الجمهورية الجزائرية الديموقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة عبد الحميد بن باديس

كلية الأدب العربي والفنون

قسم الدراسات اللغوية والأدبية

مذكرة بعنوان:

اللسانيات الحاسوبية والمعالجة الآلية للغة العربية

مذكرة لنيل شهادة الماستر في اللغة والأدب العربي

تخصص: لسانيات تطبيقية

إعداد الطالبة:

- موجب إيمان.

إشراف الأستاذة:

- بحوص نوال.

بحوص نوال
أستاذة محاضرة - أ
جامعة عبد الحميد بن باديس - مستغانم

لجنة المناقشة:

الأستاذ(ة)	الرتبة	الصفة
د. بن دحان عبد الوهاب	أستاذ محاضر (أ)	رئيسا
د. بحوص نوال	أستاذة محاضرة (أ)	مشرفا ومقررا
د. صادق فاطمة	أستاذة محاضرة (أ)	ممتحنا

السنة الجامعية: 2022م- 2023م.

الجمهورية الجزائرية الديموقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة عبد الحميد بن باديس

كلية الأدب العربي والفنون

قسم الدراسات اللغوية والأدبية

مذكرة بعنوان:

اللسانيات الحاسوبية والمعالجة الآلية للغة العربية

مذكرة لنيل شهادة الماستر في اللغة والأدب العربي

تخصص: لسانيات تطبيقية

إعداد الطالبة:

- موجب إيمان.

إشراف الأستاذة:

- بحوص نوال.

لجنة المناقشة:

الأستاذ(ة)	الرتبة	الصفة
د. بن دحان عبد الوهاب	أستاذ محاضر (أ)	رئيسا
د. بحوص نوال	أستاذة محاضرة (أ)	مشرفا ومقررا
د. صادق فاطمة	أستاذة محاضرة (أ)	ممتحنا

السنة الجامعية: 2022م– 2023م.

شكر وتقدير:

عن أبي هريرة –رضي الله عنه–قال رسول الله ﷺ: ((من لا يشكر الناس، لا يشكر الله)).

يسرني تقديم الشكر والتقدير للأستاذة المشرفة على هذا البحث المتواضع، والقيم راجية من الله تعالى أن يضيف قيمة إلى هذا العلم.

الشكر والامتنان للسادة أعضاء لجنة المناقشة الذين تفضلوا بقبول مناقشة هذه المذكرة.

والشكر والعرفان لمن ساعدني في تقديم يد العون من أساتذة ودكاترة ومنهم د. سلوى حمادة (أستاذة اللسانيات الحاسوبية بالقاهرة).

وشكر خاص لأستاذة وإدارة جامعة مستغانم ومساعدتهم في كل الأمور التي يحتاجها طالب العلم.

وشكر خاص للأهل والأصدقاء وكل من ساندني ودعمني إلى آخر دقيقة وشكرا لكل شخص منحني طاقة إيجابية.

وفي الأخير أقول شكرا لنفسي لأنها تعبت معي كل هذه السنوات من أجل هذه اللحظة لتذوق فرحة التخرج والشكر لله عز وجل على هذا اليوم، وفرحتي أتمناها لكل طالب وطالبة علم بإذن الله.

إيمان موجب.

الملخص:

منذ ظهور اللسانيات إلى يومنا هذا لازالت تشهد تطورات علمية كبيرة، تجسدت في انفتاحها على علوم عديدة من قبل علم الاجتماع، علم النفس، الهندسة والحاسوب مما أدى إلى تنوع فروعها ومنها أساسا اللسانيات الحاسوبية، هذه الأخيرة تمثل الربط بين اللغة والحاسوب.

فاللسانيات الحاسوبية تهدف إلى استغلال التقنيات التي يقدمها الحاسوب لمعالجة اللغة العربية استجابة لتطورات العصر وأمر المعالجة يتطلب معرفتين، الأولى: لسانية تلم بكل التفاصيل الدقيقة للنظام اللغوي، والثانية: حاسوبية تضع برامج وفق قواعد خوارزمية صورية.

إذ يتضح من البحث اللساني أنه لم يعد يهتم بالطرق، التي بموجبها تتكون الجمل في لغة من اللغات فقط، وإنما أصبح موضوع اللسانيات يتسع ليشمل التعرف على الكيفية التي يشتغل بها الدماغ البشري، أي تحديد الآلة اللغوية ذاتها، في محاولة التعرف على سر تكوينها وطريقة اشتغالها لدى الانسان عامة، و لعل هذا هو الإشكال الذي يثار والسر الذي استعصى فهمه حتى على العلوم الحقة التي تمتلك من الأدوات الإجرائية ما يمكنها من ذلك، وتعتبر اللغة العربية من اللغات العالمية التي استطاعت مسايرة التقدم الحضاري عبر التاريخ، إذ كانت على صلة بحضارات مختلفة، واستطاعت بفضل جهود أبنائها أن تستوعب كل تلك الحضارات وتتواصل معها. وكان استخدام اللغة العربية في الحاسوب دافعا للعديد من الباحثين اللسانيين للتعاون مع الباحثين الحاسوبيين للسعي وراء حلول جذرية تعالج مختلف إشكالات اللسانيات الحاسوبية العربية، فأصبح من الضروري إعادة صياغة مورثاتنا اللغوية من الدراسات اللغوية صياغة جديدة، وذلك على ضوء الإنجازات التي تحققت على ضوء اللسانيات الحاسوبية في اللغات الأخرى حتى يصبح بالإمكان التعامل السليم مع اللغة العربية بخصائصها المتميزة.

الكلمات المفتاحية:

الدراسات اللغوية – اللسانيات الحاسوبية – البيانات والمعلومات اللغوية – المعالجة الآلية – اللغة – الحاسوب– الآلية.

Résumé :

Depuis l'avènement de la linguistique jusqu'à nos jours, elle connaît encore de grands développements scientifiques, incarnés par son ouverture à de nombreuses sciences par la sociologie, la psychologie, l'ingénierie et l'informatique, qui ont conduit à la diversité de ses branches, dont principalement la linguistique computationnelle, cette dernière représentant le lien entre le langage et l'ordinateur.

La linguistique computationnelle vise à exploiter les techniques fournies par l'ordinateur pour traiter la langue arabe en réponse aux évolutions de l'époque, et l'ordre de traitement requiert deux connaissances, la première : une linguistique connaissant tous les détails fins du système linguistique, et le second : un ordinateur qui développe des programmes selon les règles d'un algorithme formel.

Il ressort clairement de la recherche linguistique qu'elle ne s'intéresse plus uniquement à la manière dont les phrases sont formées dans une langue, mais plutôt que le sujet de la linguistique s'est élargi pour inclure l'identification du fonctionnement du cerveau humain, c'est-à-dire la définition de la machine du langage. Elle-même, pour tenter d'identifier le secret de sa formation et de son fonctionnement. Pour l'homme en général, et c'est peut-être là le problème qui se pose et le secret difficile à comprendre même pour les vraies sciences qui possèdent les outils procéduraux qui Grâce aux efforts de ses enfants, elle a pu s'imprégner et communiquer avec toutes ces civilisations. L'utilisation de la langue arabe dans l'ordinateur a été un motif pour de nombreux chercheurs en

linguistique de coopérer avec des chercheurs en informatique pour rechercher des solutions radicales qui répondent aux divers problèmes de la linguistique informatique arabe, il est donc devenu nécessaire de reformuler nos héritages linguistiques des études linguistiques dans un nouveau formulation, à la lumière des réalisations faites à la lumière de la linguistique computationnelle dans d'autres langues afin qu'il soit possible de traiter correctement la langue arabe avec ses caractéristiques distinctes.

Les mots clés:

Études linguistiques - linguistique computationnelle - données et informations linguistiques - traitement automatique - langage - ordinateur - Automatique.

Since the advent of linguistics to this day, it is still witnessing great scientific developments, embodied in its openness to many sciences by sociology, psychology, engineering and computer science, which led to the diversity of its branches, including mainly computational linguistics, the latter representing the link between language and the computer.

Computational linguistics aims to exploit the techniques provided by the computer to process the Arabic language in response to the developments of the time, the matter of processing requires two knowledges, the first : a linguistics knowlegeable in all the fine details of the linguistic system , and the second: a computer that develops programs according to the rules of a formal algorithm.

It is clear from linguistic research that it is no longer concerned with the ways in which sentences are formed in one of languages only, but

rather that the topic of linguistics has expanded to include the identification how the human brain operates. i.e, defining linguistic machine it self, in an attemp to identify the secret of its formation and the way it operates in humans in general. The world that has been able to keep pace with civilizational progress throughout history.

keywords:

Linguistic studies - computational linguistics - linguistic data and information - automatic processing – the language – the computer - machine.

خطة البحث:

مقدمة.

مدخل.

الفصل الأول: اللسانيات الحاسوبية.

المبحث الأول: تعريف اللسانيات الحاسوبية.

المبحث الثاني: تاريخها (النشأة والتطور).

المبحث الثالث: موضوع اللسانيات الحاسوبية ومجالاتها.

المبحث الرابع: وظيفة اللسانيات الحاسوبية وفوائدها.

الفصل الثاني: المعالجة الآلية الجانب النظري.

المبحث الأول: البداية والتاريخ للمعالجة الآلية.

المبحث الثاني: تعريف المعالجة الآلية.

المبحث الثالث: دور المعالجة الآلية في الحفاظ على اللغة العربية.

المبحث الرابع: واقع المعالجة الآلية للغة العربية.

الفصل الثالث: المعالجة الآلية للغة العربية.

المبحث الأول: المستوى المورفولوجي.

- بنية الكلمة.
- الاشتقاق.
- المعالجة الآلية للمستوى المورفولوجي.

المبحث الثاني: المستوى النحوي.

- من الكلمة إلى الجملة.
- تمثيل الجملة.
- التحليل العام للجملة.
- المعالجة الآلية للمستوى النحوي.

المبحث الثالث: المستوى الدلالي.

- تمثيل المعاني.
- من القاموس إلى خريطة المفاهيم.

المبحث الرابع: نحو إنشاء قاعدة معجمية آلية في شكل معجم كوني هرمي.

- **لماذا المعاجم الآلية.**
- منهجية إنشاء معجم يبدأ من مفهوم الكون في تسلسل هرمي.
- خطوات إنشاء قاعدة البيانات المعجمية الآلية.

الفصل الرابع: طرق وبرامج المعالجة الآلية.

المبحث الأول: طرق المعالجة الآلية.

- الطرق المعتمدة على المعلومات المرتبطة بالنص.
- الطرق المعتمدة على المعطيات المعالجة.

المبحث الثاني: البرمجيات التجارية لمعالجة اللّغة العربية.

- **الدراسات الأكاديمية:**

1. نظام مسبار.
2. برنامج التقسيم.
3. معجم اللغة العربية المؤتمن.
4. أعمال شركة الصخر.

المبحث الثالث: سبيل المعالجة الآلية للغة العربية.

1. نظام OCR.
2. التدقيق النحوي الآلي.
3. البنك السعودي للأصوات نموذج تطوير بيانات لمعالجة اللغة العربية.
4. البرنامج الرقمي العربي.

المبحث الرابع: المعالجة الآلية للغة العربية ومجالات تطبيقاتها في ظل الذكاء الاصطناعي واللسانيات الحاسوبية.

1. اللغة بين الدماغ البشري والحاسوب.
2. الذكاء الاصطناعي ومجالات تطبيقاته.
3. تطبيقات الذكاء الاصطناعي.

خاتمة.

قائمة المصادر والمراجع.

الفهرس.

مقدمة

يتميز العصر الحالي بالتطور الكبير الحاصل في مجال المعلومات والمعرفة التي أحدثه الحاسوب ليصبح مقياسا لتقدم الأمم، فمنذ ظهور البرامج الحاسوبية أحدثت انقلابا حقيقيا في علاقة الانسان بالتكنولوجيا، وعززت حاجة المجتمع إلى اعتماد اقتصاد المعرفة لتجاوز اختلالات التطور والتنمية وسدّ الفجوة الرقمية في المعالجة الآلية في ظل اللسانيات الحاسوبية.

وباعتبار اللغة هوية كل أمة مثلا اللّغة الإنجليزية هي اللّغة نشأ بها جهاز الحاسوب وبهذا أصبحت لغة راقية وعالمية فصار من ضروري على اللّغة العربية أن تدخل مجال الحوسبة، فكان بذلك نشأة علم جديد هو اللسانيات الحاسوبية وكيفية تطبيقها على اللّغة العربية وما مدى التفاعل بينها وبين مجال الحاسوب، وكيفية معالجتها آليا، فاللسانيات الحاسوبية لم تعد درسا يؤسس على مناهج بسيطة محددة، تنطلق من التحليل الآلي للغات الطبيعية على المستويات الصرفية والنحوية والدلالية، بل أصبحت رهانا لتطور اللغات والتعليم والإنتاج العلمي والإبداع التكنولوجي، ولمواكبة تطور تطبيقات المعالجة الآلية للغات التي نستعملها نحن البشر، صّممت برامج ونظم محوسبة لمعالجة اللغات.

ولهذا اخترت موضوع "اللسانيات الحاسوبية والمعالجة الألية للغة العربية" ليكون مادة البحث والذي سيجيب عن التساؤلات التالية:

- ما هي اللسانيات الحاسوبية؟
- ما المقصود بالمعالجة الآلية؟
- كيف يمكن معالجة اللغة العربية آليا؟

إن قلة الدراسات الجديدة في مجال اللسانيات الحاسوبية وقلة العلوم المرتبطة بعصر المعلومات، كان من أهم الأسباب لاختيار هذا الموضوع للدراسة والبحث فيه من أجل إثرائه والاستفادة منه وإفادته للأجيال القادمة، ولأن اللسانيات الحاسوبية هي من الفروع المهمة في اللسانيات التطبيقية، وبهذا تعتبر دراسة جديدة لتفتح المجال أمام اللغة العربية للولوج إلى عالم الحواسيب وقدرتها على المعالجة الآلية، كل هذا كل حافزا ودافعًا للكشف عن أهمية اللسانيات الحاسوبية وفائدتها على اللغة العربية وعالجها آليا والحفاظ عليها.

وبهذا اعتمدت على خطة بحث، حيث تناولت في المدخل نظرة عامة عن اللسانيات وتكلمت أيضا عن اللسانيات التطبيقية وقلت أن اللسانيات الحاسوبية هي أحد أهم فروعها وعرفت العنوان مفصلا أي شرحت مصطلحات عنوان البحث، أما في الفصل الأول خصصته لتبيان اللسانيات الحاسوبية عبر أربع مباحث حيث

عرفتها وذكرت نشأتها وفوائدها والفصل الثاني خصصته للمعالجة الآلية للغة العربية وهذا عن طريق مستوياتها مثل المستوى المورفولوجي هو الذي يهتم بكيفية بناء الكلمة وجذرها والزوائد التي تلحقها. والمستوى النحوي هو الذي يتم فيه تحديد محلات إعراب الكلمة وموقعها في الجملة. والمستوى الدلالي هو الذي يركز فيه كل معاني الكلمة ومترادفاتها والاستعارة. والفصل الرابع تكلمت فيه عن الطرق والبرامج التجارية لمعالجة اللغة العربية وخاتمة عبارة عن حوصلة لنتائج البحث المتوصل إليها، ثم اعتمدت على مجموعة من المصادر والمراجع منها: اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، المجلس الأعلى للغة العربية للغة العربية، أعمال الندوة الوطنية 24. 25 ديسمبر 2019 منشورات المجلس، سناء منعم اللسانيات والترجمة الآلية بعض الثوابت النظرية والإجرائية، سلوى حمادة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي ومن أهم الصعوبات التي واجهتها هي قلة المصادر بهذا الموضوع وإن توفرت كتب إلكترونية في مواقع معينة تتكلم عن هذا الموضوع لا يسمح بتحميل الكتاب أو قراءته، وأحيانا وجود المعلومات ولكن عدم توفر معلومات ناشريها وبهذا لا نستطيع الأخذ منها لأن من شروط البحث التهميش وإرجاع المعلومة إلى صاحبها، قلة الدراسات العربية خصوصا في هذا المجال مما أد إلى قلة المصادر والأبحاث فمثلا المعالجة الآلية للغة العربية لا توجد دراسات سابقة لها وقليلة جدا وعند دراستها أو معالجتها يتطلب نمذجتها بلغة أجنبية من أجل المقارنة والدراسة.

وفي الأخير أقول أن هذه الصعوبات ما هي إلا جهود وتحديات واجهتها لأقدم وأنفع الدراسة اللغوية وأواصل البحث مستقبلا وأتمنى أن أجبت على جميع تساؤلاتكم وشكرا.

المدخل

علم اللسان أو اللسانيات أو علم اللغة هو علم حديث العهد، ظهر في بداية القرن 19م، فاللسانيات هي الدراسة العلمية الموضوعية للسان البشري أي دراسة تلك الظاهرة العامة والمشتركة بين بني البشر والجديرة بالاهتمام والدراسة بغض النظر عن كل الاعتبارات الأخرى التي لا تعد من صلب اهتمام اللسانيين، تختص بجوانب ثانوية للسان بحكمه ظاهرة معقدة ومركبة يمكن أن تتناول من زوايا عديدة اجتماعية نفسية، فيزيولوجية وفيزيائية تتكفل بها علوم أخرى مثل علم الاجتماع[1] وعلم النفس وعلوم فيزيولوجية الأعضاء وعلم الصوت الفيزيائي.

إذن يقصد بالدراسة العلمية الموضوعية للسان البشري هو إتباع الوسائل العلمية والموضوعية أي عكس الذاتية بمعنى الوسائل العلمية في الدراسة والتحليل والموضوعية بطبيعة الحال عكس الذاتية ... وبهذا تكون حاملة معها طابعا علميا جديدا لدراسة اللغة ... ومما هو متعارف عليه أن الدراسات في مجال اللسانيات قد توسعت وتعددت وذلك راجع إلى التوجهات والاهتمامات التي يعتمد عليها كل دارس في هذا المجال، وهناك فرع آخر وهي لسانيات، ولكنها تطبيقية، وتعريف هذا العلم ليس سهلا وعندما نرجع إلى المراجع العربية والأجنبية نجدها متعددة، وإنما هناك تعريفات متعددة ترتبط بعلوم اللغة النظرية والوصفية من جهة وبعض الأخرى مثل علم اللغة النفسي والتربوي والمقاميات ونظرية أعمال الكلام وعلوم الحاسب، إذن هناك جانبان جانب علوم نظرية وجانب علوم أخرى تتداخل مع بعضها البعض وهي تطبيقية لأنها تتعامل مع الحياة، فعلم اللغة التطبيقي أو ما يسمى باللسانيات التطبيقية، وهو حقل من حقول اللسانيات، ظهر سنة 1946م في الوقت الذي ظهر الاهتمام بمشاكل تعليم اللغات الحية للأجانب، إلى جانب ازدهار الدراسات التطبيقية، أو نظرية علمية يتم تمثلها عن طريق تطبيق ما هو في الإمكان، وذلك بتكوين المادة عن طريق الأنماط وترسيخ المفاهيم التي يتم فيها نقل النتائج والنظرية إلى مستوى تطبيقي[2].

من خلال تعريف صالح بلعيد للسانيات التطبيقية بأنها حقل من الحقول المعرفية الحديثة التي ساهمت في ترقية الحصيلة العلمية والمعرفية، وتطوير طرق ووسائل تعليمها وتعلمها للناطقين ولغير الناطقين بها، فظهرت مع مشكلة تعليم اللغات الحية للأجانب من أجل تحسين نوعية تعليمها ... وتشتمل لعدة مجالات منها: اللسانيات النفسية، التقابلية، الاجتماعية، الترجمة وصناعة المعاجم...

ومن أهم فروع اللسانيات التطبيقية هي اللسانيات الحاسوبية التي ترتكز في عملها على الحاسوب من منطلق أن له القدرة على الاحتفاظ بقدر كبير من المعلومات والمعارف، ولعل العلاقة التي جعلته يرتبط كل هذا الارتباط باللغات هو محاولة اللسانيين استثمار الحاسوب في مجال اللغويات بغرض تسهيل استغلال كل تلك المعارف

[1] خولة طالب الإبراهيمي، مبادئ في اللسانيات، دار القصبة للنشر، (2000- 2006)، الطبعة الثانية، ص07.
[2] صالح بلعيد، دروس في اللسانيات التطبيقية، دار هومه للطباعة والنشر، ط8، ص11.

والمعلومات، فاللسانيات الحاسوبية هي علوم حديثة تستخدم الحاسوب في تحويل النصوص، والمعلومات اللغوية إلى لغات الحاسب الرقمية لتحليلها، وترجمتها للغات أخرى، وتطوير نماذج اختبار للعمليات اللغوية، في هذا العلم تشارك اللسانيات في المساعدة على فهم خصائص المعطيات اللغوية و كذلك تقديم نظريات تفيد في كيفية بناء اللغة واستعمالها، من أهم الخدمات التي يقدمها علم اللغة الحاسبي هي المشاركة في تقديم نظريات وتقنيات تمكن من وضع برامج حاسوبية تساعد على فهم اللغة الطبيعية.

إن الدراسات والبحوث العلمية في اللسانيات الرتابية (أو الحاسوبية) ازدهرت في الوطن العربي في هذه الآونة، وتكاثر إلى حد ما الباحثون في هذا الميدان الذي تتلاقى في علوم الحاسوب (أو المعلوميات) وعلوم اللسان، وهو ميدان علمي وتطبيقي واسع جدا كما هو معروف إذ يشمل التطبيقات الكثيرة كالترجمة الآلية والإصلاح الآلي للأخطاء المطبعية وتعليم اللغات بالحاسوب والعمل الوثائقي الآلي، وتنطبق الآلات بالتركيب الاصطناعي للأصوات اللغوية وغير ذلك كثير وهي من البحوث الطلائعية وفائدتها بالنسبة للعربية عظيمة جدا [1]. إذا بعد هذا التحليل الشامل توصلنا إلى ضبط المصطلحات الآتية:

اللسانيات الحاسوبية: هي علم متخصص في دراسة اللغات ومعالجتها بشكل تطبيقي لخلق برامج[2] وأنظمة معلوماتية ذكية يتحدد دورها في مساعدة مستخدم الحاسوب على حل بعض الأمور المتعلقة باللغة والمعلومات الرقمية بشكل عام، فهي إذن نظام بيني، بين اللسانيات وعلم الحاسوب المعني بحوسبة الملكة اللغوية. إنها مجال معرفي لغوي يعتمد الحاسوب في معالجة البيانات والمعلومات اللغوية، وهي أولا وقبل كل شيء فرع من اللسانيات التطبيقية مجال واسع يضم اللغة والكلام، والتعدد اللغوي والتواصل، والتربية، والتعلم، وكذا التقنيات المعلوماتية المتقدمة وبحوث الاتصال وغيرها من مجالات الدراسة والتطبيق، وارتباط اللسانيات الحاسوبية باللسانيات التطبيقية أساسه الالتقاء الحاصل بين الاتجاهين والمتمثل في الاهتمام بالدراسات العلمية الميدانية في مجال اللغة من أجل استخلاص نتائج تكون قاعدة لبناء نظريات قائمة على التطبيق والتجريب.

إن اللسانيات الحاسوبية باختصار هي العلم الذي يبحث في اللغة البشرية كأداة طبيعية[3]لمعالجتها في الحاسوب، وتتألف مبادئ هذا العلم من اللسانيات العامة بجميع مستوياتها التحليلية: الصوتية، والصرفية، والنحوية، والتركيبية

[1] عبد الرحمن الحاج صالح، بحوث ودراسات في اللسانيات العربية، موفم للنشر، د.ط، ص230.
[2] سناء منعم، اللسانيات الحاسوبية والترجمة الآلية بعض الثوابت النظرية والإجرائية، منشورات مختبر العلوم المعرفية، عالم الكتب الحديث للنشر والتوزيع، ط01، 2015، ص93.
[3] سناء منعم، اللسانيات الحاسوبية والترجمة الآلية بعض الثوابت النظرية والإجرائية، منشورات مختبر العلوم المعرفية، عالم الكتب الحديث للنشر والتوزيع، ط01، 2015، ص93.

والدلالية، ومن علم الحاسوب، وعلم الذكاء الاصطناعي، وعلم المنطق ثم علم الرياضيات، والواقع أن تمثيل المعرفة الإنسانية في الحاسوب باعتباره آلة تكنولوجية، مرتبط ارتباطا وثيقا بتحليل اللغات الإنسانية وتركيبها، وعلى اللغويين الحاسوبيين أن يتزودوا بعدة معرفية وتطبيقية في مجال الرياضيات والإحصاء، و نظرية المعلومات، ونظرية الأوتوماتية، ونظرية المعرفة، والمنطق، وهندسة النظم وهندسة الإشارات. ويجدر الإشارة إلى أن تعريف اللسانيات الحاسوبية يختلف من باحث إلى آخر، ويعتمد ذلك على الحقل الذي يعمل به اللساني ثم التجربة العلمية التي يخوضها، فبعض الباحثين يعرف هذا العلم على أنه العمل اللغوي الذي يعالج في الحاسوب–وهو ما يعرف اختصارا بحوسبة اللغة–[1] ويعرفه بعض الباحثين الآخرين على أنه جزء من علم الذكاء الاصطناعي، وجعله يعني كباقي حقوله بدراسة كيفية تشكيل القدرة الذهنية لد المتكلمين وانشغالها باللغة الطبيعية، وطبقا لوجهة نظر هؤلاء الباحثين، فإن اللسانيات الحاسوبية هي الاستعمال الدقيق للحاسوب لإجراء بعض العمليات الرياضية فيه والتي تشبه العمليات المنطقية الرياضية التي يقوم بها الذهن البشري.

المعالجة الآلية:

أ– المعالجة: المعالجة من وجهة نظر علم اللغة الحاسوبي هي التطبيق الآلي على مجموعة من نصوص اللغة وذلك بتغيرها وتحويلها، وإبداع شيء جديد اعتمادا عليها ويتم ذلك باستعمال تقنيات وأدوات من علوم اللسانيات والإعلام الآلي والنمذجة (modélisation) ويجب التفرقة عند المعالجة بين وصف المعارف وهي وظيفة اللسانيات والتعبير عن هذه المعارف في نماذج باستخدام تقنيات[2] واستراتيجيات فعّالة مستمدة من علوم الحاسوب، وهي وظيفة علم اللغة الحاسوبي.

ب– الآلية: Automatique

العمليات الآلية هي التي تجري عن طريق الآلة والتي تقابلها العمليات التي تجري بواسطة الإنسان.

والآلة التي تستعمل في المعالجة الآلية للغة هي الحاسوب الذي اخترع لإجراء العمليات الحسابية، لذا يجب تطويره لمعالجة المعلومات ذات الطبيعة اللسانية، حيث أن المعالجة الآلية هي تتابع حركات حسابية تقوم بها الآلة وفق تسلسل زمني أي أن برنامج المعالجة الآلية (Programme automatique) يمكن أن يكون كلي (Total) أو جزئي (Partiel) حيث أن:

[1] المرجع نفسه، ص94.

[2] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي إعرابي للفعل العربي، مذكرة ماجستير في علم المكتبات والتوثيق جامعة الجزائر، كلية العلوم الاجتماعية والإنسانية، الجزائر، 2008، ص12.

