تصميم نظام ترجمة لفظية آلية باللغة الانكليزية للقرآن الكريم

سهام حسن خيون
قسم علوم الحاسبات/كلية العلوم/جامعة البصرة

الخلاصة Abstract

لقد قمنا بتصميم نظام ترجمة لفظية آلية باللغة الانكليزية للقرآن الكريم لغير متكلمي اللغة العربية ليستطيع المُستخدِم قراءة القرآن من دون حاجة لتعلم اللغة العربية لأنها صعبة التعلم لما تمتاز به من تشعبات وتداخلات وترادفات في المعنى.

ان تصميم النظام المقترح تضمن مرحلتين:الاولى عمل دراسة نظرية للقرآن الكريم لاستخلاص مجموعة من المعلومات ذات فائدة في بحثنا هذا ،اما الثانية فهي مرحلة تصميم النظام وأول خطوة في هذه المرحلة هي تصميم معجم يحتوي النص القرآني المشكل بشكل وقائمة مرتبطة بكل الجذور التي وردت في القرآن الكريم وايضاً كل حروف اللغة العربية المُشَكَلة ومايقابلها باللغة الانكليزية فضلا على معلومات أخرى تساعد في المعالجة. لقد صمم النظام ليسلك مسارين شبه منفصلين اعتمادا على نوع البيانات المدخلة اخذين بنظر الاعتبار ان المستخدم لايعرف اللغة العربية.

المسار الاول تكون المدخلات فيه عبارة عن اسم السورة وعدد الآيات وتكون المعالجة في هذه الحالة على مرحلتين: المرحلة الاولى ايجاد النص المحدد بأستعمال خوارزميات بحث والمرحلة الثانية هي ترجمة النص المطلوب لفظياً.

اما المسار الثاني فالمدخلات عبارة عن نص قرآني غير مشكل لذلك ستجري مجموعة من المعالجات على النص لأعادة تشكيله واول معالجة هي استدعاء محلل مفردات للتأكد من سلامة النص من الرموز غير العربية ومن ثم محلل صرفي يقوم بتحليل الكلمة لغرض اضافة علامات التشكيل المناسبة لها على وفق مفهوم الوزن الصرفي فعندما قمنا بدراسة نظرية للغة لاحظنا أن الكلمة المجردة يمكن تشكليها بأعتماد وزنها الصرفي والتحليل الصرفي على مرحلتين:الاولى ايجاد الجذر والوزن الصرفي للكلمة المدخلة والثانية مرحلة المطابقة للحصول على الكلمة الصحيحة المُشَكَلة واخيرا تأتي مرحلة الترجمة لكلمات النص لفظيا الى مايقابلها باللغة الانكليزية.

كلمات المفاتيح
معالجة اللغات الطبيعية حاسوبيا, الترجمة اللفظية، القرآن الكريم، اللغة العربية

1. المقدمة Introduction

بعد ولوج الحاسوب في كل مجالات الحياة ،اصبح هذا الجهاز هو الناصح والمرشد والمصمم والمترجم وغيرها من المعاني في مجالات الحياة كافة، كثيرا مانسمع بانتشار الثقافات وأصبح من يسكن الشرق قادرا على معرفة ملايين المعلومات عن شخص يسكن في اقصى بقاع الارض ليس الكتاب وحده من ساهم في نشر هذه الثقافة وانما بالدرجة الاساس الحاسوب عبر ما يدعى بشبكة الانترنت العالمية يستطيع الانسان الاطلاع على الثقافات والاديان الاخرى وكنتيجة لهذا التطور المعلوماتي دخل الاسلام الكثير ممن لايتكلمون اللغة العربية (لغة القرآن الكريم) واصبحت الحاجة ملحة الى تعلمهم القرآن وحفظه دون تعلم اللغة العربية فكما لا يخفى على الكثير ان اللغة العربية من اكثر اللغات تشعبا وتداخلا وترادفا في المعنى لذلك يتطلب تعلمها جهدا ووقتا طويلا لذلك ارتأينا ان نصمم نظاما لترجمة القرآن لفظيا باللغة الانجليزية لكي يتسنى للمُستخدِم قراءة القرآن وحفظه دون تعلم اللغة العربية.

