المعالجة الآلية للدلالة في اللغة العربية بين جهود اللسانيين وأفاق الحاسوبيين

Automatic Processing Of The Arabic Language Between The Efforts Of The Lassanians And The Horizons Of Computers

تاريخ الإرسال: 2018-01-10 تاريخ القبول: 2018-01-12

الدكتورة: راضية بن عريبة

جامعة حسيبة بن بوعلي –شلف (الجزائر)

ملخص :

تحتاج المعالجة الآلية للدلالة العربية مقداراً كبيراً من المعلومات عن مختلف جوانب اللّغة، وأفضل ترتيب لهذه المعلومات هو قواعد البيانات التي تشتمل القواعد المتعلّقة بنسق الكتابة والصّرف، وما يتعلّق به من فصل نواة الكلمة من لواحقها السابقة واللاحقة، وربط النواة على الأوزان الصّرفية المعروفة، يلي ذلك قواعد بيانات النّحو التي تحتاج إلى تقسيم اللّواصق على مكوّناتها.

الكلمات المفتاحية: المعالجة الآلية، اللسانيات الحاسوبية، اللغة، اللغة العربية، اللسانيات.

Abstract:

Automated processing of Arabic meaning requires a large amount of information on various aspects of the language. The best order for this information is the databases that include the rules related to the format and the syntax, the related word of the nucleus of the word from its previous and subsequent suffixes, Databases are grammar that you need to break up the labels on their components.

Keywords: Automated Processing, Computer Linguistics, Language, Arabic Language, Linguistics.

قواعد بيانات الدّلالة:

قاعدة البيانات ما هي إلا أفكارٌ عامّة تحتاج إلى مزيد من العناية والتوسيع، فموضوع تقنين المعاني والدّلالة على جانب كبير من الصّعوبة بالنسبة للّغات الحديثة، فكيف باللّغة العربيّة[1]" الواسعة المعاني والدّلالة،وفي قواعد بيانات الدلالة نذكر:

1. قاعدة بيانات الجذور:

فالجذور من ناحية المعاني مترابطة مع بعضها فمثلا جذور مثل: جيء، أتى، قَدِم، وصل، دبر، ولي، وكذلك: ولج، ودخل، قَبِل...إلخ، هذه المجموعات مترابطة حيث إنها تتعلق مع غيرها من المجموعات مترابطة مع بعضها، فهي تتعلق مع غيرها من المجموعات بحركة الإنسان ذهاباً وإياباً أو خروجاً ودخولاً [2].

إنّ الجذور وحدها تجتمع في غالبيتها أصولاً لمعاني عديدة، ولذلك فإنّ الاستناد إليها يعطي وجهة عامَّة غير دقيقة للدلالة[3].

2. قاعدة تمثيل بالشبكات الدَّلالية:

الشبكات الدّلالية هي شبكات هندسية تمثل بها المفاهيم ويربط بينها بخطوط الرّبط، تمثل أنواع العلاقات الموجودة التي تربط بين هذه العناصر، وشاع استخدام هذه الشبكات مع بداية السّتينيات في مجالات مختلفة أهمها: علم النفس، وعلم المكتبات، ثم تطورت إلى تمثيل النّصوص الصغيرة في جامعة كمبريدج[4]، ومع ظهور علم النفس المعرفي وهندسة المعرفة تطورت هذه الشبكات وذلك محاولة من العلماء لتقليد العقل البشري، كما تحتاج المعالجة الالية للدلالة إلى معالجة آلية للجملة بصفتها تمس معنى الجملة وتحدِّد دلالته أثناء السياق.

المعالجة الآلية للجملة العربية:

المعالجة الآلية للجملة العربية موضوع متعدد الجوانب، دو تفاصيل دقيقة، ويصب فيه نتائج كثير من النظريات وأساليب الذكاء الاصطناعي وقد نشأت الحاجة لمعالجة الجمل مع ظهور الترجمة الآلية.

تحتاج المعالجة الآلية للجملة العربية إلى دراسات إحصائية لأنواع الجمل والتعابير الاصطلاحية أو مصاحبة الصفات للأسماء، والتقديم والتأخير ...الخ.

