صورنة اللغة العربية مقاربات ونماذج

Formalization of the Arabic language approaches and models

د. يحيى بن أحمد عبد الله اللتيني

أستاذ اللسانيات الحاسوبية المساعد بقسم اللغة العربية وآدابها بجامعة الملك خالد

البريد الإلكتروني: yaabdullah@kku.edu.sa

رموز البحث

تعارض	$(\neq)$
مكافئ	$(\equiv)$
علاقة انتماء	$\ni$
علاقة جزئية	$\supset$
التسوير مهما يكن/ كل	$\forall$
النفي	$\neg$
الاتحاد	$\cap$
التقاطع	$\cup$
مجموعة فارغة	$\emptyset$
الاستلزام المنطقي	$\leftarrow$
التكافؤ الرياضي	$\leftrightarrow$
التتميم	$\backsim$
النتيجة	$\wedge$
العبارات الرياضية الخاطئة	F
العبارات الرياضية الصائبة	T
التحول إلى	$\vdash$

المستخلص:[١]

يتطرق هذا البحث إلى بيان توظيف بعض المسائل الرياضية في صورنة اللغة العربية التي تعد المرحلة الهامة قبل حوسبتها حتى نصل بها إلى قالب يتماشى ولغة الحاسوب. ويهدف البحث إلى عرض نماذج لتحويل البنى اللغوية إلى بنى رياضية مجردة. لما تتميز به الرياضيات من موضوعية صارمة يمكن أن تلج بها اللغة إلى الحاسوب ويتعامل بها بنجاعة عالية. وقد اكتفيت بثلاثة نماذج رياضية وهي المجموعات الرياضية وعلاقاتها، وحساب التفاضل والتكامل، والجداء الديكارتي (المتجه الزاوي) لحساب متجه الكلمة word vector.

وخلص البحث إلى أنه يمكن الإفادة من المجموعات الرياضية في تمثيل بعض الشروط النحوية كما فيما يجمع جمع مذكر سالم، وبيّن محلل CCG كيفية الكشف عن السلوك المنطقي للجمل والكشف عن التوليد الدلالي وكيفية التعرف على الجمل الأصلية والجمل التي تخرج عن السياق الدلالي. وأخيرا كشف البحث عن أهمية استخدام حساب الكلمات ووزنها في النص وكيف يمكن بذلك الحساب (المتجه الزاوي) تحديد المجال الذي يندرج فيه النص.

الكلمات المفتاحية: صورنة – المجموعات الرياضية – حساب لامدا – متجه الكلمات – المتجه الزاوي.

(١) شكر وتقدير

أتقدم بالشكر والتقدير لجامعة الملك خالد على الدعم المادي والإداري لإنجاز هذا البحث.

Abstract:

This research deals with the explanation of the employment of some mathematical problems in the imagery of the Arabic language, which is an important stage before its computerization, until we reach a template that matches the language of the computer by transforming linguistic structures into abstract mathematical structures. This is due to the fact that mathematics is characterized by its strict objectivity, through which language can be accessed by computer and deal with it with high efficiency. I have mentioned three mathematical models which are the mathematical sets and their collations, the calculus, and the Cartesian product (the angular vector).

The research concluded that the mathematical sets can be used to represent some grammatical conditions, as in what combines the Salem masculine plural, and the CCG Parsing to detect the logical behavior of sentences, detect semantic generation, and how to identify original sentences and sentences that go out of the semantic context. Finally, the research revealed the importance of using counting words and their weight in the text, and how this calculation (Angular velocity) can determine the field in which the text falls.

Kay words: Formalism — Mathematical Sets - Lamda calculus - Word direct – Angular vector.

المقدمة:

عدّ السيميائيون اللغة رموزا، وهذه العبارة فتحت الباب على أن تكون الرياضيات مدخلا مهما لترميز اللغة والتعامل معها من قبل الآلة باعتبارها ذلك صَورنة للغة. وعليه فالبحث يحاول الكشف عن المواضع التي يمكن أن تتداخل فيها اللغة مع الرياضيات باعتبارهما أسسا وسيرورات لها بدايات حدثية تفضي إلى نتائج محددة؛ فكما أن الرياضيات تقوم على القيم في البنى الرياضية فاللغة تتمتع بتلك الصفة بشكل ما في البنى اللغوية. ومن هنا انطلقت فكرة ما يسمى بالأرْيَضة mathematization.

مع ارتفاع الحاجة إلى إدخال اللغة الطبيعية إلى الآلة لا لمعالجتها فحسب بل لتكون سمة من سماتها في هذا القرن بخاصة؛ فالآلة تتفاعل اليوم مع المحيط الخارجي بكل أنواعه ومكوناته؛ كان العمل على صورنة اللغة أمرا ملحًّا لتُقدم للآلة بالشكل الذي يمكن أن تتفاعل معه، وهي السبيل إلى تجريد السيرورات الذهنية التي تعنى باللغة في الدماغ البشري ثم محاكاة الآلة لتلك السيرورات آليًا.

تصور بسيط: أحاول في هذا المثال البسيط شرح فكرة التقريب بين اللغة والرضيات بشكل بدائي.

لو جعلنا للأصوات العربية قيمًا وفق الجدول ١:

الصوت	القيمة	الصوت	القيمة	الصوت	القيمة	الصوت	القيمة	الصوت	القيمة
ء	١	د	٨	ض	١٥	ك	٢٢	ـَ	٢٩
ب	٢	ذ	٩	ط	١٦	ل	٢٣	ـُ	٣٠
ت	٣	ر	١٠	ظ	١٧	م	٢٤	ـِ	٣١
ث	٤	ز	١١	ع	١٨	ن	٢٥	ا	٣٢
ج	٥	س	١٢	غ	١٩	هـ	٢٦	وْ	٣٣
ح	٦	ش	١٣	ف	٢٠	و	٢٧	يْ	٣٤
خ	٧	ص	١٤	ق	٢١	ي	٢٨		

جدول ١: ترقيم الأصوات العربية الصامتة والصائتة.

فإننا إذا طلبنا من الحاسوب المعادلة إيجاد المجهول في المعادلة التالية:

س= ٢٥+٢٩+٦+٢٧

فسيعطينا قيمة س= نَحو... وهكذا.