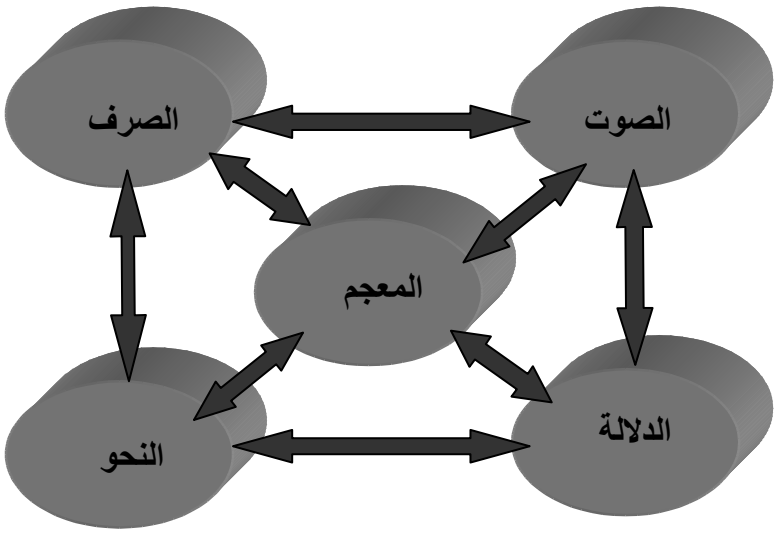
2. معالجة اللغات الطبيعية حاسوبيا Natural Language Processing

تعد معالجة اللغات الطبيعية واحدة من الجوانب المهمة في الذكاء الاصطناعي إذ إنها تعتمد على النماذج اللسانية وتكنولوجيا الحاسوب،فمعالجة اللغات الطبيعية تزداد تقدماً بتطور علوم الحاسبات واتصالات الحاسبات التي سببها الانترنيت وظهور حوسبة المعلومات [1].

وهنا يمكن تعريف نظام معالجة اللغة الطبيعية بأنه دالة طبيعية متعددة القيم ذات اتجاهين بين مجموعة المعاني الطبيعية Meanings ومجموعة التراكيب اللغوية للمعنى الواحد[2].

إن الهدف الاساس الذي يقف وراء بحوث معالجة اللغات الطبيعة هو تمكين الحاسوب من فهم اللغات الطبيعية وجعل هذه الآلة لها القدرة على الاتصال مع الانسان من دون برمجة(من دون استخدام لغات البرمجة) وباستعمال لغته الطبيعية، وذلك لتقليل الجهد والوقت الذي يستهلكه المبرمج في تعلم لغة برمجة معينة للتعامل مع الحاسوب.

يتكون نظام معالجة اللغات الطبيعية من خمس منظومات اساسية كما هو موضح في الشكل (1).وقد اقتصر بحثنا هذا على منظومة الصرف المعجم وتتضمن الترجمة اللفظية بعض مفاهيم منظومة الصوت، اما بالنسبة للنحو والدلالة لم نتعامل معهما لانتفاء الحاجة بسبب الاحتفاظ بكتاب القرآن الكريم في المعجم.



الشكل (1) يبين مكونات منظومة معالجة اللغات الطبيعية

1-2 الصوت Phonology :الصوت اللغوي هو اكثر الظواهر تأصلاً .ان ظاهرة الكلام عبارة عن نشاط حي ديناميكي يتصف بالتدفق والتلون واللامحدودية ويتفاعل فيه، ومعه وبشكل حاد

عناصر متعددة مثل الظرف والبيئة والخلفية والمتحدث والمستمع [3]. نظام معالجة الصوت مسؤول عن اخذ الاصوات وترجمة الادخال الى كلمات [4].

2-2 الصرف Morphology :الصرف هو فرع اللسانيات الذي يتعامل مع البنية الداخلية للكلمات من حيث تكوين عناصرها الاولية والتفاعل بين هذه العناصر [3]، إذ يهتم الصرف بالدرجة الاساس بمعالجة الكلمة بغض النظر عن موقعها في الجملة وهي اول مرحلة في معالجة اللغة العربية حاسوبياً[5].

3-2 النحو Syntax : النحو عند اللسانين هو علم بأصول يعرف به تغيير اواخر الكلمات اعراباً وبناءً وغاية هذا العلم صون اللسان عن اللحن في الكلام وبه يستعان على فهم الكتاب (القرآن الكريم) والسنة النبوية وكلام العرب الاوائل [6].اما عند علماء اللسانيات الحاسوبية فيعرف على انه دراسة بنية الجملة وذلك من حيث ترتيب عناصرها او مكوناتها او العلاقات التركيبية (البنائية) والوظيفية التي تربط بين هذه العناصر[3].

4-2 المعجم Lexicon :هو عبارة عن قائمة بالكلمات التي ترتبط فيها كل كلمة من هذه الكلمات مع سماتها الدلالية والنحوية والصرفية فضلاً على سمات أخرى [7]. أصبحت أهمية المعجم تزداد يوماً بعد آخر ابتداءً من عقد الثمانينيات من القرن الماضي بسبب وجوده صفة اساسية في كل انظمة معالجة اللغات الطبيعية.

