نظـام آلـي لفهـم اللغـة العربيـة

ميساء عبد الكريم ناصر كلية العلوم–جامعة البصرة
مرتضى عباس فالح كلية التربية–جامعة البصرة
صبا عبد الواحد كلية العلوم–جامعة البصرة

الخلاصة

أن المعالجة الحاسوبية للغة العربية تتطلب مجموعة من المعالجات المفرداتية والصرفية القوا عدية والدلالية والتي من خلالها يمكن الوصول لتصميم وبناء نظام قواعد محوسب يمكن المتلقي(المتعلم)للغة العربية من فهم الجملة التي تتلاءم دلالتها في سياق محدد،بحيث تحمل نفس المعنى المراد تبليغه للمتعلم.

يستثمر هذا البحث مفاهيم وأفكار النظريات الدلالية والقوا عدية للغة العربية والحاسوبية لبناء نظام حاسوبي يعمل على تحليل وتوليد بعض الجمل العربية البسيطة ,من خلال تحليل الجملة المدخلة (جملة المصدر)إلى مكوناتها الرئيسة وصولا لإيجاد مقابلاتها في جملة الهدف(الجملة المرادفة),فضلا عن شرح لمعاني المفردات وإبراز المفاهيم الأساسية للغة العربية بأجراء معالجات حاسوبية تستند في عملها بشكل أساس على المعجم اللغوي الحاسوبي.

Abstract

The computing process for the Arabic language requires group of indicative, grammatical, pure, and vocabulary processes by which is possible to design and make a data base enables the learner of Arabic language to understand the sentence that convenience indicative in a limited context has the same meaning which shall be told to the learner.

This research generates concepts and ideas of the indicative and grammatical theories of the Arabic and computing language to make a computing system works on analyzing and generating some of the simple Arabic, sentence through analyzing(infinitive sentence) to its main contents and in order to find its synonyms in the (synonymous sentence) further to explain the words meaning and to give the basic concepts for the Arabic language by making a computing process mainly depend on the computing linguistically lexicon.

المقدمة

معالجـة اللغـات الطبيعيـة تتطلـب الإلمـام بنـوعين مـن العلوم:العـلوم الإنسـانية بنوعيهـا القديمـة والحديثة,والعلوم الحاسوبية التـي تهتم بمعالجـة اللغات الطبيعية وهندستها.والتي تتوخى فتح حوار علمي بين الحاسوب وتطبيقاته والإنسان العربي بوساطة اللغة العربية بوصفها لغة طبيعية قابلة للمعالجة الآلية وهذا لن يتأتى إلا بتعميق النقاش الجاد والمثمر في القضايا الأساسية التي تطرحها المعالجة الحاسوبية للنظام العربي في سائر مستوياته الصرفة,النحوية,الدلالية,المعجمية,والبلاغية من قبيل التوليد الآلي,التحليل الآلي,اكتشاف الأخطاء الآلية للنصوص,الاسترجاع الآلي للنصوص,الاسترجاع الآلي للمعارف,الترجمة الآلية,والفهم الآلي للنصوص(عبد الستار ونعمان، 1992؛ مهديوي، 2006).

إن المعالجة الآلية للغة العربية تقتضي بالضرورة معالجة مفهوم مفرد (الكلمة) من خلال مفهوم متعدد إلى معالجة جزئية حتى المعالجة التامة إذ الانتقال إلى المفاهيم المتعددة باستعمال التصنيف المعجمي لتمييز المعاني المختلفة(عبد الباقي، 2001؛ علي، 1988؛Thompson & Mooney, 2003) .

تتضمن معالجة اللغات الطبيعية عدة مراحل هي:المعالجة الصوتية (Phonology)والمعالجة الصرفية (Morphology) ,والمعالجـة النحويـة (Syntax),المعالجـة الدلاليـة(Semantics),المعالجـة المقاميـة أو الذرائعية(Pragmatics)وأخيرا المعالجة المعجمية(Lexicon).(ضمد، 2006).

الترادف في اللغة العربية

الترادف(Synonymy)

الترادف في اللغة العربية هو ما اختلف لفظه واتفق معناه(المبارك، 1986). بمعنى امتلاك كلمات مختلفة المعنى نفسه فالكلمات(أرشد,هدى) يكون لها المعنى نفسه تقريباً في سياقات معينة,ويمكن تعريف الترادف على أنه:تضمن من جانبين (أ) و(ب) يكونان مترادفين إذا كان (أ) يتضمن (ب) و(ب) يتضمن (أ) مثل كلمة أم=والدة,أب=والد(ضمد، 2006) .

ويمكن إجمال أنواع الترادف على النحو الأتي(عمر، 1982):-

أ.الترادف الكامل(Complete Synonymy)وذلك حين يتطابق اللفظان تمام المطابقة.وهذا النوع عادة غير وارد.

ب.شبه الترادف(near Synonymy) أو ما يسمى التشابه(likeness) وذلك حين يتقارب اللفظان مثل(عام-سنة-حول,...)

