الرائعة فنياً للخطاطين الذين أولوا القرآن الكريم عناية فائقة .

يعطي البحث أمثلة عملية لبعض من المعلومات التي أدخلت إلى بعض قواعد البيانات الموصوفة أعلاه ، ويخلص إلى الإشارة إلى الجوانب والطرائق التي يمكن الاستفادة منها في مختلف الجوانب اللغوية في إعداد المعجم الآلي للغة العربية والإفادة من هذا المعجم في التعريب والترجمة وفي حوسبة النحو والصرف ونطق اللغة العربية .

١- مقدمة

تحتاج المعالجة الآلية للغة العربية معجمًا موسعًا يحوي كل ألفاظ اللغة العربية الفصحى المستعملة اليوم وجزءًا كبيرًا من اللغة العربية التي كانت مستعملة سواء قبل الإسلام أو في عصور ازدهاره. وهذا المعجم يجب أن يحوي كل المعلومات عن الألفاظ التي تحتاجها عمليات المعالجة من دلالات مختلفة وصرف وما يجوز أن تستعمل معه الكلمة وما لا يجوز والكلمات المقاربة والمضادة وكل ما يوصل إلى دقة دلالة الكلمة من وسائل.

يحتاج عمل هذا المعجم إلى أبحاث ودراسات موسعة بغية الوصول إلى تصميم مناسب له وبشكل يخدم كل ما تحتاجه المعالجة الآلية للغة من معلومات. وحيث أن الجهد اللازم لوضع المعجم هو جهد هائل يحتاج إلى مؤسسات متعاونة فيما بينها فإن إثراء المناقشة فيما يتعلق بتصميم مكونات هذا المعجم تعتبر خطوة لا بد منها.

إن دراسة مفردات القرآن الكريم هي خير بداية لمثل هذا الجهد ، وسنناقش في هذا البحث قواعد بيانات القرآن الكريم وكيف يمكن أن تساعد في مناقشات تصميم المعجم الآلي الموسع للغة العربية. إن أهمية المعجم بالنسبة للمعالجة الآلية للغة بالغة. فالمعجم يجب أن يحوي مكونات اللغة المفردة والمركبة وتعريفاتها وما تحيل إليه من المفاهيم والقيم التي تحكم وعي المجموعة اللغوية . لذلك كان الإشتغال بمعالجة المعجم من المداخل الأولية لضروب عدة من المعالجة الحاسوبية للغة (١)

٢- بين النص القرآني والكتابة العربية اليوم

يختلف التعامل مع النصوص باللغة العربية والتعامل مع نص القرآن في عدد من الأمور يمكن إيجازها على النحو الآتي :

أ- لا يورد النص القرآني غالبًا إلا مشكولاً . والتشكيل يسهل التعامل مع النص من نواحي متعددة حيث يذهب الالتباس والخطأ في المعالجات الآلية الصرفية والنحوية والدلالية . هذا مع العلم أن الكتابة العربية كانت تكتب قبل الإسلام وفي صدر الإسلام غير منقوطة وغير مشكولة ولم يضاف التنقيط إلا في وقت متأخر عن ذلك.

ب- يختلف النص القرآني عن النصوص العربية في حسن الالتزام بالرسم العثماني فجمهور الفقهاء يحبذون (بل بعضهم يعتبر ذلك ملزمًا) كتابة المصحف بالرسم العثماني وقد أجاز بعض العلماء كتابة المصحف برسم آخر غير الرسم العثماني إن كان ذلك يحقق دقة في قراءة ولفظ القرآن في عصر من العصور أفضل مما يحقق الرسم العثماني بسبب إعتياد الناس على رسم خاص لبعض الحروف.

ج- لغرض دقة قراءة القرآن الكريم ولفظه بشكل صحيح ، تحتوي المصاحف على علامات للوقف والتجويد وهي خاصة بالقرآن الكريم ولا تستعمل في أية كتابة عربية أخرى .

د- يحتوي النص القرآني على فواصل بين آية وأخرى . أي أن بداية ونهاية الآيات محددة (رغم وجود خلاف بين بعض المتخصصين برسم المصاحف حول مواقع بعض هذه الفواصل) . إن هذه الفواصل لا تقع في نهاية جمل مفيدة كاملة دائمًا بل تقع أحيانًا في وسط جملة أي أن الآية قد تكون جزءًا من جملة أو قد تكون جملة كاملة أو قد تحوي جملة وبعض جملة أو أنها قد تحتوي على عدد من الجمل .

هـ- إن الخطأ في التعامل مع النصوص غير القرآنية قد لا ينطوي عليه اعتراض يذكر أما الخطأ في التعامل مع القرآن الكريم فيلحق اعتراضًا دينيًا واضحًا ، لذلك فإن مقدار الحرص على دقة التعامل مع النص القرآني يجب أن يكون فائقًا .

٣- إعداد النص القرآني

قد يكون من المقبول في التعامل مع النصوص فيما عدا القرآن الكريم وجود نسبة بسيطة من الخطأ الإملائي أو الخطأ في التشكيل (إن كان النص مشكولاً) وقد تبلغ النسبة المقبولة من الدقة ٩٩,٩% ولكن ذلك غير مقبول على الإطلاق في نص القرآن الكريم . فهذه النسبة من الخطأ تعني وجود خطأ مقداره واحدًا بالألف . وحيث أن النص القرآني المشكول يقرب حجمه من ٧٠٠ ألف

بايت فوجود خطأ مقداره واحد بالألف يعني ٧٠٠ خطأ في القرآن كله وهذا ما لا يقبله أحد من المسلمين .

لذلك كانت المهمة الأولى هي التأكد من النص القرآني بنسبة تقرب من ١٠٠% تمامًا وهذه المهمة ليست باليسيرة فقد استدعت المهمة قراءة النص المدخل للحاسب عشرات المرات . وليس هذا فحسب ، بل دقق في مراحل تالية بعد إجراء معالجات متعددة على النص القرآني لتدارك ما تسرب من أخطاء قليلة .

لذا كان من الضروري اعتماد بعض الأسس المتعلقة بالتعامل مع النص القرآني . أول هذه الأسس هو اعتماد الرسم العثماني للمصحف واعتماد قراءة حفص عن عاصم (كونها الأكثر شيوعًا في العالم الإسلامي اليوم) واختيار مصحف المدينة النبوية كأساس لهذا الإدخال . واستنادًا لذلك اختيرت فواصل الآيات وعلامات الوقف والتجويد المستعملة في هذا المصحف ابتداءً .

كان من الضروري في البداية الإجابة على سؤال اختلف فيه الفقهاء الأقدمون كثيرًا وهو هل البسمة آية من الفاتحة؟ أم هي آية من كل سورة عدا سورة التوبة؟ أم هي ليست آية من أية سورة ، بل هي بعض آية من سورة النمل؟ . ويمكن الرجوع إلى تفاصيل هذا الخلاف في كتب التفسير والذي فصله كثير من المفسرين في بداية تفسير سورة الفاتحة (أنظر تفسير ابن كثير (٢) وتفسير القرطبي (٢٣) على سبيل المثال) .

والذي اعتمده في قواعد البيانات هذه هو اعتبار البسمة أول آية من القرآن الكريم (أي أول آية من سورة الفاتحة) دون بقية السور ، وقد وجدنا أن إثبات البسمة لمرة واحدة في أول سورة الفاتحة هو أفضل من تكرارها في كل سورة (عدا سورة التوبة) حيث أن ذلك سيطيح من تكرار لا داعي له كما أن عدم إثباتها في الفاتحة يعني أن لا تذكر البسمة إلا مرة واحدة كجزء من آية في سورة النمل .