1- **كلي**: يقوم الحاسوب بكل شيء.

2- **جزئي**: يتدخل الإنسان في بعض المراحل.

ومعالجة شيء لساني من وجهة الآلة يواجه بقيود (Contraintes) كثيرة في وصف النصوص اللغوية نفسها، لذا يجب نمذجة مكونات النصوص بطريقة واضحة ومتناسقة (Cohérence)[1].

اللغة العربية:

اللغة هي رموز منطوقة أو مكتوبة أو كلاهما معا، هذبها الإنسان وصقلها كي تعبر عن حاجته ومطالبه، ولتكون وسيلة الاتصال والتفاهم مع غيره، وهي بهذا ضرورية لكل من الفرد والمجتمع.

واللغة العربية هي لغة القرآن الكريم الذي نزل على سيد الخلق أجمعين محمد بن عبد الله النبي الأمي الأمين الذي فتح القلوب والعقول فرسم الإنسانية طريق حياتها القائمة على توسم الخير والبعد عن الضلال والاقتراب من الهداية ... واللغة العربية هي إحدى اللغات السامية وأرقاها مبنى ومعنى واشتقاقا وتركيبا[2].

إذ تعتبر اللغة قلب الأمم النابض ولسانها الناطق ومفتاح هويتها وقاموس أسرارها، وهي أهم أدوات التفكير والتواصل واللغة العربية واحدة من أهم اللغات المعاصرة، التي تنتشر في مجال جغرافي واسع يمتد على قارتين، ويتحدث بها ملايين النساء، إذ تعد واحدة من أقدم اللغات في العالم وأكثرها ثراء من حيث الألفاظ والتراكيب والأساليب البيانية، الأمر الذي جذب إليها اهتمام علماء اللغة العرب والغربيين الذين انبهروا بالجمالية المنقطعة لهذه اللغة التي لها فضائل وجماليات منها: الترادف وجمالية الفروق والدرجات، الاشتقاق(مفردات عديدة ذات أصل واحد)، المحسنات البديعية، النحو والصرف في اللغة العربية، نظام صوتي راسخ منذ القدم ومعجم لغوي ثري...

إذ تعد اللسانيات الحاسوبية علم تطبيقي حديث، يستغل ما توفره التكنولوجيا المتطورة في برمجة أنظمة المعلوماتية من أجل معالجة اللغات معالجة آلية، فهي علوم حديثة تستخدم الحواسيب في كتابة النصوص اللغوية وتحويلها إلى لغات الحاسب الرقمية لتحليلها، فاللسانيات الحاسوبية علم يربط بين اللسانيات وعلم الحاسوب، وهو مجال ينتمي إلى مجالات الذكاء الاصطناعي، ويسعى لمحاكاة الآلة للغة الطبيعية البشرية، وتعتبر اللغة العربية من اللغات العالمية التي استطاعت مسايرة التقدم الحضاري عبر التاريخ، إذ كانت على صلة بحضارات مختلفة، وكان استخدام اللغة العربية في الحاسوب، دافعا للعديد من الباحثين اللسانيين للتعاون مع الباحثين الحاسوبيين، للسعي إل حلول جذرية

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي إعرابي للفعل العربي، مرجع سابق، ص13.
[2] ينظر: زكريا إسماعيل، طرق تدريس اللغة العربية، دار المعرفة الجامعية، ط01، ص33 – 34.

تعالج مختلف إشكالات اللسانيات الحاسوبية العربية، لذلك كان الحاسوب الركيزة الأساسية في هذا الفرع اللساني التطبيقي الحديث المتصل بالذكاء الاصطناعي، باعتبار هذا الأخير يركز على قواعد المعارف الأعمق بها والأشمل من قواعد البيانات واللغة العربية استفادت منه كثيرا في إنجاز تطبيقات لغوية حققت بها قفزة نوعية مما يفسر أنه لا حل لمعضلة اللغة، وبالتالي فإن اللسانيات الحاسوبية ترتبط بالدراسات الحديثة لأجل تحقيق التكامل والمنفعة فيما بينها ومع الحقول المعرفية الأخرى وذلك بالنظر إلى الجهود اللغوية والجهود الحاسوبية.

واللسانيات الحاسوبية لها دور كبير في إدخال اللغة العربية إلى مجال المعلوماتية، إذ تعتبر من أحدث الاتجاهات اللغوية في اللسانيات العربية المعاصرة فهي تقوم على التنظيم والبرمجة[1].

وفي الأخير نقول إن ميدان اللسانيات الحاسوبية لا يزال حقلا خصبا يعوزه العمل الدؤوب الجاد، ونأمل أن نرتقب المزيد من إسهامات اللسانيين والحاسوبيين، والهدف الرئيسي للسانيات الحاسوبية يتمثل في السرعة والدقة العلمية، ومن ثم الحصول على الترجمات الآلية من اللغة العربية وإليها تضيف اللغوي القائم على نحو علمي وسريع جدا يتماشى مع التطورات الحديثة.

[1] ينظر: كنزة منديل، مجالات استخدام اللسانيات الحاسوبية في البحث اللساني العربي، جامعة محمد لمين دباغين سطيف 02 (الجزائر)، 2022/05/19، مجلة الدراسات لجامعة سطيف، ص146.

الفصل الأول:

اللسانيات الحاسوبية.

المبحث الأول: تعريف اللسانيات الحاسوبية.

المبحث الثاني: تاريخها (النشأة والتطور).

المبحث الثالث: موضوع اللسانيات الحاسوبية ومجالاتها.

المبحث الرابع: وظيفة اللسانيات الحاسوبية وفوائدها.

المبحث الأول: تعريف اللسانيات الحاسوبية.

1- اللسانيات الحاسوبية (Computational linguistics):

علم اللغة الحاسوبي أو اللسانيات الحاسوبية اسمان لعلم واحد حديث يختلف عن علمي (الحاسوب) و(اللسانيات) بشكلهما المستقل، نشأ لتمكين الحاسوب من استخدام اللغة لمخاطبة مستخدمه وتوسيع آفاق الاستخدام من جهة، وخدمة اللغة باستخدام تقنية الحاسوب من جهة أخرى[1].

ويحاول علم اللغة الحاسوبي استكناه بنية اللغة على المستوى الصوتي والصرفي والنحوي والدلالي، فهو من هذه الناحية علم لغوي، وهو من ناحية أخرى علم حاسوبي تطبيقي من حيث البحث في إمكان معالجة هذه اللغة آليا وفق البنية اللغوية، إذ يحاول هذا العلم تشكيل بنية حاسوبية افتراضية للّغة، من خلال محاولة فهم الكيفية التي يتحدث بها الناطق العاقل للغة، أي الإنسان، ومن ثم يعمل على محاكاته ضمن محددات الحاسوب التي تتطلب تعريف القواعد والمفردات، وتوصيفها للمعالج الإلكتروني بشكل دقيق، وذلك لخلوه من ميزة العقل التي يمتاز بها الناطق العاقل، ونجاح هذا العلم في أداء مهمته يؤدي بالنتيجة إلى إمكان معالج اللغة قيد البحث قيد البحث آليا، ويفتح الآفاق أمام تطبيقات واسعة وضرورية تجعل من هذا الاستخدام الآلي ميزة لهذه اللغة وللناطقين بها.

ومن هذه التطبيقات على سبيل المثال لا الحصر الترجمة وتشمل ترجمة المفردات والجمل والنصوص، ومن هذه التطبيقات تعليم اللغة عن بعد، وهذه التطبيقات وغيرها استخدامات لا تخفى أهميتها للناطقين باللغة كافة بالغة في وقت أصبح فيه الحاسوب أداة يوصف من لا يستخدمها بـ "الأمية".

2- ما بين (العربية والحاسوب) و (علم اللغة والحاسوب):

نرمي من موضوع (العربية والحاسوب) إلى ما تفيده اللغة العربية من الحاسوب، وكيف بإمكانه أن يخدمها، وفي هذا المقام لا يمكن لما أن نتكرر التقصير الكبير في حسن توظيف الحاسوب لخدمة اللغة العربية بجميع مستوياتها (صوتية وصرفية وتركيبية ودلالية)، ونمثل على ذلك بعدم وجود برنامج يخدم الأصوات العربية، -خصوصا- ويبين مخارجها وكيفية نطقها، والتقصير في هذا الجانب الكبير، وهو عائد لأبناء اللغة ونابع من ركودهم في خدمة لغتهم.

أما (علم اللغة والحاسوب) فإننا نلاحظ أن الحديث في هذا المجال يعني بعلم اللغة العام، وكيف يمكن للحاسوب أن يعين عالم اللغة أو اللساني في جمع البيانات والوصف والتحليل للوصول إلى مبتغاه.

[1] عيسى برهومة، مقدمة في اللسانيات، دار الكتب، الجامعة الهاشمية، 2005، د.ط، ص119-120.

والاسم العلمي لكلا المبحثين السابقين يدل على أنهما يفيدان من الحاسوب بوصفه آلة أو أداة لها قدرة فائقة على التخزين والتحليل والتنظيم، فقد أمكن لقرون أن تتم خدمة علم اللغة دون استخدام الحاسوب، لكن استخدامه في الوقت الحالي يفتح آفاقًا لم يكن من الممكن الوصول إليها دونه.

ونتجت اللسانيات الحاسوبية عن امتزاج: (علم اللغة) و (علم الحاسوب)، ولعلنا نلاحظ أن كلا العلمين جديد، وبإمكاننا أن نحدد سمات كل منهما على حدة، في القرن العشرين، فكيف بعلم اللغة الحاسوبي الذي ما يزال يعيش في بدايات عقده الرابع[1].

ويهدف علم اللغة الحاسوبي إلى "إيجاد الوسائل التي تمكننا من العلاج الآلي للكلام من جميع نواحيه ولشتى الأغراض، منها بناء قاعدة المعطيات، وهي الأنماط اللغوية الرياضية التي يضعها ويحررها وينتقدها المهندس واللغوي معًا".

"وتعد تقنيات المعالجة الآلية للغات هي البناء التحتي (Baik inforatructure) الواجب توفيره، حيث يتم فوقها بناء التطبيقات الأخرى كافة، ولا تقف تقنيات المعالجة الآلية حاليا عند إدخال البيانات والمعلومات واستخراجها (In put/ Out put)، بل تمتد إلى التطبيقات اللغوية التخصصية من تدقيق إملائي (Pelling checking) ونحوي (Grammar checking) وتحليل صرفي (Morphological analysis) وترجمة آلية (Machine translation)، وكذلك تقنيات التعرف الضوئي على الكتابات الآلية واليدوية وتمييز الحديث وقراءة النصوص، وما إلى ذلك من تقنيات أساسية".

من يقوم بالعمل؟

وبما أن العلم بَيْني، فبالتأكيد بن يقوم بأبحاثه عالم اللغة وحده أو عالم الحاسوب وحده، ولنا أن نتخيل نتائج الأبحاث المرجوة من عالم الحاسوب لا يعرف من علوم اللغة إلا المبادئ العامة التي تلقاها في مراحل المدرسية، أو من عالم لغة لا يفقه من الحاسوب إلا وظيفته!! فإنشاء قاعدة لغوية حاسوبية يتطلب الاهتمام بتكامل هذين العلمين وبتضافر جهود علمائهما، بل وبإلمام كل منهما بجزء من علم الآخر، على ألا يطغى أحدهما على الآخر.

3- مفهوم اللسانيات الحاسوبية:

يسير علم اللغة الحاسوبي في مسارين: أولهما محاكاة التفكير الإنساني، وذلك بدراسة النظريات المرتبطة بالمعرفة التي يحتاجها الإنسان لإنتاج اللغة وفهمها، ويعد هذا علما نظريًا يحاول العلماء فيه بناء نظام حاسوبي قادر على فهم اللغة الإنسانية وكيفية إنتاجها، تمامًا كما يفعل البشر، ثانيهما: محاكاة الأداء البشري، وهو الثمرة التطبيقية التي

[1] عيسى برهومة، مقدمة في اللسانيات، مرجع سابق، ص120- 121.

يسعى علم اللغة الحاسوبي الوصول إليها: ولقد ركزت هذه التطبيقات على مجالات لغوية محددة، وتجزئتها إلى أجزاء صغيرة (الصوتية، الصرفية، النحوية...إلخ)[1].

تتضمن اللّسانيات الحاسوبية Computational linguistics عنصرين رئيسيين هما: اللّسانيات الحاسوبيّة.

1.1. **تعريف اللسانيات**: أمّا اللّسانيات Linguistics في أبسط تعرف لها هي دراسة اللغة بذاتها ولذاتها دراسة علميّة موضوعيّة، من هذا المنطلق يقول مازن الوعر بأن اللسانيات علم يتسم بسمتين اثنتين:

الأولى: هي العلميّة، أي إن لهذا العلم-كما كان يقول علماؤنا وفلاسفتنا القدماء- حدا وموضوعا وغايّة.

الثانية: هي الاستقلاليّة، أي إن هذا العلم مستقل عن بقية العلوم الأخرى، له مقاييسه، وموازينه وقوانينه الخاصة به.

يقول مازن الوعر: "اللسانيات مصطلح أتى من اللّسان، واللّسان يعني اللغة فأضفنا الياء والألف والتّاء فأصبح علما يبحث في اللّسان أي في اللغة، فاللّسانيات إذن هي الدّراسة العلميّة للغات البشريّة من خلال لغة كل قوم من الأقوام، وعندما نقول علميّة فإننا نعني بها الملاحظة ووضع الفرضيات وفحصها والتّجريب والدّقة والشّمولية والموضوعيّة، وهذه الخصائص هي التي تميز الدّراسة اللغويّة الحديثة عن الدراسة اللغويّة القديمة".

2.1. تعريف الحاسوبية: بينما يقصد بالحاسوبية استثمار التّقانة الحديثة وخاصة الحاسوب بما يحتويه من إمكانات رياضيّة خارقة، وسعة تخزينيّة هائلة في خدمة اللغة يقول محمود إسماعيل صيني: "إن معالجة اللغات الطبيعيّة واحدة من أهم التطبيقات التي تسعى علم الحاسوب إلى التّعامل معها (الجيل الخامس) إنّ معالجة هذه اللغات يمثل حلا لمشكلات الانفجار المعلوماتي في عصرنا"[2].

تعتبر اللّسانيات الحاسوبية Linguistique Computationnelle أحد الفروع التطبيقية تهتم بالإفادة من معطيات الحاسوب في دراسة قضايا اللّسانيات المتعددة مثل: رصد الظّواهر اللغوية وفقا لمستوياتها الصوّتية، الصّرفية، النّحوية، البلاغية والعروضية، وإجراء العمليات الإحصائية، وصناعة المعاجم والترجمة الآلية وتعليم اللغات، إذن تعدّ اللّسانيات الحاسوبيّة من أبرز مجالات اللّسانيات التطبيقية لأنها تجمع بين علمين: "اللسانيات" و "علم الحاسوب"، فهو علم يبحث في اللغة البشريّة كأداة طبيعيّة لمعالجتها بالحاسوب، أي هي دراسة علميّة للغة

[1] عيسى برهومة، مقدمة في اللسانيات، مرجع سابق، ص121- 122.

[2] اللغة العربية بين اللسانيات الرتابية الحاسوبيّة واللّسانيات العرفانية في الجامعات الجزائرية، المجلس الأعلى للّغة العربية، أعمال الندوة الوطنية 24-25 ديسمبر 2019، المكتبة الوطنية العامة، الجزائر، منشورات المجلس، 2019، الجزء الأول، ص82.

الطبيعية من منظور حاسوبي، وهذه الدّراسة لا يمكن أن تتم إلا ببناء برامج حاسوبية لأنظمة اللغات البشرية، ومن هنا فإن اللسانيات الحاسوبية تقوم على جانبين مهمان:

الجانب النظري: الذي يتضمن معرفة الإطار النّظري العميق الذي يعمل في الدّماغ البشري لحل المشكلات الخاصة، والمثال على ذلك هو التّرجمة من لغة إلى أخرى.

وأما الجانب التطبيقي: فيعنى بالبحث في العمليات الرّياضية الخوارزمية التي هي عبارة عن مجموعة من القواعد المنظمة في طريقة معيّنة تنطلق من القواعد البسيطة إلى القواعد المعقدة ثم إلى القواعد التي هي أكثر تعقيدا، فإن الفكرة المهمة في الجانب التطبيقي هي أنه عندما يعمل الحاسب الالكتروني عملا لغويا ويركبه، وهذا العمل اللغوي كان قد حققه الدّماغ البشري، فإن علم اللّسانيات (الحاسوب المعلوماتي) عندها لا يمكن أن يعتبر جزءا من عملية الذكاء الاصطناعي، إذن فإن الجانب التطبيقي للسانيات الحاسوبية يعتمد على حوسبة جوانب الملكة اللغوية أي؛ تطبيق النماذج الحاسوبية على الملكة اللغوية، وفي هذا الصّدد أصبحت اللّسانيات الحاسوبية أو العلم الذي يحاول ربط علاقة بين علمي اللسانيات والمعلوماتية قصد معالجة اللغات الطبيعية معالجة آلية، أي أصبح بإمكان الحاسوب محاكاة نمط اشتغال العقل الإنساني وتقيسه من خلال لغة صورية خوارزمية أشبه ما تكون باللغة الصّناعية، كما أصبح مجالات تطبيقيا لاختيار الفرضيات حول الطريقة التي يشتغل بموجبها العقل الإنساني"[1].

تعد اللّسانيات الحاسوبية أهم فروع اللسانيات التطبيقية وهي علم يبحث في اللغة البشرية باستعمال الحاسوب، وتقوم اللسانيات الحاسوبية على جانبين نظري وتطبيقي، فالجانب النظري يبحث في كيفية عمل الدماغ الإلكتروني لحل المشكلات اللغوية والجانب التطبيقي يعنى بإنتاج برامج ذات معرفة باللغة الإنسانية.

المبحث الثاني: اللسانيات الحاسوبية (النشأة والتطور).

1- اللسانيات الحاسوبية النشأة والتطور:

لقد تم اختراع جهاز الحاسوب -حسب ما تذكره المصادر- في أواخر النصف الأول من القرن المنصرم، وتحديدا عام 1948م. ومنذ ذلك التاريخ أصبح الحاسوب متاحًا للإفادة منه في جميع مجالات الحياة ومختلف العلوم، والمعارف الإنسانية. أما بدايّة توظيف الحاسوب في دراسة اللغة على الصعيد العالمي، فمن الصعب التأريخ له؛ لأنه لم يحدث دفعة واحدة؛ بل تم وفق مجهودات أغلبها فردية وعبر مراحل مختلفة وكذا في دول متعددة، إلا أن البحث اللساني الحاسوبي اتخذ شكله الرسمي الأكاديمي في عام 1954م في جامعة جورج تاون، وقد اتخذ العمل في بداياته طابع الترجمة الآلية من اللغات الأخرى إلى اللغة الإنجليزية، ثم أخذت معالم هذا العلم تتبلور وتتشكل، ودعائمه تترسخ بعقد الملتقيات والندوات وإصدار المجلات"[2].

[1] اللغة العربية بين اللسانيات الرتابية الحاسوبيّة واللّسانيات العرفانية في الجامعات الجزائرية، المرجع السابق، ص204.
[2] سعيد فاهم، قراءة في الإسهامات اللسانية الحاسوبية العربية، مجلة دراسات لجامعة الأغواط-سبتمبر 2015- الطبعة 36، عدد 36، ص130.

وهذا يعني أن بداية الخمسينات من القرن المنصرم شهدت ولادة المعالجة الآلية للغات البشرية، وهذا طبعًا عند الغربيين، أمّا بالنسبة للعلوم النظرية عند العرب في العصر الحاضر، فقد كانت العلوم الشرعية من أسبق العلوم الإنسانية استخدامًا لتقنية الحاسبات الالكترونية، و نظم المعلومات؛ حيث شرع العمل بها والإفادة منها في السبعينات من القرن الماضي، وظلت علوم اللغة العربية في منأى عن الانتفاع بها ردحًا من الزمن، حتى قيّض الله لها من عمّم فائدة استخدام الحاسوب على العلوم العربية، وتبدأ قصة الاتصال العلمي بين الحاسوب والبحث اللغوي العربي عند لقاء الطبيب محمد كامل حسين مع الدكتور إبراهيم أنيس؛ حيث اقترح عليه إمكانية الاستفادة من الحاسوب في البحوث اللّغوية، فلقيت هذه الفكرة قبولا واستحسانا؛ لأنها كانت تداعب خاطره منذ أن سمع بإنجازات الحاسوب في شتى العلوم، ومن ثم انتهز فرصة زيارته لجامعة الكويت سنة 1971م للعمل بها أستاذًا زائرًا وهناك التقى بالدكتور علي حلمي موسى، أستاذ الفيزياء النظرية في جامعة الكويت، وطرح علي فكرة الاستعانة بالحاسوب في إحصاءات الحروف الأصلية بمواد اللغة العربية ابتغاء الوقوف على نسج الكلمة العربية. وقد رحّب بهذه الفكرة واستحسنها، وبدأ بالتخطيط لها وتنفيذها في النصف الأول من عام 1971م، وكان من ثمرة ذلك صدور الدراسة الإحصائية للجذور الثلاثية وغير الثلاثية لمعجم الصحاح الجوهري.

أما الخطوات العمل في هذا الإحصاء، فتوزعت على ثلاث مراحل الأولى: إدخال المواد اللغوية في ذاكرة الكمبيوتر، والثانية: وضع برامج له بإحدى لغات الكمبيوتر والثالثة: التنفيذ الفعلي لهذا البرنامج. وجاءت نتائج هذه الدراسة في صورة جداول إحصائية لجذور اللّغة وحروفها، وتتابع أصواتها وخصائص حروفها، مقرونة بدراسة تحليلية موجزة عن التفسير اللغوي لما ورد في تلك الجداول، وتلقى الباحثون هذا العمل العلمي بقبول حسن ولأوّل مرّة تعاون الفيزيائيين واللغويين حول إحصاء كلمات اللغة العربية. وليس من باب المبالغة في شيء أن نقول إن هذا التوجه في الفكر العربي المعاصر قد فتح بابًا واسعًا للباحثين في الدراسات اللغوية للولوج من خلاله إلى عالم الحاسوب، وتسخيره لخدمة الدرس اللغوي. وهكذا كان حقل الإحصاء اللغوي هو الميدان الأول لتطبيق اللسانيات الحاسوبية. تلك لمحة مقتضية عن نشأة اللّسانيات الحاسوبية عند الغربيين، وكذا العرب"[1].

2- جهود اللسانيين العرب في مجال اللسانيات الحاسوبية:

لكي لا نغمط حق كل من أسهم في تأسيس هذا الحقل اللساني هناك جهود اللغويين العرب نذكر على سبيل المثال لا الحصر: محمّد الحناش (المغرب) أحمد الأخضر غزال (المغرب) مازن الوعر (سوريا) عبد الرحمن الحاج صالح (الجزائر) مراياتي محمد (سوريا) نبيل علي (مصر) نهاد الموسى (فلسطين) وليد العناتي (الأردن) وهناك باحثون كثر لا يسعنا المقام هنا لسرد أسمائهم إلا أننا استأنسنا بالمشهورين والأكثر فاعلية وحضورا في الساحة اللسانية الحاسوبية العربية.

ونستهل حديثنا بالدكتور نبيل علي الذي أصدر كتابا بعنوان "اللغة العربية والحاسوب" في هام 1988م وهو أول كتاب يتناول الهندسة اللسانية العربية بجميع مستوياتها، وقد أولى المستوى الصرفي للغة العربية أهمية خاصة؛

[1] سعيد فاهم، قراءة في الإسهامات اللسانية الحاسوبية العربية، المرجع السابق، ص130-131.

حيث يقول: "إن ميكنة العمليات الصرفية بالنسبة للغة العربية تعد مدخلا أساسيًا وقاسمًا مشتركًا لمعظم نظمها الآلية، كما يشير إلى أن مدى نجاحنا في تعريب نظم المعلومات والمعارف ويتوقف بالدرجة الأولى على ما نستطيع أن نحققه على جبهة الصرف العربي آليًا مطلبًا أساسيًا لميكنة عمليات تحليل النصوص المكتوبة والمنطوقة وفهمها، واسترجاع المعلومات، وتحليل مضمون النصوص"[1].

ويعدّ هذا المصنف القيّم خطوة موفقة، نحو تأسيس لسانيات حاسوبية عربية على أساس نظري وتطبيقي في آن واحد، وهو مؤلف يتناول موضوع اللسانيات الحاسوبية مطبقة على أنظمة اللغة العربية، صوتًا وحرفًا، ونحوًا، ومعجمًا مع المعالجة الآلية لهذه النظم اللّغوية جميعها؛ إذ حالفه التوفيق في كثير من القضايا المتصلة بالحاسوب واللغة، وذلك حينما انطلق في عمله هذا من وضع دراسات تقابلية بين العربية والانجليزية شاملة لكل النظم اللغوية، بالنظر إلى الإنجليزية هي اللغة الأم لتقنيات نظم الحاسوب والمعلومات، وهذا نتج عنه معرفة أوجه الاختلاف والائتلاف بين اللغتين، وكان هذا النهج بمنزلة الأرض الصلبة والقاعدة المتينة التي هيأت للمؤلف منهجية موضوعية، ومكنته من الإسهام الإيجابي في جهود تعريب الحاسوب من جهة، والمعالجة الآلية للغة العربية من جهة أخرى.

إن هذا الكتاب يمثل حجر الأساس في مسيرة الدرس اللساني الحاسوبي؛ بل إنه كما وصفه الدكتور نهاد الموسى–بحق خطوة واسعة واثقة– تنتظم مشروعا مستوعبًا لتأسيس اللسانيات في العربية على أساس نظري وتطبيقي في آن واحدٍ، صحيح أنه لم يستوعب جميع قضايا اللغة كونه الأول الذي يشق طريق هذا الفن"[2].

إذن بدأت الدراسات اللسانية الحاسوبية تشغل حيزا كبيرا من اهتمام المشتغلين في حقل اللسانيات اللغوية والبرمجيات الحاسوبية في الآن ذاته، وهذا ما جعلها محط اهتمام من قبل الباحثين والمختصين، فعند الغرب تعود الانطلاقة الأولى لنشأة اللسانيات الحاسوبية عندما بدأ الاهتمام من أهل الاختصاص في اللسانيات والعلاج الآلي للمعلومات بشكل الصياغة المنطقية الرياضية ينبغي أن تصاغ بها النظريات اللسانية... وعند العرب فقد كانت بداية الاستفادة في هذا العلم في السبعينيات وكانت الزيارة التي قام بها الدكتور إبراهيم أنيس إلى جامعة الكويت عام 1971م، فضلا كبيرا للسانيات الحاسوبية العربية فقد التقى بالدكتور علي حلمي موسى "أستاذ الفيزياء النظرية في جامعة الكويت" حيث طرح عليه فكرة الإحصاءات اللغوية وأهميتها في البحث اللغوي... ومن ثم اتفقى على البدء بدراسة إحصائية للجذور الثلاثية وغير الثلاثية لمعجم الصحاح مثلا وانتهت الدراسة بثلاث مراحل أولها إدخال المادة اللغوية في ذاكرة الحاسوب وثانيا: وضع برنامج له بإحدى لغات الكمبيوتر وثالثا التنفيذ الفعلي لهذا البرنامج... وقد حضي هذا العمل بترحيب من قبل العلماء والباحثين.