ج.التقارب الدلالي(Semantic relation)يتحقق ذلك حين تتقارب المعاني,لكن يختلف كل لفظ عن الآخر بملمح مهم واحد على الأقل.مثل كلمتي(حلم,رؤيا) وهما من الكلمات القرآنية.

د.الاستلزام(Entailment) وهو قضية الترتيب ... ويمكن أن يعرف كما يأتي:س يستلزم سَ إذا كان في كل المواقف الممكنة التي يصدق فيها س يصدق كذلك سَ.

مثال:قام محمد من فراشه الساعة العاشرة فإن هذا يستلزم: كان محمد في فراشه قبل العاشرة مباشرة.

هـ.استعمال التعبير المماثل: أو الجمل المترادفة(Paraphrase)وذلك حين تمتلك جملتان المعنى نفسه في اللغة الواحدة(خلف، 2001). وقد قسم Nielsen هذا النوع:

1.التحويلي,وذلك بتغيير مواقع الكلمات في الجملة,بقصد إعطاء بروز لكلمة معينة.مثل:

أكل الولد الطعام

الولد أكل الطعام

الطعام أكله الولد

2.ألتبديلي أو العكس.مثل:

اشتريت من محمد آلة كاتبة : باع لي محمد آلة كاتبة

3. الاندماج المعجمي:مثال ذلك التعبير عن التجمعwith Cement Covered بكلمة واحدة هي Cemented.

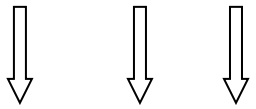
و.الترجمة(Translation) وذلك حين يتطابق التعبيران أو الجملتان في اللغتين.

ز.التفسير(Interpretation):يكون (س) تفسيراً ل(ص) إذا كان (س) ترجمة ل(ص) , وكانت التعبيرات المكونة ل(س) اقرب إلى الفهم من تلك الموجودة في (ص) وعلى هذا فكل تفسير ترجمة, ولا عكس.

أما بالنسبة إلى نوع الترادف الذي تمت معالجته في البحث فلقد أرتئي إلى تسميته (الترادف النسبي بين الجمل)وذلك من خلال معالجة الترادف بين الفعل والاسم في جملة المصدر للحصول على جملة الهدف(الجملة التي لها معنى يشابه نسبياً معنى جملة المصدر).

مثال:

أمتحن المدرس الطلاب



أختبر المعلم التلاميذ

التعبيرات المتلازمة المقيدة Restricted collocations(غازي، 2007)

وهو تلازم مفردتين أو أكثر لتستعمل في سياق مطرد،وليس في استعمال مسكوك،بحيث تتبع البناء البنيوي وتتقيد بتأليفها الاستبدالي نحوا ودلالة وتداولا نحو :

حرب ضارة : جريمة نكراء

أحرز تقدماً : خسارات جسيمة

إن المتلازمات ضرورة للتعلم والتلقين والاستيعاب والهدف منها هو الاستعمال الجيد والصائب لهذه التأليفات،التي تصبح جزءا من الأسلوب المعيار لإنتاج لغة سليمة .

تتميز اللغة العربية بما يسمى بالمشترك والترادف اللفظيين:

مفردة + مشترك لفظي

مفردة + مترادف

أنواع أصناف المتلازمات المقيدة :

1 .فعل + اسم

2.فعل + اسم +فضلة

3.اسم + صفة

وهناك تصنيفات صورية أساسية:

1.فعل + اسم

2. صفة + اسم

3.اسم + فعل

4.وحدة مجمعة مع الاسم

5.ظرف + صفة

6 .فعل + ظرف

7 . فعل + اسم/ فضلة

هذه التصنيفات معتمدة أساساً على مبادئ نحوية ومع ذلك يمكن التعرف على التعبيرات المتلازمة انطلاقا من مستوياتها الدلالية وتجمعاتها في النصوص,وفي أحيان أخرى يتم التعرف عليها انطلاقا من مستويات حدسية تجعل المتلقي يحيل إلى الألفاظ التي تتلاءم دلالتها في سياق محدد بعناصر وأجزاء تتجاذب بعضها مع بعض,كما تحمل المعنى نفسه المراد تبليغه إلى المتلقي.

الهيكل التمثيلي للمعجم

تعتمد كثير من المعاجم في اللغة العربية على استخدام الجذور كأساس وقسم يعتمد على صيغة في الكلمة,وأخر يستخدم الجذوع فضلا عن الصيغ الصرفية لها,في هذا البحث استخدمنا الجذر كمداخل معجمية إذ تم بناء معجم لغوي دلالي حاسوبي خاص باللغة العربية باستعمال برامجيات (Microsoft access),يعتمد على العلاقات الدلالية في تعريف المفردات اللغوية للحصول على قاعدة بيانات مفيدة وغنية بالمعلومات تساعد المتعلم على فهم الجملة العربية. تتكون قاعدة البيانات للمعجم الحاسوبي من الجداول الآتية:

1.جدول الأسماء

2.جدول الأفعال

3.جدول الحروف

4.جدول الصفات

5.جدول خاص بالمتلازمات اللفظية

1.جدول الأسماء:يحتوي جدول الأسماء على الحقول التالية:

-اسم المدخل:يمثل الكلمة الرئيسة للمفردة اللغوية(الاسم).

