دليل أبحاث حوسبة اللغة العربية

الجزء الأول

الطبعة الأولى

عمان - الأردن

١٤٤٠هـ/٢٠١٩م

المملكة الأردنية الهاشمـــية
رقم الإيداع لدى دائرة المكتبة الوطنية
(٢٠١٩/١/٤٣٤)

٤١٠
اللجنة الوطنية الأردنية للنهوض باللغة العربية
دليل ابحـاث مؤسسـة اللغـة العربيـة / اعـداد اللجنـة الوطنيـة الأردنيـة للنهوض باللغة العربية .- عمان: اللجنه ، ٢٠١٩

٢ج( ) ص .
ر. إ. : ٢٠١٩/١/٤٣٤.
الواصفات : /اللغة العربية//الابحاث//الحوسبة/

يتحمل المؤلف كامل المسؤولية القانونية عن محتوى مصنفه ولا يعبّر هذا المصنف عن رأي دائرة المكتبة الوطنية.

حقوق الطّبع محفوظة لمجمع اللُّغة العربيّة الأردنيّ

ويمنع تصوير هذا الكتاب أو إعادة طبعه من دون إذن المجمع

طباعة
مطبعة عمال المطابع
Printers Press

اللجنة الوطنية الأردنية للنهوض
باللغة العربية للتوجه نحو مجتمع المعرفة

دليل أبحاث حوسبة اللغة العربية

الجزء الأول

إعداد

فريـــق عمـــل

دليل أبحاث حوسبة اللغة العربية

رجب ١٤٤٠هـ/ آذار ٢٠١٩م

بسم الله الرحمن الرحيم

اللجنة الوطنية الأردنية للنهوض باللغة العربية للتوجه نحو مجتمع المعرفة

الأستاذ الدكتور

خالـد الكـركي

رئيسًا

الأعضـــاء

الأستاذ الدكتور
إسحق فرحان

الأستاذ الدكتور
كامل العجلوني

الأستاذ الدكتور
عبداللطيف عربيات

الأستاذ الدكتور
محمد حمدان

الأستاذ الدكتور
أحمد هليل

الأستاذ الدكتور
محمد عدنان البخيت

الأستاذ الدكتور
فاروق صبحي زيد الكيلاني

الأستاذ الدكتور
همام غصيب

أمين عام وزارة التربية والتعليم – مندوبًا عن الوزارة (الأستاذ محمود العكور)

أمين سر اللجنة الوطنية الأردنية للتربية والثقافة والعلوم (إبتسام أيوب)

فريق العمل

الأستاذ الدكتور
محمد زكي خضر
رئيسًا

الأستاذ الدكتور
محمد السعودي
(أمين عام مجمع اللغة العربية) عضوًا

الدكتور
مجدي صوالحة
(الجامعة الأردنية) عضوًا

الدكتور
سامي عبابنة
(الجامعة الأردنية) عضوًا

الدكتور
يوسف حمدان
(الجامعة الأردنية) عضوًا

الأستاذ
مأمون الحطاب
(شركة النص العربي) عضوًا

ملخص الدراسة

تروم هذه الدراسة بالدرجة الأولى إطلاع الباحثين العرب - اللغويين خاصة - على الجهود البحثية المكتوبة باللغة الإنجليزية والمبذولة في سبيل حوسبة اللغة العربية، وذلك بهدف ردم الهوة بين مجالي عمل الحاسوبيين واللسانيين وتوحيد جهودهما، إضافة إلى تلخيص لما نشر باللغة العربية حول الموضوع.

وقد جاءت الدراسة في مقدمة وأربعة فصول وملحق بالمصطلحات، كشفت المقدمة عن أهمية حوسبة اللغة العربية، وما يصاحبها من إشكاليات، وآفاق حلّها، ثم توضيح أهداف الدراسة ومنهجيتها وآلية التنفيذ.

وقدم الفصل الثاني دراسة مسحية لأبرز الدراسات العربية في مجال حوسبة اللغة العربية، وقد تُتبعت الجهود التي بذلت في سبيل تحقيق هذه الغاية، ونُظر إلى هذه الدراسات بالالتفات إلى بدايات تشكل الوعي بأهمية هذه القضية فعُرضت أهم الدراسات الشمولية التي سعت إلى إبراز أهمية الحوسبة للغة العربية، ثم أُظهر اهتمام الباحثين العرب بمحاولة حوسبة قواعدية اللغة في جانبي النحو والصرف، ثم أُتبع ذلك بمبحثين تناول الأول ما يُعرف بالمصادر. أما المبحث الثاني فاهتم بالأبحاث المعنية بالتطبيقات الحاسوبية ذات الصلة باللغة العربية.

وقدم الفصل الثالث ترجمة لأهم ما نشر من أبحاث ودراسات ومؤلفات عن المعالجة الحاسوبية للغة العربية في المجلات العلمية المحكمة المتخصصة والمؤتمرات العلمية وما يتوفر على الشابكة (الإنترنيت) وتبويبها حسب التخصصات الفرعية، وقد تمت العملية باختيار (٥) أوراق من الأبحاث المسحية المهمة باللغة الإنكليزية ووضع خلاصات لها بحدود ٢٠٠٠ كلمة لكل بحث، واختيار (٤٥) بحثًا من الأبحاث المهمة باللغة الإنكليزية ووضع خلاصة طويلة لها باللغة العربية بحدود ١٠٠٠ كلمة للبحث الواحد، ثم اختيار (١٠٠) بحث

باللغة الإنكليزية ووضع خلاصة مختصرة لها باللغة العربية بحدود ٤٠٠ كلمة لكل بحث. وقد وزعت الأبحاث على ثلاثة جوانب : القواعدية اللغوية التي حصرت في جانبي النحو الصرف، ثم قضايا الحوسبة ومستلزماتها في جانب المصادر وشملت خمسة مواضيع فرعية هي: المعاجم الآلية، والمكنز، والمدونات الموسومة، والأنطلوجيا، وشبكات الكلمات. أما التطبيقات الحاسوبية، فضمت جميع المباحث التأسيسية الضرورية تنظيرًا وتطبيقًا: توزعت التطبيقات على موضوعات فرعية عددها ١١ موضوعاً، هي: التعرف الصوتي وتوليد الكلام، والقارئ الآلي، والترجمة الآلية، والتحليل الدلالي، والسؤال والجواب، وتلخيص النصوص، وتحليل الرأي، والتعرف على أسماء الأشياء، والتعليم والتعلم الآلي. وقد ارتئي في اختيار الأبحاث أن تكون في غالبيتها مما نشر في السنوات الأخيرة لكي تعطي فكرة عن آخر ما توصل إليه البحث العلمي في حقول موضوعاتها. وقد وضع الفصل الثالث هذا في جزء ثان مستقل

وقد أبرز الفصل الرابع أهم النتائج والتوصيات التي خلصت إليها الدراسة...

واستلزمت الدراسة الوقوف عند المصطلحات الكثيرة التي تواجه الباحث في هذا المجال فتناول الملحق في نهاية الدراسة أكثر من (٧٠٠) مصطلح في مجال حوسبة اللغة الطبيعية مما تتطلبه دراسات من هذا القبيل؛ مما يؤمل أن يسهم في حلّ كثير من الإشكاليات التي تواجه الدارسين خاصة من اللغويين بإطلاعهم على هذه المصطلحات ومفاهيمها. وقد تضمن الجدول الذي ضم تلك المصطلحات تعريفا لها.

شكر وعرفان

يسر فريق العمل أن يتقدم بالشكر والعرفان "للجنة الوطنية للنهوض باللغة العربية للتوجه نحو مجتمع المعرفة" لتبنيها هذه الدراسة، ولرئيس اللجنة وأعضائها لتوجيهاتهم السديدة.

ويقدر فريق العمل جهود المساعدين كافة، الذين أسهموا في المساعدة في تلخيص الأبحاث وترجمة خلاصاتها، ومعالجتها حاسوبيًا، والوصول بها إلى صورتها النهائية.

قائمة المحتويات

الموضوع	الصفحة
ملخص الدراسة	١١
شكر وعرفان	١٣
قائمة المحتويات	١٥
الفصل الأول التمهيد	١٧
١-١ أهمية الدّراسة وأهدافها	١٩
١-٢ مجتمع الدراسة	١٩
١-٣ آلية التنفيذ	٢٠
١-٤ مدة إنجاز العمل ومراحله	٢١
١-٥ صعوبات الدراسة ومحدداتها	٢٢
١-٦ معايير اختيار الأوراق البحثية	٢٣
١-٧ معايير اختيار المصطلحات	٢٤
١-٨ مقدمة عن الحاسوب	٢٤
١-٩ حوسبة اللغة	٢٥
١-١٠ مناهج البحث والتطوير في حوسبة اللغة	٢٧
١-١١ التحديات التي تواجه العاملين في حوسبة اللغة العربية	٣٠
١-١٢ الموارد الأساسية في عملية حوسبة اللغة العربية	٣٢

١-١٣ محتويات الكتاب ٣٣

الفصل الثاني واقع الدراسات العربية في حوسبة اللغة العربية ٣٥

٢- ١- مقدمة ٣٧

٢- ٢-١- حوسبة اللغة العربية مقدمة تاريخية ٣٨

٢- ٢-٢- الحوسبة والمستويات اللسانية للغة العربية ٣٩

٢- ٢-٣- الدراسات ذات الطابع الشمولي ٤١

٢-٣-١- الحوسبة والمستويات اللغوية الأساسية للغة العربية ٦٠

٢-٣-٢- المستوى النحوي ٦٠

٢-٣-٣- المستوى الصرفي ٧١

٢-٣-٤- الموارد ٨٤

٢-٣-٥- التطبيقات ١١٧

٢-٤- خاتمة ١٤٤

الفصل الأول

التمهيد

المحتويات

١-١- أهمية الدّراسة وأهدافها
٢-١- مجتمع الدراسة
٣-١- آلية التنفيذ
٤-١- مدة إنجاز العمل ومراحله
٥-١- صعوبات الدراسة ومحدداتها
٦-١- معايير اختيار الأوراق البحثية
٧-١- معايير اختيار المصطلحات
٨-١- مقدمة عن الحوسبة
٩-١- حوسبة اللغة
١٠-١- مناهج البحث والتطوير في حوسبة اللغة
١١-١- التحديات التي تواجه العاملين في حوسبة اللغة العربية
١٢-١- الموارد الأساسية في عملية حوسبة اللغة العربية
١٣-١١- محتويات الكتاب

الفصل الأول
التمهيد

١-١ أهمية الدّراسة وأهدافها:

يهدف المشروع إلى استقصاء أهم ما نشر عن المعالجة الحاسوبية للغة العربية باللغة الإنكليزية وتقديم ترجمة لها باللغة العربية لتقديمه للقارئ العربي وخاصة المتخصصين باللغة العربية وذلك لتسهيل مساهمتهم في خدمة اللغة العربية عن طريق المعالجة الحاسوبية. كما يقدم مسحًا لما نشر عن المعالجة الحاسوبية باللغة العربية أيضًا.

١-٢ مجتمع الدراسة:

تتبع ما نشر من أبحاث ودراسات ومؤلفات عن المعالجة الحاسوبية للغة العربية في المجلات العلمية المحكمة المتخصصة والمؤتمرات العلمية وما يتوفر على الشابكة (الإنترنيت) وتبويبها حسب التخصصات الفرعية.

وضع ورقة مسحية لما نشر عن حوسبة اللغة العربية بالمصادر العربية مع الإشارة إلى تلك المصادر.

اختيار ٥ أوراق مسحية من الأبحاث المهمة باللغة الإنكليزية ووضع خلاصات لها بحدود ٢٠٠٠ كلمة لكل بحث

اختيار (٤٥) بحثًا من الأبحاث المهمة باللغة الإنكليزية ووضع خلاصة طويلة لها باللغة العربية بحدود ١٠٠٠ كلمة للبحث الواحد.

اختيار (١٠٠) بحث من الأبحاث الأقل أهمية باللغة الإنكليزية وضع خلاصة مختصرة لها باللغة العربية بحدود ٤٠٠ كلمة لكل بحث.

تحوي الترجمات على المعلومات الآتية:

عنوان البحث الأصلي.

عنوان البحث باللغة العربية.

أسماء المؤلف/ المؤلفين.

مصدر البحث.

تصنيف البحث ضمن حقول الأبحاث المتعددة.

خلاصة البحث باللغة العربية بحيث تحوي هدف البحث وطريقته واستنتاجاته

يجهز ملحق يتضمن قائمة بالمصطلحات الجديدة التي احتوتها الأبحاث مع تعريفاتها واقتراح أن يقوم مجمع اللغة العربية بترجمتها لكي تُقرّ تلك المصطلحات.

يلحق بالدليل كشاف موضوعي للأبحاث.

تكوين قاعدة بيانات على الشابكة (الإنترنيت) لأبحاث معالجة اللغة العربية.

١-٣ آلية التنفيذ

يقوم فريق العمل المتكون من المتخصصين بتقنية المعلومات والمتخصصين باللغة العربية بما يأتي:

وضع جدول بالتخصصات الفرعية في المعالجة الحاسوبية للغة العربية.

البحث عن الأبحاث المنشورة عن المعالجة الحاسوبية للغة العربية من مصادرها الأصلية المتمثلة بالمجلات العلمية المحكمة المتخصصة وكتيبات المؤتمرات العلمية والشابكة (الإنترنيت).

تقييم الأبحاث حسب أهمية المادة العلمية والأصالة المتضمنة فيها وتصنيفها إلى مهمة وأقل أهمية.

اختيار ٥ أبحاث من الأبحاث المسحية ووضع خلاصة وافية لها بحدود ٢٠٠٠ كلمة باللغة العربية.

اختيار ٤٥ بحثًا من الأبحاث المهمة ووضع خلاصة موسعة بحدود ١٠٠٠ كلمة باللغة العربية.

اختيار ١٠٠ بحث من الأبحاث الأقل أهمية ووضع خلاصة قصيرة لها بحدود ٤٠٠ كلمة باللغة العربية.

جمع ما يرد أثناء الترجمة من مصطلحات علمية. فإذا كانت المصطلحات معتمدة من مجمع اللغة العربية الأردني، استخدمت تلك المصطلحات، وإن لم يكن هناك مصطلح سبق اعتماده يقترح مصطلح جديد له مع تعريف للمصطلح ويوضع بجدول ملحق خاص لتقديمه لمجمع اللغة العربية لإقراره أو إيجاد بديل عنه.

كتابة دراسة مسحية لأبحاث المعالجة الحاسوبية المنشورة باللغة العربية.

تصميم قاعدة بيانات على الشابكة (الإنترنيت) لأبحاث معالجة اللغة العربية ووضع الأبحاث أو الروابط لها.

١-٤ مدة إنجاز العمل ومراحله

مدة إنجاز المشروع سنة واحدة تبدأ من تاريخ توقيع مذكرة التفاهم، ويجوز تمديد المدة إذا اقتضت الحاجة وبموافقة الطرفين.

أما مراحل العمل فكانت:

شهر واحد لوضع خطة العمل التفصيلية وإقرار التخصصات الفرعية.

أربعة أشهر لجمع الأبحاث وتبويبها.

أربعة أشهر لوضع الخلاصات وترجمتها.

شهر لتصميم قاعدة البيانات ووضع الأبحاث عليها.

شهران لإعداد التقرير النهائي وجداول المصطلحات في مجلد واحد أو مجلدين.

١-٥ صعوبات الدراسة ومحدداتها

لقد واجهت الدراسة عدد من الصعوبات التي يمكن إجمالها فيما يأتي:

أ – لقد وجد فريق العمل أن عدد الأبحاث عن حوسبة اللغة العربية المنشورة باللغة الإنكليزية تجاوز ثلاثين ألف بحث أو دراسة. وقد شكّل ذلك عبئًا ثقيلا لاختيار العدد المطلوب في هذه الدراسة. وقد اختار فريق العمل أبحاثًا ذات تواريخ نشر حديثة قدر الإمكان. لقد اطلع فريق العمل على أكثر من ٤٠٠ بحث وقام باختيار العدد المطلوب منها في هذه الدراسة.

ب- واجهت فريق العمل مشكلة اختيار المصطلحات العربية المناسبة للمصطلحات التي وردت في ترجمات الأبحاث. وقد زاد عدد المصطلحات عن ٧٠٠ مصطلح. وهي غيض من فيض من المصطلحات الواردة في الأبحاث المختارة ، حيث لم يدرج في هذا التقرير سوى المصطلحات التي وردت في الملخصات وليس كل ما في الأبحاث الأصلية كاملة. وقد وجد أن المصطلحات الواردة في موقع مجمع اللغة العربية من هذه المصطلحات هو عدد قليل، كما أن ما قُدِّم إلى المجمع في المراحل الأولى من سير العمل لم يحصل فريق العمل على إجابة بالموافقة أو الرفض أو التعديل.

ج – عند البحث عن أبحاث عن حوسبة اللغة العربية في مواضيع محددة، وجد أن هناك ندرة في بعض هذه المواضيع مثل البيانات الضخمة أو التعلم الآلي أو الذكاء الاصطناعي. لذلك لم يتضمن هذا التقرير بعض هذه المواضيع رغم أهميتها، أو جرى ضم بعض المواضيع إلى مواضيع قريبة. كما أن الاختيار بين فئات الأبحاث الثلاثة (الأبحاث المسحية والأبحاث من فئة أ (١٠٠٠ كلمة) وفئة ب (٤٠٠ كلمة) ، رغم أنه قد روعي في ذلك أهمية البحث في الموضوع المعين ، إلاّ أنه لم يكن بالإمكان الالتزام بذلك بدقة، لصعوبة تقييم أهمية بعض الأبحاث بدقة.

د – نظرًا لضخامة العمل المتضمن، فقد كان لا بد من توزيع عملية الترجمة على بعض أعضاء فريق العمل والمساعدين، ورغم محاولة الالتزام بأسس متشابهة للترجمة، إلاّ أنه كان لا بد

من وجود اختلافات في الأساليب وفي تحديد أهم الأجزاء الواجب تضمينها في الترجمة. وقد روعي في التلخيص عدم تضمين أية معادلات أو أشكال أو مناقشات حاسوبية دقيقة، نظرًا لأن الهدف الرئيس هو تقديم فكرة عن الأبحاث للغويين العرب بشكل خاص وليس للمتخصصين بتقنية المعلومات.

١-٦ معايير اختيار الأوراق البحثية

أولاً: راعت اللجنة عند اختيارها الأبحاث التي ترجمت اختلاف اهتمام الناظر إلى موضوع حوسبة اللغة باختلاف خلفيته. فاللغوي يوجه اهتمامه نحو معالجة مستويات اللغة المختلفة من خلال لسانيات الحاسوب. ومن أبرز الأمثلة على هذا النوع البحوث في نظريات تشومسكي في النحو التوليدي والتحويلي. وتشمل بحوث هذا الاتجاه مواضيع الصرف والنحو والدلالة والمعجم.

أما الحاسوبي فيهتم بخوارزميات التحليل والتركيب وتطبيقاتها في معالجة اللغات الطبيعية. ومن أمثلة هذه الخوارزميات آلية فحص الحالات المحدودة التي طبقت في تحليل الصرف.

وقد رأت اللجنة توجيه عملها إلى الجمهور الذي سيتلقاه من خارج بيئة البحث والتطوير التقني، لذلك احتوت البحوث المختارة مقدمات ومفاهيم أساسية في حوسبة اللغة.

ثانيا: راعت اللجنة المراحل التاريخية التي مرت بها بحوث حوسبة اللغة وخطوط البحث الرئيسية التي ظهرت خلالها وأهمها:

- البحث المستند إلى حوسبة بنية اللغة وقواعدها وقد بدأت به أبحاث حوسبة اللغة تاريخيا وساد حتى ثمانينات القرن الماضي.
- البحث الإحصائي الذي يبني نتائجه من التحليل الإحصائي لكميات كبيرة من النصوص للوصول إلى نتائج حول اللغة دون الخوض في بنيتها وتركيب مستوياتها، وقد ساد منذ فترة التسعينات من القرن الماضي.

- البحث الهجين الذي يجمع بين الخطين السابقين وهو أحدث الاتجاهات ظهوراً، وهو ما تبنى عليه حاليا تطبيقات الذكاء الاصطناعي وأنظمة الآلة المتعلمة.

ثالثاً: إن حوسبة اللغة هي تقنية توظف علوما وتقنيات مختلفة، لكنها ليست علما مجردا بذاتها، وهي لذلك سريعة التطور والتغير، وبناء على هذه الحقيقة :

فإن أكثر البحوث القديمة وخاصة في خوارزميات التحليل والتركيب والتطبيقات فقدت قيمتها بظهور خوارزميات وتطبيقات أحدث منها.

وحتى البحوث الجديدة فإنها عرضة لأن تصبح غير ذات قيمة في أي وقت بظهور بحوث أحدث منها.

لكن هناك بحوثا يمكن اعتبارها بحوثا مؤسِسة ساهمت في تكوين تيار بحثي أو اتجاهات ويستدل عليها من خلال كثرة العزو لها في البيئة العلمية.

١-٧ معايير اختيار المصطلحات

١. أن يكون للمصطلح تعريف في المعاجم التقنية الإنجليزية أو على شبكة الإنترنت يمكن الرجوع إليه ومقارنته بالترجمة العربية.

٢. أن يكون لترجمة المصطلح أهمية لفهم البحوث التي تحتويه.

٣. نسبة انتشار المصطلح في بحوث مجاله.

٤. عدم وجود ترجمة عربية للمصطلح، أو وجود ترجمة رأت اللجنة أنها لا تعبر بدقة عن المعنى التقني للمصطلح.

١ - ٨ مقدمة عن الحوسبة

تسارع التقدم العلمي والتقني خلال السنوات المئة الماضية بصورة غير مسبوقة في تاريخ الإنسانية. ونمت العلوم والتكنولوجيا معا يغذي واحدهما الآخر، فقد سهلت المعرفة العلمية المتجددة بناء تقنيات غير مسبوقة، وأتاحت لنا هذه التقنيات أساليب حياة لم تخطر على بال

أسلافنا من البشر. وفتحت لنا الأبواب لنسعى بجرأة نحو تغيير العالم الذي نعيش فيه، وقادتنا إلى مزيد من المعرفة العلمية التي ألهمتنا بناء تقنيات جديدة... وهكذا في دورة لا تنتهي.

لقد صمم الإنسان الحواسيب الأولى على الورق وبناها بيده، بينما تقوم اليوم حواسيب بتصميم الجيل القادم من الحواسيب، وتقوم حواسيب أخرى بإنتاجها بعد ذلك في مصانع مؤتمتة بالكامل. وهذا النموذج ليس حالة خاصة بالحواسيب، فكل العلوم وتقنياتها تشهد نموا هائلا باستخدام الحواسيب. وكل تقدم حاسوبي يفتح الطريق أمام بقية العلوم لتتقدم بسرعة ولينتج عنها تقنيات جديدة. وأهم اتجاهات البحث التي تصنع هذا التقدم هي بحوث الذكاء الاصطناعي، التي تندرج تحتها بحوث حوسبة اللغات الطبيعية.

والعلم تراث إنساني، وملك للبشرية جمعاء، لكن إنتاجه محصور اليوم في دول ومجتمعات دون دول ومجتمعات دون أخرى. والعلم أساس اقتصاد المعرفة، لذلك فقد أغنى أمما، وبقيت أمم أخرى فقيرة لأنها لم تمتلك هذا الأساس. وكلما تسارع تقدم العلوم، اتسعت الفجوة بين منتجي العلم ومستهلكي منتجاته. وأثر العلم في بناء المجتمعات لا يقل خطورة عن أثره في بناء الاقتصادات. لذلك فالأثر الاجتماعي للفقر العلمي يترافق مع الفقر الاقتصادي في رسم صورة مظلمة لمستقبل الأمم المحرومة من امتلاك العلم وإنتاجه.

لقد صنفت الدول منذ الثورة الصناعية إلى دول متقدمة صناعيا، ودول نامية تتجه نحو هذا التقدم. وفي ظل عصر ثورة المعلومات صار أساس التصنيف هو الفجوة الرقمية بين الدول، ودرجة تبني العلم، ووجود بيئة لإنتاجه ورعايته، واحترام العلماء، وإعلاء قيم البحث العلمي، والتعلم والتعليم باللغة الأولى حتى لا تكون اللغة حاجزا أمام المعرفة.

ومع أن الفضاء العلمي مفتوح لمن يشارك فيه من العلماء من جميع أنحاء العالم، إلا أن ثمرة هذه المشاركات تصب في تلك الدول التي توفرت فيها شروط ازدهار العلم. ومن هذا المنظور نطل على واقع حوسبة اللغة العربية.

١-٩ حوسبة اللغة

حوسبة اللغة هي عملية تطوير لبرامج حاسوب تقوم بثلاث وظائف: الأولى؛ تهيئة أجهزة

الحاسوب لقبول مدخلات اللغة البشرية، حتى تبدو هذه الأجهزة وكأنها تفهم اللغة. والثانية تحليل اللغة المدخلة وفهم دلالتها. والثالثة استجابة أجهزة الحاسوب على صورة مخرجات باللغة الطبيعية.

وفي منطق الحاسوب الذي يقوم على معالجة مدخلات لإنتاج مخرجات، فإن المدخلات تكون كلاما منطوقا، أو مطبوعا، أو مكتوبًا، وهذه هي وسائل التواصل البشري باستخدام اللغة. أما المخرجات فقد تكون جوابا منطوقا أو مطبوعا، أو عملية يقوم بها الحاسوب مثل قيادة السيارة دون سائق إلى المكان الذي حدده الراكب، أو تشخيص الحالة المرضية بناء على الحوار بين المريض (والروبوت) الطبيب، أو الترجمة الفورية بين اللغات.

وتتكون أنظمة حوسبة اللغة من برامج وخوارزميات حاسوبية، وقواعد بيانات وقواعد معرفة، وشبكات عصبية اصطناعية. وتعمل هذه المكونات معا لمحاكاة القدرات اللغوية للعقل البشري في استقبال اللغة وفهمها وإنتاجها. وتتداخل علوم عدة وتقنيات لبناء هذه المكونات ولتحقيق هذا الجهد. لكن العمليات الأساسية التي تقوم عليها تطبيقات حوسبة اللغة هي العمليات التي تقوم بها الحواسيب نفسها أي الإدخال والمعالجة والإخراج.

العملية	الهدف	البرنامج
الإدخال	رقمنة اللغة ليمكن معالجتها	التعرف على الحروف التعرف على الصوت
المعالجة	تحليل اللغة للوصول إلى الفهم الآلي	منظومة تحليل اللغة (التحليل الصرفي، التحليل النحوي، التحليل الدلالي، التحليل المقامي)
الإخراج	إنتاج اللغة	تحويل النص إلى صوت تحويل الصوت إلى نص

وقد تدرجت حوسبة اللغة في مراحل منذ تشأتها حوالي منتصف القرن العشرين إلى أن وصلت اليوم إلى مرحلة (الحوار بين الإنسان والآلة). وحققت حوسبة اللغات الطبيعية تقدما كبيرا ومتسارعا. وأصبح الحوار بين الإنسان والآلة باللغة الطبيعية واقعا يعيشه الناس اليوم، لا حلما مثل ما كان قبل سنوات قليلة. وتجاوزت الأبحاث هذه المرحلة إلى مراحل متقدمة تناقش أنسنة الآلات (الروبوت) بإضافة المشاعر إلى أحكامها.

وسيتمدد وجود الحواسيب في تفاصيل حياة الناس اليومية مع هذا التطور، وستغير هذه الحواسيب المتطورة المفاهيم الأساسية للتعليم ، والطب، والأعمال، وكل نشاطات الحياة الإنسانية.

وحتى تستطيع أمة أن تواكب هذا التطور، وأن تستفيد منه في بنائها الحضاري، فعلى أبنائها إتقان إنتاج هذه التقنية الحديثة واستخدامها. فلم يعد استخدام التقنية مرتبطا بتعلم مجموعة من المفاتيح لتشغيل برنامج أو آلة، بل يكفي معرفة اللغة التي نخاطب بها الآلة وتخاطبنا الآلة بها. واللغة العربية فصيحة أو عامية ليست لغة تفهمها الآلة حتى اليوم.

لقد حققت الشركات العالمية مثل مايكروسوفت وجوجل وأبل تقدما كبيرا في حوسبة اللغة العربية، لكن تطبيقات هذه الحوسبة في مناحي الحياة المختلفة بقيت متأخرة عن تطبيقات اللغات الأخرى، وبقيت تلك التطبيقات تحوم حول المعالجة السطحية للغة دون عمق في التحليل لفهم دلالاتها. وبقيت تقنيات و موارد حوسبة اللغة العربية المستخدمة في تلك التطبيقات غير متاحة في مجملها خارج هذه الشركات.

١-١٠ مناهج البحث والتطوير في حوسبة اللغة

عرف البحث والتطوير في حوسبة اللغة منذ نشأ ثلاثة مناهج هي حوسبة قواعد اللغة، والحوسبة الإحصائية ، والمنهج الهجين الذي يجمع بين المنهجين السابقين.

وقد بدأت جهود حوسبة اللغة الطبيعية تاريخيا باعتماد منهج حوسبة قواعد اللغة. واتجه اللسانيون نحو بناء وصف منطقي للغة، بهدف تحويل هذا الوصف إلى نموذج رياضي

يمكن برمجته. وكانت نظرية النحو التوليدي التحويلي بإصداراتها المتعاقبة أحد نتاجات هذه المحاولات.

واستمر زخم هذه المقاربة بحثا وتجريبا حتى الثمانينات من القرن الماضي. ومع أن نتائجها الحاسوبية كانت متواضعة مقارنة بالجهود الضخمة، والأوقات الطويلة، والكلف العالية التي أنفقت فيها، فقد فتحت هذه المقاربة الباب أمام دارسات جديدة في اللسانيات لم تتوقف عن التطور رغم ضعف إنتاجيتها الحاسوبية.

ونتيجة لشعور الجهات الممولة للأبحاث وخاصة في الولايات المتحدة الأمريكية بضعف الجدوى الاقتصادية لهذه المشاريع، فقد اتجه الدعم منذ نهاية الستينيات من القرن العشرين إلى أبحاث أنظمة الذكاء الاصطناعي. وتبع هذا ظهور الحوسبة الإحصائية التي بنيت عليها أنظمة الذكاء الاصطناعي. وبدأ الاتجاه نحو حوسبة قواعد اللغة بالفتور.

وقد تفاعل الحاسوبيون واللسانيون العرب مع منهج حوسبة القواعد، ولعل ذلك يعود إلى الدرجة العالية من المعيارية التي تسم مستويات اللغة العربية المختلفة. هذه المعيارية التي تغري بإمكانية صياغتها صوريا ونمذجتها رياضياً، فقاموا بمحاولات لنقل نظريات النحو التوليدي التحويلي وتطبيقها على حوسبة اللغة العربية. لكن هذه المحاولات لم تنتج تطبيقات محوسبة وصلت إلى السوق، ولم تترك أثرا واضحا في مسيرة حوسبة اللغة العربية. ولم تسجل المحاولات لبناء نظرية لغوية عربية حديثة أي نجاح يلفت النظر أو يمكن البناء عليه في حوسبة اللغة العربية.

وظهر خلال هذه الفترة اتجاه لحوسبة اللغة العربية اعتمادا على قواعد صرفها ونحوها التقليديين باستخدام ما سمي بالتوصيف اللغوي. وتمثل التوصيف اللغوي بإعادة صياغة المعرفة اللغوية التراثية على صورة مجموعة من القواعد التي تسهل برمجتها.

وكانت برامج التحليل الصرفي العربي أول إنتاج يعتمد هذه الطريقة بتطبيق خوارزميات الحالة المنتهية في تطويرها. لكن نجاح هذا الأسلوب لم يتجاوز برامج الصرف إلى مستويات لغوية أخرى كالنحو والدلالة. ويعود نجاح برامج الصرف المطورة بهذه الطريقة إلى أن بنية الصرف العربي منضبطة إلى درجة تسهل برمجتها، وليس إلى طريقة التوصيف اللغوي التي لم

تستطع توصيف النحو العربي بصورة يمكن برمجتها.

أما الحوسبة الإحصائية فتقوم على تحليل كمية كبيرة من البيانات إحصائيا لاستخلاص أية قوانين أو أنماط تحكم هذه البيانات أو تظهر فيها. ويمكن تطبيق هذا التحليل على أي نوع من البيانات، فقد تكون بيانات مالية وأرقاما تجمعت من حركة أسعار الأسهم والعملات، أو قد تكون بيانات الحضور والغياب لطلاب جامعة، أو غير ذلك.

وساعد النمو الكبير جدا لبيانات النصوص الإلكترونية المخزنة في الحاسوب في ظهور فكرة تحليل هذه النصوص إحصائيا لاستنتاج الأنماط والقواعد التي تحكم ترتيبها وتركيبها. وقامت برامج التحليل الإحصائي للبيانات الضخمة بتزويد الحاسوبيين بمعلومات وفرت عليهم في كثير من الأحيان عناء محاولة التقعيد اللساني لمحتوى هذه البيانات.

وظهرت النتائج الإيجابية لهذه الطريقة في تحليل أنماط جميع مستويات اللغة، وفي نسبية وجود القانون أو القاعدة إلى مجموع النصوص، وفي المقارنة بين اللغات. وسرعت هذه الطريقة تطوير تطبيقات حوسبة اللغات الطبيعية، وخاصة الترجمة الآلية. وكان أبرز نتائج العمل بها برنامج الترجمة بين اللغات الذي طورته شركة جوجل.

وقد تطورت تقنيات التحليل الإحصائي بإضافة خوارزمية جديدة تقوم بتخزين نتائج التحليل، ثم إعادة استخدام هذه النتائج في عمليات التحليل اللاحقة للوصول إلى عمق أكثر ودقة أكبر، بصورة تحاكي عمل الخلايا العصبية في الدماغ، لذلك أطلق عليها اسم الشبكة العصبية الاصطناعية. وسميت الحواسيب التي تشغل هذه الخوارزميات بأنظمة التعلم الآلي، وذلك لقدرتها على تخزين معلومات جديدة واتخاذ قرارات بناء عليها دون تدخل بشري.

إلا أن استخدام الطريقة الإحصائية في حوسبة اللغة كانت له أعراض جانبية، فقد أضعف شعور الحاسوبيين بضرورة العمل المشترك مع اللسانيين، والاستفادة من جهودهم في عمليات حوسبة اللغة الطبيعية. وسهل التوجه نحو حوسبة اللهجات العامية التي يمكن تحليلها إحصائيا لتوفر كميات كبيرة من نصوصها في وسائل التواصل الاجتماعي، دون الحاجة إلى البحث في قواعدها.