[1] سعيد فاهم، قراءة في الإسهامات اللسانية الحاسوبية العربية، المرجع السابق، ص134.
[2] المرجع نفسه، ص134.

المبحث الثالث: موضوع اللسانيات الحاسوبية ومجالاتها.

1/ موضوع اللسانيات الحاسوبية:

تعتبر اللسانيات الحاسوبية **L'inguistique Computationnelle** أحد فروع اللسانيات التطبيقية تهتم بالإفادة من معطيات الحاسوب في دراسة قضايا اللسانيات المتعددة مثل رصد الظواهر اللغوية وفقا لمستوياتها، الصّوتية، الصرفية النحوية، البلاغية والعروضية وإجراء العمليات الإحصائية، وصناعة المعاجم والترجمة الآلية، وتعليم اللغات، فهو علم يهتم باللغة، بحيث يبحث في اللغة البشرية كأداة طبيعية لمعالجتها في الآلة (الحاسبات الإلكترونية، الكمبيوتر) وتتألف مبادئ هذا العلم من اللسانيات العامة بجميع مستوياتها التحليلية والصوتية والنحوية، والدلالية، ومن علم الحاسوب الإلكتروني (الكمبيوتر)، ومن علم الذكاء الاصطناعي، وعلم المنطق، ثم علم الرياضيات حيث تتناسق هذه الفروع وتتألف لتشكل مبادئ علم اللسانيات الآلي"[1].

2/ مجالات اللسانية الحاسوبية:

إن مجالات اللسانيات الحاسوبية متعددة وهي قابلة للتوسع والدخول في مجالات أخرى أما أهم المجالات التي تستخدم فيها اليوم فهي[2]:

1/2. تحليل النصوص آليا: إن الكم الهائل من النصوص اللغوية يستدعي استخدام برمجيات متخصصة في النصوص، ومن أهم هذه البرمجيات الكلمات المتوافقة وتعمل هذه البرمجيات على إنجاز المهام التّالية:

- الإحصاء العددي وتتضمن استخراج النسبة المثوى لتكرار كلمة معينة في النص؛
- البحث عن الشيء ويشمل هذا الأمر إيجاد كلمة معينة أو معرفة معانيها؛
- تصنيف الكلمات أي ترتيبها وتبويبها في فئات نحوية كالأسماء والأفعال والصفات؛
- السياق الذي ترد فيه الكلمة في النص: وتشمل الإحصاء العددي لكلمات النص اللغوي غير معالج الخطوات التالية:
- إحصاء العدد الإجمالي للكلمات في النص؛
- إحصاء مقدار تكرار كلمة بجوار كلمة أخرى وتسمى هذه الخطوة بالتكرار التوافقي للكلمة؛
- إحصاء العدد الإجمالي للجمل في النص.

2/2. الترجمة الآلية:

تعريف الترجمة الآلية: إن المصطلح الترجمة يشير إلى الترجمة بمعناها القديم والشائع أي إلى ما يبتغي أن ندعوه الآن الترجمة البشرية بعد أن ظهر أسلوب جديد في الترجمة وهي الترجمة الآلية وهي من أهم مجالات اللسانيات الحاسوبية وإحدى غاياتها الأساسية وهي التي تفجر معظم مشكلات حوسبة اللغة.

[1] اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، أعمال الندوة الوطنية: 24-25 ديسمبر 2019، الجزء الثالث، ص68.

[2] اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، أعمال الندوة الوطنية: 24-25 ديسمبر 2019، الجزء الثاني، ص34.

جاء في مصطلح الترجمة تحت مادة الترجمة الآلية "وهي الترجمة بواسطة برنامج معلوماتي معدة التحليل النص لتحليل النص المصدر والإنتاج النص والهدف من غير أي تدخل بشري في الترجمة الآلية ويكون المترجم في خدمة الآلة، فيما تكون الآلة في خدمة المترجم حال اللجوء إل الترجمة بمساعدة الحاسوب"[1].

وتعتبر الترجمة الآلية أول تطبيقات بحوث الذكاء الاصطناعي، والمقصود **بالذكاء الاصطناعي**: يطلق على محاكاة ذكاء الإنسان وكيفيّة استخدام خبرته المكتسبة في مجال معين بواسطة الآلة، وخاصة باستخدام أنظمة الحاسبات وتتضمن **عملية المحاكاة**: التّعلم أو الحصول على المعلومات، و قواعد استخدام تلك المعلومات للوصول إلى استنتاجات محددة أو تقريبية وتصحيح الأخطاء ذاتيا، وهي تتطلب معرفة وخبرة في حقلين هما علم الحاسوب وعلم اللغة واللسانيات، وتحاول تطبيقات الذّكاء الاصطناعي أن تجعل الآلات تفعل أشياء تتطلب إذ ما فعله البشر.

إنّ ذلك يتطلب من الحاسوب تحويل المعنى من اللّغة المترجم منها (اللّغة المصدر) إلى اللغة المترجم إليها (لغة الهدف):

3/2. الإحصاء اللغوي: ويمكن أن يكون للجذور اللّغوية والأسماء والأفعال والمشتقات وغير ذلك[2].

4/2. التحليل الصرفي الآلي: إن هذه المهمة تفيد جدا الباحثين والدّارسين في القرآن والمعاجم والكتب اللغوية وذلك في عمليات الإحصاء للجذور ومشتقاتها واستخلاص نتائج دقيقة تصف أداء المؤلف أسلوبيا.

5/2. دراسات المقارنة والتقابلية: هي من مناهج الدراسة اللسانية والاستفادة من القدرة التّخزينية للحواسيب وسرعة المعالجة تفيد جدا في هذا المجال.

6/2. التدقيق الإملائي والنّحوي: وهي من أصعب العمليات في اكسابها للحاسوب إذ لا يمكن الاعتماد الكامل على الحاسوب في التدقيق اللغوي ولا بد أن يراجعه مختص بعد ذلك.

7/2. تعليم العربية للناطقين بغيرها: إن هذا المجال سيفيد كثيرا جدا من نتائج الدراسات التّقابلية بين اللغات والدّراسات المقارنة أيضا مما يجعل تعليم اللّغة أبسط وأسرع[3].

8/2. تحويل النص إلى كلام والكلام إلى نص: تعد هذه العلمية أكثر العمليات نفعا من تطبيقات اللسانيات الحاسوبية لأنها ستكون مجالا للاستخدام من قبل جميع النّاس بينما ينتفع بالعمليات الأخرى بعض الفئات دون بعض.

إذ تهتم اللسانيات الحاسوبية بدراسة قضايا اللسانيات المتعددة مثل رصد الظواهر اللغوية وفقا لمستوياتها، وتبحث في اللغة البشرية كأداة طبيعية لمعالجتها في الآلة ومجالاتها متعددة وقابلة للتوسع، فهي تستخدم تحليل النصوص آليا، الترجمة، التدقيق الإملائي والنحوي، الدراسات المقارنة والتقابلية وتحويل النص إلى كلام والعكس صحيح.

[1] اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، الجزء الثاني، مرجع سابق، ص34- 35.
[2] المرجع نفسه، ص35.
[3] المرجع نفسه، ص36.

المبحث الرابع: وظيفة اللسانيات الحاسوبية وفوائدها.

1. وظيفة اللسانيات الحاسوبية:

تجاوزت علاقة اللغة العربية بالحاسوب الصورة التقليدية التي كان يتم النظر إليه به باعتباره بديلا عن الآلة الكاتبة لكتابة المقالات والكتب، ولكنه بديل راق يمكنك من التحكم في النصوص المنسوخة وتعديلها، ربما تكون هذه هي النظرة الأولى لهذه العلاقة، ولكن بلا شك اتفق الجميع على خطأ هذه النظرية وإن شئت الدقة قبل تخلفها، فإن إمكانيات الحاسوب وما حدث له من تطور في علوم اللغة جعله يقف في مصاف المعجزات التي حققها العلم الحديث[1].

وليس مطلوبا من اللسانيات الحاسوبية أن تنتج جهازا يمكنه الاستغناء التام عن البشر، والتحدث مثله، أو استخدام تقنية اللغة بنفس كفاءة البشر، بل المطلوب أن تكون البرامج الحاسوبية مشابهة لطريقة عمل الإنسان، وفي هذا الإطار يكفي أن تكون هذه اللسانيات قادرة على مساعدة الإنسان، وتخفيف الصعوبات التي يواجهها في مجال بحوثه اللغوية والعلمية، مثلا لو أردت أن أبحث عن كلمة السلام في القرآن الكريم كم مرة وردت؟ وسياقته المختلفة؟ يستغرق هذا عددًا كبيرًا من الساعات، نختصرها في لحظات بواسطة الحاسوب، فنقول: ذكرت السلام خمس مرات في النساء والمائدة والأنعام ويونس والحشر، ونستطيع الآيات مَشْكُولة بخط المصحف إلى ملف **word** بواسطة القص واللصق؛ للاستفادة منها، ودراستها، كما أستطيع الوقوف على معانيها كلها في التفاسير كافة، بمجرد الضغط على زر، وهذا يوفر جهدًا كبيرًا في القيام من مكاني، والبحث عن الجزء من الكتاب الذي فيه الكلمة ثم البحث عن الكلمة نفسها، ونقل الكلام عنها فمعنى آية واحدة يمكن البحث عنه في ساعات بينما لا يأخذ لحظات بواسطة الحاسوب ونظرا لأن الموضوع جديد فما زال الجدل دائرا حول تعريف دقيق اللسانيات الحاسوبية، وقد قدمت جمعية اللسانيات الحاسوبية بتعريف اللغويات الحاسوبية بأنها تركز على الجوانب النظرية لمعالجة اللغة الطبيعية، وتعلم استخدام الآلة ومن ثم إمكانية دراسة الأبحاث اللغوية بها.

2. فوائد اللسانيات الحاسوبية:

لقد فتحت اللّسانيات الحاسوبيّة آفاقا كثيرة للدّرس اللّغوي بحيث ذلّلت له الكثير من الصّعوبات فاعتماده على أجهزة إلكترونية والتي تتمثل في الحواسيب، مكّنه من استيعاب هذا الكمّ الهائل من النصوص، إضافة إلى سرعة المعالجة والاستجابة لأيّ سؤال.

[1] عصام محمود، اللسانيات الحاسوبية العربية، دار الوفاء لدنيا الطباعة والنشر، ط1، 2015، ص38.

وبانتشار المعالجة الآلية للّغات الطبيعيّة انتشارا واسعًا، فقد تهيّأ للّغة دخولها مجالات التّحليل الرياضي والمنطقي والإحصائي، وتهيأ الحاسوب للقائه مع اللّغة بالسرعة الفائقة وسعة الذاكرة، وصغر الحجم، وسهولة النّقل والاستخدام، يُضاف إلى ذلك ظهور لغات البرمجة الحديثة، وأساليب الذّكاء الاصطناعي، وهذا سيسهم في تعليم اللّغة للناطقين بها والنّاطقين بغيرها، وبناء برامج الترجمة الآلية للنصوص، وتطبيقات اكتشاف الأخطاء وتصحيحها، والتعرف الآلي على الكلام المنطوق، وقراءة النّصوص المكتوبة، وضغط النّصوص، وفهرستها، وضبطها بالشّكل[1].

إذ وظيفة اللسانيات الحاسوبية تعلم استخدام الآلة والتمكن من دراسة الأبحاث اللغوية وتركز على الجوانب النظرية لمعالجة اللغة الطبيعية وتكمن برامج الحاسوبية مشابهة لطريقة عمل الإنسان وتخفف عنه الصعوبات التي يواجهها في مجال بحوثه اللغوية والعلمية فمثلا لو أردت معرفة معنى كلمة ما بمجرد ضغط على زر لوحة المفاتيح تظهر النتيجة مباشرة وهذا عكس ما كان سابقا كان عند البحث عن كلمة ما أو تفسيرها يجب الرجوع إلى الكتب ومن ثم البحث عن الجزء من الكتاب الذي فيه الكلمة ثم البحث عن الكلمة نفسها، أي البحث كان يستغرق ساعات وربما أيام ولكن بفضل اللسانيات الحاسوبية أصبح في لحظات قصيرة فقط، وبهذا للسانيات الحاسوبية فوائد كثيرة لأنها فتحت آفاقا كثيرة للدّرس اللغوي فاعتمادها على الأجهزة الإلكترونية والتي تتمثل في الحواسيب مكنت من استيعاب هذا الكم الهائل من النّصوص وسرعة المعالجة والاستجابة لأي سؤال.

[1] اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، أعمال الندوة الوطنية: 24-25 ديسمبر 2019، الجزء الأول، ص108-109.

الفصل الثاني:

المعالجة الآلية الجانب النظري.

المبحث الأول: البداية والتاريخ للمعالجة الآلية.

المبحث الثاني: تعريف المعالجة الآلية.

المبحث الثالث: دور المعالجة الآلية في الحفاظ على اللغة العربية.

المبحث الرابع: واقع المعالجة الآلية للغة العربية.

المبحث الأول: البداية والتاريخ للمعالجة الآلية.

1. أهمية الموضوع:

أصبحت المعالجة الآلية للغة العربية نتيجة حتمية لكثرة النصوص الالكترونية المكتوبة بهذه اللغة مما تطلب وجود نظام حاسوبي عربي يكون فعال من ناحية إدارته للمعلومات وتخزينها في ذاكرة الحاسوب محدودة السعة كما يقوم بالتكشيف عبر كامل النص والاستخلاص الآلي لهذه النصوص[1].

وحتى يتم إنشاء هذا النظام الحاسوبي الذي يسمح بالتكشيف الآلي والاستخلاص واستخراج الكلمات المفتاحية يجب انجاز تحليل سطحي للغة العربية عبر المستويات الثلاث (تحليل مورفولوجي، نحوي ودلالي) وبعد ذلك يتم انجاز نموذج لساني يمكننا من تمثيل كل مكونات اللغة (الفعل، الاسم، الحرف) وأخيرا يتم إنجاز نموذج لوغارتيمي.

واللغة العربية تنتمي إلى مجموعة اللغات السامية التي تتميز بخصائص عديدة تشترك بهامع اللغات الطبيعية كما أنها تتفرد بخصائص خاصة بها، لذا تحتم على الباحثين في هذا الميدان مراعاة هذه الخصائص لنمذجتها (Modélisation) واستثمار تقنية المعالجة الآلية في ذلك.

2. البداية والتاريخ للمعالجة الآلية:

لم تظهر المعالجة الآلية فجأة بل كانت جهودا متواصلة متراكمة حققت تقدما لا بأس به حتى الآن، وإذا ما[2] استعرضنا تاريخ هذه القضية المعقدة سنجده يعود إلى أربعينيات القرن الماضي مع أول ظهور للحاسبات، فساعتها ساد الشعور بالتفاؤل في الأوساط السياسية والعسكرية اتجاه إمكانية استخدام قدراته في التحليلي اللغوي والترجمة الآلية، ففي منتصف الخمسينات اتجهت الجهود العسكرية في الولايات المتحدة الأمريكية إلى استخدام الحاسب في الترجمة الآلية للنصوص من اللغات الأجنبية وخاصة الدوريات العلمية الروسية إلى الإنجليزية، ولأن الحاسبات قد أثبتت قدرتها الفذة على القيام بالعمليات الحسابية بسرعة أكبر ودقة أعلى من الإنسان، كان هناك اعتقاد شائع بأنه لن يمضي وقت طويل قبل أن نتمكن من معالجة اللغة بنفس كفاءة العقل البشري، وأن الأمر سهل للغاية من خلال إنشاء قواميس آلية ثنائية اللغة، وسرعان ما باءت المحاولات الأولى بفشل ذريع، لأن الحاسب أو أنظمة تشغيله أو لغات برمجته وقتها لم تكن قد بلغت من النضج والتطور ما يؤهلها لدراسة وتحليل تعقيد اللغة وغموضها ومجازها، لأن ذلك لا يتطلب فهم معاني الألفاظ فحسب، بل يتطلب كذلك معرفة عميقة بالعالم الخارجي مع القدرة على استغلال هذه المعرفة بشكل ذكي، واعترف العلماء بصعوبة إخضاع اللغات للمعالجة الآلية قبل أن تتوفر القاعدة المعرفية الكافية، وهكذا ولدت علوم ودراسات اللغويات الحاسوبية ومعالجة اللغات بواسطة الحاسب[3].

---

[1] فارش شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي إعرابي للفعل العربي، مذكرة لنيل شهادة الماجستير في علم المكتبات والتوثيق جامعة الجزائر، 2008م، ص06.

[2] شاعر التميمي، المعالجة الآلية للغة العربية، جهود وتحديات، 2010/01/02، د.ت، 2023/02/12 س15.30 د، مجلة ستار تايمز.

[3] المرجع نفسه.

المعالجة الآلية للغة Automatic language processing:

مما لا شك فيه أن الهيئة العليا العربية للمعالجة الآلية الجديرة بأن تمنحها ارتقاء وكفاية ومنطقية على المستويين النظري والتطبيقي إلى جانب دعمها بعوامل القوة والصمود لتقلبات الزمن أسوة بما تم إنجازه في كثير من لغات الأمم الأخرى كالإنجليزية، والفرنسية، والألمانية، واليابانية، والعبرية، والفنلندية والروسية، والسويدية، ولا يخفى على أحد منا ما للسانيات الحاسوبية Computational linguistics من دور فاعل مؤثر في الخروج بنا من أزمتنا الثقافية الطاحنة التي أفقدت الكثير من ثباتهم وتماسكهم، فعلى مستوى التنظير لا يخفى دور استخدام التكنولوجيا الآلية المتقدمة في التعامل مع النصوص في النهوض بأبحاثنا اللغوية بشتى فروعها، وإخراجها من سذاجتها وسطحيتها، أما على المستوى التطبيقي Applied approach فدور اللسانيات الحاسوبية فاعل وشامل، فعلى المستوى التربوي لا يخفى دور الحاسوب إذا ما أحسن استغلاله في تعليم اللغة العربية لأبنائها: صغارًا وكبارًا[1] ولغير أبنائها أيضا Computer based training.

إذا أصبحت المعالجة الآلية للغة العربية نتيجة حتمية لكثرة النصوص الإلكترونية المكتوبة بهذه اللغة، مما تطلب وجود نظام حاسوبي عربي يكون فعّال من ناحية إدارته للمعلومات وتخزينها في ذاكرة الحاسوب، والمعالجة الآلية للغة لم تظهر فجأة بل كانت جهودا متراكمة حققت نقدا لا بأس به، وتاريخ هذه القضية يعود إلى أربعينيات القرن الماضي مع أول ظهور للحاسوب فساعتها ساد الشعور بالتفاؤل في الأوساط السياسية والعسكرية اتجاه إمكانية استخدام القدرات في التحليل اللغوي والترجمة الآلية وكمثال على ذلك اتجهت الجهود العسكرية في الو م أ إلى استخدام الحاسوب في الترجمة الآلية للنصوص من اللغات الأجنبية، ومما لا شك فيه أن الهيئة العليا للغة العربية تمنح ارتقاء وكفاية للمعالجة الآلية، فاللسانيات الحاسوبية لها دور فعال في الخروج بنا من أزمتنا الثقافية الطاحنة، فعلى مستوى التنظير لها دور استخدام التكنولوجيا المتقدمة في التعامل مع النصوص والنهوض بالأبحاث اللغوية بشت فروعها.

[1] سلوى حمادة، المعالجة الآلية للغة العربية معناها ومبناها، جامعة طيبة، المدينة المنورة، مجلة فكر الثقافية، نشر بتاريخ 19/08/16، ت د، 2023/02/12 على 16:49 د.

المبحث الثاني: تعريف المعالجة الآلية.

1. المعالجة الآلية للغة:

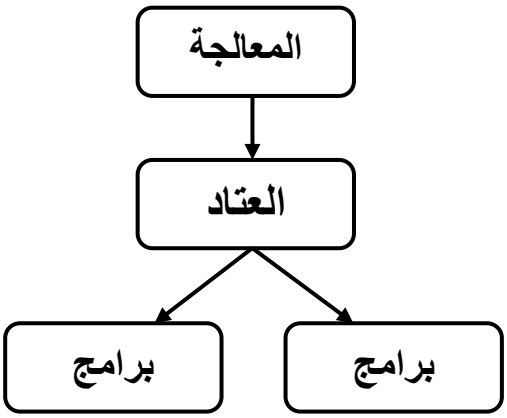
تهتم المعالجة الآلية بدراسة الجوانب الحاسوبية للغة والمشاكل اللسانية والحاسوبية التي تواجه هذه المعالجة سواء أكانت هذه اللغة منطوقة أو مكتوبة وبناء نظام معالجة اللغة العربية مهمة معقدة وصعبة، وذلك لصعوبة إدماج المعارف الصوتية والصرفية والنحوية والدلالية في هذا النظام.

1.1. تعريف بالمعالجة الآلية[1]:

المعالجة: هي التطبيق الآلي على مجموعة من نصوص اللغة وذلك بتغييرها وتحويلها، وإبداع شيء جديد اعتمادا عليها، ويتم ذلك باستعمال تقنيات وأدوات من علوم اللسانيات والإعلام الآلي والنّمذجة، ويجب التّفرقة بين وصف المعارف التي من وظيفة اللسانيات والتعبير عن هذه المعارف في نماذج باستخدام تقنيات واستراتيجيات فعالة مستمدة من علم الحاسوب، وهي وظيفة علم اللّغة الحاسوبي.

الآلية: العمليات الآلية هي التي تجرى عن طريق الآلة، والتي تقابلها العمليات التي تجرى بواسطة الإنسان والحاسوب هو الآلة التي تستعمل في معالجة المعلومات ذات الصّلة بالطبيعة اللسانية؛ حيث أن المعالجة الآلية تتابعه حركات حسابية تقوم الآلة وفق تسلسل زمني أي إن برنامج المعالجة الآلية يمكن أن يكون كليا أو جزئيا حيث إن:

1. **كلي:** يقوم الحاسوب بكل شيء.
2. **جزئي:** يتدخل الإنسان في بعض المراحل.



2.1. مجالات البحث في المعالجة الآلية للغة[2]:

تنقسم مجالات البحث في المعالجة الآلية للغة إلى ثلاثة مجالات أساسية:

الأول: مجال مشترك وتقصد به العتاد اللساني.

الثاني: برامج بحثية.

الثالث: برامج تعليمية.

[1] اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، الجزء الثاني، مرجع سابق، ص28.
[2] المرجع نفسه، ص29-30.

1- البحث العلمي:

- التعرف الآلي للعلاج للكلام العربي؛
- الترجمة الآلية والترجمة المسعفة بالحاسوب؛
- التعرف الآلي للنّصوص.

2- التعليم:

- تصميم برامج تعليمية علت عتاد إلكتروني؛
- تصميم برامج على الإنترنت برامج عالمية.

وإذا خصصنا أي برنامج تعليمي مصمم لأي مرحلة من مراحل التعليم سنجده يحصر خبرته في الجانب التربوي والجانب الحاسوبي.

إذا تهتم المعالجة الآلية بدراسة الجوانب الحاسوبية للغة والمشاكل اللسانية والحاسوبية التي تواجه هذه المعالجة سواء أكانت هذه اللغة منطوقة أم مكتوبة، فالمعالجة هي التطبيق الآلي على مجموعة من نصوص اللغة وذلك بتغييرها وتحويلها، ويتم ذلك باستعمال تقنيات وأدوات من علوم اللسانيات والإعلام الآلي، والآلية هي عبارة عن تلك العمليات الآلية التي تجرى عن طريق آلة وتقابلها العمليات التي تجري بواسطة الإنسان والحاسوب وهو الآلة التي تستعمل في معالجة اللغة فالمعالجة تقوم على جزء كلي وهو عندما يقوم الحاسوب بكل شيء وجزئي عندما يتدخل الإنسان في بعض المراحل.

ومجالات البحث في المعالجة الآلية للغة تشمل برامج بحثية علمية مثل: التعرف الآلي للعلاج للكلام العربي والتعرف الآلي للنصوص وبرامج تعليمية مثل تصميم برامج تعليمية على عتاد إلكتروني وتصميم برامج على الإنترنت.

المبحث الثالث: دور المعالجة الآلية في الحفاظ على اللغة العربية.

دور المعالجة الآلية في الحفاظ على اللغة العربية وتطويرها:

تشكل اللغة العربية تراثا مشتركا للأمة العربيّة، وتعدّ مسألة الحفاظ على هذه اللغة وتطويرها من الضرورات الملحة والأمور الواجبة على جميع أبنائها لأجل النهوض بها لمواكبة التطور العلمي والتكنولوجي الذي يشهده عالم اليوم في شتى مجالات الحياة، ولا يمكن القيام بذلك إلا من خلال معالجتها آليا بأحدث الوسائل والتقنيات ولعلّ من أبرز الطرق للحفاظ عليها وتطويرها:

1. التوثيق الآلي للمعلومات[1]:

ويعرف التوثيق الآلي للمعلومات بأنه أحد أشكال العمل الببلوغرافي الذي يسهل الوصول إلى المعلومة عن طريق ترتيب النصوص وهندستها لغويا؛ بحيث تسّهل على الباحثين والدّارسين الوصول إليها دون جهد أو عناء عند الحاجة إليها.

ويتم خلال عملية التوثيق هذه معالجة النصوص إلكترونيا، وعمل فهارس كاشفة وتوضيحيّة لها بطريقة مشابهة لطريقة عمل الدماغ البشري.

ونظرا لأهمية البالغة للتوثيق الآلي فإنه يتوقع منه معالجة الوثائق والمعلومات بشكل يسهل على الباحثين الوصول إليها، واستدعاءها عن طريق التّجميع والاستخلاص والنّشر، وذلك كالتصنيف الأوتوماتيكي للملفات والمؤلفات تبعا لفهارسها ومؤلفيها أو موضوعاتها، ذلك أنّ مع الثورة التكنولوجية والمعلوماتية الهائلة التي يشهدها عالمنا اليوم، سهل الحاسوب كثيرا من العمليات المتعلقة بتبادل المعلومات وأرشفتها كيفما كان نوعها، ومهما بلغ حجمها.

2. صناعة المعاجم الإلكترونية وتطوير العمل المصطلحي:

يتطلب بناء المعاجم الإلكترونية صياغة المصطلحات وتعميم استعمالها ونشرها وتداولها، إلّا أن هذا الصّنف من المعاجم يجب أن يلمّ بجميع مستويات اللّغة؛ ذلك أنه يعتمد على أدوات مشفّرة، وقاعدة معطيات مُرمزة تخصّ مستويات اللغة؛ نحوًا ودلالة، بحيث تخضع المادة المعجمية فيه لبناء يلم إلمام دقيقا بفروع اللّغة.

ويشترط في المعجم الإلكتروني أن يكون شاملًا وعامًا، لأن البرنامج اللّساني المُعدّ لمعالجة آليا لا ينبغي أن يفشل في العثور على أيّة معلومة مهما كان نوعها لأن أيّ خطأ في المعلومات المُدخلة من شأنه أن يتسّرب إلى باقي مفردات الجملة أو النّص برُمته، ومن ثمة يعرقل عملية اشتغال البرنامج.