٤- المعالجة الأولية للنص القرآني

مما تقدم كان لا بد من اعتماد النص القرآني آية آية . إن كل آية تعتبر قيدًا مستقلًا (record) أو سطرًا كاملًا . ولا يفصل بين كلمة من كلمات الآية سوى فراغ واحد . أما بين آية وأخرى فلا داعي لأية فاصلة لأن الانتقال من سطر لآخر إشارة إلى انتهاء الآية والبدء بآية جديدة وقد أدى ذلك إلى انتفاء الحاجة إلى ترقيم الآيات الواحدة تلو الأخرى أو وضع الترقيم بين آيتين . فالمعالجة الحاسوبية عند البرمجة يمكن أن تقوم بالعد والترقيم دون الحاجة إلى إثبات ذلك في النص . وعلى هذا فقد احتوى النص على (٦٢٣٦) سطر أو قيد .

أما الفواصل بين سورة وأخرى فكان لا بد من اختيار رمز يشير إلى ابتداء سورة جديدة . وقد اختيرت العلامة # كرمز لبداية سورة جديدة .

أما علامات الوقف والتجويد فقد اختير لكل علامة رمز من الرموز المستعملة في الحاسوب مثل \$ & % + - وغيرها . كما اختيرت علامة خاصة لمواقع سجود التلاوة . وقد لوحظ أن بعض علامات الوقف والتجويد تقع بين كلمتين من كلمات المصحف وبعضها يقع في آخر الكلمة بينما وقع البعض الآخر في وسط الكلمة . وهذا جعل الكلمات بعضها غريبًا في شكله نظرًا لأن الحروف تختلف في هياتها إن وقعت في وسط الكلمة أو أولها أو آخرها إذا ما أدخل وسط الكلمة رمز من هذه الرموز . وعلى أية حال سوف يتبين فيما بعد أن المعالجة الأخيرة كانت بحذف كل هذه الرموز عند استخلاص الكلمات كلمة كلمة وذلك لسببين : السبب الأول صعوبة استخدام هذه العلامات في برامج معالجة النصوص الشائعة حيث يستدعي ذلك استخدام بنط (ترميز للحروف) خاص بالمصحف لا يتوفر عند الغالبية العظمى من مستخدمي الحاسوب والسبب الثاني والأهم أن الكلمة القرآنية غالبًا تنطق بمعزل عن علامات الوقف والتجويد (إلا إذا كانت العلامة وسط الكلمة وهي نادرة) والتي غالبًا ما تكون بين الكلمات المتعاقبة أو نهاية الجمل أو الآيات . أما الأمر المهم الآخر في هذا المضمار فهو الألف الخنجرية التي توضع فوق بعض الحروف ولا تكتب ألقًا مثل "الرحمن" و "هذا" . وقد تم استخدام علامة التعجب ! لتتوب محل هذه الألف . وسيتبين فيما بعد

أن هذه الألف كذلك قد تم الاستعاضة عنها بالألف الممدودة تارةً في مثل كلمة "العالمين" وبالفتحة تارةً أخرى في مثل كلمة "الرحمن" و "هذا" وذلك تماشيًا مع الخط العربي الشائع اليوم بكتابة هذه الكلمات خلأً للقاعدة بدون ألف ممدودة . وكان هذا الحذف لضرورات الطباعة باستعمال معالج نصوص عادي . هذا مع الإشارة أن هذه المعالجة لم تتم في هذه المرحلة ولكن في مرحلة لاحقة كما سيتبين في ما بعد . كما استعملت علامة التعجب ؟ بدل همزة الوصل.

وعلى هذا فإن النص القرآني في هذه المرحلة كان نصًا استبعدت منه علامات الوقف والتجويد ومواقع سجود التلاوة ولكن أقيمت فيه علامات الألف الخنجرية . أي أن النص في هذه المرحلة هو نص بالرسم العثماني خال من العلامات (عدا الألف الخنجرية وهمزة الوصل) وليس نصًا مطابقًا للنطق ، حيث أن هناك عددًا من الكلمات القرآنية التي لا يمكن نطقها بشكل صحيح إلا بمساعدة العلامات المثبتة في المصاحف حاليًا والتي استثنيت في هذه المرحلة من المعالجة . وقد حذفت العلامات المشار إليها باستخدام أحد برامج التحرير (edit).

٥- قاعدة بيانات رسم الكلمات

كان لا بد من البدء بتكوين قاعدة بيانات لرسم المصحف وفق الرسم العثماني . فالكلمات الأساس في قاعدة البيانات هذه هي حقل رسم الكلمة بالرسم العثماني ، محذوفًا منه علامات الوقف والتجويد فقط . ويلاحظ أن هناك كلمات في الرسم العثماني مكتوبة في أماكن مختلفة في المصحف بأشكال مختلفة فقد تختلف فيما بينها في رسم التاء مرة كتاء مربوطة ومرة كتاء طويلة وكذلك في وجود ألف واو الجماعة مرة وعدم وجودها مرة أخرى وهكذا .

أما الملاحظة الثانية فهي عن رسم الهمزة . فرسم الهمزة في المصحف مختلف عن الرسم الحديث للهمزة في اللغة العربية المتداولة اليوم في كثير من المواضع .

لا شك إن معظم الكلمات القرآنية متطابقة في كتابتها بين القديم والحديث . ولكن هناك عدد لا يستهان به من الكلمات المختلفة في شكل كتابتها . ويمكن إجمال الاختلافات بما يأتي :

أ- استعمال الألف الخنجرية في كثير من الكلمات مثل العالمين (تكتب العلمين) .
ب- كتابة الهمزة بشكل مختلف عما هو متعارف عليه في الخط الحديث مثل الذكزين التي تكتب اليوم الذكزين باستعمال المدة فوق الألف (التي هي همزة بعدها فتحة طويلة (أي ألف) رغم أن المصاحف تستعمل المدة كعلامة من علامات التجويد).
ج - ألف واو الجماعة تكتب أحيانًا بالألف وأحيانًا بدون الألف . كما أن هناك أحيانًا واوًا أصلية في الكلمة مضافًا لها ألف مثل أدعوا .

د - كتابة التاء تارة كتاء مربوطة وتارة كتاء طويلة .

هـ - كتابة بعض الكلمات بشكل خاص مثل الربا تكتب (الربوا) .

و - دمج بعض الكلمات أحيانًا وعدم دمجها أحيانًا أخرى مثل أم من (تكتب أحيانًا أمَّن) وأن لا تكتب أحيانًا (ألا) وإن لم (ألم) .

ز - دمج بعض الكلمات بطريقة مغايرة لما متعارف عليه اليوم مثل " يا ابن أمي " (بينوم) و " فما للذين " (فمال الذين) .

ح - الحروف في فواتح بعض السور التي تسمى بالنورانية غير متعارف عليها في الخط الحديث . ولكنها ربما تشبه استعمال الرموز في المعادلات الرياضية أو في الأشكال الهندسية أو تبويب الفقرات هجائيًا .

وعلى هذا فإن أول خطوة في تكوين قاعدة بيانات رسم الكلمات هو إعادة كتابة الكلمات المختلف عليها برسم حديث فتكون قاعدة البيانات هذه من حقلين الأول هو الرسم العثماني الدقيق والثاني هو الرسم الحديث .