5-2 الدلالة Semantic :المقصود بالدلالة تحديد معنى الجملة ،وهو العلم الذي يدرس الشروط الواجب توافرها في العنصر (اسم او فعل او حرف)حتى يكون قادراً على حمل المعنى، والتحليل الدلالي فرعان، الاول يهتم ببيان معاني الكلمات وهنا يرتبط مع المعجم، اما الثاني فيهتم بمعاني الجمل والعبارات [7،8].ان فهم معنى الجملة اللغوية يتشكل من تجميع معاني الكلمات المكونة للجملة [2].وان معنى الكلمة يعتمد على المعرفة المرتبطة بهذه الكلمة وايضاً بقية المعاني التي سترتبط بالكلمة او بالجملة نفسها خلال المعالجة. وللدلالة دور مهم في رفع الغموض عن الكلمات او الجمل.

3. **الترجمة Translation**

ان ظاهرة الترجمة كانت نتيجة الظاهرة اللغوية لدى البشر ، فما ان تفرق البشر الى شعوب وقبائل وتطورت لديهم الظاهرة اللغوية حتى برزت الحاجة للترجمة لتوصيل الافكار والمشاعر والرغبات ولتحقق التفاهم الذي هو الوظيفة العليا للغة.

ان اكثر التطورات تنوعا وتسارعا قد حصلت في اساليب انجاز الترجمة، فبعدما كانت مفردتا "مترجم" و"ترجمة" تشيران الى الانسان وانجازه في هذا الحقل ،دخلت الآلة لتقدم العون الى الانسان في الترجمة او لتطلب معاونته في انجازها.[9]

إن الترجمة تأخذ محورين اولهما ترجمة معنى النص من لغة إلى لغة أخرى وهو ما متعارف عليه في مفهوم الترجمة اما المحور الآخر فيهتم بطريقة لفظ كلمات النص الاصلي (الفونيمات الصوتية) وتسمى ترجمة لفظية، تستعمل مع اسماء الاعلام او النصوص(مثل القران الكريم) المراد الحفاظ عليها عند نقلها الى لغة اخرى في الاول تكون الترجمة كلمة من لغة المصدر مقابل كلمة من لغة الهدف او بأستخلاص معنى الجملة ككل وترجمته الى اللغة الهدف اما في الثاني فتتم الترجمة حرفيا (حرف-حرف) للحفاظ على فونيمات النص الاصلي.

4. **النظام المقترح**

يعتمد تصميم النظام المقترح على مرحلتين هما:

المرحلة الاولى First Stage:

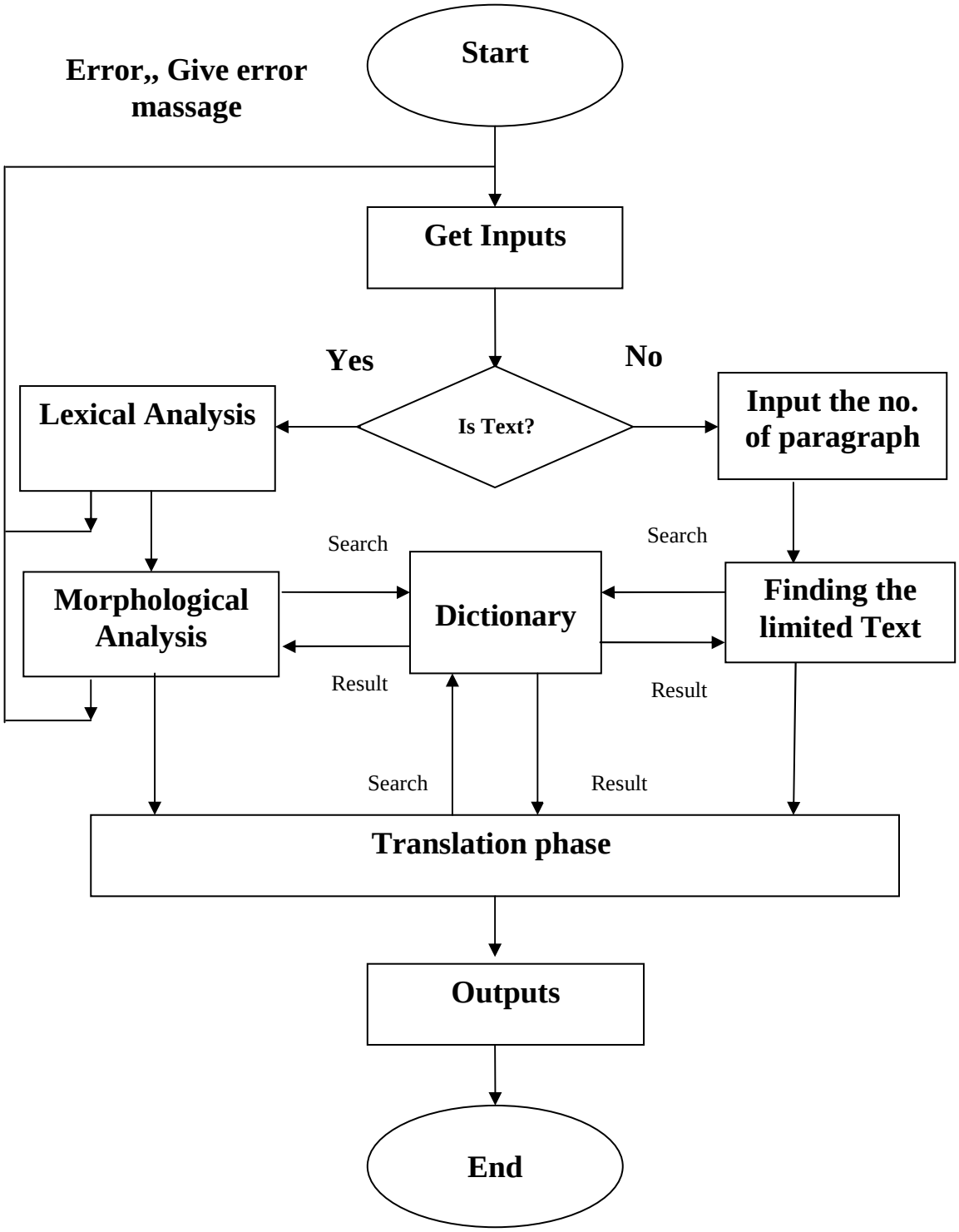
تضمنت دراسة نظرية لكتاب القرآن الكريم وكانت النتائج المستخلصة كمايلي:

- ان جذور الكلمات التي يتكون منها القرآن الكريم هي احادية وثنائية وثلاثية و...الخ وكانت نسبة الجذور الثلاثية هي الاكبر من بين بقية الجذور.
- تم تحديد جميع الاوزان الصرفية التي يمكن استخلاصها من كل جذور القرآن الكريم فمثلا(كتب) جذر ثلاثي ورد في القرآن بالكلمات التالية(كتب، كتاب، كاتب، مكتوبة،...)حيث تأخذ هذه الكلمات الاوزان التالية وحسب التسلسل (فعل، فعال، فاعل، مفعولة،...).
- احصاء اسماء الاعلام التي وردت في القرآن وتشمل أسماء أشخاص ومدن وأماكن ووقائع.

المرحلة الثانية Second Stage :

تضمنت تصميم البرنامج المقترح استنادا الى المعلومات التي استخلصت في المرحلة السابقة وقد استعملنا لغة البرمجة C++ لبرمجة النظام المقترح والشكل (2) يوضح المخطط الكتلي للنظام المقترح. ان مدخلات النظام اما ان تكون عبارة عن اسم السورة ومن ثم يحدد رقم الآية او مجموعة الآيات او كل السورة التي يرغب في ترجمتها، أو نص قرآني غير مشكل(التشكيل يعني اعطاء حركات لحروف الكلمة) آخذين بنظر الاعتبار ان معظم مستخدمي النظام من غير متكلمي اللغة العربية.اما المخرجات فتكون عبارة عن نص قرآني مكتوب باللغة الانجليزية مترجم لفظيا للنص المطلوب.نظرا لأختلاف نوع المدخلات ينحى النظام مسارين شبه منفصلين فالمسار الاول يعتمد النوع الاول من المدخلات وهي اسم السورة وتكون المعالجة في هذه الحالة مرحلة تحديد النص المطلوب وترسل النتائج مباشرة لطور الترجمة لأن النص القرآني المحفوظ في المعجم يكون مشكلاً ولايحتاج الى معالجة صرفية، اما المسار الثاني الذي يعتمد النوع الثاني من المدخلات فسنحتاج معالجة النص للتأكد من صحته واضافة علامات التشكيل له على وفق نظام صرفي دقيق ومن ثم يأتي طور الترجمة الذي يعطي النتائج النهائية.

قبل الخوض في تفاصيل خوارزمية النظام المقترح سنتكلم اولا على المعجم بوصفه المرجع والاساس وقاعدة البيانات الذي بني على اساسه النظام الصرفي والمترجم.



الشكل (2) يوضح المخطط الكتلي للنظام المقترح

1-4 المعجم Lexicon

من النتائج التي استخلصت من مرحلة الدراسة النظرية وحسب متطلبات النظام المقترح صممنا المعجم ليحتوي على مايلي:

- نص قرآني مشكل(كتاب القرآن الكريم) على شكل ملف نصي Text File يحتوي كل سور القرآن حصلنا عليه من احد برامجيات التفسير(جامع التفاسير). ان فكرة الاحتفاظ بالقرآن في المعجم جاءت للمحافظة على سلامة النص القرآني، كما لايخفى على القارئ إن أي خطأ وان كان بسيطا في نقل النص القرآني يعد غير مقبول شرعاً.
- قائمة مرتبطة Linked List تحتوي كل الجذور التي وردت في القرآن الكريم سواء كانت احادية او ثنائية اوثلاثية او...الخ.مع كل جذر ترفق قائمة مرتبطة اخرى تمثل ارقام الاوزان الصرفية التي اشتقت من هذا الجذر إذ أُعطي كل وزن صرفي رقما معينا وذلك لتسهل معالجته من قبل النظام الصرفي كما ان الاوزان الصرفية المشتقة من الجذر (كتب)على سبيل المثال ليس بالضرورة ان نعالجها كلها انما نعالج التي وردت في القرآن فقط.فمثلا الوزن الصرفي:فعل، رقمه:1 والوزن الصرفي: مفاعيل، رقمه:11 ... وهكذا.والشكل(3) يوضح بعض جذور القرآن الكريم مع اوزانها الصرفية.
- قائمة بكل حروف اللغة العربية ومايقابلها باللغة الانكليزية مع مراعاة علامات التشكيل. والشكل (4) يوضح ذلك.
- ايضا نحتاج قائمة مرتبطة بأسماء الاعلام التي وردت في القرآن حصرا مثل محمد، يونس، موسى، مكة، سبأ، بدر...الخ وذلك لأن المترجم يحتاج معرفة اسماء الاعلام في معالجته النص.

الجذر	الاوزان الصرفية															
	(1)فعل				(2)أفعل			ف	فع	5	6	.	يفعلون	فعلل	فعلول	..
ا								أ								
أحد	أَحَدَ	أَحَدٍ	أَحَدٌ	...												
...																
برز	بَرَزَ															
...																
حلقم															الحُلقُوَم	
...																
زهق	زَهَقَ															
...																
سمع	سَمِعَ	السَّمعَ	السَّمعِ	لِلسَّمعِ	أسَمَع	اَسَمَعُ	أسمِع									
طه									طه							
كتب	كُتِبَ	كُتُبٍ	كُتُبٌ	كَتَبَ	أكتُب	...	...									
...																
هجع													يَهجَعُونَ			
...																
يونس														يُونُس		

الشكل (3) يوضح بعض الجذور وتصريفاتها

الحرف العربي	ما يقابله باللغة الانكليزية
أ	a
أُ	au
ا	a
إ	ae
. .	. .
د	D
دْ	D
دَ	Da
دِ	De
دُ	Du
دٌ	Dun
دّ	Dd
دٍ	Den
. .	. .

الشكل (4) يوضح جدول بالحروف العربية ومايقابلها باللغة الانكليزية

2-4 المسار الاول First Path:

ذكرنا سابقا ان النظام المقترح سيخطو بأتجاهين اعتمادا على نوع المدخلات فاذا كانت المدخلات اسم السورة فستكون المعالجة كمايأتي:

1.**ادخال بعض المعلومات** وهذه المعلومات هي عدد الآيات او رقم الآية اذا كان المطلوب ترجمة آية واحدة أوكل السورة.

2.**ايجاد النص المطلوب Finding the Limited Text** أن المسألة لاتتطلب معالجة صرفية وماشابه وانما استعملنا خوارزمية البحث التسلسلي في النص القرآني المحفوظ في المعجم عن طريق اسم السورة استعملنا هذه الخوارزمية لأن سور القرآن غير مرتبة ابجديا وانما حسب ظهورها في الكتاب الكريم ،ولأن عدد سور القرآن ليس كبيرا فيمكن ان تعطي هذه الخوارزمية نتائج جيدة وبعد الوصول الى السورة المطلوبة استعملنا خوارزمية البحث السريع للبحث عن الآيات المطلوب ترجمتها لأن السورة مرتبة حسب رقم الآية بعد أن حصلنا على النص المطلوب يتم تكوين قائمة بكلمات النص المحدد التي تكون مشكلة اصلا.

3.**طور الترجمة Translation Phase** يأتي طور الترجمة مباشرة بعد المرحلة السابقة وسنتكلم لاحقا على هذه المرحلة عند شرح المسار الثاني لأن هذه المرحلة مشتركة بين المسارين.

3-4 المسار الثاني Second Stage:

عندما تكون المدخلات عبارة عن نص فان ذلك يتطلب مجموعة من المعالجات وكمايأتي

4-3-1 محلل المفردات Lexical Analyzer

لايخفى ان القرآن الكريم يتكون من (114) سورة وكل سورة تحتوي على عدد محدد من الآيات، لقد اعتمدنا النقطة كفاصلة بين آية واخرى للنص المدخل لنظامنا المقترح. يمكن للمستخدم ادخال اي عدد يشاء من الآيات من دون تشكيل لتقليل الجهد عليه ولأنه لايعرف اللغة العربية جيدا وانما مجرد نقل نص من الكتاب الى الحاسوب.