لقد أسهم المنهج الإحصائي في تسريع وتيرة تطوير تطبيقات حوسبة اللغات. إلا أن دقة عمل هذه البرامج اعتمدت على كميات النصوص التي جرى تحليلها إحصائيا، وعلى مدى استيعاب هذه النصوص للظواهر اللغوية المختلفة.

وبما أن النظام الكتابي لأي لغة لا يستوعب كل ظواهرها اللغوية، فإن التحليل الإحصائي للنصوص المكتوبة أو المطبوعة لا يكفي للوصول إلى وصف شامل ودقيق للغة. لذلك ظهرت الحاجة إلى مقاربة هجينة تستفيد من عمليات الإحصاء لتسريع تطوير البرامج، ومن عمليات حوسبة القواعد لاستكمال جوانب النقص بالتقعيد الجزئي للظواهر اللغوية التي لم تظهر في التحليل الإحصائي للنصوص.

وتظهر هذه الحاجة بقوة في حالة اللغة العربية، فنظام كتابة العربية يقوم في الأصل على الاختزال، ولا يعبر عن ظواهر لغوية كامنة في النص تعين على فهم المكتوب، لذلك استحدث فيه التنقيط والتشكيل. لكن النصوص المكتوبة والمطبوعة اليوم تخلو من علامات التشكيل، وهذا يمثل تحديا كبيرا لعمليات حوسبة اللغة مهما كانت طريقة المعالجة المستخدمة لتحليل النصوص، فمخرجات الحواسيب تعتمد على درجة اكتمال المدخلات.

١-١١ التحديات التي تواجه العاملين في حوسبة اللغة العربية

ينتمي العاملون في حوسبة اللغة العربية إلى واحدة أو أكثر من الجهات الآتية:

- الشركات الكبيرة مثل جوجل ومايكروسوفت وأبل
- الجامعات سواء أكانوا أساتذة أم طلابا
- المؤسسات البحثية
- الشركات الخاصة
- الأفراد أصحاب المبادرات الذاتية

ورغم كثرة العاملين في حوسبة اللغة العربية، والكثرة النسبية لأبحاث حوسبة اللغة العربية، فإن ما يلفت النظر أن جل المنتجات العربية المتاحة اليوم في السوق من إنتاج الشركات الكبيرة مثل جوجل ومايكروسوفت وأبل وسامسونج.

أما الجهود التي يقوم بها الباحثون الأكاديميون وطلابهم والعاملون في المؤسسات البحثية، والعدد الكبير من الأبحاث التي تنشر سنويا حول حوسبة اللغة العربية، فإنها لم تنتج تقنية متاحة، ولا منتجات تتناسب مع عدد تلك الأبحاث، ولم تضف منتجات جديدة ذات أثر إلى التي أنتجتها الشركات الكبيرة.

ويعود هذا إلى أسباب منها عدم نضج أكثر هذه الأبحاث لتصبح منتجات. فهذه المشاريع البحثية تتسم بقصر العمر لارتباطها عادة برسالة للحصول على درجة جامعية، أو ورقة بحثية للترقية. كما أنها غالبا ما تعالج جزئية من اللغة لا تكفي لتكون منتجا قائما بذاته. ومن هذه الأسباب شح الموارد المادية الماثل في نقص التمويل. وكذلك شح موارد البحث والتطوير الحاسوبي اللغوي التي تمثل البنية التحتية لعمليات البحث والتطوير.

ومن الأسباب غياب الجانب العملي، إذ لا نجد حول تلك البحوث البيئة المتكاملة لدورة حياة المنتج منذ الولادة بحثا نظريا وصولا إلى التطبيق ثم إلى السوق، وما يتطلبه هذا الأمر من عمل مجموعات مختلفة خارج إطار مجموعة البحث. وهو ما يتحقق من خلال العلاقة بين بيئة البحث وبيئة الأعمال.

ويواجه الباحثون الحاسوبيون غياب الإطار النظري اللغوي الحديث لحوسبة اللغة العربية، مما يضطرهم، بحثا عن هذا الإطار، إلى التعامل مع الفكر التراثي اللغوي دون امتلاك العدة العلمية للتعامل معه. كما يعانون غياب نظرية مصطلحية عربية موحدة – أو حتى جهود منسقة لإيجاد مصطلحات جديدة تعريبا ووضعا – مما أوجد ظاهرة فوضى المصطلحات في أبحاث حوسبة اللغة المنشورة باللغة العربية. ولا يوجد حتى اليوم معجم موحد أو قاعدة بيانات موحدة لهذه المصطلحات.

وأما جهود الشركات والأفراد، فمع أن بعض الشركات العربية والغربية طورت منتجات تقوم على حوسبة اللغة العربية مثل برامج الترجمة البسيطة والمعاجم الآلية، والتعرف الآلي على الحروف، والقراءة الآلية، إلا أن التحديات التي واجهت هذه الشركات حالت دون استمرارها.

وأهم هذه التحديات ضعف الطلب على هذه المنتجات في السوق العربي والعالمي. وانتشار ثقافة قرصنة البرمجيات في العالم العربي، وغياب حماية حقوق الملكية الفكرية، وعدم القدرة على

منافسة الشركات الكبيرة في الدخول إلى الأسواق العالمية والعربية، ونقص الكوادر المؤهلة في حوسبة اللغة لأن تلك الكوادر تهاجر للالتحاق بالشركات العالمية الكبيرة، وعدم توفر موارد البحث والتطوير الحاسوبي اللغوي الأساسية، مما يعني البدء من الصفر في كل مرة لإنتاج البنى التحتية المطلوبة لحوسبة اللغة العربية والمستخدمة في تلك المنتجات. وربما كان هذا التحدي هو الأكبر الذي يحول دون تقدم حوسبة اللغة العربية ويحصر جهود تطويرها بجهات محددة. وقد أدت هذه التحديات إلى العجز المادي عن الاستمرار في تطوير المنتج وتحسينه، والخروج من السوق.

١-١٢ الموارد الأساسية في عملية حوسبة اللغة العربية

إذا كان هدف حوسبة اللغة هو محاولة بناء عقل لغوي آلي، يمكن بالقياس إدراك الحاجة إلى عناصر أساسية تبني هذا العقل، فلا بد من إغناء هذا العقل بالمفردات، ومعانيها، ثم بناء المعرفة اللغوية لاشتقاق تلك الكلمات وتصريفها، وبعدها بناء العقل النحوي لتركيب الجمل. ولكل عنصر من هذه العناصر، مورد من موارد تطوير حوسبة اللغة، يمكن تطويره وإتاحته لاستخدام الباحثين والمطورين، ليقوموا بإجراء التحسينات عليه أو باستخدامه في منتجات وتطبيقات حوسبة اللغة.

ومما تمتاز به بيئة البحث والتطوير لحوسبة اللغة الإنجليزية هو توفر تلك الموارد الأساسية التي تعين الباحث على بناء منتجات جديدة دون البدء من نقطة البداية. ومن أهم هذه الموارد شبكة الكلمات، والمدونات الموسومة، ونتائج التحليل الإحصائي لتلك المدونات بصورة يمكن استخدامها دون العودة إلى مرحلة التحليل مرة أخرى، والمحللات النحوية والدلالية. وفي الجانب اللساني وجود إطار رياضي يصف اللغة وصفًا يمكن حوسبته.

ويعرض الجدول الآتي بعض الموارد اللغوية المحوسبة الأساسية التي يجب أن تكون متاحة لتطوير أنظمة حوسبة اللغة العربية:

التطبيق	المورد اللغوي
التعرف على الحروف	أشكال الحروف
التعرف على الصوت	الأصوات العربية
منظومة تحليل اللغة (التحليل الصرفي، التحليل النحوي، التحليل الدلالي، التحليل المقامي)	شبكة الكلمات المدونات الموسومة نحويا ودلاليا ومقاميا. المحلل الصرفي المحلل النحوي المحلل الدلالي المحلل المقامي
تحويل الصوت إلى نص تحويل النص إلى صوت	الأصوات العربية

وفي ظل عدم وجود هذه الموارد الأساسية للبحث والتطوير لحوسبة اللغة العربية، ولارتفاع كلفة تطويرها، ولاحتفاظ الجهات التي تبني هذه الموارد بما أنتجته وعدم إتاحتها لاستخدام الباحثين والمطورين، فقد انصرفت معظم الأبحاث نحو تجريب تطبيق تقنيات حوسبة اللغة التي وصل إليها الغرب على جزئيات من اللغة العربية، دون وضعها في إطار يمكن تحقيقه عمليا لتصبح منتجات.

١-١٣ محتويات الكتاب

يحوي هذا التقرير أربعة فصول. الفصل الأول هو هذا التمهيد. والفصل الثاني هو الدراسات العربية في حوسبة اللغة العربية والفصل الثالث هو ترجمة خلاصات الأبحاث باللغة الإنكليزية والفصل الرابع يعرض الاستنتاجات والتوصيات التي خلصت لها الدراسة. يحوي الملحق الأول المصطلحات.

الفصل الثاني

واقع الدراسات العربية في حوسبة اللغة العربية

٢- ١- مقدمة

٢-٢-١- حوسبة اللغة العربية مقدمة تاريخية:

٢-٢-٢- الحوسبة والمستويات اللسانية للغة العربية:

٢-٣-١- الحوسبة والمستويات اللغوية الأساسية للغة العربية

٢-٣-٢- المستوى النحوي

٢-٣-٣- المستوى الصرفي

٢-٣-٤- الموارد

٢-٣-٥- التطبيقات

٢-٤- خاتمة

٢- ١ مقدمة

يهدف هذا الفصل إلى تقديم دراسة مسحية تظهرالجهود العربية المختصة بحوسبة اللغة العربية، وإذ تأتي في إطار تقديم دليل أبحاث حوسبة اللغة العربية المعني -أساسًا - بترجمة دراسات حوسبة اللغة العربية المنشورة باللغة الإنجليزية وعرضها، فإنها ستلتزم ببعض المحددات المنهجية الضرورية لتحقيق هدفها؛ ولعل أهم ذلك:

١. إظهار حجم هذه الدراسات وانتشارها؛ وذلك من خلال تتبع الدراسات العربية في مجال حوسبة اللغة العربية.

٢. تحديد بدايات الاهتمام بحوسبة باللغة العربية، ومجالاتها وتطورها إجمالاً.

٣. توزيع الدراسات على المستويات اللغوية وفق وجهة نظر توفق بين الاهتمامات اللسانية والحاسوبية.

٤. عرض ملخص لثلاث دراسات مميزة في كل مستوى من المستويات اللغوية، علمًا أن الاهتمام بالدراسات المبكرة -أحيانًا- يأتي من أجل إظهار بدايات الاهتمام بالمجال الحاسوبي أو المستوى اللساني في حوسبة اللغة العربية.

٥. عرض أبرز النتائج التي ظهرت في كل مستوى لساني أو مجال حاسوبي.

٦. تقديم خلاصة كلية تظهر أبرز النتائج والتوصيات التي خلص إليها الفصل.

وإذا كان هذا الفصل من الدراسة معنيًا بعرض الدراسات العربية في حوسبة اللغة، فإن وجهة النظر المؤسسة لهذا الجهد تنطلق من نظرة عامة لنظام الأنساق اللغوية في العربية التي تفسح المجال لمزيد من الاهتمام بحوسبة اللغة العربية، فاللغة العربية شأنها شأن اللغات الطبيعية قابلة لأن تخضع لنمط البنية الصوريّة التي تختزل أنسقة اللغة في وجوه تمثيلية في حدودها الدنيا، ثم ما يسبقها ويلحقها من مكونات لغوية في بُعديها الإفرادي والتركيبي، وهو ما يعني إمكانية ضبط اللغة ضبطًا رياضيًا يتجاوز التوصيف إلى القدرة على التنبؤ بما يحقق فكرة الذكاء الاصطناعي ليكون محاكيًا للعقل الإنساني، إن لم يكن متفوقًا عليه من جهتي سرعة الأداء ودقة التحليل.

إنّ طبيعة اللغة العربية قد ضبطت ضبطًا معياريًا ووصفيًا منذ ما يزيد على ألف عام من الزمان، وقد طال هذا الضبط بنية اللغة في مستويات عدة: الصوتية، والمعجمية، والتركيبية، والدلالية، وقد أظهر ذلك السعة التي تمتاز بها اللغة العربية من جهة عدد مفرداتها، وتنوع أنماطها التركيبية، وتداخل نتائجها الدلالية تبعًا لسياقات استخدامها، وهو ما جعل حوسبة اللغة العربية ليس بالأمر اليسير.

وعلى الرغم من ذلك فإن الإنجاز الذي حقق في حوسبة اللغة العربية لا يمكن إنكاره، ولعل ذلك جاء حصيلة جهود كبيرة بدأت منذ الثلث الأخير من القرن العشرين، ولعل إلقاء نظرة على البعد التاريخي لهذه لجهود كفيلة بالكشف عن أبرز ما قدم.

٢-٢-١- حوسبة اللغة العربية مقدمة تاريخية

لقد بدأ الاهتمام -عالميًا - بحوسبة اللغة منذ منتصف القرن العشرين مع ظهور الجيل الأول من الحاسوب، وتشير أغلب الدراسات المعنية بحوسبة اللغة العربية إلى أنّ بداية الاهتمام بحوسبة العربية قد بدأت بداية متواضعة - شأن وضع الحاسوب - منذ بداية السبعينيات من القرن العشرين، واقتصرت على الاستفادة من الحاسوب في إحصاء الجذور الثلاثية وغير الثلاثية للغة العربية كما وردت في بعض المعاجم مثل: "الصحاح" للجواهري، ولسان العرب لابن منظور، وتاج العروس للزبيدي [(١)].

ويشير نبيل علي إلى أربعة مراحل شهدتها جهود الدارسين في حوسبة اللغة العربية:

١. دخول الحاسوب لأول مرة إلى مؤسسة عربية، هي معهد التخطيط القومي في مصر، عام ١٩٦٢، واقتصر الأمر على طباعة الكتابة العربية، وذلك بإحلال الحروف العربية محل الحروف الإنجليزية.

١- العارف، عبد الرحمن، توظيف اللسانيات الحاسوبية في خدمة الدراسات اللغوية العربية، مجلة مجمع اللغة العربية الأردني، ع٧٣، ٢٠٠٧، (٤٩-٥٠).

٢. تصميم نظام آلي لاختيار شكل الحرف العربي تلقائيًا على يد دكتور سيد حيدر (باكستاني الأصل) في جامعة مونتريال عام ١٩٧٣.

٣. ظهور سلسلة من التطبيقات النمطية، في المدة ١٩٧٣-١٩٨٥، مثل: استخدام "العربية" في نظم قواعد البيانات واسترجاع المعلومات، وظهور برمجة عربية مثل "البيسك" و"اللوغو" العربيتين...

٤. نجاح إدارة البحوث والتطوير في شركة "العالمية" للبرامج بتطوير أول معالج (محلل ومولد) صرفي.

٥. ويعد كتاب نبيل علي "اللغة العربية والحاسوب" - الصادر سنة ١٩٨٨م - أول دراسة جادة تتجاوز الإحصاء إلى محاولة تحليل مستويات اللغة العربية الصوتية والتركيبية والدلالية من جهة صلتها بعلم الحاسوب.

٢-٢-٢- الحوسبة والمستويات اللسانية للغة العربية

تتعدد المستويات اللغوية التي تبحثها اللسانيات، ومن أهمها: المستوى الصوتي، فالصرفي، فالنحوي، فالدلالي. وإذا كان هذا التقسيم معهودًا في اللسانيات الحديثة فإن علم الحاسوب اللساني يعيد توزيع هذه المستويات وفق نظرته المختلفة منهجًا وأداء على الرغم من اشتراكه مع اللسانيات في المادة المبحوثة، وإذا كانت اللسانيات تتعامل مع اللغة تحليليًا وفق هذه المستويات فإن المعالجة الحاسوبية للغة تتجاوز التحليل إلى التركيب أو التوليد؛ ذلك أن علم الحاسوب اللساني يتعامل مع اللغة وفق مجال افتراضي يؤسس لتقنين اللغة آليًا، ومن ثَمّ فهو يبدي اهتمامه في المستويات اللسانية على أنها مجالات عمل حين يوضع "الحاسوب" مكان العقل الإنساني؛ ليدخل هذا الفرع المعرفي البيني في إطار ما يعرف باللسانيات التطبيقية على أساس أن اهتمام المجالين المعرفيين: اللساني والحاسوبي هو اللغة، مادة اللسانيات الأولى بامتياز.

من هذا المنطلق سيجمل هذا الفصل جهود الباحثين العرب في حوسبة اللغة العربية، ولا بد من توضيح النهج الذي سيتبع في إجمال ما قدموه، حيث سيركز على الأبحاث العلمية دون

المقالات التثقيفية العامة، وستنتظم هذه النظرة المُجملة لجهود الدراسين وفق التصور التنظيمي الآتي:

أولاً: الدراسات ذات الطابع الشمولي، التي اهتمت بالكشف عن هذا المجال المعرفي الجديد – على نحو ما– معتمدة على البحث اللساني العربي في محاولة استظهار أهميته في دراسة اللغة العربية، وما الآفاق الممكنة لخدمة اللغة العربية والدفع بها لتتلاءم مع مستجدات العصر، ومحاولة رسم الآليات الضرورية لتحقيق ذلك، والصعوبات التي تواجه الباحث في هذا المجال؟

ثانيًا: توزيع الدراسات وفق أبرز المستويات اللغوية المعروفة في الدرس اللساني، وهي: المستوى النحوي، والمستوى الصرفي. أما المستويات الأخرى: الصوتي والمعجمي والدلالي فستعرض من خلال وجهة نظر علم الحاسوب، إذ يدرج المستوى الصوتي وما يتبعه من رسم الحرف العربي في إطار الجانب التطبيقي، فكلاهما يتعلقان بالتعرف الآلي على الصوت العربي والكتابة العربية، وهو الجزء المشار إليه بالموارد. ويعرض المستوى الدلالي من خلال ما يعرف في علم الحاسوب (بالأنطولوجيا) وما يرتبط بها من عناية بالمصطلحات والمفاهيم وما يتعلق بهما من البحث الآلي عن المحتوى اللغوي على الشابكة، وما يعرف بالذخيرة اللغوية، وهو يصنف في علم الحاسوب اللساني تحت مفهوم "الأدوات".

ثالثًا: لمّا كان من المتعذر عرض وإجمال الدراسات جميعها، فقد كان لا بدّ من الاكتفاء بعرض أبرز هذه الدراسات على ألّا تقل عن ثلاث دراسات في كل مستوى من هذه المستويات.

رابعًا: إظهار تطور التفكير في هذا المجال المعرفي البيني؛ بحيث يظهر مسرد الدراسات حجمها وموضوعاتها تبعًا لظهورها تاريخيًا.

خامسًا: إظهار الفرق بين تعامل اللسانيين المعنيين بحوسبة اللغة مع الظاهرة اللغوية ومكوناتها، ومعالجة الحاسوبين المعنيين بالظاهرة ذاتها.

وأخيرًا عرض أبرز النتائج، والتوصيات.

٢-٢-٣- الدراسات ذات الطابع الشمولي

امتازت بعض الدراسات العربية في حوسبة اللغة العربية بالاهتمام بالقضايا الكلية والشمولية التي تبلور أهم ما يستلزمه تأسيس هذا المجال المعرفي وتطويره، وقد تنوع تناول هذا المبحث بين الحاسوبيين واللغويين، وكان فضل السبق فيه للحاسوبيين، على الصعيدين النوعي والتاريخي، فعلى الرغم من بدء الوعي بأهمية حوسبة اللغة كما ظهر لدى إبراهيم أنيس منذ مطلع السبعينيات من القرن الماضي[(١)]، إلا أنّ دراسة علي حلمي موسى "دراسة إحصائية لجذور معجم الصحاح (باستخدام الكمبيوتر)" (١٩٧١)، أول محاولة فعلية في توظيف الحاسوب في التعامل مع اللغة العربية[(٢)]، وإذا كانت هذه البداية متواضعة للغاية سواء على صعيد التنبيه إلى هذا المجال الجديد في دراسة اللغة، أو تطبيقه فعليًا؛ إذ انحصر في إطار الدراسة الإحصائية، لكن ذلك يتفق مع السياق العربي العام الثقافي والعلمي في بدء الاهتمام بالحاسوب، لكن مرحلة الثمانينيات أظهرت وعيًا جديدًا في علاقة اللغة بالحاسوب بدأه نبيل علي، ثم ظهر الاهتمام من عدد من اللسانيين أمثال مازن الوعر وعبد الرحمن الحاج صالح، ومحمد الحناش، ونهاد الموسى وغيرهم، وستعرض الدراسة الحالية ثلاثة نماذج من الدراسات الشمولية في تناول موضوع اللسانيات والحوسبة؛ دراستين للحاسوبيين، هما: "اللغة العربية والحاسوب "لنبيل علي، و"المعالجة الآلية للغة العربية المشاكل والحلول" لسلوى حمادة، ودراسة لسانية لنهاد الموسى.

١ - إبراهيم أنيس، "المحاسبات الإلكترونية في البحوث اللغوية"، مجلة المجمع المصري للثقافة العلمية، ع٤٢، ١٩٧٢، ص.

٢ - علي حلمي موسى، دراسة إحصائية لجذور معجم الصحاح (باستخدام الكمبيوتر)، الهيئة المصرية العامة للكتاب، ط٣، ٨٧٩١.

٢-٢-٣-١- اللغة العربية والحاسوب لنبيل علي، مؤسسة تعريب، الكويت، ١٩٨٨:

ليس في الإشارة إلى جهود نبيل علي المبكرة والمؤسسة لبحث حوسبة اللغة من جديد، فقد أجمع جلّ الدارسين إلى أهمية دراسته الأولى المميزة نوعيًا المنشورة سنة ١٩٨٨م، تحت عنوان: "اللغة العربية والحاسوب دراسة بحثية"، لقد كانت دراسة شمولية على نحو واضح، فقد جاء في التمهيد للدراسة قول نبيل علي: "إنّ المعالجة الآلية لا يمكنها إلا أن تتعامل مع الدقيق والمضبوط والمكتمل، لذا فهي تتطلب الكشف عن دخائل البنية الدفينة للغة العربية، وتحديد خصائصها ذات المغزى لأمور معالجتها آليًا، وهي مهمة عويصة - بلا شك - حيث عليها أن تتعامل مع اللغة على اتساعها: فقهًا وتنظيرًا وتأريخًا وتعليمًا واستخدامًا، وهو الأمر الذي فرض على هذه الدراسة أن تتحرك ما بين العموميات والشموليات أحيانًا، والتفصيلات والتفريعات أحيانًا"، وقد وضع نبيل علي اللبنة الأولى في تأسيس هذا الفرع المعرفي مبينًا أن دراسته ليست: دراسة في اللسانيات، أو في الحاسوب، بل هي دراسة حقل اللسانيات الحاسوبية مطبقة على اللغة العربية"، مدشنًا بذلك هذا المفهوم الهجين الذي يجمع بين مجالين معرفيين، وموظفًا خبرة طويلة في حقل المعلوماتية وهندسة الطيران على مدى عشرين عامًا، وعلى الرغم من اختصاصه بالمعلوماتية إلا أنه جعل اللغة منطلقه لا الحاسوب، وحقيقة يمثل ذلك حالة وعي غاية في الأهمية. وقد أعلن عن أن من دوافعه بدء الحوار بين اللغويين والحاسوبيين، وأن يكون للغوي في هذا الجهد الدور الريادي، ولعل هذا الدافع ما زال مطلبًا ملحًا حتى الساعة في ضرورة الاهتمام بالعلاقة البينية بين اللغة والحاسوب.

أما محتوى الدراسة فقد تشكل من مدخل وثمانية فصول:

١ – منظومة "اللغة العربية" من منظور الحاسوب.

٢ – منظومة الحاسوب من منظور اللغة العربية.

٣ – المعالجة الآلية للغة العربية.

٤ – المعالجة الآلية لمنظومة الكتابة العربية.

٥ – المعالجة الآلية لمنظومة الصرف العربي.

٦ – المعالجة الآلية لمنظومة النحو العربي.

٧– المعالجة الآلية للكلام العربي.

٨– ميكنة المعجم العربي.

بعد أن عرض نبيل علي تعدد عناصر المنظومة اللغوية وتنوعها، وبروز لا محدوديتها خاصة في السياقات النحوية والدلالية والمقامية، يقترح عددًا من الوسائل للسيطرة على لا محدوديتها، فيذكر منها:

١ – "تصنيف العدد غير المحدود من الكلمات في هيئة فصائل متمايزة (مثل أقسام الكلم، وأبواب الفعل، إلخ)، وتصنيف المكونات النحوية التي لا حصر لها في هيئة عدد محدود من المقولات (جملة تامة، شبه جملة اسمية، شبه جملة فعلية، شبه جملة وصفية، شبه جملة مجرورة، إلخ)".

٢– صوغ القرائن في شكل قواعد.

٣– تحويل الصيغ اللغوية إلى أنماط ونسق.

٤– تركيز القواعد في كتلة مضغوطة، أو عدد محدود من المبادئ والقيود.

٥– استخدام التعميم كأداة للاختزال والربط بهدف دمج الظواهر ذات الصلة في أطر موحدة.

٦ – رد الفروع إلى أصلها المشترك.

وقد ظهر على الدراسة سمة الدراسة التأسيسية، إذ بذل جهدًا واسعًا في محاولة إظهار الأهمية والعلاقة المتبادلة بين اللغة والحاسوب، وبين اللسانيات وعلم الحاسوب، وهو ما طبع الدراسة بطابعها الشمولي، وجعل منها منطلقًا وأرضية خصبة لما تلاها من دراسات التي اتكأت عليها اتكاء مباشرًا في كثير من الأحيان.

وقد خلصت الدراسة إلى أن اللسانيات الحاسوبية تتضمن المعالجات التالية: معالجة الكتابة العربية آليًا، ومعالجة الصرف العربي آليًا، وميكنة المعجم العربي، ومعالجة النحو العربي آليًا، ومعالجة الدلالة آليًا، ومعالجة السياق العربي المتصل آليًا. وقد نظر إلى طرق المعالجة هذه من جانبين: جانب تحليلي، وآخر تركيبي (توليدي)، ليشار إلى حالتي تحليل الحاسوب لبعض مستويات اللغة العربية ومباحثها، ومحاكاة إنتاج هذه المستويات والمباحث آليًا، لتكون هذه المعالجات سبيلاً لتطبيقات الحاسوب على اللغة العربية في تطبيقاتها المتعددة كما تتطلبها الحوسبة الآلية، من قضايا تتعلق بإدخال النص العربي، والتعرف إليه، ومعالجته، وما يمكن أن تسفر عنه من النظم الآلية للنشر المكتبي، وتعليم اللغة العربية بوساطة الحاسوب، والترجمة الآلية من اللغة العربية إلى اللغات الأخرى، ومن اللغات الأخرى إلى اللغة العربية. وانتهت الدراسة بإيراد قائمة ببحوث مقترحة في مجال اللسانيات الحاسوبية مطبقة على اللغة العربية.

٢-٢-٣-٢- العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية، نهاد الموسى، ط١، المؤسسة العربية للدراسات والنشر، بيروت، ٢٠٠٠.

تعد هذه الدراسة من أكثر دراسات اللسانيين شمولاً وجدية، وقد اعتمدت عليها جلّ دراسات اللسانيين في حوسبة اللغة العربية، وقد بدأ الكتاب بتقديم توصيف للعربية يتلاءم مع الحاسوب الذي يفتقر للكفاية اللغوية التي يمتلكها البشر، ويخلص من ذلك - في تقديمه - إلى تفرقة أساسية بين الوصف والتوصيف، فالوصف، الذي قام عليه النحو العربي، يعتمد على مقدرة الإنسان وحدسه في معرفة أوجه اللغة المقبولة والممكنة، وتوصيف اللغة يتخذ بعدين: كميًا ومنهجيًا، أما الكمي فيتعلق بالذاكرة الحافظة؛ ذلك أن ذاكرة الحاسوب تفوق الذاكرة

الفردية، وأما البعد المنهجي فيتعلق بالبعد الأول من جهة أنه يمثل سياقًا لازمًا لتحقيق الكفاية اللغوية.

يشتمل الكتاب على فصول عدة فضلا عن البيانات التمهيدية، وهي: نحو تمثيل النظم، نحو تمثيل النص، نحو تمثيل الإعراب، نحو تمثيل البنية، نحو الدليل على أخطاء النظم والإعراب والبنية، نحو تمثيل المعجم، نحو تمثيل المنطوق والمكتوب، وأخيرًا: مشكلات تمام التوصيف.

يقيم نهاد الموسى التفرقة بين الوصف والتوصيف بتحديد الوصف بما وقع للعرب من قواعد مستنبطة من الأداء اللغوي الواقعي، أما التوصيف فيقوم على الوصف مضافًا إليه كل الاحتمالات التي ليس بمقدور الحاسوب تعرّفها دون توصيف دقيق، ولا شك أنّ أول من أشار إلى مفهوم "التوصيف"، الذي يعد ركيزة أساسية فيما يقدمه نهاد الموسى ومن تبعه من اللسانيين، هو نبيل علي في كتابه: "اللغة العربية والحاسوب"[(١)].

وقد عرض نهاد الموسى البحوث التي قدمت في مجال حوسبة العربية، واهتم بترسيخ مصطلحاته التي تعينه في إقامة اللسانيات الحاسوبية في العربية، فاستخدم مصطلح "التمثيل" ليدل على ضبط اللغة في بناها المتعددة على هيئة ثابتة تُقدم للحاسوب ليستخدمها في منظومة رياضية، واستخدم مصطلح "العيار" لتعريف المادة وما تتضمنه من معطيات تفصيلية تُعرفه تعريفًا جامعًا لينتظم في عناصر صوتية وصرفية ونحوية ومعجمية.

١- نحو تمثيل النظم: يستخلص في هذا الفصل القواعد التي قررها النحاة، والهيئات التي تتحقق على أساسها الجملة العربي، فحاول استيعاب أنماط الجملة الفعلية والاسمية، وذلك بهدف توصيف أنماط التركيب العربي بكل تجلياته وتحققاته الأسلوبية الممكنة في العربية للحاسوب، وبدأ عمله في الجملة الفعلية من أبسط أشكالها المتمثل في النمط القائم على الفعل الماضي المجرد المثبت الصحيح اللازم، والفاعل.

ثم يعمل على مدّ الجملة من يسارها بإضافة عناصر ليست من أساس الجملة مما عدّ (فضلةً) من

١- (نبيل علي، ١٩٨٨، ١٤)..

ملحقات الجملة، ثم يبدأ بمدّ الجملة من اليمين لاستيعاب كل ما يعرض للجملة من أنماط خطية أفقية، مضيفًا للجملة كل ما تحتمله في نظام العربية من عناصر تسبق الفعل.

ثم يعمد – بعد ذلك – إلى جرد الأنماط التركيبية المحتملة والمنبثقة عن كل نمط تركيبي إضافي، مستهديًا بمرونة الإعراب في العربية، فيرصد الاحتمالات الممكنة...، دون أن يكتفي بالإمكانيات المقرّ بصحتها، فينبه إلى أنماط التركيب التي لا تجيزها العربية، ويتحفظ عليها، ويدل على لا نحويتها.

ويهتم – من جهة أخرى – بدلالة الفعل وأثرها في التركيب، فالأفعال الحسية كما في: (رأى، وسمع، ولمس...إلخ) غالبًا ما تقتضي فاعلاً ومفعولاً، وأفعال الاتجاه مثل: (ذهب، وارتحل، وسافر، وهاجر) غالبًا ما تقتضي شبه جملة (ظرفًا أو جارًّا ومجرورًا)، وأفعال التحويل كجعل وصيّر تقتضي مفعولين.

وقد اتبع النهج ذاته في توصيف الجملة الاسمية، فبدأ من الحدّ الأبسط لنمط الجملة الاسمية المكون من المبتدأ والخبر في صورتيهما الأبسط (الاسم المفرد)، ثم عمل على مدّ الجملة يمينًا ويسارًا كما فعل مع الجملة الفعلية.

٢- في الفصل الثاني الذي حمل عنوان: نحو تمثيل النص، انتقل من تمثيل الجملة البسيطة إلى تمثيل النص، وذلك بتوصيف الروابط المنطقية الدلالية، واللغوية الشكلية بين الجمل، متخذًا من (الواو) نموذجه في توصيف الروابط، فتتبع حالات ورودها أكانت عاطفة أم معية أم حالية...إلخ.

ثم وقف عند الحالات المحتملة تركيبيًا لوقوع الواو من غير أن تكون مقبولة، فتكون الواو مقحمة، ويكون ورودها من الأخطاء الشائعة.

٤ – تناول في الفصل الثالث تمثيل الإعراب: واعتمد التوصيف على تمثيل الإعراب للحاسوب اعتمادًا على نموذج هيكلي يرتكز على أربع مقولات، هي:

- نوع الكلمة أكانت اسمًا أم فعلاً أم حرفًا، وما يتبع كل صنف منها.
- الوظيفة، ويقصد بها المعاني النحوية أو الوظائف النحوية، مثل: مبتدأ، خبر، اسم

كان، فاعل...إلخ.

- الحالة الإعرابية: الرفع، النصب، الجرّ...
- علامة الإعراب، وتوّصف علامات الإعراب في العربية على أساس كونها أصلية أو فرعية أو علامة بناء.

٥- نحو تمثيل البنية: ويقدم فيه توصيف أبنية اللغة العربية واطرادها في الصرف، واعتمد في توصيفها على بعدين:

- تشكيلي تركيبي، ويتمثل في صوغ الأبنية بأدلة وخطوات إجرائية.
- تحليلي يتمثل في تعيين الأمثلة وتبيين معانيها الصرفية حسب سياق الجملة أو سياق النص الذي ترد فيه.
- ويشير نهاد الموسى - انسجامًا مع نهجه - إلى حالات اللبس مما قد يقع في دائرة المحمل غير المقبول في الأبنية اللغوية.
- وينبه إلى أخطاء النظم والإعراب والبنية، وهذه الأخطاء هي مما يقع للبشر في استخدام العربية، ومما عدّ من الأخطاء الشائعة.

٦- نحو تمثيل المعجم: يرى نهاد الموسى أن معجم الحاسوب ينبغي أن يكون متوفرًا على المعطيات المختلفة التي يعتمدها الحدس البشري، فينبغي على المعجم "أن ينتظم كل تشكيل لفظي ممكن تسمح به قواعد "وضع" الكلمة العربية سواء أكان مستعملاً أم لم يكن إلا ما لا تسمح به القواعد الفونولوجية من تتابع أصوات بأعيانها".