فنحن بالصورنة ننقل اللغة من إطارها المحسوس إلى إطار مجرد حيث تكون رموزًا مجردة تقنن الخطاب الطبيعي وتضبط تصوراته ونتائجه.

مشكلة البحث:

إن الرصد الدقيق للغة يجعلها أداة للتفكير الصحيح والدقيق، كما يجعلها أداة قانونية لصياغة القوانين والدساتير دون تعارض بين اللوائح. كما أن مَنطَقَة اللغة وربطها بالبرهان الرياضي يجعلها طيعة للأتمتة Automationومن ثّم يستطيع الحاسوب أن يتعامل معها بكل سهولة. ومن هنا يحاول البحث الكشف عن العلاقة بين اللغة باعتبارها أداة للتفكير، والرياضيات باعتبارها أداة للصورنة منضبطة في منطلقاتها وبراهينها ونتائجها، كما يهدف إلى الوصول إلى تصور رياضي **لبعض مسائل** اللغة الطبيعية نستطيع من خلاله صورنة تلك المسائل وتهيئتها للمعالجة الآلية للغات الطبيعية.

أولا المقاربات:

١- الفلسفة والمنطق.

عندما تواجهنا مشكلة في صورنة بعض الوحدات اللغوية على هيئتها الحالية الموروثة (بفلسفتها الحالية) فإننا بحاجة إلى إعادة النظر في فلسفة اللغة وكيفية عملها في المخ للكشف مجددا عن نواح أخرى تمييزية يمكن من خلالها التعرف على سيرورات لغوية ذهنية جديدة موضوعية يمكن صورنتها ثم محاكاتها. ولذا نحن بحاجة ماسة إلى مواصلة التوصيف اللغوي المستمر والتعرف على إمكانيات الفلسفة الحديثة وعلم الأعصاب اللغوي في هذا المجال. وعليه يمكن أن نرصد العلاقة بين مقولات اللغة ومقولات العقل أو بوصف آخر بين علاقات البنى اللغوية وقوانين العقل.

- **لماذا ندرس المنطق في اللسانيات؟**

نحن ندرس المنطق في اللسانيات لاعتقادنا أن اللغة في بنيتها العميقة قبل إنجاز الجملة عبارة عن قضايا تترابط وفق اشتراطات معينة لتحقيق المصادقات، وأنها تحتكم إلى قواعد تستجيب لحتمية العمليات المنطقية [(١)]. وأن العلاقة التي تربط الدال بالمدلول وتربط الدلالة بالمرجع أو تصديق الحكم على القضية أو أيا كانت تسميتها لا تحتكم إلى قوانين لغوية –وأنى لها! – وإنما تحتكم إلى قوانين منطقية رياضية يمكن صورنتها. فالخطاب وهو الغاية اللغوية مؤسس على الشرطيات المنطقية.

كما أن احتكامنا إلى اللغة الطبيعية ذات القضايا غير المحدودة سيحوجنا إلى براهين غير محدودة أيضًا، وهذا لا يعني أن اعتمادنا اللغات الصورية سيجنبنا كثيرا من العراقيل؛ غير أن الأمر سيكون أقل بكثير مما يواجهنا عند اعتماد اللغات الطبيعية[(٢)].

---

(١) وهذا ما تؤكده مدرسة نحاة بور- رويال Por-Royal. ويرى ذلك أيضا ديكرو في كتابه (الدليل والمقول) (la preuve et le dire:١٩٧٢) حين يعتقد أن النسق المنطقي والرياضي يشكل البنية العميقة للغات الطبيعية. انظر العزاوي، أبو بكر: المنطق واللغة مدخل نظري ٢٠١٤. مطبعة طوب بريس. الرباط. ص:٩

(٢) انظر: **باهي، حسان: اللغة والمنطق** ص:٤٦

- **وظيفة المنطق الصوري في اللغة**[1]:

١- وظيفة الصورنة formalization: نمذجة السيرورات اللغوية بلغة صورية يمكن التعامل معها من قبل الحاسوب.وتمكن من العمليات الآلية كالتوليد والتحليل. وعمليات الإحصاء الرياضي، والتطبيقات اللغوية المختلفة.

٢- وظيفة استكشافية heuristic: وتشمل عمليات الاسترجاع والفهرسة من ناحية، كما تشمل عمليات سياقية ونتائج أكبر تعقيدًا، تتعلق بتحليل الخطاب وكشف أساليبه وتلازماته. كما يمكن الإفادة منها في معرفة الحجاج اللغوي وبناء النتائج عليه.

- **سمات اللغة الصورية:**

١- أنها تبتعد عن كل ما لا يمكن صورنته مثل الدلالة والتأويل والصوت. ويمكن للآلة التعرف على النتائج الدلالية والصوتية في نهاية المرحلة الصورية بناء على السمات التمييزية المدمجة في النموذج الصوري؛ ولذلك وجدنا تشومسكي في التوليدية قد قصد إلى إسقاط المكون الدلالي والمكون الصوتي من جميع مراحل النظرية التوليدية واعتبره نتيجة عندما جعل له مرحلة في البرنامج الأدنوي minimalist program وإنما يجب على اللغة الصورية أن تتكفل بمجموعة من السمات المميزة التي تفضي عند توافرها بشروط معينة إلى المكونين الدلالي والصوتي باعتبارهما نتيجة.

٢- أنها تركز على التركيب وشكل بناء الجملة الصورية بغض النظر عن القيم التي تحملها فهي تكافئ القيمة المطلقة في الرياضيات.

٣- أن اللغة الصورية قابلة لإعادة الكتابة، وهي قاعدة توليدية نقلها تشومسكي من الرياضيات عن العالم الرياضي (إميل بوست) فالنحو التوليدي يندرج رياضيا تحت نظرية الدوال التكرارية recursivity أو ما يسمى بالعودية. وقد

(١) انظر: **العزاوي، أبو بكر: المنطق واللغة مدخل نظري** ٢٠١٤. طوب بريس الرباط. ص٢٠

وظفها تشومسكي في البنى التركيبية[١]. وهي الخاصية التي تميز بها الملكة اللغوية وتمثل الفرق بينه وبين غيره من الحيوانات.

٤- يجب أن تكون اللغة الصورية قابلة للحوسبة computable؛ ولا يمكن ذلك إلا بوضع معايير كافية شاملة بسيطة أدنوية كما يشترط النحو التوليدي الذي يعد أحد الأنحاء الصورية.