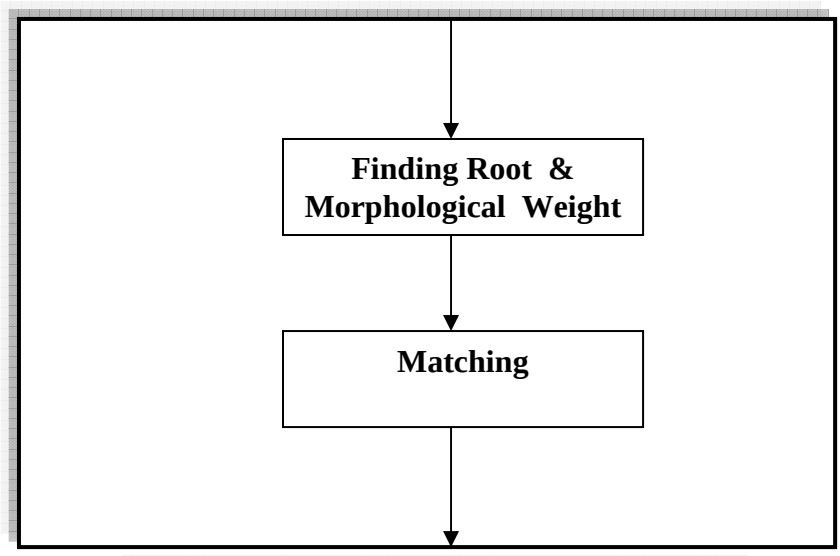
يعمل محلل المفردات على تقطيع النص المدخل وفحص حروفه فاذا كان هناك رمزا لاينتمي للغة العربية يعطي رسالة خطأ "ان النص المدخل غير صحيح" ويطلب اعادة المحاولة والا يقوم المحلل بتكوين قائمة مرتبطة لكل الكلمات التي وردت في النص إذ يعتمد المحلل الفراغ بين كلمة واخرى وستكون هذه القائمة مدخلات للمرحلة اللاحقة وهي المحلل الصرفي.سنأخذ مثالا مفصلا لتوضيح نتائج كل مرحلة على حدة.

المدخلات:"الحمد لله رب العالمين.الرحمن الرحيم."

المخرجات: ["الحمد" ، "لله" ، "رب" ، "العالمين" ، "." ، "الرحمن" ، "الرحيم" ، "."]

4-3-2 المحلل الصرفي Morphological Analyzer

ان النص المدخل للنظام المقترح غير مشكل والقاريء يعلم أن لفظ كلمات اللغة العربية يعتمد على علامات التشكيل الملحقة بالكلمة لأن الكلمة المجردة من علامات التشكيل يمكن ان تأخذ عدة معان وعدة قراءات (مثلا "كُتُب" و "كَتَبَ" يتم ادخالهما "كتب" في حين الاولى اسم وتعني جمع "كتاب" والثانية فعل وتعني قام بفعل الكتابة)ويعد هذا تشويه للقرآن لذلك سيقوم النظام الصرفي باضافة علامات التشكيل ورفع الالتباس والشكل(5) يوضح مهام هذا المحلل .



الشكل (5) يوضح مهمام المحلل الصرفي

1.استخلاص الجذر والوزن الصرفي Finding Root &Morphological Weight: يقوم النظام باستخلاص جذر الكلمة(هو مدخل الكلمة في المعجم المقترح) ليستطيع البحث في المعجم لأن الاخير لايحتوي على كلمات وانما جذور فضلا عن معلومات اخرى كما ذكرنا سابقا عندما تكلمنا على المعجم.أما بالنسبة للوزن الصرفي فيتم استخلاصه لإضافة علامات التشكيل للكلمة المدخلة المجردة. ويتم استخلاصهما وفق الخطوات التالية:

1. البحث في المعجم هل توجد مطابقة للكلمة مع قائمة الجذور اذا حصلت المطابقة يحفظ مع الكلمة جميع الكلمات ذات الاوزان الصرفية المطابقة استعملنا خوارزمية البحث السريع Quick Sort لأن حجم قائمة الجذور كبير جدا.
مثال:الكلمة المدخلة: **(كتب)**
سيحدث تطابق لأن للكلمة نفس حروف الجذر ويستنتج الوزن الصرفي وهو "فعل" وبما ان الكلمة المدخلة غير مشكلة فيمكن ان تأخذ الشكل "فَعَلَ" او "فُعِلَ" او...الخ لذلك سيتم الاحتفاظ بجميع الاوزان المطابقة ذات الاشكال المختلفة فتكون المخرجات **["كَتَبَ" ،"كُتِبَ" ،"كُتِّب"،.....]**
2. اذا فشلت الخوارزمية اعلاه في ايجاد الجذر فهذا يعني ان الكلمة المدخلة تحتوي على زوائد او فيها خطأ املائي وهنا يقوم النظام بأستدعاء خوارزمية مقترحة تقوم برفع جميع الزوائد المحتملة للكلمة لأيجاد الجذر. تبنى فكرة هذه الخوارزمية على دراسة كل وزن صرفي وتحديد الحروف الزائدة وموقعها نسبة للكلمة إذ يكون الاساس حروف الجذر الاصلية وكل مايضاف لها فهو من الزوائد مثلا يكون وزن الجذر الثلاثي **"فعل"** وكمثال على جذر ثلاثي **"سمع"** فاذا كانت الكلمة المدخلة هي**"السميع"** فأن الزوائد هي **"ال"** و"ي" ويشتق النظام الوزن الصرفي لهذه الكلمة وهو**"فعيل"** لأن الحرف المزيد "ي" جاء بعد **"ع"** في الوزن الصرفي التي تقابل **"م"** في الجذر **"سمع"**.بعد ان حصلنا على جذر الكلمة يمكن البحث في المعجم بالطريقة السابقة نفسها فاذا حصلت مطابقة فيقوم النظام بحفظ جميع الكلمات ذات الوزن الصرفي الذي اشتق مع الكلمة المدخلة والا يتوقف ويعطي رسالة خطأ مناسبة ويطلب اعادة المحاولة في ادخال النص.