إحتوت قاعدة بيانات القرآن الكريم الرئيسة على ٨٣٨٢٩ قيدًا تحوي بينها ١١٤ علامة # لبداية السور و ٦٢٣٦ علامة @ لبداية الآيات حيث أن عدد كلمات المصحف بالرسم العثماني هو ٧٧٤٧٩ فيها ٤ كلمات هي كلمات البسملة في بدء سورة الفاتحة . وعند استخلاص الكلمات القرآنية

ووضعها في قاعدة بيانات ثانوية خاصة بالكلمات القرآنية فإن عددها سيكون ١٨٨٤١ كلمة. أما عند حذف تشكيل الآخر والشدة على أول حرف إن وجدت بسبب التنوين في الكلمة السابقة فعددها ينزل إلى ١٥٢٦٣.

وسّعت قاعدة البيانات الثانية بخطوات متعاقبة وذلك بحذف الشدة على الحرف الأول لبعض الكلمات نتيجة وقوعها بعد كلمة محتوية على تنوين مثل تشديد اللام في (هدىً لِلْمُتَّقِينَ) وتشديد الميم في (ريب مّمّا) وقد تأتي هذه الشدة حتى في بداية آية إن كان في نهاية الآية التي قبلها تنوين مثل قوله تعالى (عدواً مبيناً - رَبِّكُمْ أَعْلَم) وتتحصر هذه الحالات إن كانت بداية الكلمة لاماً أو ميماً أو راءاً أو نوناً أو ياء أو واواً وكانت الكلمة التي تسبقها منونة كما في الأمثلة السالفة . ولهذا حذفت هذه الشدة على مرحلتين المرحلة الأولى بواسطة برنامج يدقق إن كانت هناك شدة على الحرف الأول من كل كلمة تبتدى بهذه الحروف فيقوم بفصلها والمرحلة الثانية كانت بأن تدقق يدويًا ومن ثم تحذف . وهناك مواضع أخرى للشدة التي كان يجب أن تحذف من بداية الكلمة مثل " لهم مّا يشاؤون " نتيجة التقاء ميمين في كلمتين متعاقبتين . كما عولجت مشكلة تشكيل اللام ألف . وتجدر الإشارة هنا إلى أن معالجات النصوص الشائعة تكتب اللام إلف بشكل غريب إن كانت اللام مشكولة أو عليها شدة (تكتب "لًا" أو "لًا" وليس "لا") ولذلك استعويض عن تشكيل اللام بتأخير الفتحة إلى ما بعد الألف بحيث تصبح "لًا" . وهذا خطأ كان لا بدّ من القبول به في هذه المرحلة . وبالطبع فإن المشكلة تزيد سوءًا إذا كانت الألف بعد اللام هي همزة مثل "لأنتم" حيث كتبت "لأنتم" أو "إلى الله" حيث كتبت "إلى الله" وقد حذفت الحركة على اللام في مثل هذه الحالات . وقد يلتقي اللام ألف والشدة على اللام نتيجة تنوين الكلمة السابقة لها وبذلك تجتمع المسألتان أعلاه في موضع واحد .

هناك حقل ثالث مهم يجب إضافته لقاعدة البيانات هذه ، ألا وهو الرسم بالخط العادي بدون تشكيل . فالكتابة الشائعة اليوم غير مشكولة بصورة عامة ولا شك أن ذلك يحدث كثيرا من اللبس في كتابة العديد من الكلمات وهذا الحقل ضروري لكي تؤخذ إمكانية التعامل مع هذا اللبس في أي معالجة آلية للغة العربية ذات معنى .

تكوين قاعدة البيانات هذه قد مرّ بمرحلتين . مرحلة آلية وذلك باستحداث قاعدة بيانات للكلمات المتشابهة ثم يجري تعديلها إلى الرسم العادي يدويًا وفي مرحلة لاحقة يحذف منها التشكيل وتعاد هذه الحقول إلى قاعدة البيانات الرئيسية لتتكون قاعدة بيانات مكونة من نص بالرسم العثماني وآخر بالرسم الحديث (العادي) وثالث بدون تشكيل .

إن كتابة الهمزة بالرسم الحديث تستند إلى قواعد مقننة من قبل مجامع اللغة العربية ويمكن فصلها في قاعدة بيانات خاصة بالهمزة والتعامل معها ببرنامج وفق هذه القواعد .

كما تجدر الإشارة إلى مسألة الشواذ في الكتابة . فكلمات مثل هذا والرحمن تكتب بالرسم العادي بلا ألف كما أن واو الجماعة لا تلفظ كما أن مسألة دمج بعض الكلمات في الرسم العثماني لها ما يقابلها من كلمات منفصلة في الرسم العادي . وبذلك فإن عدد كلمات المصحف بالرسم العثماني ستختلف عن عدد كلمات المصحف بالرسم العادي أي أن تقابل هذه الكلمات بعضها تجاه البعض يحتاج إلى معالجات خاصة . أنظر شكل (١) كما يبين الشكل (٢) جدولاً بالجذور الأكثر ترددًا في القرآن الكريم والتي يزيد عدد مرات تكرارها عن ١٢٠ مرة. لاحظ أن بعض الكلمات (خاصة الحروف) قد أدرجت مع الجذور أيضاً. ومن هذا الجدول يمكن أن تستنبط الأفعال الأكثر ترددًا في اللغة العربية والتي يمكن البدء بها عند تعليم العربية لغير العرب. يبين الشكل (٣) قاعدة بيانات الكلمات القرآنية .

٦- قاعدة بيانات الجمل القرآنية

كما سبق ذكره فإن الجمل القرآنية متداخلة مع فواصل الآيات. أي أن الجمل القرآنية يمكن أن تكون بعض آية أو آية كاملة أو جزءًا من آية مع جزء من آية أخرى أو آية كاملة مع أجزاء من آية أخرى أو عددًا من الآيات.

لذلك كان من الضروري وضع إشارات ضمن النص تشير إلى نهاية جملة وبدء جملة جديدة . وذلك للقرآن كله. وتجدر الإشارة إلى أن علامات الوقف والتجويد قد يشير بعضها إلى مثل هذه المواضع حيث أن إشارة (الوقف اللازم أو الوقف أولى) تشير إلى بعض من هذه المواضع. خذ مثلاً سورة الفاتحة ، فالجمل فيها :

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ @
الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ @ الرَّحْمَنِ الرَّحِيمِ @ مَالِكِ يَوْمِ الدِّينِ @
إِيَّاكَ نَعْبُدُ

وَإِيَّاكَ نَسْتَعِينُ @
اهدنَا الصِّرَاطَ الْمُسْتَقِيمَ @ صِرَاطَ الَّذِينَ أَنْعَمْتَ عَلَيْهِمْ غَيْرِ الْمَغْضُوبِ عَلَيْهِمْ وَلَا الضَّالِّينَ @
وهذه التجزئة هي إستقرائية لا غير. وقد تختلف مواقع تجزئة النص إلى جمل بحسب فهم من يقوم بها من المفسرين. فمثلا فاتحة سورة البقرة يمكن أن تجزأ كما يأتي:

@الم
ذَلِكَ الْكِتَابُ لَا رَيْبَ
فِيهِ هُدًى لِّلْمُتَّقِينَ @
ويمكن أن تجزأ كالآتي:

@الم
ذَلِكَ الْكِتَابُ لَا رَيْبَ فِيهِ
هُدًى لِّلْمُتَّقِينَ @

وهناك بعض المواضع التي تشير إليها علامات خاصة من علامات الوقف والتجويد التي تبين إمكان الوقف على أي من الكلمتين (فيه أو هدى) بشكل متبادل (أي أحدهما فقط) . ولكل من الصيغتين إعراب مختلف لكل جملة. وهكذا تتكون قاعدة بيانات للجمل القرآنية يكون أساساً لقاعدة بيانات النحو.