- **التراتبية اللغوية:**

يجب أن نعلم قبل الشروع في أي صورنة حقيقة التراتبية اللغوية؛ فهناك لغة في الذهن ولغة عبرت عنها صوتيا أو كتابيها وهي اللغة الطبيعية المنجزة ولغة وصفت تلك اللغة الطبيعية وهي طبيعية أيضا [٢]، ثم لغة مصورنة حاولت أن تصف هذه اللغة الطبيعية وفق الشكل١ الآتي:

(١) انظر: المالكي، طارق: الاستدلال في المنطق وتطبيقاته في اللسانيات. دار كنوز. الأردن. ٢٠١٨.ص:١٩٧

(٢) انظر مثلا: **باهي، حسان: اللغة والمنطق بحث في المفارقات**. دار الأمان. الرباط.٢٠١٥م. ص:٨٠. وهو يتبع تارسكي في كتابه (the semantic conception of truth and the foundations of semantics) في أن اللغة الطبيعية هي اللغة الشيئية، واللغة الطبيعية الواصفة سماها اللغة الفوقية.

الشكل ١: مخطط التراتبية اللغوية

٢- اللغة والرياضيات

لماذا الرياضيات؟

تُعد الرياضيات الضامن العلمي لعلمية العلوم، والمقياس الذي يُحتكم إليه في تجريبية المعارف. كما أن الانضباط الرياضي يتمثل هدفه في تطوير الأفكار ودراستها التي تشكل أساس جهاز الصورنة formalism لوصف بنية اللغات الطبيعية.

ما اللسانيات الرياضية؟: هي المجال العلمي الذي يستخدم أساليب ومفاهيم رياضية لدراسة النظم اللغوية [١]؛ لنمذجة الظواهر اللغوية في نظم منطقية أكثر ثراء ومرونة يمكن للآلة التعامل معها [٢]. وهي فرع من فروع علم المنطق الرياضي.

(1) Geoffrey K. Pullum and Andr´as Kornai Final version: Mathematical Linguistics.p2 http://www.kornai.com/MatLing/matling3.pdf

(2) van Benthem, Johan, and Alice G. B. ter Meulen. 2010. *Handbook of logic and language*. 2d ed. Amsterdam and New York: Elsevier Science

والرياضيات بطبيعتها تحقق قدرا كبيرا من الموضوعية؛ فإن المنوال الرياضي يتعهد بتوفير خصائص البنية الدقيقة وهي:

١- قابلية الدحض falsifiabilityالتي تعتمد على الاستلزام المنطقي في القضية الرياضية المشهورة (إذا كان أ صحيحا فإن نفي أ خطأ) علمًا أن النتيجة هنا تعتمد على صحة الافتراض الأصلي.

٢- القدرة على التنبؤ predictability وهذا يعتمد على المنهج الإحصائي الذي يبني النتيجة على كثرة التواتر والورود. ويعطي ملامح عامة من خلال دراسة مجموعة من النماذج يتعقب فيها العنصر المراد حسابه، وبعد الإحصائية يصل الحاسوب إلى نتائج التعقب التي ترسم الملامح العامة للعنصر. وعليه يمكن الصول إلى نماذج ذكية Smart Paradigms تعطي الحاسوب قدرة على التنبؤ(١). وهذه العملية من أهم مقومات العلوم فيجب على النظرية أن توجد حلولا لما لم تواجهه ولم تدرسه كما وجدت حلولا لما واجهته ودرسته، وهو من أهم ما تبنته التوليدية منذ نشأتها ١٩٥٦م على يد تشومسكي.

٣- الموضوعية Objectivity: وهي من أهم أسس الاستقراء بغض النظر عن مصدر العنصر. ومن المسلم به أن اللغات الطبيعية ذات طبيعة معقدة ونتائج متناقضة لا يمكن أن تكون موضوعية بذاتها فيجب اللجوء إلى الرياضيات لتطوير لغة خاصة ومبسطة من التقنيات الرياضية للوصول إلى الهدف المنشود.

بين اللسانيات الرياضية والجبرية: فرقت بار هليلBAR-HILLEL بين اللسانيات الرياضية mathematical linguistics واللسانيات الجبرية algebraic linguistics باعتبار أن الثانية هي التي تكتفي من الرياضيات بالإحصاء؛ أما الأولى فهي باب أكبر من رديفه(٢). ويرى سولمون ماركوس Solomon Marcus أن

(1) Gregoire D ´ etrez and Aarne Ranta: Smart Paradigms and the Predictability and Complexity of Inflectional Morphology.p645

(2) Y. Bar-Hillel, Four lectures on algebraic linguistics and machine translation. A revised version of a series of lectures given in July, 1962, before a NATO advanced summer institute of automatic translation of languages, in Venice, Italy. p:1

اللسانيات الجبرية تغطي نوعين أساسيين من النماذج اللسانية هما: التوليد والتحليل[1]. والتوليد في أبسط صوره هو الذي يعتمد أساسًا تنطلق منه، بينما التحليل عكسه وهو العودة بالعملية من النتائج المولدة إلى الأساس الذي يُعد نقطة التوليد.

- **الإجراءات العلمية [2]scientific method** :

١- ملاحظة الظاهرة.

٢- صوغ فرضية لوصفها.

٣- إجراء تجربة أو جمع ملاحظات لاختبار الفرضية.

٤- تحليل النتائج لمعرفة ما إذا كانت تؤكد الفرضية أو تدحضها.

علمًا أنه في بعض الأشكال اللغوية لا يمكن أن تنطبق كل هذه الخطوات كاملة؛ لأن المعالجة تختلف من مجال إلى آخر ومن هدف إلى آخر. في حين أننا في بعض مسائل التحليل الصرفي الذي يعتمد على مدخلات كافية وضرورية يمكن أن نطبق هذه الخطوات كاملة كما نجد في شروط جمع المذكر السالم معادلة مطردة.

- **المسَلّمات والبرهان والحدس:**

المسلمات Axioms : ويمكن تسميتها بالبديهيات وهي العبارة الرياضية أو اللغوية التي تعد مثالا صحيحا لا تحتاج إلى برهنة. ويمكن أن ننطلق منها لصناعة الجملة الصورية ونحاكم إليها الجمل المولدة صحة ولحنًا.