٧- نحو تمثيل المنطوق والمكتوب: يسعى في تمثيله للمنطوق والمكتوب حاسوبيًا إلى وضع عيار يضبط مواقع حدوث اللبس الصوتي كما يقع بين (سوط وصوت)، بما يؤدي إلى تمثل الحاسوب للفرق بينهما فيعرفه ويدل عليه.

ويرى أن المعجم أظهر أدلة التصحيح الإملائي بما قد دوّن فيه من هيئات رسم الكلمات، ولكنه يظل ناقصًا، ولا بدّ أن يُلجأ إلى دليل صوتي محدد يتمثل في بيان احتمالات

تشكيل الكلم وفقًا لقوانين الأصوات العربية (قيود التتابع)، وذلك بوضع جدول التتابع غير الممكن بين الأصوات (الحروف) ليكون أداة إضافية، على الرغم من أنه لا يكون إلا في الكلمة المفردة، مع الأخذ بعين الاعتبار إمكانية وقوعه في كلمتين متتابعتين.

وينهي كتابه بعرض ما يعترض تمام التوصيف من فوائت تحول دون محاكاته للملكة اللغوية عند العربي، هذه الفوائت هي: اللبس - تحولات المجاز - فضاء النص - صيرورة المعجم.

ولعل من أبرز ما يظهر قيمة هذه الدراسة هو شموليتها وتنويعها في التعامل مع المستويات اللغوية وتمثلاتها، ومحاولة توصيفها توصيفًا يقترب إلى حدٍّ ما من أن ينقل إلى عبارات منطقية تحاكي المنطق الرمزي الذي تتطلبه عملية الحوسبة، وعلى الرغم من ذلك فما زال هذا التوصيف يعتمد مصطلحات النحو العربي التقليدية، ويعتمد الأمثلة المصنوعة التي يسهل التعامل معها، ولم يصل إلى حدّ من التجريد القابل للتعميم، الأمر الذي يجعله قاصرًا حتى في حدود التوصيف اللساني للمعالجة الحاسوبية، فتوصيف الموسى ما زال يعتمد مبدأ الصواب والخطأ، ومبدأ القابلية الذي يستشعره البشر حتى وإن صح على المستوى التجريدي الحسابي.

٢-٢-٣-٣- المعالجة الآلية للغة العربية المشاكل والحلول لسلوى حمادة، دار غريب للطباعة والنشر والتوزيع، ٢٠٠٩.

جاء هذا الكتاب ليضطلع بدور قومي في خدمة اللغة العربية كما تشير الباحثة في مقدمة كتابها، وقد توزعت مباحثه على خمسة أبواب، اهتم الباب الأول بالمعجم وكيفية عمل معجم عصر المعلوماتية؛ المعجم اللغوي الحاسوبي للعربية، وقد عرضت فيه المشاكل وحلولها التي تواجه هذا العمل، بالكشف عن ظواهر الترادف والتضاد والاشتراك في العربية وكيفية التغلب عليها، واقترحت في الفصل الأول من هذا الباب منهجية لكيفية إثراء المعجم العربي تقوم على عمل معاجم فرعية تعمل معًا بالتنسيق الحاسوبي فيما بينها، كالمعجم الصرفي، ومعجم المترادفات والأضداد والمشترك اللفظي، ومعجم الأخطاء الشائعة، ومعجم التعبيرات المتداولة

في العربية، ومعجم المركبات الاسمية، ومعجم المركبات الفعلية، ومعجم التطور الدلالي للألفاظ، ومعجم الألفاظ العامية الدارجة.

وترى أن معجم عصر المعلومات يجب أن يتسم بـ:

١ - المحافظة على سلامة اللغة وجعلها وافية متماشية مع تقدم العلوم، والفنون، وسادة لحاجات العصر الحاضر.

٢- تبويب الألفاظ تبويبًا هرميًا كونيًا تذكر فيه المعاني متدرجة من الأصلي إلى الفرعي، ومن الحسي إلى المعنوي، ومن المألوف إلى الغريب.

٣- تحديد جذر اللفظ ومشتقاته، كما يعطي ألوان المعاني والمدلول الحقيقي والمدلول المجازي، وأن يحدد المحيط الخاص بكل لفظ كي لا يختلط بما سواه.

٤- أن يشمل جانبًا لغويًا يجمع بين القديم والحديث من ألفاظ اللغة بحيث توضع الألفاظ مرتبة ترتيبًا تاريخيًا يكون من السهل لمستخدمه تتبع تطور الألفاظ منذ أقدم العهود حتى تاريخ صدور المعجم.

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

وعلى الرغم من أهمية الجهد الذي قدمته سلوى إلا أن توصيفها اللغوي لظواهر اللغة العربية ومكوناتها يشوبها الخلط والقصور في مواقع كثيرة، من ذلك - مثلاً- حديثها عن أنواع المركبات حيث تقسمها من حيث التركيب إلى:

- مركبات نحوية عامة: مضاف ومضاف إليه، مثل: عين أحمد، يد الكرسي...، ويشوب هذا التقسيم عدم شرعيته من جهة تشكله ضمن منظومة التحليل اللغوي، فالتركيب الإضافي هو سمة لفظية يحلّ فيها المضاف إليه مكان التنوين أو ما هو في حكمه للفظة المفردة، ويماثله تراكيب لفظية أخرى مثل التركيب المزجي، وتركيب البناء كما في "أحدَ عشرَ".
- مركبات نحوية اصطلاحية: مضاف ومضاف إليه مثل قرن الشمس، تباشير الصباح، شرخ

الشباب، وهو توصيف قاصر لأن هذه التراكيب ليست اصطلاحية، وإنما هي مركبات لفظية ناجمة عن حالات استخدامها تعرف في الدراسات اللغوية بمصطلح "المتلازمات اللفظية".

- مركبات ثنائية التركيب: صفة وموصوف - جار ومجرور - مضاف ومضاف إليه، وما يشوب هذا القسم أيضًا أنّ هذه التراكيب جزء من فكرة المعاني النحوية التي تنتج هذا الشكل من التراكيب.
- مركبات ثلاثية التركيب أو أكثر: وتحددها الباحثة بأنها نفس المركبات السابقة لكن بتركيب متعدد، وهنا يظهر وجه الخلل في عدم الاهتمام بالنحو العربي كأساس البنية التركيبية في اللغة العربية، حيث يمكن الاستفادة من الدراسات اللغوية العربية في هذا المجال.

إن تسجيل هذه الملحوظات تظهر الحاجة الماسة لتجسير الهوة بين دراسات اللغويين ودراسات الحاسوبيين، وهو ما ألمحت إليه سلوى حمادة معللة ذلك بأنها لا تدعي الدقة في التعريفات والسمات، وأن بمقدور أهل اللغة تحري الدقة في ذلك، ولكنها مهتمة بالمنهجية التي تتبعها لإثراء المعجم العربي.

وجاء الباب الثاني في مجال حوسبة الصرف والاشتقاق، وذلك بالبحث في توليد الكلمات من خلال أوزان الكلمات واشتقاقاتها المختلفة.

واهتم الباب الثالث بمشكلات التحليل اللغوي للغة العربية، وركز على مشكلات اللغة العربية على الشابكة، وكيفية استخدام الحاسوب في التحليل اللغوي للنصوص، بالاستفادة من المعاجم المذكورة في البابين السابقين لحل مشكلة اللبس اللغوي من وجهة نظر حاسوبية تقوم على بناء قواعد بيانات لغوية صرفية ومعجمية عن كل ما هو مستعمل من أوزان الكلمات العربية وقواعدها في شكل رياضي رقمي، وبناء قاعدة بيانات للأعلام، وبناء قاعدة بيانات لأنماط الأخطاء الإملائية والنحوية الشائعة للنصوص، وبناء قاعدة بيانات لاحتمالات استعمال الكلمات العربية ومدى دورانها وشيوعها.

ونبهت -أيضًا - إلى طرق فك اللبس التركيبي في اللغة العربية اعتمادًا على بعض توصيفات

اللغويين مثل: الرتبة النحوية، وعوامل الربط، ودلالة السياق، والقرائن المعجمية.

وتقوم عملية فك اللبس اللغوي في المعالجة الحاسوبية باتباع الخطوات الآتية:

١- البحث في صور تحليل الجمل المعروفة، وعندما يتعرض التحليل إلى نوع من أنواع اللبس يقوم باختيار القيد المناسب مرتبطًا بالقيد الإحصائي لأعلى نسبة ورود، فإذا فشل يبدأ في بحث النسبة الأقل فالأقل.

٢- البدء أولاً بالقيد النحوي حيث يمكن أن يرفض كثيرًا من الحالات مما يقلل نسبة اللبس أو ينهيها.

٣- البحث من خلال السياق الداخلي، وإذا فشل يبحث من خلال السياق الخارجي.

٤- تطبيق بعض جزئيات من هذه القيود من الباحثة، لاستحالة إخضاع هذا العمل للعمل الفردي فقط، وجاءت النتائج مُرضية إلى حدٍّ ما، من خلال جمل مفردة سواء أكانت بسيطة أم مركبة.

٥- المطلوب لتعميم هذا العمل: إنشاء المعاجم بالصورة الموضحة بالهياكل مع مراعاة القيود الإحصائية التي تحتاج في حدّ ذاتها إلى فرق عمل متكاملة.

وعرضت في الباب الرابع مشكلات الترجمة الآلية، فعُرضت الجهود والمشكلات في هذا المجال عند غير العرب، ثم توضيح الحاجة الماسة للترجمة الآلية نتيجة حاجات السوق العالمي والسوق العربي لبرمجيات الترجمة من الإنجليزية إلى العربية، وانتهى بذكر بعض النظم والبرمجيات التي تسهم في الترجمة من الإنجليزية إلى العربية والعكس، وتناول - أيضًا - تقييمًا لبعض الترجمات الآلية في السوق العربية.

وانتهى الكتاب بمقترحات العمل المستقبلي، التي يمكن اختصارها بما يأتي:

١- إنشاء قاعدة بيانات نصية تخدم جميع مجالات الدراسة اللغوية.

٢- تدعيم المعجم العربي بسمات وخصائص جديدة.

٣- إمكانية تحليل النصوص وإعرابها آليًا، والعودة إلى التشكيل تدريجيًا.

٤- إمكانية الإجابة عن أي مسألة فيما بعد.

٥- سهولة الترجمة.

٦- تصحيح الأخطاء (الصرفية - النحوية - الدلالية).

٧- الارتقاء بالأساليب التعليمية للغة العربية.

٨- تحليل النصوص الأدبية وحفظ سماتها، مما يمكننا فيما بعد من توليد نصوص بالصياغة نفسها وبالأداء نفسه.

٩- قراءة الكتب آليًا، وعمل ملخصات لأهم النقاط فيها.

١٠- استعادة المعلومات عن طريق المحادثة باللغة الطبيعية مع الحاسب وشبكات المعلومات.

١١- ترجمة الكتب والمقالات من جميع اللغات.

١٢- إمكانية اختصار النصوص.

يمكن القول عما جاء في هذه الدراسة يمثل مرحلة مهمة في حوسبة اللغة العربية، لكنه مبني على تصورات مفترضة أكثر من أن تكون عملية تمامًا.

٢-٢-٣-٤- مسرد الدراسات الشمولية والعام في حوسبة اللغة العربية

الرقم	عنوان الدراسة	المؤلف	مكان النشر	تاريخ النشر
١	الثورة التكنولوجية واللغة	محمد صالح عمر	دار الشؤون الثقافة العامة -وزارة الثقافة والإعلام، بغداد-العراق	١٩٨٦
	استخدام اللغة العربية في علوم الحاسوب	أحمد الأخضر غزال	المجلة العربية للتربية، تونس، المجلد٦، العدد١	١٩٨٦م
٢	اللغة العربية والحاسوب	نبيل علي	مجلة عالم الفكر، م ١٨، ع ٣	١٩٨٧

	التقنيات الحديثة واللغة العربية	محمد ظافر الصواف	مجمع اللغة العربية الأردني	١٩٨٧م
٣	اللغة العربية والحاسوب	نبيل علي	مؤسسة تعريب، الكويت	١٩٨٨
٤	اللغة العربية والحاسوب (تعليق على بحث للدكتور نبيل علي)	قاسم السارة	مجلة عالم الفكر، م١٩، ع٢	١٩٨٨
٥	آخر التطورات في مجال تقييس تعريب الحاسوب	أحمد أبو الهيجاء	مؤتمر الكويت الأول للحاسوب -جمعية الحاسب الآلي الكويتية	١٩٨٩
٦	معالجة اللغة العربية بالحاسوب	محمد عبد المنعم حشيش	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	١٩٩٢
٧	استخدام اللغة العربية في تقنية المعلومات (تسجيل مرئي)	إعداد عبد الرحمن الحاج صالح؛ تقديم محمد بن عبد الرحمن الربيع	إنتاج مكتبة الملك عبد العزيز العامة. بيانات النشر الرياض : مكتبة الملك عبد العزيز العامة	١٩٩٢
٨	بحوث ودراسات في اللسانيات الحاسوبية	عبد الرحمن الحاج صالح	موفم للنشر - الجزائر	د. ت
٩	اللسانيات وبرمجة اللغة العربية	محمد الزركان	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات- مكتبة الملك عبد العزيز - الرياض	١٩٩٣

١٠	استخدام اللغة العربية وتقنية المعلومات	إعداد محمد الحناش	مجلة التواصل اللساني. المجلد الأول	١٩٩٣
١١	العرب وعصر المعلومات – الفصل التاسع	نبيل علي	سلسلة عالم المعرفة – الكويت	١٩٩٤
١٢	الحاسوب واللغة العربية	عبد ذياب العجيلي	جامعة اليرموك الأردن	١٩٩٦
١٣	اللغة العربية والتقنيات المعلوماتية المتقدمة	محمد الحناش	وقائع المؤتمر الدولي الثاني، مجلة التواصل اللساني. المجلد الثالث	١٩٩٦
١٤	استخدام اللغة العربية في المعلوماتية	مجموعة من المؤلفين	المنظمة العربية للتربية والثقافة والعلوم تونس	١٩٩٦
١٥	الحوار إنسان – آلة	نورالدين اللوز	مجلة التواصل اللساني – مؤسسة العرفان للاستشارات التربوية والتطوير المهني – المغرب	١٩٩٦
١٦	التقنيات الحديثة وآفاقية اللغة العربية	إسحق علي حبيبي	الموسم الثقافي الرابع عشر لمجمع اللغة العربية الأردني	١٩٩٦
١٧	المشروع اللساني الحاسوبي العربي	محمد الدهبي	مجلة التواصل اللساني – مؤسسة العرفان للإستشاراتالتربوية والتطوير المهني – المغرب	١٩٩٦
١٨	اللسانيات الحاسوبية مفهومها وتطورها ومجالات تطبيقها	بلقاسم اليوبي	مجلة مكناسة ع١٢	١٩٩٩
١٩	العربية نحو توصيف جديد في ضوء اللسانيات الحاسوبية	نهاد الموسى	المؤسسة العربية للدراسات والنشر، بيروت	٢٠٠٠

٢٠	تهيئة اللغة العربية لمواجهة طوفان المعلومات والعولمة	سلوى حمادة السيد	مؤتمر لسان العرب بمقر جامعة الدول العربية القاهرة	٢٠٠٠
٢١	اللغة العربية والحاسوب	هادي نهر	مجلة التواصل (جامعة عدن) العدد ٤،	٢٠٠٠
٢٢	حوسبة التراث العربي	علي حلمي موسى	محاضرة ألقيت في مجمع اللغة العربي الأردني في ١٧ نيسان (أبريل)	٢٠٠١
٢٣	اللغة العربية والحاسوب (قراءة سريعة في الهندسة اللسانية العربية) أو مقاربة في محاكاة الدماغ العربي لغويا	محمد الحناش	محاضرة بكلية العلوم الإنسانية والاجتماعية ، قسم اللغة العربية وآدابها جامعة الإمارات العربية المتحدة. مؤسسة العرفان للاستشارات التربوية والتطوير المهني	٢٠٠٢
٢٤	التجارب الراهنة حول حوسبة النصوص التي تعتمد اللغة العربية	موسى زمولي	مجلة اللغة العربية –العدد السابع الجزائر	٢٠٠٢
٢٥	تحليل اللغة العربية بواسطة الحاسب الآلي	سعد بن هادي القحطاني	جلة علوم اللغة ، م ٥، ع ٣، (القاهرة : دار غريب.)	٢٠٠٢
٢٦	قضايا اللغة العربية في عصر الحوسبة والعولمة	نبيل علي	مجمع اللغة العربية الأردني	٢٠٠٢
٢٧	الصياغة المنطقية الخليلية وفق القواعد التوحيدية	عشيت عبد المجيد	مجمع اللغة العربية الأردني	٢٠٠٢

٢٨	دور النظرية الخليلية الحديثة في النهوض بالبحوث الحاسوبية الخاصة باللغة العربية	عبد الرحمن الحاج صالح	مجمع اللغة العربية الأردني	٢٠٠٢
٢٩	اللغة العربية والحاسوب قراءة سريعة في الهندسة اللسانية العربية	محمد الحناش	مجلة التواصل اللساني مج٩	٢٠٠٣
٣٠	اللسانيات والحاسوب واللغة العربية	مازن الوعر	صحيفة رؤى ثقافية (سورية) العدد ٤	٢٠٠٣
٣١	تمثيل الكفاية اللغوية للحاسوب	نهاد الموسى	ورقة مقدمة إلى ندوة الهوية اللغوية والعولمة جامعة البترا الأردنية الخاصة	٢٠٠٣
٣٢	الثنائيات في قضايا اللغة العربية المعاصرة الفصل من الوصف إلى التوصيف.. مقاربة في حوسبة العربية	نهاد الموسى		٢٠٠٣
٣٣	اللسانيات: المجال والوظيفة والمنهج الفصل السابع	سمير استيتية	عالم الكتب الحديث، الأردن	٢٠٠٥
٣٤	اللسانيات الحاسوبية العربية...المفهوم، التطبيقات، الجدوى	وليد العناتي	مجلة الزرقاء للبحوث والدراسات، عمادة البحث العلمي بجامعة الزرقاء الأهلية الأردنية، المجلد السابع، العدد الثاني	٢٠٠٥

٣٥	اللسانيات الحاسوبية واللغة العربية"	عز الدين غازي	الحوار المتمدن – ع١٦٣٩	٢٠٠٦
٣٦	الفجوة الرقمية في اللغة العربية	عبد المجيد نصير	مجمع اللغة العربية الأردني موسمه الثقافي الرابع والعشرين	٢٠٠٦
	اللسانيات الحاسوبية العربية رؤية ثقافية	وليد العناتي	مجلة فكر ونقد – المغرب	٢٠٠٦
٣٧	الحاسوب وميكنة اللغة العربية	رأفت الكمار	القاهرة : دار الكتب العلمية للنشر والتوزيع	٢٠٠٧
٣٨	دليل الباحث إلى اللسانيات الحاسوبية العربية	وليد العناتي وخالد الجبر	دار جرير للنشر والتوزيع عمان، الأردن	٢٠٠٧
٣٩	توظيف اللسانيات الحاسوبية في خدمة الدراسات العربية : جهود ونتائج	عبد الرحمن بن حسن العارف	مجلة مجمع اللغة الأردني، ع.٧٣	٢٠٠٧
٤٠	اللغة العربية والحاسوب	ندى غنيم وأميمة الدكاك	الاحتماع الثاني لخبراء المعجم الحاسوبي التفاعلي– الرياض	٢٠٠٨
٤١	اللُّغة العربيّة والشّبكة العنكبوتيّة / قضايا وحلول	أحمد حياصات	مجمع اللغة الأردني	٢٠٠٨
٤٢	المعالجة الآلية للغة العربية (المشاكل والحلول)	سلوى حماده	دار غريب للطباعة والنشر والتوزيع	٢٠٠٩
٤٣	أساسيات الحاسب المعالجة الآلية للغة العربية	شريف عصام خطاب		٢٠٠٩

٤٤	فعاليّة اللسانيات الحاسوبية العربية	ديدوح عمر	الأثر-مجلة الآداب واللغات، جامعة قاصدي مرباح، الجزائر، ع٨	٢٠٠٩
٤٥	اللسانيات الحاسوبية العربية: الإطار والمنهج	وجدان محمد صالح كنالي	المؤتمر الدولي الثاني للغة العربية - دبي	٢٠١٣
٤٦	الفجوة الرقمية وتعليم اللغة العربية الواقع والمأمول	إبراهيم صلاحالهدهد	نشر في موقع مركز الشيخ زايد لتعليم اللغة العربية لغير الناطقين بها	٢٠١٣
٤٧	اللغة العربية ومجتمع المعرفة	محمد الحناش	محاضرة في جمعية النبراس للثقافة والتنمية - وجدة - المغرب	٢٠١٣
٤٨	اللسانيات الحاسوبية العربية: الإطار والمنهج	وجدان محمد صالح كنالي	المؤتمر الدولي الثاني للغة العربية- دبي	٢٠١٣
٤٩	اللغة العربية في عصر الحوسبة والمعلوماتية: مقاربة من الهيكل العام لأنظمة العربية	صالح أبو صيني	حوليات جمعية كليات الآداب - الأردن	٢٠٠٤
٥٠	حوسبة اللغات وبعض إشكالياتها: العربية أنموذجًا	حسن كون	مجلة جيل الدراسات الأدبية و الفكرية	٢٠١٥
٥١	قراءة في الإسهامات اللسانية الحاسوبية العربية : آفاق ورهانات	فاهم سعيد	دراسات -الجزائرع٣٦	٢٠١٥
٥٢	اللسانيات الحاسوبية مشكل المصطلح والترجمة	أحمد رضا بابا	مخبر المعالجة الآلية للغة العربية جامعة تلمسان	د.ت

٥٣	اللسانيات الحاسوبية العربية	عصام محمود	الإسكندرية، مصر : دار الوفاء لدنيا الطباعة والنشر	٢٠١٥
٥٤	المشاريع الحاسوبية على اللغة العربية والقرآن بجامعة ليدز	عبد الباقي شرف وآخرون		
٥٥	تقنيات اللغة العربية الحاسوبية – معايير التقييم ورؤى التطوير	عمرو جمعة	مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية	٢٠١٦
٥٦	اللسانيات الحاسوبية: رقمنة اللغة العربية ورهان مجتمع المعرفة	إبراهيم مهديوي	شبكة الألوكة	٢٠١٦
	التحديات التي تواجه حوسبة اللغة العربية وبعض الحلول المقترحة	محمد عدني السيد وإريك اتول	الدورة العاشر للمؤتمر الدولي لعلوم وهندسة الحاسوب (ايكا ICCA)	٢٠١٦
٥٧	محاضرات في اللسانيات الحاسوبية	بن عريبة راضية	القاهرة : المجموعة العربية للتدريب والنشر،	٢٠١٧
٥٨	اللسانيات الحاسوبية في المجال الأكاديمي المغربي رفع اللبس وتحديد المفهوم	سالم الرامي	معهد الدراسات والأبحاث للتعريب بالرباط	د. ت
٥٩	اللسانيات الحاسوبية: مشكل المصطلح والترجمة	رضا بابا أحمد	مخبر المعالجة الآلية للغة العربية – جامعة تلمسان)الجزائر	د. ت
	اللغويات الحاسوبية في المغرب: دراسة في الجيل الأول	عمر مهديوي	اللسانيات الحاسوبية واللغة العربية إشكالات وحلول – كنوز المعرفة – الأردن	٢٠١٨

٢-٢-٣-٥- خلاصة:

تكشف الدراسات ذات الاهتمام الشمولي عن إدراك الباحثين العرب ضرورة الاهتمام بحوسبة اللغة العربية بعدّ ذلك أمرًا لا مناص منه لمواكبة التطور العلمي، ومتابعة السياق الحضاري العالمي المبني على معطيات التكنولوجيا، ومن الواضح أن إحساس الحاسوبيين واللغويين يرجع إلى مرحلة واحدة تقريبًا؛ هي منتصف الثمانينيات من القرن الماضي، وأن كليهما تعامل مع هذه الحالة برؤية واحدة تشعر بالتمسك باللغة العربية بما تمثله من قيمة للأمة في ماضيها وحاضرها ومستقبلها، وهو ما يظهر من الإشارات الكثيرة الواردة فيها إلى نصوص دينية، أو تراثية، أو تقدير لنظام اللغة العربية الفصيحة دون الاهتمام باللهجات الدارجة منها.

لكن الاختلاف يتجلى في منهج التعامل مع الظاهرة اللغوية، وهو ما يمكن ملاحظته في طريقة التعامل مع عناصر الظاهرة، فما يبحث على أنه إحدى مكونات اللغة البنيوية التي يمكن معالجتها وفق تصميمات علوم اللغة المعروفة، وجد اهتمامًا مغايرًا على صعيد المنهج في دراسات الحاسوبيين، على الرغم من أن بداية الاهتمام في دراسات الحاسوبيين لم تخرج عن الأطر القواعدية التي ضمنها تاريخ طويل من جهود علماء اللغة، وهو ما سيظهر تاليًا ظهورًا جليًا للعيان.

٢-٣-١- الحوسبة والمستويات اللغوية الأساسية للغة العربية

تتعامل اللسانيات مع الظاهرة اللغوية بتحليلها إلى مستويات ليسهل وصفها وتجريد قواعدها الضابطة؛ بما يكفل فهم الظاهرة وتأطيرها علميًا، ولعل أبرز مستويين يمثلان نظام اللسان في اللغة العربية هما: المستوى النحوي، والمستوى الصرفي.

٢-٣-٢- المستوى النحوي

يشكل المستوى النحوي أبرز ما يمثل جانب القواعدية للغة العربية، ويشكل مفهوم الجملة أساس هذا المبحث، وقد استطاع الدرس النحوي قديمًا ضبط هذا المفهوم وتوصيفه

توصيفًا دقيقًا؛ فكان أساس الجملة يعتمد على مبدأ "الإفادة"، لتكون الجملة أساس أصغر وحدة تداولية في عملية التواصل اللغوي، وكان مفهوم "الإسناد" هو ما يجسد الجانب الشكلي القادر على تحقيق مبدأ "الإفادة" هذا، وكانت "نظرية العامل" من أهم ما يبلور الفهم العملي والتحليلي القائم على تصور منطقي لمفهوم الإسناد، وقد امتاز نظام اللغة العربية بوجود علامات صوتية تظهر موقع المفردة اللغوية ووظيفتها في الجملة.

ويعد هذا المستوى من أهم المباحث التي تستلزمها عملية حوسبة اللغة العربية، فهو يمثل أساس عملية التواصل اللغوي، ومن ثم يصبح ذا ضرورة لكثير من العمليات الحاسوبية كالترجمة الآلية، والتصحيح اللغوي، والتدقيق اللغوي، والقدرة على توليد الجمل بما يحقق حالة التنبؤ اللغوي.

ومن الملحوظ على الدراسات التي عُنيت بالمستوى النحوي في حوسبة اللغة اهتمامها بمباحث جزئية دون الاهتمام بتوصيف كامل يجسد أساس منطق نظام اللغة العربية في حالة من التجريد الكلي، فقد أظهرت الدراسات ذات النزوع الشمولي كما في دراسة نبيل علي "اللغة العربية والحاسوب" ودراسة نهاد الموسى "نحو توصيف جديد للعربية في ضوء اللسانيات الحاسوبية" جوانب مهمة في حوسبة النظام التركيبي النحوي للغة العربية، لكن ذلك لم يكن كافيًا، واهتمام الدراسات بمباحث محددة مثل التركيب الإضافي أو قواعد العدد لا يمكن أن يقدم تصورًا جامعًا يمكنه أن يتفرع إلى مباحثه الجزئية، ومن ثمّ كان لا بد من الاهتمام بدرجة عالية من التجريد في منطق اللغة العربية التركيبي على صعيد الجملة بما يمكن تمثيله بمنطق رياضي مجرد، بحيث يمثل قواعدية النحو الكلي، وقد كان بالإمكان الاستفادة من بعض التصورات اللغوية العربية القديمة، ومن ذلك مفهوم "التعليق" في نظرية النظم عند عبد القاهر الجرجاني؛ ذلك أن هذا المفهوم أقيم على توضيح العلاقة بين اللغة والذهن، فارتباط أي لفظة بأخرى يقوم على أساس تصور ذهني، وعلى أساس ذلك يجرد الجرجاني أساس نظام التركيب النحوي في اللغة العربية على أساس أربعة أنماط تركيبية: تعليق اسم باسم، وتعليق اسم بفعل، وتعليق فعل بفعل، وتعليق حرف بهما، وتقوم العلاقة بينها على أساس ما يسميه

"المعاني النحوية" كالفاعلية والمفعولية...إلخ، ثم يفرع من هذه العلاقات المجردة ما يقع تحتها من المباحث النحوية، وقيمة هذا التوصيف أنه يسبق مفهوم الإسناد الذي يمكن أن يشكل الخطوة الثانية لتحديد العامل في الجملة، ثم يتلوه تحديد المعمولات والتوابع.

من منطلق الرؤية السابقة فقد جاء اختيار الدراسات التي ستعرض هنا على أساس مدى اقترابها من تقديم توصيف كلي مجرد، ولذلك اختيرت ثلاثة أعمال لعرضها، هي:

٢-٣-٢-١- نظام محوسب لمحلل نحوي في اللغة العربية لجمل فعلية غير مشكولة من الفعل الماضي المبني للمعلوم، تمارى القبلان، رسالة ماجستير، جامعة آل البيت، الأردن، ٢٠٠٤.

هدفت الدراسة إلى تقديم تحليل نحوي لجمل فعلية غير مشكولة من الفعل الماضي المبني للمعلوم، وذلك ببناء برنامج حاسوبي يؤدي عمل المحلل النحوي لهذا النمط من الجمل، وقد بني معجمان بسيطان للغة العربية؛ يحتوي الأول على عينة من الكلمات مثل الأسماء والأفعال والضمائر والحروف وما يتبعها من سمات وخصائص تساعد في معرفة موقع الكلمة من الإعراب، وتوفير البيئة الملائمة لإجراء التحليل النحوي، ويحتوي الثاني محللاً صرفيًا بسيطًا يقوم بتحديد السوابق واللواحق المرتبطة بالكلمة ونوعها لتربط - بعد ذلك - بالمعجم لمعرفة صفاتها.

تكونت الدراسة من خمسة فصول، الفصل الأول نوقشت فيه أهمية الدراسة وأهدافها والتحديات التي تواجه معالجة النحو العربي حاسوبيًا. واهتم الفصل الثاني بعرض الدراسات السابقة التي اعتنت بمعالجة اللغة العربية آليًا، وموضوع التحليل النحوي. وناقش الفصل الثالث مفهوم اللسانيات الحاسوبية في مستويات التحليل الأربعة: المعجمي والصرفي والنحوي والدلالي، أما الفصل الرابع فتناول الجملة الفعلية غير المشكولة من الفعل الماضي المبني للمعلوم والتقسيمات المقترحة والقواعد التركيبية المقترحة، ومجموع التراكيب التفصيلية لحالات الجملة الفعلية غير المشكولة المحتملة من الفعل الماضي المبني للمعلوم، ومجموعة قواعد المعرفة المعتمد

عليها في التحليل النحوي. وعرض الفصل الخامس واجهة البرنامج وكيفية استخدامه، ومجموعة الجمل التي اختبرها البرنامج، والنتائج التي انتهت إليها الدراسة.

وقد أشارت الدراسة إلى نجاح البرنامج في إعراب ٨٥٪ من عينة الجمل المدخلة إعرابًا تامًا وصحيحًا، و١٥٪ بإعراب غير تام؛ ويعزى عدم النجاح إلى الحاجة إلى محلل دلالي ومحلل صرفي، وهما خارج إطار الدراسة.

وتتبدى أهمية هذه الدراسة بأمرين: اهتمامها بتأسيس تصور كلي قبل الولوج في موضوعها الخاص، واعتمادها على التطبيق العملي والتجريب في المعالجة الآلية لنحو اللغة العربية في إطار محدد هو من أكثر الأطر التي يعتمد عليها في الجملة العربية؛ الجملة الفعلية المكونة من الفعل الماضي المبني للمعلوم، فضلاً عن تقديمها عددًا من التوصيات لاستكمال ما قدم مثل: إجراء التحليل الدلالي للجمل الفعلية غير المشكولة من الفعل الماضي المبني للمعلوم لمساندة التحليل النحوي، وتوسيع قواعد المعرفة المقترحة لإجراء التحليل النحوي للجملة الفعلية من الفعل الماضي المبني للمجهول، وتوسيع قواعد المعرفة المقترحة وقواعد إعادة الكتابة المقترحة لإجراء التحليل النحوي للجملة الفعلية من الفعل المضارع وفعل الأمر، ودراسة الحالات التي يتقدم بها المفعول به على الفعل والفاعل، ودراسة الممنوع من الصرف وكيفية إعطاء الكلمة الممنوعة من الصرف الحركة الإعرابية التي تناسب موقعها الإعرابي في الجملة.

٢-٣-٢-٢-توليد الجمل في اللسان العربي (دراسة لسانية حاسوبية)، بابا أحمد رضا، وهي رسالة دكتوراه قدمت في جامعة أبي بكر بلقايد - تلمسان في الجزائر سنة ٢٠١٤.

بدأت الدراسة بتعريف المفاهيم الأساسية التي تتطلبها عملية توليد الجمل في العربية، بدءًا بالجملة واللسان العربي، ثم أُسس الدراسة الحاسوبية وموضوعها ومنهجها، وبما أن عملية توليد الجمل هي عملية لسانية ولسانية نفسية كذلك؛ فقد عرض الباحث نماذج التوليد من المنظورين اللساني واللساني النفسي، وبنى فكرة التوليد الآلي للجمل في العربية على أساس

أن الجانب اللساني الحاسوبي واللساني النفسي يتقاطعان في عملية التوليد الآلي للنصوص، وقدم تصورًا لإنشاء برامج حاسوبية لتوليد الجمل العربية باستخدام لغة البرولوغ.

أشارت الدراسة إلى توصيف الجملة العربية من وجهة نظر بنيوية، فعدّ الجملة "قطعة من السلسلة الكلامية مستقلة من الناحية التركيبية؛ أي هي شكل لساني مستقل غير متضمن في شكل لساني أكبر بموجب بنية نحوية"، ثم من وجهة نظر التوليديين، فالجملة "هي أولية قاعدية تتفرع عنها سلسلة من الرموز بواسطة قواعد تركيبية"، ويشير – بعد ذلك – إلى ارتباط الجملة في بنائها بفكرة العامل ارتباطًا وثيقًا؛ فهو الذي يقوم بوصل مكوناتها ويجعلها نسيجًا متلاحمًا، شكل معها وحدة معنوية وبناء متكاملاً.