٧- قاعدة بيانات الصرف

هناك من الخطوات التي ينبغي إجراؤها قبل البدء بتكوين قاعدة بيانات الصرف لنص القرآن الكريم . أول هذه الخطوات هو تجزئة الكلمات القرآنية إلى لواصقها الأولى ولواصقها النهائية وأدوات التشكيل ويمكن كحد أعلى اعتبار القطع ٤ لواصق أولية كحد أعلى و ٢ من اللواصق الأخرية وجذع الكلمة وتشكيل للجذع وتشكيل آخر .

وبذلك فإن الكلمة يحجز لها ٩ حقول . خذ مثلاً كلمة " أفبالباطل " تحتوي ٤ لواصق أولية هي " أ - ف - ب - أل " وجذعها هو باطل يعقبها كسرة في الأخير . أما كلمة " فسيكفيكم " فتحتوي على لواصق أولية عددها ٢ هما " ف - س " والجذع يكفي واللواصق الأخرية هي " ك - و " هم " والتشكيل على الجذع هو السكون (غير الظاهر على الياء) وتشكيل آخر اللواصق هو السكون أيضا (الذي يظهر أو يختفي حسب الكلمة التالية للميم)

وتجدر الإشارة إلى أن ألف لام التعريف عند حذفها يكون الحرف الأول من الكلمة مشدداً إن كان من الحروف الشمسية لذلك يجب حذف هذه الشدة من جذع الكلمة . أما إن كان من الحروف القمرية فإن اللام (من ألف لام) تحمل سكوناً يجب حذفه .

أما جذع الكلمة فيمكن أن تكون مشتقة من فعل ثلاثي على وزن ما . ويمكن أن يتم ذلك بواسطة برنامج خاص يقوم باستنباط الوزن الذي اشتقت منه هذه الكلمة . وهذا الاشتقاق يمكن أن يكون بسيطاً إذا كانت أحرف الجذر ليس بينها حرف علة . أما عندما يكون أحدها (أو أكثر من واحد) حرف علة فإن قواعد تحويل الواو إلى ياء أو ألف أو بالعكس تجعل القواعد أكثر تعقيداً كما هو معروف .

أما إذا لم يكن الجذع مشتقاً فيعني ذلك عدم وجود جذر للكلمة . وعلى هذا فإن عدد حقول قاعدة بيانات الصرف تصبح ١٢ حقلاً هي

الكلمة ، ٤ حقول للواصق الأولية ، جذع الكلمة ، تشكيلها ، لواصق أخرية ، تشكيل الآخر ، جذر الجذع ، وزن الجذع .
يبين الشكل (٥) نموذجًا لقاعدة البيانات الصرف كما يعطي الشكل (٦) قاعدة بيانات اللواصق وهي قد قلصت فيها اللواصق الأولية إلى اثنتين فقط (دمجت اللواصق التي تزيد عن اثنتين إلى اثنتين وهي حالات محدودة) . هذا وقد استعين في تحديد جذور الكلمات القرآنية بالمعجم المفهرس لألفاظ القرآن الكريم (٤) ولسان العرب (٥).

٨- قاعدة بيانات النحو

يمكن تكوين قاعدة بيانات النحو هذه مشتقة من قاعدة بيانات الصرف وتختلف قاعدة البيانات هذه عن قاعدتي البيانات السابقتين (الكتابة والصرف) بأن عدد كلماتها أكثر من القاعدتين السابقتين حيث أن كل جزء من أجزاء اللواصق (الأولى والأخيرة) هي كلمة مستقلة ، إضافة إلى جذع الكلمة . كما يلحق بكل كلمة من هذه الكلمات الجديدة تشكيلها . كما أنه من الضروري إضافة حقل خاص يحدد الكلمة القرآنية التي يعود لها جزء الكلمة هذه وموقعها التسلسلي ضمن الكلمة . فكلما فسيفيكيفهم التي رقمها في قاعدة بيانات كتابة المصحف هو (٢٥٩٧)

اللواصق	الحركة	رقم الكلمة	تسلسل جزء الكلمة
ف	فتحة	٢٥٩٧	١
س	فتحة	٢٥٩٧	٢
يكفي	سكون	٢٥٩٧	٣
ك	فتحة	٢٥٩٧	٤
هم	سكون أو ضم عند التقاء الساكنين	٢٥٩٧	٥

وهنا بعد هذا يضاف إعراب الكلمات واحدة واحدة مع تبيان إن كانت الكلمة مبنية أو معربة وعلامة الإعراب أو أية معلومات تفصيلية عن الإعراب.
وتجدر الإشارة إلى أن بعض أجزاء الكلمة قد لا يكون له محل من الإعراب مثل ألف التعريف فيشار إليه بذلك .

يبقى هنا كيف يمكن التعبير عن إعراب أشباه الجمل أو الجمل ، وقاعدة البيانات المذكورة غير مناسبة لذلك فيصبح من الضروري تكوين قاعدة بيانات فرعية تحدد رقم الكلمة وتسلسل جزء الكلمة الذي تبدأ به شبه الجملة ومداهما ثم موقعها من الإعراب وعلامة الإعراب مع نص شبه الجملة أو الجملة .

هناك أمر آخر على جانب من الأهمية في موضوع النحو وهو عودة الضمير . فالضمائر في اللغة العربية تعود إلى كلمة سابقة أو شبه جملة . وقد تكون الضمائر ظاهرة أو مستترة . لذلك من الضروري اعتبار الضمير المستتر كلمة فارغة بين أجزاء الكلمة ضمن تسلسلها مع إضافة حقل خاص لتبيان ما يعود له الضمير .

أن تكوين قاعدة البيانات هذه ليس أمرًا واحدًا ثابتًا كنص القرآن الكريم فهناك خلافات نحوية كثيرة بين المتخصصين في النحو في إعراب الكثير من الكلمات القرآنية واستنادًا إلى تلك الخلافات يختلف المعنى والتفسير أيضا ولغرض استيعاب مثل هذا الخلاف بالرأي يمكن أن توضع قاعدة بيانات النحو بشكل يقبل تعدد الآراء بحيث تصبح كقاعدة بيانات ثانوية تحتوي ما يأتي :

رقم الكلمة القرآنية

رقم جزء الكلمة

حركة جزء الكلمة

نوع جزء الكلمة مبنية أم معربة وهل هي اسم أو فعل أو حرف.

موقعه من الإعراب

رمز الرأي الإعرابي

عودة الضمير

ملاحظات إن وجدت

وبذلك يمكن أن يكون للكلمة الواحدة عدد من أشكال الإعراب تجمع في قاعدة بيانات واحدة ويمكن الرجوع إلى أي منها عند الحاجة .