ومن المسلمات في النحو العربي:

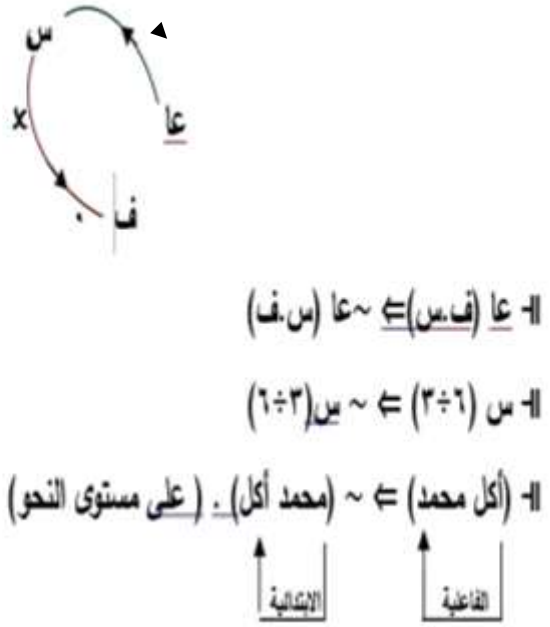
- أن العمل غير متناظر وهو بهذا يشبه عمليتي القسمة والطرح في الرياضيات. فإذا كان الفعل عاملا في الاسم فلا يمكن أن يكون الاسم عاملا في الفعل. ويمكن صياغتها رياضيا بهذا الشكل:

ف.س≠ س.ف

(1) Solomon Marcus: Algebraic Linguistics; Analytical Models.1967. Academic press new york and London. p vii

(2) David Eddington: linguistics and the scientific method. Southwest journal of linguistics, volume 27, number 2 (2008). p:1

أو بالمنطق الرياضي[١] كما في الشكل٢ :



الشكل٢: المنطق الرياضي لنظرية العامل

البرهان: في أصله إجراء رياضي حُجيّ/جدلي argumentum يستند إليه في الحكم على صحة العبارة أو لحنها.

الحدس: هو ما صححه التجريب، وعجز الرياضيون عن برهنته[٢].

(١) انظر: **المالكي، طارق: الاستدلال في المنطق وتطبيقاته في الرياضيات**.٢٠١٨ م. دار كنوز المعرفة. الأردن. ص:٢٦٩

(2) David N. Perkins: Software Goes to School: Teaching for Understanding with New Technologies. New York Oxford university Press 1995.p:93 https://books.google.com.sa/books?id=JyKelnvECc4C&pg=PA93&lpg=PA93&dq=%22although+counterpoint+between+the+particular+and+the+general%22&source=bl&ots=hIvE9Onw06&sig=U3DTIi0tkie-oZdllilyAT4y1DE&hl=en&sa=X&ei=rAEwUJeZLazqigK3_IHABQ&redir_esc=y#v=onepage&q=%22although%20counterpoint%20between%20the%20particular%20and%20the%20general%22&f=false.

- **الحوسبة والصورنة:**

تنقسم محاولات الحوسبة عموما **إلى** طريقين: **الأولى** تعتمد على بحوث اللسانيين في اللغات والفلسفة والمنطق. وتراعي هذه الطريقة المنوالات اللسانية في اللغات ومحاولة محاكاتها Simulation ويقوم **فيها** الباحثون بالأعمال المشتركة بين اللسانيين والحاسوبيين ولا تعتمد هذه الأعمال على المدونات اللغوية بشكل أساسيّ.

والثانية تعتمد على العمليات الإحصائية [1] لتواتر الظواهر اللغوية، والاحتمالات الرياضية بشيء من التنبؤ ثم يتعلم الحاسوب بواسطة تلك النتائج. وهذه الطريق تعتمد المدونات اللغوية الموسومة Tagged بشكل أساسيّ، المعتمد على التعليم تحت الإشراف، وإنما تراعي المنهج اللساني في بداية تدريب الآلة بالتعليم الموجه supervised ثم تتعلم الآلة بواسطة الإحصاء والتكرار.

(١) انظر: **الحباشة، صابر: علاقة اللسانيات بالرياضيات، رهانات أم عقبات؟**: مجلة المخاطبات. تونس. العدد ٥، فبراير ٢٠١٣م. ص:١٠٩

ثانيا – النماذج:

وسأعرض في هذا الجزء لبعض النماذج التي يمكن صورنتها وتهيئتها للبرمجة:

أولا– نظرية المجموعات ودورها في (صياغة جمع المذكر السالم):

نعلم أنه يشترط لصوغ جمع المذكر السالم من المفرد أن يكون المفرد إما علما مذكرا ليس فيه علامة تأنيث، أو أن يكون وصفا لمذكر خال من علامة التأنيث لا يساوى فيه المذكر والمؤنث وليس على أفعل الذي مؤنثه فعلاء.(١)

وعليه يمكن أن نصوغ البرهان الرياضي كما يلي:

نفرض أن س = اسم مفرد تتوفر فيه شروط جمع المذكر السالم

ونفترض أن ص = جمع المذكر السالم.

فإن ص= س+ ون

والسؤال الآن: كيف تستطيع إثبات أن س تتوفر فيه كل شروط جمع المذكر السالم؟

وهنا نلجأ إلى نظرية المجموعات الرياضية، حيث نبني في مدخلات الحاسوب مجموعة رياضية تحوي كل الكلمات التي تتوفر فيها كل شروط جمع المذكر السالم وذلك من خلال البناء الأنطولوجي الذي يصنف كل الموجودات المعجمية باستعمال المعجم المحوسب.(٢)

وعليه نفترض أن هذه المجموعة المرادة رمزها ج = ج١ Uج٢.

حيث ج١ = {س = (علم، مذكر، عاقل، ليس فيه علامة تأنيث)}

---

(١) يشترط في جمع المذكر السالم:١– إذا كان جامدا أن يكون علما، لمذكر، عاقل، خاليا من تاء التأنيث، ومن التركيب.