1. خطوات المعالجة الآلية للجملة في الحاسوب:

اقترح نبيل على الخطوات التالية لمعالجة الجملة العربية في الحاسوب:[5]

الخطوة الأولى: تحليل صرفي لكلمات الجملة، لاستخلاص الجذر والصيغة الصرفية والحالة التصريفية والإعرابية، والأدوات والضمائر المتصلة بها.

الخطوة الثانية: تصنيف معجمي لعناصر الجملة.

الخطوة الثالثة: تطبيق قواعد استبعد البدائل المستحيلة.

الخطوة الرابعة: تطبيق قواعد النَّحْو لتكوين أشباه الجمل وما يعلوها من مكونات نحويّة، وذلك للحصول على بدائل محتملة لبنية الجملة، ويساعد أسلوب المخططات الشجرية على تقليص هذه البدائل.

الخطوة الخامسة: تطبيق قواعد الضبط النّحوي والإعرابي، وذلك للتأكد من صحة المكونات النحوية.

الخطوة السادسة: ربط الضمائر بمراجعها، وذلك بمطابقة السمات النّحويّة الصرفية للضمائر مع مقابلاتها.

الخطوة السابعة: تطبيق القيود الدّلالية، للحصول على بنية الجملة الصحيحة وينتهي بذلك تحليل الجملة ليتم على إثرها تمثيل بنية الجملة في صورة هيكليّة متعددة المستويات.

الخطوة الثامنة: ضبط كلمات الجملة (وضع الحركات على حروف كلماتها) ويتضمّن ذلك شكل ساق الكلمة، وعلامات الضبط الإعرابي، والتعديلات اللازمة على نهاية الكلمة نتيجة اتصالها بالضمائر، أو نتيجة التقاء السّاكنين.

2. مشاكل معالجة الجملة العربية آلياَ:

1. نشير إلى بعض المشاكل التي تعترض سبيل معالجة الجملة العربية آلياً، ومنها[6]

2. غياب الشكل (الحركات) فخاو معظم النصوص العربية نم علامات الشّكل وعدم استخدام علامات الترقيم على نحو صحيح، وإقفال كتابة الهمزة على الألف بعض الأحيان، وعدم التعريف بين الياء والألف المقصورة، كل ذلك يعدّ مشكلة تواجه الحاسوب في معالجته للجملة.
3. تعدّد المعاني المعجمية للكلمة الواحدة: فأكثر الكلمات لها أكثر من معنى معجمي وعلى الحاسوب أن يختار من هذه المعاني ما يناسب سياق الجملة أو النص.
4. اللّبس الناجم عن التراكيب
5. اللبس الصرفي: فمن الكلمات ما يشترك في:
   أ- الاسمية والفعلية: نحو كَاتِبُوه.
   ب- اسم الفاعل والمفعول: نحو: المختار.
   ت- الفعلية والحرفية: نحو: لعلَّ ، لما ....الخ
6. اللّبس النحوي: فمن الأفعال ما يكون مسنداً إلى ضميريْن نحو: نَحْوَ: قاضٍ، كِتَابِي.

والحاسوب في هذا كلّه قد لا يكتشف اللّبس إلّا في مرحلة متأخرة من المعالجة[7]

تجتمع الجذور وحدها في غالبيتها أصولاً لمعاني عديدة، ولذلك فإنّ الاستناد إليها يعطي وجهة عامّة غير دقيقة للدّلالة[8]

تمثيل المعرفة بالشبكات الدّلالية:

الشبكات الدّلالية هي شبكات هندسية تمثّل بها المفاهيم، ويربط بينها بخطوط الربط، تمثّل أنواع العلاقات الموجودة التي تربط بين هذه العناصر، وشاع استخدام هذه الشّبكات مع بداية لستّينات في مجالات مختلفة أهمها: علم النفس، علم المكتبات، ثمّ تطوّرت إلى تمثيل النصوص الصّغيرة في جامعة كمبريدج [9]

ومع ظهور علم النفس المعرفي، وهندسة المعرفة تطوّرت هذه الشّبكات، وذلك محاولة من العلماء لتقليد العقل البشري.

كما تحتاج المعالجة الآلية للدّلالة إلى معالجة آلية للجملة بصفتها تمسّ معنى الجملة، تحدّد دلالته أثناء السياق.