كما أن الاصطلاحات التي يتم اعتمادها يمكن أن تكون قواعد بيانات فرعية مثل قاعدة بيانات الكلمات المبنية وقاعدة بيانات علامات الإعراب وقاعدة بيانات المواقع من الإعراب . هذا وإن هناك حقلين آخرين تجدر إضافتهما إلى قاعدة البيانات الرئيسية أيضا وهما حقل الجنس (التذكير والتأنيث) إن وجد ، حقل الإفراد والتثنية أو الجمع إن وجد وبالنسبة للجمع نوع الجمع إن كان جمع مذكر سالم أو جمع مؤنث سالم أو جمع تكسير . أنظر الشكل (٥)

٩- قواعد بيانات المعاني والدلالة

قاعدة البيانات هذه ما هي إلا أفكار عامة تحتاج إلى مزيد من العناية والتوسيع فموضوع تقنين المعاني والدلالة على جانب كبير من الصعوبة بالنسبة للغات الحديثة فكيف باللغة العربية الواسعة المعاني والدلالة وكيف إن كان الموضوع يتعلق بصفوة البيان العربي ألا وهو القرآن الكريم . إلا أن قاعدة البيانات هذه تكاد تكون صفوة قواعد البيانات التي يرجى إكمالها لتخدم تفسير القرآن ومعاني مفرداته .

أول قاعدة بيانات للمعاني هي قاعدة بيانات الجذور .

فالجذور من ناحية المعنى مترابطة مع بعضها فمثلا جذور مثل : جيء ، أتى ، قدم ، حضر ، وصل ، ولي ، دبر ، وكذلك : دخل ، ولج ، قبل ورد ، وصل وهذه المجموعات مترابطة مع بعضها حيث أنها تتعلق مع غيرها من المجموعات بحركة الإنسان ذهابًا أو إيابًا أو خروجًا أو دخولًا أو وصفا لكيفية هذه الحركة من سرعة أو انحدار أو ارتفاع أو تغير أو خوف أو إجبار أو توقيت أو حلول أو تكرار .

وعلى هذا فالأفعال تتجمع في مجموعات يمكن وصفها بصفة جامعة عامة ثم تنقسم إلى مجموعات أدنى من ذلك تتصف بجامع أخص من الأول وهكذا .

لكن المجموعات النهائية لا تتفق في المعنى فهي مؤتلفة أحيانا ومختلفة أحيانا أخرى . وهناك الكثير من المؤلفات في التراث العربي التي تذكر مثل هذه الألفاظ (٦) .

كما هناك قواميس حديثة التي تعطي مثل هذه المجموعات وما يقابلها من مفردات أجنبية (٧).

إن الجذور وحدها تجمع في غالبيتها أصولاً لمعاني عديدة . ولذلك فإن الاستناد إليها يعطي وجهة عامة غير دقيقة للدلالة . لذلك يجب الانتقال من مستوى المجموعات العامة للجذور إلى المجموعات المتخصصة إلى مجموعات الجذور المتقاربة إلى الأوزان المختلفة التي تدخل على الجذر نفسه . فلجذر ذهب صيغ أفعال مثل ذهب ، أذهب ، ذهب وللجذر قدم : صيغ مثل : أقدم (أفعل) وقدم (فَعَل) واقتدم (افتعل) تقادم (تفاعل) تقدم (تفعل) واستقدم (استفعل) (٨) .

ويلاحظ أن بعضًا من هذه الصيغ ذات علامة بالسير . أما بعضها مثل تقادم فليس لها علامة بذلك بل تتعلق بمعنى الزمن ، كما أن قدم قد لا تشير إلى حركة الإنسان وحده بل إلى حركة الأشياء أيضا . وهكذا فإن سلسلة الأفعال المتقاربة (المزيدة) بالمعنى يجب أن ينظر إليها وليس إلى الجذور المتقاربة فقط .

هناك حاجة أيضا إلى الإشارة إلى الفروق بين الأوزان المختلفة . فمثلاً قدم واستقدم تشيران إلى فعل القدوم أو التقديم على الغير وللغة العربية قواعد راسخة في مثل هذه المعاني تحتاج إلى إدخال إلى قواعد البيانات بطرق مناسبة ذات ترميز دقيق .

الأمر الآخر الذي ينبغي أن تشملته قواعد البيانات هو تعدد المعاني للفظ الواحد .

الأفعال باللغة العربية من ناحية المعاني يمكن تقسيمها إلى :

أفعال لازمة : اركعوا واسجدوا

أفعال متعدية لمفعول واحد : ورد ماء مدين

أفعال متعدية بحرف الجر : واسجدوا لله
أفعال متعدية لمفعول واحد وبحرف الجر : أوى إليه أخاه
أفعال متعدية لمفعولين : فسيفيكمهم الله
أفعال متعدية لمفعولين وبحرف الجر : وجد عليه أمة من الناس يسقون
وقد يكون الفعل تارة من أحد الأفعال المتعدية بحروف الجر بحيث يعطيها معاني مختلفة عند اختلاف حرف الجر ، فالفعل ذهب مثلا ورد مع إلى مثل " اذهب إلى فرعون " . وقد تأتي الجمل أو أشباه الجمل في موقع المفعول به في أفعال أخرى.
ومع الباء : " تذهبوا به " ومع عن : " يذهب عنكم الرجس " .
وقد يأتي لازماً بلا حرف جر مثل " اذهب أنت وربك " والفعل آتى أكثر من ذلك تشعباً .
مثل " يأت آباءهم " ، " فأتوا بآياتنا " و " آتاهم الله من فضله " ، " وأتوا البيوت من أبوابها " .
وأوتيت من كل شيء تعدي الفعل لمفعول معين يشير إلى معنى مثل " يبغونها عوجاً " . كذلك فإن تضاد المعاني بين أفعال (أو كلمات) وهذا التضاد قد يكون من جهة معينة . فقد يكون التضاد يتعلق بالاتجاه مثل ذهب وأتى أو قام وقعد أو مثل فوق وتحت . وقد يكون التضاد من جهة الزمن مثل الماضي والمستقبل وقد يكون من جهة التبادل بين شيئين مثل أخذ وأعطى وقد يكون من جهة صفة معينة مثل أمات وأحى وهكذا وهذا التضاد يعطي المعنى الدقيق أحياناً إذا عرف معنى الضد . وفي القرآن كما في اللغة العربية أصلاً عبارات كاملة تشير إلى معنى محدد كان يمكن أن يشار له بفعل واحد (لكن ربما تحمل معنى إضافياً)

قدرناها من الغابرين = أهلناها

خذوا ما آتيناكم بقوة = اتبعوا ما آتيناكم بدقة

واضرب لهم مثلاً = اقصص عليهم مثلاً

مما تقدم يتبين أن قواعد البيانات التي تخدم المعنى والدلالة واسعة جداً ومعقدة ، نظراً لصعوبة الموضوع أصلاً ولكنها ضرورية جداً للتقدم في مجالات معالجة بيانات اللغة العربية آلياً والترجمة والدلالة . ويمكن أن تذكر هنا بعض قواعد البيانات التي تخدم المعاني والدلالة في القرآن الكريم .

- أ - قاعدة بيانات مجموعات الجذور العامة .
- ب - قاعدة بيانات مجموعات الجذور المتخصصة .
- ج - قاعدة بيانات مجموعات الجذور المتسلسلة المعنى .
- د - قاعدة بيانات الأفعال المزيدة .
- هـ - قاعدة بيانات الأفعال من ناحية اللزوم والتعدي .
- و - قاعدة بيانات تعلق حرف الجر بالأفعال .
- ز - قاعدة بيانات الكلمات المتضادة .
- ح - قاعدة بيانات العبارات التي تشير إلى معنى محدد .
- ط - قاعدة بيانات المعاني المعقدة .