لقد بدا على الدراسة هذا الاهتمام الموسع في الإحاطة بمفهوم الجملة في اللغة العربية، وذلك نتيجة الإحساس بضرورة أن تهتم الدراسات اللسانية المعمقة والدراسات الخاصة بالمعالجة الآلية ببلورة نحو صوري للغة العربية، غير أن هناك مشاكل عدة تواجه الباحثين في هذه العملية كما يشير الباحث، منها قلة الموارد والمدونات والذخائر اللغوية المخزنة في شكلها الرقمي التي يعتمد عليها في هذا الشكل من الدراسات، "وكذا الإطار الصوري الذي ينبغي اتباعه لإنجاز تلك العملية".

ويرى الباحث أن النمذجة الحاسوبية هي أهم مميزات اللسانيات الحاسوبية، حيث تهدف النمذجة إلى وصف السيرورات التي يمكن حسابها، التي تركب فيها اللغة وتحلل، بالإضافة إلى تحليل المشكلات في مستواها الحاسوبي، ويكون ذلك بصياغة نماذج وأنظمة عمليات تحاكي بنيتها العلائقية سيرورة معينة كسيرورة اللغة.

ويرى أن المعالجة الآلية للغات هي حقل بحث متعدد الاختصاصات يساهم فيه لسانيون وحاسوبيون ومناطقة ونفسانيون ووثائقيون ومعجميون ومترجمون، وهو يندرج في ميدان الذكاء الاصطناعي.

من الواضح أن الدراسة ترتكز على بعدين أساسيين: الأول تقديم توصيف شمولي لنظام الجملة في اللغة العربية، أو ما سماه "النحو الكلي وهو عبارة عن مجموعة من القواعد التركيبية"،

ويندرج تحت التحليل اللغوي، والثاني: القدرة على إنتاج الجمل العربية، وهو ما يندرج تحت التوليد اللغوي. وهما البعدان اللذان أشار إليهما نبيل علي في دراسته الرائدة "اللغة العربية والحاسوب".

اعتمد الباحث على أن يكون نظام الجملة العربية على شكل: "عامل معمول١ معمول٢ مخصص"، وذلك لتتوافق مع "لغة برولوغ" لمحاكاة بعض الأنشطة اللغوية، وتشغيلها ضمن برامج حاسوبية، وحصر تطبيقه في نمطي الجملتين الفعلية والاسمية في صورتهما البسيطة.

وبناء على ذلك قام بصياغة برامج حاسوبية تعمل على تكوين جمل عربية تكوينًا آليًا، وتنطلق مما زود به برنامج برولوغ من وقائع وقواعد تتعلق بالمطابقة بين وحدات الجمل العربية، فيقوم البرنامج بتكوين جمل صحيحة التركيب سليمة المعنى، إلا أن بعضها بدا غير مقبول عند المتكلمين العرب، وهو ما يؤمل تطويره في المستقبل بإضافة بعض القيود المتعلقة بالجوانب الثقافية والدينية لتجنب توليد جمل صحيحة شكلاً، ولكنها غير مقبولة دلاليًا أو تداوليًا.

واستنتج – على الصعيد اللغوي – أن دراسة ظاهرة المطابقة بين الفعل وفاعله في الجملة الفعلية، وبين المبتدأ والفعل الواقع خبرًا له في الجملة الاسمية الكبرى وتشغيلها حاسوبيًا، تؤكد نظرية النحاة التي تتضمن عدم تماثل هاتين البنيتين من حيث الشكل، وبعض اللسانيين المحدثين من أن الفاعل يمكن تقديمه على فعله من غير تغيير في بنية الجملة.

٢-٣-٢-٣- اللغة الداخلية وحوسبة البنية النحوية، سرور اللحياني، وقائع الندوة العلمية الدولية الثالثة للسانيات وإعادة البناء، كلية الآداب والفنون والإنسانيات بمنوبة، تونس، ٢٠١٤.

استندت هذه الدراسة إلى فرضية مفادها أن البنية النحوية في شكلها البسيط أو المركب عملية حركية قصدية منتظمة تتعالق وتتمازج في أشكال تكرارية تختزن المعنى في كل عملية نظم لطور من أطوار اشتقاق البنية. وتتعين خصائصها من خلال البحث في كيفية اشتغال اللغة الداخلية باعتبارها فضاء ذهنيًا مجردًا تتحكم فيه علاقات تركيبية داخلية.

تتفرع عن هذه الفرضية فرضيتان: الأولى تعتبر أن اللغة الداخلية تفسر تفاعل النظام الحسي- الحركي مع النظام الإدراكي القصدي والتواجه بين الصورة الصوتميّة والصورة المنطقية. فتتجلى البنية التركيبية بالبنية الدلالية بحثًا في كيفية انتظام هذه الأنظمة في الذهن.

الفرضية الثانية تعتبر أن البحث في هندسة البنية النحوية وحوسبتها يمكّن من تفسير خصائصها التصورية ومساءلة طبيعة الجهاز اللغوي وأنظمة اشتغاله لتعيق البحث فيما وراء الكفاية التفسيرية.

ولحل الإشكالية التي تواجه فرضيات البحث المتمثلة في: كيف يُدرك القصد والمضمون الدلالي في التركيب؟ فقد سار البحث في هذه الإشكالية شيئًا فشيئًا ليؤكد تواجه اللفظ والمعنى، فتجلى هذا التواجه بين الصورة الصوتمية والصورة الدلالية بحثًا في مدى دلالة المبنى على المعنى وفي إمكانية اشتقاق المعنى من المبنى، ويهدف البحث - بناء على ذلك - بتعريف طبيعة الجهاز اللغوي من خلال البحث في إمكانية حوسبة اللبس لأن التركيب الواحد يمكن أن يقبل أكثر من تأويل.

ويؤسس الباحث تصوره عن اللغة الداخلية استنادًا إلى تصور تشومسكي عن الملكة اللغوية لاختبار قدرة المنوال الأدنوي على تحقيق الكفاية التفسيرية؛ إذ عدتْ ممثلة للحالة الأولى التي توافق (النحو الكلي)، وهي متماثلة لدى البشر مرّت في تطورها الطبيعي بعدة أحوال متعاقبة لتصل إلى حالة ثانية تتميز بكونها ثريّة بما يكفي لتصف كل الألسنة الخاصة، وافترض الباحث أن الملكة اللغوية تتوفر على ثلاثة أنظمة متشارطة هي الداخلي والفردي والقصدي، فيعالج النظام الداخلي الحالة الذهنية للمتكلم باعتباره منجز الخطاب، ويتعلق الذاتي بالخصائص الاشتقاقية للسان مخصوص، أما القصدي فيعيّن قصد المتكلم من إنتاج الخطاب.

وانتهى البحث إلى تحديد مقاييس حوسبة البنية النحوية، وهي تتعلق ببحثين متعالقين: الأول هندسة البنية النحوية المجردة باعتبارها ممثلة لتصميم أمثل للغة يختزن كل المحلاّت الوظيفية والمعجمية، والثاني اشتقاق هذه البنية الذي يشترط الجدوى والاقتصاد.، وكانت على النحو الآتي:

١- المقياس (١): تتشكل البنية النحوية المجردة من محلاّت قابلة للتعجيم، مثل:

- محل المركب الفعلي المصدري يقبل رأسه التعجيم بمقولة المصدري.
- محل المركب الفعلي الضامر يقبل رأسه التعجيم بفعل ضامر.
- محل المركب الفعلي المعجمي يقبل رأسه التعجيم بفعل معجمي.
- تكتسي هذه المحلاّت قيمة موجبة أو سالبة، ويمكن التمييز بين مقولة الظهور والضمور من تمثيل مختلف أصناف الرؤوس الفعلية.

٢- المقياس (٢): تتشكل كل بنية من مقولات معجمية تتحكم فيها مقولات وظيفية وتوجهها سمات نحوية (السمات الزمانية، والسمات الصرفية والسمات الإعرابية والسمات الدلالية)...

٣- المقياس (٣): يحتاج اشتقاق مكونات البنية النحوية وحوسبتها إلى العمليات التركيبية التالية: النظم، والمطابقة، والنقل، والتهجية.

٤- المقياس (٤): الاشتقاق توليد تكراري قصدي لمجموعة لامتناهية من البنى اللغوية من مجموعة متناهية من العناصر يحكمها قيد المقروئية

٢-٣-٢-٤- مسرد الدراسات في المستوى النحوي:

الرقم	عنوان الدراسة	المؤلف	مكان النشر	تاريخ النشر
١	منطق النحو العربي والعلاج الحاسوبي للغات	عبد الرحمن الحاج صالح	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢

٢	التوليد الصوتي والنحوي والدلالي لصيغ المبني للمجهول	مازن عوض الوعر	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢
٣	الفعل العربي وطرق معالجته	صلاح الدين صالح حسنين	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢
٤	الحاسوب والنحو العربي	نبيل علي	الموسم الثقافي الرابع عشر لمجمع اللغة العربية الأردني	١٩٩٦
٥	توصيف الضمير المتصل للحاسوب : المعالجة والإشكال	مهدي أسعد عرار	أعمال المؤتمر العام للغة العربية بعنوان: قضايا الأدب واللغة والتحديات المعاصرة كلية الآداب- الجامعة الإسلامية بغزة – فلسطين	٢٠٠٠
٦	نموذج محوسب لمحلل نحوي للجمل الاسمية غير المشكولة في اللغة العربية	معتصم الحمدان	رسالة ماجستير – جامعة آل البيت الأردن	٢٠٠٢
٧	توصيف مركب العدد في اللغة العربية للحاسب الآلي	محمود الديكي	وقائع الندوة الدولية : المعالجة الآلية للغة العربية – معهد الدراسات والأبحاث للتعريب – المغرب	٢٠٠٧
٨	التركيب الإضافي في العربية: نحو توصيف جديد في ضوء اللسانيات الحاسوبية	جنات علي محمد أحمد	رسالة ماجستير – الجامعة الاردنية الأردن	٢٠٠٧

٩	توصيف النحو العربي في ضوء اللسانيات الحاسوبية : الفعل الماضي نموذجًا	أحلام عامر شريف الزبن	رسالة ماجستير – الجامعة الهاشمية الأردن	٢٠٠٨
١٠	المعالجة الآلية للغة العربية: إنشاء نموذج لساني صرفي إعرابي للفعل العربي	ارس شاشة	رسالة ماجستير – جامعة الجزائر الجزائر	٢٠٠٨
١١	توصيف نحوي للأفعال الواردة في شعر محمود درويش في ضوء اللسانيات الحاسوبية	أحمد أنيس شحادة عامر	رسالة ماجستير –الجامعة الهاشمية الأردن	٢٠٠٩
١٢	محاولة في توصيف الجملة الفعلية حاسوبيًا الملة البدوءة بالفعل الماضي التام المجرد الثلاثي الصحيح المبني للمعلوم	ريم فرحان المعايطة	المجلة الأردنية في اللغة العربية وآدابها مج ٦ ع٣	٢٠١٠
١٣	إسناد الأفعال إلى الضمائر في ضوء اللسانيات الحاسوبية	محمود مصطفى عيسى خليل	رسالة ماجستير – الجامعة الاردنية الأردن	٢٠١١
١٤	متطلبات المعاجلة الآلية للجمل الفعلية المتكلسة	عماد اللحياني	جملة بحوث جامعية كلية الآداب والعلوم الإنسانية– صفاقس – تونس	٢٠١٢
١٥	التحليل الآلي للوقف والابتداء في نصوص اللغة العربية	مجدي صوالحة وآخران	المؤتمر الدولي للحاسوب في اللغة العربية – القاهرة	٢٠١٢

١٦	طريقة تعتمد على المدونات اللغوية لتجهيز بيانات تدريب واختبار أنظمة لتجهيز الوسوم النحوية	عبدالمحسن عبيد الثبيتي وآخرون	مدينة الملك عبد العزيز للعلوم والتقنية معهد بحوث الحاسب المملكة العربية السعودية	د. ت
١٧	بناء نظام تحديد أقسام الكلام في النصوص العربية باستخدام منهجية تعلم نصف آلي	رياض سنبل وآخران	المجلة الدولية لعلوم وهندسة الحاسب باللغة العربية	د. ت
١٨	اللغة الداخلية و حوسبة البنية النحوية	سرور اللحياني	وقائع الندوة العلمية الدولية الثالثة للسانيات بعنوان اللسانيات وإعادة البناء – كلية الآداب والفنون والإنسانيات بمنوبة – تونس	٢٠١٤
١٩	توصيف الجملة الأسمية حاسوبياً (دراسة في الجملة المبدوءة بضمير المفرد المتكلم وخبرها مفرد نكرة جامد غير مضاف – للمفرد المذكر)	نبال نبيل نزال	دراسات العلوم الإنسانية والاجتماعية الأردن	٢٠١٤
٢٠	توليد الجمل العربية باستخدام لغة البرولوغ	رضا بابا أحمد	المجلة العربية لعلوم و هندسة الحاسوب، مج٦، ع١	٢٠١٥
٢١	تصنيف الأفعال والأسماء في نظرية أصناف الأشياء	عماد اللحياني وعبد الحميد عبد الواحد	مجلة اللسانيات العربية – مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية – السعودية	٢٠١٥

٢-٣-٢-٥- خلاصة:

إن متابعة جهود الباحثين العرب في معالجة المستوى النحوي وفق منظور حوسبة اللغة يظهر جملة من النتائج، لعل أبرزها:

١- أشار الدارسون العرب إلى صعوبة حوسبة اللغة العربية في مستواها التركيبي، ونظامها النحوي؛ نظرًا إلى ما يمتاز به نظام اللغة العربية من حيوية وتعدد، مما تعذر معه القدرة على ضبط ذلك كله.

٢- انشغلت أغلب دراسات اللسانيين بالمفهوم الذي أسس له نبيل علي في دراسته الرائدة، وتبناه نهاد الموسى تحت مصطلح "التوصيف"، لكن هذه الدراسات لم تستطع الفكاك من مصطلحات النحو العربي القديمة، ولذلك ظهرت بمستوى المقاربة لحوسبة اللغة دون أن تصل حدّ إقامة نسق منطقي مجرد يمكن تقييسه ليرفد حالات النمذجة النمطية الكلية التي تستلزمها عملية حوسبة اللغة العربية.

٣- استطاع بعض الدارسين أن ينقل اهتمامه في المستوى النحوي من إطار التوصيف النظري إلى الإطار العملي التجريبي، وأظهرت دراساتهم نجاح ملحوظًا، لكنها ظلت محصورة في أنماط تركيبية محدودة من بنية اللغة العربية التركيبية، وغالبًا ما أشار الدارسون - في توصياتهم - إلى ضرورة استكمال جهودهم في توسيع النمط التركيبي المبحوث، أو الاهتمام بمستويات أخرى لتجنب الأخطاء التي ظهرت وتقليل نسبتها.

٢-٣-٣- المستوى الصرفي

يبحث علم الصرف في بنية الكلمة وما يطرأ عليها من تغيرات سواء صوتية أم بنائية أم في بعض جوانبها الوظيفية كالتمكن والتصرف وغير ذلك، ويعدُّ المورفيم الوحدة الصغرى في اهتمامات هذا العلم، وهي أصغر وحدة صوتية ذات معنى، ثم ما يشهده المورفيم من امتدادات وتغيرات.

٢-٣-٣-١- حوسبة الصرف العربي: الموارد والخبرات اللسانية، عبد الرزاق

تواربي، وخالد الأشهب، ومعد عبد الفتاح، وقائع الندوة الدولية: المعالجة الآلية للغة العربية – معهد الدراسات والأبحاث للتعريب – المغرب، ٢٠٠٧.

قدمت هذه الدراسة مقاربة لسانية نسقية لحوسبة الصرف العربي تندرج في إطار قاعدة الصور المعجمية العربية المولدة (GENFO)، وهو مشروع تمت بلورة بعض جوانبه اللسانية والحاسوبية بمعهد الدراسات والأبحاث للتعريب. واعتمدت هذه المقاربة الخبرات اللسانية لتحديد تصور المعالجة الآلية وأسسها، وكشف نسقية المواد اللغوية، وعدم الاكتفاء بما هو موجود، بل التنبؤ بما هو ممكن في اللغة أيضًا.

ويقدم مشروع قاعدة الصور المعجمية العربية المولدة أساسًا تصوريًا لتوليد الموارد اللسانية وتنظيمها وتسويغها باعتماد خبرات لسانية تترجم تقدم البحث اللساني المقارن بالمغرب، وتقدم النظرية اللسانية عمومًا في جوانبها النظرية والتطبيقية،تهتم بتحديد النساج العام للمشروع، وتدقيق آلية التوليد، بحيث تحدد أهم المصافي الصوتية والصرفية والخطية، بشكل مبسط للتمثيل فقط. ثم يُقترح المولد الصرفي للصور المعجمية العربية، مع تطبيق أول على تصريف الأفعال العربية، لمناقشة بعض جوانب المعجم الذهني مثل مسألة الابتكار اللغوي التي يسعى المشروع إلى رصدها. وأهمية هذه الجوانب تتجلى في عكسها من جهة للخبرات اللسانية، وتصور الموارد اللسانية الموجودة باعتبارها تمثل جزءًا قليلاً من المعرفة المعجمية التي يمتلكها متكلم اللغة العربية ومستمعها. وتُختم هذه الورقة بالحديث عن الكلمات الممكنة من خلال طرح العلاقة بين الصرف والمعجم.

ترتكز الدراسة على جانبين: التحليل والتوليد، وتكشف – بداية – عن الصعوبة التي تواجه التحليل الصرفي لبنية الكلمة العربية، فهي تتجلى، من جهة، "في طابعها اللاسلسلي والمرتبط أساسًا بإدراج الجذور في الصيغ، والأواسط، ومن جهة أخرى، بنمط الكتابة العربية الذي لا يكتب، في الغالب، الحركات والعلامات الإعجامية، مثل الشدّة. بالإضافة إلى هذا، تشكو المعاجم العربية التي يمكن استعمالها كمرجع في التحليل من مجموعة من الثغرات، منها أنها لا تدرج جميع الصور النسقية في اللغة، ممثلة المشتقات الفعلية والجموع، وأسماء الأعلام

والأماكن، وقد تدرج بعض الأدوات والضمائر ملتصقة بالكلمة، ثم لا تلتزم بالشكل التام للصور".

تعرض الدراسة نموذجين في التحليل الصرفي:

الأول: نموذج (قطع - لصق)، وتقوم المقاربة المعتمدة على هذا النموذج على "تقطيع الكلمة السطحية إلى صرفيات ممكنة، ويكون التقطيع من بداية الكلمة ومن نهايتها بعزل القطع التي يمكن أن تكون سابقة أو لاحقة إلى أن تبقى ثلاثة صوامت أو أربعة تكون جذرًا ثلاثيًا أو رباعيًا. وأخيرًا، ينظر في القاموس هل توجد هذه الجذور أم لا؟" (٦٧). وهذه الطريقة سهلة لكنها لا تنجح في حالات كثيرة خاصة في حالة وجود أواسط في الكلمة العربية.

الثاني: نموذج الصرافة ذات المستوين: ويقوم على نموذجين متمايزين ومترابطين معًا، هما: المستوى المعجمي المجرد والمستوى السطحي الخطي، فالمستوى الأول؛ المعجمي يمثل تجميع الصرفيات؛ أي متتالية الصوامت، كما تظهر في المعجم. ويمثل المستوى الثاني سلسلة الصوامت الممثلة للكلمة كما تظهر في السطح والخط. وطبقًا لمواضعات التمثيل، يُكتب المستوى المعجمي دائمًا فوق المستوى السطحي. وفي النموذج، لا تحول السلسلة المعجمية أو تغير خلال التوليد إلى سلسلة معجمية، كما أن السلسلة السطحية لا تحول إلى سلسلة معجمية خلال التعرف إلى الكلمة. فكلا المستويين يوجدان في الوقت نفسه، والهدف هو الربط بينهما ساكنًا بساكن من خلال ما يسمى بالأزواج المحسوسة، وهي عبارة عن حرف معجمي وحرف سطحي يحققه.

والملحوظ على هذه الدراسة اعتمادها على المعاجم المخزنة في الحاسوب أكثر من اعتماد قواعد للتصريف، وهي سمة غالبة في دراسات حوسبة اللغة العربية، فالقواعد -حسب رأي الباحثين- "مجرد مصافٍ، والمعاجم هي التي تقود عملية البحث".

نموذج التوليد الصرفي: اعتمد في هذا النموذج على النظام الاشتقاقي التصريفي للكلمات العربية الذي صدرت النسخة الأولى منه سنة ١٩٩٨ في المعهد العالي للعلوم التطبيقية والتكنولوجيا في دمشق، وأنجزه الأساتذة: مروان البواب ومحمد مرياتي، ومحمد حسن الطيان، وأسامة رجب، ويحيى مير علم. ومشكلة هذا النظام أنه غير توليدي بالمعنى الدقيق للكلمة؛

لأنه لا يسعى إلى إيجاد الكلمات العربية النسقية سواء كانت مثبتة في المعاجم أو غير المثبتة.

والمقترح لحل هذه الإشكالية الاعتماد على ما سمي بالمعجم الذهني، بحيث يعمل على الابتكار المعجمي وسدّ الثغرات المعجمية، والغرض من ذلك محاكاة العقل البشري في معرفة ما إذا كان تسلسل عدد من الأصوات يشكل كلمة في اللغة أم أنها ليست كذلك، وذلك اعتمادًا على التوليد كأساس للمعجم الذهني لتجاوز ما هو موجود إلى ما هو ممكن عبر ما أطلق عليه "الصرافة".

٢-٣-٣-٢- ثانيًا: محلل صرفي للكلمات العربية خارج السياق وداخله، عز الدين مزروعي وعبد الوافي مزيان، وعبد الحق لخواجة، ومحمد ولد عبد الله ولد بباه، وعبد الرحيم بو دلال، ورشيد بلحبيب، اجتماع خبراء المحللات الحاسوبية الصرفية للغة العربية، دمشق، ٢٦-٢٨/ ٤/ ٢٠٠٩

تقدم هذه الورقة نبذة عن محلل صرفي للكلمات العربية قادر على تحليل الكلمة العربية بالاعتماد على سياقها في الجملة. يعمل هذا البرنامج على معالجة الجملة العربية من خلال مستويين: يُعنى المستوى الأول من المعالجة بتحليل كل كلمة على حدة بغض النظر عن الكلمات المجاورة لها. وفي هذا المستوى يقوم البرنامج بتقطيع الكلمة إلى لبناتها الأساسية من سابق وجذع ولاحق، بغية تحديد مجموعة من المعلومات الصرفية المحتملة للكلمة.

أما المستوى الثاني من التحليل فيُعنى باختيار الحلّ الأنسب لكل كلمة بالنظر إلى سياقها في الجملة، ولهذا الغرض فقد انتهجت الدراسة أسلوبًا إحصائيًا باستخدام نماذج سياقها في الجملة. واعتمد لتحقيق هذا الغرض أسلوب إحصائي باستخدام نماذج ماركوف الخفية.

وقد اعتمدت الدراسة في تدريب المحلل الصرفي واختباره على ذخيرة لغوية مكونة مما يربو على ٥٠٠٠٠٠ كلمة، وقد أُصدرت هذه الذخيرة من معهد RDI في القاهرة، حيث استخدم ما يصل إلى ٩٢٪ من هذه الذخيرة لغرض تدريب البرنامج، واستبقي ٨٪ لغرض اختبار كفاية البرنامج.

وقد أشارت الدراسة في خاتمتها إلى جملة من الملحوظات، هي:

- **إنّ التوصيف، الذي اعتمد في قاعدة بيانات الأوزان العربية الخاصة بالمحلل الصرفي المعد، ينطلق من الاهتمام في مجال التشكيل الآلي بإنجاز محلل صرفي قادر على إعطاء جميع التشكيلات المحتملة للكلمة خارج سياقها من الجملة. ثم يطبق على النموذج الإحصائي الذي جُرّب من أجل الاستخلاص الآلي لجذور الكلمات العربية داخل السياق.**
- **إنّ جانبًا مهمًا من الجهد انصب على ضبط تشكيل الأوزان العربية تشكيلاً تامًا، وتم التعامل مع الظواهر الصرفية من إعلال وإبدال وغيرها باستحداث أوزان اصطناعية لتسهيل التعامل معها.**
- **إن مرحلة التحليل الصرفي خارج السياق في البرنامج المقترح نجم عنها عدد من المخرجات التي تستلزم، من أجل مزيد من التدقيق، مجهودًا من جانب اللغويين في تجاه تعريف الأوزان التي يقبلها كل جذر عربي على حدة، مثل تحليل الفعل (فَرْمَتَ).**
- **هناك حاجة إلى إنجاز توصيف أعم لتوصيف أوزان الأسماء يسمح بإضافة معلومات صرفية من قبيل (اسم الفاعل، اسم المفعول،..إلخ)، والتذكير والتأنيث، وغيرها.**

٢-٣-٣-٣- توظيف قواعد النحو والصرف في بناء محلل صرفي للغة العربية، مجدي صوالحة وإيرك أتوك، جامعة ليدز- المملكة المتحدة.

تهدف الدراسة إلى بناء ذخيرة لغوية معنونة بالتحليل الصرفي والنحوي لجميع كلماتها؛ فتقييم المحللات الصرفية التي طورت مسبقًا سلط الضوء على بعض القصور فيها، حيث إنّ ربع كلمات الذخيرة اللغوية لم تُحلل بالشكل المطلوب. إن عملية تعيين العناوين الصرفية والنحوية للذخائر اللغوية للعربية ليست عملية سهلة، وقد طورت بعض المحللات الصرفية لاستخراج الجذر أو الجذع للكلمة العربية، ولكن هذا غير كافٍ لعمل المحللات النحوية للغة العربية بالشكل المراد تحقيقه.

وتبنى الدراسة على أساس أن الكلمة العربية تتركب من خمسة أجزاء: زوائد في بداية

الكلمة ثم سوابق الكلمة ثم الجذر أو الجذع ثم لواحق الكلمة ثم زوائد في نهاية الكلمة، ويقوم المحلل الصرفي بإضافة الخصائص اللغوية المناسبة لكل جزء من أجزاء الكلمة، وعوضًا عن عنوان واحد للكلمة نحن بحاجة إلى عنوان صرفي (subtag) لكل جزء من أجزاء الكلمة (أو أكثر من عنوان (subtags multiple) لكل جزء إذا احتوى على أكثر من زائدة أو سابقة أو لاحقة للكلمة.

وأشار الباحثان إلى وجود تحديات كثيرة تواجه تطوير المحللات الصرفية للغة العربية؛ مثل: الخاصية الصرفية الغنية "الجذر والوزن" والنظام المعقد لاشتقاق الكلمات من الجذر والوزن خاصة إذا احتوت هذه الجذور على حرف أو حرفين من حروف العلة كحروف أصلية، فضلاً عن ذلك، الخواص الإملائية كالحركات (حروف العلة القصيرة – الفتحة والضمة والكسرة)، والهمزة، والتاء المربوطة والهاء في آخر الكلمة، والياء والألف المقصورة، والتشديد (الحروف المضعفة)، والمد (آ) الذي يعدُّ حرفًا مركبًا من الهمزة متبوعًا بألف، تعتبر جميعها من التحديات الصعبة لتطوير المحللات الصرفية للغة العربية.

وقد طُور محلل صرفي يوظف قواعد النحو والصرف، ويوظف الذخائر اللغوية للتحقق من هذه القواعد المستخدمة، ولفهم المشكلة بشكل جيد؛ قام الباحثان بتحليل الجذور العربية وتصنيفها لتحديد نسبة الجذور والكلمات التي تحتوي على حرف أو حرفين من حروف العلة أو همزة أصلية في جذور هذه الكلمات، عن طريق تحليل الكلمة وجذرها لكلمات القرآن الكريم وكلمات القاموس المختلفة الذي بُني من خلال تحليل خمسة عشر معجمًا عربيًا، ثم قام الباحثان باستخلاص قوائم شاملة للزوائد والسوابق واللواحق والأوزان وبنائها من الكتب القيمة للنحو والصرف وقواعد اللغة العربية، وقد فُحصت هذه القوائم عن طريق تحليل الذخائر اللغوية (القرآن الكريم، والذخيرة اللغوية العربية (Arabic Contemporary of Corpus The)، والذخيرة اللغوية العربية المطورة بجامعة بنسلفانيا (Treebank Arabic Penn The)، بالإضافة إلى نصوص المعاجم العربية الخمسة عشر كذخيرة لغوية رابعة.

واعتمد المحلل الصرفي على هذه القوائم لتحليل الكلمات، وتطوير خوارزمية جديدة

لتحديد وزن الكلمة الصحيح إذا احتوت هذه الكلمات على حروف علّة أو همزة أو إعلال أو إقلاب. ويقبل المحلل الصرفي النصوص أو الكلمات المفردة سواءأكانت مشكولة كليًا أو جزئيًا أو غير مشكولة كمدخلاتٍ للنظام، وللتعامل مع هذه الأنواع المختلفة من النصوص اقترحت هيكلية واحدة للبيانات، فبداية يقوم برنامج خاص بتقطيع كلمات النص المدخل إلى كلمة عربية مشكولة كليًا أو جزئيًا أو غير مشكولة، أو إلى رقم أو عُملة أو علامة ترقيم، بعدها يقوم البرنامج بمعالجة الكلمات العربية المستخرجة من النص المدخل، وتقوم عملية المعالجة بإرجاع الحروف المضعّفة والمدّ إلى أصلها، حيث يُستبدل الحرف المضعّف والشّدّة الظاهرة عليه بحرفين الأول ساكن والثاني متحرك بنفس حركة الحرف الأصلي.

ودرس الباحثان المحللات الصرفية والنحوية لمعرفة إمكانية استخدامها في بناء ذخيرة لغوية معنونة بالعناوين الصرفية والنحوية للكلمات، وعرضا نتائج مقارنة محللات صرفية وخوارزميات استخراج الجذور اعتمادًا على معيار ذهبي احتوى ألف كلمة من نصوص القرآن الكريم، وألف كلمة من نصوص الصحف والمجلات اقتبستْ من الذخيرة اللغوية العربية (Arabic Contemporary of Corpus) (CCA)، ولم تكن نتائج المقارنة إيجابية، حيث إن هذه الأنظمة لم تستطع تحديد التحليل الصحيح لما يقارب ربع كلمات الاختبار، وعلى أساس ذلك بدأ البحث عن طرق أخرى لتطوير محلل صرفي قادر على تجاوز الأخطاء التي ظهرت في المحللات الصرفية التي تمت مقارنتها، ولفهم المشكلة بشمولية، فقد أجري تحليلٌ للجذور الثلاثية لكلمات القرآن الكريم والكلمات المختلفة المخزنة في القاموس، وأظهرت الدراسة أن ما يقارب ٤٠٪ من هذه الجذور الثلاثية هي جذور معتلة تشكل تحديًا للمحللات الصرفية.

لقد طُور محلل صرفي للغة العربية اعتمادًا على قوائم معدة مسبقًا للزوائد والسوابق واللواحق والجذور والأوزان، استخرجت من كتب قواعد اللغة العربية وكتب النحو والصرف، وتم التحقق من هذه القوائم عن طريق فحصها ومقارنتها بكلمات الذخائر اللغوية العربية، واحتوت قائمة السوابق على ٢١٥ سابقة، وقائمة اللواحق على ١٢٧ لاحقة، واحتوت قائمة الأوزان على ٢٧٣٠ وزنًا للأفعال و٣٩٠ وزنًا للأسماء.

ولقد طور المحلل الصرفي لتحليل الكلمات وتحديد خواصها اللغوية، ومُيز بين العديد من الخصائص اللغوية للكلمات، على أمل أن تكون المحللات الصرفية قادرة على تحديدها، وقد طورت مجموعة العناوين الصرفية والنحوية للغة العربية (Morphological features part-of speech tag set)، التي يمكن استخدامها في المحللات الصرفية والنحوية وعنونة الذخائر اللغوية، وتتكون عناوين هذه المجموعة من متسلسلة من الرموز طولها تسعة عشر رمزًا، بحيث يمثل كل رمز في موقع معين في العنوان الصرفي والنحوي قيمة لخاصية لغوية للكلمة المحللة.

ولتقييم نتائج المحللات الصرفية المختلفة ومقارنة نتائجها، اقترح بناء المعيار الذهبي لهذا الغرض، بحيث يتكون هذا المعيار من جحم مناسب من الكلمات تغطي معظم الحالات التي يجب على المحللات الصرفية والنحوية تحديد خصائصها، واختيرت النصوص المكونة للمعيار الذهبي من نصوص متعددة المصادر والأشكال، بحيث تحتوي على نصوص مشكولة كليًا وجزئيًا وأخرى غير مشكولة.