-المرادف: الاسم المرادف للكلمة الرئيسة.

المعنى الصرفي للاسم

-السمات الدلالية:قائمة بالسمات الدلالية جميعها التي تمثل(الجنس,مفرد أو مثنى أو جمع ,...)

2.جدول الأفعال:يتضمن مدخل الفعل على وصف للإطار الدلالي للفعل مع أصناف الفعل الدلالية ,يمثل الإطار الدلالي المعلومات التالية:

-اسم المدخل(الفعل):يمثل جذر الكلمة.

-المرادف: الفعل المرادف للفعل الرئيس.

-المعنى الصرفي للفعل

-السمات الدلالية للفعل

3.جدول الحروف

يتضمن مدخل الحرف على:

-اسم المدخل:يمثل الحرف نفسه.

-نوعه:أي هل الحرف يدل على زمان أو مكان.

مثال:

.[(مكان),على]Art

4.جدول الصفات

يحتوي مدخل الصفات على:

-اسم المدخل:يمثل الصفة نفسها.

-المرادف: ألكلمه المرادفة للكلمة الرئيسة

5.جدول المتلازمات اللفظية المقيدة

يحتوي الجدول على اللفظ واللفظ المرادف له .

عمل النظام المقترح

عملية المعالجة الحاسوبية تتكون من:

1. تحليل جملة المصدر(الجملة المدخلة) : والتي تتضمن مجموعة أطوار رئيسه تبدأ بطور البحث المعجمي,ثم طور التحليل الصرفي,طور التحليل ألقواعدي,وأخيرا طور التحليل الدلالي,و في حالة فشل أي طور من هذه الأطوار فسيتم إنهاء عملية التحليل.

– طور تحليل المفردات(المميز):

الغرض من هذه المرحلة هو تقطيع الجملة المدخلة إلى كلمات وفرز كل كلمة لتهيئتها للأطوار اللاحقة تعتمد عملية الفرز بين الكلمات على الفراغ بين الكلمات ووضعها في قائمة.

مثال: التلاميذ يذهبون إلى المدرسة

تقسم الجملة إلى كلمات وتوضع في قائمة بالشكل:

["التلاميذ","يذهبون","إلى","المدرسة"]

– طور البحث المعجمي:

بعد انتهاء طور تحليل المفردات تبدأ عملية البحث عن كل كلمة مكونة للقائمة في الطور السابق ,نتيجة هذه العملية هي الحقائق الممثلة للكلمات مع مواصفاتها القوا عدية والدلالية,و في حالة كون المدخل المعجمي فارغا يتم استدعاء طور التحليل الصرفي لاحتمال كون الكلمة المدخلة تحتوي على لواصق صرفية.وأما في حالة عدم القدرة للوصول إلى أحد المداخل المكونة للجملة تتم عملية الانتقال للبحث في القاموس الخاص بالمتلازمات المقيدة وفي حالة عدم وجود الكلمة في هذا الجدول تتم عملية إيقاف التحليل.

– طور التحليل الصرفي: عند انتهاء طور البحث المعجمي, إذ يتم تحليل كل كلمة في الجملة بغض النظر عن موقعها في الجملة لإزالة اللواصق(الحروف الزائدة) التي من الممكن أن تلتصق بالكلمة والتي لكل منها تأثير على معنى الجملة ومكوناتها.

مثال: الإدخال: التلاميذ يذهبون إلى المدرسة

الإخراج : التلاميذ (ي+ ذهب + ون) + إل +إلى+ مدرسة

– طور التحليل ألقواعدي:

في هذا الطور تتم عملية فحص الجملة المدخلة قوا عديا,من خلال اختيار نمط للجملة المدخلة من الأنماط القوا عدية المخزونة فإذا وجد نمط الجملة المدخلة,فالجملة صحيحة قوا عديا,وبخلافه تكون الجملة غير صحيحة قواعد يا,وسيتم إصدار رسالة خطأ قواعدي وإنهاء عملية التحليل.