١٠ - قاعدة بيانات اللفظ

كانت العناية بقراءة وترتيل وتجويد القرآن فائقة على مر العصور . وقد عني المسلمون بقواعد التجويد ووضع الرموز المساعدة للوقف واللفظ الصحيح أثناء القراءة . وتحتوي المصاحف الشائعة اليوم على الكثير من هذه الرموز وفق القراءة القرآنية الشائعة (هناك ثلاث قراءات تشيع المصاحف المكتوبة وبقها وهي برواية حفص وورش وقالون) .
ويمكن لقاعدة بيانات تجويد القرآن الكريم أن تحتوي على ما يأتي :

١ - قاعدة بيانات القراءات القرآنية (٨)

فالمعروف أنه على الرغم من التزام كل القراء بالاعتماد على الرسم العثماني إلا أن هناك اختلافات طفيفة بينها ويمكن جمع هذه الاختلافات في قاعدة بيانات تبين ذلك .
يضاف إلى ذلك الرموز المختلفة التي تشير إلى طريقة نطق بعض الحروف كالهزمة والتنوين .

ولا يكفي أن يوضح نطق كل كلمة منفردة لوحدها بل يجب تبيان نطق إرتباط الكلمات المتتالية مع بعضها حيثما كان ذلك ضروريًا للإيضاح.

٢- قاعدة بيانات الألفاظ الشاذة وهي كلمات خاصة تلفظ بصيغة معينة في بعض القراءات دون أخرى .

٣- قاعدة بيانات الأصوات .

فالمعروف إن وحدة الكلام المنطوقة هي الأصوات phonemes وليس الحروف وتحتوي اللغة العربية على أصوات لا توجد في غيرها كما أن أصوات التجويد وخاصة المدود ولفظ النون الساكنة والميم الساكنة والإدغام والإمالة والتخفيف تنطق بشكل معين . وجمع رموز هذه البيانات يجعلها مفيدة للاستعمال في دراسات وأبحاث عن اللغة العربية .

إن نطق بعض الحروف أو الأصوات يخضع في القراءات القرآنية لكثير من القواعد غير الظاهرة في المصاحف. فهناك أصوات مختلفة لبعض الحروف مثل الراء المخففة والمفخمة والنون والنون مع الغنة المخففة والنون مع الغنة المدغمة والواو والواو الإختكاكية والواو الصائتة والياء والياء الإختكاكية والياء الصائتة وصوت الإمالة وصوت التفخيم واللام المرققة واللام المفخمة. هذا بالإضافة إلى الحركات الثلاثة الفتحة والضمة والكسرة.(١٠)

٤- قواعد البيانات الصوتية .

لقد أصبح بالإمكان تجزئة الصوت بالحاسوب وتقطيعه وإعادة تركيبه والتعرف عليه. فإذا ما خزن صوت نموذجي لقارئ ما وكان هذا النموذج وافيًا وقطع إلى أجزاء مناسبة فإن بالإمكان إعادة تركيب هذه الأجزاء لقراءة نص آخر مغاير للنص الأصلي لكنه يحتوي على تلك القطع والأجزاء وبصوت القارئ نفسه .

إن قواعد بيانات اللفظ لا تكون مكونًا مباشرًا من مكونات المعجم الموسع لكنها تساعد في خدمة هذا المعجم.

١١- ماذا يمكن أن تقدم قواعد بيانات القرآن الكريم للغة العربية

إن صغر حجم بيانات القرآن الكريم بالمقارنة مع اللغة العربية جعل بالإمكان إكمال بعض قواعد البيانات بجهد معقول . أما البيانات الكاملة للغة العربية فتحتاج إلى جهود كبيرة . لكن الأمثلة التي عرضت في هذا البحث تعطي إشارة واضحة إلى سهولة الكثير من هذه الجهود وإمكان برمجتها . ويمكن تلخيص ما يمكن أن تقدم بيانات القرآن الكريم للغة العربية على النحو الآتي :

١- تعطي قواعد بيانات القرآن الكريم أمثلة واضحة للكثير من المعالجات التي يمكن أن تتم على اللغة العربية . وما تحتاجه اللغة العربية سوى توسيع بعض هذه القواعد لكي تشمل كل الكلمات والمعاني المستعملة في اللغة .

٢- إن البرامج التي كتبت وفق قواعد معينة لمعالجة كلمات القرآن الكريم يمكن استخدامها نفسها على اللغة العربية بشكل عام .

٣- إن نصوص القرآن الكريم التي تمت معالجتها كانت دقيقة لكنها تحتوي على الكثير من أشكال الكتابة غير القياسية . أما اللغة العربية الحديثة فهي تستعمل أدوات قياسية أكثر، لكن احتمال الخطأ في الكتابة يمكن أن يستفيد من أشكال الكتابة غير القياسية في المصاحف . وعندها يمكن تصحيح أخطاء الكتابة مثل الهمزة أو التنوين وغيرها .

٤- إن نصوص القرآن الكريم التي تمت معالجتها كانت مشكولة واشتق منها نص غير مشكول . وحيث أن الكتابة العادية في غالبيتها غير مشكولة لذلك فإن بالإمكان الاستفادة من قواعد بيانات القرآن الكريم لإعادة بعض التشكيل بصورة آلية إلى الكتابة العربية .

٥- إن الكثير من الأبحاث على قواعد بيانات القرآن الكريم لم تؤت ثمارها بعد لحاجتها إلى الكثير من المعلومات اللغوية . لكن تصميم هذه القواعد يعطي فوائد للغة وفي الوقت نفسه يعود بالفائدة على أبحاث القرآن الكريم وخاصة في علم التفسير .

٦- إن دراسات الصوت التي يمكن أن تتم على القرآن الكريم يمكن أن تؤتي ثمارها في الترجمة الآلية وآلات الإملاء الآلي وتركيب الأصوات .

١٢- المعجم الآلي الموسع للغة العربية

المعجم اللغوية الموسعة المتوفرة الآن عبر الشبكة العالمية (الأنترنت) للغات العالمية كالإنكليزية والفرنسية واليابانية تركز على جانبين : المعجم الكتابية والفهارس اللفظية. المعجم الكتابية تأخذ كل كلمة من كلمات اللغة وتعطي تركيبها الصرفي المحتمل في اللغة وإمكانية ورودها نحويًا والمعاني المختلفة لها مع الحروف أو التراكيب التي يمكن أن تدخل معها لتكوين معان جديدة . وبعد كل ذلك تعطي أمثلة على ذلك. أما في اللغة العربية فإن المزيد فوق ذلك يجب أن يؤخذ بعين الإعتبار. فالمعجم يجب أن يحوي جذور الكلمات ووزنها الصرفي كما يجب أن يحوي إمكانيات دخول اللواحق المختلفة عليها.

إن الموقع الإعرابي في اللغة الإنكليزية مثلاً هو غير الموقع الإعرابي في اللغة العربية. فكون الكلمة إسمًا أو صفة متعلق بالمعنى أن تكون الكلمة تسمية لشيء ما أو أنها تصف شيئاً ما . أما في اللغة العربية فإن موقع الكلمة كفاعل أو صفة لا يتعلق بالمعنى بالموقع الإعرابي في جملة معينة. لذلك فإن وسيلة جديدة يجب أن تبتكر لتصنيف الكلمات العربية من ناحية المعنى غير التصنيف النحوي.

