مجلة علمية دولية محكّمة نصف سنوية
تصدر عن مخبر الدراسات النقدية والأدبية المعاصرة جامعة تيسمسيلت/الجزائر
ISSN 2571-9882 EISSN 2600-6987
https://www.asjp.cerist.dz/en/PresentationRevue/297
Contemporary Studies is a bi-annual open access International double-blind journal. It is published by the University of Tissemsilt, Algeria.

المجلد: 05 / العدد: 02 / (2021)، ص143/160

الخط العربي من الكتابة اليدوية إلى المعالجة الآلية
–دراسة حوسبية في ضوء اللسانيات التطبيقية–

Arabic calligraphy from handwriting to automated processing
-A computational study in the light of applied linguistics-

د. بن جلول مختار
mokhtar.bendjelloul@univ-tiaret.dz
جامعة ابن خلدون تيارت
(الجزائر)

تاريخ الاستلام:2021/07/01 تاريخ القبول:2021/08/30 تاريخ النشر:2021/12/02

ملخص:

تسعى هذه الورقة البحثية إلى دراسة الخط العربي أنطولوجيا من خلال تتبع مراحل نشأته وصولا إلى معالجته الآلية، وكيف واجه التحديات الكبرى التي حالت بينه وبين تصميمه الآلي، وذلك بعد التطرق إلى مفهوم اللسانيات التطبيقية وعلاقتها بعلم اللغة العام من جهة، ومن جهة أخرى علاقة اللسانيات التطبيقية باللسانيات التعليمية واللسانيات الحاسوبية باعتبار أن الخط العربي تتم معالجته يدويا في التعليمية وآليا في الحاسوبية، وذلك لتحقيق أهداف مسطرة مسبقا تمثلت في محاولة الوصول باللغة العربية إلى الاستعمال الآلي الأمثل ليواكب أبناؤها ما وصل إليه الانسان المعاصر من تطور تكنولوجي على مستوى محاورة الآلة التي أضحت تحاكي العقل البشري في وظائفه، وذلك كون استعمال اللغة الأم في مثل هكذا مجالات له الحظ الأوفر في الاستيعاب والفهم، وكذلك توظيفها في البرمجة الآلية باعتبارها اللغة الأمثل كونها تحوي نظاما علائقيا يتسم بالمنطق والإحكام؛ سواء كان ذلك على مستوى علاقة اللفظ بالمعنى أم على مستوى العلاقات الداخلية بين الأصوات والمفردات والتراكيب فيما بينها.

كلمات مفتاحية: الخط ؛ الحوسبة؛ المعالجة الآلية؛ اللسانيات التطبيقية؛ اللسانيات الحاسوبية.

Abstract:

This research seeks to study Arabic calligraphy ontology by studying the stages of its inception and its automatic treatment, and how it faced the major challenges that prevented it from its automated design, This is after addressing the concept of applied linguistics and its relationship with general linguistics on the one hand, and on the other hand the relationship of applied linguistics to educational linguistics and computer linguistics, given that Arabic calligraphy is handled manually in educational and automatically in computer, In order to achieve pre-established goals represented in the attempt to reach the Arabic language to the optimal

mechanical use of its children to keep pace with the technological development that modern man has reached at the level of the machine dialogue that has become mimics the human mind in its functions, and that the use of the mother tongue in such areas has the best chance in Comprehension and understanding, and employing it in automatic programming as it is the ideal language as it governs a relational system characterized by logic and precision. Whether this is at the level of the relationship of the word meaning, or at the level of the internal relationships between sounds, vocabulary and the structures between them.
Keywords: Computing; Automated processing; Applied linguistics; Computational linguistics.

مقدمة:

الكتابة صناعة بشرية برزت إلى الوجود متأخرة جدا عن ظهور اللغة عالج فيها الانسان بأنامله رسم الحرف اللغوي، وأصبغ عليه رؤاه وتصوراته الذهنية، وجسد فيه علاقة الأثر بالفكر؛ علاقة المادة بالروح، علاقة اللفظ بالمعنى. وكباقي الصناعات البشرية لم تَشُذ الكتابة في نشأتها عن التدرج في النمو ككل الموجودات فلقد مرت بمراحل عدة في تشكلها؛ مرحلة البدايات التي اتسمت بالبساطة والعموم من خلال مقاربة كلية للظاهرة، ومرحلة النضج التي اكتست فيها الكتابة أهم ملامحها وصفاتها الوجودية، وثالثة الأثافي مرحلة الإبداع التي أُلتفت فيها إلى البعد الجمالي بعدما اكتمل البناء العضوي لها.

وبنفس الوتيرة تمت معالجة هذا الحرف على المستوى الحوسبي الآلي حيث مر بالمراحل نفسها وذلك في تماثل ثنائي بينه وبين مراحل تطور برامج المعالجة الآلية للغة ومع بدايات ظهور الحواسيب لم تكن برامجها قادرة على استيعاب تفاصيل الحرف بأشكاله المتعرجة والمتصلة المرسومة يدويا الأمر الذي فرض مقاربته بالرسم الصناعي؛ أي بخطوط مستقيمة وانحرافات زاوية وتقوسات دائرية من خلال رص مجموعة من النقاط، الأمر الذي جعل شكله بعيدا كل البعد عن ملامح الخط اليدوي. ولكن بعدما تطورت الحواسيب وأنتج علماء الاعلام الآلي برامج متطورة بدأت الآلة تقارب الكتابة البشرية لدرجة أصبحنا غير قادرين على التمييز بين ما كتبته الأنامل البشرية وما صورته البرامج الآلية.

وعلى ضوء هذه المقدمة فإننا في هذه الورقة البحثية سنسعى إلى الكشف عن مراحل تطور الخط العربي منذ نشأته إلى غاية استوائه، ثم الحديث عن المعالجة الآلية لهذا الخط والتحديات التي تواجهه، كما ترنو الورقة إلى الكشف عن طرق المقاربة الآلية للخط العربي، ولا يمكن الكشف عن هذه القضايا إلا بعد تحديد معالم البيئة التي تنتعش فيها وهي اللسانيات الحاسوبية كفرع من فروع اللسانيات التطبيقية، فكان لزاما الحديث عن العلاقة بينهما لتتضح معالم المعالجة الآلية للكتابة العربية.