كما يشترط في المعلومات أن تكون في المعجم الإلكتروني واضحة وموافقة للمداخل المعجمية المراد معالجتها آليًا، ومن هنا يظهر توظيف واستغلال التّقنيات المعلوماتية في الأعمال التطبيقيّة لصناعة المعاجم المصطلحية باعتماد برامج معيّنة لمعالجتها، والتي أصبح كثيرٌ منها متوافرا ومتاحا للاستعمال على مواقع شبكة الأنترنت[2].

[1] اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، الجزء الثاني، مرجع سابق، ص59.
[2] المرجع نفسه، ص60.

وهي تقدم خدمات بحث مجانية في أهم القواميس والمراجع اللّغوية العربيّة كلسان العرب، ومقاييس اللّغة، والصّحّاح في اللّغة، والقاموس المحيط والعباب الزاخر وغيرها...وتضّم هذه المواقع الإلكترونية ملايين الكلمات المُستقاة من أهم معاجم اللغة العربية.

وتسعى اللسانيات الحاسوبية من خلال المعالجة الآلية للغة إلى إنشاء بنوك للمصطلحات، انطلاقا من عمليات التخزين المُرفقة بمعلومات عن كل مصطلح مفرد، والتي للمترجمين والمحررين والمتعلمين مع دعم الترجمة الآلية وبناء المعاجم المختصّة، وتخوّل البنوك المصطلحية تخزين معطيات دقيقة عن كل مصطلح في ضوء نصّوص موثقة، مع ذكر مقابلاته باللّغات الأخرى وتوضيح مجالات استخدامه، وأساليب توظيفه، مع الإشارة إلى مرجعه.

3. الترجمة الآلية (Automatic traduction):

الترجمة الآلية هي أحد فروع المعالجة الآلية للغة، وهي تهتم بترجمة نصوصّ مكتوبة أو خطابات شفويّة من لغة طبيعيّة إلى لغة أخرى أو أكثر باستعمال الحاسوب، والهدف منها تسهيل المهام أما فئة غير النّاطقين بتلك اللّغة وتعرف التّرجمة الآلية للغة –كذلك– بأنّها: تدّخُل الذكاء الاصطناعي لأداة الترجمة عن طريق الأنماط اللغوية بفعل تراكيب معينة مخزنة مسبقا في ذاكرة الحاسوب يسترجعها عند الحاجة[1].

وتتطلب عملية الترجمة الآلية نقل النصوص والأبحاث والأعمال من لغتها الأصليّة (المصدر) إلى لغات أخرى فرعية (الهدف)، وتعد اللّغة الإنجليزية اللّغة الطبيعية الأولى التي خضعت لهذه العملية، ويتجلى موضوع التّرجمة الآلية في تحليل النّص الأصلي ونقل عناصره من اللّغة التي سيترجم إليها، ثم توليد هذا النّص اعتمادا على التحليل والنقل.

وبهذا القول تكون التّرجمة هدفها بناء نظريّة في النّقل (نقل المحتوى) من اللّغة (المصدر) إلى اللّغة الهدف مع مراعاة خصوصيات النقل التي تفرضها اللّغة المستهدفة ويلعب الحاسوب دورا فاعلا في هذه العمليّة إذ يعمل على تزويد المستخدم بالترجمة المطلوبة إذا كانت مسجلة في ذاكرته، وفي حال غيابها فإنه يقترح عليه مقابلا لها؛ وذلك فإن الأمر يتطلب توفير المصطلحات والتراكيب المتقاربة، نظرا لتعدد معاني ومقاصد ألفاظ في اللغات الطبيعية، وبخاصة في اللغة العربية؛ والتي يشترط في ترجمة ألفاظها وتراكيبها معرفة سياقها الاستعمالي التّداولي، ذلك أنه مهما كانت الترجمة الآلية للنّص فإنه لا بد من تدخل الإنسان لإعادة ضبط وتحرير جمل النّص وكلماته والتعبيرات والمصطلحات الواردة فيه؛ والتي لم يتمكن الحاسوب من ترجمتها بشكل يتناسب مع السياق.

وتبقى الترجمة الآلية من وإلى اللغة العربية من أنجع الوسائل في توظيف المعرفة العلمية والتّقنية في المجتمع العربي؛ لأنّ قدرًا كبيرًا من المعرفة قد أنتج ونُشر وحُفظ باللغة الإنجليزية، وللوصول إليه لابدّ من تفعيل دور التّرجمة الآلية من اللّغات الأخرى إلى العربيّة أو العكس تعتبر من السبل الكفيلة بسدّ الفجوة العلميّة النّاتجة عن تضخم الإنتاج العلمي والثقافي العالمي قياسا بالإنتاج العربي في هذا المجال.

[1] اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، الجزء الثاني، مرجع سابق، ص61.

4. الإنتاج الآلي للنصوص[1]:

الإنسان كائن مفكر ومنتج، وفي الوقت ذاته يُولّد ويطور ويعدل إبداعاته وهو يحتاج في ذلك إلى كثير من الجهد والوقت؛ لذلك كان في أمسّ الحاجة إلى وسيلة تساعده وتجنّبه عناء التّصفح والمراجعة الكثيرة والمتكررة للمعلومات فكان جهاز الحاسوب من خلال محاكاته للذّهن البشري أحسن بديل عن جهد الإنسان في هذا الشّأن، فهو يعتمد على برامج وأنظمة تُمكنه من إعادة تصحيح ومراجعة النّصوص وتخزينها في ملفات خاصّة، تفتح كلما دعت الحاجة إليها وبذلك يكون قد وفر على الإنسان إمكانية الإنتاج النّصي المتعددّة كالمراسلات والدّعوات والتعليقات الصّحفية والإعلانات الإشهاريّة، وغيرها... بمعنى أنّ مثل هذه النّصوص التي كان الإنسان يبذل في سبيل صياغتها وإنتاجها كثيرا من الجهد والوقت صارت اليوم –بفضل تخزينها في ذاكرة الحاسوب–متاحة يمكنه استعمالها في الوقت الذي يريد.

5. تعليم وتعلّم اللغات:

يُعد تعليم وتعلّم اللّغات من أهم مجالات المعالجة الآلية للّغة العربية وبخاصة للمتعلمين المبتدئين والنّاطقين بغير العربية، إذ يعتمد في تحقيق الغايات المرجوة في هذا المجال وبشكل كبير على جهاز الحاسوب؛ باعتباره من أهم الوسائط التي تمزج بين الصّوت والصورة والكتابة، والهدف الأسمى من هذه العملية هو تجاوز الطّرق التعليمية التقليدية المعروفة التي تعتمد بشكل أساسي على أسلوب التّلقين والحفظ؛ في نقل المعرفة وإيجاد طرق جديدة وناجعة تُمكن من استغلال قدرات الحاسب من لدن المتعلمين، الشيء الذي أدى إلى إعداد برامج حاسوبية تعليمية تتماشى مع النّظريات البيداغوجية والتعليمية في عصرنا الحاضر.

هذه إذن أهم تطبيقات اللسانيات الحاسوبية التي تشكل حقلا جديدا في اللسانيات التطبيقية والنظرية؛ يعالج اللغات البشرية وترجمتها عن طريق الحاسوب، وقد أدى تفجر المعلومات في العصر الحالي –نتيجة التطورات السريعة في مجالات العلوم والتكنولوجيا–إلى ضرورة اللجوء إلى وسائل التقنية الحديثة في سبيل الإسراع بعملية نقلها وتناقلها بين الشعوب المختلفة، ولهذه الغاية عمد معظم الباحثين والمتخصصين إلى المعلوماتية إل بناء برامج وأنظمة خاصّة متسولين بالحاسوب والتّقنيات المتقدمة لتحطيم الحواجز اللغوية[2].

المعالجة الآلية لها دور في الحفاظ على اللغة العربية وتطويرها، وهذا من خلال معالجتها آليا بأحدث الوسائل والتقنيات مثل التوثيق الآلي للمعلومات فالتوثيق الآلي يسهل الوصول إلى المعلومة عن طريق ترتيب النصوص وهندستها لغويا، حيث تسهل على الباحثين الوصول إليها دون جهد أو عناء، وتطوير العمل المصطلحي عن طريق صناعة المعاجم فمن شروط المعجم الإلكتروني أن يكون شاملًا وعامًا وأن تكون المعلومات في المعجم الإلكتروني واضحة وموافقة للمداخل المعجمية المراد معالجتها آليا ومن يظهر التوظيف واستغلال المعلومات والتقنيات في الأعمال

[1] اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، الجزء الثاني، مرجع سابق، ص62.
[2] المرجع نفسه، ص63.

التطبيقية لصناعة المعاجم، الترجمة الآلية هي أحد فروع المعالجة الآلية والتي تهتم بترجمة نصوص مكتوبة أو خطابات شفوية من لغة طبيعية إلى أخرى أو أكثر باستعمال الحاسوب، الإنتاج الآلي للنصوص، تعليم وتعلم اللغات.

المبحث الرابع: واقع المعالجة الآلية للغة العربية:

مواقع المعالجة الآلية للغة العربية:

بالرغم من الخدمات الجليلة التي أحدثتها المعالجة الآلية للغة العربية في مختلف المجالات فإنّها ما تزال تعاني جملة من المشكلات والصّعوبات التقنية بسبب خصوصية اللغة العربية على مستوى كافة مستوياتها؛ وبخاصّة في المستوى الدلالي؛ والذي يعدّ من أكثر المستويات اللّغوية تعقيدا وأشدها تعصيّا على جهاز الحاسوب ويعود ذلك إلى كون الدّلالة من أقل المستويات اللّغوية فيما يخصّ التبايّن اللّغوي، كما يشيع فيها عدة ظواهر تخرجها من واقع الاستخدام اللّغوي وحقيقته إلى المجاز كالاستعارة، والكناية، والتشبيه وهو أمر يتطلب تحديد تلك التّعابير غير الحقيقية وتصنيفها دلاليا بما يساعد النّظام الحاسوبي على تمثلها، ومن ثم معالجتها آليا، كما أنّ المعنى يمثل مشكلة كبرى للأنظمة الآلية، فتعدد المعنى للكلمة الواحدة، وحساسية السيّاق في تحديد دلالة الكلمة، واختلاف الدّلالة تنطوي على مفارقات يصعب بسببها تمثيل هذا المستوى أو توصيفه حاسوبيا، وهذا لا يعني إغفال المعالجة الآلية للجانب الدّلالي في اللّغة العربية بحال من الأحوال لأن حضوره موجود ضمن المستويات اللغوية الأخرى، كالمستوى الصوتي والصرفي، والنحوي، والمعجمي وضمن القضايا اللغوية ذات العلاقة بالدّلالة كالترجمة الآلية، وهو ما نلمسه في جهود بعض الدّارسين والباحثين في هذا المجال[1].

الحلول والمقترحات[2]:

أظهرت بعض الدّراسات التي أجريت من طرف عدد من الباحثين في هذا المجال أن المشكلات اللّغوية والتقنية النّاجمة عن هذه المعالجة الآلية للغة العربية كثيرا بالنظر إلى خصوصياتها من حيث ثراء وعراقة تراثها من جهة وأهميّة وضرورة تحديثها لمواكبة مستجدات العصر من جهة ثانية، ومن الحلول المقترحة في هذا الشأن:

– ضرورة تطوير عمل المجامع اللغويّة لمواجهة كافة التحديات، والشروع في البرمجيات لوضع إطار تقانة المعلومات من منظور اللغة العربية وإقامة النماذج اللغوية وتحليل فروعها المختلفة في ميادين الصّرف الحاسوبي والنّحو الحاسوبي والدّلالة الحاسوبية والمعجميّة الحاسوبية وعلم النفس اللغوي الحاسوبي للمواءمة بين المنظومات البرمجية وطبيعة اللغة العربية؛

– ضرورة تضافر كافة الجهود المبذولة والأعمال المنجزة في مجال اللسانيات الحاسوبية[3] العربية من خلال تشكيل فرق عمل كاملة من المختصين اللغويين والخبراء الحاسوبيين في أي مشروع علمي يهدف إلى برمجة الأنظمة اللغوية للغة العربية وتحليلها ومعالجتها آليا؛

[1] اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، الجزء الثاني، مرجع سابق، ص64.
[2] المرجع نفسه، ص64.
[3] المرجع نفسه، ص65.

– ترجمة جميع الأعمال والدراسات العلمية في مجال اللسانيات الحاسوبية العربية من اللغات الأجنبية إلى العربية بغرض استثمارها والاستفادة منها؛

– توحيد المصطلحات في هذا المجال من خلال إعداد وصناعة معجم موّحد للمصطلحات اللسانية الحاسوبيّة بالعربية والإنجليزية، وفق القواعد المتعارف عليها علميا في هذا النّوع من المعاجم المصطلحية؛

– الاهتمام بالبرامج الأكاديمية بالجامعات، وبخاصة في أقسام اللّغة العربية في مجالات حوسبة اللغة العربية وحوسبة المعاجم العربية والنّهوض بتطوير مناهج اللغة العربية؛

– الاهتمام بوضع سياسات لغوية وطنية وتوحيد الجهود بين الباحثين في المجال اللغوي والحاسوبي للإفادة من الإنترنت وإعداد مواقع لتعليم اللغة العربية؛

– ضرورة نشر الرسائل العلمية والأبحاث الجامعية التي كتبت باللغة العربية عن قضايا استخدام اللغة العربية في الحاسب الآلي.

– ضرورة تفعيل النشر الإلكتروني باللغة العربية[1].

إذ بالرغم من الخدمات التي أحدثتها المعالجة الآلية للغة العربية في مختلف المجالات إلا أنها مازالت تعاني جملة من المشكلات والصعوبات التقنية خصوصا على المستوى الدلالي لأنه يعد من أكثر المستويات اللغوية تعقيدا وأشدها تعصبا على جهاز الحاسوب، فتعدد المعنى للكلمة الواحدة، وحساسية السياق في تحديد دلالة الكلمة صّعب تمثيل هذا المستوى وتوصيفه حاسوبيا... وبهذا أجريت بعض الدراسات من طرف عدد من الباحثين لعل المشكلات اللغوية الناجمة عن المعالجة الآلية للغة العربية منها: ضرورة تطوير عمل المجامع اللغوية لمواجهة كافة التحديات، ترجمة جميع الأعمال والدراسات العلمية في مجال اللسانيات الحاسوبية، الاهتمام بالبرامج الأكاديمية في الجامعات، ضرورة نشر الرسائل العلمية والأبحاث الجامعية، وضرورة تفعيل النشر الالكتروني.

[1] ينظر: اللغة العربية بين اللسانيات الرتابية الحاسوبية واللسانيات العرفانية في الجامعات الجزائرية، الجزء الثاني، مرجع سابق، ص66.

الفصل الثالث:

المعالجة الآلية للغة العربية.

المبحث الأول: المستوى المورفولوجي.

- بنية الكلمة.
- الاشتقاق.
- المعالجة الآلية للمستوى المورفولوجي.

المبحث الثاني: المستوى النحوي.

- من الكلمة إلى الجملة.
- تمثيل الجملة.
- التحليل العام للجملة.
- المعالجة الآلية للمستوى النحوي.

المبحث الثالث: المستوى الدلالي.

- تمثيل المعاني.
- من القاموس إلى خريطة المفاهيم.

المبحث الرابع: نحو إنشاء قاعدة معجمية آلية في شكل معجم كوني هرمي.

- **لماذا المعاجم الآلية.**
- منهجية إنشاء معجم يبدأ من مفهوم الكون في تسلسل هرمي.
- خطوات إنشاء قاعدة البيانات المعجمية الآلية.

المبحث الأول: المستوى المورفولوجي.

منذ ظهور الحاسوب في أواخر الأربعينيات وصلته باللغة تتوثق وتتأصل؛ فاللغة تقع في قمة الموضوعات التي تهتم بها العلوم الإنسانية. والحاسوب هو ذروة التقنيات الحديثة، لذلك كان من المنطقي بل من الحتمي أيضا أن تلتقي اللغة والحاسوب، وذلك لسبب رئيسي وبسيط، وهو كون اللغة تجسيدا لنشاط الإنساني الذهني في الوقت نفسه الذي يتجه فيه الحاسوب نحو محاكاة بعض وظائف الإنسان وقدراته الذهنية.

واتجه أخصائيو الحاسوب واللسانيين إلى معالجة مستويات اللغة (المستوى المورفولوجي النحوي والدلالي) كل على حدا وذلك اعتمادا على نماذج لغوية محددة ودقيقة، حتى ينتج لدينا برمجيات تهدف إلى معالجة اللغة سطحيا.

1- المستوى المورفولوجي[1]:

الدراسة المورفولوجية هي الدراسة الشكلية للوحدات التي لها معنى والتي تشكل الكلمات فمعلوم أن هذه الأخيرة تتكون من وحدات دالة صغيرة تعرف بالمورفيمات (Morphémes)، فكلمة [يذهب] تتكون من مورفيمين.

ذهب: القاعدة.

ي: السابقة وهي محل إعراب المذكر الغائب في الماضي.

وتحليل الكلمة إلى وحداتها الدالة الصغيرة يعطي لها نوعين من المورفيمات:

المورفيم المعجمي [Lexéme]: وهو جذر الكلمة منزوع منه علامات التصريف والحالة نجده داخل المعجم.

المورفيم الإعرابي: هي جميع التحويلات التي تطرأ على الجذر حتى تميز أشكال التصريف، الشخص، الزمن والنوع للفعل، أما الاسم فهي النوع، الحالة والعدد.

1. بنية الكلمة:

كل كلمة لديها في شكلها الخطي ثلاث مستويات مستوى مورفولوجي، مستوى نحوي، ومستوى دلالي، والمستوى المورفولوجي هو المستوى الذي يظهر في الشكل الخطي للكلمة (Graphique) ويعتبر المدخل إلى الوحدات الأخرى النحوية والدلالية.

إن التحليل المورفولوجي لوحدة لغوية –كلمة– ينتج عنها المعلومات اللغوية المورفولوجية التالية:

1.1. الجذر Radical: هو الذي يحمل المعنى الأساسي للكلمة، وتوجد عدّة مصطلحات متقاربة المعنى وتؤدي نفس الغرض، القاعدة (Base)، الجذر (Radical)، الجذع (Racine).

القاعدة: هي ما يتبقى من الكلمة بعد حذف الزوائد (Affixes) الإعرابية.

الجذع: وهي وحدة خيالية تحوي على المعنى المشترك لكل الكلمات التي تتكون من هذا الجذع.

الجذر: وهو ما يتبقى من الكلمة بعد حذف الزوائد.

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص23.

2.1. السابقة (Préfixe): هي كل الزوائد التي تلحق بالجذر في أوله.

3.1. اللاحقة (Suffixe): هي الزوائد التي تضاف آخر الجذر.

4.1. الواسطة (Infixe): هي زائدة تضاف إلى وسط الجذر وتتبع القواعد الفونولوجية ومثلها الألف في الفعل تضارب.

2. الاشتقاق:

من أهم خصائص اللغات السامية والهندو-أوروبية خاصية الاشتقاق التي تعرف بأنها: "عملية تتمثل في إنتاج كلمات جديدة وذلك بتغيير المورفيمات مقارنة بالجذر وإضافة السوابق واللواحق"[1].

كما يعرف الاشتقاق بأنه: "عملية زيادة حروف الجذر وتغيير بعض حروفه للحصول على وحدات معجمية جديدة".

وتتمثل عملية الاشتقاق إما بإضافة حروف للقاعدة الاشتقاقية وهي الجذر الذي يضاف إليه الزوائد، وهنا يسمى اشتقاقا بالزوائد (Dérivation affixale) ويكون على ثلاثة أنواع:

Courage	→	Courageuse	الاسمي
Laver	→	Lavage	الفعلي
Aride	→	Aridité	النعتي

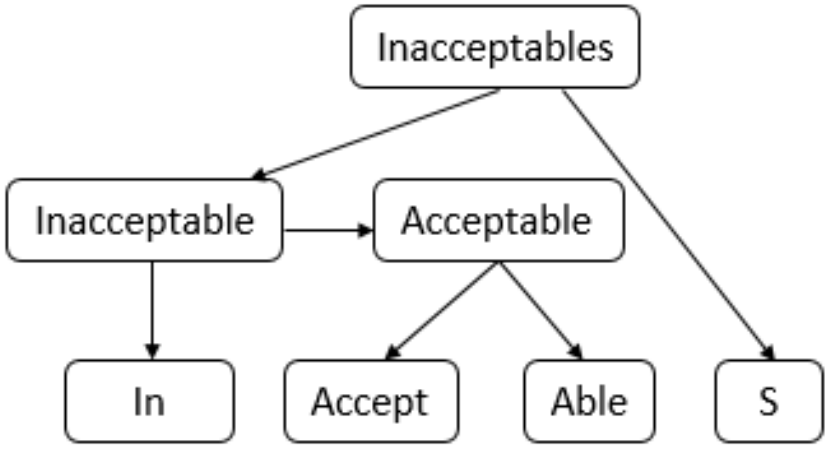
أما النوع الثاني من الاشتقاق فهو الاشتقاق المعكوس Dérivation conversionelle وهو لا يغير في بنية الكلمة -مورفولوجيا- إنما يغير وظيفتها النحوية ومثال ذلك الفعل: Boire – manger في الجملة التالية[2]:

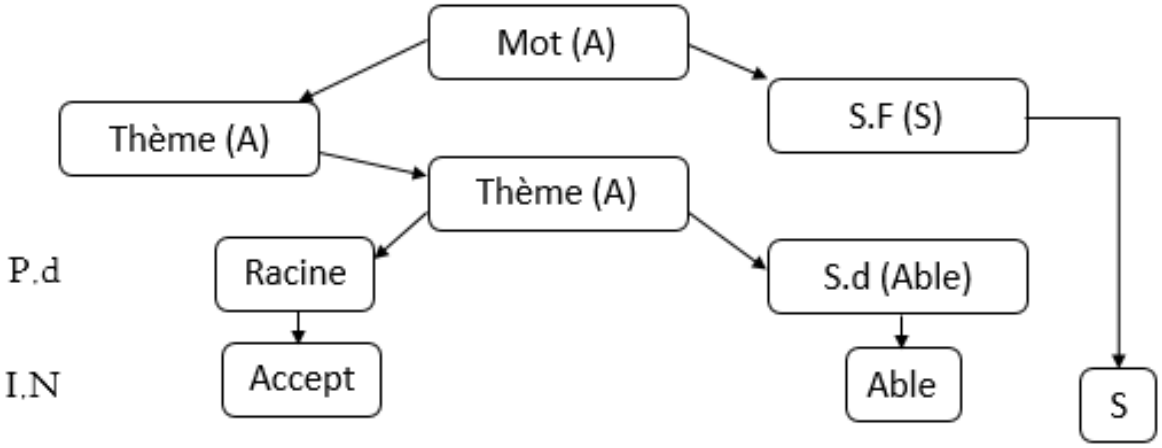
Boire	→ Boire	
[Verbe]		[Substantif]
Manger	→ Manger	

وخاصية الاشتقاق ليست عامة لكل كلمات اللغة فقد قسم اللغويين كلمات اللغة حسب ميزة الاشتقاق إلى نوعين:

1.2. المنتجة Productive: وهي التي يمكن الاشتقاق منها: بإضافة السوابق التي تغير معنى القاعدة الاشتقاقية واللواحق التي لا تغير معناها.

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص24.
[2] المرجع نفسه، ص25.





2.2. الجامدة: هي الجذور التي لا يمكن اشتقاق كلمات جديدة منها ولم تأخذ من غيرها ودلت على حدث أو معنى من غير ملاحظة صفة كرجل وشجر وبقر في اللغة العربية والفأرة Souris، Clavier في اللغة الفرنسية.

وعملية الاشتقاق تنتج عنها مداخل معجمية متقاربة المعنى، وهذا ما أدى إلى إنجاز قواميس لتعلم اللغات باستخدام هذه الخاصية مثل قاموس DAFAواستخدامها في محركات البحث في الإنترنت، ومثال ذلك Google فيمكن أن تدخل في أيقونة البحث، فالجذر علم* ونضيف نجمة فيقوم المحرك بالبحث عن كل الكلمات التي بها [علم] ويضيف إليها جميع اللواحق الممكنة.

3.2. محلات الإعراب: هي تغيرات شكلية تطرأ على جذر الكلمة بإضافة حروف أو إنقاص يعضها أو تغيير الجذر وذلك حسب موقع الكلمة في الجملة.

ويأخذ الاسم محلات الإعراب التالية: الجمع، الجنس والحالة (Cas) وهذه الأخيرة خاصة باللغات الجرمانية. أما القعل فمحلات إعرابه هي: الأصل (Mode)*، الزمن، الشخص[1].

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص25.

3. المعالجة الآلية للمستوى المورفولوجي:

1.3. القائمة المعجمية:

1.1.3. نموذج القاعدة: يعمل هذا النموذج على إعطاء كل كلمة جذرها ومشتقاتها المختلفة ولكل شكل جميع المعلومات المورفولوجية الإعرابية، وتعرض الكلمات مرتبة ألفبائيا[1].

مثال: AMI

Forme	Base	Cat	Flexion
Ami	Ami	N	MS
Amif	Ami	N	PS
Amis	Ami	N	MP
Amies	Ami	N	FP

الجدول رقم (02): قائمة معجمية.

MS : المذكر المفرد.

MP : الجمع المذكر.

FS : المؤنث المفرد[2].

Fp : الجمع المؤنث.

ويسمح هذا النموذج بإيجاد لكل كلمة الجذر وحالتها ومحلها من الإعراب.

(Ami, mp) ← Amis

كما يقدم لكل جذر ومحله من الإعراب الكلمة المناسبة.

(Ami, ms) ← Ami

المزايا: التحليل الصرفي بمساعدة القوائم المعجمية سهل من ناحية الإعلام الآلي تطبيقه حيث توجد عدّة تقنيات وبرمجيات في اللغات الأوروبية لإنشاء هذه القوائم، ويتطلب ذلك حواسيب ذات ذاكرة كبيرة لأن المعجم يكون ثقيل، حيث كل مدخل معجمي نحدد له الجذر الممكن له مع تحديد جميع المشتقات التي ترتبط به وتحديد جميع المعلومات المورفولوجية والإعرابية لكل كلمة فجذر الفعل كتب سوف نحدد له المشتقات المرتبطة به وعلاماتها الإعرابية.

العيوب: أول مشاكل التي تواجه هذه التقنية هي ثبوتها (Standard) فلا يمكنها إنشاء أشكال المداخل الجديدة، كما أن الأفعال لديها عدّة أشكال وهذا ما يجعل المعجم كبير جدا ومكلف من الناحية المالية، فقاموس اللغة

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص26.
[2] المرجع نفسه، ص27.

الفرنسية التي يعرف بـ (BDLEX*) يحتوي على ما يقارب 25000 مدخل (Lexéme) تطلب إنجازه سنة كاملة وبتعاون 100 عالم متعددي التخصصات: علم الحاسوب، اللسانيات، المعاجم، علم النفس المعرفي.