إذا كان صفة أن تكون لمذكر، عاقل، خالية من تاء التأنيث، ليست من باب أفعل فعلاء، ولا من باب فعلان فعلى، ولا مما يستوي فيه المذكر والمؤنث. انظر شرح ابن عقيل على ألفية ابن مالك: دار التراث–دار مصر للطباعة ط٢٠، ١٤٠٠هـ. ١/٦٠

(٢) **انظر: اللتيني، يحيى ٢٠١٩: أسس صناعة المعجم المحوسب**. دار كنوز المعرفة. لأردن.

وج٢ = {س= (وصف، مذكر، ليس على أفعل الذي مؤنثه فعلاء، وليس مما يتساوى في المذكر والمؤنث)}

وعليه حتى تكون ص صحيحة/صادقة فإنه يجب أن ينتمي س لمجموعة ج

ونعبر عنها رياضيا: س∋ ج

ونصوغ البرهان كما يلي:

ص T فقط إذا كانت س∋ ج [1]

$$s\ t\ if\ C \epsilon G$$

أو بعبارة الاستلزام التالية:

$$s \Rightarrow C \epsilon G$$

وعليه نرمز للحاسوب بالرموز التالية:

ص= س+ ون

$$s = c + wn$$

جمع معلم = معلم + ون

جمع معلم = معلمون.

وهذا يعني أن (معلم) تنتمي إلى مجموعة الأسماء التي تتوفر فيها شروط جمع المذكر السالم ج

ويمكن أن نصوغه بالاستدلال المنطقي هكذا: علم مذكر (معلم) ∧ علم مذكر (معلمون)

برهان عكسي:

قد تبدو لنا بعض الأسماء مذكرة لخلوها من علامة التأنيث مثل: (زينب)، ولا يصدق عليها جمع المذكر السالم لأنها لن تكون في مجموعة ج

وعليه: $G \not\ni$ زينب

ومن هنا ف: علم مذكر (زينب)F

(١) (T)يرمز للعبارات الرياضية الصائبة، و (F) كما سيأتي يرمز للعبارات الرياضية الخاطئة.

ثانيا– حساب التفاضل والتكامل لامدا Lambda[١] ودوره في صورنة الدلالة:

وحساب التفاضل والتكامل Lambda يشكل جانبا مهما في تحويل اللغات الطبيعية إلى لغات صورية؛ حيث يمكن أن نعبر عن المعاني بالمشتقات الدالية التي تتغير في كل تشكيل لغوي الأمر الذي تتغير معه الدلالة من تركيب إلى آخر. وفي مبدأ فريجيه Fregeفي التأليفية compositionality[٢] أن معنى أي تعبير معقد يتكون من معاني أجزائه التي يتم دمجها في الجملة. وقد اعتمدته نظرية *مونتيجو* Montague[٣] في الدلالات الشكلية formal semantics . وقد نقلت نظرية النوع *لمونتيجو* إلى اللسانيين إستراتيجية فريجيه جاذبةً تطبيق ***حجة–دالة*** بوصفها أساسًا لما يسمى (***اللصق الدلالي***) لدمج المعاني الجزئية في الجملة[٤].

(١) حساب التفاضل والتكامل لامدا Lambda: ويرمز له بالرمز (λ) وهو نظام صوري وظيفته تجريد المتغيرات وحساب الدوال واستبدالها في التطبيق. وهو النظام الأساسي للمنطق الرياضي. ويعد نموذجًا عاما للحوسبة عموما وحوسبة اللغة خصوصا في كثير من قضاياها. انظر:

Turing, A. M. (December 1937). "Computability and λ-Definability". The Journal of Symbolic Logic. 2 (4): 153–163. doi:10.2307/2268280. JSTOR 2268280

(2) Dever.J.2006. "compositionality" In E. Lepore & B. Smith (eds.), The Oxford Handbook of Philosophy of Language. Oxford University Press: pp. 633-666.

(٣) يعد ***مونتيجو*** أول من اقترح بجدية في أطروحة علمية أن العلاقة بين التركيب والدلالة في اللغات الطبيعية لا يختلف جوهريًا عن العلاقة بين التركيب والدلالة في اللغات الصورية مثل لغة FOL (First-Order-Logic) انظر :

Raffaella Bernardi: Computational Linguistics: Lambda Calculus and NL. KRDB, Free University of Bozen-Bolzano. http://old.unibuc.ro/prof/dinu_a_d/docs/2016/mar/28_09_50_08lambdaBernardi.pdf

(4) look: Bach, Emmon, Oehrle, Richard, and Wheeler, Deirdre eds. 1987. Categorical Grammars and Natural Language Structures. Dordrecht: D. Reidel.

وقبل التمثيل على ذلك يجب أن نعرف ببعض الاختصارات الهامة في جدول ٢:

الرمز	مدلوله
جملة	S
مركب اسمي	NP
مركب فعلي	VP
مركب حرفي	PP
مركب اسمي+ أداة تعريف (تحديد)	DET CNP
مركب اسمي مشترك	CNP
الأعلام: زيد، عمرو، أحمد...	ProperN: *John, Mary, Bill, ...*
فعل متعدٍ: رأى، أحب، أكل...	TV *sees, loves, catches, eats*
فعل لازم: يذهب، يسير، يركب...	IV *walks, talks, runs*
عملية تحويل	T
المحددات / المعرفات:أل، بعض،كل...	DET: *some, a, the, every*
الصفات: اللاحمة، العاشبة، الطويل، القديم، سعيد	ADJ: *carnivorous, happy, skillful, tall, former, alleged, old, ...*
أسماء عامة: رجل، ملك، جراح، عازف...	CN: *man, king, violinist, surgeon, fish, senator, ...*

جدول٢: الفئات النحوية لمنتيجوMontague

١- القيم التركيبية التي اعتمدها منتيجو في [1]:

$S \rightarrow NP\ VP$
$NP \rightarrow DET\ CNP$

(1) Partee, Barbara H. 1999. "Semantics" in R.A. Wilson and F.C. Keil, eds., The MIT Encyclopedia of the Cognitive Sciences. Cambridge, MA: The MIT Press. 739-742.