وككل الأعمال والبحوث الأكاديمية في سعيها إلى تحقيق أهداف وغايات متوقعة، فإن بحثنا هذا يهدف إلى الوصول باللغة العربية إلى الاستعمال الآلي الأمثل ليواكب أبناؤها ما وصل إليه الانسان المعاصر من تطور تكنولوجي على مستوى محاورة الآلة التي أضحت تحاكي العقل البشري في وظائفه لدرجة أنها؛ أي الآلة، أصبح بمقدورها كتابة قصيدة شعرية من تصوراتها الشخصية عن طريق منافذ مسحها للصورة والصوت، فضلا عن كتابتها على منوال شاعر بعينه بعدما تزود بأعماله، وذلك لغايتين اثنتين:

الأولى: كون استعمال اللغة الأم في مثل هكذا مجالات له الحظ الأوفر في الاستيعاب والفهم.
الثانية: توظيف اللغة العربية في البرمجة الآلية باعتبارها اللغة الأمثل كونها تحوي نظاما علائقيا داخليا وخارجيا يتسم بالمنطق والإحكام؛ سواء كان ذلك على مستوى علاقة اللفظ بالمعنى أم على مستوى العلاقات الداخلية بين الأصوات والمفردات والتراكيب فيما بينها.

اللسانيات التطبيقية / مفهوم اللسانيات الحاسوبية وعلاقتها بعلم اللغة العام

من إشكالات اللسانيات التطبيقية عدم الإجماع على تحديد ماهيتها؛ أ هي علم قائم بذاته مستقل له مجالات يبحث فيها أم أنها من العلوم البينية التي تربط بين علمين آخرين وإذا تتبعنا فروع هذا العلم فإننا سنرجح الرأي الثاني، ففروعه تتناول مواضيع لعلوم أخرى، فاللسانيات الاجتماعية على سبيل المثال لا الحصر تعالج القضايا الاجتماعية وإسقاطاتها على العملية التعلمية فهي " تطمح ... إلى إيجاد حلول للمشكلات اللغوية في المجتمع،"[1] وبذلك وتبعا لمثل هذه التعريفات تكون منحصرة في المجالات التعلمية؛ أي أنها " تبحث في التطبيقات الوظيفية التربوية للغة من أجل تعلمها وتعليمها للناطقين بها، وتبحث أيضا في الوسائل البيداغوجية المنهجية لتقنيات تعليم اللغات البشرية وتعلمها"[2] على حد تعبير الدكتور مازن الوعر. كما يمكن رؤية ذلك تجسيدا من خلال حالتين اثنتين وهما:

" أ. ... تطوير لهجات معينة بهدف جعلها عاملة وسط محيطات جديدة ونفس الشيء بالنسبة للسانيات النفسية واللسانيات العصبية واللسانيات الحاسوبية وغيرها من الفروع.

ب. ... تدريس العدد الكبير من الجماعات البشرية لغات لا يعرفونها لكي يكون بمقدورهم التعامل بواسطتها مع محيطهم الجديد،"[3] حسب ما أشار إليه ج. فيشمان J. Fishma, ، ونفس الشيء بالنسبة للفروع الأخرى كاللسانيات النفسية والعصبية والحاسوبية وغيرها من الفروع. ومن هذا المنطلق يتضح لنا أن اللسانيات التطبيقية تستثمر مجالات معرفية أخرى في العملية التعلمية من خلال حل إشكالات ومعيقات تربوية تعجز الدراسة اللغوية عن حلها كما أشار إلى ذلك عبده الراجحي قائلا "علم ذو أنظمة علمية متعددة يستثمر نتائجها في تحديد المشكلات اللغوية وفي وضع الحلول لها.

وإذا كان علم اللغة لا يمثل العنصر الوحيد في ميداننا، لأنه يستقي من علوم أخرى، فلا شك أنه يمثل أهم عنصر فيه،"[4] وعليه فإن هذه التعاريف تحيلنا إلى وجود علم آخر يقابل هذا العلم؛ علم نظري منغلق على نفسه، قضاياه مضمرة لا يمكن الوقوف على حيثياتها إلا إذا وجد لها فضاء تتمدد فيه، هذا العلم هو علم اللغة النظري؛ أي اللســـانيات العامة، وتكون " اللسانيات التطبيقية من هذا القبيل في علاقة تبعية مع اللسانيات البحتة شأنها شأن تقنيات المهندس والطبيب في علاقاتها مع معطيات العلوم الأساسية التي يقوم عملها عليها. لكن تجربتهما الخاصة عن طريق نوع من رد الفعل، تسيطر على معطيات العلوم الأساسية التي تستند إليها، أو تدحضها، وتساهم في التطوير النظري لهذه العلوم."[5]

ولذلك يرى الكثير من العلماء أنه " من المهم أن نشير إلى أن علم اللغة ينهض على دعامتين؛ نظرية لغوية Linguistic theory ووصف لغوي Linguistic description، تقدم النظرية الإطار المعرفي العام عن اللغة وعن طبيعتها، ويقدم الوصف المعالجة العلمية لظواهر اللغة على مستوى الأصوات والصرف والنحو والدلالة، على تنوع ... في الاتجاهات والمدارس،"[6]

فروع اللسانيات التطبيقية:

تتضح أهم فروع اللسانيات التطبيقية من مسمياتها ويتبين لنا المجال الذي يشتغل عليه علم اللغة التطبيقي، وفي هذه الجزئية سنتناول كل فرع وما يشتغل فيه به :

اللسانيات التعليمية:

الفرع الأول، يُشْتغل فيها على نظريات تعلم اللغة من جهة، وتوظيفها؛ أي اللغة، في الاكتساب المعرفي في تخصصات أخرى الرياضيات والطبيعة وما إلى ذلك من التخصصات، لأن " اللغة مثلما تكتسب تطورها بعامل التوليد الداخلي فإنها تحصل على عناصر جدتها بالاقتراض الخارجي،"[7] هذا من جهة، من جهة أخرى تعمل على إيجاد الروابط اللغوية المشتركة بين التنوع الاستعمالي للتراكيب ووظائف المفردات في اللهجات واللغات الأخرى، فهي تدرس العلاقات بين اللغات، وتبحث في عملية التأثر والتأثير بين اللغات التي تنتعش في فضاء لغوي واحد كالازدواجية اللغوية في مجتمع واحد، ولنا مثال واضح على ذلك في المجتمع الجزائري بين اللغة العربية واللغة الأمازيغية.[8]

اللسانيات الاجتماعية:

الفرع الثاني، فهذا الفرع يتناول مؤثرات التفاعل الاجتماع بين البشر في التغيرات التي تطرأ على البنية اللغوية وكيف يتم توظيف الممارسات اليومية الناتجة عن هذا التفاعل في قوالب وتراكيب لغوية لا تفهم إلا باستحضار هذه الصور الحياتية في ذهن السامع والمتكلم على حد سواء، وبالعودة إلى أعمال دي سوسير في بنيته النظرية للغة نجده يقرر – على حد تعبير جولييت غارمادي – " أن اللغة منظومة علامات أودعها مراس الكلام في الجمهور المتكلم، وأن المنظومة ناتجة عن تبلور اجتماعي، و أن الطبيعة الاجتماعية هي طابع داخلي للمنظومة، وأنه لا توجد حقيقة لسانية خارج الديمومة والجمهور المتكلم، وأن الزمن وحده يأذن للقوى الاجتماعية بممارسة تأثيراتها على اللغة."[9]

علاقة دي سوسير بهذه الثنائية (اللغة – علم الاجتماع) لا تبدو مستساغة لأول الأمر باعتبار أن دي سوسير درس اللغة لذاتها ولأجلها، وسعى لإيجاد مبررات تثبت أن اللغة كفيلة بنفسها ولا يمكنها الاستناد إلى فضاءات أخرى كعلم الاجتماع مثلا، وذلك أنه " لم يكن شاغل سوسير الأول إرساء الأسس الكافية لدرس العلاقات بين اللغات والمجتمعات، وإنما كان هاجسه بالأولى تحديد موضوع وإيجاد طرائق اللسانة ذاتها. إن اللغة منظومة لا تعرف سوى نظامها الخاص بها،"[10] ولكن عدم تمكن دي سوسير من إثبات هذه النظرية والعدول عن ثنائيته (الدال — المدلول) إلى ثلاثية (الدال — المدلول – المرجع) من قبل أقرب تلامذته والمتأثرين به، لدليل على أنه أثبت علاقة اللغة بالمجتمع، وذلك بسبب كون اللغة ظاهرة اجتماعية ولذلك ربط بيتر ترادجل Trudgill Peter بين اللغة والمجتمع في تعريفه لعلم اللغة الاجتماعي sociolinguistique باعتباره " الشعبة المتفرعة عن علم اللسانيات والتي تعنى باللغة باعتبارها ظاهرة اجتماعية وحضارية. فهي تدرس التبادل الجدلي القائم بين اللغة والمجتمع، وتستند إلى علوم أخرى مساعدة، كعلم الاجتماع والأنثروبولوجيا والجغرافيا البشرية وعلم النقس الاجتماعي وغيرها."[11]

اللسانيات النفسية واللسانيات العصبية:

الفرع الثالث والرابع، هذان الفرعان لا يمكن دراستهما بمعزل عن بعضهما البعض لارتباط الدماغ بالممارسات السيكولوجية للإنسان، فكل تغير في مزاج الانسان يترتب عليه تغيير في بنيته المعرفية والعكس صحيح، ولذلك أُنتقد تصور تشومسكي للعلاقة بين علم النفس المعرفي واللغة عندما وصف " مقاربته في علم اللغويات على أنها جزء مما يسميه « الثورة المعرفية » التي حدثت في حوالي منتصف [القرن الماضي][12] بالنسبة لتشومسكي، كانت السمة الأساسية لهذه الثورة هي الاعتقاد الجديد بأن المعرفة كانت قابلة للاستقصاء أو البحث العلمي.

إن المعرفة اللغوية هي نوع واحد فقط من المعرفة، ولكن يمكن دراستها تجريبيا ويمكن صياغة الفرضيات حول بنية المعرفة اللغوية في العقل البشري. ويميز تشومسكي بين معرفة لغة معينة يتم وصفها من خلال قواعد النحو التوليدي Generative Grammar لتلك اللغة، وبين معرفة اللغة بشكل عام والتي تشملها قواعد النحو العالمية Universal Grammar،"[13] حيث أن تشومسكي أخلط بين الأساليب التي تستخدم في مجال علم النفس وأسس المعرفة ذاتها لأن الأساليب التي اعتمد عليها "لغوية بحتة، على الرغم من أن الفرضيات التي تطرحها متأثرة بأساساتها المعرفية : فالتطوير المعروف باسم نظرية المبادئ والمعايير أو القيم الوسيطة (Principles and Parameters Theory) هو مثال واضح على ذلك . ليس لدى تشومسكي ما يقوله حول كيفية استخدام المعرفة اللغوية: وبعبارة أخرى، هو لا يحاول أن يربط بين اللغة والعملية النشطة من التفكير."[14]

اللسانيات الحاسوبية:

الفرع الخامس، هذا الفرع هو الذي ستنطلق منه دراسة الخط العربي arabic calligraphy في شقه الرقمي، ولكن قبل الخوض في مفهوم اللسانيات الحاسوبية وعلاقتها بالنمذجة الآلية للخط العربي يستوجب منا الحديث عن الحوسبة الآلية والتفكير البشري؛ بعبارة أخرى الحديث تقابليا عن دماغ الانسان أو عقله والحاسوب.