٢-٣-٣-٤-مسرد الدراسات في المستوى الصرفي:

الرقم	عنوان الدراسة	المؤلف	مكان النشر	تاريخ النشر
١	محلل صرفي للكلمات العربية المشتقة	آمال عبد اللطيف الرزوق	المؤتمر الثاني حول اللغويات الحسابية العربية-معهد الكويت للأبحاث العلمية	١٩٨٩
٢	التحليل الصرفي للعربية	يحيى هلال	وقائع مختارة من ندوة استخدام اللغة العربية في الحاسب الآلي- دار الرازي بيروت	١٩٨٩
٣	تمثيل الدلالة الصرفية في النظم الآلية لفهم اللغة العربية	محمد غزالي خياط	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢

٤	الاستكشاف الآلي للفظة الاسمية	شافية بن طامة	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢
٥	بحث تمثيل الدلالة الصرفية في النظم الآلية لفهم اللغة العربية	محمد غزالي خياط	(السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات) مطبوعات مكتبة الملك عبد العزيز العامة، الرياض	١٩٩٣
٦	إحصاء الأفعال العربية في المعجم الحاسوبي	مروان البواب وآخرون	مكتبة لبنان	١٩٩٦
٧	الاصطلاح المولد GENTERM: نظام للتوليد الآلي للمصطلحات والمولدات	عبد القادر الفاسي الفهري	معهد الدراسات والأبحاث للتعريب بالرباط	١٩٩٦
٨	التحليل الصرفي للغة العربية باستخدام الحاسوب	مأمون حطاب وآخرون	الموسم الثقافي الرابع عشر لمجمع اللغة العربية الأردني	٤ أيار- ٢٢ حزيران ١٩٩٦
٩	معجم تصريف الأفعال العربية	حسن بيومي وآخرون	دار إلياس العصرية- القاهرة	
١٠	قيود تأليف الأبجديات الصرفية: لواصق تطابق الفعل والفاعل نموذجًا، في التوليد والنسقية والترجمة الآلية	عبد الرزاق تواربي	معهد الدراسات والأبحاث للتعريب بالرباط	٢٠٠١

١١	حول المولد الصرفي للكلمات المعجمية العربية، في التوليد والنسقية والترجمة الآلية	عبد الرزاق تواربي وسالم الرامي	معهد الدراسات والأبحاث للتعريب بالرباط	٢٠٠١
١٢	المنظومة الصرفية للغة العربية رؤية حاسوبية	عبد المجيد ضوة	مجمع اللغة العربية الأردني	٢٠٠٢
١٣	النظام الصرفي للعربية في ضوء اللسانيات الحاسوبية مثل من جمع التكسير	هدى سالم عبد الله آل طه	رسالة دكتوراه الجامعة الأردنية قسم اللغة العربية وآدابها	٢٠٠٥
١٤	حوسبة الصرف العربي: الموارد والخبرات اللسانية	عبد الرزاق تواربي، وخالد الأشهب، ومعد عبد الفتاح	وقائع الندوة الدولية: المعالجة الآلية للغة العربية – معهد الدراسات والأبحاث للتعريب – المغرب	٢٠٠٧
١٥	التحليل الصرفي للأسماء العربية	عبد الفتاح حمداني	مجلة أبحاث لسانية – معهد الدراسات والأبحاث للتعريب بالرباط، ع٢٣، ٢٤	٢٠٠٧
١٦	توليد الأسماء من الجذور الثلاثية الصحيحة في اللغة العربية – مقاربة لسانية حاسوبية	عمر مهديوي	جامعة الحسن الثاني – كلية الآداب والعلوم الإنسانية – الدار البيضاء	٢٠٠٨
١٧	محلل صرفي للكلمات العربية خارج السياق وداخله	عز الدين مزروعي وآخرون	اجتماع خبراء المحللات الحاسوبية الصرفية للغة العربية، دمشق	٢٠٠٩
١٨	المحللات الصرفية للغة العربية	سلوى السيد حماده	اجتماع خبراء المحللات الحاسوبية الصرفية للغة العربية دمشق	٢٠٠٩

١٩	تقرير في المحللات الصرفية للغة العربية	محمد زايد	اجتماع خبراء المحللات الحاسوبية الصرفية للغة العربية، دمشق	٢٠٠٩
٢٠	المحلل الصرفي للغة العربية لمخبر ﴿ميراكل﴾	عبد المجيد بن حمادو	اجتماع خبراء المحللات الحاسوبية الصرفية للغة العربية، دمشق	٢٠٠٩
٢١	توظيف قواعد النحو والصرف في بناء محلل صرفي للغة العربية	مجدي صوالحة وإيرك أتوك	اجتماع خبراء المحللات الحاسوبية الصرفية للغة العربية، دمشق	٢٠٠٩
٢٢	توصيف لغوي صرفي لشعر بدر شاكر السياب في ضوء اللسانيات الحاسوبية	عزت جهاد العجوري	رسالة ماجستير الجامعة الهاشمية – الأردن	٢٠٠٩
٢٣	التحليل الصرفي للقرآن الكريم: قاعدة بيانات مفهرسة كاملة لكامل النص القرآني	يحيى محمد الحاج وآخرون	المجلة العربية لعلوم و هندسة الحاسوب، مج٣، ع١	٢٠٠٩
٢٤	تجذير اللغة العربية باستعمال المسوقات ذات الحالات النهائية: كلمة واحدة، عدة جذور	ياسين اليونسي وآخرون	المجلة العربية لعلوم و هندسة الحاسوب، مج٣، ع١	٢٠٠٩
٢٥	استخدام النماذج الإحصائية في التعرف الآلي على خصائص المفردات العربية	يحيى محمد الحاج وآخران	المجلة العربية لعلوم و هندسة الحاسوب، مج٣، ع٢	٢٠١٠

٢٦	المقاربة الحاسوبية للصرف العربي: قراءة في الحصيلة والآفاق	عمر مهديوي	مؤتمر المحتوى العربي في الانترنت (التحديات والطموح)، مج٢ جامعة الإمام محمد بن سعود الإسلامية	٢٠١٠
٢٧	بناء نظام تحديد أقسام الكلام في النصوص العربية باستخدام من، جية تعلم نصف آلي	رياض سنبل وآخران	المجلة العربية لعلوم و هندسة الحاسوب، مج٤، ع٢	٢٠١١
٢٨	المقاربة الحاسوبية للصرف العربي : قراءة في الحصيلة والآفاق	عمرمهديوي	لسجل العلمي لمؤتمر المحتوى العربي ، صص٩٩٩-١٠٢٧.	١٤٣٢هـ ٢٠١١
٢٩	التحليل الصرفي لنصوص اللغة العربية الحديثة والكلاسيكية	مجدي صوالحة وإرك أتوك	المؤتمر الدولي للحاسوب في اللغة العربية – الرياض	٢٠١١
٣٠	معالجة الوحدات الإسمية في اللغة العربية: الأعداد المركبة نموذجاً	عزالدين غازي	أشغال الندوة الدولية(CITALA)، معهد الأبحاث والدراسات للتعريب ، الرباط	٢٠١٢
٣١	بناء نظام تحديد أقسام الكلام باستخدام منهجية تعلم نصف آلي	محمد سعيد دسوقي وآخرون	المجلة الدولية لعلوم وهندسة الحاسوب باللغة العربية، العدد الثاني، المجلد الرابع، آب	٢٠١٢
٣٢	مقاربة صرفية إحصائية للتشكيل الآلي	عز الدين مزروعي وآخرون	Communications of the Arab Computer Society, Vol. 5، No. 1	٢٠١٢

٣٣	دراسة الإبدال الصرفي في ضوء اللسانيات الحاسوبية	إلهام أبو فريحة	دكتوراه جامعة العلوم الإسلامية العالمية	٢٠١٣
٣٤	تصنيف الأفعال والأسماء في نظرية أصناف الأشياء	عبد لحميد عبد الواحد	مجلة اللسانيات العربية – مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية – السعودية	٢٠١٥
٣٥	تصنيف الأفعال والأسماء في نظرية أصنافالأشياء	عماد اللحياني وعبد الحميد عبد الواحد	اللسانيات العربية مركز الملك عبد الله بن عبد العزيز لخدمة اللغة العربية– الرياض	ذو القعدة ١٤٣٦هـ – سبتمبر ٢٠١٥
٣٦	التحليل الصرفي لمكونات الكلمات العربية : دراسة لغوية حاسوبية	أحمد راغب أحمد	مجلة جيل الدراسات الأدبية والفكرية	٢٠١٦
٣٧	لغويات حاسوبية: دراسة صوتية صرفية في أبواب الفعل الثلاثي في المعجم الوسيط	محمد جواد النوري	دار الجندي للنشر والتوزيع	٢٠١٦
٣٨	أقسام الكلم العربي نحو تقييم المحللات الصرفية العربية في ضوء منهج تمام حسان	عمرو حمدي الجندي	مجلة الكلمة، س٢٣، ع٩٢ – بيروت	٢٠١٦
٣٩	لغويات الحاسوبية، دراسة صوتية صرفية في جذور الأفعال الثلاثية	محمد جواد النوري	دار الجندي للنشر والتوزيع	٢٠١٦
٤٠	لغويات حاسوبية، دراسة صوتية صرفية في الأفعال الثلاثية المزيدة	محمد جواد النوري	دار الجندي للنشر والتوزيع	٢٠١٦

٢-٣-٣-٥- خلاصة:

إن أبرز نتيجة يمكن ملاحظتها عند تتبع جهود الباحثين العرب في حوسبة اللغة العربية في المستوى الصرفي تكمن في كثرة الدراسات فيه قياسًا إلى المستوى النحوي، ويلحظ - كذلك -أن معالجة المستوى الصرفي حاسوبيًا أسهل من معالجة المستوى النحوي؛ نظرًا إلى أن النظام الصرفي يبقى قابلاً للضبط الدقيق سواء في إطار نمط البنية الصرفية للكلمات، أم في إمكانية الاعتماد على ذخيرة لغوية تسهم في تصحيح وتعديل الأخطاء على مستويي التحليل والتوليد.

وقد أظهرت جلّ الدراسات أن الإشكاليات التي واجهت معالجة المستوى الصرفي للغة العربية كانت محدودة في بعض الحركات ومثيلاتها مما يلحق بالكتابة مثل الشدة والمد والتنوين، وكان ثمة مقترحات لتجاوز هذه الصعوبات.

٢-٣-٤- الموارد

٢-٣-٤-١- الأنطولوجيا:

تعرف (الأنطولوجيا) بأنها تنظيمات ومواصفات تصور المفاهيم لتتخذ شكل خرائط مفاهيمية، ويمكن تحديدها بأنها كلمات ومفردات ومصطلحات مترابطة تعكس المفاهيم في صورتها المجردة، بحيث تعمل على تنظيم المعلومات على الشابكة، لا أن تبقى ركامًا غير منتظمة، ليسهل الربط بين محتوياتها والرجوع إليها، وتتعدد المجالات التي تُستخدم فيها؛ مثل: هندسة البرمجيات، والشبكة الدلالية، وتشابك الكلمات، ومعالجة اللغة، وعلوم إدارة المعلومات وغيرها.

ولعل أقرب التصورات اللسانية إلى مفهوم (الأنطولوجيا) هو "نظرية الحقول الدلالية"، وقد قدم في إطار الدراسات اللسانية والدراسات الدلالية الثقافية تصورات ونظريات كثيرة في هذا المجال، لكن المعاين لدراسات حوسبة اللغة لا يلمح اهتمام الباحثين في الحوسبة بهذه الصلة العميقة.

ويكاد يكون مفهوم "الشبكة الدلالية" (الويب دلالي) أكثر المفاهيم المرتبطة

بـ(الأنطولوجيا) فاعلية في إدراك قيمة المعرفة اللسانية في تطوير عمليات البحث على الشابكة، ذلك أن هذا المفهوم يتعامل مع أحد أهم المستويات اللسانية؛ وهو المستوى الدلالي، وهدف "الشبكة الدلالية" (الويب دلالي) ربط كل صفحة من صفحات الشابكة بحقل مفاهيمي يبين علاقة المفاهيم فيما بينها؛ مما يتيح للمستخدم الوصول إلى صفحات قد لا يكون المصطلح المراد أو المقصود قد ورد فيها حرفيًا، ولكن "الأنطولوجيا" (الأنطولوجيا) أشارت إلى وجود ارتباطات بين محتوى الصفحة من مفاهيم وموضوع البحث.

وستُعرض -هنا - ثلاث دراسات تظهر طبيعة اهتمام الدارسين العرب بمبحث "الأنطولوجيا"، واهتمام الدراسين في اللسانيات الحاسوبية بعلم الدلالة لاستظهار مدى إدراك الطرفين للقواطع المشتركة بين اهتماماتهما، فعلم الدلالة قادر على تقديم التصورات الأولية لأية عملية تهدف إلى إقامة "حقول مفاهيمية" للحوسبة.

٢-٣-٤-١-١- نحو منهجية لبناء هندسة الأنطولوجيا التصنيف بالصفات، جمال ضاهر ومصطفى جرار، المؤتمر الفلسطيني الدولي الثالث لتكنولوجيا المعلومات والحاسوب، ٢٠١٠.

تعد هذه الدراسة من الدراسات القليلة المبكرة التي اعتنت بمجال "الأنطولوجيا"، وهي تسعى إلى تقديم منهجية جديدة للتبادل السليم والفعال للبيانات بتعريف المعنى الدلالي للبيانات تعريفًا دقيقًا اعتمادًا على صفات الأشياء لا على أساس أجناسها، وذلك أن هذه المنهجية أسهل في التطبيق وتقود إلى نتائج سليمة وأكثر اتساقًا على المستوى الدلالي.

تطرح الدراسة مشكلة دراسات (الأنطولوجيا) المعتمدة في تصنيف البيانات على أجناس الأشياء؛ فكيف يمكن التمييز بين "طالب" و"موظف" إذا عُرّف كلٌ منهما على أنه صنف من أصناف الإنسان؟ إذ لا يمكن في - هذه الحالة - التمييز بينهما إلا اعتمادًا على الصفات المميزة لكل منهما.

إن العملية بمجملها تعتمد على أن تكون المعاني في (الأنطولوجيا) مكتوبة بطريقة تستطيع الأنظمة فهمها وحسابها من خلال تعريف المعاني وتمثيلها بلغة المنطق، وتكمن الصعوبة في أن تصنيفات الأشياء لا يمكن أن تكون أحادية وكلية تمامًا، إنما هي متفرعة ومتشابكة في اتجاهات عدة، ولذلك تعتمد المنهجية المقترحة من الباحثين على الصفات العرضية لا الجوهرية، وتأتي هذه المحاولة للتخلص من التصنيف القائم على الصفات الجوهرية التي يصعب تحديدها عندما تقارن بالصفات العرضية للأشياء.

٢-٣-٤-١-٢- نحو التكامل المعرفي من واقع توظيف الأنطولوجيات في إطار التنقيب عن البيانات: دراسة تحليلية، مؤمن النشرتي، Cybrarians Journal. - ع ٣٥، سبتمبر ٢٠١٤.

سعت هذه الدراسة إلى توضيح مفهوم التكامل المعرفي في سياق تقني، وأشهر المبادرات التقنية التي ساهمت في تحقيقه على صعيد المحتوى المتاح على الشابكة، ثم رصدت القدرة على توظيف (الأنطولوجيا) في إطار التنقيب عن البيانات سعيًا بذلك إلى الاستفادة منهما فيما يعرف بنظم (الأنطولوجيا) للتنقيب على الشابكة وتحقيق التكامل المعرفي لمحتواها.

واعتمدت الدراسة على المنهج المسحي لرصد التحديات الخوارزمية واللغوية التي تواجه البحث على الشابكة في تحقيق التكامل المعرفي، وتعتمد -أيضًا - على المنهج التحليلي في رصد واقع قدرات الأنطولوجيا وتحليله في أن توظف لتحقيق التكامل المعرفي للمحتوى المتاح على الشابكة.

لقد عرضت الدراسة لأقسام التنقيب على الشابكة: التنقيب اعتمادًا على بنية الروابط المتاحة، والتنقيب عن محتوى، والتنقيب اعتمادًا على أنماط الإفادة والاستخدام.

ثم بيّنت الدراسة آليات التنقيب، مثل ما يعرف بزواحف الشابكة، والتأخيذ والتحليل، وهذه الآلية تقوم باستخلاص المعلومات باستعمال نمطين: النمط الأول هو استخلاص المعلومات من نص اللغات الطبيعية للمحتوى، والنمط الثاني هو استخلاص المعلومات من

البيانات المهيكلة من صفحات الشابكة، ويحظى النمط الأول بالاهتمام من جانب دراسات معالجة اللغة الطبيعية، بينما يمثل النمط الثاني جوهر المعالجة في هذه الدراسة، وعادة ما يسمى البرنامج الذي يضطلع باستخراج هذه البيانات باسم المجمع Wrapper[١٠].

ويهتم النمط الثاني بعمليات كثيرة ذات الصلة بمباحث لسانية؛ فالاعتماد في كشف محتوى النصوص بأنساقها الحاسوبية المتنوعة على الكلمات الدالة لطبيعة النصوص التي تشمل مجموعة من العمليات الفرعية، مثل:

- (التأخيذ) (**Tokenization**) الذي يقوم بتفكيك النصوص الكاملة إلى كلمات وتحديدها للوصول إلى كيان مستقل للكلمة الذي يتحقق من خلال المسافات وعلامات الترقيم داخل النص، وغيرها من العناصر التي تعمل على تحديد ملامح الكلمة.
- قائمة الاستبعاد: وتتمثل في استبعاد الكلمات التي تحمل دلالات معلوماتية ولغوية ضئيلة.
- جذور الكلمات: وتعمل هذه المنهجية على استخراج الجذور الصرفية لكل كلمات الوثيقة.
- هيكلة البيانات: وتعمل على أساس توفير نماذج وتقنيات تعمل على تحقيق التكامل المعرفي والدلالي بصورة آلية، بحيث تهدف إلى تمثيل المعرفة البشرية في صورة لغة مهيكلة ذات بناء منطقي.

وكان من أهم نتائج الدراسة:

١- وجود العديد من التحديات التي تواجه الترابط على الشابكة في تحقيق التكامل المعرفي على صعيد المحتوى، والتي تتعلق بـ:

- عدم قدرة زواحف الشابكة (crawler) على تحميل كافة الصفحات المتاحة على الشابكة وتجميعها؛ نظرًا إلى ضخامة الشابكة وعدم وجود سياسة واضحة للتجميع لدى الزاحف.
- صعوبة الكشف عن وجود مكررات على صعيد محتوى الشابكة.
- لا تتناسب عمليات التأخيذ، وهي إحدى المراحل الأساسية في التنقيب على الشابكة التي تعتمد على تفتيت المحتوى الكامل إلى كلمات مستقلة، مع طبيعة بعض المحتويات الخاصة بالوثائق، ففرضًا إذا فُتت معادلة حسابية فلايمكن لأدوات البحث أن تسترجع المعادلات الحسابية أو الرياضية وفقًا لهذا المبدأ.

ومن أبرز توصيات الدراسة:

١- أن يعمل مصممو محتوى الشابكة على نمذجة المحتوى المتاح على الشابكة في مجموعة من نماذج البيانات اعتمادًا على الأنطولوجيا.

٢- أن تتخلى أدوات البحث والاسترجاع ومختلف تطبيقات الشابكة عن منهجيات التكشيف التي تعتمد على تطويع النصوص وتفتيتها بهدف الفصل في الاسترجاع.

٣- أن تعمل أدوات البحث على توظيف نماذج التنقيب وتقنياته على الشابكة لاستخراج المعرفة من مختلف موارد المعلومات.

٤- أن يعمل مطورو الشابكة على إيجاد نماذج مفاهيمية وأنطولوجية بهدف التنقيب عن المعرفة في الملفات صوتية والفيديو والصور الثابتة، وذلك في سياق التكامل المعرفي على الشابكة.

٥- أن تتوجه نظم التنقيب على الشابكة إلى تطوير محركات للاستدلال والمنطق التي تكفل

القدرة على التحليل والتكشيف المفاهيمي للوثائق.

٦- أن تعمل المنظمات الإقليمية ذات التوجه المعلوماتي على إثراء المحتوى العربي بالأنطولوجيات العامة والمتخصصة.

٧- أن توفر رابطة الشابكة العالمية نطاقات أسماء عامة ومتخصصة بحيث يراعى فيها التحديث وجعلها مظلة للمحتوى المتاح على الشابكة.

٨- أن تعمل شركات تقنيات المعلومات على تطوير برمجيات خاصة بتحرير الأنطولوجيات تتسم بدعمها للغة العربية.

٩- أن تثري المحافل العربية، من مؤتمرات وندوات متخصصة بالمحتوى العربي على الشابكة، البحوث الأصيلة العلمية في مختلف المجالات بهدف رفع الرتبة اللغوية العالمية للغة العربية على الشابكة.

٢-٣-٤-١-٣- علم الدلالة والأنطولوجيا من منظور حوسبة اللغة العربية، تحرير:هند بنت سليمان الخليفة، مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، الرياض، ١٤٣٨هـ-٢٠١٧م.

ضم هذا الكتاب خمس دراسات توزعت على فصوله الخمسة، وختمت بفصل يبين النتائج والتوصيات.

- في الفصل الأول قدمت نوال الحلوة -من الجانب اللغوي- تعريفًا بالدلالة كعلم مستقل في اللغة وما يحوم حولها من مفاهيم مثل المعنى والمفهوم وعلاقة الدلالة بالأنطولوجيا كشق من علم الدلالة، ثم تناولت بشيء من التفصيل نظرية الحقل الدلالي والتحليل الدلالي والعلاقات الدلالية بمستوياتهاالمختلفة، وذلك لما للنظريات الدلالية الحديثة من دور في خدمة الأنطولوجيا.
- في الفصل الثاني تناولت عريب العويشق -من الجانب الحاسوبي- مفهوم الأنطولوجيا

وأنواعها ولغاتها وأساليب بنائها والأدوات المعينة على ذلك، ثم تعمقت بشيء من التفصيل خطوات تصور المفاهيم وتشكيلها في الأنطولوجيا.

- في الفصل الثالث، استعرضت هند الخليفة واقع الأبحاث والجهود المبذولة في إنشاء أنطولوجيا تعتمد اللغة العربية في بنائها وذلك لتحقيق أهداف عدة؛ مثل استخدامها في استرجاع المعلومات، وبناء نظم الإجابة الآلية عن الأسئلة التي تعتمد على الأنطولوجيا، وتوسيع الاستعلام، والتلخيص الآلي للوثائق الإلكترونية وغيرها، ثم ختمت الفصل بتلخيص للتحديات التي تواجه مجال هندسة الأنطولوجيا العربية، فضلاً عن الإشارة إلى بعض موارد الأنطولوجيا العربية المشاعة للاستخدام.
- في الفصل الرابع تناولت عالية باحنشل بعضًا من أشهر الإطارات الحاسوبية التي أنشئت في مجال اللغة، وذلك بتعريف كل إطار وأهدافه والجهود المبذولة لإنشاء مثيلاته في اللغة العربية تحديدًا، وذلك لفتح المجال للمهتمين في توظيف الأنطولوجيا في أحد مجالات معالجة اللغة العربية.
- في الفصل الخامس قُدم درس تطبيقي في بناء أنطولوجيا باستخدام أحد البرامج الحاسوبية المتخصصة، وهو برنامج بروتيجي (Protégé)، لتختتم بعد ذلك هند الخليفة الكتاب بطرح بعض التوصيات للمختصين باللغة العربية والحوسبة لتجاوز التحديات الحالية التي تواجه مجال الأنطولوجياالعربية، ثم سُلط الضوء على أهم البحوث المستقبلية في هذا المجال.
- في الفصل السادس،"خاتمة ونظرة مستقبلية"،تعرض هند الخليفة بعض التحديات والمجالات البحثية المستقبلية التي تكتنف مجال الأنطولوجيا لعلها تفتح آفاقًا أرحب للباحثين وطلبة الدراسات العليا لدراستها.

وكان من أبرز التحديات التي تواجه مجال الأنطولوجيا العربية، سواء كانت في هندسته، أم في أبحاثه الحالية،أم في الأطر اللغوية المستخدمة للتمثيل المعرفي وبناء أنطولوجيا للغة العربية، ما يلي:

- التحدي الأول: يكمن في تعريب كلمة أنطولوجيا وما يتبعها من مصطلحات، فبعضها يوجد لها مقابل في اللغة العربية وإن وجدت فهناك تعدد للمعنى، فكلمة أنطولوجيا مستقاة من الفلسفة الإغريقية وتعني علم الوجود، وهذه الكلمة قد تجعل دارسينيترد دونفي استخدام مثل هذه الكلمة في أبحاثهم؛ لذا يظهر في بعض الأبحاث المكتوبة باللغة العربية استخدام ألفاظ مختلفة للدلالة على المصطلح، وهذا يؤدي إلى صعوبة البحث عن الدراسات السابقة في المجال.
- التحدي الثاني: في بناء الأنطولوجيا وتعديلها، إن اكتشاف مفاهيم أو علاقات متجددة قد يؤدي إليتغيير بنية الأنطولوجيا؛ لذا يتوجب أن يكون هناك اتفاق على ماهية الإضافة أو التغيير في بنية الأنطولوجيا قبل الشروع بتعديلها.
- التحدي الثالث: قلة الموارد اللغوية للغة العربية بشكلها الرقمي، ويُقصد بالموارد اللغوية القواميس والمدونات وغيرها التي تساعد في بناء الأنطولوجيا، فالجهودالموجودة هي جهود مبعثرة غير متكاملة، تبدأ عادة من الصفر وتتوقف عندما تَشِحُّالموارد مثل الوقت أو التمويل. وهذا مما يصعب مهمة الحاسوبي عند قيامه ببناء أنظمة تعتمد الأنطولوجيافي عملها، فيضيع وقته وجهده في البحث عن هذه الموارد وتحويلهاإلى صيغ رقمية، يضاف إلى ذلك عند وجود موارد لغوية مثل المدونات فإنها تأتي بصيغتها الخام مما يتطلب العمل عليها وتزويدها بمعلومات لغوية تسهل معاجلتها آليا.
- التحدي الرابع: ضعف مشاركة الموارد اللغوية والبرمجيةلأنظمة الأنطولوجيا العربية، وهذا ما يلحظفي كثير من الأبحاث التي تستعرض أعمالها دون الإشارة إلى كيفية إعادة استخدامها كليًا أو جزئيًا، يضاف إلى ذلك الضعف فيالمحتوى العربيبالأنطولوجيات العامة والمتخصصة.
- التحدي الخامس: ضعف حركة النشر والترجمة والتعريب لأنظمة الأنطولوجيا باللغة العربية، فعلى سبيل المثالما وردفي الفصل الخامس عن برنامج (البروتيجي)، سيلحظ

القارئ أن واجهة البرنامج كانت باللغة الإنجليزية ولمتُعرب الواجهة مع أن البرنامج مفتوح المصدر، وقد يعزى ذلك إلى قلة مستخدمي مثل هذه البرامج التخصصية من مختصي اللغة العربية تحديدًا، أما ما يتعلق بقلة النشر والترجمة، فهناك شبه ندرة في المجلات العربية المهتمة بنشر أبحاث الأنطولوجيابشكل خاص والحوسبة العربية بشكل عام، فضلاً عن أنه لا يوجد اهتمامفيترجمة الكتب المتميزةفي هذاالمجال لنقل العلم للباحث العربي.

- التحدي السادس: عدم وجود مجموعات تهتم أو تتحالف – سواء كانت محلية أو إقليمية – في الوطن العربي لدفع حركة البحث في مجال الأنطولوجيا العربية أسوة بما نراه في أوروبا وغيرها، وعدم وجود تمثيل لأبحاث اللغة العربية في منظمات دولية أيضا.

أما على صعيد النظرة للمستقبل، فقد أشير – في الدراسات المضمنة في الكتاب – إلى أن مجال أبحاث الأنطولوجيا العربية مازال وليدًا، فضلاً عن أنه يتقدم ببطء، ومع ذلك فإن هذا لم يمنع الباحثين من الاهتمام فيمجالات متقدمة في أبحاث الأنطولوجيا، وذلك بإسقاط المنهجيات البحثية المستخدمة للغات الأخريعلى اللغة العربية، وحاليًا، يعدُّمجال أبحاث الأنطولوجيا العربية من المجالاتالخصبة والواسعة التي يمكنلها أن تفتح آفاقًا مستقبلية لدراسات بينية، من هذه الأفكار البحثية ما يلي:

- هناك كثير من الدراسات اللغوية التي ينتجها الباحثون فيمجال اللغة العربية منها رسائل علمية أو بحوث ترقية، وهي مما يمكن الاستفادة منها في تنشيط أبحاث الأنطولوجيا، غير أن الوصول لمثل هذه الأبحاثفي الوقت الراهن صعب المنال لغياب ثقافة المشاركة ونشر العلم.
- بناء موارد لغوية دلالية مثل فكرة بناء (ووردنت) عربية للقرآن الكريم، وذلك عن طريق استخراج المرادفات بشكل آليبالاعتمادعلى المعاجم العربية القديمة.
- بناء أدوات ومحررات خاصة بمعاجلة الأنطولوجيا بواجهة عربية وتدعم التعامل مع النص العربي بسهولة، ودعم حركة تعريب الأدواتالحالية.

- بناء موقع مرجعي متخصص لتجميع الأنطولوجيا العربية وتصنيفها وفهرستها تمكن العامل ينفي هذا المجال من مشاركة أنطولوجياته، ويحتوي الموقععلى محرك بحث متخصص وكشاف بالأنطولوجياالمتاحة بالموقع.
- عمل دراسات مستفيضة للأدواتالمتاحة حاليًا لبناء الأنطولوجياوتحريرها وغيرها من وظائف لاختبار مدى ملاءمتها للغة العربية واقتراحتحسينات عليها.

٢-٣-٤-٤- مسرد الدراسات:

الرقم	عنوان الدراسة	المؤلف	مكان النشر	تاريخ النشر
١	دراسة إحصائية لجذور معجم الصحاح (باستخدام الكمبيوتر)	علي حلمي موسى	الهيئة المصرية العامة لكتاب	١٩٧٨
٢	علم المصطلحات وبنوك المعطيات	ليلى المسعودي	مجلة اللسان العربي، العدد ٢٨،	١٩٨٧م
٣	المعجم الحاسوبي في نظام خبير للغة العربية	محمد مراياتي وزملاؤه	بحوث المؤتمر العلمي الأول حول الكتابة العلمية باللغة العربية - واقع وتطلعات) الذي نظمته جامعة العرب الطبية، بنغازي - ليبيا	١٩٩٠م
٤	نحو معجم عربي للتطبيقات الحاسوبية	محمود إسماعيل (الصيني)	لسجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات المنعقدة في مكتبة الملك عبدالعزز بالرياض (٨-١٢ / ١١/ ١٤١٢هـ) ، صص ٥١١-٥٢١	١٩٩٢

٥	البحث من العنوان في قواعد البيانات العربية	بخيت سليمان البخيت	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢
٦	تعريب المصطلحات المستعملة في الحواسيب الصغروية	أحمد بوعزي	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢
٧	في سبيل نظرية مصطلحية عربية ممكنة	محمد رشاد الحمزاوي	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢
٨	مصطلحات المعلوماتية واللغة العربية	سعد الحاج بكري وآخران	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢

٩	التخطيط لخدمات معلوماتية باللغة العربية	عبد الله الضلعان وآخران	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢
١٠	بنوك المصطلحات الآلية والمعاجم الألكترونية	محمود إسماعيل صالح (الصيني)	لسجل العلمي للندوة الثانية لتعريب الحاسوب (٢٧- ٣. مارس ١٩٩٤م)	١٩٩٤م
١١	برنامج لساني-حاسوبي للتعرف الآلي على التعابير المسكوكة في اللغة العربية	محمد الحناش	مجلة التواصل اللساني ، ملحق سلسلة الندوات ، المجلد ٣	١٩٩٦
١٢	المعاجم العلمية العربية المتخصصة ودور الحاسوب	إبراهيم بن مراد	الموسم الثقافي الرابع عشر لمجمع اللغة العربية الأردني	١٩٩٦
١٣	كفاءة التحليل الصرفي في استرجاع النصوص العربية	مساعد الطيار	لة مكتبة الملك فهد الوطنية، المجلد ٤ ، العدد١	١٩٩٨
١٤	بنوك المصطلحات الآلية	محمود إسماعيل صيني	مجلة اللسان العربي، العدد ٤٨	١٩٩٩م
١٥	البنك الآلي السعودي للمصطلحات (باسم)	عبد الرحمن بن عبد العزيز الفاضل	مجلة اللسان العربيالعدد ٤٧	١٤٢٠هـ- ١٩٩٩

١٦	نظام حاسوبي لرومنة الأسماء العربية	عبدالملك السلمان ومنصور الغامدي وحسن الصبي	لندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية ، صص ٢١٥-٢٢٨ .	٢٠٠٧
١٧	المتطلبات اللغوية لمعالجة التعابير الاصطلاحية العربية معالجة آلية	وفاء كامل فايد	لندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية (٢٩/ ١٠-٢/ ١١/ ١٤٢٨هـ/ ١٠-١٢/ ١١/ ٢٠٠٧). الرياض: مدينة الملك عبدالعزيز للعلوم والتقنية.	٢٠٠٧
١٨	الهيكلية الآلية للنصوص العربية باقتباس المفاهيم الشكلية المثالية واستعمالها للبحث في النصوص	علي الجوّة وآخرون	المجلة العربية لعلوم و هندسة الحاسوب، مج١، ع٢	٢٠٠٧
١٩	أداة ويب معتمدة على عملية التحليل الهرمي للحصول على معجم عربي موحد لتقنية المعلومات	عبدالله شرف الغامدي و بدرية سليمان الفرهود	ندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية	٢٠٠٧
٢٠	رهانات نظرية حاسوب-لسانية في بناء معجم آلي للتعابير المسكوكة	عز الدين غازي	مجلة الدراسات المعجمية المغرب، ع ٦	٢٠٠٧
٢١	التعرف الموضوعي للنصوص العربية	مراد عباس وآخران	المجلة العربية لعلوم وهندسة الحاسوب، مج١، ع٣	٢٠٠٨

٢٢	التعرف الموضوعي للنصوص العربية	مراد عباس وآخران	المجلة العربية لعلوم وهندسة الحاسوب، مج٢، ع٣	٢٠٠٨
٢٣	البحث عن النصوص	علي الجوّة وآخرون	المجلة العربية لعلوم وهندسة الحاسوب، مج٢، ع٢	٢٠٠٨
٢٤	الجوانب التقييسية للمعاجم الحاسوبية	عبد المجيد بن حامدو	المنظمة العربية للتربية والثقافة والعلوم ومركز الملك عبد العزيز للعلوم والتقنية	٢٠٠٨
٢٥	الهيكلية الآلية لنتائج محركات البحث العربية والإنجليزية باقتباس المفاهيم الشكلية	علي الجوّة وآخران	المجلة العربية لعلوم وهندسة الحاسوب، مج٣، ع١	٢٠٠٩
٢٦	محركات البحث الدلالي في ظل تطبيقات الويب الدلالي	فاتن سعيد بامفلح	المجلة العربية للأرشيف والتوثيق والمعلومات. مج ١ ، ع ٢٧	٢٠١٠
٢٧	محركات البحث الدلالي في ظل تطبيقات الويب الدلالي	فاتن سعيد بامفلح	المجلة العربية للأرشيف والتوثيق والمعلومات. مج ١ ، ع ٢٧	٢٠١٠
٢٨	استخدام الشبكات العصبية في تصنيف النصوص العربية بالاستناد على طريقة تحليل القيم المفردة	فوزي حراق	المجلة العربية لعلوم وهندسة الحاسوب، مج٣، ع٢	٢٠١٠
٢٩	دراسة أدوات استخراج المصطلحات النصية و مدى تكيفها مع اللغة العربية	صورية زايدي ومحمد الطيب العسكري	المجلة العربية لعلوم و هندسة الحاسوب، مج٣، ع٢	٢٠١٠