NP → ProperN
CNP → ADJ CNP
CNP → CNP REL
CNP → CN
VP → TV NP

يحدد الجدول التالي تحليل الجملة حسب مونتيجو Montague بتحديد العناصر النحوية والدلالات الأساسية وحاول إيجاد علاقات بين الفئة النحوية الأنواع الجالية فكل فئة نحوية ترتبط بنوع دلالي معين يأتي في تعبيرات أساسية لتشكيل معجم أساسي مصغر وبهذه العلاقات نحصل على لغة منطقية، وهو تقليد يشبه إلى حد كبير ما انتهجه تشومسكي في تحديد الدلالة المرتبطة بالتفريع المقولي sub-categorization [1]فكل فعل يأخذ نوعا محددا من المفاعيل والفاعلين بسمات دلالية معينة تتغير بتغير الفرع المقولي، وعند منتيجو تتغير بتغير الفئة النحوية، وبمعنى أوضح نجد منتيجو يربط الفئة النحوية بفئة التعبير وأي تغير في طرفي المعادلة سيغير في الطرف الآخر ومن هنا نشأت قائمة منتيجو للتمييز بين دلالات الجمل كما في الشكل٣ التالي:

(١) انظر اللحياني، سرور (٢٠١٠): خصائص الرأس الفعلي وظواهر من انتظام المعجم، منشورات كلية الآداب والفنون والإنسانيات، منوبة. ص: ٦٨

الفئات النحوية Syntactic category	Semantic type النوع الدلالي	Expressions الفئات التعبيرية
S	t	sentences
ProperN	e	names (John جون)
CN(P)	e → t	common noun phrases (cat قط)
NP	(i) e	"e-type" or "referential" NPs
	(ii) e → t	(John جون, the king الملك) NPs as predicates (a student طالب, a king الملك)
ADJ(P)	e → t	predicative adjectives (carnivorous, happy)
REL	e → t	relative clauses (who(m) Mary loves)
VP, IV	e → t	verb phrases, intransitive verbs (loves Mary, is tall,
TV	<e,e> → t	transitive verbs (loves, sees)
Is	none	temporary treatment as in first-order logic: pretend it isn't there!
DET	(i) function from type e→t to type e	the
	(ii) function from type e→t to type e→t	a

شكل٣: قائمة منتيجو للفئات النحوية والفئات التعبيرية

- **استخدام حساب التفاضل والتكامل في تحليل CCG Parsing :**

وهو اختصار Combinatory Categorical Grammar[1] القواعد الفئوية التجميعية. ويستخدم تحليل CCG لتحليل جمل اللغات الطبيعية وذلك من خلال مجموعة من العناصر المعجمية التي تتبع مجموعة مبسطة من القواعد، وهذه القواعد تحدد التعبيرات الممكنة المرتبطة بهذا العنصر المعجمي. ويقوم تحليل CCG بتحديد السلوك المنطقي للجمل الذي يدمج فيه دلالات تلك الجمل للحصول على الدلالة التأليفية للجملة الناتجة عن الربط بين الكلمة المعجمية ونوعها الصرفي ودلالتها كما في الخطاطة التالية:

f: الجملة ← الشكل المنطقي

حيث حيث يقوم محلل CCGبالوصول إلى صورنة للجملة مبنية على علاقة الفئات النحوية بالدلالات ومن ثم يمكن لنا بناء نموذج تعليم الآلة وفق المخطط التالي:

نمذجة الدلالة ← تعليم ← تحليل

(1) S. Clark and J. R. Curran, "Wide-coverage efficient statistical parsing with CCG and log-linear models." Computational Linguistics, vol. 33, 2007.

وعليه فمثال: محمد يأكل الخبز

محمد ---- NP

يأكل ----- (S\ NP)/NP

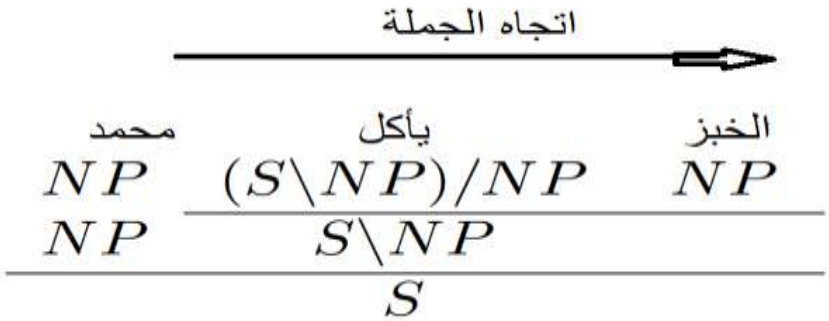
الخبز ---- NP

"محمد" و "الخبز" مركبان اسميان لهما سلوك محدد، و"يأكل" مركب فعلي معقد يتولى أكثر من مهمة في التركيب حيث ترتبط به المركبات الاسمية في الجملة وهذا الدور يجعل له سلوكا معينا يختلف عن بقية المركبات.

وتمثل العلاقة بين اللغة الطبيعية واللغة المنطقية المعتدة على تحليل CCG كما في الشكل٤و٥ التاليين:

اللغة الصورية لمخرجات CCG (القوالب Templates)	اللغة الطبيعية (الوحدات المعجمية Lexemes)
$\vdash S\backslash NP: _\lambda x.$يأكل(x)	يأكل
$\vdash NP: _\lambda x.$محمد(x)	محمد
$\vdash NP: _\lambda x.$الخبز(x)	الخبز

الشكل٤: مقابلة اللغة الطبيعية باللغة الصورية بتحليل CCG



الشكل٥: يمثل العلاقة بين اللغة الطبيعية واللغة الصورية المعتمدة على محلل CCG

حيث إن:NP,S, رموز الفئات النحوية.

و (\، /) رموز عوامل التجميع النحوي.

يوضح الجدول أعلاه والشكل كيف يمكننا صورنة اللغة الطبيعية باستخدام محلل CCG في قوالب محددة وعليه تنتج لنا معجم للمدخلات ومعجم للقوالب يشكل السلوك المنطقي لتلك المدخلات، وذلك بالكشف عن الاختلافات المنهجية في استخدام الكلمات، ثم يمكن بعد ذلك تطبيق كل اختلاف على الوحدات المدمجة ذات المعنى المعجمي.