الإدخال : التلاميذ يذهبون إلى المدرسة

الإخراج : التلاميذ اسم

يذهبون فعل

إلى حرف جر

أل أل التعريف

مدرسة أسم

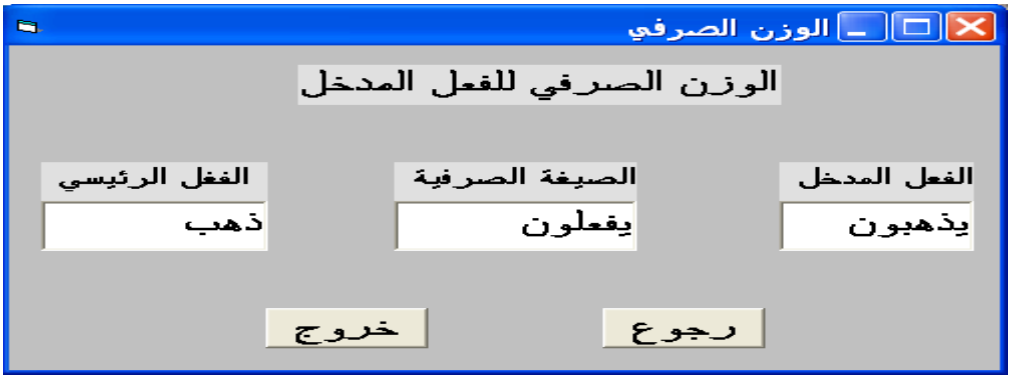
- طور التحليل الدلالي:

هدف هذا الطور تحديد صحة الجملة دلاليا,باستعمال الإطار الدلالي للفعل ومقارنته مع سمات الفاعل والمفعول به الموجودة في الجملة, ومن خلال هذه السمات يتم تحديد صحة الجملة بحيث تكون الجملة مقبولة دلالياً أو لا(عملية فحص الجملة دلاليا).

مثال: التلاميذ يذهبون إلى المدرسة جملة صحيحة دلاليا

الفعل درس يحتاج إلى فاعل حي إنساني ,ومفرد,....

2. مرحلة اكتشاف الوزن الصرفي للفعل المدخل: تتم عملية مطابقة بين الفعل والصيغة الصرفية الخاصة به سواء أكان الفعل مجرد أم مزيد مع الصيغ الصرفية الموجودة في المعجم ا لغرض من هذه المرحلة إتمام عملية إعطاء فهم كامل لمعنى الجملة كما في الشكل رقم(1).



شكل رقم (1)

3.مرحلة أيجاد الاشتقاقات الصرفية للفعل المدخل(ناصر، 2008):

يتم اشتقاق الفعل إلى الصيغ الصرفية المطلوب اشتقاقها, إضافة إلى السمات الصرفية للفعل كونه (مذكر أم مؤنث , مفرد أم مثنى أم جمع ,الضمير).تم معالجة هذه السمات من خلال :

1.إذا كان الفعل يحتوي في نهايته (ت) التأنيث فقط فان الفعل مؤنث وغير ذلك يكون مذكر.

2. إذا كان الفعل يحتوي في نهايته على (ا,ن) فانه مثنى في حالة المضارع أو (ا) في حالة الماضي ,وإذا كان الفعل يحتوي على (و,ن) فانه جمع مذكر سالم في حالة المضارع ,(و,ا) في حالة الماضي وإذا كان الفعل يحتوي على نون النسوة فأنه جمع مؤنث سالم عدى تلك الحالات فأنه مفرد .

3. إذا كان الفعل مثنى فان ضميره (هما) وإذا كان الفعل جمع مذكر سالم فان ضميره (هم) وإذا كان الفعل جمع مؤنث سالم فان ضميره (هن) وإذا كان الفعل مفرد مذكر فان ضميره (هو) وإذا كان الفعل مفرد مؤنث فان ضميره (هي) .

شكل رقم (2)

4.مرحلة أيجاد الفعل المرادف للفعل المدخل:-

يأخذ الفعل المدخل من المرحلة الأولى ويتم البحث عنه في المعجم المكون في (Microsoft access)وإيجاد الفعل المرادف له والمعنى الصرفي للفعل كما في الشكل رقم(3).

شكل رقم (3)

5.مرحلة أيجاد الجملة المرادفة للجملة المدخلة:

يتم الاستفادة من نتائج البحث المعجمي التي حصلت في مرحلة التحليل والتي من خلالها أمكن الوصول إلى المداخل المعجمية لمكونات جملة المصدر ففي الوقت نفسه يتم الوصول إلى مقابلاتها في جملة الهدف لوجودها في المداخل نفسها وبذلك يمكن تقليص زمن البحث إلى النصف تقريباً.وكذلك يتم الوصول إلى المواصفات القوا عدية والدلالية للكلمات في لغة الهدف,يستفاد من هذه المواصفات في بناء جملة الهدف(الجملة المرادفة لجملة المصدر) من خلال التأكد من الصيغ القوا عدية الوصفية المخزونة لجملة الهدف لبناء جملة صحيحة قوا عديا ودلاليا من خلال أجراء معالجات بسيطة تتضمن تنسيق الكلمات في قوالب قوا عدية وفق ضوابط تناسب تراكيب لغة الهدف لإنتاج الجملة بشكلها النهائي.