٣٠	بناء المحتوى العربي الدلالي الأول و إسهاماته في تطوير معالجة اللغة آليًا	أمل السيف وكاتجا ماركرت	المجلة العربية لعلوم و هندسة الحاسوب، مج٤، ع١	٢٠١١
٣١	نحو منهجية ومدونة في خدمة المعلوماتية العربية في الواب الاجتماعي الدلالي	إبراهيم بو نحاس ويحيى سليمان	المجلة العربية لعلوم و هندسة الحاسوب، مج٣، ع٣	٢٠١١
٣٢	نحو منهجية ومدونة في خدمة المعلوماتية العربية في الواب الاجتماعي الدلالي	إبراهيم بو نحاس ويحيى سليمان	المجلة العربية لعلوم و هندسة الحاسوب، مج٣، ع٣	٢٠١١
٣٣	بناء المحتوى العربي الدلالي الأول و اسهاماته في تطوير معالجة اللغة آليًا	أمل السيفو كاتجا ماركرت	المجلة العربية لعلوم و هندسة الحاسوب، مج٤، ع١	٢٠١١
٣٤	الاستخلاص الآلي للمحتوى العربي على شبكة الإنترنت بين الواقع والمأمول	محمد فتحي الجلاب	السجل العلمي لمؤتمر المحتوى العربي ، صص ٨٥٣-٨٩.	٢٠١١
٣٥	التحديات التي تواجه محركات البحث في استرجاع المحتوى العربي على الإنترنت: دراسة تحليلية	مؤمن النشرتي	لسجل العلمي لمؤتمر المحتوى العربي ، صص ٧١٥-٨١٤.	٢٠١١
٣٦	اقتراح خدمات الويب الدلالي كحل لاحتياجات المعمارية الموجهة للخدمات	ثابت سليمانيو المحسن رواشد	المجلة العربية لعلوم و هندسة الحاسوب، مج٤، ع٣	٢٠١٢

٣٧	محركات البحث في النصوص العربية	مروان البواب	المؤتمر السابع لمجمع اللغة العربية بدمشق	٢٠١٢
٣٨	تطبيقات الـويـب الـدلالـي في بيئة المعرفة	علي بن ذيب الأكلبي	مجلة مكتبة الملك فهد الوطنية مج ١٨، ع٢	ذو الحجة ١٤٣٣هـ / مايو - نوفمبر ٢٠١٢م
٣٩	الحاسوب في البحث اللغوي : لسانيات المدونات اللغوية أنموذجا	محمود إسماعيل صالح	الرياض: جامعة الأميرة نورة .	٢٠١٢
٤٠	المعجم التفاعلي للغة العربية	عبد الحق الخواجة وآخران	المؤتمر الدولي لعلوم وهندسة الحاسوب باللغة العربية في دورته الثامنة- مصر	٢٠١٢
٤١	اقتراح خدمات الويب الدلالي كحل لاحتياجات المعمارية الموجهة للخدمات	ثابت سليماني والمحسن رواشد	المجلة العربية لعلوم و هندسة الحاسوب، مج٤، ع٣	٢٠١٢
٤٢	نحو آلية لتوليد جذاذات المعاجم العربية	حامد السحلي	ندوة معجم اللغة العربية التاريخي المركز العربي للأبحاث - الدوحة	٢٠١٢
٤٣	تطبيقات الـويـب الـدلالـي في بيئة المعرفة	علي بن ذيب الأكلبي	مجلة مكتبة الملك فهد الوطنية مج ١٨، ع٢	ذو الحجة ١٤٣٣هـ/ مايو - نوفمبر ٢٠١٢م

٤٤	نحو معجم مفهرس للمصطلحات العربية الموحدة في ضوء اللسانيات الحاسوبية ومشروع الذخيرة العربية	عمر محمد أبو نواس	مجلة الدراسات اللغوية والأدبية، ع.١، ماليزيا	٢٠١٣
٤٥	نحو تأصيل منهجي لبناء أنطولوجيا اللغة العربية	مصطفى جرار	المؤتمر الدولي الثاني للغة العربية – دبي	٢٠١٣
٤٦	تطبيق العنقدة المتعددة المستويات على نص القرآن الكريم	محمد سعيد دسوقي	المجلة الدولية للتطبيقات الإسلامية في علم الحاسب والتقنية ، المجلد ٢ ،العدد ١ مارس	٢٠١٤
٤٧	المعجم الإلكتروني العربي المختص: قراءة نقدية في نماذج مختارة	أنور الجمعاوي	المؤتمر العربي الخامس لترجمة، فاس – المغرب	٢٠١٤
٤٨	أنطولوجيا الويب الدلالي ودورها في تعزيز المحتوى الرقمي: دراسة في المفاهيم والبنية الهيكلية والخدمات التفاعلية في البوابات الدلالية للتعلم الإلكتروني	أحمد فرج أحمد	مجلة المكتبات والمعلومات العربية، ع٤، مج٢٥	أكتوبر ٢٠١٥
٤٩	نحو شبكة للكلمات العربية لأغراض الصناعة المعجمية	المعتز بالله السعيد	المجلة المصرية لهندسة اللغة، ع٣	٢٠١٥

٥٠	البنك الشجري النحوي: بناؤه وتوظيفه في إطار تقنيات الذكاء الاصطناعي	أحمد روبي محمد	مركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية	٢٠١٧
٥١	استخدام تقنية المعلومات للبحث في القرآن العظيم بالرسم العثماني دراسة تقييمية للمواقع القرآنية	أمير الحمامي	المجلة الدولية للتطبيقات الإسلامية في علم الحاسب والتقنية، مج٥، ع١	٢٠١٧
٥٢	منهجية فهم التعبيرات الاصطلاحية وطرق تجهيزها للمعالجة الآلية	سلوى حمادة	اللسانيات الحاسوبية واللغة العربية إشكالات وحلول - كنوز المعرفة - الأردن	٢٠١٨
٥٣	المصطلحية بين رهان المعرفية والمعالجة الرقمية	خالد اليعبودي	اللسانيات الحاسوبية واللغة العربية إشكالات وحلول - كنوز المعرفة - الأردن	٢٠١٨
٥٤	الدخلة البسيطة والدخلة المركبة من منظور لساني حاسوبي	دكيكي عبد الواحد	اللسانيات الحاسوبية واللغة العربية إشكالات وحلول - كنوز المعرفة - الأردن	٢٠١٨

٢-٣-٤-٥- خلاصة:

يظهر أن اهتمام اللسانيين في هذا الجانب كان محدودًا للغاية، على الرغم من الصلة الواضحة بين مجال الأنطولوجيا ودراسات علم الدلالة اللغوي والثقافي، ويلحظ كذلك أنّ اهتمام الحاسوبيين ظل بمعزل عن اهتمام اللسانيين بالمستوى الدلالي للغة العربية.

ويبدو أن هذا المبحث الذي يضم المستوى المعجمي والدلالي من مستويات اللغة في البحث اللساني يعمد إلى التعامل مع هذا المستوى من جانب عملي يجمع عمليتي التحليل والتوليد في

آن؛ أو بتعبير أدق، يفترض هذا المجال ضرورة وفرة (الأنطولوجيا) بما تضمه من عمليات لتحقيق غاية عملية أخرى في إطار حوسبة اللغة، ولذلك يميز علم الحاسوب اللساني التعامل مع المستوى الدلالي بعده غاية في ذاته عن الاهتمام به بعده أداة تعين على عمليات حاسوبية أخرى مثل: البحث عن النصوص والمحتوى على الشابكة، أو تدقيق عمليات الترجمة بتحليل الروابط بين مجموعة من الكلمات في الجمل بما يتجاوز المعجم إلى الروابط المفاهيمية المشكلة تداوليًا باختلاف سياقات استخدام الدلالات تاريخيًا، أو معرفيًا وما إلى ذلك. وعلى هذا الأساس نظر إلى المستوى الدلالي في اللغة على أنه جزء مما يعرف بالموارد.

وقد أظهرت الدراسات أن هذا المبحث الذي وجد عناية واسعة في كثير من جزئياته؛ مثل الدراسات حول المعجم، أو آلية البحث عن النصوص والموضوعات، إلا أن الاهتمام بتأسيس "خرائط مفاهيمية" (أنطولوجيا)، لم تدرك قيمته إلا في السنوات القليلة الماضية، ولعل هذا ما يفسر ما كان يُشتكى منه في دراسات المستوى النحوي أو الصرفي من مشكلات الدلالة التي تواجه عملية توليد الجمل، وكانت الإشارات التي وردت في بعض هذه الدراسات تتحدث عن الدلالة لا عن (الأنطولوجيا).

٢-٣-٤-٢- الذخيرة اللغوية:

عادة ما يشار إلى الذخيرة اللغوية بأنها مجموع النصوص المحفوظة والمدونة على الحواسيب والشابكة في صورها المتعددة كتابة ونطقًا، وقد كانت جهود اللساني الجزائري عبد الرحمن الحاج صالح مبكرة في الالتفات إلى هذا المجال وإدراك أهميته في حوسبة اللغة العربية، وقد تباينت تسميات هذا المبحث، من ذلك: "الذخيرة اللغوية"، و"المدونات اللغوية".

٢-٣-٤-٢-١- الذخيرة اللغوية العربية، عبد الرحمن الحاج صالح، مجلة مجمع اللغة العربية الأردني، ع٣٠، ١٩٨٦.

يعد بحث عبد الرحمن الحاج صالح أول بحث أثار الاهتمام بفكرة "الذخيرة العربية"؛فقدم تصوره عن مشروع الذخيرة العربية في المؤتمر الخامس للتعريب الذي عقد في عمان سنة ١٩٨٥، ونشر في مجلة المجمع الأردني للغة العربية سنة ١٩٨٦، وكانت فكرته تتمثل في تخزين عدد هائل من النصوص العربية لا مجرد كلمات، حتى يتمكن الباحث العربي أيًا كان وأينما كان من العثور عليمعلومات شتى من واقع استعمال العربية بكيفية آلية وفي وقت وجيز، وذلك بإنجاز مخزن آلي للغة العربية المستعملة بالفعل؛ يتضمنأمهات الكتب التراثية الأدبية والعلمية والتقنية وغيرها، والإنتاجالفكري العربي المعاصر،إضافة إلى عدد كبير منالخطابات والمحاورات العفوية بالفصحى في شتى الميادين.

ورأى أن الذخيرة اللغوية "عبارة عن قاموس جامع للألفاظ العربية"، يختلف عن غيره من القواميس في صفات أساسية هي:

١ - سيكون له ثلاثة أشكال:

- شكل تسجيل في ذاكرة الحاسوب.
- شكل جذاذية عادية من جهة ومصغرة (ميكرو فيشات تحتوي كل واحدة على ٦٠ صفحة) من جهة أخرى.
- شكل كتاب عادي (موسوعة لغوية).

٢- يحصر جميع الألفاظ التي استعملت في أي نص من النصوص التي وصلتنا من أمهات الكتب القديمة والحديثة والآثار الأدبية والعلمية والتقنية منذ العصر الجاهلي حتى عصرنا الحاضر، مع الإشارة إلى انتماء الكلمة أو العبارة إلى الفصيح المسموع عن الفصحاء، أو المولّد الذي جاء على قياس كلام العرب، إضافة إلى المعاجم العربية.

٣- يذكر كل السياقات (الحقيقية) التي ورد فيها اللفظ ولا يخترع الأمثلة كما تفعله القواميس

الحديثة، ويثبت جميع سياقاته من أمهات الكتب والآثار الأدبية والعلمية التي ورد فيها اللفظ مع ذكر المرجع بدقة، ولا يكتفى بالسياق الواحد.

٤- ترتب فيه الأوضاع اللغوية في ذاكرة الحاسوب بطرق عدة: ترتيب أبجدي عام (الانطلاق من الألفاظ)، وترتيب أبجدي بحسب مجالات المفاهيم (الانطلاق من المعاني)، وترتيب بحسب درجة تواتر الكلمة (عدد المرات التي ظهرت في النصوص)، وترتيب بحسب درجة شيوع الكلمة أو ذيوعها في البلدان العربية، وترتيب بحسب العلوم والفنون.

ويرى أن الذخيرة تنقسم إلى قسمين:

- بنك المعلومات اللغوية، ويندمج فيه بنك المصطلحات، وهو عبارة عن رصيد ضخم جدًا جمعت فيه المادة الخام (الألفاظ وسياقاتها) التي دونها وجردها الباحثون مع ذكر كل المعلومات الإضافية الضرورية (التواتر والشيوع والمرجع أو مصدر الأخذ).
- المعجم المحرّر، وهو عبارة عن موسوعة يحرر فيها العلماء بحوثًا حول كل لفظة، فكل باب أو مدخل من هذا المعجم يحتوي على ما يلي:

١. تحليل دلالي للفظة انطلاقًا من السياقات وحدها ثم تحديدات علماءاللغة القدامى إن وجدت وذلك بـ:

- التوضيح الدقيق:
  - للمعنى الوضعي للمادة الأصلية(الجذر).
  - للمعنى الوضعي والمعاني الفرعية لكل كلمة اشتقت من تلك المادة (بالتمييز بين المعانيالفنية وغير الفنية).
- ذكر المقابل الإنكليزي والفرنسي لكل كلمة إن وجد أو ما يقرب منهمع بيان الفوارق التصورية.

٢. تعليق نحوي صرفي وجيز (وصوتي وهجائي إن اقتضى الحال)بالاعتماد على ما

ذكره علماء اللغة قديمًا (مع ذكر المراجع).

٣. تعليق تاريخي للمادة وفروعها (انطلاقًا) من تحليل النصوص أوالمقارنة بينها:

- بيان أصل الكلمة إن كانت من الدخيل وتفسير تكييفها.
- ذكر تاريخ أول ظهور للكلمة في النصوص (الأصلية والدخيلة).
- ذكر تاريخ أول تحول دلالي للكلمة (والسياقات التي ظهرت فيهاالمعاني المستحدثة).
- ذكر تاريخ آخر ظهور لها إن اختفت في الاستعمال.
- وصف إجمالي تفسيري للتطور اللفظي والدلالي للكلمة.
- بيان نظائر الكلمة في اللغات السامية (مع ذكر المواد الأصلية).

٤. ذكر درجة تواتر الكلمة حسب العصور والبلدان، وبالنسبة للآثار العلمية أوالأدبية إن اقتضى الحال.

٥. بيان شيوع الكلمة الجغرافي (حسب العصور أيضًا).

٦. ذكر المترادفات والأضداد للكلمة إن وجدت، والألفاظ التي تجانسها في المفهومكذلك.

٧. ذكر الدراسات التي خصصها العلماء لهذه الكلمة أو تلك المادة.

وقد أتبع ذلك ببيان أهمية الذخيرة اللغوية، ومقترح لتوزيع الجهود لإنجاز ذلك، ولا شك أنه بذلك قد قدم تصورًا مهمًا لهذا المشروع الذي ظل يسعى إلى تحقيقه حتى تبنته جامعة الدول العربية سنة ١٩١٠.

٢-٣-٤-٢-٢-نحو آلية لتطوير المدونات لتوليد جذاذات المعاجم العربية، حامد السحلي، أعدت هذه الورقة البحثية لندوة معجم اللغة العربية التاريخي التي يرعاها المركز العربي لأبحاث ودراسات السياسات في الدوحة ٩/ ١١/ ٢٠١٢.

تهدف هذه الدراسة إلى بناء آلية رقمية لتخزين المضمون الدلالي والاشتقاقي للمفردة باللغة العربية ومعالجتها، وتندرج - إلى حدما - ضمن المدونات المحوسبة وسبل الاستفادة منها، وتقترح آلية لإنشاء الجذاذة المعجمية آليًا انطلاقًا من تطوير المدونات، ومن ثم فالدراسة تقترح آلية جديدة لبناء المعاجم.

وإذا كان المعجم الاشتقاقي (التأثيلي)يهتمبالمفردة وجذرها ووزنها وأصلها وصلتها بشقيقات العربية أو اللغات الأخرى والدلالات الممكنة لها مجردة من سياقها، وكان المعجم الحاسوبي يهتم بدلالة الكلمة ضمن سياقها أي عمليًا بالجملة، فإن المعجم التاريخي يركز على تطور دلالات المفردة تاريخيًا، أي دراسة دلالة المفردة ضمن كلام في زمن معين. ومن ثم فإن الحاجة لوسم زمني لمقاطع النصوص إضافة للوسمين الاشتقاقي والإعرابي ذو أهمية بالغة.

والهدف هو التوصل إلى تصميم صيغة رقمية تتضمن في بنيتها هذه المفاهيم الدلالية المترابطة التي هي اشتقاقية وإعرابية وتاريخية للنصوص بحيث تشكل أساسًا لتوليد الجذاذات المعجمية الثلاث، في الوقت نفسه الذي تشكل فيه هذه النصوص الذخيرة العربية المنشودة.

يتطلب هذا المسعى صيغًا رقمية وسيطة، للانتقال من الصيغة الحالية للنصوص العربية غير المشكولة، والخالية من أي وسوم اشتقاقية أو إعرابية أو تاريخية إلى الصيغة المتضمنة كل ذلك، أي الحصول على نص عربي يتضمن الوزن والمعلومات الدلالية والإعرابية ومؤرخًا بدل الصيغة الحالية التي هي الاحتفاظ بالنص مع إضافة لواحق له بعضها يجزأه إلى مفردات، ويلحق به وسومًا للوزنوأخرى للمعنى،ووسومًا إعرابية للنص مجزأً إلى جمل مفيدة، ويضاف لها اجتزاء الشواهد لإلحاقها يدويًا بنص جديد مقروء إنسانيًا هو المعجم التاريخي، بينما الصيغة المقترحة ستجعل النص مقروءًا بما يحمله من دلالات لكل من الإنسان والآلة.

قُسمت الدراسة إلى ثلاث مراحل:

- المرحلة الأولى تعرض تصورًا لتوسيع الترميز المقترح للكلمة العربية ليشمل الدلالات الإعرابية والتاريخية، وهو التوسيع المتوقع فيالتصميم الأولي أساسًا.
- المرحلة الثانية تعرض تصورًا للآليات والترميز الوسيطة للانتقال من النص البسيط في المدونات إلى الترميز المقترح.
- المرحلة الثالثة تعرض تصورًا لتوليد الجذاذة المعجمية من المدونات المرمزة وفق البنية المقترحة.
- ختامًا تقوم الدراسة بمقارنةمختصرةللمحاسن والمساوئ بين استخدام ترميز خاص بالعربية واتباع نهج اللغات الأوربية واستخدام أدواتها نفسها، ولكن هذه المقارنة ستكون تصورية لأنها تتطلب إثباتًا من خلال أدوات برمجية لم تُطور، وكان من أبرزها:

١. تبقى التعديلات المدخلة على الأدوات الغربية قاصرة ومشتتة لعدم وجود استراتيجية ملائمة أو جهة واحدة قائدة.

٢. الصعب جدًا تجميع هذه الأدوات المعدلة في منتجات حقيقية تفيد الناطقين العرب وتحسن استفادتهم من التقانة واللغة، وغالبًا تقوم بهذا شركات عالمية عمالقة على هامش مشاريعها المتعلقة باللغات.

٣. في الغالب يختلف توجه هذه الشركات عن حاجة العرب، ويسيطر حاليًا اتجاه لفصل الفصحى الحديثةعن العربية الفصحى القديمة التي اصطلح بتسميتها عربية القرآن، وهي خطوة ستؤدي إلى إزالة الإعراب وتفكيك البنية الصرفية المطردة كما فعل الإسرائيليون بالعبرية، حتى جامعة الملك سعود اتبعت نفس النسق الذي بات دارجًا.

٤. الترميز يحتاج جهدًابرمجيًا وخبرات تقنية أكبر بكثير من الجهد والخبرات اللغوية

خلافًا للآلية التقليدية التي تعتمد على جهوداللغويين، فضلاً عن أن خبراء حوسبة اللغة التقنيين العرب أقل نسبيًا من الخبراء اللغويين العرب.

٥. لكن خبراء حوسبة اللغة التقنيين أكثر قدرة على الاستفادة من التقنية من اللغويين، ويمكنهم العمل ضمن شبكة متباعدة دون اتصال فيزيائي، وهذا يجعل قدرتهم على إنتاج عمل موحد أكبر بكثير من قدرة اللغويين وأقل كلفة من حيث قيمة الاستثمار المطلوب.

٦. ستنتج الآلية التي يقترحها الباحث نمذجة حاسوبية للغة العربية، إضافة إلى المعاجم والذخيرة العربية المعلّمة، وهو أمر العربية بأشد الحاجة له كيلا تبقى على هامش التطبيقات العالمية.

٧. سيكون هذا النظام متاحًا للجميع مفتوح المصدر، وليس ملكًا للشركات العمالقة غير العربية، كما هي النتيجة المتوقعة للجهد الحالي.

٢-٣-٤-٢-٣- الذخيرة اللغوية لتراكيب القرآن الكريم، محمد زكي خضر، الفصل السادس في كتاب: "اللسانيات الحاسوبية واللغة العربية إشكالات وحلول"، إعداد وتنسيق عمر مهديوي، دار كنوز المعرفة، عمان، ط١، ٢٠١٨.

سعت هذه الدراسة إلى الاهتمام بالذخيرة اللغوية من جهة التراكيب لا الكلمات المفردة في القرآن الكريم، واستندت، لتحقيق هذ الغاية، إلى التراكيب المتشابهة لفظًا في القرآن الكريم مرتبة وفق جذور الكلمات المكونة للتركيب لا الكلمات نفسها؛ أي أن التراكيب التي تشترك كلماتها بالجذور نفسها تعتبر متشابهة ومجتمعة مع بعضها، لذلك فقد استثنيت التراكيب التي تظهر في القرآن الكريم مرة واحدة لعدم وجود تكرار لتلك التراكيب.

أشارت الدراسة إلى أنّ الأسماء والأفعال في اللغة العربية غالبًا ما تشتق من جذور الكلمات الثلاثية وأحيانًا الرباعية والخماسية، لذلك فإن هذه الذخيرة يمكن أن تكون أساسًا لكثير من

الأبحاث المستقبلية المتعلقة بعلاقة كلمات اللغة العربية مع بعضها، خاصة في الأبحاث ذات العلاقة بالمعاني والدلالة. ويؤمل أن تخدم هذه الذخيرة ربط المفاهيم في القرآن الكريم مع كتب الحديث النبوي والفقه والعلوم الشرعية بالإضافة إلى اللغة العربية.

تعرض الدراسة - بداية- أهمية الذخيرة اللغوية، وطرق إعدادها، والجهد الذي بذل لإعداد ذخيرة القرآن الكريم باللغة الإنجليزية بمبادرة من جامعة ليدز البريطانية، ثم تشير إلى المعجم المفهرس لكلمات القرآن الكريم الذي ألفه محمد فؤاد عبد الباقي، ثم المعجم المفهرس للتراكيب المتشابهة لفظًا في القرآن الكريم من تأليف الباحث نفسه، وأظهر خطته في جمع التراكيب المشتركة في جذورها، ثم أعاد تصميم البيانات التي احتواها الكتاب لكي تكون مناسبة لوضعها على شكل ذخيرة لغوية تحوي مداخل كل تركيب منها على سطر مع مراعاة ما يأتي:

١- وضعت التراكيب وفق ورود أول تركيب منها في المصحف. عند ورود تركيب جديد يدرج معه كل التراكيب التي تشترك معه في سلسلة الجذور بدءًا من التراكيب التي تتفق معه في الكلمات والتشكيل...

٢- كل مدخل يحوي معلومات عن محله وروده في المصحف **RF**، والإشارة إلى أول تركيب ورد في المصحف متفق معه بالجذور **RPH**.

٣- يعاد التركيب عند وروده في محله من آيات المصحف وسوره بحيث يشير إلى وجود مجموعة التراكيب المتفقة بالجذور عند أول ورودها لكي يكون بالإمكان الرجوع إليها. وقد أشير إلى هذه المداخيل بالرمز **RF:REF**.

٤- أدرج نص التركيب بالرمز **PH** وجذوره بالرمز **RT**.

٥- على هذا الأساس فقد حوى كل مدخل ستة أرقام تحت رمز **RF**، وستة أرقام تحت الرمز RPH، ونص التركيب تحت الرمز PH، ونص الجذور تحت الرمز **RT**.

وقد خلصت الدراسة إلى عدد من الاستنتاجات؛ فذخيرة التراكيب المكونة من أكثر من

كلمة مهمة في معالجة اللغات الطبيعية، وبالأخص في الترجمة الآلية. وقد ركزت هذه الدراسة على التراكيب من أكثر من كلمة في القرآن الكريم بالاستناد إلى جذور الكلمات المكونة للتركيب فقط.

جاءت التراكيب في البحث على ثلاثة أنواع: النوع الأول هو التراكيب المتشابهة بالكلمات مع تشكيلها، والنوع الثاني الكلمات المتشابهة لكن تشكيلها مختلف، والنوع الثالث المتشابهة في جذورها لكنها مختلفة الكلمات والتشكيل. والأنواع الثلاثة تشكل مجموعة واحدة مشتركة بالجذور نفسها.

٢-٣-٤-٢-٤- مسرد الدراسات:

الرقم	عنوان الدراسة	المؤلف	مكان النشر	تاريخ النشر
١	استخدام الآلات الحاسبة الإلكترونية في دراسة ألفاظ القرآن الكريم	علي حلمي موسى	مجلة عالم الفكر (الكويت) المجلد ١٢،العدد ٤	١٩٨٢
٢	الذخيرة اللغوية العربية	عبد الرحمن الحاج صالح	مجلة مجمع اللغة العربية الأردني	١٩٩٨٦
٣	مسألة المصطلحات في تعريب الحاسبات	سعدي الحاج بكري	المجلة العربية للعلوم (تونس)	١٩٨٨
٤	المعالجة الآلية للكلمات وبحث النص في الأعمال المصطلحية	حسين الهبائلي	أشغال المؤتمر الرابع للسانيات: اللسانيات العربية والإعلامية- الجامعة التونسية تونس	١٩٨٩

٥	مشروع نظرية حاسوب لسانية في بناء معاجم آلية للغة العربية	محمد الحناش	مجلة التواصل اللساني، المجلد٢، العدد٢	١٩٩٠م
٦	نظرية حاسوبية لسانية لبناء المعاجم الآلية	محمد الحناش	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	١٩٩٢
٧	استخدام نظام المستشار في بناء المكانز العربية (النظم العربية المتطورة)	عبد الجبار العبد الجبار	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	١٩٩٢
٨	الحاسب الآلي وصناعة المعجم العربي	محمود فهمي حجازي	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	١٩٩٢

٩	تعريب المصطلحات المستعملة في الحواسيب الصغروية	أحمد بوعزي	السجل العلمي لندوة استخدم اللغة العربية في تقنية المعلومات) مطبوعات مكتبة الملك عبد العزيز،الرياض	١٩٩٢
١٠	نظام هبير عن اللغة العربية	سلوى أحمد الجمل	ندوة استخدام اللغة العربية في تقنية المعلومات- مكتبة الملك عبد العزيز العامة الرياض	١٩٩٣
١١	المعاجم العلمية العربية المختصة ودور الحاسوب	إبراهيم بن مراد	الموسم الثقافي الرابع عشر- مجمع اللغة العربية الأردني	١٩٩٦
١٢	المعجم الحاسوبي: إحصاء الأفعال العربية في المعجم الحاسوبي	محمد مراياتي، يحي مير علم ، محمد حسن طيان	يروت: مكتبة لبنان	١٩٩٦
١٣	الحاسوب والصناعة المعجماتية	عبد الغني أبو العزم	اللسان العربي العدد السادس والأربعون	١٩٩٨
١٤	الحاسوب والصناعة المعجماتية، اللسان العربي	عبد الغني أبو العزم	اللسان العربي العدد السادس والأربعون	١٩٩٨
١٥	البنك الآلي السعودي للمصطلحات (باسم)	عبد الرحمن بن عبد العزيز الفاضل	جلة اللسان العربي، العدد ٤٧،	١٩٩٩
١٦	المعجم العربي التأريخي الآلي	علي حلمي موسى	مجمع اللغة العربية الأردني	٢٠٠٢

١٧	الجوانب البرمجية في إعداد المعجم المفهرس للتراكيب المتشابهة لفظًا في القرآن الكريم	محمد زكي خضر	مجلة دراسات، الجامعة الأردنية، العدد ١	٢٠٠٤
١٨	رهانات نظرية حاسوب-لسانية في بناء معجم آلي للتعابير المسكوكة	عزالدين غازي	مجلة الدراسات المعجمية المغرب، ع ٦	٢٠٠٧
١٩	استخدام ذخائر النصوص لاستخلاص المصطلحات المتخصصة	عبدالمحسن عبيد الثبيتي	لندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية ، ص ٣١-٣٨.	٢٠٠٧
٢٠	المعجم العربي في ضوء اللسانيات الحاسوبية	محمد غاليم	الندوة الدولية الأولى عن الحاسب واللغة العربية : الأوراق البحثية، صص ١٥٧-١٦٦.	٢٠٠٧
٢١	المعجم العربي التفاعلي التصميم المفاهيمي للمشروع	حسين إبراهيم وآخرون	مدينة الملك عبد العزيز للعلوم والتقنية والمنظمة العربية للتربية والثقافة والعلوم- الرياض -تونس الرياض	٢٠٠٨

٢٢	نحو معجم عربي معاصر	محمد محمد حلمي هليل	بحوث الاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية. الرياض: مدينة الملك عبدالعزيز للعلوم والتقنية.	٢٠٠٨
٢٣	المعجم الحاسوبي للغة العربية	محمد حسن عبد العزيز، محمد يونس الحملاوي والمعتز بالله السعيد طه	لاجتماع الثاني لخبراء المعجم الحاسوبي للغة العربية المنعقد بمدينة الملك عبدالعزيز للعلوم والتقنية في الرياض	٢٠٠٨
٢٤	عمل المدونات والفهرسة الآلية	سلوى السيد حمادة	المجلة العربية لعلوم و هندسة الحاسوب، مج٢، ع٣	٢٠٠٨
٢٥	منيج إعداد المعجم العربي الحاسوبي	مروان البواب	الاجتماع الثاني لخبراء المعجم الحاسوبي التفاعلي للغة العربية-مدينة الملك عبد العزيز للعلوم، الرياض	٢٠٠٨
٢٦	معمارية المعجم العربي الالكتروني: رؤية جديدة لمعجم آلي مُبنين	عزالدين غازي ومحمدهلال	مجلة الدراسات المعجمية، المغرب، ع٨	٢٠٠٩

٢٧	استخراج ومعالجة المصطلحات: تجربة البيئة المجانية المفتوحة المصدر نوج (NooJ)	عزالدين غازي ومحمد لهلال	المنتدى المصطلحي، سوسة، تونس	٢٠٠٩
٢٨	قراءة في تركيبة المعاجم الحاسوبية التفاعلية بَيْئَةُ التطوير اللغوية (NooJ) نموذجاً	عزالدين غازي	ورقة قدمت في الندوة العلمية حول "المعجمية والقاموسية والمصطلحية والمقاربات اللسانية الحديثة بجامعة منوبة"، تونس	٢٠١٠
٢٩	المدونات اللغوية لمتعلمي اللغة العربية: نظام لتصنيف و ترميز الأخطاء اللغوية	عبدالله يحي الفيفي	السجل العلمبي للمؤتمر الدولي لعلوم وهندسة الحاسوب في اللغة العربية في دورته الثامنة (٢٦-٢٨ ديسمبر، ٢٠١٢) جامعة القاهرة.	٢٠١٢
٣٠	التشاركية في إغناء معجم اللغة العربية التفاعلي	ندى غنيم وآخرون	مجلة جامعة دمشق للعلوم الهندسية، المجلد ٢٩ – العدد الأول	٢٠١٣
٣١	حوسبة المعجم العربي: الواقع والآفاق	جيلالي بن يشو	مجلة أرتين – المرجع الأول لطلاب الأدب. Art-En.com	٢٠١٣

٣٢	لسانيات المتون وعلوم اللغة	صالح فهد العصيمي	مجلة كلية الآداب والعلوم الإنسانية (فاس ، المغرب) العدد ١٩ ، السنة الخامسة والثلاثون	٢٠١٣
٣٣	معجم محوسب لمعاني الأفعال الثلاثية المجردة في اللغة العربية	إيمان دلول	الجامعة الإسلامية غزة (رسالة ماجستير)	٢٠١٤
٣٤	أبحاث جامعة ليدز في مجال لسانيات المدونات العربية	إيريك أتوك وعبد الله بن يحيى الفيفي	الحرف العربي والتقنية أبحاث في حوسبة العربية – مركز الملك عبد الله بن عبد العزيز الدولي – الرياض	٢٠١٥
٣٥	الذخيرة اللغوية لتراكيب القرآن الكريم	محمد زكي خضر	اللسانيات الحاسوبية واللغة العربية إشكالات وحلول – كنوز المعرفة – الأردن	٢٠١٨

٢-٣-٤-٢-٥- خلاصة:

كشفت الدراسات التي اهتمت بالذخيرة اللغوية وما يتصل بها عن أهمية هذا المجال، وقد بينت أن الأهمية المناطة بتشكيل الذخيرة اللغوية للغة العربية ستخدم المجالات التطبيقية في عمليات الحوسبة الخاصة باللغة العربية مثل الترجمة والتلخيص والبحث عن الموضوعات، ويبدو – من مجمل ما أشير إليه في كثير من الدراسات – أن الجهود المبذولة حتى اللحظة الراهنة ما زالت دون المدى الضروري لنجاح عمليات حوسبة اللغة العربية تلك. والملحوظ على هذه الدراسات اعتمادها على ما هو مخزن في الحاسوب أكثر من اعتماد قواعد للتعريف.

٢-٣-٥- التطبيقات

لا شك أن الغاية التي تسعى إليها دراسات حوسبة اللغة لا تأتي من أهمية اللغة واستيعابها باعتبارها منجزًا إنسانيًا فحسب، وإنما من أجل إلحاق اللغة في مختلف حالات استخدامها بركب التطورات المتلاحقة في عصر أهم ما يميزه دخول التقنية في مجالات الحياة المتعددة، ولذلك فإن الغاية من الدراسات السابقة في مستويات اللغة ومجالات بحثها يكمن في الانتقال بها إلى التطبيق الحاسوبي، والارتقاء به ليعضد الجهد الإنساني في نشاطه وسلوكه اللغوي شأن نشاطاته الحياتية الأخرى.

ولعل أبرز المجالات التي تحتاجها اللغة العربية تكمن في مجال التواصل اللغوي عبر الحاسوب والشابكة، ومن ثم فإن من أهم مجالات التطبيقات الحاسوبية الاهتمام بالتعرف إلى الحرف العربي نطقًا وكتابة؛ لتحقيق غايات التواصل مع الآلة والحواسيب على وجه الخصوص، ثم الاهتمام بالترجمة الآلية بما تمثله من قيمة في التواصل الحضاري المعاصر.