أما العيب الثاني يتمثل في كون القوائم المعجمية لا تعالج مشاكل الاشتقاق الذي به يمكن إنشاء أشكال غير محدودة من جذر واحد، كما أن الكلمات المركبة ليس لها مكان في هذه القوائم[1].

2.1.3. القائمة السهمية (Formes fléchies):

حتى يتمكن الحاسوب من تكوين القائمة المعجمية يلجأ إلى جدولين.

جدول الجذور Bases: الذي يحتوي على الجذور جميعها وجميع حالاتها الإعرابية ودليل قد يكون حرف أو رقم.

جدول النماذج Modéles: الذي يربط بكل جذر جميع العمليات التي تطرأ عليه للحصول على أشكال جديدة.

Table de base

Base	Cat	Modèle
Ami	N	01

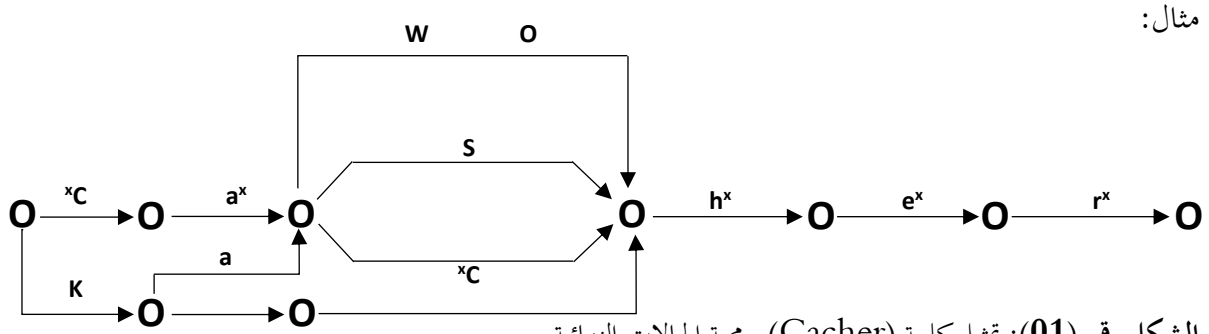
Modéles	Flexions	Enlever	Ajouter
01	MS	0	O
01	MP	0	S
01	FS	0	E
01	FP	0	ES

الجدول رقم (03) يبين قائمة السهمية لكلمة (Ami).

2.3. برمجيات الحالة النهائية:

هي برمجيات آلية صغيرة تعمل علة نمذجة تتابع الحروف وتركيب الكلمات.

مثال:



الشكل رقم (01): تمثيل كلمة (Cacher) برمجية الحالات النهائية.

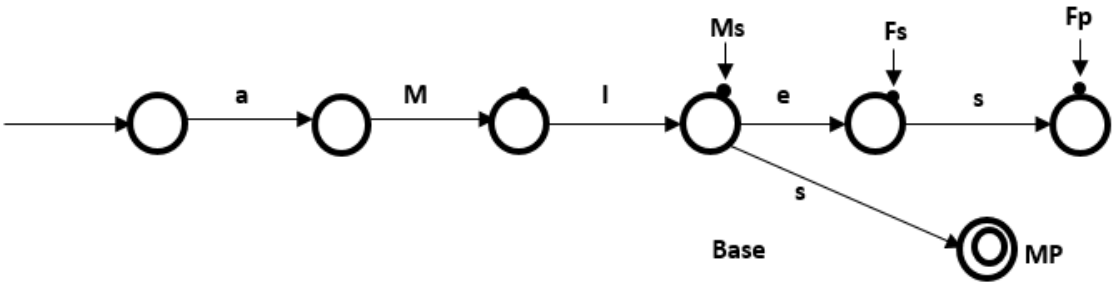
[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص28.

في هذا الشكل تستطيع البرمجية استخراج كلمة (Cacher) وذلك باستعمال نظام المطابقة كما يمكنها أن تشكل كلمات جديدة، حتى وإن كانت غير مقبولة من جهة اللغة.

1.2.3. تطبيق برمجيات الحالة النهائية على محلات الإعراب[1]:

مجموعة كلمات المعجم (Lexique) يمكن تمثيلها ببرمجية ذات حالة نهائية حيث كل خط (arc) يمثل حرف، كما تظهر في المثال التالي:

يمكن للبرمجية أن تحوي معلومات مورفولوجية لاستخدامها في المعالجة الضرورية كالتعميم (Génération) والتحليل (Analyse) ويكون ذلك بعدة طرق، ويبين الشكل الموالي طريقة الأسهم حيث أضيف إلى البرمجية السابقة علامات (étiquettes) للدلالة على محلات الإعراب كما أضيفت علامة تبين المعلومات المورفولوجية الجذر.



الشكل رقم (02): تطبيق البرمجيات على محلات الإعراب.

MS : المذكر المفرد.

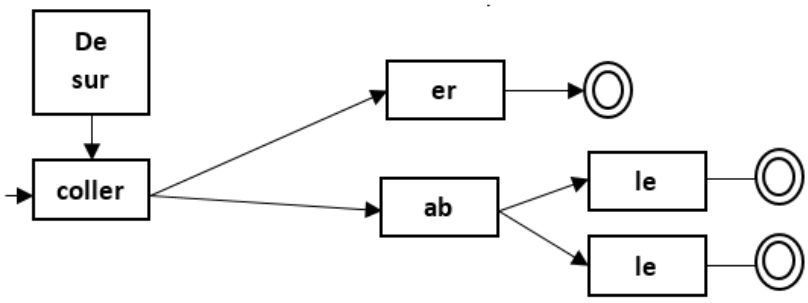
MP : المذكر الجمع.

FS : المؤنث المفرد.

FP : المؤنث الجمع.

2.2.3. التطبيق على الاشتقاق:

تستطيع البرمجية في هذه الحالة تشكيل كلمات جديدة اعتمادا على الجذر وإضافة اللواحق.



[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص29.

الشكل رقم (03): تطبيق البرمجيات على الاشتقاق.

في هذا المثال تستطيع البرمجية تشكيل الكلمات التالية: Recoller, recollage, décoller

المبحث الثاني: المستوى النحوي.

المستوى النحوي[1] له أهمية كبيرة في الدراسات اللسانية، لأن معرفة المركبات اللغوية التي يتألف منها التركيب اللغوي –الذي يشتمل جملة مفهومة أساسية أو مشتقة– هي أمر مهم – والمركبات اللغوية في الجملة وما ينتج عنها من دلالات مختلفة مهمة كذلك إذن إن معرفة البنية النحوية التي تفرزها اللغة العربية بهذه اللغة العربية لهذه البنية يسهل عملية التعلم والتعليم والتوصيل.

كما يعمل هذا المستوى على معرفة التراكيب اللغوية التي يتألف منها النص لأن هذا الأخير هو عبارة عن وحدة لسانية قائمة بذاتها تتشكل من ضوابط لسانية تؤلف أجزاء هذه الوحدة اللسانية.

والمستوى[2] النحوي هو العلم الذي يبحث في التراكيب حيث يطلق عليه اسم (علم التركيب) ويدرس هذا العلم تنظيم الكلمات في شكل مجموعات.

فالمستوى النحوي[3] لا يكتفي بالقواعد الصرفية بل يحتاج إلى دراسة بنية الجملة اللغوية وأنماطها والعلاقات بين الكلمات وآثارها والقواعد التي تحكم تلك العلاقات، فيدرس هذا المستوى دور كل جزء في بناء الجملة، وعلاقة الأجزاء ببعضها البعض، وأثر كل جزء في الآخر.

الدراسة النحوية للغة تعني دراسة طرق تركسي الكلمات في جملة، حيث أن الكلمات تأخذ أشكال وتضاف إليها علامات عند جملة، يهتم علم النحو بهذه الظاهرة ويعمل على استخراج هذه العناصر اللسانية المكونة للجمل.

وعلى المستوى النحوي يظهر مصطلح الإعراب Grammaire الذي يأخذ ثلاث مفاهيم هي: أثر ظاهر أو مقدر يجلبه العامل في آخر الكلمة ويعني حالات الرفع والنصب والجر والجزم الناجمة عن عامل.

عدم لزوم آخر الكلمة حالة واحدة وتأثرها بالعوامل أي اختلاف آخر الكلمة باختلاف العوامل وهو يقابل البناء وتقسم الكلمات بحسبه إلى مبنية ومعربة.

الوظيفة النحوية للكلمة في الجملة أو للجملة في النص كأن تكون فاعلا أو مفعولا به أو حالا[4].

1. من الكلمة إلى الجملة:

1.1. التركيب:

في اللسانيات التحليلية عندما نريد معرفة تركيب الجمل من الأسفل [الكلمة] إلى الأعلى [الجملة] فإننا نجدها تتكون من مكون (*Syntagme) ويمكن أن يكون اسمي، فعلي، وصفي، اعتراضي (Préposition).

---

[1] صفية مطهري، التفاعل الدلالي بين المستويات اللسانية، التراث العربي، مجلة فصيلة محكمة، اتحاد الكتاب العرب بدمشق (العدد 112)، 2008، ص269.

[2] عيسى برهومة، مقدمة في اللسانيات، مرجع سابق، ص137.

[3] ينظر: محمود عكاشة، التحليل اللغوي في ضوء علم الدلالة، دار النشر للجامعات، القاهرة، ط1، ص14.

[4] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص30-31.

أما العناصر الأخرى التي نجدها في الجملة هي: المخصصات Spécifieurs (مثل أداة التعريف أل)، المميزات Qualificateurs (مثل النعت، الحال)، المتممات Compléments (المفعول به).

2.1. الجملة: هي مفهومها التقليدي هي جملة مفيدة ذات تركيب مكثف بنفسه وتامة الإفادة وهي مؤلفة من كلمتين أو أكثر فإذا كانت الكلمة الأولى اسما اعتبرت جملة اسمية، أما إذا كانت فعلا فتعتبر جملة فعلية.

ولدراسة البنية السطحية للجملة اقترح **مارتيني** (Andres Martenet) ثلاث مقاييس لذلك وهي:

- مفهوم الاستقلال التركيبي.
- الوحدات الوظيفية.
- موقع الكلمة في التركيب.

1.2.1. الاستقلال التركيبي:

أعطى مارتيني نوعا معينا من التراكيب مكانة في نظريته حيث جعله النواة الأساسية للجملة وأقل ما يكون عليه الكلام وأقل ما يفيد ويتمثل في التركيب الإسنادي الذي يتكون من المكون الأساسي (النواة) والمكمل (المحكوم عليه)، فنجد في الجملة: Les enfants (du voisin) jouent (dans le jardin).

التركيب المكون من Les enfants jouent هو النواة الذي يشكل أساس الجملة ولا يمكن أن يزول فإذا زال فسدت الجملة وهو تركيب مستقل لأنه يدل بنفسه على وظيفته، أما بقية العناصر فمتعلقة به، تضاف لتحديد الزمان والمكان.

2.2.1. الوحدات الوظيفية: هي الوحدات التي تحدد وظيفة الوحدات الأخرى داخل التركيب مثل: dans، du في المثال السابق الذكر.

3.2.1. الوظيفة الإعرابية:

الوظيفة الإعرابية هي الوظيفة النحوية التي تأخذها الكلمة بناء على موقعها في الجملة ونوعها، حيث أن كلمات اللغة العربية تقسم إلى ثلاث مجموعات كبرى هي: الاسم، الفعل، الحرف، ولكل مجموعة من هذه المجموعات وظائف إعرابية خاصة بها تحدد بعلامات تعرف بأنها حركات صرفية تتمثل في إضافة حركات وحروف إلى جذر الكلمة لتشكل المجموعات الإعرابية (الجنس، العدد، النوع)، وكذلك وظائف نحوية كالفاعل والمفعول به بالنسبة للاسم، والفعل ماضي أو المضارع أو الأمر[1].

2. تمثيل الجملة:

هناك عدّة تقنيات استخدمت لتمثيل الجملة نذكر منها:

- تقنية العلبة.
- التمثيل الشجري.

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص32.

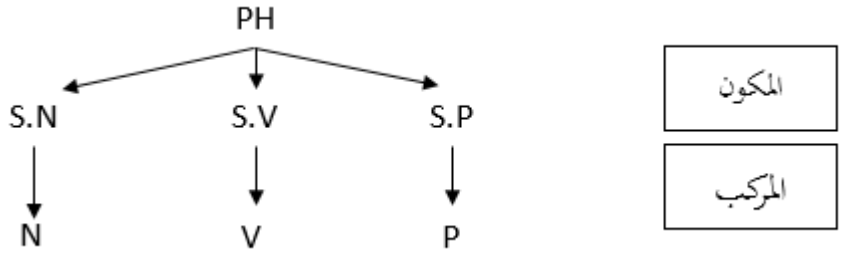
1.2. تقنية العلبة: تدرج الجملة في مستطيل كبير، ثم تقسم إلى وحداتها الجزئية بتقسيم المستطيل الكبير وذلك على مراحل حتى الوصول إلى مكوناتها الجزئية.

مثال: La fillette regardait le chat

La fillette regardait le chat						
La fillette		Regardait		Le chat		
La	Fillette	Regardait		Le		Chat
La	fill	ette	Regard	ait	Le	Chat

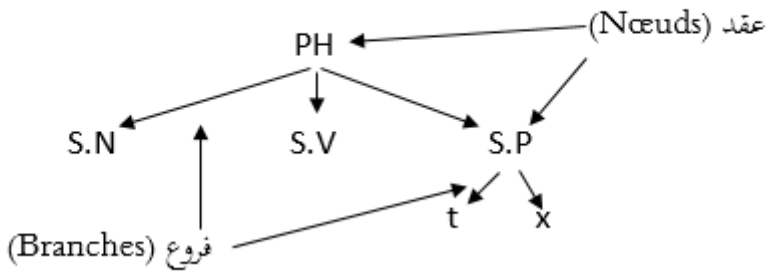
الجدول رقم 04: يمثل تقنية العلبة.

2.2. التمثيل الشجري: يسمى كذلك تمثيل المحددات النحوية حيث يعتمد على تقسيم الجملة إلى وحداتها الجزئية الأولى والتي تسمى المكونات (Les syntagmes) وتقسم هذه الأخيرة إلى وحدات أصغر تعرف بالمركبات Les constituants[1].



الشكل رقم (04): التمثيل الشجري للجملة.

وتتكون كل شجرة من عقد وفروع والشكل التالي يوضح ذلك:

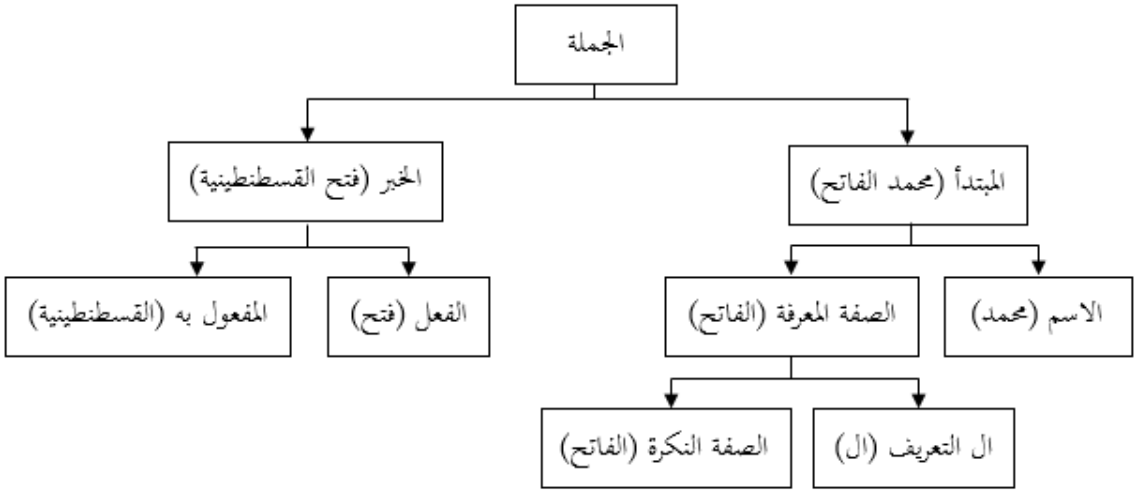


الشكل يبين عقد وفروع الشجرة.

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص33.

التحليل العام للجملة:

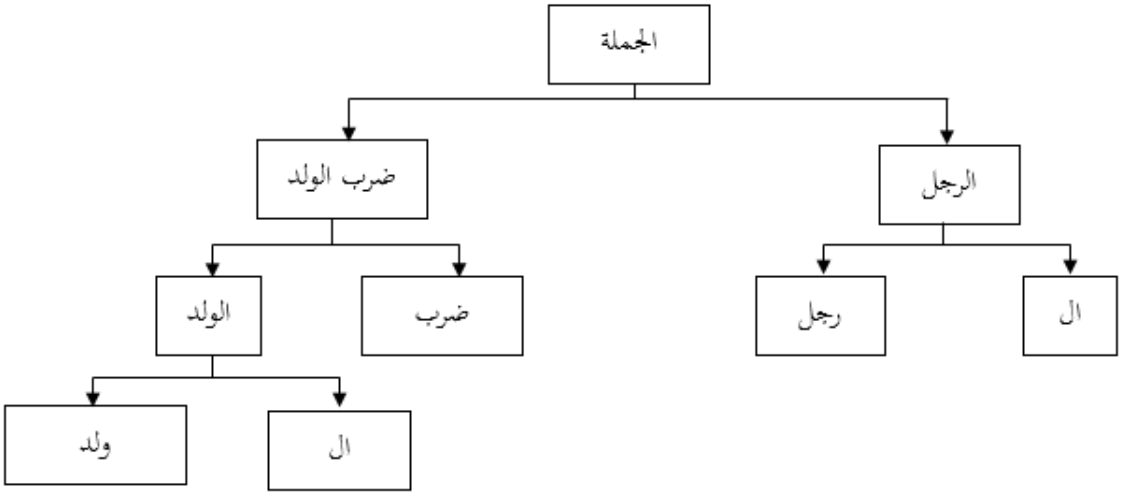
هي شكل لغوي مستقل[1]، أو أنها تتألف من مسند ومسند إليه مثلا نقوم بتعديل هذا المفهوم بناء على نتائج البحث والدراسة، أي أن يكون التعريف الفعلي للجملة هو نتيجة البحث والدراسة، أي أن يكون التعريف الفعلي للجملة هو نتيجة البحث كله. لا مقدمة له ... وهذا ما نراه في نتائج المدرسة الوصفية التشكيلية التي كان رائدها **بلومفيد** في أمريكا، فهذه المدرسة اتبعت منهج رئيسي لتحليل الجملة وتمثيلها، لقد اتبعوا ما يدعى بالتحليل إلى المكونات المباشرة (Immédiate constituent analysis) الذي يمكن بموجبه تحليل الجملة ليس على أساس أنها مؤلفة من كلمات مرصوفة بعضها بجانب بعض أفقيا[2].



هكذا:

محمد / الفاتح / فتح / القسطنطينية.

بل على أساس أنها مؤلفة من طبقات من مكونات الجملة بعضها أكبر من بعض إلى أن يتم تحليلها إلى عناصرها الأولية من الكلمات وحتى مورفيمات.



[1] ينظر: نايف خرما، أضواء على الدراسات اللغوية المعاصرة، المجلس الوطني للثقافة والفنون والآداب، الكويت، سبتمبر1978، سلسلة عالم المعرفة، العدد 9، ص234.

[2] نايف خرما، أضواء على الدراسات اللغوية المعاصرة، مرجع سابق، ص236.

القاعدة رقم 1: الجملة مبتدأ + خبر.

القاعدة رقم 2: المبتدأ: ال + اسم.

القاعدة رقم 3: الخبر فعل + اسم.

القاعدة رقم 4: أل آل التعريف.

المفردات:

الاسم: رجل، ولد.

الفعل: ضرب.

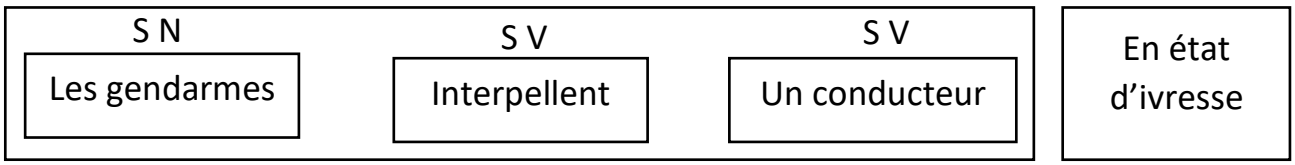
←هذه الميزة ترتب وتنظم الخطوات التحليلية وتربطها معا باعتماد كل خطوة منها على خطوة سابقة لها[1].

3.2. المعالجة الآلية للنحو:

1.3.2. التحليل السطحي: يهدف هذا التحليلي إلى معرفة المكونات (Les syntagmes) النحوية للجملة دون معرفة العلاقات التي ترتبط بها ومختلف علاقات التكرار والترابط.

الهدف من هذا التحليل هو الحصول على نتائج قليلة الغنى عن مكونات الجملة ولكن بسرعة كبيرة ومؤكدة[2].

وهذا التحليل ليس له فائدة كبيرة من حيث التطبيقات الآلية للمعالجة النحوية مثل الترجمة والتلخيص الآلي وفهم الجمل دلاليا، لكن فائدته تكمن في تقطيع الجمل ومعرفة مكوناتها ثم استخراج الواصفات منها.



2.3.2. الإعراب خارج السياق (Les grammaire hors-contexte):

الإعراب خارج السياق يتكون من مجموعة قواعد على الشكل التالي: $E \rightarrow E_1 \ldots E_N$ التي تعني أن عدد غير محدد من المعاني ($E_1 \ldots E_N$) يمكن كتابتها بكلمة واحدة وتعوض جميع هذه المعاني، وأول من تحدث عن هذه النظرية هو العالم اللغوي **تشومسكي** (Chaumescky) حيث حاول نمذجة اللغة الإنجليزية وخاصة الناحية النحوية التركيبية منها والتعبير عما يشاء بقواعد محدودة لذا اقترح قواعد إعادة كتابة الجملة وهذه القواعد هي:

PH→(SP) SN SV (SP)

SN→(det) N (SP) (SA)
SN→(SN) (SP) : Complément du nom
SV→(aux) V (SN) (SP) (SA)
SP→ Prép SN
SA→ S (Adi) A (SP)

[1] نايف خرما، أضواء على الدراسات اللغوية المعاصرة، مرجع سابق، ص246.
[2] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص34.

S (adm) → S (adm) adv
N→ Pro

*العناصر التي بين قوسين اختيارية.

وطبق Tanier هذه القواعد على الجملة البسيطة في المثال التالي وكان التحليل كما هو مبين في الشكل الموالي:

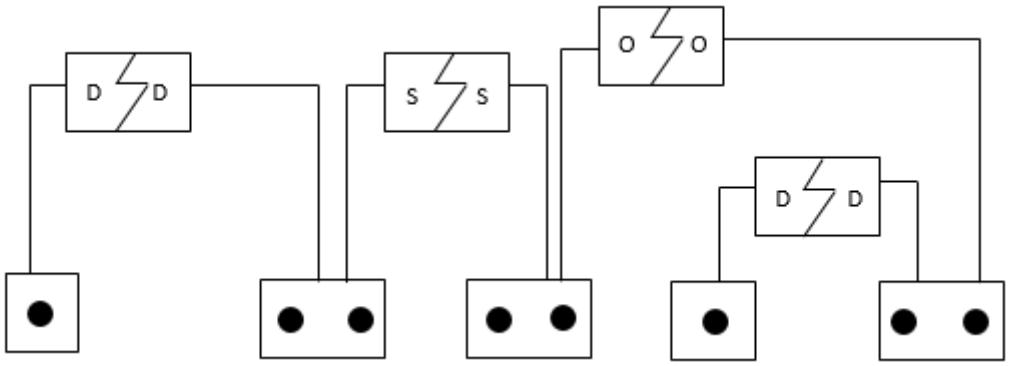
	Règles	Exemples
PH	SN SV	[Le loup] $_{SN}$ [Sort de la fort] $_{SV}$
SN	Pronom	
SN	Nom propre	Paul
SN	Det adj nom adj	[Le] $_{dit}$ [petit] $_{adj}$ [Chaperon]$_{nom}$ [Rouge]$_{adj}$
SN	SN SP	[La fille]$_{SN}$ [Minos et de Pasiphaé]$_{SP}$
SV	Verbe SP	[Sort]$_{Verbe}$ [De la forte]$_{SP}$

ورغم أهمية هذا النموذج في التحليل النحوي إلّا أنه يقدم معلومات على بنية الجملة ظاهريا ولا يتعمق في البنية الداخلية لها حيث لا يقدم طبيعة العلاقات الموجودة بين هذه المكونات[1].

3.3.2. التحليل بالمتابعة (L'analyse en dépendance):

يفتلق هذا التحليل عن التحليل التركيبي (مثل قواعد خارج السياق) في طريقة تمثيل المعلومات، وتكمن فكرة هذا التحليل في ربط كلمات الجملة مع بعضها حيث كل كلمة نحدد مكوناتها (أدوات التعريف، ضمائر الربط).

وأهم مثال عن هذا التحليل هو القواعد السهمية الموضحة في الشكل التالي ونذكر أن هذه الأسهم يجب أن لا تتصالب.



الشكل رقم (05) يبين التحليل النحوي باستخدام القواعد السهمية.

4.3.2. تقنية العلامات (Les traits):

إن التقنيات الثلاث السابقة لا تستطيع تحليل اللغة الطبيعية بطريقة فعالة حيث توجد قواعد وظيفية بين الكلمات المكونة للجملة لا تحللها هذه التقنيات مثل الأفراد والجمع، التأنيث والتذكير. حيث أن الكلمات التالية

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص35.

Les loup، La fille غير مقبولة نحويا، لكن هذه التقنيات توفر معلومات عن كيفية تحليل الجملة وليس معلومات نحوية عن الجملة تمكن الآلة من فهمها فهما دقيقا.

هذان المشكلان أدى باللسانيين الحاسوبيين إلى وضع علامات مرتبطة بالمداخل المعجمية تحدد نوعها: التذكير والتأنيث وحالتها: المفرد، الجمع، وهذه العلامات تكون خارج سياق الجملة.

$$SN \rightarrow Det\ N$$

$$S \rightarrow Det\ et\ Nom\ S'accordent\ en\ genre\ et\ en\ nombre.$$

Soit

$$SN \rightarrow Det\ Nom$$

$$\begin{bmatrix} Gerne: G \\ Nombre: N \end{bmatrix} \quad \begin{bmatrix} Gerne: G \\ Nombre: N \end{bmatrix} \quad \begin{bmatrix} Gerne: G \\ Nombre: N \end{bmatrix}$$

وفي هذه الحالة لا تقبل البرمجيات Les automates إلا ما يوافق هذه القواعد[1].

المبحث الثالث: المستوى الدلالي.