ت	الجملة المدخلة(جملة المصدر)	الجملة المرادفة(جملة الهدف)
1	كتبت سعاد الدرس بجد	خطت البنت الواجب بتفوق
2	أكل احمد الفاكهة	مضغ الولد الثمار
3	سمعت الولد يبكي	علمت إن الصبي يصرخ
4	أكملت هند الواجب	أتمت البنت درسها
5	اختبر المدرس الطلاب	امتحن المعلم التلاميذ
6	احمد طالب مجد	الولد تلميذ متفوق
7	ذهب سعد إلى حتفه	مضى الرجل إلى موته

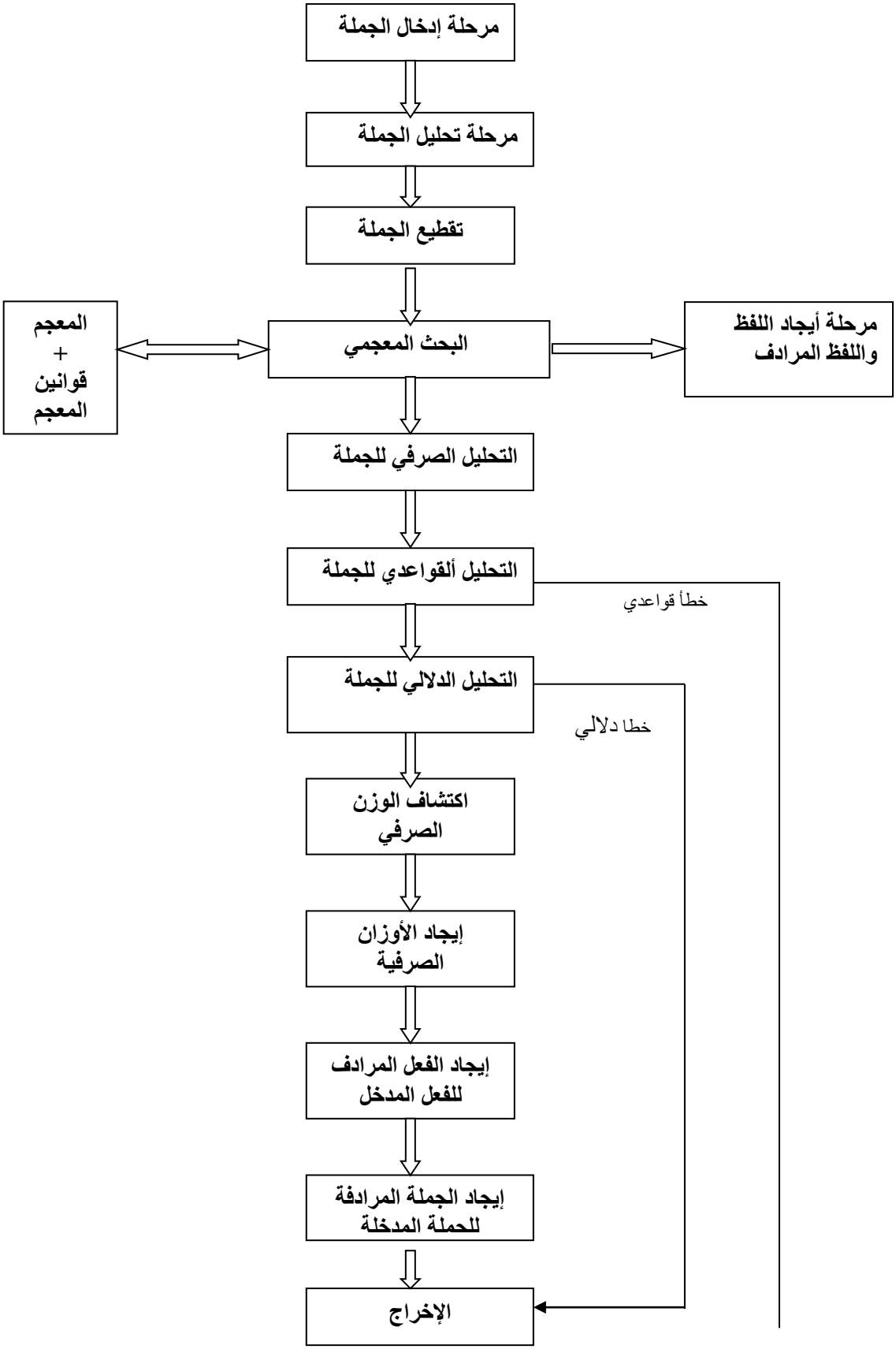
شكل رقم(4) يوضح الجملة المدخلة(جملة المصدر) والجملة المرادفة لها(جملة الهدف) في اللغة العربية

5.مرحلة أيجاد المرادفات اللفظية للمتلازمات المقيدة:

تم إضافة هذه المرحلة إلى البحث بحيث يكون عملها منفصلاً عن المراحل السابقة,إذ يتم إيجاد المرادفات اللفظية للمتلازمات المقيدة,من خلال المعجم الخاص بها كما موضح في الملحق رقم(1).