المعالجة الآلية للدلالة:

تحتاج المعالجة الآلية للدّلالة العربية مقداراً كبيراً من المعلومات عن مختلف جوانب اللّغة، وأفضل ترتيب لهذه المعلومات هو قواعد البيانات التي تشتمل القواعد المتعلّقة بنسق الكتابة والصّرف، ما يتعلّق به من فصل النواة الكلمة من لواحقها السابقة واللاحقة. وربط النواة على الأوزان الصّرفية المعروفة. يلي ذلك قواعد بيانات النّحو التي تحتاج إلى تقسيم النواصف إلى مكوّناتها.

قواعد بيانات الدّلالة:

قاعدة البيانات ما هي إلا أفكار عامّة تحتاج إلى مزيد من العناية والتوسيع، فموضوع تقنين المعاني والدّلالة على جانب كبير من الصّعوبة بالنسبة للغات الحديثة، فكيف باللّغة العربيّة [10] الواسعة المعاني والدّلالة، وقواعد بيانات الدّلالة نذكر:

قاعدة بيانات لجذور: فالجذور من ناحية المعاني مترابطة مع بعضها، فمثلاً: جيء، أتى، قَدِمَ، حَضَر، وَصَلَ، دَبَرَ، ولّى، وكذلك: وَلَجَ، دَخَلَ، قَبِلَ، ...إلخ، هذه المجموعات مترابطة حيث إنّها تتعلّق مع غيرها من المجموعات مترابطة مع بعضها، فهي تتعلّق مع غيرها من المجموعات بحركة اللّسان ذهاباً وإياباً؛ خروجاً أو دخولاً [11]

المعالجة الآلية للجملة العربية:

المعالجة الآلية للجملة العربية موضوعٌ متعدّدُ لجوانب،ذو تفاصيلَ دقيقة، تصبُّ فيه نتائجُ كثيرٍ من النظريّاتِ وأساليبِ الذّكاءِ الاصطناعيِّ، فقد نشأت الحاجة لمعالجة الجمل مع ظهور الترجمة الآلية.

تحتاج المعالجة الآلية للجملة العربية إلى دراساتٍ إحصائيةٍ لأنواعِ الجملِ والتعابيرِ الاصطلاحية، ومصاحبة الصّفات للأسماءِ ، ولتقديمِ والتأخيرِ ... إلخ

1. **خطوات المعالجة الآلية للجملة في الحاسوب:**

اقترح نبيل على الخطوات التالية لمعالجة الجملة العربية في الحاسوب: [12]

الخطوة الأولى: تحليل صرفي لكلمات الجملة لاستخلاص الجذر والصّيغة الصّرفية والحالة التصريفية والإعرابية، والأدوات والضّمائر المتّصلة بها.

الخطوة الثانية: تصنيف معجمي لعناصر الجملة.

الخطوة الثالثة: تطبيق قواعد استبعاد البدائل المستحيلة.

الخطوة الرابعة: تطبيق قواعد النّحو لتكوين أشباه الجمل وما يعلوها من مكوّناتٍ نحويّةٍ، وذلك للحصول على بدائلَ محتملة لبنية الجملة، ويساعد أسلوب المخطّطات الشجريّة على تقليص هذه البدائل.

الخطوة الخامسة: تطبيق قواعد الضّبط النّحوي والإعرابي، وذلك للتأكّد من صحّة المكوّنات النحوية.

الخطوة السادسة: ربط الضّمائر بمراجعتها، وذلك بمطابقة السّمات النّحويّة الصّرفية للضّمائر مع مقابلاتها.

الخطوة السّابعة: تطبيق القيود الدّلالية للحصول على بنية الجملة الصّحيحة، وينتهي بذلك تحليل الجملة ليتمّ على إثرها تمثيل بنية الجملة في صورة هيكليّة متعدّدة المستويات.

الخطوة الثّامنة: ضبط كلمات الجملة (وضع الحركات على حروف كلماتها)، ويتضمّن ذلك شكل ساق الكلمة، وعلامات الضّبط الإعرابي، والتعديلات اللازمة على نهاية الكلمة نتيجة اتذصاهل بالضّمائر، أو نتيجة التقاء السّاكنين.