ان مخرجات هذه المرحلة للمثال السابق هي
["الحَمـــدُ" ، "لِلَّهِ" ، ["رَبَّ"، "رَبِّ" ، "رَبُّ" ، "رَبِّ" ، "رَبٌّ"]، "العـــالَمِينَ" ، "." ، ["الرَّحمنِ" ، "الرَّحمَنَ" ، "لِلرَّحمَنِ" ، "الرَّحمَنِ" ، "الرَّحَمنُ"] ، ["الرَّحِيمِ" ، "الرَّحِيمُ" ، "رَحِيمٍ"] ، "."]

2.مرحلة المطابقة Matching:بعد ان حصلنا في المرحلة السابقة على مجموعة من الاوزان الصرفية المطابقة للكلمة المدخلة يتم في هذه المرحلة ايجاد الكلمة ذات الوزن الصرفي المناسب من بين المجموعة التي تم الحصول عليها استعملنا خوارزمية البحث التسلسلي مع مفهوم العودة ثانية Back Tracking إذ تُطبقّ الأوزان مع النص القرآني المحفوظ واذا حصلت مطابقة يتم التحقق هل الكلمة التالية تطابق ما موجود في النص القرآني وهكذا الى ان يتم التأكد من جميع كلمات النص المدخل. ان عملية البحث والمطابقة مع النص القرآني يأخذ جميع الاحتمالات المتوفرة لكل كلمة عند حصول تطابق يتقدم الى الامام لأتمام عملية المطابقة وعند عدم حصول مطابقة يعود الى الخلف لأختبار الاحتمال التالي وحين استنفاذ جميع الاحتمالات لكل الكلمات ولم تحصل مطابقة شاملة للنص يتوقف ويعطي رسالة خطأ مناسبة والا يكون قائمة مرتبطة تحتوي على كل كلمات النص المدخل بعد اضافة علامات التشكيل.
مخرجات هذه المرحلة للمثال السابق هي
["الحَمدُ" ، "لِلَّهِ" ، "رَبِّ" ، "العالَمِينَ" ، "." ، "الرَّحمنِ" ، "الرَّحِيمِ" ، "."]

3-3-4 طور الترجمة Translation Phase

إنّ مدخلات هذه المرحلة هي عبارة عن قائمة مرتبطة بكلمات النص المراد ترجمته سيقوم المترجم بتقطيع كل كلمة الى حروفها وكل حرف مع حركته اخذين بنظر الاعتبار ان الحرف الذي لايحمل حركة هو حرف ساكن (مثلا **"كُتِبَ"** تقطع الى **["كُ"،"تِ"،"بَ"]**) ومن ثم يقوم المترجم بترجمة كل حرف الى مايقابله في المعجم من حرف او مقطع باللغة الانكليزية (لنفس الكلمة في المثال السابق تكون **["ba", "te", "ku"]**) بعد ذلك يبحث عن الكلمة في المعجم فاذا كانت اسم علم يغير الى الحرف الاول الى حرف كبير وكذلك يفعل اذا كانت الكلمة الاولى في الآية ومن ثم يقوم بدمج الكلمة(**"kuteba"**) وهكذا يفعل لكل كلمات القائمة المرتبطة.

خطوات ترجمة المثال السابق

1. تقطيع الكلمات إلى حروفها وستكون المخرجات [["ا" ، "ل" ، "حَ" ، "م" ، "دُ"]، ["لِ" ، "لَّ" ، "هِ"] ، ["رَ" ، "بِّ" ،....].
2. ترجمة كل حرف الى مايقابله باللغة الانكليزية فتكون المخرجات [["a" , "l" , "ha" , ,.....] ["du" , "m" ,
3. تحويل الحرف الاول في الآية الى حرف كبير وكذلك اذا كان اسم علم.
[["A" , "l" , "ha" , "m" , "du"] ,...., "." ,["A" , "l" , "ra" , "h" , "ma" , "ne"],...,"."]
4. تجميع الحروف الى كلمات مجددا. وستكون المخرجات النهائية كما يلي

Alhamdu lelahe rabe ala'almeena. Alrahmane alraheeme.