لم تنفك اللغة في دراستها من قبل العلماء على اختلاف مشاربهم وتوجهاتهم وتخصصاتهم، وعلى اختلاف بيئاتهم وعصورهم من علاقتها بالعقل، وذلك كون هذا الأخير هو الجهاز المنتج لها، أما جهاز التصويت فهو مجرد أداة يتم بواسطتها إخراج اللغة في شكل أصوات مادية كما يمكن ذلك بأدوات أخرى كالكتابة والإشارة.

اللسانيات الحاسوبية مجال من مجالات اللسانيات التطبيقية وهي جسر ما بين جملة من العلوم أو ما يعرف بالعلم البيني؛ إلا أن أبرز نشط لها ذلك الربط الذي تؤديه ما بين اللغة و الحاسوب، هذا الربط الذي من أهم أولوياته التحديد الأنطولوجي للمكونات اللسانية من خلال تماثلها في مختلف المواضع من التشكيل اللغوي ومن مختلف السياقات التي تنشأ فيها تلك المكونات اللغوية، وذلك لتتحول إلى خوارزميات حاسوبية يماثل بها العقل الآلي العقل البشري، هذه الأخيرة التي تخضع للمنطق والمماثلة الرياضية كون اللغة قيم مضبوطة لمتوالية من المفاهيم المعقولة والمعبرة عن الممارسات الفعلية والحسية لأي كائن ما.

كما أن اللسانيات الحاسوبية تهتم " بالناتج العملي لنمذجة الاستعمال الإنساني للغة، وهو يهدف إلى إنتاج برامج ذات معرفة باللغة الإنسانية. وهذه البرامج مما تشتد الحاجة إليه لأجْل تحسين التفاعل بين

الإنسان والآلة؛ إذ إن العقبة الأساسية في طريق هذا التفاعل بين الإنسان والحاسوب إنما هي عقبة التواصل،"[15] ذلك لجعل الآلة قادرة على تحقيق كفاءة لغوية تحاكي الوظائف اللغوية عند الإنسان، هذه الكفاءة التي ينعتها تشومسكي وقبله ابن خلدون بالملكة اللغوية.

ولا يمكن لهذا الجهاز الأصم أن يتفاعل مع تعقيدات اللغة فهما وإنتاجا إلا إذا تسلَّم بين ثنايا تلك الخورزميات سيلا من الوسوم النحوية المرافقة لكل المفردات التي تم تزويده بها وفي شتى وظائفها التركيبية.

إن علم اللسانيات الحاسوبية أعقد مما يتصوره الباحثون أحاديو التكوين، فهو علم تتزاوج فيه لدرجة التناسل جملة من التخصصات كالرياضيات والتكنولوجيا وعلم الاجتماع وعلم النفس وغيرها؛ هذه المجالات المعرفية التي تعيش تسارعا في عملية فهم ما يحيط بالإنسان وفي نقاط تقاطع مشتركة كثيرة أفرزت " اهتماما متزايدا ببناء أنسقة حاسوبية قادرة على استعمال اللغات الطبيعية ومعالجتها، أي أنسقة قادرة على محاكاة الطريقة التي تعمل بها الأنسقة البشرية المتميزة ببنيتها الداخلية البالغة التعقيد، ويعتبر بناء الأنسقة الحاسوبية أحد المسالك العلمية الهامة المؤدية إلى نمذجة الأنساق المعرفية الإنسانية."[16]

إن تجسيد اللسانيات الحاسوبية كواقع تفاعلي بين الآلة والممارسات اللغوية الإنسانية يستوجب بالإضافة لما ذكر سابقا أن تتم دراسة اللغة دراسة موضوعية بدأ بماهيتها وكيفية نشأتها وعلاقتها بالإشارة إلى الأحداث والممارسات الفيزيائية ويعتبر هذا جانبا نظريا يتناول " النظريات الصورية للمعرفة اللغوية التي يحتاج إليها الإنسان لتوليد اللغة وفهمها."[17]

وعليه فإن اللسانيات الحاسوبية Computational Linguistics هو " حقل متعدد الاختصاصات يهتم بمعالجة اللغة بواسطة الحاسوب،"[18]هذا التعدد جعله قبلة لعدد هائل من الباحثين على اختلاف تخصصاتهم مما أكسبه " متانة من حيث الكتب والدوريات المنشورة عنه والمؤتمرات المقامة حوله والأساتذة المتخصصون فيه وكثرة الجمعيات والروابط المتعلقة به وما شاكل ذلك."[19]

هذا التصور العام والمتشاكل لماهية اللسانيات الحاسوبية والذي بني من عدة جهات تبعا للتخصصات المتعددة في بناء هرميته أفرز تعاريف عديدة ومتنوعة، ولكن هذا الاختلاف ليس اختلاف تضاد بقدر ما هو اختلاف تنوع، فقد عرفه دفيد كرستال David Crystal على أنه فرع من فروع علم اللغة يتم فيه تطبيق التقنيات والمفاهيم الحاسوبية لتوضيح المشكلات اللغوية والصوتية.

« computational linguistics is A branch of linguistics in which computational technics and concepts are applied to the elucidation of linguistic and phonetic problems. »[20]

ويتضح من هذه الملامسات المختلفة الاتجاهات اتفاق جوهري لهذا العلم " بوصفه دراسة علمية للغة الطبيعية من منظور حاسوبي، وهي دراسة تهدف إلى إنتاج أنظمة حاسوبية قادرة على فهم اللغة الطبيعية وإنتاجها،"[21] وذلك كون تواجد ثنائية تشمل أولا علم اللغة الحاسوبي الذي " يعد مجالا وسيطا بين علم اللغة الذي تضم مادته جميع مظاهر اللغة العامة ومستوياتها المختلفة –الصوتية والصرفية والنحوية والدلالية والمعجمية–"[22] و ثانيا علم الحاسب الآلي الذي " يتم من خلاله بناء نظم معالجة اللغة بمستوياتها المختلفة."[23]