أما التصنيف النحوي فهو ضروري أيضاً لكثير من الكلمات مثل الضمائر النحوي معظمها تحوي الموقع الإعرابي فأنا وأنت للموقع الإعرابي المرفوع و ياء المتكلم للنصب والجر وهكذا. ومن الضروري أن يحوي المعجم الصيغ والإرتباطات الممكنة بشكل متكامل. فالفعل أكل مثلاً لا يأخذ مفعولاً معنوياً فلا يقال أكل فلان حريته أو قوله . أي أن كل فعل يجب أن تذكر له عند ذكر دلالته ما يجوز أن يرتبط به من أنواع الفاعل أو المفعول به أو الجار والمجرور وما يجوز. إن وصف اللغة العربية بشكل مقنن في قواعد بيانات موسعة هو الأسلوب الأمثل للتفاعل الآلي مع اللغة العربية ومكوناتها داخلياً ، وكذلك بينها وبين اللغات الأخرى خارجياً. وذلك ضروري جدًّا لوضع اللغة العربية في المكانة التي تستحقها اليوم بين اللغات الحديثة.

١٣- المصادر:

- ١- قاعدة البيانات المعجمية العربية – حصيلة وآفاق : دز عبد الفتاح ابراهيم و د. سالم غزالي – مجلة المنظمة العربية للتربية والثقافة والعلوم - تونس
- ٢- تفسير القرآن العظيم للحافظ ابن كثير
- ٣- الجامع لأحكام القرآن العظيم للقرطبي
- ٤- المعجم المفهرس لألفاظ القرآن الكريم لمحمد فؤاد عبد الباقي
- ٥- لسان العرب لابن منظور
- ٦- الألفاظ المختلفة في المعاني المؤتلفة – جمال الدين الجباني دار عمار ١٩٩١
- ٧- قاموس المفردات المتضادة – انكليزي – عربي د. كاظم عادل ناصر دار البشير ١٩٨٩
- ٨- إحصاء الأفعال العربية في المعجم الحاسوبي : مروان البواب وغيره، مكتبة لبنان ١٩٩٦
- ٩- معجم القراءات القرآنية – د. عبد العال سالم – جامعة الكويت
- ١٠- القراءات القرآنية بين الدرس الصوتي القديم والحديث للدكتورة مي فاضل الجبوري – دار الشؤون الثقافية العامة – بغداد ٢٠٠٠

شكل (١) نموذج للكلمات القرآنية مرتبة هجائياً

ءَأَشْكُرُ	ءَأَسْجُدُ	ءَأَلْهَيْتُنَا
ءَأَفْرَزُ	ءَأَسْلَمْتُ	ءَأَتَّخِذُ
ءَأَلِدُ	ءَأَسْفَقْتُ	ءَأَرْبَابُ

ءَأَنْزَلْنَاهُمْ

ءَأَنْتَ
ءَأَنْتُمْ

ءَأَمِنْتُمْ
ءَأَمِنْتُمْ

شكل (٢) الجذور والكلمات(*) التي يزيد ترددها عن ١٢٠ مرة في القرآن الكريم

رعي ٣٢٩	مثل ١٧٠	عزز ١٢١
علي ٣٣٢	قتل ١٧١	عمم ١٢٢
رحم ٣٤٠	تبع ١٧٥	يدي ١٢٣
ثم ٣٤٣	سبل ١٧٧	أهل ١٢٤
جعل ٣٤٧	صلح ١٨١	بشر ١٢٤
عمل ٣٦١	خرج ١٨٣	رزق ١٢٤
عذب ٣٧٤	سمع ١٨٦	خوف ١٢٥
كل ٣٧٧	حيي ١٩٠	دخل ١٢٧
أبيه ٣٨٣	ضلل ١٩٢	بل ١٢٨*
قد ٤٠٧	حسن ١٩٥	خلف ١٢٨
لأ ٤٢١	خير ١٩٧	عظم ١٢٩
هم ٤٤٥	نور ١٩٧	جمع ١٣٠
عرض ٤٦٢	جنن ٢٠٢	سأل ١٣٠
عن ٤٦٥	عند ٢٠٢*	طوع ١٣٠
يوم ٤٧٦	ملك ٢٠٧	لعل ١٣٠
ذلك ٤٨١	حكم ٢١١	نظر ١٣٠
هو ٤٨٢*	دعو ٢١٢	لكن ١٣١
رسل ٥١٤	أنت ٢١٨*	نذر ١٣١
ب ٥١٨*	أم ٢٢١*	دنو ١٣٤
شيء ٥٢٠	نوس ٢٣١	قدر ١٣٤
بين ٥٢٤*	غفر ٢٣٥	موسى ١٣٧*
كفر ٥٢٦	بعد ٢٣٦*	سلم ١٤١
أتي ٥٥٠	أمر ٢٤٩	حتى ١٤٣
لم ٥٧٥*	أخر ٢٥١	دون ١٤٥
قوم ٦٦١	أبي ٢٥٣	نعم ١٤٥*
أذ ٦٩٤*	وقى ٢٥٩	لقي ١٤٧
ألى ٧٤٣*	ولى ٢٦٠	بصر ١٤٩
له ٨٠٦*	خلق ٢٦٢	رود ١٤٩
من ٨٢٩*	أخذ ٢٧٤	وعد ١٥٢
علم ٨٥٥	عبد ٢٧٦	وحد ١٥٤
أمن ٨٨٠	لو ٢٧٦*	غير ١٥٥*
رب ٩٨٢	هذا ٢٧٦*	صدق ١٥٦
كون ١٣٥٩	جيء ٢٧٩	بعض ١٥٩
ألى ١٤٤٣*	كذب ٢٨٣	نصر ١٦٠
ألى ١٤٧١	أحق ٢٨٨	شهد ١٦١
في ١٧٠٢*	ذكر ٢٩٣	نبأ ١٦١
قول ١٧٢٣	نزل ٢٩٤	كبر ١٦٢
لا ١٧٣١*	أقبل ٢٩٥*	بنو ١٦٤
ما ٢١٨١*	نفس ٢٩٩	مع ١٦٥*
أله ٢٨٥٢	ظلم ٣١٦	موت ١٦٦
من ٣١٠٠*	سمو ٣١٨	سوء ١٦٨
ن ٤٠٣٨*	كتب ٣٢٠	كثر ١٦٨
	أعو ٣٢٦	شرك ١٦٩
	أهدي ٣٢٧	قلب ١٦٩