عند إدخال اللفظ المقيد يتم البحث عنه في المعجم الخاص به وإيجاد اللفظ المرادف له ,أما في حالة فشل البحث أي عدم وجود اللفظ المقيد سيتم أظهار رسالة خطأ وإنهاء البحث.كما في الشكل رقم(5)

شكل رقم (5)



شكل رقم(6) يمثل المخطط الانسيابي لسير عمليات النظام

الاستنتاج

1. التعبيرات المتلازمة هي المكون المفتاح في معجم كل لغة طبيعية، حيث استعمالها في المعجم السياقي الثنائي للغة ضرورة في تعليم اللغة العربية وفي الترجمة من والى اللغة العربية.
2. يمكن تقسيم التوليد إلى نوعين :توليد داخلي ,لتوليد المداخل المعجمية من جذورها وفق قواعد مضبوطة والنوع الأخر توليد خارجي(أو ظاهري) من خلال توليد الجمل كنتيجة للتوليد الداخلي.
3. محاولة أيجاد طريقة مثلى لبناء المعاجم العربية,يمكنها معالجة اللبس ألقواعدي من خلال بناء آلية للمفردات اللغوية العربية البسيطة منها والمركبة,على غرار ما هو موجود في اللغات الأجنبية.
4. ضرورة انتماء كل وحدة معجمية إلى حقل دلالي واحد فقط يؤدي إلى جمع عدة مفاهيم تمثل المعاني المتعددة للوحدة المعجمية,وهذا يساعد على تقليص الوقت المستخدم في عملية البحث.
5. أن قاعدة البيانات يمكنها استيعاب مدخل معجمي جديد,مما يعني أنها تعكس الكفاية اللغوية للمستخدم.
6. أن هذا النظام برنامجا تعليميا يمكن استخدامه مستقبلا في مجال ترجمة الجمل البسيطة,بواسطة عمليات الترادف يمكن اشتقاق كل قراءات الجمل التي يمكن الحصول عليها من الجمل المصدرية المدخلة وكيف أن تتولد نفس دلالة الهدف من هذه القراءات أي اعتبار هذه العمليات وسيلة للترجمة من والى اللغة العربية.
7. يمكن تطوير النظام ليشمل عدد اكبر من الجمل البسيطة والمركبة بإضافة القوانين التوليدية التحليلية لفلمور للآخذ بنظر الاعتبار كل الاحتمالات الممكنة للصيغة البنيوية.
8. تصميم بيئة أكثر جاذبية وتحاوريه مع المستعمل تلبي حاجاته وتكون مقروءة ومألوفة حتى لغير المختصين بالحاسوب ,كلغة(visual basic).

المصادر

جومسكي، نعوم, (1987)، "البنى النحوية",ترجمة:د.يؤيل يوسف عزيز,دار الشؤون الثقافية العامة,بغداد.

خلف، زينب علي (2001) ,"تطبيق محو سب لمعالجة الجمل العربية بتفسير علاقات الترادف"رسالة ماجستير,جامعة البصرة,كلية العلوم,علوم الحاسبات.

ضمد، يسرى مالك, (2006)، "تصميم وبناء معجم دلالي حاسوبي لدعم المعالجة الآلية للغة العربية",رسالة دكتوراه,جامعة البصرة,كلية العلوم,علوم الحاسبات,2006.

عبد الباقي، .سلمى , (2001) "معجم عربي معرفي لفهم اللغة العربية",رسالة دكتوراه,جامعة البصرة,كلية العلوم,علوم الحاسبات.

عبد الستار، سعد ,نعمان محمد نعمان, (1992) "فهم الحاسوب للجمل العربية",وقائع بحوث المؤتمر الوطني الأول للحاسبات في العراق,بغداد.

علي، نبيل, (1988) ، "اللغة العربية والحاسوب",تعريب للطباعة والنشر,القاهرة,مصر.

عمر، احمد مختار , (1982) ، "علم الدلالة",مكتبة دار العروبة للنشر والتوزيع,الكويت.