1. المستوى الدلالي[2]:

التحليل الدلالي باب علم الدلالة، وهو اعلم الذي يعكف على دراسة المعنى، ويعد علم الدلالة جمّاع الدراسات الصوتية والنحوية والمعجمية؛ فوظيفة اللغة الجوهرية تكمن في الإبلاغ والتبليغ أو في التعبير عن المقاصد، فهي وظيفة دلالية أساسا، فالجانب الدلالي يهم أصحاب اللغة جميعا على اختلاف اختصاصاتهم ومستوياتهم الفكرية وطبقاتهم الاجتماعية. وفي هذا المستوى يمكن دراسة الكلمات، الكلمة والسياق الذي تقع فيه وعلاقاتها الإستبدالية والمتجاورة، الاختيار، المصاحبات اللغوية والصيغ الاشتقاقية[3]. وقد ارتبطت اللغة من المنظور الدلالي بفكرة المواضعة، و هي فكرة مترسخة عالجها بعض العلماء أمثال الفارابي والقاضي عبد الجبار والخفاجي. يقول القاضي عبد الجبار:"وعلى هذا الوجه تختلف اللغات والمراد لا يختلف وقد تتفق الألفاظ في اللغات المختلفة والفائدة المختلفة، ولو كانت المواضعة هي المعتبرة في هذا الباب، وقد حصلت المواضعتان في الكلمة الواحدة، لم يكن إذا وقعت من المتكلم بأن يكون خبرا عن أحد الأمرين بأولى من أن يكون خبرا عن الأخر".

يهتم الجانب لدلالي في التحليل اللغوي بالمعاني التي تحملها الكلمات وذلك بعد تحليلها مورفولوجيا ونحويا[4].

واتجهت الدراسة الدلالية للغة في اتجاهين رئيسيين:

- اتجاه يهتم بدراسة مفردات اللغة فيبحث في معانيها وأصلها وتغير المعاني وتطورها وذلك هو علم المعجم.

---

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص.36

[2] موسى حبيب، التلازم الدلالي داخل المستويات اللسانية وأثره في تعليم اللغة العربية، مجلة اللغة الوظيفية، جامعة مصطفى اسطمبولي، معسكر، الجزائر، ت. د: 20/06/30 المجلد 07 العدد 01، ص304.

[3] المرجع نفسه، ص305.

[4] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص37- 38.

- اتجاه يهتم بدراسة الدلالات اللغوية في حد ذاتها محاولا اكتشاف العوامل التي تساهم في بلورتها والقوانين التي تنتظم بفعلها ويغلب على هذا الاتجاه ثلاث مدارس كبرى:

1- نظرية المقام: وصاحبها بلومفيلد الذي يقول: "إن دلالة صيغة لغوية ما إنما هي في هذا المقام الذي يفصح فيه الدلالة والرد اللغوي والسلوكي الذي يصدر عن المخاطب".

2- النظرية السياقية: لتحديد وحدة لغوية يجب أن نتبع ونستقرئ جميع السياقات التي تحققت فليس للكلمات دلالة خاصة بل لها استعمالات إلا.

3- نظرية السمات المعنوية: أكبر ممثل لهذه المدرسة هو اللغوي بامسلاف فقد انطلق هذا اللغوي من الفكرة التي تؤمن بوجود توازن مطلق بين مستوى اللفظ ومستوى المعنى في اللغة، حيث يمكن تقطيع الكلمات دلاليا إلى أصغر وجدة تسمى السمات المعنوية.

1.1. تمثيل المعاني: حتى تتمكن البرمجيات من معالجة اللغة دلاليا يجب تمثيل المعرفة باستخدام الطرق التالية:

- تمثيل المعرفة في هيئة قواعد.
- تمثيل المعرفة بالشبكات الدلالية.
- أسلوب الدلالة الرسمية.

1.1.1. تمثيل المعرفة في هيئة قواعد: تستخدم هذه الطريقة في تمثيل المعرفة العلمية حيث تصاغ في شكل قواعد منطقية من الشكل:

إذا كان الإنسان يموت

وسقراط إنسان

فإن سقراط يموت

وتكتب بالرموز المنطقة كالتالي:

كل إنسان يموت ∩ سقراط إنسان ⟵ سقراط يموت

2.1.1. تمثيل المعرفة بالشبكات الدلالية:

الشبكات الدلالية هي شبكات هندسية تمثل بها المفاهيم ويربط بينها بخطوط الربط تمثل أنواع العلاقات الموجودة التي تربط بين هذه العناصر، وشاع استخدام هذه الشبكات مع بداية الستينات في مجالات مختلفة أهمها علم النفس وعلم المكتبات، ثم تطورت إلى تمثيل النصوص الصغيرة في جامعة كمبريدج Combridge. مع ظهور علم النفس المعرفي وهندسة المعرفة تطورت هذه الشبكات وذلك محاولة العلماء لتقليد العقل البشري[1].

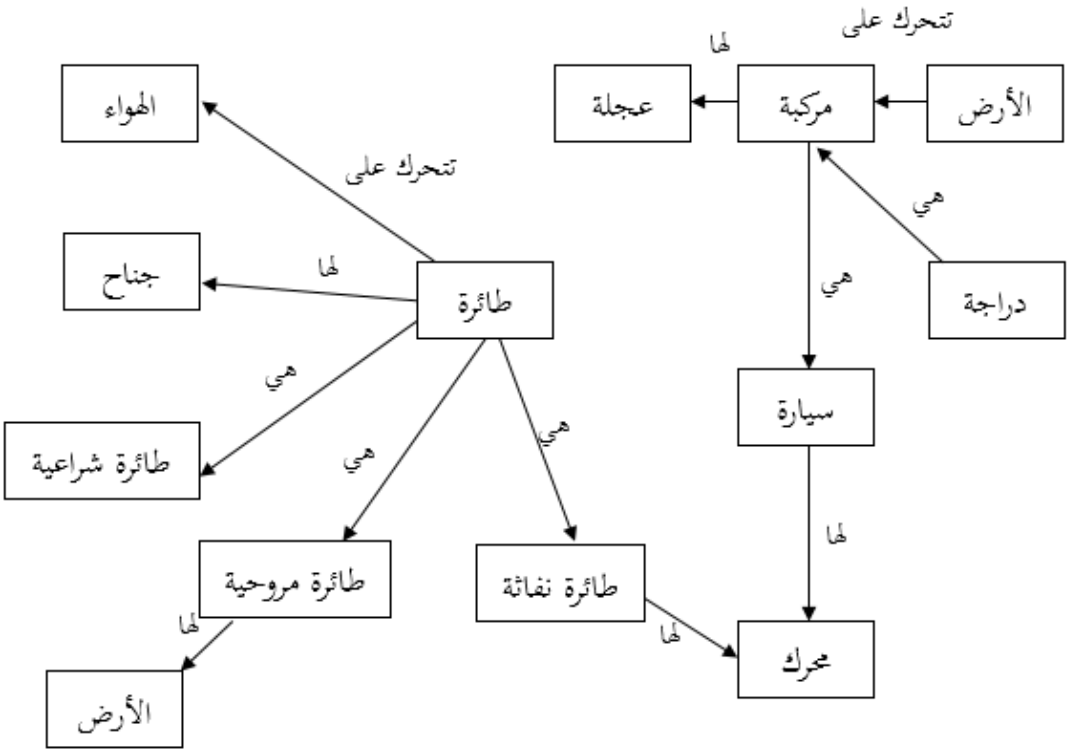
وأهم العلاقات الموجودة في الشبكات الدلالية:

- علاقة تضمين Implosion (هو، هي).
- علاقة امتلاك Possesion (له، لها).

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص38.

- علاقة ارتباط Association.

الشكل التالي يمثل شبكة دلالية على وسائل النقل[1]:



مثال عن شبكة دلالات للمركبات:

وما يمكن ملاحظته في الشبكات الدلالية أنها عبارة عن تطوير لتمثيل المكانز سهميا ويكمن الفرق في أن المكانز لا تمثل الأشياء بل تمثل ما له تعريف محدد، كما ذكر جاك شومي (Jacques chaumier) في كتابه.

3.1.1. أسلوب الدلالة الرسمية:

وتستخدم عادة في تحويل العبارات اللغوية إلى علاقات منطقية من دوال الإسناد باستخدام أساليب الدلالة الصورية (المنطق) وذلك لتسهيل الإنتاج ونقدم المثال التالي لتوضيح ذلك:

المثال: صحح عمر امتحان الفهرسة.

ففي قاعدة المعرفة تكون الاستنتاجات التالية من هذه الجملة:

- عمر أستاذ الفهرسة.
- عمر متحصل على ماجستير في علم المكتبات.

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص39.

– أجرى عمر امتحان الفهرسة لطلبة الجامعيين.

ورغم هذه الاستنتاجات بديهية للإنسان إلا أن الآلة لا تستطيع استنتاجها إلا بعد بناء قاعدة معرفة من هذه الحوادث[1].

2.1. من القاموس إلى خريطة المفاهيم:

1.2.1. تعريف خريطة المفاهيم (Ontologie):

يعود مصطلح **Ontologie** إلى اليونان وتعني العلم، وهي كلمة استخدمت في الفلسفة الأرسطية قبل 2000 سنة ويعني أرسطو بهذا المصطلح "الجانب الفيزيقي الذي يطبق على النفس البشرية في حد ذاتها".

وتطور استخدام هذا المصطلح واستخدامه في مجالات مختلفة أهمها هندسة المعرفة والإعلام الآلي وذلك لحل مشاكل تمثيل المعاني وفهمها من طرف الآلة وتعرف خريطة المفاهيم لأنها: "توضيح مخصص وأكيد لمفهوم مشترك" ويكون هذا التوضيح بواسطة جملة من المصطلحات مرتبة شجريا في ميدان معين وتستخدم في بناء قواعد المعرفة (Bases de connaissance).

والتعريف السابق يشير إلى أن خريطة المفاهيم هي تطور في تمثيل الكلمات والمفاهيم، ولكنه لا يبين العلاقات الموجودة بين هذه المصطلحات ولا كيفية الترتيب.

وهناك تشابه بين: القاموس والكشاف، قاموس المصطلحات، المكنز وخريطة المفاهيم في بعض الجوانب.

فالقاموس هو عبارة عن كلمات مرتبة ألفبائيا وذكر معناها باستعمال اللغة الطبيعية.

أما الكشاف فهو قائمة ألفبائية للكلمات، المواضيع، أسماء الأعلام التي وردت في نص ما مع ذكر صفحة ورودها.

أما قاموس المصطلحات فهو مجموعة من المصطلحات المستخدمة في التكشيف وتكون مرتبة ألفبائيا كما تشرح العلاقات الموجودة بين هذه المصطلحات.

إذا نلخص أن خريطة المفاهيم هي قائمة بالمصطلحات في ميدان ما مع ذكر العلاقات الدلالية بين هذه المصطلحات، وكذلك لتمثيل المعلومات الموجودة في نص ما.

2.2.1. مكونات خريطة المفاهيم[2]:

تعمل خريطة المفاهيم على نمذجة المعرفة وذلك بالاعتماد على:

أ– المفاهيم: تستعمل لتمثيل معاني الأشياء والمعرفة التي تحملها ويمكن أن تقسم إلى التعريفات، المصطلحات، الامتداد Extension، التركيز Intention.

التركيز: هي التعريف المحدد لمفهوم ما:

الامتداد: هي جميع المصطلحات التي لها علاقة بمصطلح ما.

التعريف: هي المداخل التي يتم استخراجها من نص.

[1] المرجع نفسه، ص40.

[2] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص41.

المصطلح: هي جميع المصطلحات التي تعرف المدخل أي الكلمات متساوية المعاني (علاقة التساوي).

مثال:

Cencept décrivant la nation de voiture.

Indentifiant : C- voiture.

Terminologie : Véhicule, Auto, Bagnome, Automobile, caisse.

Intension : Véhicule de transport automobile motorisé a quatre noues et pour transporter de une a six personnes.

Extension : Exaltai immatriculé 9658 G 44. La coralla verso de ma mère la tauing jaune devant nous, la golf TDI d'écric.

ب- العلاقات:

هي جميع العلاقات التي تربط بين المفاهيم السابقة وكذلك بين المفاهيم بعضها مع بعض وهذه العلاقات تكون ثنائية[1].

إذا اتجهت الدراسات الدلالية في اللغة في اتجاهين هما اتجاه يهتم بدراسة مفردات اللغة فيبحث في معانيها واتجاه يهتم بدراسة الدلالات اللغوية في حد ذاتها، وحتى تتمن البرمجيات من معالجة اللغة دلاليا يجب تمثيل المعرفة باستخدام تمثيل المعرفة في هيئة قواعد وتمثيل المعرفة بالشبكات لدلالية وأسلوب الدلالة الرسمية، وخريطة المفاهيم تستخدم لحل مشاكل تمثيل المعاني وفهمها من طرف الآلة وتستخدم في بناء قواعد المعرفة وتسهم في تطوير تمثيل الكلمات والمفاهيم، وتعمل خريطة المفاهيم على نمذجة المعرفة بالاعتماد على التعريفات والمصطلحات فالتعريفات هي المداخل التي يتم استخراجها من نص وهي عبارة عن مفاهيم تربط بعضها بالبعض.

المبحث الرابع: نحو إنشاء قاعدة معجمية آلية في شكل معجم كوني هرمي (مكة) "معجم عصر المعلوماتية".

إن العمل المعجمي شاق ومرهق لاسيما في مجال اللغة العربية لأسباب كثيرة، ولسنا الآن بصدد الحديث عن روعة أساليبها ودقة معانيها وتعدد ألفاظها، ومن ثم كان من الصعب أن يقوم بالعمل المعجمي فرد واحد أو مجموعة من الأشخاص. إن معجم عصر المعلوماتية لا يصنعه اللغويين فحسب وإنما يصنعه العلماء في اللغة والمتخصصون في علوم كثيرة؛ لذلك يجب تكوين فوق عمل من تخصصات مختلفة فريق لغوي، وفريق منطقي، وفريق بلاغي، وفريق هندسي... إلخ متعاونين مع أفراد يتقن الواحد منهم ضربا من الفن الذي لا بد منه لإخراج معجم يرضي أساطين اللغة، وغيرهم من العلماء، ويفيد الطلاب، وسائر القراء.

يجب على واضعي المعجم في حقل ما أو فرع ما من فروع المعرفة الإلمام بالألفاظ والمصطلحات ولا يكتفي بمجرد المعرفة اللغوية بل لا بد من التخصص الدقيق في الحقل نفسه في جميع فروع المعرفة.[2]

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص42.

[2] سلوى حمادة، المعالجة الآلية للغة العربية المشاكل والحلول، دار غريب للطباعة والنشر والتوزيع، القاهرة، ت. إ: 2009/01/01، الطبعة الأولى، ص05.

لماذا العمل المعجمي بداية:[1]

إن شبابنا المثقف يعاني كسلا ذهنيا عجيبا بحيث يرضى بالفهم الناقص للنص الأدبي الذي يحتوي مفردات لغوية جديدة عليه، أو يقنه بما يوجبه السياق من معنى مهتز أو مخطئ لهذه الألفاظ ويردد الكلمات دون الإحساس بالحاجة إلى فهم ما تعنيه هذه الألفاظ.

إن وعينا المعجمي –إن وجد– أدنى بكثير من المستوى الذي ينبغي أن يكون عليه لأبناء أمة لها تاريخ أدبي طويل، وتدخر كنوزًا لغوية في القرآن الكريم. لذا ينبغي تنمية الوعي المعجمي عند أبناء العربية، فما من شك في أنه كلما استحكم ربط استعمال اللغوي بالمعجم كانت اللغة أقل عرضة للتفتت اللغوي والتحريف اللفظي أو الغموض المعنوي.

إن المعجم المرجو سيربط اللغة بتراثها الأدبي ويقوي الوحدة اللغوية بين الناطقين بها، ويكون مستوى لغويا مشتركا لمستعمليها.

إن المعالجة المعجمية تتيح معالجة بسيطة وأنيقة في ذات الوقت للغة حيث يجب أن توضح مستويات مختلفة من المفاهيم من خلال تحليل يتيح تفهمًا أعمق للبنية الداخلية للنصوص يصعب النفاذ إليه بواسطة أيه معالجة أخرى، في الوقت ذاته يضع المحلل نصب عينيه ارتباط التحليل بالمعالجات الأخرى للتطبيقات اللغوية.

المعجم الذهني:[2]

المعجم الذهني هو المعجم الذي عن طريقه تتعرف على الأشياء ومعانيها كبشر من خلال حصيلة الخبرة البشرية المتراكمة في شتى المجالات والميادين. مداخل هذا المعجم يمكن أن تكون صوتية أو تصويرية أو حسية أو لفظية بخلاف المعاجم الآلية التي يفترض في مداخلها أن تكون لفظية فقط.

لماذا المعاجم الآلية:

يفضل معظم الناس عدم الرجوع للمعجم، أو القاموس في البحث عن الكلمات الصعبة؛ لصعوبة البحث في أوراق المعجم مع طول الوقت المطلوب، خاصة في حالة ظهور أكثر من خمس كلمات في الصفحة الواحدة. هذه الصعوبة تؤدي إلى إهمالهم لكثير من الكلمات، ومحاولة تخمينها؛ الأمر الذي يبعدهم عن المجرى الأساسي للمعنى دون أن يشعروا. محاولات التخمين هذه تؤدي للفهم الخاطئ للمعنى، وربما المعنى العكسي تماما، وتؤدي كذلك لحفظ معنى خاطئ للكلمات يصعب التخلص منه حتى بعد معرفة المعنى الصحيح، أما في حالة البحث الآلي فبمجرد كتابة الكلمة على لوحة المفاتيح يمكننا الحصول على معناها، ومرادفاتها، وسياقها، وأية معلومة أخرى يمكن أن يمدنا بها المعجم الآلي.

[1] سلوى حمادة، المعالجة الآلية للغة العربية المشاكل والحلول، مرجع سابق، ص06.
[2] المرجع نفسه، ص07.

أهم مشكلات المعاجم الحالية:[1]

1- أحادية تنظيم مواد المعجم على أساس الجذر يضع حاجزا عازلا بين المعجم ومستخدمه والذي لا يهمه أصل الاشتقاق بقدر ما يهمه المشتق ذاته، علاوة على ذلك فإن هذا التنظيم المبني على افتراض معرفة المستخدم بالتفاصيل الدقيقة للتحليل الصرفي، وقواعد الإبدال، والإعلال التي يعتذر بدونها استخلاص جذور الكلمات تهيئة لعملية الكشف عنها وهي قدرات لغوية لا تتوفر للكثيرين، أي أن الصرف قد طغى على المعجم على حساب النحو والدلالة وربما المستخدم أيضا.

2- إهمال العلاقات المعجمية التي ترتبط بين المفردات بعضها ببعضها الآخر نذكر منها:

أ- عدم توفر علاقات الترادف، والتضاد.

ب- عدم تجميع أزواج الكلمات ذات الصلة الخاصة بالأسماء، والصفات المتوافقة معها.

3- شبه إغفال للكلمات المركبة، مثل إنسان العين، وصفر اليدين، وابن السبيل.

4- إغفال البعد الزمني تطور دلالة الألفاظ.

5- ظهور عنصر الدلالة للأفعال، والأسماء، والصفات، والظروف، وأقسام الكلام الأخرى وغياب الدراسات الدقيقة لظاهرتي الترادف، والمشترك اللفظي.

6- فوضى لغة التعريف في المعجم.

فوضى لغة التعريف: فمن المنتقد أن يجد المرء في معجم حديث تفاوتًا في التعريف مثل تعريف الأبراج مثلا:

الأسد: أحد بروج السماء.

الجدي: برج من السماء بجوار الدلو.

الدلو: برج من بروج السماء.

القاموس: برج في السماء وهو البرج التاسع.

الميزان: برج سماوي وكذلك بالنسبة لكثير من الحقول الدلالية مثل الأيام والشهور وغيرها.

لذلك يجب عدم تحديد اللفظة بألفاظ مشتقة من نفس جذرها أو بمفاهيم مجهولة هي نفسها تحتاج للرجوع إلى المعجم.

مثال: "المنجد في لغة العلوم والآداب" الصفحة السادسة نبحث عن ائتدام فنجد الشرح كالآتي:[2]

ائتدام: أكل الخبز مع الإدام، الإدام: كل موافق وملائم.

إن الطائرة توافق وتلائم السفر فهل هي إدام؟!!

ويضيف إدام الطعام هو ما يجعل مع الخبز فيطيبه.

اليانسون يجعل مع الخبز ويطيبه فهل هو إدام؟!!

[1] سلوى حمادة، المعالجة الآلية للغة العربية المشاكل والحلول، مرجع سابق، ص08.
[2] المرجع نفسه، ص09.

كذلك في صفحة مائة وخمسمئة وسبعين نجد تعريف (أخس): فعل فعلا خسيسا، وأخسه: وجده خسيسا.
كذلك صفحة ثلاث مائة وثماني وسبعون: ساغ الشراب: ساغ يسوغ سوغًا والسيغ من الشراب: السائغ!!
انظر حل هذه المشكلة في التعليق.

- **منهجية إنشاء معجم يبدأ من مفهوم الكون في تسلسل هرمي:**

من الواضح لمن يدرس ويتمعن النظر في حقائق اللغة العربية، ويوغل في البحث عن خصائصها ومكوناتها أن لكثير من معانيها ألفاظا خاصة لا يغني عنها غيرها، وأن لكل مسمى من مسمياتها اسما، أو صفة، لا يلائمه سواها إذا أراد الكاتب فعلا أن يوضح مضمونه ويضع المعاني في مواضعها، ويجب أن يستعمل لكل معنى ما وضع له من اللفظ، ولا يفضل استعمال سواه. ومن ثم وجب على واضعي المعاجم الآلية الحديثة وحتى المطبوعة، الاهتمام بخصائص الكلمات، ومدلولاتها، وتحديدها بدقة لحل هذه المشكلة، خاصة المعاجم الآلية بالذات لأنها لا تعاني من مشكلتي الحجم، وسرعة البحث، وغيرها من مشاكل المعاجم المطبوعة، ويوضح هذا الفصل كثيرا من السمات والعلاقات بين الكلمات التي يمكن أن تثري معاجم الكلمات، ومن ثم يجب على واضعي المعاجم الاهتمام والإشارة لمثل هذه العلاقات، والسمات بين الكلمات وما يمكن أن تضيفه لعمل معاجم التحليل الآلي.

متطلبات معجم عصر المعلوماتية[1]:

معجم عصر المعلوماتية له دور أساسي، ومن الضروري تأليفه طبقًا لمناهج منطقية، وموضوعية، وتربوية وإلا تسبب في فوض لغوية، ولا بد من الحفاظ على اللغة يعمل على تشويهها بل وعرقلة الفكر السليم. لذلك يجب على واضعي المعجم حبك الرباط اللغوي حول الألفاظ، والمعاني ومد المعجم بمواد تناسب الحضارة الحديثة حتى يمكننا أن نأمل في جيل جديد يفهم لغته ويعي مدلولاتها ويستطيع التعبير عما بداخله من فكر. إذن معجم عصر المعلوماتية يجب أن:

1. يحافظ على سلامة اللغة ويجعلها وافية متمشية مع تقديم العلوم، والفنون، وسادة لحاجات الحياة في العصر الحاضر.

2. يبوب في صورة تبويبا هرميًا كونيًا يتم فيه ذكر المعاني متدرجة من الأصلي إلى الفرعي ومن الحسي إلى المعنوي ومن المألوف إلى الغريب.

3. يحدد جذر اللفظ ومشتقاته كما يعطي ألوان المعاني والمدلول الحقيقي والمدلول المجازي وأن يحدد المحيط الخاص بكل لفظ لكي لا يختلط بها سواه، إذ أن لكل لفظ حقل دلالي خاص به.

4. يشمل جنيا لغويا يجمع بين القديم والحديث من ألفاظ اللغة بحيث توضع الألفاظ مرتبة ترتيبا تاريخيا يكون من السهل لمستخدمه تتبع تطور الألفاظ منذ أقدم العهود حتى تاريخ صدور المعجم. إن هذا التطور سيجلي للمستخدم صورة من صور الحضارة الإنسانية وهي تدرج مدارج التقدم والارتقاء. إذن لا بد من وجود جزئية تتبع المراحل التي مر بها اللفظ والأحوال التي اعترته فتضخمت معانيه الواقعية والاصطلاحية.

[1] سلوى حمادة، المعالجة الآلية للغة العربية المشاكل والحلول، مرجع سابق، ص11.

5. عدم إهمال العلاقة بين الألفاظ والمعاني لأن هذه العلاقات توضح المفاهيم من جهة وتثري المعاجم من جهة أخرى.

العلاقة بين الألفاظ والمعاني (المشترك والترادف والتضاد):

1. يسمى الشيئان المختلفان بالسمين المختلفين وذلك أكثر الكلام.

2. تسمى الأشياء الكثيرة باسم واحد وذلك **اللفظ المشترك** من أنواعه التضاد حيث يسمى الشيئان الضدان باسم واحد. إذن المشترك اللفظي يقع على شيئين ضدين كاللون (الأبيض/أسود) ويقع على مختلفتين غير ضدين كالعين (عين الماء وعين الإبرة) وعامة أمثلة التضاد في اللغة العربية لا تزيد عن عشرين كلمة[1].

3. يسمى الشيء الواحد بأسماء كثيرة وذلك **الترادف**.

4. تسمى أزواج الكلمات حسب العلاقة النسبية للاختلاف بينها وذلك التضاد فإن كان يحمل مقارنة بين سمة ما "كالحجم (صغير وكبير)" أو في النسبة الحجمية أيضًا "أصغر وأكبر" سمي **تخالف**. أما إذا كانت العلاقة بين الكلمتين أن نفي أحدهما تأكيد الآخر مثل: "متزوج وأعزب لأن ليس متزوج تأكيد لأعزب" وذلك يسمى **تباين** أما إذا استلزمت كلمة وجود الأخرى مثل "زوج تستلزم وجود زوجة/ بائع تستلزم وجود مشتري سمي ذلك **بالتعاكس**."

5. وتسمى أزواج الكلمات التي ترتبط ببعضها مصاحبة في نفس المجال الدلالي ارتباطا قويًا **بالمصاحبات اللفظية**.[2]

محتويات معجم عصر المعلوماتية:

لتحقيق المتطلبات السابقة يجب أن يحتوي المعجم على:

1. الشرح والتعريف والتفسير: يشرح معنى اللفظ من خلال تعريف سماته الموجودة في حقله الدلالي وقد يفسر بلفظ أو بعدة ألفاظ.

2. سرد الخواص الصرفية والاشتقاقية للفظ وبيان صيغه ومعانيه الصرفية، كالطلب، والصيرورة، والمطاوعة، وبيان طابعها النحوي، كالتكلم والخطاب والغيبة والعدد والجنس، ولأي حالة من الأسماء الجامدة تذكر ويشار لذلك.

3. ذكر أصول الألفاظ وهو ما يعرف بالاتيمولوجيا.

4. سرد المعاني مرتبة على أساس التسلسل التاريخي.

5. الفصل بين المداخل المتماثلة ذات الأصول المختلفة، أو المجالات الدلالية المختلفة، "المشترك اللفظي".

6. اتباع المداخل بفقرة تسرد فيها مترادفات المداخل وتبين ما قد يكون بينها من فروق لغوية، وربما أمكن أن نضيف أسباب الترادف، وأنواعه إن وجدت.

[1] سلوى حمادة، المعالجة الآلية للغة العربية المشاكل والحلول، مرجع سابق، ص11.
[2] المرجع نفسه، ص12.