" فرع عن علمي اللغة والحاسب، يهدف إلى تصميم نماذج رياضية للتراكيب اللغوية؛ للتمكن من معالجة اللغة آليا عن طريق الحاسب، كما يعرفه من وجهة نظر لغوية على أنه تشكيل للنظريات والنماذج اللغوية أو تنفيذها على الآلة، ويرى أنه بإمكاننا النظر إليه على أنه وسيلة لتطوير نظريات لغوية جديدة بمساعدة الحاسب."[24]

ويعرف هذا العلم بعدة مسميات أهمها " معالجة اللغات الطبيعية (Natural Language Processing _ NLP)، وتقنيات اللغة الطبيعية (الإنسانية) (Human Language Technologies _ HLT)"[25]وهذا التعدد في التسميات راجع لتنوع المشارب الفكرية والعلمية لرواده حيث أنهم مختلفو التخصصات فمنهم أصحاب التخصصات الانسانية؛ اللسانيات علم الاجتماع وعلم النفس، ومنهم أصحاب التخصصات التقنية؛ الرياضيات، الإعلام الآلي التكنولوجيا، ولكن اتجاهات هذه التعاريف تصب في نقطة واحدة وهي " ذلك العلم الذي تُوَجَّهُ من خلاله أنظمة الحاسوب إلى فهم لغة الإنسان ومحاكاة الذكاء البشري."[26]

مجالات اللسانيات الحاسوبية:

تم تطوير العديد من مجالات البحث، بما في ذلك معالجة اللغة الطبيعية، وتركيب الكلام ، والتعرف على الكلام ، والترجمة الآلية ، وصنع التوافق ، واختبار القواعد النحوية ، والعديد من المجالات التي تتطلب عمليات التعداد والتحليل الإحصائي (على سبيل المثال في الدراسات النصية الأدبية).

« Several research areas have developed, including natural language processing, speech synthesis, speech recognition, automatic translation, the making of concordances, the testing of grammars, and the many areas where statistical counts and analyses are required (e.g. in literary textual studies). »[27]

" تقوم حوسبة اللغة على ثلاث محاور أساسية هي :

تقنيات معالجة النصوص (Text Processing). ومن أمثلتها : الترجمة الآلية، والتلخيص الآلي، والتنقيب في النصوص.

تقنيات معالجة الكلام المنطوق (Speech Processing). ومن أمثلتها : التعرف الآلي على الكلام المنطوق، وتحويل النص المكتوب إلى كلام منطوق.

تقنيات معالجة الصور (Image Processing). ومن امثلتها : التعرف الآلي على الكتابة (Optical Character Recognition)"[28]

الخط العربي النشأة والتطور:

لم يكن العرب يولون أهمية بالغة للكتابة والخط في بدايات تاريخهم، فقد كانوا يعتمدون على المشافهة والحفظ؛ بل وكانوا يرون الكتابة والتدوين دليلا على ضعف العقل والنقص، يقول ابن قتيبة (ت276هـ) في هذا الشأن عن عبد الله بن عمرو عندما أذن له النبي ﷺ بكتابة الحديث " وكان غيره من الصحابة أميين، لا يكتب منهم إلا الواحد والاثنان، وإذا كتب لم يتقن ولم يصب التهجي،"[29]وأكد من المحدثين هذا القول الدكتور إبراهيم أنيس (ت1397هـ) حين قال " فإذا وجد فيهم من يكتب ويقرأ فإنما هو نزيل هبط

إليهم، أو آيب من سفر بعد طول إقامة في أرض متحضِّرة، أو آخذ عن هذين، وهو نادر."[30] ولكن هذا الرأي لم يستسغه كثير من العلماء فقد رُوي عن ابن فارس (ت 395هـ) في كتابه الصاحبي أنه قال " فإنا لا نزعم أن العرب كلها مدرا ووبرا قد عرفوا الكتابة كلها والحروف أجمعها، وما العرب في قديم الزمان إلا كنحن اليوم: فما كلٌّ يعرف الكتابة والخط والقراءة،"[31] وهذا ديدن الشعوب كلها.

وعلى الرغم من أن تاريخ العرب القديم لا يوجد فيه من التدوين والأثر المادي على استعمال الكتابة، إلا أن هناك حفريات ونقوشا في الجزيرة العربية موغلة في التاريخ تكشف أن العرب كانوا على دراية بالكتابة والخط، كما يمكن أن يكون تاريخ اليمن القديم الذي ازدهرت فيه الحياة الاجتماعية والذي قامت فيه دول عظمى سادت العالم آنذاك، أقول يمكن أن يكون دليلا على وجود الكتابة؛ بل والدليل من القرآن الكريم فالرسالة التي بعث بها النبي سليمان عليه السلام إلى ملكة سبأ كانت كتابا، قال تعالى " اذهب بكتابي هذا فألقه إليهم ثم تولَّ عنهم فانظر ماذا يرجعون، قالت يا أيها الملؤا إني ألقي إلي كتاب كريم، انه من سليمان وإنه بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ، ألا تعلوا علي واتوني مسلمين."[32] وهذا الكتاب قرأته الملكة بلقيس وهذا دليل آخر على معرفة العرب الكتابة والقراءة.

نشأة الخط العربي

وردت روايات كثيرة في تاريخ نشأة الخط العربي وأكثرها لا يمت بصلة إلى الموضوعية والدقة العلمية ولا حتى إلا المنطق، فقد جاء عن البلاذري (ت279هـ) ما رواه ابن الكلبي عن الشرقي بن القطامي (ت155هـ) قوله " اجتمع ثلاثة نفر من طيء ببقة، وهم مرامر بن مرة وأسلم بن سدرة وعامر بن جدرة، فوضعوا الخط، وقاسوا هجاء العربية على هجاء السريانية، فتعلمه منهم قوم من أهل الأنبار، ثم تعلمه أهل الحيرة من أهل الانبار،"[33] وأرجع كعب الأحبار (ت32هـ) أصل الخط العربي والسرياني إلى سيدنا آدم عليه السلام،[34] فهذه الرواية على سذاجتها – خاصة ما تعلق بالأسماء التي جاءت على وزن واحد والتي يستبعد أن تكون من باب الصدفة –[35] إلا أنها تقودنا إلى أن الخط العربي قديم جدا، وهي دعم لبعض الروايات الأخرى التي تتسم بالموضوعية والمنطق، ومن هذه الروايات رواية السجستاني (ت 255هـ) التي نقلها ابن دريد (ت 321 هـ) في جمهرته، يقول " كان خط العرب يسمى قديما في الجاهلية الجزم،"[36] واختلف في أصل هذه التسمية، يقول السجستاني " إنما سمي هذا الخط بالجزم لأنه جزم من المسند، أي أخذ منه،[37] والمسند هو خط حمير أيام ملكهم."[38]