الشكل (٣) قاعدة بيانات رسم الكلمات القرآنية

k	n	oth	no_shada	no_beg	no_sh_end	word	rnum	root	verb	root2
17	#	#	#	#	#	#	0	#	#	#
27	@	@	@	@	@	@	0	@	@	@
٣٠	بِسْمِ	بِسْمِ	اسْمِ	بِسْمِ	بِسْمِ	بِسْمِ	38	سمى	اسْمِ	سمى
٤٠	اللّٰهُ	اللّٰهُ	اللّٰهُ	اللّٰهُ	اللّٰهُ	اللّٰهُ	٥٠	ءله	اللّٰهُ	اللّٰهُ
٥٠	الرَّحْمٰنِ	الرَّحْمٰنِ	رَحْمٰنِ	رَحْمٰنِ	الرَّحْمٰنِ	الرَّحْمٰنِ	٧٧	رحم	رَحْمٰنِ	رحم
٦٠	الرَّحِیْمِ	الرَّحِیْمِ	رَحِیْمِ	رَحِیْمِ	الرَّحِیْمِ	الرَّحِیْمِ	٧٧	رحم	رَحِیْمِ	رحم
٧٧	@	@	@	@	@	@	٠	@	@	@
٨٠	الرَّحْمٰدِ	الرَّحْمٰدِ	رَحْمٰدِ	رَحْمٰدِ	الرَّحْمٰدِ	الرَّحْمٰدِ	١٦	حمد	رَحْمٰدِ	حمد
٩٠	اللّٰهُ	اللّٰهُ	اللّٰهُ	اللّٰهُ	اللّٰهُ	اللّٰهُ	٥٠	ءله	اللّٰهُ	اللّٰهُ
١٠٠	رَبِّ	رَبِّ	رَبِّ	رَبِّ	رَبِّ	رَبِّ	١١١	ربب	رَبِّ	ربب
١١٠	الرَّعٰلَمِیْنَ	الرَّعٰلَمِیْنَ	عٰلَمِیْنَ	عٰلَمِیْنَ	الرَّعٰلَمِیْنَ	الرَّعٰلَمِیْنَ	١٦١	علم	عٰلَمِیْنَ	عالم
١٢٧	@	@	@	@	@	@	٠	@	@	@
١٣٠	الرَّحْمٰنِ	الرَّحْمٰنِ	رَحْمٰنِ	رَحْمٰنِ	الرَّحْمٰنِ	الرَّحْمٰنِ	٧٧	رحم	رَحْمٰنِ	رحم
١٤٠	الرَّحِیْمِ	الرَّحِیْمِ	رَحِیْمِ	رَحِیْمِ	الرَّحِیْمِ	الرَّحِیْمِ	٧٧	رحم	رَحِیْمِ	رحم
١٥٧	@	@	@	@	@	@	٠	@	@	@
١٦٠	مٰلِكِ	مٰلِكِ	مٰلِكِ	مٰلِكِ	مٰلِكِ	مٰلِكِ	٥٩	ملك	مٰلِكِ	ملك
١٧٠	یَوْمِ	یَوْمِ	یَوْمِ	یَوْمِ	یَوْمِ	یَوْمِ	٣٨	يوم	یَوْمِ	يوم
١٨٠	الرَّذِیْنِ	الرَّذِیْنِ	ذِیْنِ	ذِیْنِ	الرَّذِیْنِ	الرَّذِیْنِ	٢٩	دين	ذِیْنِ	دين
١٩٧	@	@	@	@	@	@	٠	@	@	@

الشكل (٤) قاعدة بيانات الصرف

Wazin	oth	no_shada	no_beg	no_sh_end	word	root	verb	root2
فعلان	الرَّحْمٰنِ	الرَّحْمٰنِ	رَحْمٰنِ	رَحْمٰنِ	الرَّحْمٰنِ	رحم	رَحْمٰنِ	رحم
فعلیل	الرَّحِیْمِ	الرَّحِیْمِ	رَحِیْمِ	رَحِیْمِ	الرَّحِیْمِ	رحم	رَحِیْمِ	رحم
فاعلین	الرَّعٰلَمِیْنَ	الرَّعٰلَمِیْنَ	عٰلَمِیْنَ	عٰلَمِیْنَ	الرَّعٰلَمِیْنَ	علم	عٰلَمِیْنَ	عالم
فاعل	مٰلِكِ	مٰلِكِ	مٰلِكِ	مٰلِكِ	مٰلِكِ	ملك	مٰلِكِ	ملك

الشكل (٥) قاعدة البيانات النحوية

علامة الإعراب	الإعراب	no_sh_end	Word
		#	#
		@	@
	حرف جر	بِ	
الكسرة	اسم مجرور وهو مضاف	اسْمِ	بِسْمِ
الكسرة	مضاف إليه	اللَّهِ	اللَّهِ
الكسرة	صفة	الرَّحْمَنِ	الرَّحْمَنِ
الكسرة	صفة	الرَّحِيمِ	الرَّحِيمِ
		أَبْتَدَأُ	
نهاية جملة	+	+	+
		@	@
الضمة	مبتدأ مرفوع	أَلْحَمْدُ	أَلْحَمْدُ
	حرف جر	لِ	
الكسرة	إسم مجرور	اللَّهِ	بِاللَّهِ
الكسرة	صفة وهو مضاف	رَبِّ	رَبِّ
الياء والنون	مضاف إليه	العَالَمِينَ	العَالَمِينَ
		@	@
الكسرة	صفة	الرَّحْمَنِ	الرَّحْمَنِ
الكسرة	صفة	الرَّحْمَنِ	الرَّحْمَنِ
		@	@
الكسرة	صفة وهو مضاف	مَالِكِ	مَالِكِ
الكسرة	مضاف إليه وهو مضاف	يَوْمِ	يَوْمِ
مبني	مضاف إليه	الدِّينِ	الدِّينِ
		@	@
نهاية جملة	+	+	+
مبني	مبتدأ وهو مضاف	إِيَّاكَ	إِيَّاكَ
ضمير متصل مبني	مضاف إليه	كَ	
الضمة	فعل مضارع مرفوع	نَعْبُدُ	نَعْبُدُ
	فاعل مستتر مقدر	نَحْنُ	
+	+	+	+
	حرف عطف		وَ
مبني	مبتدأ وهو مضاف	إِيَّاكَ	إِيَّاكَ
ضمير متصل مبني	مضاف إليه	كَ	
الضمة	فعل مضارع	نَسْتَعِينُ	نَسْتَعِينُ
	فاعل مستتر	نَحْنُ	

الشكل (٦) قاعدة بيانات الكلمات مفصلة مع اللواحق

WORD	ROOT	K_A	PRE1	PRE2	CENTER2	SIGN1	LAST1	LAST2	OTH	K_B
#	#				#				#	1
@	@				@				@	2
بِسْمِ	سمى	11602	بِ		اسْمِ	◌ِ			بِسْمِ	٣
اللَّهِ	له	12011			اللَّهِ	◌ِ			اللَّهِ	٤
الرَّحْمَنِ	رحم	١١٨٠٦	ال		رَحْمَنِ	◌ِ			الرَّحْمَنِ	٥
الرَّحِيمِ	رحم	١١٦٨٢	ال		رَحِيمِ	◌ِ			الرَّحِيمِ	٦
@	@				@				@	٧
الْحَمْدُ	حمد	٧٥٢٢	ال		حَمْدُ	◌ُ			الْحَمْدُ	٨
بِاللَّهِ	له	١٢٠٠٨	لِ		اللَّهِ	◌ِ			بِاللَّهِ	٩
رَبِّ	رب	١٠٣٧٧			رَبِّ	◌ِ			رَبِّ	١٠
الْعَالَمِينَ	علم	٦٢٦١	ال		عَالَمِ	بَيْنَ			الْعَالَمِينَ	١١
@	@				@				@	١٢
الرَّحْمَنِ	رحم	١١٨٠٦	ال		رَحْمَنِ	◌ِ			الرَّحْمَنِ	١٣
الرَّحِيمِ	رحم	١١٦٨٢	ال		رَحِيمِ	◌ِ			الرَّحِيمِ	١٤
@	@				@				@	١٥
مَالِكِ	ملك	١١٤٦٥			مَالِكِ	◌ِ			مَالِكِ	١٦
يَوْمِ	يوم	١١٦١٧			يَوْمِ	◌ِ			يَوْمِ	١٧
الدِّينِ	دين	١١٩٦٣	ال		دِينِ	◌ِ			الدِّينِ	١٨
@	@				@				@	١٩
إِيَّاكَ	علي	٤٢٥٣			إِيَّاكَ		كَ		إِيَّاكَ	٢٠
تَعْبُدُ	عبد	٧٥٤٧			تَعْبُدُ	◌ُ			تَعْبُدُ	٢١
وَإِيَّاكَ	علي	٤٢٥٤	وَ		إِيَّاكَ		كَ		وَإِيَّاكَ	٢٢
نَسْتَعِينُ	عون	٨٧٥٤			نَسْتَعِينُ	◌ُ			نَسْتَعِينُ	٢٣
@	@				@				@	٢٤