مشاكل معالجة الجملة العربية آلياً:

نشير إلى بعض المشاكل التي تعترض سبيل معالجة الجملة العربيّة آلياً –كخاتمة للبحث – ومنها: [13]

1. غياب الشّكل(الحركات) فخلوُّ معظمِ النّصوصِ العربيّةِ من علاماتِ الشّكلِ وعدمِ استخدامِ علاماتِ التّرقيم على نحوٍ صحيحٍ، وإقفالِ كتابةِ الهمزةِ على الألفِ بعضَ الأحيانِ، وعدمُ التّعريفِ بين الياءِ والألفِ المقصورةِ، كلُّ ذلك يعدُّ مشكلةً تواجه الحاسوبَ في معالجتهِ الجملةَ.
2. تعدّدُ المعاني المعجميّة للكلمةِ الواحدةِ: فأكثر الكلماتِ لها أكثر من معنىً معجميٍّ وعلى الحاسوبِ أن يختارَ من هذه المعاني ما يناسب سياقَ الجملةِ أو النصِّ.
3. اللبسُ الناجمُ عن تطابقِ الصّفةِ والموصوفِ: مثلاً "الأساتذةُ والطلابُ منتخبون" هناك لبسٌ في كلمة "منتخبون" هل هي صفة للأساتذة، أم الطّلاب، أم كليهما.
4. اللّبس النّاجم عن التراكيب.
5. اللّبس الصّرفيُّ: فمن الكلمات ما يشترك في:

أ. الاسمية والفعلية: نحو كاتِبُوه.

ب. اسم الفاعلِ والمفعولِ: نحو: المختار.

ت. في صيغة الماضي والأمر: اصطادوا.

ث. الفعلية والحرفية: نحو: لعلَّ، لما، ... إلخ

6. اللّبس النحويُّ: فمن الأفعال ما يكون مسنداً إلى ضميرين نحو:
7. تَحْمِلُ وَ يَغْزَوْن، ومنها ما يحتمل حالتين إعرابيّتين أو أكثر: نحو: قَاضٍ، كِتابِيٌّ.

والحاسوب في هذا كلّهِ لا يكتشف اللّبس إلا في مرحلة متأخّرةٍ من المعالجة [14]

التهميش :

[1] ينظر: نحو معالجة الدّلالة في اللّغة العربية عبر قواعد البيانات، دراسة أولية لنص القرآن الكريم، محمد زكي الخضر، الجامعة الأردنية، المؤتمر الوطني 17 للحاسب الالي، جامعة الملك عبد العزيز، المدينة المنورة، أبريل 2004، ص:415.

[2] ينظر: المرجع نفسه، ص:429.

[3] ينظر: إحصاء الأفعال العربية في المعجم الحاسوبي، مروان البواب وغيره، مكتبة لبنان، 1996، ص32.

[4] L'intelligence Artificielle et la longue, Sabah, Gérard, Edition hermès, paris, 1988, p : 152

[5] ينظر: اللغة العربية والحاسوب، نبيل علي، دار تعريب، الكويت، دط، 1988، ص: 37.

[6] ينظر: استخدام اللغة العربية في المعلوماتية، المنظمة العالمية للترجمة، ص: 54.

[7] ينظر: استخدام اللغة العربية في المعلوماتية المنظمة العالمية للترجمة، ، ص55.

[8] ينظر: إحصاء الأفعال العربية في المعجم الحاسوبي، مروان البواب وغيره، مكتبة لبنان، 1996، ص: 32.

[9] L'intelligence artificielle et la langue, Sabah, Gérard, édition hermès, paris, 1988, p : 152.

[10] ينظر: نحو معالجة الدلالة في للّغة العربية عبر قواعد البيانات- دراسة أوّلية لنص القرآن الكريم، محمد زكي الخضر، الجامعة الأردنية ، المؤتمر الوطني 17 للحاسب الآلي، جامعة الملك عبد العزيز، المدينة المنوّرة، أبريل 2004، ص: 415.

[11] ينظر: المرجع نفسه، ص: 429.

[12] ينظر: اللّغة العربية والحاسوب، نبيل علي، دار تعريب، الكويت، د ط، 1988، ص: 37.
[13] ينظر: استخدام اللّغة العربيّة في المعلوماتية، المنظّمة العالمية للتّرجمة، ص: 54.
[14] ينظر: استخدام اللّغة العربيّة في المعلوماتية، ص: 55.