٢-٣-٥-١- التعرف إلى اللغة العربية صوتيًا وكتابيًا:

لقد نظر إلى الصوت اللغوي (الفونيم) على أنه حزمة مائزة من السمات الصوتية، وقد عني علم الأصوات اللغوية بتوصيف الأصوات اللغوية توصيفًا فيزيائيًا وفسيولوجيًا، وكانت حصيلة ذلك تقديم صورة دقيقة لأصوات اللغات الطبيعية، ومن جملة ذلك أصوات اللغة العربية.

وقد وجد المستوى الصوتي اهتمامًا محدودًا من الدارسين العرب في حوسبة اللغة العربية؛ على الرغم مما لهذا المستوى من أهمية بالغة في إدخال النص العربي إلى الحاسوب بعد تطور عمليات الحوسبة الآلية إذ لم تعد تكتفي بطرق الإدخال التقليدية المعتمدة على الإدخال عن طريق لوحة المفاتيح، وإنما بات من المعهود التحدث وترك أمر تحويل المنطوق اللغوي إلى كتابة مباشرة، ومن الملحوظ قلة الدراسات التي اعتنت بهذا المستوى.

ويأتي إلى جانب ذلك نقل أصوات العربية (الفونيمات) كتابيًا، وما يعرف بالأبجدية التي

تمثل الرموز الكتابية لأصوات العربية في التدوين، ولكلا الجانبين المنطوق والمدون أهمية بالغة في معالجة اللغة العربية حاسوبيًا سواء في المدخلات أو المخرجات.

٢-٣-٥-١-١- الحروف العربية والحاسوب، محمد زكي محمد خضر، الموسم الثقافي لمجمع اللغة العربية، حزيران ٢٠٠١.

تبدأ الدراسة بمحاولة تأسيس الرؤية الكامنة خلف طرح هذا الموضوع، وتبدو كما لو أنها استكمال لدراسة مبكرة للباحث نفسه قدمت في مجمع اللغة العربية الأردني ١٩٩٦، تشير الدراسة إلى المراحل التي تظهر تطور كتابة الحروف العربية إلى أن وصلت مرحلة الحوسبة، وتركز على الإشكاليات التي صاحبت حوسبة الحروف العربية؛ مثل اعتبار التشكيل رمزًا مستقلاً، شأنه شأن الحروف. وقد أعطي لبعض الحروف رموزًا عدة كالهمزة بحسب وضعها من الكلمة. وعلى هذا فإن استعمال الرموز في الحاسوب استند إلى خدمة الكتابة بعدها رمزًالا شكلاً، ولم يكن هذا الاستعمال خدمة أمينة لمحتوى الكلمة العربية أيضًا.

وإذا ما أريد لهذا الوضع أن يصحح، فإن مقدارًا من الذكاء يجب أن يدخل للبرامج العربية التي يتعامل معها المستخدم بحيث تساعده على الاقتراب من اللغة العربية السليمة.

تعتمد عملية قراءة النص العربي من الحاسوب على المسح الضوئي للنصوص المطبوعة أو المكتوبة، وهي عملية سهلة ومتيسرة اليوم، لكن الصور التي تنتج نتيجة هذا المسح الضوئي قد تنتج بعض التشويش وما يسمى بالضوضاء التي قد تشوش عملية التمييز في المراحل اللاحقة، ومن أجل ذلك يقترح الباحث عددًا من خطوات عملية للتمييز:

١- عملية المسح الضوئي، وذلك بتحويل الصفحة إلى مجموعة من النقط السوداء والبيضاء، ويعتمد عدد هذه النقط على دقة الماسح الضوئي وتتراوح دقة الماسحات حوالي ٣٠٠ نقطة لكل بوصة أو أكثر من ذلك.

٢- تتضمن الخطوة التالية عمليات تهيئة تدعى بمرحلة ما قبل التمييز، وهذه المرحلة تتضمن معرفة المواصفات العامة للوثيقة؛ فهل تحتوي على رسوم وأشكال؟ وهل تحتوي على عدد

من الأعمدة أم عمود واحد؟ وهل كان تصويرها مائلاً أم عموديًا؟ وهل تحوي جداول؟ إلى غير ذلك من المواصفات. وخلال هذه العملية يجري تحديد الأسطر ووضع حدود فارقة بين سطر وآخر وتحديد ارتفاع الكتابة في السطر الواحد وهكذا.

٣- أما الخطوة التالية فتتعلق بالتعرف إلى الكلمات والمقاطع والحروف بمختلف أشكالها ومواقعها في الكلمة، ثم استخلاص سمات كل جزء من المقطع، وذلك تمهيدًا للمرحلة اللاحقة التي تقوم بعملية مطابقة هذه السمات مع ما يعرف من مواصفات للحرف أو المقطع الواحد. وهذه أهم وأصعب خطوة في عملية التمييز.

وقد أجريت أبحاث معمقة فيما يخص الكتابة المطبوعة، وكان من السمات التي تميز حرفًا عن غيره: عدد النقاط ومواقعها، واستقامة الحرف وعرضه وارتفاعه وعلوه عن متوسط السطر، والميل، وعدد الحلقات المقفلة وشكلها والحجم الكلي، ومركز ثقل الحرف، ومواقع نقاط الاتصال مع ما قبله أو مع ما بعده، وغير ذلك من السمات الأخرى.

أما الكتابة المكتوبة بخط اليد فلا تزال تعاني من انخفاض في الدقة وكثرة الأخطاء ولبس بين حرف وآخر نظرًا لأن الذكاء الاصطناعي المدخل في هذه البرامج لم يبلغ درجة عالية من النضج والعمق، وتتبع في عملية التمييز وسائل متعددة رياضية ومنطقية وشكلية وقواعد بيانات وغير ذلك. إلا أن ما تجدر الإشارة إليه أنه كلما كانت البرامج أقرب إلى ما يفعله الإنسان في قراءة النصوص كلما كانت الدقة أعلى. ولكن استنباط ما يقوم به الإنسان ومحاكاته في الآلة ليست عملية سهلة على الإطلاق. وتجدر الإشارة إلى أن بعض مناهج تمييز الكتابة تعتمد تمييز المقطع أو الكلمة أصلاً وليس الحرف. ويشبه ذلك ما تدعو إليه بعض المناهج التعليمية للأطفال من تعليم الطفل الكلمات قبل تعريفه بالحروف.

وما يميز اللغة العربية هو تشابه حروف كثيرة مع اختلاف في عدد النقط ومواقعها، وعلى الرغم من أن تلك ميزة مفيدة، إلا أنها قد تكون عائقًا في بعض الأحيان، وذلك عند عدم قدرة الآلة على تمييز النقطة الضعيفة أصلاً، أو الخلط بين النقطة والنقطتين والثلاث نقاط، أو الظن بوجود نقطة نتيجة عدم دقة الماسح الضوئي إلى غير ذلك من صعوبات.

ويفيد في هذه الحقل وجود قاموس بالكلمات المتداولة وإعطاء احتمالات لماهية الكلمة المعنية إن كان لها معنى بوجود النقطة أو عدمه، وكذلك التفريق بين أن يكون فيها نقطة واحدة أو أكثر، فضلاً عن أن سياق الجملة قد يفيد في ذلك؛ لذلك فإن عملية تمييز الكتابة لا تنفصل عن معالجة الجملة العربية ككل ولا عن معالجة الكلمة والمقطع.

وتشير الدراسة إلى أنّ عملية تمييز الحروف المطبوعة تعتبر أسهل وأقل تعقيدًا من عملية تمييز الكتابة المكتوبة بخط القلم، فالكتابة المطبوعة ما هي إلا صور متشابهة تمامًا للحرف نفسه حسب موقعه من الكلمة؛ فحرف الجيم الذي في هذه الصفحة له الشكل نفسه تمامًا إن وقع في أول الكلمة أو حينما يسبقه حرف غير متصل مثل الواو أو الراء. والجيم نفسها تكون متطابقة مع أية جيم أخرى في الصفحة إن وقعت في وسط الكلمة واتصلت من جهتيها. وهكذا. وتتم عملية تمييز الحروف المطبوعة عادة بمطابقة شكل الحرف مع مخزون شكله في الحاسوب، ومن ثم تستطيع الآلة التعرف إلى الحرف بمقارنة شكله مع الشكل المخزون فيها. ويمكن أن تقوم الآلة بتخزين أنواع مختلفة من أشكال الحروف التي تدعى Fonts.

وتشير الدراسة إلى الصعوبات في تمييز الحروف المكتوبة بخط اليد، ولعل أبرزها:

- أن الكاتب نفسه لا يعيد كتابة الحرف بالشكل نفسه تمامًا عند إعادة الحرف نفسه في موقع آخر من الكتابة. فضلاً عن أن الأشخاص المختلفين لا تتطابق كتاباتهم نهائيًا مهما حاولوا ذلك، لذلك فإن عملية تدريب الحاسوب على التعرف إلى الكتابة اليدوية لا يزال في مراحل التطوير والبحث ولم يصل مراحل التطبيق الفعلي.
- أن عملية تمييز الكتابة المكونة من حروف مقطعة كالإنكليزية مثلاً أسهل من تمييز الكتابة المتصلة، لذلك فإن تمييز الكتابة العربية أصعب من تمييز الكتابة الإنكليزية نظرًا إلى أن اللغة العربية لا تكتب إلا متصلة. وقد أجريت أبحاث علمية عديدة على تمييز الكتابة العربية في جامعات عربية وأجنبية ومراكز أبحاث عديدة.
- أن إحدى الصعوبات المهمة في تمييز الكتابة العربية بالإضافة إلى مسألة فصل الحروف

عن بعضها مسألة كتابة بعض الحروف متداخلة عموديًا؛ أي بعضها أعلى من بعض، وقد يؤدي ذلك إلى تداخل الحروف والكلمات مع بعضها.

وترى الدراسة أنه عندما تُلاحق يد الكاتب لحظة بلحظة، يمكن تمييز الحروف بسهولة أكثر من ملاحظة صفحة مكتوبة مسبقًا، فالحاسوب يقوم بالتعرف إلى الحروف أولاً بأول، وهذه المعلومات التي تُنقل إلى الحاسوب هي معلومات مفيدة تسهل عملية التمييز، لكن الكتابة يجب أن تتم بقلم خاص وعلى لوحة خاصة متصلين بأسلاك، وهذا بحد ذاته عائق كبير جدًا ولن يكون بديلاً لبرامج كفؤة للتعرف إلى الكتابة بخط اليد.

وتهتم الدراسة – من جانب آخر– بتحليل الحاسوب أصوات العربية والتعرف إليها، وهي عملية تشهد عددًا من الصعوبات، فتمييز الأصوات يحتاج إلى تحليل للإشارة الصوتية ومعالجة كل حيز من الترددات على حدة لمعرفة الخواص الدقيقة لكل مكون من مكونات الصوت، وتعتبر عملية تقسيم الصوت إلى قطع متتالية من أصعب هذه المهام نظرًا لاختلاف مواصفات الصوت بين متكلم وآخر، واختلاف نطق المتكلم نفسه وسرعة كلامه ودقة نطقه وكيفية ربط الكلمات مع بعضها.

وتجري عملية التمييز بعد عملية التقطيع، وهي لا تقل صعوبة عن سابقتها للأسباب نفسها. إن عملية تمييز الأصوات لا تقل تعقيدًا عن تمييز الحروف المكتوبة بخط اليد نظرًا إلى أن كليهما مختلف بين شخص وآخر، وللشخص نفسه بين حرف وآخر.

إن إحدى التطبيقات المهمة المؤمل منها تمييز الأصوات العربية هو تطوير وسيلة إملاء آلية بحيث يتم تحليل الصوت وتحويله إلى كتابة صحيحة، وبالطبع فإن تخليص الصوت المنطوق من شوائبه وإضافة ما لم ينطق بشكل كامل أو صحيح إلى الكتابة هو أمر في غاية الأهمية.

إن المصاحف اليوم بما تحويه من علامات وقف وترتيل هي أمثلة على ما تحويه اللغة من مكنونات لفظية غير ظاهرة للعيان عند تدوينها بالحروف فقط. فهناك الحركات الخاصة بالشكل وهناك العلامات الأخرى اللازمة لضمان دقة اللفظ.

والنظرة السريعة للحركات في اللغة العربية توحي بوجود ثلاث حركات هي الفتحة

والضمة والكسرة. إلا أنه – في الحقيقة–هناك عدد أكبر من ذلك نتيجة تعاقب الحركات مع حروف العلة (الألف والواو والياء) قبلها وبعدها، فضلاً عن وجود الشدة على حرف العلة، فالكسرة قبل الواو المفتوحة (مثل سوى) هي غير الكسرة قبل الياء المفتوحة (مثل عليّا).

وتشير الدراسة – كذلك – إلى أن بعض الحروف العربية لها أكثر من فونيم؛ كاللام المرققة والمفخمة والراء المرققة والمفخمة، فضلاً عن أن الياء والواو هما حرفان صامتان، وهما حركتان في الوقت نفسه. أما الهمزة فهي حرف صامت بينما الألف هو حركة فقط. وبعض الحروف تنقلب إلى حروف أخرى وفق موقعها في الكلمة وحركتها مثل أسفَرَ (سين)، وأسدَلَ (زاي)، ويسطَع (صاد).

إضافة إلى ذلك تشير الدراسة إلى ضرورة تحليل اللهجات العامية، فمن الناحية العملية لا تقل أهمية عن فهم الفصحى، وذلك أن عملية فهم العامية واستنباط ما يقابلها من فصحى هو أمر في غاية الأهمية، فالمتكلمون بالعامية اليوم هم أكثر بكثير ممن يتقن اللغة الفصحى، لذلك يمكن أن يكون الحاسوب وسيلة فعالة للقضاء على العامية من خلال التصحيح الفوري للحن بالفصحى أو استبدال العامية بالفصحى أو تصحيح الأخطاء اللفظية في تشكيل أواخر الكلمات أو غير ذلك من المعالجات اللفظية للغة العربية. وكل هذا يحتاج إلى دراسات وأبحاث للغة العربية وكيفية برمجة أصواتها.

وتهتم الدراسة –أيضًا – بتكوين الأصوات العربية وتوليدها، وبهذا الصدد فإن تكوين مثل هذه الأصوات في غاية الأهمية نظرًا لأنه وسيلة لتعليم اللغة العربية الصحيحة نطقًا وتشكيلاً وتداولاً وهو وسيلة للعودة إلى الفصحى بدل العامية.

وتنتهي الدراسة إلى بيان أهمية الدراسات الإحصائية للغة العربية، وأن هناك كثيرًا من المهام يجب أن يقوم بها الباحثون العرب في حقل اللغة العربية نحوًا وصرفًا ودلالة وبلاغة ونطقًا من المتخصصين باللغة العربية، وفي حقل برمجة اللغة العربية من أخصائيي الحاسوب، وعلماء اللسانيات فيما يتعلق بعلاقة اللغة العربية مع اللغات الأخرى ترجمة ومفاهيم.

وتنبه الدراسة إلى خطر ترك الأبحاث عن اللغة العربية بيد غير العرب، إذ أصبحكثير من الدراسات المتعلقة باللغة العربية والحاسوب ذات طابع تجاري، وقد أدركت ذلك الشركات

الكبرى ومراكز الأبحاث الغربية؛ لذلك فهي تقوم بأبحاث ثرية متعلقة باللغة العربية وتنتج برامج تباع لمن يتعامل مع اللغة العربية على الحاسوب أو عبر الشابكة، لذلك فإن خدمة اللغة العربية عبر التقنيات الحديثة لا تتم إلا بأيدي أبنائها من خلال مراكز أبحاث رصينة أو جامعات عربية تفهم اللغة وتدافع عنها وتحمل همومها.

إنّ هذه الدراسة – على الرغم من سمتها العمومية – قد أثارت أهم المشكلات التي تعترض هذا الجانب من حوسبة اللغة العربية المتعلق بأصوات اللغة العربية، والتعرف إليها نطقًا وكتابةً، وقدمت تصورًا كان مؤسسًا لدراسات أخرى تالية.

٢-٣-٥-١-٢- تصميم رموز حاسوبية لتمثل الألفبائية العالمية، لمنصور بن محمد الغامدي (٢٠٠٦)، مجلة جامعة الملك عبد العزيز: العلوم الهندسية، م١٦، ع٢، ص ص٢٧-٦٤، (٢٠٠٦م/ ١٤٢٧هـ).

هدفت دراسة الغامدي إلى "تدوين الأصوات الصادرة عن الإنسان على شكل رموز مقروءة تحفظها مواد قابلة للبقاء لفترة زمنية طويلة"، وتبدأ الدراسة بالإشارة إلى نظام "الألفبائية الصوتية الدولية"، والصعوبة التي يجدها الدارس واللساني العربي في استخدام هذه الرموز لأنها قائمة على رسم الحرف الروماني، ولا تنسجم مع رسم الحرف العربي عند وضعها في نص مكتوب بالحرف العربي، ولذلك لا بدّ من وضع ترميز يناسب الحرف العربي من جهة، ويتلاءم مع طبيعة الصوت في اللغة العربية وحالاته من مهموس ومجهور وتفخيم، وقد عرض الباحث محاولات الباحثين لتغيير نظام الترميز الصوتي وتعديله بما يناسب اللغة العربية؛ فعرض مقترح خليل محمد عساكر الذي قدمه في مجمع اللغة العربية في القاهرة سنة ١٩٥٥م، والذي قدم فيه مقترحًا لطريقة كتابة اللهجات العربية الحديثة، وعرض مقترح عبد العزيز السويل لعلماء الصوتيات العرب لتقديم "ألفبائية عربية موحدة" سنة ١٩٨٦، ثم مقترح السفروشني لإضافة أحد عشر رمزًا جديدًا تضاف إلى الحروف العربية، فأضاف رموزًا جديدة للفتحة والكسرة مع الإبقاء على رمز الضمة، وغيرها من المقترحات.

ويرى الباحث أن المقترحات السابقة كانت – في الغالب – محاولات لحل مشكلة محدودة، ومن ثمّ كان لا بد من إيجاد رموز صوتية تغطي جميع أصوات اللغات البشرية بما فيها اللهجات العربية واللغات التي تكتب بالحرف العربي كالفارسية والأوردو وبعض اللغات الإفريقية، ويسمي نظام هذه الرموز بـ (ألفبائية صوتية دولية عربية)، ويُختصر بـ(أصدع).

وقد نتج عن هذه الورقة الكشف عن خطين عربيين للرموز الصوتية الدولية (أصدع) يغطيان جميع رموز الألفبائية الصوتية الدولية البالغ عددها ١٨٦ رمزًا، وقد أخذ في الحسبان عند وضعهما التوزيع الشائع للحروف العربية على لوحة المفاتيح في نظام مايكروسوفت.

ولم يؤخذ في الاعتبار التوزيع الإحصائي للحروف الكتابية لأسباب عدة منها: أولاً: أن الدراسات السابقة لتوزيع الحروف العربية على لوحة المفاتيح لم يكن لها تأثير على واقع الاستخدام اليومي، فعلى الرغم من الدراسات السابقة في هذا المجال إلا أنها لم تخرج إلى حيز التطبيق، وبقيت مجرد بحوث يرجع إليها الباحثون. ثانياً: العدد الكبير لرموز (أصدع) الذي بلغ ١٨٦ رمزًا، هذا العدد يجعل من الصعوبة بمكان تحديد تردد استخدام كل رمز خاصة أنها تستخدم من عدد محدود من المتخصصين للتوضيح والبحث والدراسة فقط، وليس لإدخال نصوص كما هي الحالة في استخدام حروف الكتابة. ثالثاً: عند الرجوع إلى كيفية استخدام الألفبائية الصوتية في الخط SIL، كان استخدامها أكثر تعقيدًا من الطريقة الدولية التي عرضت في هذه الورقة، وعند الرغبة في إدخال أي رمز يجب أولاً الضغط على مفتاح Alt ثم إدخال رمز الحرف المكون من عدة أرقام، فعلى سبيل المثال، عند إدخال رمز العين /ع/ الذي يمثله في الألفبائية الصوتية الدولية الرمز: ʕ، فإنّ المستخدم يضغط على Alt ثم يُدخل الأرقام (٠١٩٢).

ويمكن إدخال رموز (أصدع) في النص العربي بسهولة حسب رأي الباحث؛ فهي تنسجم مع الكتابة العربية من حيث الاتجاه؛ من اليمين إلى اليسار، وتتسم رموز (أصدع) كذلك باستقلالية كل رمز حيث لا تشبك مع بعضها كما يحدث في حالة الكتابة بالحرف العربي، وهذا يسهل إضافتها إلى الموجات الصوتية سواء في الأشكال التوضيحية أم في الكتابة الصوتية.

لقد قدمت هذه الورقة خطين حاسوبيين للرموز الصوتية الدولية العربية تقوم على الحرف

العربي، وهي متوافقة مع طبيعته، فضلاً عن سهولة استخدامها وربطها بالألفبائية الصوتية الدولية IPA، وهذا يسهل على المختصين في هذا الميدان من الدارسين والباحثين العرب الرجوع إليها ومعرفة الخصائص الدقيقة لكل رمز من حيث مخرج الصوت الدال عليه وطريقة إخراجه، وبهذا فإن مجموع الرموز يساوي ٢٨ صائتًا و٨٠ صامتًا و٧٨ رمزًا، أي ما مجموعه ١٨٦ رمزًا صوتيًا، وهو ما يؤمل أن يستفيد منه الباحثون في الدراسات الصوتية واللسانية والحاسوبية بما في ذلك الكتابة الصوتية للعبارات المنطوقة بغض النظر عن اختلاف اللغة أو اللهجة.

٢-٣-٥-١-٣- الصوت اللغوي والحوسبة الآلية، لراضية بن عربية، مجلة الأكاديمية للدراسات الاجتماعية الإنسانية، ع ٢، الجزائر، (٢٠٠٩).

تندرج هذه الدراسة في إطار الدراسات التأسيسية التي تسعى إلى توصيف المستوى الصوتي للغة العربية؛ ليكون ذلك سبيلاً إلى الاستفادة منها في حوسبة اللغة العربية.

وقد بدأت الدراسة بتعريف الصوت بأنه "ظاهرة طبيعية، وشكل من أشكال الطاقة، وهو يستلزم وجود جسم في حالة اهتزاز أو تذبذب، وهذه الاهتزازات والذبذبات تنتقل عبر وسط معين حتى تصل أذن الإنسان"، ويأتي هذا التحديد ملائمًا لمسعى الحوسبة بالتركيز على الجانب المادي الفيزيائي للصوت، ثم يُميز الصوت اللغوي عن الصوت الطبيعي غير اللغوي بالاستعانة بآراء الدارسين قدماء ومحدثين، مع الإشارة إلى ما يمثله الصوت اللغوي من ظاهرة طبيعية وسيكولوجية في الوقت ذاته.

وتشير الدراسة إلى مراحل الصوت ومروره بثلاث مراحل:

١- مرحلة الإصدار: يمثلها الجانب النطقي للمتكلم، أي خروج الأصوات من فم المتكلم. وهذا مجال علم الأصوات النطقي أو العضوي أو الفزيولوجي.

٢- مرحلة الانتشار: وتتمثّل في الذبذبات أو الموجات الصادرة من فم المتكلم والمنتشرة في الهواء ليصل إلى المستمع، وهذا مجال علم الأصوات الفيزيائي أو الأكوستيكي.

٣- مرحلة الاستقبال: وتتمثل في الجهاز السمعي لدى السامع وكيفية وصول الصوت

الصادر عن المتكلم والمنتشر في الهواء ليصل إلى أذنه، وهذا مجال اختص به علم الأصوات السمعي.

ويأتي هذا التوصيف للمستوى الصوتي في اللغة ليبدو جهازًا قابلاً للتأطير والتجريد، ثم تُعرض أبرز آراء اللغويين العرب الأوائل والمحدثين في تنظيراتهم لأصوات اللغة العربية وصفاتها.

أما الجزء المهم في الدراسة فهو "المعالجة الآلية للصوت اللغوي"، واهتم فيه بثلاثة جوانب:

- الكتابة الصوتية العالمية: وهو نظام ترميزي وضعته المنظمة العالمية للصوتيات يسمى "الألفبائية الصوتية الدولية"[1]، وهو يقدم نظامًا ترميزيًا اعتمادًا على رسم الحروف اللاتيني لتمثيل الأصوات المنطوقة في اللغة العربية بما في ذلك الحركات.
- وطبيعة الصوت وخصائصه، تقدم الدراسة ثلاث خصائص لطبيعة الصوت اللغوي: التردد الفيزيائي الذي يقابل الحدة من الناحيتين السمعية والإدراكية، والشدة الفيزيائية التي يقابلها علو الصوت، والشكل الموجي الفيزيائي الذي يقابله نوع الصوت. لتخلص إلى نتيجة تؤكد طبيعة الصوت الفيزيائية.
- أجهزة تسجيل الصوت: وتسعى الدراسة إلى إيجاد طريقة معالجة الصوت عن طريق ما تسميه "السلسلة الفيزيائية"، وهي تقوم بإدخال الصوت اللغوي إلى جهاز الحاسوب وتحليله باستخدام برنامج محلل الصوت (Speech Analyzer)؛ ليبين المحلل الطيفي كيف يتغير طيف التركيب الصوتي، ثم تُجرى خطوات للتعريف الآلي، وحفظها ثم إخراجها عند الحاجة...، وترى الدراسة في "السلسلة الفيزيائية" في معالجة الصوت آليًا "همزة وصل بين الحاسوبي واللساني، حيث تتضافر الجهود بينهما لخدمة اللغة عامة واللغة العربية بصفة خاصة.

١- أشير إليه في الدراسة باسم "الأبجدية الأصواتية العالمية" مع أن التسمية مقترنة بالترتيب الهجائي (أف باء تاء..) وليس الأبجدي.

فمن الواضح أن هذه الدراسة هي محاولة لرسم خطة للمساهمة بحوسبة اللغة في المستوى الصوتي إلى جانب الحاسوبيين، فالدراسة تندرج - شأنها في ذلك شأن مجمل دراسات اللسانيين - في التوصيف لما يمكن أن ينجز، دون أن تظهر المشكلات الحقيقة التي تواجه عملية حوسبة الأصوات في اللغة العربية على وجه الخصوص.

٢-٣-٥-١-٤- مسرد الدراسات:

الرقم	عنوان الدراسة	المؤلف	مكان النشر	تاريخ النشر
١	استخدام دوالّ الأرجحية في تمييز الأرقام العربية المكتوبة باليد	إيمان القيسي وحسن ناصر	مؤتمر الكويت الأول للحاسوب-جمعية الحاسب الآلي الكويتية	١٩٨٩
٢	خبرات في التعرف على الكلمات العربية المنطوقة المنفصلة	م. أ. حشيش وآخران	وقائع مختارة من ندوة استخدام اللغة العربية في الحاسب الآلي- دار الرازي بيروت	١٩٨٩
٣	معالجة اللغة العربية الطبيعية آليًا	محمد مراياتي	وقائع مختارة من ندوة استخدام اللغة العربية في الحاسب الآلي- دار الرازي بيروت	١٩٨٩
٤	تحويل نص عربي مكتوب إلى نص كتابي	ع. مرادي وآخران	وقائع مختارة من ندوة استخدام اللغة العربية في الحاسب الآلي- دار الرازي بيروت	١٩٨٩

٥	الإدراك الآلي للتضعيف	منصور محمد الغامدي	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢
٦	القراءة الآلية للنص العربي	حازم يوسف عبد العظيم	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢
٧	دراسة صوتية وتمييز حروف العلة الفصحى	عويزرات حاج	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢
٨	نظام تصحيح الهجاء	حسام الدين حسن محجوب	السجل العلمي لندوة استخدام اللغة العربية في تقنية المعلومات مكتبة الملك عبد العزيز العامة الرياض	مايو ١٩٩٢

٩	المعالجة الآلية للكلام المنطوق	سالم الغزالي	ندوة استخدام اللغة العربية في المعلوماتية، المنظمة العربية للتربية والثقافة والعلوم تونس	١٩٩٦
١٠	تعامل الأجهزة والمعدات مع الحرف العربي	محمد مراياتي	ندوة استخدام اللغة العربية في المعلوماتية، المنظمة العربية للتربية والثقافة والعلوم تونس	١٩٩٦
١١	الحروف العربية والحاسوب	محمد زكي خضر	الموسم الثقافي الرابع عشر مجمع اللغة العربية الأردني	١٩٩٦
١٢	تمييز حروف اللغة العربية المكتوبة آلياً باستخدام الشبكات العصبية ذات الانتشار الرجوعي	عاصم عبد الفتاح نبوي، و صبري عبد الله محمود	مجلة جامعة الملك سعود (علوم الحاسب والمعلومات) المجلد٩	١٩٩٧م
١٣	اللسانيات الحاسوبية والترجمة الآلية	سعد عبد الستار مهدي	بيت الحكمة – بغداد	١٩٩٩
١٤	القارئ الآلي للنصوص العربية والأعداد	خضير بن بلبل	مجمع اللغة العربية الأردني	٢٠٠٢
١٥	حفص ملتقى الأصوات والحاسوب	عمرو الجندي	رسالة ماجستير جامعة القاهرة قسم اللغة والدراسات السامية والشرقية	٢٠٠٧
١٦	نظام للتعرف آلي على العناوين البريدية المخطوطة بالعربية	منصف الشرفي٨ وآخرون	المجلة العربية لعلوم و هندسة الحاسوب، ع١	٢٠٠٧

١٧	إصلاح انحناء وتقوس الكتابة في صور الوثائق العربية القديمة	منصف الشرفي وآخرون	المجلة العربية لعلوم و هندسة الحاسوب، ع٢	٢٠٠٧
١٨	مقاربة جديدة لاسترجاع التسلسل الزمني للكتابة المنفصلة باستعمال الخورزميات الجينية	عبد الكريم البعتي وآخرون	المجلة العربية لعلوم و هندسة الحاسوب، مج١، ع٣	٢٠٠٨
١٩	إصلاح انحناء وتقوس الكتابة في صور الوثائق العربية القديمة	أحمد الشريف ومحمد كحيلي	المجلة العربية لعلوم و هندسة الحاسوب، مج٢، ع٢	٢٠٠٨
٢٠	معالجة الاضطرابات التواصلية في إطار الفهم الآلي للغة العربية في الخطاب الشفوي	يونس باجو وآخرون	المجلة العربية لعلوم و هندسة الحاسوب، مج٣، ع٣	٢٠١١
٢١	القراءة الآلية للكلمات العربية المخطوطة باستعمال شبكة عصبية شفافة	سامية السنوسي المدوري	المجلة العربية لعلوم و هندسة الحاسوب، مج٤، ع٣	٢٠١٢
٢٢	تقنية التعرف الآلي على الكلام العربي – طرق جديدة لتعزيز الأداء	ضياء أبوزينة وآخرون	المجلة العربية لعلوم و هندسة الحاسوب، مج٤، ع٣	٢٠١٢
٢٣	توجيه الضوابط اللغوية والصورية للتعرف الآلي على الخط اليدوي العربي: دراسة لساني حاسوبية	سهام موساوي	رسالة دكتوراه – جامعة أبي بكر بلقايد – الجزائر	٢٠١٣

٢٤	حول نظام تمثيل الحرف العربي	مأمون حطاب	الحرف العربي والتقنية أبحاث في حوسبة العربية – مركز الملك عبد الله بن عبد العزيز الدولي – الرياض	٢٠١٥
٢٥	تقنيات التعرف الآلي على الكلام المنطوق وتطبيقاتها في القرآن الكريم: واقع وطموح	يحيى محمد الحاج	الحرف العربي والتقنية أبحاث في حوسبة العربية – مركز الملك عبد الله بن عبد العزيز الدولي – الرياض	٢٠١٥

٢-٣-١-٥- خلاصة:

اهتمت الدراسات العربية بمجال التعرف إلى الحرف العربي نطقًا وكتابة على حدّ سواء، وقد أشارت جل الدراسات إلى الصعوبات التي توجه الباحثين في تطوير برامج قادرة على تمييز الحرف العربي، وإلى ضرورة بذل مزيد من الاهتمام في هذا المجال لما له من أثر عملي على نحو واضح.

٢-٣-٥-٢- الترجمة الآلية:

تعد الترجمة – عمومًا – مطلبًا حضاريًا مهمًا على مرّ العصور، وقد باتت الحاجة إليها في العصر الحالي أشد من أي وقت مضى، وقد تولدت هذه الحاجة من سياق العصر الحضاري المرتبط بالتقنية والحوسبة والإمكانيات الثرية التي تتيحها الشابكة، ولذلك لم تعد الترجمة البشرية مجدية في هذا السياق، وأصبحت الترجمة الآلية مطلبًا في غاية الأهمية.

وتكاد تكون الترجمة غاية أساسية في دراسات حوسبة اللغة العربية في مستوياتها ومجالاتها المذكورة سابقًا، ولذلك فقد لقيت عناية واسعة من الباحثين العرب.

٢-٣-٥-٢-١- التّرجمة الآليّة للُّغة العربيّة/ قضايا وحلول، مأمون الحطّاب، الاثنين ١٢ جمادى الآخرة ١٤٢٩هـ- ١٦ حزيران ٢٠٠٨م.

تسعى هذه الدراسة إلى الكشف عن جوانب مهمة في عملية الترجمة الآلية، وتتناول هذه القضية من أبعاد عدة:

١- خطاب لأهل اللُّغة العربيّة.

٢- لماذا نحتاج إلى ترجمة آليّة؟

٣- أين انتهى العاملون عليها؟

٤- كيف نترجم آليًّا؟

٥- مستقبل التّرجمة الآليّة العربيّة، نظرة تاريخيّة.

٦- مستقبل حوسبة اللُّغة العربيّة خارج الإطار العربيّ: أبعاد الواقع.

يكشف الباحث -بداية - عن ضرورة اهتمام الباحثين في اللغة العربية بقضايا حوسبتها من خلال إعادة الاتّصال بالرّياضيّات والمنطق وسائر العلوم الأخرى المتجلّية في علم الحاسوب واللّسانيّات وتوظيفها لخدمة العربيّة؛ كي تظهر في ثوب جديد يشهد بصلاحيّتها لكلّ زمان ومكان. والتّرجمة الآليّة واحدة من الغايات القصوى لحوسبة اللّغة، إذ إنها تأتي ثمرةً لتحقّق ما يسمّى بالفهم الآليّ للّغة. ولا تستطيع الآلة أن تحوّل نصًّا من لغة إلى أخرى من دون تحليل هذا النّصّ إلى عناصر تكوينه، ثمّ بناء النّصّ المقابل في اللُّغة الأخرى.