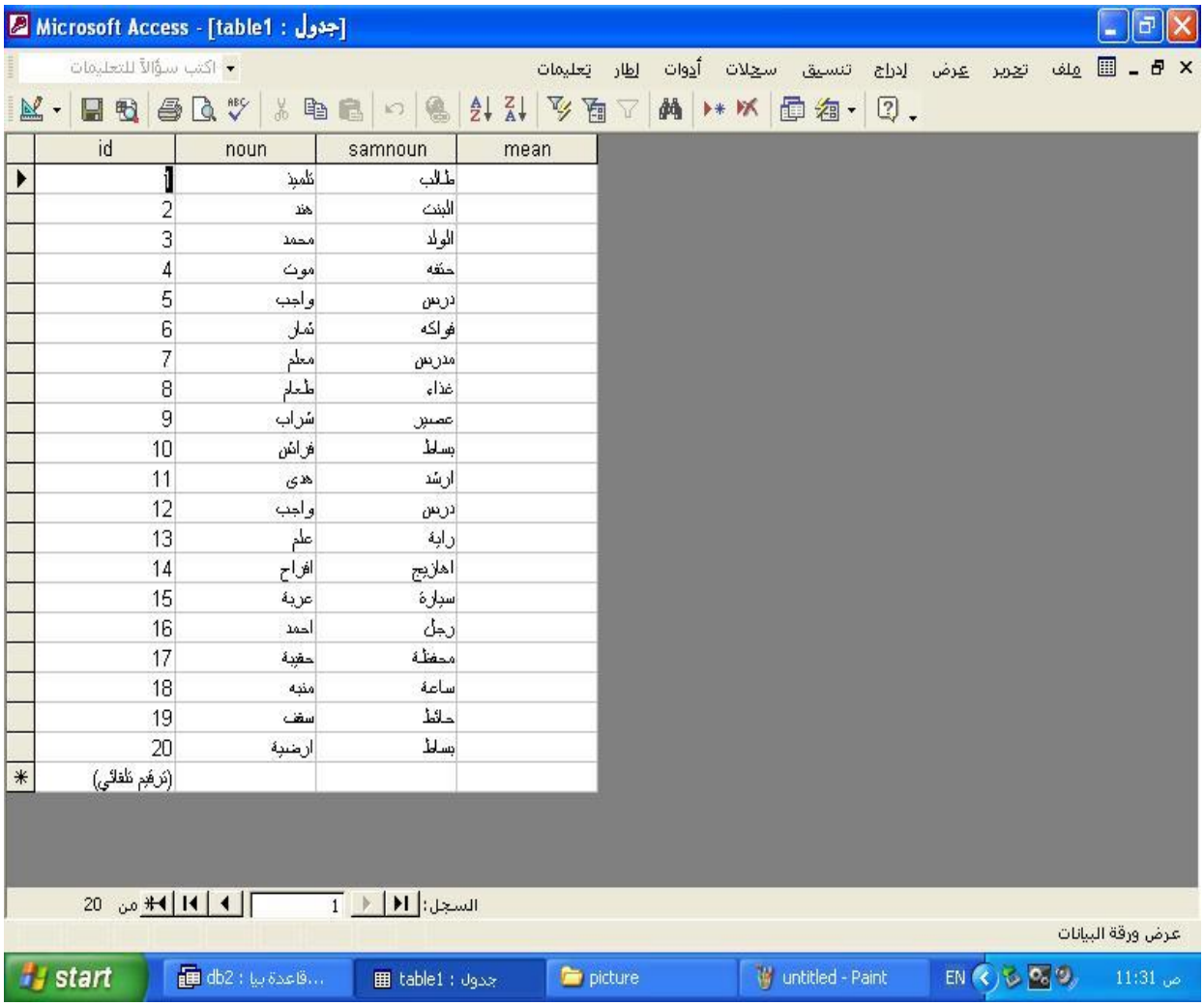
غازي، عز الدين، (2007) ، "المتلازمات في اللغة العربية ومعالجتها في القواميس الثنائية اللغة",الحوار المتمدن ,العدد: 2039.

المبارك، عبد الحسين, (1986) ، "فقه اللغة",طبع على نفقة جامعة البصرة.

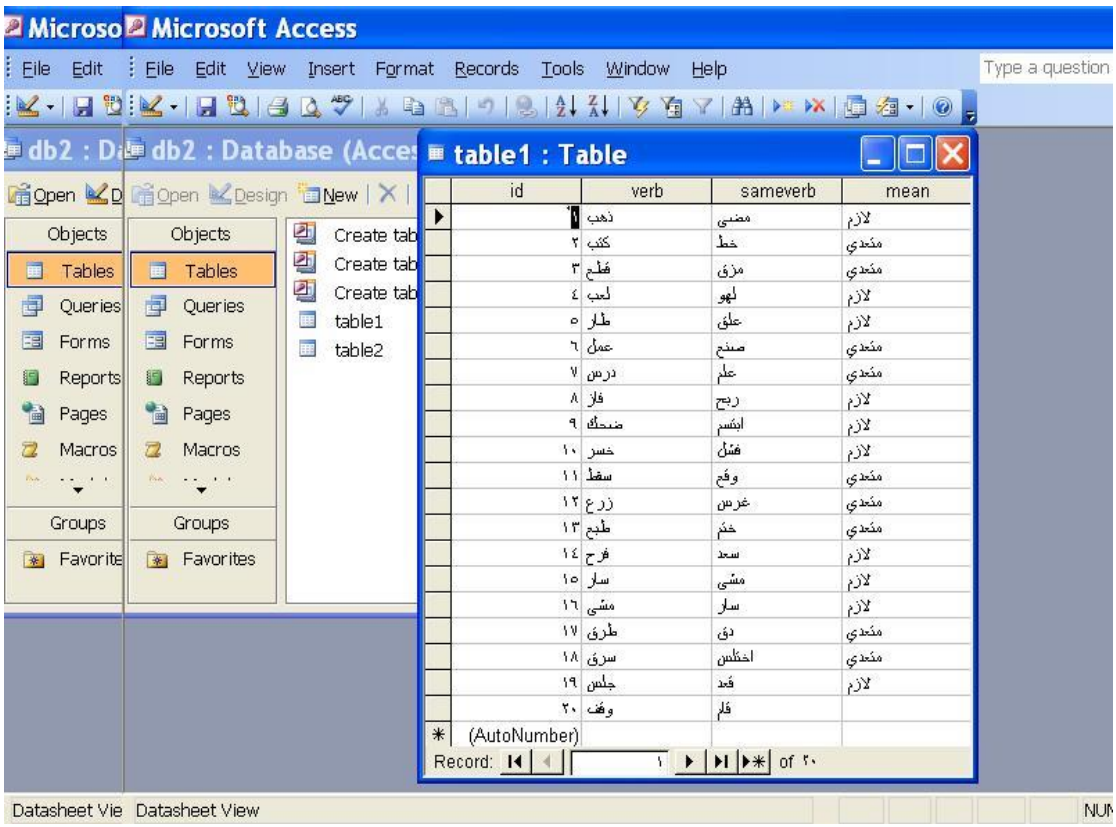
مهديوي، عمر , (2006) "خوارزميات توليد الأسماء في اللغة العربية",المغرب.