وقد ذكر ابن النديم (ت 395 هـ) عن ابن عباس رضي الله عنهما (ت 68 هـ) أنَّ " أول من كتب بالعربية ثلاثة رجال من بولان، وهي قبيلة سكنوا الانبار، وأنَّهم اجتمعوا فوضعوا حروفا مقطعة وموصولة، وهم مرامر بن مرة وأسلم بن سدرة وعامر بن جدرة، ويقال مروة وجدلة، فأما مرامر فوضع الصور، وأما أسلم ففصل ووصل، وأما عامر فوضع الإعجام."[39]

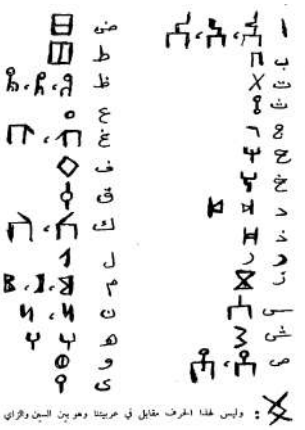
ويرى ابن خلدون (ت 808 هـ) أن جودة الخط إنَّما تكون على قدر الاجتماع و العمران والتناغي في الكمالات، فقد جاء في مقدمته ما مفاده " أن الخط العربي قد انتقل من اليمن، حيث كان بالغا مبالغة من الإحكام والإتقان والجودة في دولة التبابعة إلى الحيرة، حيث دولة آل المنذر نسباء التبابعة، ولم يكن

الخط عندهم من الإجادة كما كان عند التبابعة لقصور ما بين الدولتين، ومن الحيرة لقنه أهل الطائف وقريش."[40]

إلا أنَّ عبد الصبور شاهين يرى أنَّ وجه الشبه بين الخط المسند يكاد يكون منعدما، مستندا في ذلك إلى آراء بعض القدماء في أن أشكال الخطين مختلفين.[41]

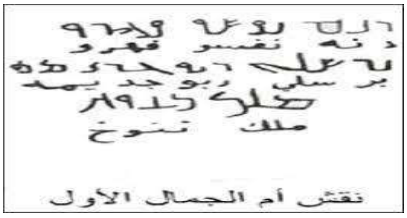
ويعرف الجوهري ت 393 هـ في هذا الصدد خط المسند بأنَّه"خط لحمير مخالف لخطنا هذا"[42]، ويشاطره الرأي في ذلك ابن النديم ت 395 هـ حيث يقول في مصنفه الفهرست " إن حمير كانت تكتب بالمسند على خلاف أشكال ألف وباء وتاء."[43]

نماذج من تشكلات الخط العربي

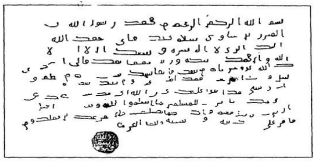


الشكل 1 : حروف الخط المسند في مقابل الحرف العربي المعاصر[44]

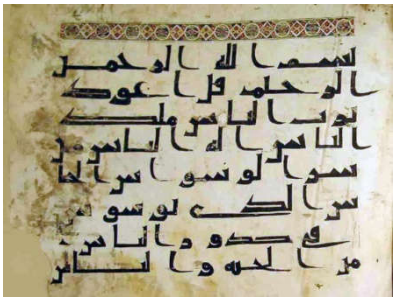
الشكل 2 : نماذج من القلمين النبطي المتأخر والقلم العربي القديم[45]

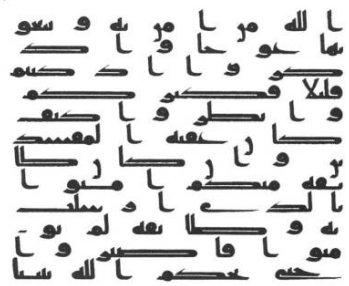


شكل 3 : نموذج للخط النبطي المبكر[46]



الشكل 4 : الخط الحجازي اللين زمن البعثة





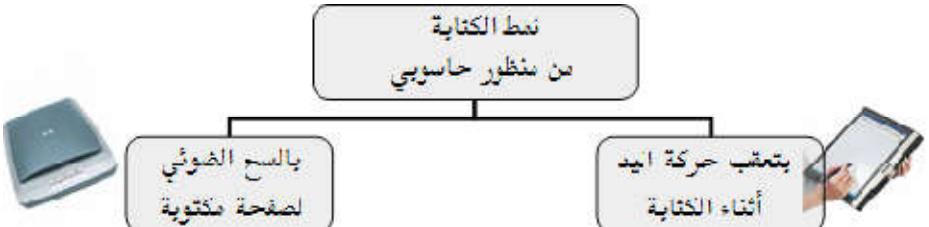
الشكل 5 : الخط العربي في العهد الاموي **الشكل 6** : الخط العربي في العهد العباسي