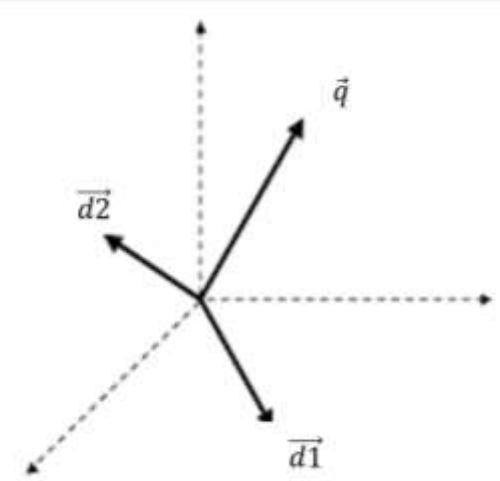
وهذه صورة مختصرة لكيفية الإفادة من حساب التفاضل والتكامل في صورنة اللغة وتهيئتها لتعلم الآلة.

ثالثا- الجداء السلمي والتمييز الدلالي:

يمكن أن نرصد الحقول الدلالية التي تنتمي لها النصوص بالاعتماد على مفهوم الفضاء الشعاعي بحساب نسبة توارد كلمات ذلك الحقل في النص وعليه اختيار المعنى السياقي المراد بما يخدم في الترجمة الآلية وغيرها من التطبيقات اللغوية مثل تصنيف المدونات وفهرستها واسترجاع المعلومات والبحث فيها.

ويكون الحساب باستخدام آليات الجداء السلمي. ونحسب البعد والقرب من الحقل الدلالي بحساب ما يسمى المسافة الزاوية بالقياس التشابهي sim(x,y)؛ فكلما تغير القيمة الشعاعية للمدونة تغير الحقل الذي تنتمي إليه؛ وذلك بعد حساب الفضاء الشعاعي Vector Spaces الذي يمثل الحقول الدلالية في مساحة فضاء

متعدد الأبعاد، بحيث تمثل المفردات المميزة لكل حقل دلالي بعدا معينا يرمز له بالرمز($\overrightarrow{d1}$) و ($\overrightarrow{d2}$)وهكذا بحسب عدد الحقول الدلالية المتشابهة. وتمثل مفردات المدونة المدخلة لغرض الترجمة –مثلا– متجها يرمز له بالرمز ($\overrightarrow{q}$) كما في الشكل(٦)؛ وعليه فالمدونة التي تحوي مجموعة من المفردات المميزة لحقل دلالي معين فستظهر نتيجة التشابه بقيمة أكبر من الأخرى ومن ثَمّ يمكن تحديد المجال الذي تنتمي إليه المدونة المدخلة بحسب قيمة الحساب الزاوي كما سيأتي.



الشكل٦: فضاء المتجه الذي يمثل المدونة المدخلة ومفردات الحقول الدلالية المتشابهة.

- **طريقة حساب التشابه:**

نستعمل معامل جاكارد $d(c_j,c_i)$ [١] لمقارنة التشابه والاختلاف بين مجموعات العينات المحدودة. كلما زادت قيمة معامل Jaccard ، زاد تشابه المدونة المدخلة مع مفردات حقل دلالي معين وذلك باحتساب نتيجة قسمة تقاطع عينتين على مجموع

(١) هناك العديد من مقاييس التشابه للمسافة في معالجة اللغات الطبيعية، وهنا بعض المقاييس شائعة الاستخدام، بما في ذلك خوارزمية تعديل المسافة ومعامل جاكارد وخوارزمية سميث ووترمان. وتعتمد مكتبة nltk قانون zipf وحساب متجه الكلمات AVG-W2V.

المكونات وستقوم الآلة بالترجمة وفقا للحقل الذي انتمت إليه المدونة المدخلة [١] كما في المعادلة التالية:

$$d(ci,cj) = \frac{b+c}{a+b+c}$$

حيث:

a: مجموع قيم مفردات المدونة المدخلة.

b: مجموع المفردات المميزة لكل حقل من الحقول المتشابهة.

- **طريقة حساب المسافة الزاوية:**

بعد حساب القياس التشابهي(x,y) sim بمعامل جاكارد بين مفردات المدونة المدخلة والمفردات المميزة للحقول الدلالية المتشابهة، ونعتبر الحقول الدلالية قاعدةَ الفضاء الشعاعي كما في الشكل(٧). حيث يمثل حساب خط مفردات الحقل الدلالي متجها معينا. وتمثل زاوية البعد بين المدونة والحقل الدلالي بالقياسات التشابهية كما يلي:

$$D(C, H_k) = \frac{1}{H_k} \sum_{(C_j \in H_k)} d(C, C_i)$$

حيث:

H_k: الحقل الدلالي ويمثل محورا في الفضاء الشعاعي

N_k: عدد الكلمات في الحقل الدلالH_k .

C : كلمة تنتمي إلى الحقل الدلاليH_k . (الحواس ، الطعام ، القضاء).

(١) ويمكن استخدام ذلك في معالجة اللغات الطبيعية من خلال لغة Python باستدعاء حزمة CountVectorizer من مكتبة sklearn.feature_extraction.text مع حساب متجه الكلمات word2vec واستخدام AVG-W2V لحساب متوسط متجه الجملة. للمزيد انظر: موقع المبرمج العربي https://arabicprogrammer.com/article/79741264314/ بتاريخ ٢٤ \ ١٠ \ ٢٠٢١م. الوقت ٢٠:١١

$d(C_i, C_j)$: نتيجة القياس التشابهي.

مثال: كلمة (شهد) [١].

له معاني مختلفة بحسب السياق الذي يرد فيه. فقد يكون في أحد الحقول الآتية:

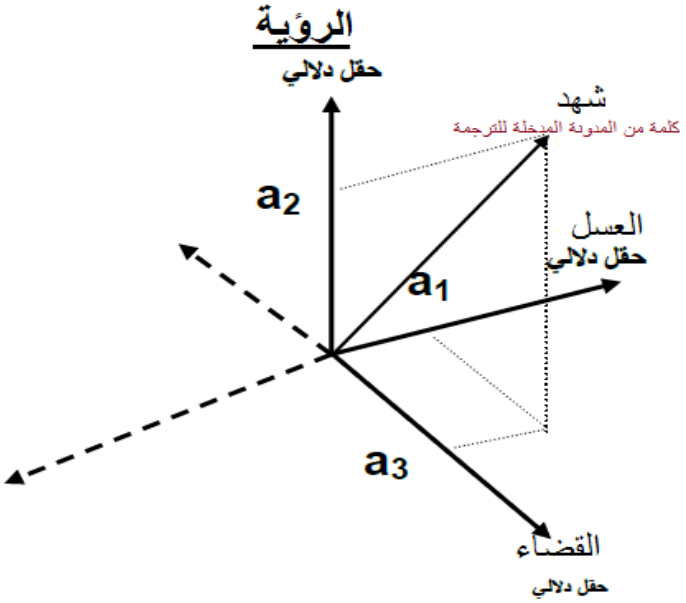
١- الحواس

٢- الطعام

٣- القضاء

عيون	عسل	حضور
نظر	شمع	محكمة
مشاهدة	نحل	قضية
منظر	شفاء	دعوى
رؤية	نحل	قاض
إبصار	دوق	خصومة
حدة		