7. اتباع المداخل بفقرة نسرد فيها الكلمات التي ترتبط بها بعلاقات (التباين/ التعاكس/ التضاد/ التخالف)[1].

8. اتباع كل معنى مفرد من معاني المدخل يمثل أو أكثر يوضح ذلك المعنى أو يحدده على نحو يمتنع معه كل لبس ويتبين منه أدق الفروق والظلال.

9. اتباع كل المداخل بما يندرج تحتها من عبارات اصطلاحية والتي تدل على معاني لا يمكن أن تدرك من مجرد فهم معاني مفرداتها، مثل "رجع بخفي حنين". هذه العبارات يجب أن يشار إليها أيضا في الحقول الدلالية التي تتبعها، وفي هذا المثال يرد التعبير تحت حقل "الخسارة/الخيبة" كذلك يذكر المصاحب اللفظي مثل (ورقة وقلم).

10. إيراد الألفاظ، والعبارات العامية، أو الدارجة، وإيضاح مدلولاتها وضرب الأمثلة عليها من التمييز بين ما هو عامي، وفصيح، وفي حالة العامي يوضح اللفظ الفصيح المرادف له ليستخدم بدلا منه.[2]

ومن هذه الفقرة نقترح أن تكون الألفاظ العامية أيضا مدخلًا من مدخلات من مدخلات المعجم حتى يتمكن المستخدم بسهولة من الاستغناء عنها واستخدام مرادفها الفصيح.

11. إيراد الألفاظ الشائع استخدامها الخاطئ، تحت مجلاتها الدلالية، مع تصويبها حتى نتمكن من التخلص من هذه الظاهرة مثال (رقَم: رقْم/ جَلْطة: جلطة). ومن هذه الفقرة نقترح أن تكون الألفاظ الخاطئة أيضا مدخلا من مدخلات المعجم حتى يتمكن لمستخدم بسهولة من الاستغناء عنها واستخدام مرادفها الصحيح.

12. عند عمل أي مجال حيث يذكر فيه ما يخصه مع إضافة جانب موسوعي يقدم ألوانا من العلوم، والمعارف تحت أسماء المصطلحات، والأعلام، والمشاهير من الرجال، والنساء منذ بدء الخليقة حتى الآن. أيضا يجب أن يحتوي على إحاطة بشتى المذاهب، والديانات المختلفة، والنظريات العلمية في شتى المجالات، وجميع أسماء الأشياء، والبيئات، والفترات التاريخية، يتم ذلك من خلال لجان متخصصة بكل فرع من فروع العلوم في مستويات مختلفة من جمع المادة/ كتابتها/ مراجعتها ثم مستوى أعلى من المراجعة حتى تجيء التعريفات على أدق وجه وأكمله.

13. تزويد المعجم بالصور، والرسوم، والأصوات التي تساعد على الشرح، ونقل المعاني بدقة، وأمانة وهو ما يسمى في مجال الحاسب بالوسائط المتعددة.

الإتيمولوجيا: وهي ذكر أصول الألفاظ وتتلخص فيما يلي:[3]

أ- إذا كان اللفظ مجهول الأصل يشار إليه "الأصل المجهول/ الأصل موضع الشك".

ب- بيان أصول الكلمات المركبة بذكر لفظتيها مثل (لاسلكي تتكون من لاسلكي ومثلما تتكون من مثل وما).

ج-بيان أصول الكلمات المزجية بذكر لفظتيها.

د- بيان أصول الكلمات المختصرة بذكر أصلها مثل (درعمي أصلها ار علوم).

هـ-بيان العلة في التسمية.

[1] سلوى حمادة، المعالجة الآلية للغة العربية المشاكل والحلول، مرجع سابق، ص12.
[2] المرجع نفسه، ص13.
[3] المرجع نفسه، ص14.

و-وضع الكلمة في إطارها التاريخي.

توضيح الأصل الأجنبي الذي أخذت منه الكلمة.

خطوات إنشاء قاعدة البيانات المعجمية الآلية:

شجرة الكون:[1]

1. نبدأ بالكون كنواة لهذه الشجرة بمعنى أن كل المفاهيم يمكن أن ترى من خلال مفهوم الكون.

2. يتم التصنيف لكل الموجودات تحت هذا المفهوم في حقول دلالية، مع حصر جميع السمات لهذه الحقول.

3. تبدأ هذه الحقول في التفرع إلى مستويات أدنى بنفس الطريقة، حتى نصل لأقصى المفاهيم تدنيا في بساطته، والذي لا يمكن أن يتفرع من غيره. عن هذه النقطة يكون هذا الفرع قد انتهى العمل فيه ولكن في التطبيق الفعلي لا نصل لنهاية أي مفهوم خاصة مع اكتشافات العلم الحديثة مع ملاحظة توارث المستوى الأدنى لسمات المستويات الأعلى.

مداخل المعجم: ما هي مداخل المعجم التي تحتوي عليها؟ هل هي مفردات فقط ويتم اشتقاق الأخرى بواسطة قواعد اشتقاقية قبل الدمج المعجمي؟ أم يحتوي المعجم على جميع مداخل المفردات؟ كيف يتعامل المعجم مع المركبات؟

إن اللفظ المفرد لا يكتسب معناه من نفسه وإنما بما يحيط به من ألفاظ أخرى سواء على مستوى التركيب أو الجملة أو النص. هذه الألفاظ تربطه بها علاقات متعددة ومتداخلة تكسبه معاني مختلفة في بعض الأحيان أو تضيف له معاني أخرى لم تكن لتفهم منه بمفرده. إن مدخلات المعجم ليس مجرد قائمة من الكلمات التي لا روابط بينها، بل إن بينها علاقات متعددة ومتداخلة كلما زاد التطرق إليها وإيضاحها كلما كان المعجم له قدرة أكبر على جلاء غموض الكلمات وتوضيحها بدقة.

الخواص الصرفية الاشتقاقية:[2]

من البحث في كتب النحو يمكن أن نجد أن تحت الجذر الواحد يمكن أن يصرف الفعل في صور متعددة أكثرها انتشارا وتداولا صور كما هو موضح بجدول (1). أيضا لكل تصريف من هذه الصور يوجد أشكال عديدة للمشتقات الإسمية بعمليات من التباديل يمكن أن يكون تحت الجذر الواحد حوالي مائة مشتقة ليست جميعها تنتمي لنفس مجاله الدلالي أو ليست ذات معنى أساسًا. وبالنسبة للأفعال قد يخرج الفعل عن نوعه في بعض حالات من المشتقات.

[1] سلوى حمادة، المعالجة الآلية للغة العربية المشاكل والحلول، مرجع سابق، ص14.

[2] المرجع نفسه، ص15.

مثال:[1]

جذر	وزن	المشتقة	المصدر	السياق	النوع
ع ر ف	فعل	عرف	معرفة	عرف أنور واجباته	متعدي
	افتعل	اعترف	اعتراف	اعترف اللص بالجريمة	لازم
	تفعل	تعرف	تعرف	تعرف الرجل على اللص	لازم

أيضا الدلالة الصرفية لأوزان الأفعال قد تكون متعددة المعاني مثل وزن استفعل الذي يستخدم بمعاني كثيرة:

مثال:

1. الطلب: استسمح أنور أخاه.
2. التحول: استحجر قلب الرجل.
3. اعتقاد الصفة: استكرمه بمعنى اعتقده كريمًا.
4. القيام بالفعل استأنس بمعنى أنس.

لذلك يجب حصر الأوزان والمشتقات الإسمية التي لها معان بالطريقة الإحصائية السابقة ثم وضع رموز رقمية توضح نوع الوزن والمشتقة.

مثال:

ضرب م (1.1)، ضارب م (1.2).

أنواع المركبات تنقسم المركبات من حيث التركيب إلى:

مركبات نحوية: مضاف ومضاف إليه مثل: عين أحمد، يد الكسي.

مركبات اصطلاحية: مضاف ومضاف إليه مثل: قرن الشمس، تباشير الصباح، شرخ الشباب.

تنقسم المركبات من حيث العدد إلى:

مركبات ثنائية التركيب: صفة وموصوف –جار ومجرور– مضاف ومضاف إليه.

مركبات ثلاثية التركيب أو أكثر: نفس المركبات السابقة لكن بتركيب متعدد.

مداخل المعجم يمكن أن تكون:[2]

1. **مفردات:** اسم/ فعل/ صفة.

لا تنحصر مداخل المعجم في معجميات بسيطة فحسب وهي ما يطلق عليه المفردات بل تتكون من معجميات مركبة ومعقدة مثل:

2. **مركبات دلالية:** مثل إنسان العين، البيت الأبيض.

[1] سلوى حمادة، المعالجة الآلية للغة العربية المشاكل والحلول، مرجع سابق، ص15.
[2] المرجع نفسه، ص16.

3. **مركبات اصطلاحية:** مثل "لأول وهلة".

4. معاني نبحث عن لفظ يلائمها.

5. ألفاظ عامية أو دارجة نبحث عن نظائرها الفصيحة.

6. ألفاظ من الأخطاء الشائعة نبحث لها عن تصويب.

7. ألفاظ وعلاقات نبحث لها عن الألفاظ التي تقابل العلاقة مثل **مرادف** (كتب).

الخواص المعجمية:

الخواص الصرفية: تداولت كثير من الكتب والبحوث بشتى الأشكال والصور والخواص الصرفية لكلمات المعجم سواء كانت بحوث لغوية أو بحوثا لغوية آلية.

الخواص النحوية: وهي التي تحدد شكل الفعل في الجملة بمعنى أن نقسم الأفعال مثلا إلى:

أ– أفعال تتطلب فاعلا مثل الأفعال اللازمة.

ب– أفعال تتطلب فاعلا ومفعولا مثل الأفعال المتعدية.

وفي بعض الحالات الخاصة:

ج– أفعال تتطلب دملا اسمية مثل: (ثبت، اتضح، تبين): ثبت أن المسألة معقدة.

د– أفعال تتطلب جملا استفهامية مثل (سأل، استفهم): سألته هل جاء زيد؟

وقد لا تتطلب مثل سألته عن الكتاب.

هـ– أفعال تتطلبهما معا مثل (علم، درى، عرف): علمت أن زيدًا قائم.

لا أعلم هل زيد قائم؟

بطبيعة الحال أن الخواص النحوية هي التي ستحدد أيا من السمات التالية سيظهر مع الفعل في الجملة مع مراعاة أيضا أيّها إجباري. المقصود بالإجباري هو ما لا يصح أو يتم المعنى بدونه مثال: جاء فعل لازم لا بد له من فاعل إذن الفاعل إجباري وجوده.

أما الفعل المتعدي فيستلزم فاعلا ومفعولا مثل أكل الولد الطعام.

هذا فقط على سبيل المثال إذ ربما تستلزم بعض الأفعال اللازمة والمتعدية أركانا أخرى.[1]

إذا للمعجم الآلي دور كبير لأنه يسهل عملية البحث، فقبل المعجم الآلي كان البحث عن مفردة واحدة يأخذ وقتا طويلا، خاصة في حالة ظهور أكثر من خمس كلمات في الصفحة الواحدة، هل الأمر الذي أدى بالكثير إلى عدم البحث، أما في حالة البحث الآلي فبمجرد كتابة الكلمة على لوحة المفاتيح يمكننا الحصول على معناها ومرادفاتها وأية معلومة أخرى يمكن أن يمدنا بها المعجم الآلي.

[1] سلوى حمادة، المعالجة الآلية للغة العربية المشاكل والحلول، مرجع سابق، ص17.

الفصل الرابع:

طرق وبرامج المعالجة الآلية.

المبحث الأول: طرق المعالجة الآلية.

- الطرق المعتمدة على المعلومات المرتبطة بالنص.
- الطرق المعتمدة على المعطيات المعالجة.

المبحث الثاني: البرمجيات التجارية لمعالجة اللّغة العربية.

- **الدراسات الأكاديمية:**

1. نظام مسبار.
2. برنامج التقسيم.
3. معجم اللغة العربية المؤتمن.
4. أعمال شركة الصخر.

المبحث الثالث: سبل المعالجة الآلية للغة العربية.

1. نظام OCR.
2. التدقيق النحوي الآلي.
3. البنك السعودي للأصوات نموذج تطوير بيانات لمعالجة اللغة العربية.
4. البرنامج الرقمي العربي.

المبحث الرابع: المعالجة الآلية للغة العربية ومجالات تطبيقاتها في ظل الذكاء الاصطناعي واللسانيات الحاسوبية.

1– اللغة بين الدماغ البشري والحاسوب.

2– الذكاء الاصطناعي ومجالات تطبيقاته.

3–تطبيقات الذكاء الاصطناعي.

المبحث الأول: طرق المعالجة الآلية.

1. الطرق المعتمدة على المعلومات المرتبطة بالنص[1]:

إن هذه الطرق التي ترتكز على النظريات اللغوية هي الأولى في الظهور وذلك حوالي 1950 حيث أن الإعلام الآلي لم يكن متطورا كفاية وحجم ذاكرته صغير وقوة الحساب بطيئة لهذا اتجه الباحثين لدراسة اللغويات الحاسوبية اعتمادا على النظريات اللغوية.

2. طريقة القوانين والمقاربات اللغوية:

يعود ظهور أول محلل نحوي إلى بداية الخمسينات نتيجة تمثيل الجملة في نماذج (Forms) واقترح بارهيل [Barhaille] تحديد الجملة نهائيا في أشكال والتعبير عنها بلغة الآلة، كما أن ظهور البرمجيات الصغيرة (Automate) أدت إلى ظهور طريقة تعتمد على حالة نهائية تعالج الجمل فتحللها إلى مكوناتها الجزئية وطريقة تتابعها وجميع المعلومات الصرفية والنحوية فيها.

وفي هذه الطريقة يتم تمثيل المعلومات في قوانين تكتب من طرف لغويين ويتم الفصل بين المعلومات المعجمية والنحوية، ومميزاتها أنها عملية ولا تحتاج إلى موارد كثيرة ولكن العيب الكبير فيها هو الوقت اللازم لتحضير القواعد يدويا كما أن هذه القواعد يتطلب استخراجها وقتا طويلا وهي عملية معقدة.

3. الطرق المعتمدة على المعرفة:

كان ميدان المعالجة الآلية مرتبط بالذكاء الصناعي والأنظمة الخبيرة حتى بداية السبعينات ولهذا كانت الطرق المعتمدة على المعرفة تستفيد من تقنيات الأنظمة الخبيرة.

واعتبارا من أن الآلة يجب عليها معالجة اللغة فقد دعت الحاجة إلى التمثيل الدلالي حتى تصبح مفهومة من طرف الآلة وفي هذا المجال عمل الباحثين في ميدان الذكاء الصناعي على إنشاء أنظمة خبيرة تساعد على حل مشاكل جزئية ومنها المشاكل اللغوية.

ورغم مزايا هذه الطريقة إلّا أن أكبر عائق لها هو تمثيل المعرفة وخاصة المعرفة الضمنية لدى الأشخاص في قواعد رياضية.

وأهم المشاريع التي اعتمدت على هذه الطريقة هي عمل فريق GETA* حيث تم دمج خبير في قاعدة لغوية للترجمة لتصحيح الترجمة الآلية والتصحيح الآلي.

4. الطرق المعتمدة على المعطيات المعالجة: عرفت هذه الطرق نجاحا أحسن من الطرق الأولى نتيجة الطبيعة التقنية التي تمتاز بها حيث عاصرت الحاسوب من الجيلين الأول والثاني الذي كانت ذاكرته صغيرة وطاقة حسابه بطيئة وأهم هذه الطرق:

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص42.

1.4. نموذج ماركوف (Markov):

تعرف اختصار HMM [Hidden Markov Models] واستخدمت أول الأمر في التعرف على الكلام آليا.

ونموذج ماركوف هو طريقة إحصائية تسمح بنمذجة الظواهر الصوتية وتستعمل في عدّة مجالات: التعرف على الكلام وتلخيصه وتكشيف الوثائق، التعرف على الصور.

2.4. الطرق التصنيفية (Méthode de classification):

هي إحدى نتائج تطور تطبيقات الذكاء الصناعي في المعالجة الآلية للغة حيث اعتبرت المعالجة الآلية للغة إحدى طرق التعلم الإلكتروني، كما أن الشبكات العصبونية إحدى أهم هذه الطرق واستخدمت منذ 1985 في التعرف على الكلام.

والمبدأ الأساسي لهذه الطرق هي أن المعلومات موجودة في البيانات لذا يجب معالجة هذه الأخيرة وتصنيفها لتحديد مبادئ التصنيف وعند التحديد فإن الآلة تقوم بالتصنيف الآلي للمعلومات الجديدة التي تدخل لها[1].

3.4. الطرق الإحصائية [Méthodes statististiques]: عرفت منذ 1950 واستعملت في البحث عن المعلومات والمعالجة الآلية للغة والتحليل النحوي ثم ارتبطت بإنجاز قواعد بيانات نصية ضخمة وترتكز هذه الطرق على حساب N-Gramme أي نسبة تزدد الكلمات في نص ما. وهذه الطرق لها أهمية كبيرة عندما تكون النصوص ضخمة حيث تمكننا من معرفة تردد الكلمات داخل النص ثم يتم ترتيب النتائج المقترحة حسب هذا التردد ونذكر هنا أن معظم محركات البحث في الإنترنت تستعمل هذه الطريقة.

يتحدد الاتجاه العام في معالجة اللغة بالحاسوب في ثلاثة جوانب لغوية رئيسة، هي: (الصرف والنحو والدلالة) التي تعمل متداخلة مع بعضها، فضلا عن عنصر أساس مشترك لها جميعا يغذي بالمعلومات الخاصة بما يتطلبه كل جانب من الجوانب اللغوية الثلاث، ويعرف هذا العنصر بالمعجم. ومعالجة اللغة حاسوبيا ليس بالشيء السهل أو اليسير، وذلك لما تنفرد له لغات البشر من خصائص قواعدية كالمرونة النحوية في التقديم والتأخير والحذف والإبدال النحوي وسواها، فضلا عن الصرف العربي وقواعده. لذلك فإن تطبيقات معالجة اللغة حاسوبيا يجب أن تتحدد ضمن مجال محدد من مجالات العلم والمعرفة دون الخوض في المجالات الأدبية. وتخضع اللغة في معالجتها إلى شقين أساسين هما:

التحليل: أي تحليل النصوص والجمل بمفرداتها وكلماتها إلى عناصرها الأولية (الجذر، الوزن الصرفي).

التوليد: أي توليد أو تكوين الكلمات ومن ثم الجمل والنصوص اللغوية من المكونات الأساسية للغة.[2]

لذلك فإن تطبيقات معالجة اللغة حاسوبيا بشمولية يجب أن تتناول الشقين.

---

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص43.

[2] المرجع نفسه، ص44.

المبحث الثاني: البرمجيات التجارية لمعالجة اللّغة العربية.

البرمجيات التجارية لمعالجة اللّغة آليا:

- **الدراسات السابقة[1]:**

إن الدراسات السابقة التي اعتمدنا عليها تنوعت ما بين دراسات أكاديمية عالجت جانبا من المعالجة الآلية للغة العربية أو بعض البرمجيات التجارية التي قدمت برمجيات تقوم بتحليل وتشكيل وتصنيف النصوص في اللغة العربية.

وقد ذكرنا في هذا العنصر بعض الدراسات التي انتهت إلى إنشاء برمجيات تعالج تحليل وتلخيص النصوص العربية، كما تحدثنا على تجارب شركة صخر في هذا الميدان.

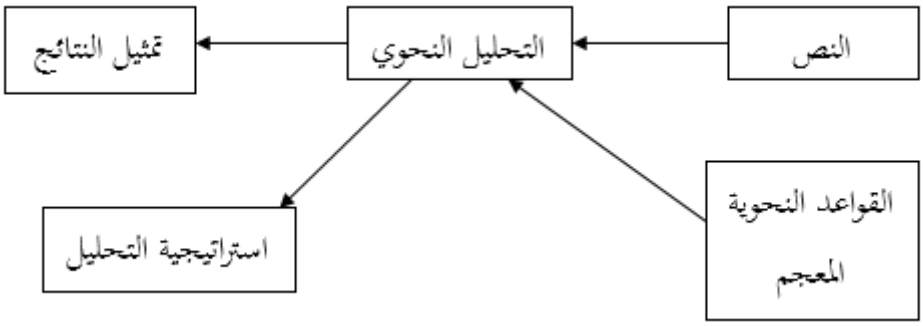
- **الدراسات الأكاديمية:**

1. نظام مسبار MASPAR:

أ- يهدف هذا النظام إلى:

- تقسيم النص إلى جمل وتقسيم الجمل إلى الكلمات الجزئية المكونة لها.
- معرفة البنية المورفولوجية لكل كلمة.
- كيفية ارتباط الكلمات مع بعضها.

ب- بنية النظام العام: يعمل على النظام كما يلي:



ج- مشاكل النظام:

- الالتباس النحوي خاصة عند عدم وجود الحركات القصيرة (يكتب النص العربي دون حركات).
- مشكل ترتيب الكلمات داخل الجملة حيث يمكن تقديم بعض الوحدات النحوية على بعض إذا توفرت بعض الشروط مثل: تقديم المفعول به على الفاعل.
- مشكل تقسيم النص.
- مشكل الارتباط (Proclitiques).

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص16-17.

د- هيكلة البرمجية: يتكون برنامج المسبار:[1]

1. المقسم (Agent segmenteur): يعمل على تقسيم النص إلى فقرات ثم إلى جمل.

2. المعجم (Agent lexical): يقسم الجمل إلى الكلمات المركبة لها ثم يقدم جميع الاحتمالات لكل كلمة (السوابق+الجذر+اللواحق).

3. العامل الصرفي (Agent morphologique): يقدم نتائج التحليل المورفولوجي وعامل الإضمار (Ellipse) والاستعارة (Anaphore) لتحديد النوع النحوي للكلمة: فعل، حرف، اسم.

4.عامل الإضمار (Agent Ellipse): يعمل على البحث عن الوحدات المضمرة في الجملة خاصة الضرورية منها (مثل: حذف الفاعل وجوبا إذا كان ضمير منفصل في الجملة: ضُرِبَ عُمرُ).

وهذا البرنامج ساعدنا كثيرا في معرفة الخلفية اللسانية للفعل العربي وكيفية تحديد محلات إعرابه إلا أن نتائج التحليل التي يقدمها غير دقيقة لأنه يهمل الحركات قصيرة المدى (الفتحة، الضمة، الكسرة، السكون) وهي معايير شكلية لمحلات إعراب الفعل العربي.

2. برنامج التقسيم Glor arab:

يعمل على تقسيم الكلمة إلى مكوناتها الجزئية: السوابق، اللواحق والجذر وذلك وفق مرحلتين:

1- المعالجة الآلية: ويتم ذلك وفق الخطوات التالية:

تقسيم الجملة إلى وحداتها الجزئية مع إضافة معلومات عنها مثل ترتيب الكلمة داخل الجملة وترتيب الجملة داخل النص.[2]

- إعادة بناء النص بعد حذف الوحدات المكررة.
- تشكيل كشاف كلمات النص.

2.المعالجة اللسانية: ويتم ذلك وفق ثلاث خطوات.

- البحث عن المداخل المعجمية لكل كلمة.
- البحث عن جذر كل مدخل معجمي متحصل عليه.
- التأكد من سلامة التحليل اعتمادا على مصادر خارجية (المكانز اللغوية).

تكمن أهمية هذا البرنامج في دوره بحث يقدم تحليل مورفولوجي للكلمات العربية ويحدد جذر كل كلمة وقد استفدنا منه عندما درسنا جذور الفعل العربي وكيفية الحصول عليها وإن كان هذا النظام غير متطور وذلك لاعتماده على مبدأ المطابقة أي يبحث على الجذر الذي يطابق الكلمة التي نريد تحليلها مورفولوجيا وسلبيات هذه الطريقة أن الفعل العربي تحدث له تغيرات مورفولوجية كثيرة عند تصريفه لذا عند استخدام مبدأ المطابقة يجب الأخذ بعين الاعتبار هذ التغيرات.

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص17.
[2] المرجع نفسه، ص18.

3. معجم اللغة العربية المؤتمن (DIINAR): هو مصدر هام للمعالجة الآلية للغة العربية يقدم المداخل المعجمية مع معلومات مورفولوجية ونحوية لكل مدخل.

يتكون معجم DIINAR من مجموعة قواعد البيانات (الفعل، الاسم، الحروف، أسماء الآلة).[1]

	الجذر	المشتقات
الفعل	19457	5970488
الاسم	39099	1781316
أسماء الأعلام	1384	11403
أسماء الآلة	445	11731
المجموع	60385	7774938

الجدول رقم (01): يبين مكونات قاموس DIINAR

ويمكننا هذا المعجم من:

- الحصول على جميع مشتقات جذر ما: سواء الأفعال، الأسماء، الصفات (Adverbe).
- اشتقاق الأسماء الممكنة من الجذور المختلفة.
- تصريف الأفعال.
- لكل اسم أو فعل يمكن الحصول على معلومات مورفولوجية نحوية خاصة بالمدخل، وهذا المعجم ساعدنا في حصر أنواع الفعل العربي ومعرفة عددها وخصائص كل نوع ولكن اعتماد هذا المعجم على مبدأ الاشتقاق جعل حجمه كبير.

4. أعمال شركة الصخر SAKHR:[2]

1- التحليل الصرفي: يتيح المعالج الصرفي المتعدد الأطوار من صخر المعالجة العميقة للكلمة العربية المفرد، ويعطي هذا المحلل نطاق الكلمات العربية بالكامل الحديث منها والقديم.

يقوم المحلل بالتعرف على جميع أشكال جذر الكلمة أي أنه يقوم باستخلاص أصل الكلمة بعد تجريدها من اللواصق ولا يتوقف عند هذا بل يتخطى ذلك لاستخلاص البيانات الصرفية للكلمة مثل الجذر والميزان الصرفي لها.

كما يعمل المركب بشكل عكسي حيث يتم توليد الكلمة من أشكالها الصرفية (الجذر، واللواحق والميزان الصرفي).

2- المكشف الآلي: يقوم هذا البرنامج بتحليل أي نص عربي والتعرف على عبارات النص وعناصر البيانات الرئيسية التلقائية.

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص19.
[2] المرجع نفسه، ص20.

وهي تساعد الشركات على تطوير فرص العمل الرئيسية عن طريق تحديد المعلومات الهامة الموجودة في البيانات الضخمة ثم استخلاص التفاصيل الأكثر أهمية بذكاء ليتم استخدامها بالمؤسسة.

تساعد الكلمات المفتاحية في تصنيف المستندات والوثائق في شجرة الموضوعات المحددة من قبل المستخدم لتسهيل استعراض المعلومات والوصول إليها، كما تساعد هذ الأداة كذلك على ربط المستندات ببعضها بطريقة ديناميكية وتقليل الوقت المطلوب للبحث في محتواها.

3- الملخص الآلي: يقوم محرك الملخص الآلي بتحديد أكثر الجمل أهمية في النص وعرضها كملخص نصي قصير، يسهل الملخص إجراء مسح شامل للجمل الهامة بالنصوص، مما يوفر الوقت المطلوب.