المعالجة الآلية للخط العربي

لم يكن بالسهل أن يتوصل الانسان إلى وسيلة كالكتابة لتدوين أفكاره وتاريخه فعملية إدراك الوحدة الصوتية للكلام ضمن كل معقد ليس بالأمر اليسير، لذلك مر الإنسان بعدة مراحل لشكل الكتابة لدرجة أن بين المرحلة والمرحلة قد يكون بآلاف السنين، يقول يوهانس فريدريش " إن طريقة الانتقال من الكلمة إلى المقطع أمر صعب جدا ... على الأغلب إن مبتكر الكتابة لم يدرك الصوت كعنصر من الكلمة" فالطريق كان معقدا أيضا في الانتقال من الكتابة الصورية Pictography التي تحاول نقل الخطاب بالمعنى العام بواسطة مجموعة من الرسوم إلى الكتابة بالفكرة Ideography التي تدون الخطاب عن طريق صور رمزية مفردة Ideogram تمثل شيئا أو فكرة."[47]فهذا الجهد المبذول لإخراج الكتابة بهذا الشكل الذي نعايشه حاليا أفنى جيوشا من العلماء على فترات زمنية طويلة من تاريخ البشرية ليصبح في العصر الحالي عملية جد بسيطة لا تأخذ مع طفل عمره ست سنوات عاما واحدا من الدراسة ليكتب ويقرأ بها، إذا ما يقوم به علماء اللسانيات الحاسوبية حاليا في معالجة الكلام البشري عن طريق التقنية لا يعدو أن يكون عملا مماثلا لقدماء السومايين، وعليه فإنه ليس من المستبعد أن يتعامل طفل ست سنوات القرن الثاني والعشرين أو دونه مع الذكاء الاصطناعي في المعالجة الآلية للغة فيحاكي به شخصيات تاريخية كأن ينشأ خوارزمية لرجل آلي بفكر سيبوية أو عبد الرحمن حاج صالح ، والسؤال المطروح بأي لغة سيتمكن هذه الطفل من مبتغاه ؟

إن كوليرا اللغات وموت اللغات وتوقف إنجاب اللغات بسبب الانتشار الواسع للشابكة يوحي بأن العربية هي التي ستكون الأداة الوحيدة لذلك التقدم الرقمي المذهل، لأنها محافظة عبر الأزمنة على نظامها العلائقي الذي يعتبر؛ أي ثبات النظام الداخلي للغة، شرطا أساسيا لأن تحاكي الآلة الإنسان كلاميا.

فعلى الرغم من التعثر الذي شهدته العربية في مجال الرقمنة مقارنة باللغات الأخرى إلا أنها ستتفوق على جميع تلك اللغات لأن تلك اللغات تسير بحركة تباطئية لدرجة التوقف النهائي. وإذا رجعنا إلى تاريخ الطباعة بالحرف اللاتيني والحرف العربي في القرة الخامس عشر فإننا سنجد أن العربية تعثرت في بداياتها ولكن سرعانما تداركت الأمر ولحقت بالركب، لذلك " يمكن النظر لحوسبة خطاطة اللغة بصفة عامة على أنها امتداد لسالفتها من التقنيات الأقدم في هذا الصدد وهي الآلات الكاتبة الكهربائية ومن قبلها الميكانيكية ومن قبلها تقنيات الطباعة بأجيالها المتعاقبة. ومنذ بداية رحلة العالم مع الطباعة لم تكن اللغة العربية في هذا الصدد متخلفة عن نظيراتها من اللغات الأوروبية، فقد أدخل العثمانيون آليات الطباعة مبكرا إلى حاضرتهم المركزية ((الآستانة)) في النصف الثاني من القرن الخامس عشر، وهناك خلاف عما إذا كانوا سبقوا بشكل مستقل أو نقلوها عن الأوربيين الألمان، وبغض النظر عن ذلك فإن الأمر المهم هو أن العثمانيين وقتذاك كانوا يخطون اللغة التركية بالحرف العربي كما أن اللغة العربية نفسها كانت لغة سائدة في العلم والثقافة يتقنها جل المتعلمين الأتراك.."[48]

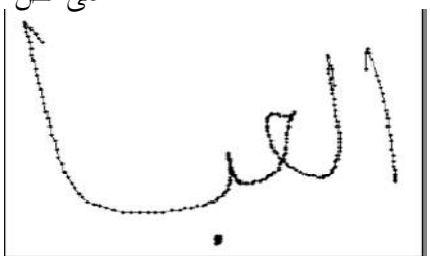


الشكل 7 : تصنيف أنماط إنتاج الكتابة حسب التعامل معها حاسوبيا[49]

أشكال أولية للكتابة الحاسوبية

الشكل 8 : المسح الضوئي لصفحة تحتوي على كتابة بخط اليد

الشكل 09 : المسح الضوئي لصفحة تحتوي على نص مطبوع



الشكل 10 : مثال على رقمنة الكتابة اليدوية أثناء الكتابة

تخضع المعالجة الآلية للغات إلى توصيف مستوياتها؛ الصورية والمعنوية، الشكل والدلالة، المكتوب والمنطوق، وذلك بجملة من الوسوم التي بإمكانها إنشاء تغطية شبه كلية لمكونات أي جانب منها حتى تتمكن الآلة من الموازنة بين المدخلات والمخرجات لتنشِئ خوارزميات تقاربية لفهم طبيعة اللغة البشرية وتجسيدها آليا؛ إما من خلال كتابة المسموع أو نطق المكتوب، بالإضافة إلى فهم وإفهام ما يعطى كإنتاج لغوي ضمن حقل معرفي معين ومحدد، ولا يمكننا ذلك إلى إذا حددنا وحدة قياس مركباتها، فالكتابة تعالج وفق وحدة معينة " وهي الجرافيم Grapheme أو المحرَف، الذي يعد أصغر وحدة خطية للغة المكتوبة؛ ويقابله في اللغة المنطوقة الفونيم Phoneme الذي يعبِّر عن الصوت اللغوي. وتشمل الوحدات الكتابية

(الجرافيمات) – عموما – جميع محارف اللغات الانسانية، سواء كانت حروفا أم أرقاما أم رموزا رياضية، أم غير ذلك،"[50] وقد يكون للجرافيم أكثر من شكل وذلك لاعتبارات مختلفة كنظام الكتابة لأي لغة، أو تبعا لموضعه داخل النظام الخطي، " ويعرف الشكل الواحد للجرافيم بـ (ألوجراف Allograph). وعلى سبيل المثال، يأخذ جرافيم (A) في الانجليزية ألوجرافين فقط، هما (A, a)، بينما يأخذ جرافيم (ء) [الهمزة] في العربية خمسة ألوجرافات، هي (ء، أ، إ، ؤ، ئ)."[51]ولذلك يجب مراعات هذه التنوعات وضبطها بشكل محكم لتتمكن الآلة من معالجتها في جميع حالاتها.