(١) كلمة (شهد) تتعدد دلالاتها بحسب السياق فقد تكون فعلا (شَهِدَ) في المحكمة أو بمعنى (حَضَرَ) أو اسما بمعنى العسل(شَّهْدٌ)



الشكل٧: تمثيل الحقول الدلالية على محاور الفضاء الشعاعي

الخاتمة:

لقد حاولت في هذا البحث أن أعرض مناطق التقاطع بين الرياضيات واللسانيات التي تتمثل في صورنة اللغة بأشكال مختلفة مفيدا من الموضوعات الرياضية مثل المجموعات الرياضية والتفاضل والتكامل وخاصة حساب لامدا Lamda الذي يعد الأساس الذي اعتمد عليه حساب CCG للكشف عن السلوك المنطقي للجمل. وأخيرا الجداء السلمي الذي يهتم بحساب المتجه للكلمات word direction. فمن خلال المجموعات الرياضية يمكن معالجة بعض القضايا في النحو العربي التي تجعل مجموعة من الوظائف النحوية ترتبط بمجموعة من الخصائص المحددة لها. ومن خلال التكامل والتفاضل يمكن معالجة بعض القضايا الدلالية التي يمثل اشتقاق الدالة وتحولها أهم علامات التواؤم بينها وبين تغير الدلالة بتغير أشكال التركيب وأنواع المركبات. وأخيرا يمثل الجبر الخطي وحساب الجداء السلمي أهم أدوات حساب وزن الكلمات في النصوص لتصنيفها وتحديد مجالاتها وتوجيه المعنى المراد من بين المعاني المترادفة وتحديد المصطلحات لكل تخصص، وهو الأمر ذاته المستخدم في تحليل المشاعر.

المصادر والمراجع

المراجع العربية:

باهي، حسان: اللغة والمنطق بحث في المفارقات. دار الأمان. الرباط.٢٠١٥م

الحباشة، صابر: علاقة اللسانيات بالرياضيات، رهانات أم عقبات؟: مجلة المخاطبات. تونس العدد ٥، فبراير ٢٠١٣م

العزاوي، أبو بكر ٢٠١٤.: المنطق واللغة مدخل نظري مطبعة طوب بريس. الرباط.

اللحياني، سرور (٢٠١٠): خصائص الرأس الفعلي وظواهر من انتظام المعجم، منشورات كلية الآداب والفنون والإنسانيات، منوبة.

اللتيني، يحيى ٢٠١٨: أسس صناعة المعجم المحوسب. دار كنوز للتوزيع والنشر. الأردن.

المالكي، طارق: الاستدلال في المنطق وتطبيقاته في اللسانيات. دار كنوز. الأردن.٢٠١٨.

المراجع الأجنبية:

Bach, Emmon, Oehrle, Richard, and Wheeler, Deirdre eds. 1987. Categorial Grammars and Natural Language Structures. Dordrecht: D. Reidel

David Eddington 2008: linguistics and the scientific method. Southwest journal of linguistics, volume 27, number 2 (2008).

David N. Perkins: Software Goes to School: Teaching for Understanding with New Technologies. New York Oxford university Press 1995

Dever.J.2006. "compositionality" In E. Lepore & B. Smith (eds.), The Oxford Handbook of Philosophy of Language. Oxford University Press.

Geoffrey K. Pullum and Andras Kornai2008 Final version :Mathematical Linguistics

Gregoire D ´ etrez and Aarne Ranta: Smart Paradigms and the Predictability and Complexity of Inflectional Morphology

Partee, Barbara H. 1999. "Semantics" in R.A. Wilson and F.C. Keil, eds., The MIT Encyclopedia of the Cognitive Sciences. Cambridge, MA: The MIT Press

Raffaella Bernardi: Computational Linguistics: Lambda Calculus and NL .KRDB, Free University of Bozen-Bolzano.

Solomon Marcus 1967: Algebraic Linguistics; Analytical Models. Academic press new york and London

Turing, A. M. (December 1937). "Computability and λ-Definability". The Journal of Symbolic Logic. 2 (4): 153–163. doi:10.2307/2268280. JSTOR 2268280

Van Benthem, Johan, and Alice G. B. ter Meulen. 2010. Handbook of logic and language. 2d ed. Amsterdam and New York: Elsevier Science

Y. Bar-Hillel, 1962: Four lectures on algebraic linguistics and machine translation. A revised version of a series of lectures given in July, 1962, before a NATO advanced summer institute of automatic translation of languages, in Venice, Italy.

المواقع الالكترونية:

موقع المبرمج العربي

https://arabicprogrammer.com/article/79741264314/

Bibliography

Al-ʿAzāwi, Abu Bakr : **Logic and language theoretical introduction 2014.** (in Arabic) Rabat: publications of Toub Bris.

Al-Habāsha, Ṣābir: **The relationship of linguistics to mathematics, bets or obstacles?** (in Arabic) al-Mukhātabāt Journal. Tunisia: issue 5, Febuary 2013.

Al-Latīnī, Yahya 2018: Foundations of a Computerized Arabic Lexicon. (in Arabic) Jordan: Dār Kunūz.

Al-Lihyānī, Surur (2010): **Actual head characteristics and phenomena of lexical regularity,** (in Arabic). Manouba: publications of college of Arts and Humanities.

Al-Maliki, Tariq: Inference in Logic and its Applications in Linguistics. (in Arabic). Jordan: Dār Kunuz, 2018.

Bāhi, Hasān: **Language and logic: a search for paradoxes**. (in Arabic). Rabat: Dār al-Amān, 2015.