يوفر هذا المحرك للشركات إمكانية التلخيص النصي لجميع جمل النص باستخدام قائمة كلمات أساسية مرتبة حسب الأهمية، مع إمكانية تحديد مستوى معين للتخليص.

4- المصحح الآلي: يقوم المصحح الآلي من صغر باكتشاف وتصحيح الأخطاء الإملائية العربية والأخطاء العربية الشائعة، إضافة إلى الأخطاء النحوية ويمتاز بـ:

التصحيح التلقائي: يمكن اقتراح بدائل للكلمة الخطأ بحيث يقوم المصحح بتصحيح هذه الكلمة تلقائيا للكلمة التي اقترحها.

تقليل عدد الاقتراحات: يتميز المصحح بأنه يقلل عدد الاقتراحات البديلة للكلمة الخطأ مع مراعاة ترتيب هذه البدائل حسب علاقتها بالكلمة الخطأ، بحيث يكون الاقتراح الأول هو الأقرب للصواب.

مراعاة السياق: يراعي المصحح السياق الذي وردت فيه الكلمة الخطأ فعلى سبيل المثال إذا أدخلت عبارة "مائدة الإفطار" أو "عجلات القطار" فسيعطي البرنامج بدائل صحيحة للكلمات بما يتوافق من سياق العبارة، بالنسبة للعبارة الأولى سيعطي كلمة "الإفطار" أما العبارة الثانية فسيعطي كلمة "القطار"، وذلك أنه مزود بذخيرة لغوية واسعة أي مجموعة كبيرة من النصوص العربية تساعده في تحديد سياق الكلمة في الجملة.[1]

إذا أعادت البرامج التجارية المحلية لمعالجة اللغة العربية آليا وذلك بالنظر في معايير القوة الحقيقية للمعالجة الآلية للغات، إذ لم تعد اللسانيات الحاسوبية درسا يؤسس على مناهج بسيطة محددة، تنطلق من التحليل الآلي للغات الطبيعية عن مستويات صرفية والنحوية والدلالية، بل أصبحت رهانا لتطور اللغات والتعليم والإنتاج العلمي والإبداع التكنولوجي، ولمواكبة تطور التطبيقات المعالجة لآلية اللغات التي نستعملها نحن البشر ضمن البرامج التجارية الآلية ونظم محسوبة لمعالجة اللغات.

[1] فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي للفعل العربي، مرجع سابق، ص21.

المبحث الثالث: سبل المعالجة الآلية للغة العربية.

البرامج الحاسوبية:

"هي مجموعة التعليمات والبيانات التي توضع في القسم الإلكتروني والتي يتبعها بتنفيذ مهامه". إن تكييف البرمجيات الحرّة ذات المصدر المفتوح تطويعها للعربية أمر ميسر ويسمح بزيادة القدرات المحلية في مجال البرمجيات، وينشر التقنيات الحديثة دون القيود التي تضعها الدول المتقدمة، وباستخدام هذه البرمجيات المعربة يمكن نشر التطبيقات العربية على نطاق واسع في مجالات والحكومة الالكترونية والتعلم والتعليم والتجارة الالكترونية وغيرها، مما يدفع بعجلة تطوير اللغة العربية لخدمة مجالات هامة ذات طيف واسع ومتجدد.

سبل المعالجة الآلية للغة:

1. نظام OCR: (**Optical character necognition**) هو التعرف الضوئي على الحروف نوع من البرمجيات الحاسوبية، يمكن تحويل صور النصوص المكتوبة باليد أو بآلة مثل الماسحة إلى نصوص قابلة للتحرير، لذلك يستخدم هذا البرنامج كثيرا في الحياة اليومية، لأنه يمكن استخدام النصوص من الصور أو PDF ممسوحة ضوئيا، وربما لا يدعم اللغة العربية، لكن يوجد برنامج OCR يدعم اللغة العربية. بالإضافة إلى ذلك، هو يدعم تحويل PDF إلى نصوص قابلة للتحرير، وهو برنامج قاعدي ضروري تمتلكه كل لغة، يسمح بتحويل النصوص المصورة بكاميرا أو ماسح ضوئي Scanner إلى نص رقمي يمكن فتحه بناشر الكتروني مثل: WORD، أو أرشفته كملف على الكمبيوتر، وبدونه يلزم من جديد إعادة الطباعة كل ما كتب العربية على الكمبيوتر[1].

2. التدقيق النحوي الآلي:

هو نظام الإعراب الآلي للجملة العربية واحد من التطبيقات الرئيسية للحاق اللغة العربية بالجيل الثاني من تطبيقات معالجة اللغات الإنسانية التي تتضمن نظام الفهم الآلي وتحليل بنية السرد المستمر للغات.

3. البنك السعودي للأصوات نموذج تطوير بيانات لمعالجة اللغة العربية:

يعمل هذا البنك على تطوير قاعدة بيانات لمعالجة اللغة العربية آليا منها تطوير نظام حاسوبي للقراءة الآلية، وإخراج النصوص المكتوبة بشكل صوتي، وما يشفع لهذا البنك الآلي العربي أنه يحتوي على نظام تكنولوجي لوضع علامات التشكيل على الحروف- وهذا لا محالة يساعد المحقق والفهرس في عملية التحقيق المرقمن- وغير ذلك.

[1] ينظر: جميلة روقاب، محمد حاج هني، المعالجة الآلية للغة وتطوير البرمجيات والنظم الإلكترونية في الضوء الدرس اللساني الحاسوبي، قسم اللغة العربية، جامعة شلف الجزائر، مجلة الكترونية فصلية محكمة، أقلام الهند، www.aqlamlhind.com ت. د 02/04/2023 علاء س: 01:00h السنة السادسة، العدد الثاني، أبريل،يونيو 2021.

أضف إليه مشروع مكتبة فهد الوطنية في رقمنة المخطوطات، كما أنشأت دولة الإمارات العربية المتحدة مدينة عالمية للإنترنت، فضلا عن هذا يوجد مركز الماجد بالإمارات وهو مركز تكنولوجي عربي له جهود طيبة في حوسبة المخطوطات العربية والإسلامية، لكن هذه الجهود البشرية والمشاريع التقنية على تنوعها ما يؤخذ عليها أنها تبقى مجرد جهود وأعمال مبعثرة.

4. البرنامج الرقمي العربي (Digital arabic content)[1]:

يهدف هذا البرنامج إلى سد الفجوة بين متحدثي اللغة العربية، وبين توافر المحتوى الرقمي باللغة العربية (Digital arabic content) يعبر مفهوم هذا الأخير عن مجموعة من تطبيقات تعالج وتخزن وتعرض المعلومات باللغة العربية، والبرمجيات ولإعداد تطبيقات تتلاءم مع اللغة العربية إلكترونيا ويشمل المحتوى كل معلومة متوافرة باللغة العربية بصيغة رقمية، أي كل ما يتم تداوله رقميا من معلومات مقروءة، أو مرئية، أو مسموعة. وتنشأ أهميته من عاملين: أولهما نشر بسرعة الوصول، ومدى الانتقال إلى المتلقي، وثانيهما: كثافة المحتوى الرقمي الذي أصبح من أهم عوامل التعبير عن الثقافة والحضارة على الصعيد العالمي، حيث قلم بتأسيس البرنامج كل من شركة يملي (Yamli) إذ يعتبر **حبيب حداد** مؤسسها الفعلي، ويسمح هذا البرنامج الرقمي بتحويل الأحرف الإنجليزية إلى كلمات عربية لتسهيل الطباعة بالعربية، ويشرح حبيب حداد فيقول: "لطالما شعرت بالإحباط في عدم وجود لوحة مفاتيح عربية وكلما حاولت إيجاد أخبار عربية على الشابكة، فمثل الملايين من المستعملين أجد سهولة كبرى في التعبير عن كلمات عربية باستعمال الأحرف الإنجليزية، ولكنني أجد صعوبة في طبعها؛ فبدت لي لأن الحاجة لتقنية تحول اللفظات الصوتية هذه يشكل فعال إلى كلمات عربية ضرورية، يجدر أن تكون هذه التقنية، سلسلة بحيث يطبع المستعملون للكلمات العربية وينسون أنهم يستعملون الأحرف الإنجليزية، وهكذا ولد هذا البرنامج المحوسب حيث يحوّل (Yamli) على الأنماط الشعبية والألفاظ الأكثر شيوعا مما يزيد من دقته. ويبدي أحد المؤسسين وهو **عماد جريد يني** رأيه على أن الشبكة العربية تقلع بشكل حقيقي عندما يصبح المستعملون قادرون على إيجاد المعلومات العربية وكتابتها بسهولة فيها هناك عدد من التطبيقات لتقنيتها هذه، ارتأينا الاستجابة لهاتين الحاجتين يسمح (Yamli search) للمستعملين بالبحث على الإنترنت من خلال محرك البحث غوغل (Google) فيستطيع مستعملو الشابكة الالكترونية البحث في مضمون الويب، من فيديوهات، وصور وأخبار ومدوّنات إلكترونية أما (Yamli editor) فيسمح للمستعملين بكتابة نص عربي يمكن نقله إلى أي وسيط الكتروني"[2].

[1] ينظر: جميلة روقاب، محمد حاج هني، المعالجة الآلية للغة وتطوير البرمجيات والنظم الإلكترونية في الضوء الدرس اللساني الحاسوبي، مرجع سابق.
[2] المرجع نفسه.

المبحث الرابع: المعالجة الآلية للغة العربية ومجالات تطبيقاتها في ظل الذكاء الاصطناعي واللسانيات الحاسوبية.

1- اللغة بين الدماغ البشري والحاسوب:

تعتبر الدراسات اللسانية البيولوجية من الحقول المهمة في المعرفة العلمية الحديثة، وتأتي أهميتها من العلاقات التجريدية التي تربط الفاعليات اللغوية بالفاعليات البيولوجية في الدماغ البشري، وإذا أخذنا بوجهة النظر البيولوجية اللغوية، فإنه ينبغي أن تعتبر دراسة المقدرة البيولوجية في جميع مراجلها وأوضاعها الأولية والمتوسطة ثم الأخيرة، مشابهة تماما لدراسة الوحدات الحسابية المتفاعلة في الأدمغة الإلكترونية في جميع مراحلها وأوضاعها[1].

وفي هذا الإطار برهن علم الدماغ Cerveau إلى جانب علوم أخرى كعلم الأعصاب وعلم النفس عن طبيعة العلاقة بين الدماغ Cerveau البشري والحاسوب، فقد أثبت ك.لازورث G.Lazorthe 1989م أن الدراسة السيكولوجية توقفت في تحصيل التقارب بين الدماغ والحاسوب، من منطلق أن كلا منهما يعتمد في إدخال المعلومات على وحدات إدخال، وفي إخراجها على وحدات إخراج، وفي معالجتها على وحدات المعالجة المركزية، وفي تخزينها على وحدات تخزين، هذه الوحدات المختلفة تعتبر من المكونات الأساسية للحاسوب[2].

وهناك لغات برمجة متنوعة، كل منها قد صمم لتسهيل تشفير نوع معين من الاستخدامات، ويمكن تصنيف لغات البرمجة إلى لغات البرمجة الدنيا (لغة التجميع) التي تتعامل مع التفاصيل الداخلية للحاسوب، ولغات البرمجة العليا أو المتطورة التي تستخدم صياغا تقترب من اللغات الطبيعية، ولكنها تحتاج إلى مترجم لتحويل أوامرها إلى لغة الحاسوب، مما أدى إلى ظهور لغات ذات قابليات متفاوتة، وفيما يلي جدول يضم بعض من أشهر هذه اللغات، إضافة إلى المجالات التي صممت للتطبيق فيها[3]:

لغة البرمجة	مجالات استخدامها
فورتران (Fortran formula translation)	المسائل الهندسية والعمليات الرياضية.
الكول Algol (Algorithmic language)	المسائل الهندسية والأعمال الرياضية.
كوبول Cobol (Common business oriented language)	إدارة الأعمال وتشمل عددا من أنظمة المكتبات.

[1] سناء منعم، اللسانيات الحاسوبية والترجمة الآلية بعض الثوابت النظرية والإجرائية، منشورات مختبر العلوم المعرفية، الأردن، الطبعة الأولى، 2015م، ص74.
[2] المرجع نفسه، ص75.
[3] المرجع نفسه، ص79.

بي. ل/ 1. 1/ P.1 (Programming language)	الأعمال التجارية والعلمية والنصوص.
أر. بي. جي. RPG (Report program generator)	التطبيقات التجارية مثل: الحسابات والفواتير.
بيسك Basic (Beginners All-purpos Sympolic instruction code)	المدارس الأولية لتعليم البرمجة وحل المسائل.
باسكال Pascal	المدارس الأولية لتعليم البرمجة وحل المسائل.
لوكو Logo	المدارس الأولية لتعليم البرمجة وحل المسائل.

2- الذكاء الاصطناعي ومجالات تطبيقاته:

يقال إن الحاجة أم الاختراع، ولعل ذلك ينطبق تمام الانطباق على اختراع وتطوير حواسيب تستجيب من جهة لأساليب الذكاء الاصطناعي التي من أهم مقوماتها القدرة الآلية للتعامل مع اللغة: مبانيها ومعانيها، أصواتها وخطوطها، رموزها ومدلولاتها، معطياتها ومفاهيمها، ومن جهة أخرى هندسة المعرفة، فالإنسان منذ فجر التاريخ يتطلع إلى تطوير أدوات الإنتاج التي يمكنها أن تخفف عنه أعباء العمل الجسمي والفكري أيضا[1].

الذكاء الاصطناعي: تعريف وتحديد:

لقد ظهر الذكاء الاصطناعي في الخمسينيات من القرن الماضي نتيجة الثورة التي حدثت في مجال المعلومات والسيبرنتيقا Cybornétique، وترمي أبحاثه إلى تحقيق هدفين رئيسيين[2]:

الأول: الوصول إلى فهم عميق للذكاء الإنساني عن طريق محاكاته.

الثاني: الاستثمار الأفضل للحاسوب والعمل على استغلال إمكاناته كافة.

وحتى الآن لم يوضع تعريف محدد للذكاء الاصطناعي، ويعود ذلك إلى غموض مفهوم الذكاء، ولكن يمكننا تعريفه من خلال ما سبق بأنه أحد علوم الحاسوب الحديثة التي تبحث عن أساليب متطورة لبرمجته للقيام بأعمال

[1] سناء منعم، اللسانيات الحاسوبية والترجمة الآلية بعض الثوابت النظرية والإجرائية، مرجع سابق، ص82.
[2] المرجع نفسه، ص84.

واستنتاجات تشابه ولو في حدود ضيقة تلك الأساليب التي تنسب لذكاء الإنسان، فهو بذلك علم يبحث أولا في تعريف الذكاء الإنساني وتحديد أبعاده، ومن ثم محاكاة بعض خواصه كالتعلم والاستنباط واتخاذ القرارات. بشكل أكثر تحديدا يعرف الذكاء الاصطناعي بأنه ذكاء يظهر عند كيان اصطناعي غير طبيعي من صنع الإنسان، ويعنى بتطوير خوارزميات وتقنيات ذكية لتطبيقها في الحواسيب والروبوتات بحيث تمتلك سلوكا ذكيا في أداء المهام أو في حل المشاكل.

3-تطبيقات الذكاء الاصطناعي:

اكتسب علم الذكاء الاصطناعي أهمية بالغة في السنوات الأخيرة لتطبيقاتها العديدة في مجالات حيوية كالدفاع، والاستخبارات، والحاسوب والترجمة الآلية وغيرها، وهو أحد العلوم التي نتجت -كما سبقت الإشارة- عن الثروة التكنولوجية المعاصرة.

إن لعلم الذكاء الاصطناعي مجالات تطبيقه واسعة جدا وإذا توفرت الشروط التالية:

- البعد عن الخيال والتوقعات المبنية على غير الحقائق العلمية المثبتة مثل ما يردنا من كتب وأفلام خيالية عن أنظمة تفوق ذكاء الإنسان أو تتحكم فيه.
- استيعاب حدود هذا العلم الجديد وإمكاناته واستخداماته في المجالات المناسبة.
- العمل على حث علمائنا وباحثينا على دراسته، وبحثه، وتطويره، وتطويعه لاحتياجاتنا.

ويمكن تصنيف تطبيقات الذكاء الاصطناعي إلى أربع مجالات رئيسية، تمس معظم مجالات حياتنا اليومية وهي:

1. النظم الخبيرة Les systèmes experts.
2. نظم التعليم الآلي Les systèmes d'apprentissage automatique.
3. نظم المعالجة الآلية للغات Les systèmes de traitement automatique des langues.
4. نظم محاكاة القدرات البصرية والحركية الإنسان[1].

[1] سناء منعم، اللسانيات الحاسوبية والترجمة الآلية بعض الثوابت النظرية والإجرائية، مرجع سابق، ص89.

خاتمة

من أهم فروع اللسانيات التطبيقية هي اللسانيات الحاسوبية، فالدراسات والبحوث العلمية ازدهرت في الآونة الأخيرة وتكاثر الباحثون في هذا الميدان الذي تتلاقى فيه علوم الحاسوب وعلوم اللسان وهو ميدان علمي تطبيقي واسع كما هو معروف إذ يشمل التطبيقات الترجمة الآلية والإصلاح الآلي للأخطاء المطبعية وتعليم اللغات بالحاسوب... وبهذا كان عنوان بحثي "اللسانيات الحاسوبية والمعالجة الآلية للغة العربية"، إذ بعد الاطلاع عليه وجمع المراجع والمصادر والتجول في الفصول والاطلاع على مباحث توصلت منه إلى نتائج تالية:

1. تعتبر اللسانيات الحاسوبية أحد أهم فروع اللسانيات التطبيقية، تهتم بالإفادة من معطيات الحاسوب في دراسة قضايا اللسانيات المتعددة مثل: رصد الظواهر اللغوية وفقا لمستوياتها الصوتية، الصرفية، النحوية وإجراء العمليات الإحصائية وصناعة المعاجم في اللسانيات الحاسوبية، تجمع بين علمين اللسانيات وعلم الحاسوب، فهو علم يبحث في اللغة البشرية كأداة طبيعية لمعالجتها بالحاسوب.

2. للسانيات الحاسوبية فوائد كثيرة لأنها فتحت أفاق كثيرة للدرس اللغوي بحيث سهلت له الكثير من الصعوبات فاعتماده على أجهزة الكترونية التي تتمثل في الحواسيب، مكنه من استيعاب الكم الهائل للنصوص وسرعة المعالجة والاستجابة لأي سؤال وبانتشار المعالجة الآلية للغات الطبيعية انتشارا واسعا هيئ للغة دخولها مجالات التحليل الرياضي والمنطقي وهيئ الحاسوب للقائه مع اللغة بسرعة فائقة وسعة الذاكرة.

3. تهتم المعالجة الآلية بدراسة الجوانب الحاسوبية للغة والمشاكل اللسانية والحاسوبية التي تواجه هذه المعالجة سواء كانت لغة منطوقة م مكتوبة كلية عن طريق الحاسوب أو جزئية بتدخل انسان في بعض المراحل.

4. للمعالجة الآلية دور في الحفاظ على اللغة العربية وتطويرها وتعد مسألة الحفاظ على هذه اللغة من الأمور الواجبة على جميع أبنائها لأجل النهوض بها لمواكبة التطور العلمي والتكنولوجي.

إذا أصبحت المعالجة الآلية للغة العربية نتيجة حتمية لكثرة النصوص الالكترونية المكتوبة لهذه اللغة، مما تطلب وجود نظام حاسوبي عربي يكون فعال من ناحية إرادته للمعلومات وتخزينها في ذاكرة الحاسوب، وتعد المعالجة الآلية للغة الطبيعية من أهم وسائل تسيير النصوص الضخمة وخاصة في عصر الانفجار المعلوماتي لذا صار الاهتمام بتطبيق تقنياتها في معالجة اللغة العربية أمر في غاية الأهمية واللسانيات الحاسوبية من العلوم الجديدة التي أحدثت ثورة على الساحة اللغوية، فقد استطاعت اللغة العربية بفضل هذا العلم أن تضمن لنفسها مكانة في علم الحاسوب حتى يكون شأنها شأن اللغات المتطورة والراقية وهذا عن طريق كافة تضافر كافة الجهود المبذولة والأعمال المنجزة في مجال

اللسانيات الحاسوبية، من خلال تشكيل فرق عمل متكاملة من المختصين اللغويين والخبراء الحاسوبيين في أي دراسة علمية تهدف إلى برمجة الأنظمة اللغوية للغة العربية ومعالجتها آليا.

قائمة المصادر والمراجع

الكتب:

1. اللغة العربية بين اللسانيات الرتابية الحاسوبيّة واللّسانيات العرفانية في الجامعات الجزائرية، المجلس الأعلى للّغة العربية، أعمال الندوة الوطنية 24-25 ديسمبر 2019، المكتبة الوطنية العامة، الجزائر، منشورات المجلس، 2019، الجزء الأول.
2. اللغة العربية بين اللسانيات الرتابية الحاسوبيّة واللّسانيات العرفانية في الجامعات الجزائرية، المجلس الأعلى للّغة العربية، أعمال الندوة الوطنية 24-25 ديسمبر 2019، المكتبة الوطنية العامة، الجزائر، منشورات المجلس، 2019، الجزء الثاني.
3. اللغة العربية بين اللسانيات الرتابية الحاسوبيّة واللّسانيات العرفانية في الجامعات الجزائرية، المجلس الأعلى للّغة العربية، أعمال الندوة الوطنية 24-25 ديسمبر 2019، المكتبة الوطنية العامة، الجزائر، منشورات المجلس، 2019، الجزء الثالث.
4. خولة طالب الإبراهيمي، مبادئ في اللسانيات، دار القصبة للنشر، (2000- 2006)، الطبعة الثانية.
5. سلوى حمادة، المعالجة الآلية للغة العربية المشاكل والحلول، دار غريب للطباعة والنشر والتوزيع، القاهرة، ت. إ: 2009/01/01، الطبعة الأولى.
6. سناء منعم، اللسانيات الحاسوبية والترجمة الآلية بعض الثوابت النظرية والإجرائية، منشورات مختبر العلوم المعرفية، عالم الكتب الحديث للنشر والتوزيع، ط01، 2015.
7. صالح بلعيد، دروس في اللسانيات التطبيقية، دار هومه للطباعة والنشر، الطبعة الثامنة.
8. عبد الرحمن الحاج صالح، بحوث ودراسات في اللسانيات العربية، موفم للنشر، د.ط.
9. عصام محمود، اللسانيات الحاسوبية العربية، دار الوفاء لدنيا الطباعة والنشر، ط1، 2015.
10. عيسى برهومة، مقدمة في اللسانيات، دار الكتب، الجامعة الهاشمية، 2005، د.ط.
11. محمود عكاشة، التحليل اللغوي في ضوء علم الدلالة، دار النشر للجامعات، القاهرة، ط1.
12. نايف خرما، أضواء على الدراسات اللغوية المعاصرة، المجلس الوطني للثقافة والفنون والآداب، الكويت، سبتمبر1978، سلسلة عالم المعرفة، العدد 9.
13. زكريا إسماعيل، طرق تدريس اللغة العربية، دار المعرفة الجامعية، ط01.

المجلات والمواقع الإلكترونية:

1. جميلة روقاب، محمد حاج هني، المعالجة الآلية للغة وتطوير البرمجيات والنظم الإلكترونية في الضوء الدرس اللساني الحاسوبي، قسم اللغة العربية، جامعة شلف الجزائر، مجلة الكترونية فصلية محكمة، أقلام الهند، www.aqlamlhind.com ت. د 02/04/2023 علاء س: 01:00h السنة السادسة، العدد الثاني، أبريل، يونيو 2021.
2. سعيد فاهم، قراءة في الإسهامات اللسانية الحاسوبية العربية، مجلة دراسات لجامعة الأغواط–سبتمبر 2015– الطبعة 36، عدد 36.
3. سلوى السيد حمادة، المعالجة الآلية للغة العربية معناها ومبناها، جامعة طيبة، المدينة المنورة، مجلة فكر الثقافية، نشر بتاريخ 19/08/16، ت د، 2023/02/12 على 16:49 د.
4. شاعر التميمي، المعالجة الآلية للغة العربية، جهود وتحديات، 2010/01/02، د.ت، 2023/02/12 س15.30 د، مجلة ستار تايمز.
5. صفية مطهري، التفاعل الدلالي بين المستويات اللسانية، التراث العربي، مجلة فصيلة محكمة، اتحاد الكتاب العرب بدمشق (العدد 112)، 2008.
6. كنزة منديل، مجالات استخدام اللسانيات الحاسوبية في البحث اللساني العربي، جامعة محمد لمين دباغين سطيف 02 (الجزائر)، 2022/05/19، مجلة الدراسات لجامعة سطيف.
7. موسى حبيب، التلازم الدلالي داخل المستويات اللسانية وأثره في تعليم اللغة العربية، مجلة اللغة الوظيفية، جامعة مصطفى اسطمبولي، معسكر، الجزائر، ت. د: 20/06/30 المجلد 07 العدد 01.

رسائل الماجستير:

1. فارس شاشة، المعالجة الآلية للغة العربية، إنشاء نموذج لساني صرفي إعرابي للفعل العربي، مذكرة ماجستير في علم المكتبات والتوثيق جامعة الجزائر، كلية العلوم الاجتماعية والإنسانية، الجزائر، 2008.

الفهرس

الصفحة

شكر وتقدير.

الملخص.

خطة البحث.

مقدمة..... أ

مدخل..... 1

الفصل الأول: اللسانيات الحاسوبية..... 6

المبحث الأول: تعريف اللسانيات الحاسوبية..... 7

المبحث الثاني: تاريخها (النشأة والتطور) 10

المبحث الثالث: موضوع اللسانيات الحاسوبية ومجالاتها 13

المبحث الرابع: وظيفة اللسانيات الحاسوبية وفوائدها 15

الفصل الثاني: المعالجة الآلية الجانب النظري..... 17

المبحث الأول: البداية والتاريخ للمعالجة الآلية..... 18

المبحث الثاني: تعريف المعالجة الآلية..... 20

المبحث الثالث: دور المعالجة الآلية في الحفاظ على اللغة العربية..... 22

المبحث الرابع: واقع المعالجة الآلية للغة العربية..... 25

الفصل الثالث: المعالجة الآلية للغة العربية..... 27

المبحث الأول: المستوى المورفولوجي..... 28

المبحث الثاني: المستوى النحوي..... 34

المبحث الثالث: المستوى الدلالي..... 40

المبحث الرابع: نحو إنشاء قاعدة معجمية آلية في شكل معجم كوني هرمي..... 44

الفصل الرابع: طرق وبرامج المعالجة الآلية..... 53

المبحث الأول: طرق المعالجة الآلية..... 54

المبحث الثاني: البرمجيات التجارية لمعالجة اللّغة العربية..... 56

المبحث الثالث: سبل المعالجة الآلية للغة العربية..... 60

المبحث الرابع: المعالجة الآلية للغة العربية ومجالات تطبيقاتها في ظل الذكاء الاصطناعي واللسانيات الحاسوبية..... 62

خاتمة..... 65

قائمة المصادر والمراجع..... 68

الفهرس..... 71