ناصـر، ميسـاء عبـد الكـريم, (2008) ، "نظـام صـرفي نحـوي دلالـي لـبعض المفـاهيم الأساسـية فـي اللغـة العربية",مجلة بابل,المجلد(الرابع عشر) العدد الثالث.

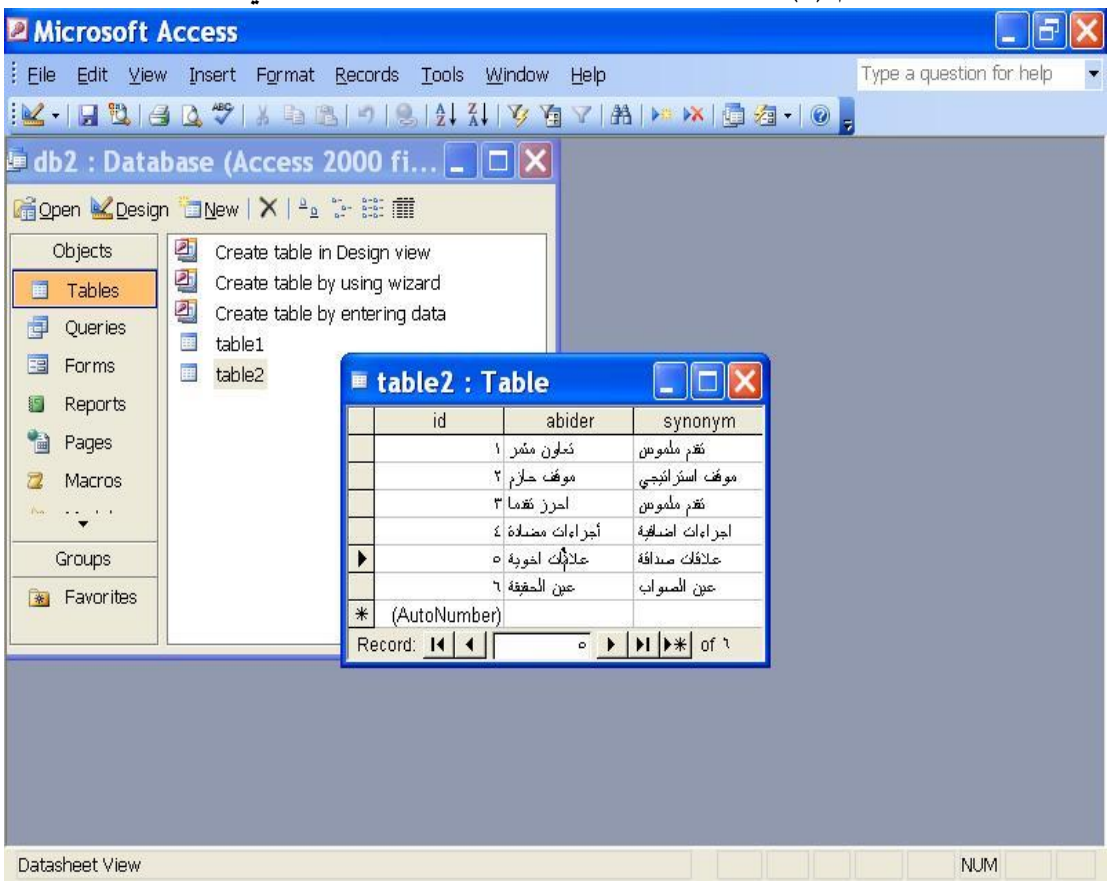
Thompson,C.,A.&Mooney,R.,J.," (2003) Acquiring word-meaning mappings for natural language interfaces", journal of artificial intelligence research 18.



جدول رقم (1) يمثل جدول الاسم والاسم المرادف



جدول رقم (2) يمثل الفعل والفعل المرادف فضلا عن المعنى الصرفي للفعل



جدول رقم (3) يمثل اللفظ واللفظ المرادف