وتأتي حاجتنا إلى الترجمة الآلية كي ندرك ما فات وما يجري الآن من حولنا، فضلاً عن أننا نحتاج إلى سرعة الإنجاز وتوفير الوقت والجهد، إضافة إلى الحاجة لدقّة التّرجمة والاطّراد في ترجمة المصطلحات والاستفادة من مصادر المعرفة المختلفة.

وتشير الدراسة إلى نظم التّرجمة الآليّة وتصنيفها وفقًا لِلطُّرق الأساسيّة في بناء نظمها إلى ما يأتي:

- **التّرجمة الآليّة المباشرة: نظم الجيل الأوّل**

وتقوم على تنفيذ التّرجمة كلمة بكلمة من خلال المقارنة المعجميّة المباشرة في قاموس ثنائيّ اللّغة. وهذه الأنظمة تفتقر إلى التّحليل العميق لمكوّنات الجمل، وهي عادة ما تعمل بين لغتين وباتّجاه واحد.

- **التّرجمة الآليّة الوسيطيّة: نظم الجيل الثاني**

وتعد أولى طرق التّرجمة الآليّة غير المباشرة، وكان هدفها إيجاد لغة وسيطة واحدة لكلّ لغات العالم، بحيث تمثل المعاني بين أكثر من لغة في نماذج وبنى وسيطيّة مشتركة مما يسمح بتصميم النّظم متعدّدة اللُّغات.

وتعمل هذه الطّريقة من خلال تحليل نصّ اللّغة المصدر ونقله إلى نماذج وبنى ممثّلة في اللُّغة الوسيطة، ثمّ توليد نصّ إحدى اللُّغات الهدف انطلاقًا من هذه البنى والنّماذج الوسيطة، ومن صفات التّمثيل الوسيطيّ أنّه يشمل كلّ المعلومات لتوليد النّصّ الهدف من دون الرُّجوع للنّصّ المصدر، ويعدُّ تمثيلاً مجرّدًا للنّصّ المصدر والنّصّ الهدف في الوقت نفسه، ويعدُّ تمثيلاً حياديًّا بين اللُّغات المختلفة أيضًا.

وقد واجه هذا الاتّجاه صعوبات عدة في تعريف اللّغة الوسيطة الواحدة، وواجهتها صعوبات في تعريف اللّغة الوسيطة للغات من عائلة واحدة أيضًا؛ وذلك لوجود الاختلافات الصّرفيّة والنّحويّة والدّلاليّة بين اللّغات من العائلات اللُّغويّة المختلفة.

- **التّرجمة الآليّة التّحويليّة: نظم الجيل الثّالث**

هذه النّظم تستخدم لغتين وسيطيّتين لترجمة نصّ مصدر إلى نصّ هدف؛ لغة وسيطة تتعلّق باللّغة المصدر فقط لتمثيل أيّ نصّ مكتوب بها، ولغة وسيطة تتعلّق باللّغة الهدف فقط لتمثيل أيّ نصّ مكتوب بها، وتعمل من خلال عمليّات متتالية هي:

- التّحليل: تحويل النّصّ المصدر إلى بنية وسيطة المصدر.

- التّحويل: نقل البنية الوسيطة المصدر إلى مكافئتها الهدف.
- التّوليد: استخراج النّصّ الهدف انطلاقًا من البنية الوسيطة الهدف.

ويشير الباحث – بعد ذلك – إلى عدد من المشروعات العربيّة الحديثة والأنظمة العاملة والمتوفرة حاليًّا في الأسواق مثل:

١- نظام "المترجم العربي" الذي طوّرته شركة ATA في لندن، ولدى الشركة المذكورة فرع في مسقط بعُمان.

٢- نظام "عربترانز"، وقد طوّرته شركة عربيّة أيضًا في لندن، وكان متوافرًا في الأسواق العربيّة.

٣- نظام "النّاقل العربيّ" الّذي طوّرته شركة سيموس العربيّة في باريس، وهذا النّظام أكثر الأنظمة طموحًا، حيث تملك الشركة المذكورة أربعة برامج للتّرجمة بين الإنجليزيّة والعربيّة وبين الفرنسيّة والعربيّة – برنامج لكلّ اتّجاه.

ثم يكشف الباحث عن التحدّيات التي تواجه الترجمة الآلية، والتي تتفق – إجمالاً – مع تحديات مستقبل حوسبة اللّغة أيضًا، ويرى أنّ أهل اللّغة وحدهم هم القادرون على حلّها، ومن هذه التّحدّيات حلّ إشكاليّات مصادر اللّبس في اللّغة مثل:

- التّطوّر.
- الفيض المعلوماتيّ والتّفاعل بين اللُّغات.
- التّعامل مع العربيّة غير المشكولة.
- قابليّة المفردات اللُّغويّة للعمليّات الإلصاقيّة.
- غياب علامات التّرقيم.
- ظاهرة التّحوّل الدّلاليّ.
- اللّبس النّاتج عن التّعقيد التّركيبيّ لبعض الجمل.

أخطاء في استعمال اللّغة.

- طبيعة الجملة العربيّة في ترتيب مكوّناتها مما يجعل من الصّعب التّيقّن من المعنى المقصود إلا في نهايتها.

وينهي الباحث دراسته بعدد من التوصيات جاءت على النحو الآتي:

١- حشد الجهود لإقناع الجهات القادرة كي تتبنّى هذه القضيّة.

٢- إعداد مشاريع قابلة للتّحقق ومقنعة علميًّا لتقديمها لتلك الجهات.

٣- توجيه أساتذة الجامعات أبحاثهم وطلابهم نحو هذا الموضوع.

٤- إدخال مناهج البحث اللّغويّ العلميّ الحديثة إلى الجامعات العربيّة.

٥- إيجاد وعي لدى عامّة النّاس بأهمّيّة هذه القضيّة في النّهوض الحضاريّ.

ومن الواضح أن هذه الدراسة تأتي في إطار الدراسات التأسيسية التي تحاول تأطير واقع الترجمة الآلية في اللغة العربية، فتكشف عن الإشكاليات الماثلة، والإمكانيات المتاحة، وأوجه الحلول المقترحة.

٢-٣-٥-٢-٢- المعالجة الآلية للغات الطبيعية والترجمة الآلية: تطبيقات نظام"نوج NooJ"، عزالدين غازي، جامعة محمد بن عبد الله، فاس- المغرب، مؤتمر "الترجمة الآلية والحاسوب"، أيام ١٥ و ١٦ و١٧ ماي ٢٠١٤.

تحاول هذه الدراسة الإجابة عن المقترحات الملامسة للمشاكل اللغوية المطروحة عمليًا، وذات أثر بالغ على الترجمة الآلية التي يمكن اختزالها في ثلاث منها:

١- كيف يجب على الآلة أن تبني خصائص جديدة لتمثيل المعطى الدلالي / المعرفي لنص ما في لغة ما؟

٢- وما النماذج الإدراكية-المعرفية المناسبة، لتقييس السيرورة البيو- عصبية، بمعنى أن مسألة

تمثيل "المعرفة اللغوية بالنسبة للآلة" محدد بشكل مجرد، في ظل مسافة تمثيل المعرفة القائمة بين العنصرين البشري والآلي؟

٣- ثم ما الإجراءات التي يجب اتباعها لتحويل نص مدخل، حسب هذا التمثيل، إلى نص مخرج يمكن استعماله من الإنسان/ المستخدم؟

تبدأ الدراسة بالإشارة إلى أنظمة المعالجة الآلية للغات الطبيعية ومدى جدواها في الاستفادة من المفاهيم والأدوات الرياضية العقلانية والهندسة المعلوماتية كمفهوم الخوارزم في معالجة ميكانزمات اللغة من حيث الوصف والتحليل والتوليد،وذلك اعتمادًا على نظريات قامت بدورها بتطوير أنحاء صورية في أشكال خوارزمية تشبه إلى حدّ بعيد ما يسمى بلغات البرمجة الاصطناعية، وأصبحت أسلوبًا إجرائيًا لبناء برمجيات وتطبيقات حاسوبية ومنها الترجمة الآلية المماثلة للترجمة البشرية/ الطبيعية،وقد برزت في سوق الترجمة نماذج، أكثر تعقيدًا وتطورًا لتقييس الكفاية اللغوية الطبيعية-منها ما هو مشهور بالترجمة التلقائية المتاحة على منصة غوغل، عبر فتح المجال لإدخال قواعد لغوية-منطقية إلى الآلة/ الحاسوب، تحتوي على السمات الأساسية الدلالية والنحوية والصرفية للمتواليات اللسانية.

ومع تطور هذه النماذج التي تسمح للجهد البشري بالتدخل في الترجمة الآلية بدرجة ما، فإن الصعوبات لا زالت مطروحة خاصة بالنسبة للأنظمة اللغوية الإعرابية كما هو شأن اللغة العربية حيث نصادف مجموعة من المشاكل لاسيما الدلالية والتداولية والتأويلية منها، وتكاد تجمع معظم المقاربات على وضع نظام آلي لتحليل المادة اللغوية تحليلاً معجميًا وتركيبيًا ودلاليًا، في اللغة الأصل كما في اللغة الهدف. مما جعل مصممي البرمجيات التطبيقية يفكرون في معمارية جديدة تمكن الآلة من التوفر على عناصر لاستقبال المادة اللغوية وتحليلها، وأخرى للتعامل مع القواعد المخزونة في ذاكرتها ثم تحويل هذه المادة إلى لغة ثانية، ولقيامها بذلك عليها أن تُجهز بأدوات هندسية لتمثيل القواعد الطبيعية في أشكال صورية قد تكون عبارة عن نحويات إلكترونية واصفة لكل البنيات اللغوية الموجودة في النص المراد ترجمته.

وتورد هذه الدراسة نموذجًا حيًّا للبيئة اللغوية والتطويرية نوج **NooJ** بصفتها نظامًا

آليًا لمعالجة اللغات الطبيعية ومعجمًا إلكترونيًا، يتضمن أدوات مصممة منهجيًا ووظائف توفر للمترجم فرصًا لعلاج اللغة، ناهيك عن إجراءات البحث والاختبار والتدريب مساعدة للمستخدم.

وتقدم الدراسة أمثلة تطبيقية لهذه البيئة مركزة على جانب من البنيات المركبة والمتلازمة ومدى ترجمتها الآلية، والهدف تطوير نشاط الترجمة باستخدام هذه البيئة أولاً، وبتطوير وظائفها ثانيًا، وذلك بشكل تفاعلي. سعيًا وراء تقييم أمثل وأفضل للكفاية اللغوية لدى الناطقين باللغة العربية سواء في المراحل المتقدمة و/ أو كلغة أجنبية.

وقد انتهت الدراسة إلى أن تصميم نظام ترجمة متعدد اللغات يستدعي مترجمين ولغويين ومهندسين. وعلى الترجمة بالآلة أن تُدخل في اعتباراتها الهندسية المعرفة الإنسانية والأنساق الثقافية حتى تتمكن من فهم سليم وطبيعي للغة؛ وذلك بألا تقتصر على نوع معين من النصوص. ويجب أن تأخذ بالاعتبار البنيات المجازية المعقدة وكل مستويات التأويل. وتمثل الجملة الوحدة الأساسية في عملية الترجمة كما يفعل نظام نوج، مستندًا إلى قواعد، ويعمل مستعينًا بقاعدة بيانات معرفية، فاستخدام الآلة في التطبيقات ضاعف بشكل كبير لا مثيل له من مقدرات طرفي عملية الترجمة، وساهم في تطوير البرامج وتسريع خطواتها كما هو الشأن مع الترجمة التلقائية والتفاعلية المفتوحة المصدر، إذ صار بالتفاعل إمكان التحكم في الأخطاء وحساب النتائج الدقيقة بسرعة توفيرًاللوقت، مما سيكون له الأثر البالغ في إدماج السيرورات المعرفية المعقدة في العملية وتمثيلها في أساليب ميسرة وهادفة وسريعة شكلاً ومضمونًا.

ويبقى –من الضروري– الآن تطوير برامج عربية أصيلة مستفيدة من النتائج المذهلة التي وفرتها المعالجة الآلية للغات الطبيعية، وهي الاستراتيجية الحقيقية المرجوة من وراء البرامج اللغوية، إن بيئة نوج كما يبدو، تبدو واعدة لترجمة اللغات آليًا، ميزتها الرئيسة بساطتها لأنها تسمح لكل مستخدم ببناء الموارد اللغوية باستخدام واجهات بسيطة فتصبح عبارة عن مشاريع لتطوير نشاط الترجمة، وليس من الضروري استخدام سلسلة معقدة من المعالجة بعد تطوير العديد من الميزات للتطبيقات، وهذا يضمن التنسيق بين المعالجة اللغوية– أولاً – والبحث

عن المعلومات وتوليدها بصفة تلقائية في السياق. الذي يُمكّن المستعملين من بناء الطلبات الصحيحة، وبذلك يصبح ذا قدرة على التمييز بين أنواع مختلفة من التحويلات والاشتقاقات والمتغيرات الدلالية والثقافية؛ إذ تُنشأ تلقائيًا بأخذ العينات عشوائيًا لضمان جودة البيانات المستخرجة، حيث يمكن التحقق منها وتصحيحها إذا لزم الأمر في كل حالة على حدة، وهناكمميزات أخرى تُقدم في (نوج)، تساعد في تصنيف المبنى الذي يتم التعامل معه في برنامج التوافق الآلي، وبناء على مرحلة التقييم يمكن اقتراح الأنشطة الجديدة التي بدورها تحدد التطورات اللاحقة للبيئة.

حاولت الدراسة- من خلال التطبيقات السالفة - مقاربة بعض المتواليات المتلازمة في اللغات العربية والفرنسية والانجليزية، لتبين مدى أهمية إجرائها في التحليل والمعالجة والترجمة، واكتفي بتوضيح قوة الإنجاز والجودة واتساع الاستغلال المبنين لبيئة نوج التطويرية لما لها من أهمية في:

١- التنمية اللغوية بإتاحتها بناء التعريفات الشكلية واختبارها وإبقائها بغطاء أوسع للغة، في شكل معاجم حاسوبية أو نحويات إلكترونية متممة لبعضها؛ نظراً لطبيعة عناصر اللغة المشكلة للمادة المعجمية.

٢- تطوير التطبيقات لمعالجة اللغة صرفيًا وتركيبيًا وتحليل المدونات بالنسبة للمختصين في الترجمة الآلية.

٣- وضع النماذج المفاهيمية والمنطقية والمادية للتعبيرات والروابط التركيبية والصرفية والمعجمية، ووضع حلول معلوماتية وبرمجية حاسوبية مع وضع الآليات والأدوات الكفيلة بالنفاذ إلى المعاجم، والقيام بعمليات البحث والمعالجة بالتفاعل البيداغوجي المباشر.

٤- وتتضمن هذه البيئة المجانية، أدوات للمعالجة الآلية للغة مثل نظام إدارة البيانات ولغة البرمجة مفتوحة المصدر، بما تمكن من معالجة المدونات المعجمية بطرق البحث والاختبار، وهذا ما يتطلب تقديم منطق صوري مكثف للغة.

٥- من جانب آخر، تعتبر البيئة محللاً آليًا لمنظومة المعاجم، وللتخلص من بعض الصعوبات أثناء العمليات لاسيما التحليل الصرفي، رُفع مستوى أداء هذه البيئة لتوحيد المداخل المعجمية (lemmas-super) من جهة، ولتدبير منظومة المعاجم جنبًا إلى جنب مع جداول المعجم- التركيبي من جهة أخرى.

٦- استطاعت التقنية الجديدة، فيما يتعلق بمحولات الأوضاع النهائية، أن تشفر المورفيمات وجميع اللواصق في وحدة معجمية ذات شكل موحد، وتفهرس كل لغة حرفها الأول اللاتيني مثل(A).

٢-٣-٥-٢-٣- الهندسة اللغوية والترجمة الآلية المفهوم والوظيفة، عمر مهديوي، بحث مقدم المؤتمر السنوي للمنظمة العربية للترجمة حول الترجمة والحاسوب، نحو تطوير بنية تحتية للترجمة بفاس في، ١٥ -١٧ ماي ٢٠١٤.

يندرج هذا البحث في إطار العلاقة البينية القائمة أو الممكن قيامها بين الترجمة الآلية (Translation Machine) والهندسة اللغوية (LinguisticEngineering). فيفرض ذلك، بطبيعة الحال، تحديدهما مفاهيميًا أولاً، والكشف عن الوظيفة أوالوظائف التي يقوم بها كل مجال ثانيًا، ثم رصد الروابط التي تربط بين المجالين ومجالات معرفية أخرى ثالثًا.

وبعد أن يعرض الباحث لتاريخ حركة الترجمة الآلية وتطورها، وأبرز نظمها، يشير إلى الترجمة الآلية في اللغة العربية وما تواجهه من مشكلات؛ فالترجمة الآلية من اللغة العربية إلى لغة أخرى أو العكس، تواجه مشاكل عدة، يمكن تصنيفها إلى صنفين:

- مشكلات تقنية؛ إذ يمكن القول إن علوم الحاسوب تشهد تطورًا مستمرًا، وتقدمًا سريعًا، إلى درجة أنه غطى جميع المشكلات والمسائل المتعلقة بالتطبيقات الحاسوبية، وتأتي الترجمة الآلية والترجمة بمساعدة الحاسب في مقدمة المجالات التي استفادت كثيرًا من هذه التقنية.
- ومشكلات لغوية، وهي كثيرة، وتتباين من اللغة المصدر إلى اللغة الهدف، لهذا الاعتبار،

لابد من مواجهتها، وإيجاد الحلول المناسبة لها، وإلا فما الفائدة من البحث العلمي المتخصص في الترجمة الآلية؟

وهكذا، يمكن حصر أوجه التقصير بشأن الترجمة الآلية في اللغة العربية في الأمور التالية:

- طريقة التجزيء والتشتت التي تهيمن على الجهود العربية المتخصصة في التطبيقات الآلية العربية عامة، وفي ميدان الترجمة الآلية على الخصوص، مما جعلها تنحرف عن الهدف المنشود، ذلك أن معظم الحلول المتوفرة حالياً، تعتبر حلولاً جزئية، كالمعاجم الإلكترونية، وبرامج التشكيل الآلي والصرف، وغيرها. على أن إنجاز ترجمة آلية من اللغة العربية وإليها، وبجودة عالية تتطلب توحيد تلك الحلول كلها، بالإضافة إلى عمليات التحليل الدلالي والنحوي والصرفي ونقل المعاني، ربحًا للوقت، وادخارًا للمال العربي الذي ينفق، في بعض الأحيان، في مشاريع نفعية وتجارية، قد لا تعود على لغة الضاد بطائل.
- قصور في الدعم المادي الحكومي أو الخاص المخصص لمثل هذه المشاريع التطبيقية من البحث والتطوير، وإن وُجد فهو دعم قليل لا يساعد على تغطية تكاليف المشروع من بدايته حتى نهايته، أو أنه قصير الأمد بسبب البحث عن الربح السريع، وتلك ثغرة من ثغرات اللغويات الهندسية العربية.
- غياب رؤية علمية لمفهوم الترجمة عامة، والترجمة الآلية خاصة، من شأنها أن تؤسس فعلاً لـ "بيت حكمة" عربي، قادر على إثراء الخزانة العربية بالمؤلفات والكتب في مختلف العلوم والفنون والآداب. ومن هنا، لابد أن من التنويه إلى العمل القيم الذي تقوم بـه المنظمة العربية للترجمة، إلى جانب مؤسسات أخرى، في هذا الصدد، إذ استطاعت في ظرف وجيز أن تفتح أفق القارئ العربي على عوالم ممكنة من المعارف والأفكار من خلال ما يترجمه باحثون عرب في مجالات معرفية مختلفة.

ويحدد الباحث أسس العمل في إطار الهندسة اللغوية، التي ينبغي أن تأخذ بعين الاعتبار ثلاثة أساسيات:

- الأس الأول: يتمثل في ضرورة الاعتماد على إطار لساني صوري، قادر على توصيف الأنظمة اللسانية للغة المدروسة.
- الأس الثاني: ضرورة الانطلاق من قاعدة المعطيات أو المعارف أثناء المعالجة الآلية، لأنها تعتبر من أساسيات اللغويات الرقمية، التي تختلف جذريًا عن اللغويات النظرية التقليدية.
- الأس الثالث: ضرورة الجمع بين الخبرات اللغوية والخبرات الهندسية في معالجة أنظمة اللغات الطبيعية.

ويخلص في هذه الدراسة عدد من النتائج والتوصيات، وهي:

النتائج:

- إن تطوير بنية تحتية للترجمة في اللغة العربية، يتطلب توفير الموارد البشرية والمعرفية والمادية الضرورية، التي تساعد على تحقيق صناعة للترجمة، باعتبارها فرعًا رئيسًا من فروع صناعات اللغة في مجتمع المعرفة.
- تعد الترجمة عامة، والترجمة الآلية خاصة، وسيلة من وسائل توطين المعرفة العلمية والتقنية في الوطن العربي.
- تعد الترجمة الآلية أداة فعالة وناجعة لدعم المحتوى الإلكتروني العربي على الشابكة وتعزيزه.
- إن تجاوز مشكلات الترجمة الآلية في اللغة العربية يتطلب تطوير هندسة لغوية قادرة على فهم خصائص النقل واستيعابها من اللغة المصدر إلى اللغة الهدف.

التوصيات

- إن النهوض باللغة العربية في ظل مجتمع المعرفة والمعلومات لم يعد من قبيل الترف الفكري،

بل أضحى استراتيجية علمية وحضارية بامتياز، لا محيد عنها لترقية لغة الضاد إلى مصاف اللغات المتقدمة أولاً، ولتنمية الإنسان العربي على مستويات اقتصادية واجتماعية وثقافية ثانيًا.

- إن وضع الترجمة في الوطن العربي غير مشرف، مما يستدعي منا أفرادًا ومؤسسات، بذل جهود مضاعفة، وتوحيد المشاريع من أجل سدّ هذه الثغرة بين العربية واللغات المتفوقة في هذا المجال.
- الدعوة إلى إنشاء مرصد عربي للترجمة، من شأنه أن يجمع الخبرات العربية ويوحد بينها.
- وأخيرًا، ضرورة الانفتاح على مؤسسات الترجمة (الخاصة والعامة) في العالم المتقدم بغرض الاستفادة من تجاربها، ومن ثم الوصول إلى بلورة سياسة لغوية جديدة للترجمة العربية.

٢-٣-٥-٢-٤- مسرد الدراسات:

الرقم	عنوان الدراسة	المؤلف	مكان النشر	تاريخ النشر
١	الحاسب الآلي والترجمة	عبد الفتاح أبو السيدة	مجلة اللسان العربي، العدد ٢٨	١٩٨٧
٢	الترجمة الآلية واللغة العربية	محمود إسماعيل الصيني	وقائع مختارة من ندوة استخدام اللغة العربية في الحاسب الآلي- دار الرازي بيروت	١٩٨٩
٣	بعض الصعوبات في الترجمة الآلية من الإنجليزية إلى العربية ومن العربية إلى الإنجليزية	داود عبده	المؤتمر الثاني حول اللغويات الحسابية العربية - معهد الكويت للأبحاث العلمية - الكويت	١٩٨٩

٤	الترجمة الآلية للغة العربية	محمود إسماعيل صيني	مجلة الفيصل(الرياض)، العدد، ٢٣٩	١٩٩٦
٥	التفاعل بين الإنسان والآلة في الترجمة الحاسوبية	سلمان داود الواسطي	مجلة التعريب، المركز الثقافي العربي للترجمة والتعريب، دمشق، ع٢٠	ديسمبر ٢٠٠٠
٦	مقدمة في الترجمة الآلية	عبد الله بن حمدان الحميدان	مكتبة العبيكان – الرياض	١٤٢١هـ – ٢٠٠٠م
٧	الترجمة الإلكترونية: آفاق الحاضر والمستقبل	روحي بعلبكي	ضمن كتاب العربي (مستقبل الثورة الرقمية) – الكويت	٢٠٠٤
٨	أهمية التعريب في حوسبة اللغة العربية	عمر مهديوي	مجلة التعريب، المركز العربي للتعريب والترجمة والتأليف والنشر العدد ٢٦	٢٠٠٧
٩	الترجمة ا لآلية: كبنية أساسية في صرح التعريب	سلوى السيد حمادة	المجلة العربية لعلوم وهندسة الحاسوب، ع١	٢٠٠٧
١٠	الترجمة الآ لية: تمثيل المعلومات لفك اللبس	سلوى السيد حماده	المجلة العربية لعلوم وهندسة الحاسوب، ع١	٢٠٠٧
١١	الترجمة الآلية ذات الصلة بالعربية من منظور تاريخي	محمد راجي زغلول وعواطف أبو الشعر، ترجمة: محمد راجي زغلول	المجلة العربية لعلوم و هندسة الحاسوب، مج٢، ع١	٢٠٠٨
١٢	في اللغة والحاسوب: الترجمة وتدقيق الإملاء بين الإنسان والآلة	داود عبده	عمان: دار جرير للنشر والتوزيع	١٤٣١هـ ٢٠١٠

١٣	حوسبة اللغة والترجمة الآلية: مسار عقدين لشركة آي تي آي	هيئة التحرير	مجلة العربية والترجمة – لبنان مج١٥، ع١٣	٢٠١٣
١٤	الترجمة الآلية من العربية وإليها		الحرف العربي والتقنية أبحاث في حوسبة العربية – مركز الملك عبد الله بن عبد العزيز الدولي – الرياض	٢٠١٥
١٥	اللسانيات الحاسوبية والترجمة الآلية	سناء منعم	إربد، الأردن : عالم الكتب نشر–توزيع– طباعة	٢٠١٥
١٦	تطويع التقنية الحديثة لخدمة الترجمة الإسلامية، نظام حرف للترجمة نموذجً	إبراهيم صالح النمي وآخران	المجلة الدولية للتطبيقات الإسلامية في علم الحاسب والتقنية، مج٤، ع٣	٢٠١٦
١٧	التقنية في خدمة الترجمة والمترجمين	محمود إسماعيل صالح	اللسانيات الحاسوبية واللغة العربية إشكالات وحلول – كنوز المعرفة – الأردن	٢٠١٨

٢–٣–٥–٢–٥– خلاصة:

تكشف الدراسات المعنية بالترجمة الآلية عن إشكاليات كثيرة لا تفترق في مجملها عما أشارت إليه الدراسات السابقة، وهي تشير من جهة أخرى إلى حاجتها إلى الاهتمام بحوسبة اللغة في مستوياتها ومجالاتها السابقة لتشكل أرضية مناسبة لتحقيق ترجمة آلية مرضيًا عنها تماشيًا مع السياق الحضاري المعاصر.

٢–٤– خاتمة

إن استعراض الدراسات السابقة يظهر أن اهتمامات الباحثين العرب من لسانيين وحاسوبيين في مجال حوسبة اللغة العربية قد بدأ بداية جادة منذ ثمانينيات القرن الماضي، وهي

المرحلة التي شهدت دخول الحاسوب في مجالات الاختصاص العلمي وانتشاره ثقافيًا على المستوى العالمي، أما المرحلة التي سبقت ذلك فقد اعتمدت على الإحصاء للمفردات اللغوية وجذورها، وهي أبسط عمليات الحوسبة. ويمكن من متابعة جلّ هذه الدراسات ملاحظة عدد من النتائج التي تظهر واقع الدراسات العربية في حوسبة اللغة، ولعل أهمها:

١- أشار جلّ الدارسين إلى مفهوم اللسانيات الحاسوبية، وميز بعضهم علم اللغة الحاسوبي عن علم الحاسوب اللغوي، ولعل هذا التمييز لم يأتِ على أساس اختلاف الباحثين وطريقة المعالجة فحسب، وإنما بالارتباط بمجال العمل أيضًا، فاهتمام اللسانيين انصب على التصورات المنطقية التي تتجاوز الوصف إلى التوصيف باعتبار ذلك ممهدًا لتهيئة أنظمة اللغة العربية لتتواءم مع الحوسبة، واهتمام الحاسوبيين انصب على المعالجة الآلية القائمة على الخوارزميات.

٢- إن عددًا من الدراسات-بما فيها المتأخرة - قد ذيلت عنوانها بإشكاليات وحلول؛ مما يشعر بأن البحوث في حوسبة اللغة العربية ما زالت لم تستقر على ثوابت واضحة على صعيد المصطلحات والمفاهيم طرق المعالجة، ولعل ذلك راجع إلى حالة التأخر المتتابعة والملازمة لانتقال النظريات والمفاهيم والمناهج الغربية، وهو شأن عام، يجعلها غير مستوعبة، ويدفع الدارسين - في الآن ذاته- إلى مواكبة الجديد المتطور.

٣- إنّ من أهم الإشكاليات التي تواجه دراسات حوسبة اللغة عدم وجود نظرية لسانية شاملة في الدراسات اللغوية قادرة على ضم أطراف الظاهرة اللغوية وعناصرها بقدر من الشمول يسمح باختزال أنساق اللغة إلى عدد محدد من الأنساق الرياضية الحسابية، ومن ثمّ بات الاعتماد على الحاسوبيين في تقديم الجانب التقني لحل إشكاليات الجانب المعرفي في النظرية اللسانية، والملحوظ أن جهود الحاسوبيين تفوق ما قدمه اللسانيون في هذا المجال، وهو ما يجعل سير العمل معكوسًا؛ إذ تسبق الأداة الظاهرة والمادة، وتصبح المسألة وكأن دراسات حوسبة اللغة تسعى إلى تطويع اللغة لتناسب الأداة بخلاف الأصل في هذا الشأن.

٤- إن أغلب اللسانيين لم يستطع أن يضع الأصول الضرورية لحوسبة اللغة؛ إذ بقيت

تصوراتهم بمنأى عن فهم العمليات الرياضية التي يتعامل معها الحاسوب، وقد جاءت مجمل تصوراتهم ذات طابع عام أو تقليدي غير ممنهج بدقة علمية كافية، ولعل من أبرز المفاهيم التي ارتكزت عليها بعض الدراسات التركيز على الفرق بين الوصف والتوصيف باعتبار الأخير يقدم بديلاً للحدس عند الإنسان وهي الفكرة التي طرحها نبيل علي وتبناها نهاد الموسى، وعلى الرغم من أهمية هذه الفكرة لتأسيس حوسبة اللغة؛ على الصعيد الاصطلاحي التأسيسي على الأقل، إلا أنها بقية في الإطار العام دون أن توضح أساسيات أصول النظام اللغوي للغة العربية بشكل دقيق، ولذلك جاءت كثير من الدراسات حالة أكثر منها واقعية عملية.

٥- في المقابل فإن جهود الحاسوبيين- في الأغلب - انحسرت في معالجة مستويات دون أخرى مثل المستوى الصرفي، أو مباحث محددة هي التي يمكن إخضاعها للقياس مثل: الجملة البسيطة، والمركب العددي، وهي أسهل مجالات نظام اللغة العربية في المستوى التركيبي النحوي.

٦- إذا كانت بعض النظريات قد أتاحت المجال إلى الضبط المعياري الدقيق للنظام اللغوي مهما اختلفت تمثلات النسق فيه؛ مثل: الميزان الصرفي، والصيغ الصرفية، ونظرية العامل، وتحديد المعاني النحوية وما يرتبط بها من علامات صوتية إعرابية، فإن ما بقي مستعصيًا على الضبط والتقنين السياقات اللغوية التي تعين الرتبة للمفردات اللغوية، وتحدد الارتباطات الدلالية التداولية لها، ثم تتحدد الدلالة وفقًا لذلك كله، وهو ما حاولت دراسات الحوسبة تأصيله وبحثه في دراسات (الأنطولوجيا).

٧- اتسم منهج الحاسوبين باعتماد المنطق الرياضي المحض، وقياسه تجريبيًا، وهوما ميز جهودهم بالعلمية والعملية في آن، في حين بقيت جهود اللسانيين أسيرة التصورات القواعدية للغة الطبيعية المرتكزة على تصنيفات النحو التقليدي وتقعيده، والغريب أنها عوملت كما لو كانت خوارزمية تمامًا مع أنها مبنية على الحدس الإنساني. ويظهر هذا

الفرق – مثلاً –من أنّ اعتماد الحاسوبيين على الذخيرة اللغوية كان يعني الاعتماد على حالة منظمة من الاستخدامات اللغوية، ومن ثم فدراساتهم تقوم على الواقعي والمتحقق، في المقابل كانت دراسات اللسانيين افتراضية، وتعتمد على الأمثلة المصنوعة، وتوقعات الخطأ والصواب، ومن ثم اتسمت بالعمومية وعدم إدراك طبيعة المجال البيني معرفيًا ومنهجيًا.

٨- ثمة اختلافات واضحة في اهتمامات الحاسوبيين واهتمامات اللسانيين؛ إذ تحتل صنعة المعجم والصرف والترجمة أولوية للحاسوبيين، يلحظ اهتمام اللسانيين بقضايا النحو اهتمامًا ملحوظًا.

٩- ثمة نظريات لغوية ولسانية قديمة أو حديثة في التعامل مع الظاهرة اللغوية تؤطر لبعض المستويات اللغوية، مثل: "نظرية الحقول الدلالية" لم يطّلع عليها الحاسوبيون، في الوقت الذي اهتموا بهذا المستوى في مجال (الأنطولوجيا)، مما يشير إلى وجود فجوة بين جهود الطرفين، وذلك ما يكشف عن ضرورة التفكير –جديًا – بالاهتمام بمجال معرفي بيني متخصص باستحداث برامج أكاديمية متخصصة فيه، لا مجرد الاعتماد على جمع جهود الطرفين والتقريب بين خبراتهما.

١٠- على الرغم من عناية بعض المؤسسات العربية بحوسبة اللغة العربية، مثل: مجمع اللغة العربية الأردني الذي عقد أكثر من ندوة في مواسمه الثقافية للبحث في هذا المجال، والمنظمة العربية للتربية والثقافة والعلوم، ومعهد التعريب والترجمة في المغرب، ومركز الملك عبدالله بن عبدالعزيز الدولي لخدمة اللغة العربية في السعودية، وغيرها، إلا أنها لم تتجاوز حدود المؤتمرات والندوات إلى تبني مشاريع طويلة الأمد في حوسبة اللغة العربية ببناء وحدات وأقسام ثابتة، تهتم بإنجاز برامج تطبيقية في مجالات المعالجة الحاسوبية للغة العربية، وتكون قادرة على التنبؤ بالحاجات المستقبلية لا أن تهتم بما يطرأ لحظيًا.