5. الاستنتاجات

عندما صممنا النظام المقترح فرضنا المدخلات عبارة عن نص قرآني مشكل مع عدم الاحتفاظ بالنص القرآني في المعجم لاحظنا أن الأمر يتطلب معالجة صرفية ونحوية ودلالية ويمكن الحصول على نتائج جيدة لكن حتما ستكون هناك نسبة خطأ ويعد هذا تشويها للقرآن الكريم وهذا لا يجوز شرعا لأنه عند التعامل مع النص القرآني لايمكن تقبل اي خطأ وان كان بسيطا اما اذا احتفظنا بالنص القرآني في المعجم وفرضنا المدخلات نصاً مشكلاً أيضا فستكون المعالجة يسيرة ولاتتطلب محللا صرفيا ونحويا ودلاليا، لكن إدخال نص مشكل يعد عبئا ثقيلا على من يجيد اللغة العربية فكيف بمن لا يجيدها لذلك ارتأينا ان تكون المدخلات عبارة عن نص غير مشكل ويقوم النظام المقترح باعادة تشكيله مع الاحتفاظ بالنص القرآني المشكل لسلامة القرآن من التشويه والتحريف. لكن الذي لايتكلم العربية لا يجيد استخدام الحروف العربية لذلك جعلنا النظام يسلك مسارين فمن الممكن ان يستعمل النظام للترجمة اللفظية بدون ادخال النص وانما بمجرد ادخال اسم السورة وتحديد عدد الآيات المطلوبة فاصبح النظام يخدم فئة اكبر من السابق وكانت النتائج التي حصلنا عليها جيدة جدا، وإن النظام المقترح قد عالج نصوص القرآن بنسبة 100%.

6. العمل المستقبلي

يمكن تطوير هذا النظام لنصوص أُخرى لغرض ترجمتها لفظيا من والى العربية وذلك للمساعدة على تعلم لفظ كلمات اللغات الاخرى ،لكن ستتطلب المسألة معالجة نحوية ودلالية بالاضافة الى معالجة صرفية.

المصادر

1. G. H. S AL-Gaphari,. "A Constraint – Based Object Lexicon for Supporting Natural Language processing (Arabic Language)" Ph.D. Dissertation, Iraq, 1999.

2. E. Gibson, & N. J. Pearlmutter, "Constraints On Sentence Comprehension ", Cognitive Sciences, July, Volume 2 No. 7, 1998.

3. نبيل علي " اللغة العربية والحاسوب "، مؤسسة التعريب للنشر ، القاهرة ، 1988.

4. E. Charniak, & D. McDermott, "Introduction to Artificial Intelligence" ,Addison-Wesley Publishing Company, 1988.

5. محمد نعمان مراد " اللسانيات واللغة العربية والتطبيقات الحاسوبية "، مجلة آفاق عربية، مايس – حزيران ، السنة الثالثة والعشرون.

6. احمد قصير العاملي " متن الاجرومية ودروس في النحو "، دار مكتبة الهلال ، بيروت، 1985.

7. G. Gazdar, & Ch. Mellish, "Natural Language Processing in Prolog: An Introduction to Computational Linguistics ", California, USA, Addison- Wesley Publishing Company, 1989.

8. نايف خرما " اضواء على الدراسات اللغوية "، مطبعة عالم المعرفة، الكويت، 1978.

9. سلمان داود الواسطي "التفاعل بين الانسان والآلة في الترجمة الحاسوبية" ، كلية الآداب ،الجامعة المستنصرية ،2002.

Abstract:

An automatic verbal translation system for the holy Quran into English Language is designed to help users who can not speak Arabic Language in order to read the holy Quran without necessity to learn Arabic Language.

The suggested system consist of two stages, the first is studying the holy Quran to emerge many features which are necessary to process the holy text. The second stage is designing the system, the first step in this stage we designed a special lexicon which involves diacritical text of Quran as text file, roots of words, a linked list of all diacritical letters in Arabic and their mapping in English.

The system have two tracks depending on the type of inputs. The first track when input is the name of Surah, then the process is to find the specified text in the lexicon and translate it verbally. The other track when the input is non-diacritical text, the suggested system will translate the entire text during several phases. The first phase is **Lexical Analysing** to be sure that the text has not any non-Arabic characters, the second phase **Morphological Analysing** to make diacritical text by some steps which are **Finding Root, Morphological Pattern,** and **Matching**. Then the last phase is translating the text into English verbally.