من القضايا المهمة التي يجب أخذها بعين الاعتبار في توسيم الجرافيم علاقته بالجرافيمات الأخرى من حيث الشكل داخل الكلمة من لغة إلى أخرى، فاللغات الأوروبية يمكن كتابة حروفها منفصلة، بينما اللغة العربية تتصل حروفها فيما بينها، "ومن وجهة نظر أي آلية حاسوبية للتعرف على الأنماط الرسومية، فإنه مع تثبيت جميع الظروف الأخرى يكون التعرف على الأنماط المنفصلة أيسر من التعرف عليها وهي متصلة ببعضها البعض، حيث يتوجب في حالة الاتصال حل مسألة تعيين حدود كل رمز رسومي (جرافيم) وهو ما يطلق عليه الباحثون في هذا المجال اسم (التقطيع segmentation) إضافة بالطبع إلى مسألة (التعرف *recognition*)على الحرف الذي يرمز إليه كل جرافيم."[52]

المرحلة التمهيدية preprocessing phase:

وهي بمثابة حجر الزاوية في العملية ككل لأنها تتمحور حول " إلتقاط الصور وتحديد مناطق الكتابة فيها بالمعالجة المسبقة للصور preprocessing وذلك لتحسين جودة ووضوح النصوص فيها،"[53] وذلك خدمة للمراحل التي تليها. تخضع هذه المرحلة لنظامين مختلفين تمام الاختلاف في عملية ادخال البيانات المراد معالجتها؛ أحدهما مباشر والآخر غير مباشر، ففي النظام الاول غير المباشر Offline تتم معالجة النص المكتوب باليد أو الآلة على الورق عن طريق مسح ضوئي كمدخلات للحاسوب على شكل صور وبالتالي فإن مفهوم الكتابة غير مدرك على مستوى الحاسوب " من هذا المنطلق ظهرت الحاجة إلى حل هذه المشكلة وإيجاد برامج تقوم بتحويل صورة المستند إلى مستند قابل للتعديل والبحث. فظهرت برامج عديدة تقوم بالتعرف الآلي على الحروف والأرقام والكلمات وغيرها من الرموز المكتوبة على الورق،"[54] أما في النظام المباشر Online " يتم التعرف على الكتابة اليدوية مباشرة أثناء الكتابة، باستخدام قلم ضوئي على شاشة خاصة مربوطة بالحاسب، أو باستخدام الفأرة للكتابة يدويا على الشاشة. رغم أن هذه الطريقة مخصصة للتعرف على الكتابة اليدوية، إلا أن النظام يمكنه التسجيل اللحظي لكيفية وتسلسل الكتابة، مما يسهل عملية التعرف،"[55] غير أن كلى النظامين معرض لجملة من النقائص تحول دونهما وتحقيق الحاسوب محاكاة اللغة البشرية المكتوبة كما يحاكيها العقل البشري،" ففي النظام غير المباشر، من الممكن حصول تشوهات على البيانات المدخلة في الصورة الناتجة من عدة أسباب مثل انخفاض جودة الماسح الضوئي Scanner أو انخفاض جودة الورق أو التشوهات المتكونة من الورق وغيرها،"[56] أما ما تعلق بالنظام المباشر فإنه " تظهر مشاكل التشوهات من خلال عيوب في اجهزة الادخال مثل الألواح الرقمية Digital Tablets،"[57]

ومن أهم التشوهات التي يتم معالجتها في هذه المرحلة التمهيدية –وذلك قصد التقليل منها وتهيئة البيانات في حلة لائقة للمراحل اللاحقة– ما يتعلق بحجم النص المدخل *Text Size لأن* الاختلاف في حجم النص الواحد قد يعطي أبعادا مختلفة للنص المراد التعرف عليه مما يترتب عليه اختلافا في خصائصه ومن ثم يتسبب ذلك في عدم الحصول على مخرجات صحيحة من النظام، وموقع النص *Text Position* حيث أنه قد تختلف إحداثيات النص المدخل عندما يقوم الكاتب بالكتابة في المساحة المخصصة للكتابة، حيث أن هذا الاختلاف قد يعطي النظام أبعادا مختلفة للنص الواحد مما قد يقلل من إمكانية التعرف على النص الصحيح، وعدم تساوي المسافات بين جزئيات الكلمة وبين الكلمة وما تليها نتيجة لسرعة الكتابة على سطح جهاز الادخال، [حيث] من المحتمل الحصول على نقاط غير موزعة بشكل متساو، مما يقلل من انتظام النقاط المطلوبة للحصول على الشكل المطلوب، كذلك من التشوهات التي تلحق بالمدخلات الاختلالات في الكتابة المتعلقة بما يراد ادخاله كبيانات تذبذب النص *Zigzag Shape* قد يواجه بعض مستخدمي الأنظمة الضوئية صعوبة في التحكم بالقلم والكتابة بالشكل الصحيح كما هو معتاد على الورق، مما ينتج عنه تذبذب وتشوه في الكتابة وهذا بدوره يقلل من دقة النص المدخل ويزيد في عدم وضوحه،مشكل آخر يتعلق بانحراف النص المكتوب عن مساره الصحيح *Slant and Skew of* the Handwriting ميل الكتابة هو الكتابة بشكل مائل على المحور الأفقي، أما انحراف الكتابة فهو الكتابة بشكل مائل على المحور العمودي. وهو من الأخطاء والمشاكل الشائعة عند إدخال النصوص اليدوية المباشر، مشكل آخر تعاني منه هذه المرحلة متعلق بالزوائد عند أطراف الكتابة *Hooking* تعتبر الزوائد إحدى المشاكل التي تقلل من كفاءة النظام. تنتج هذه الزوائد عند البدء في الكتابة على السطح الضوئي بدون التركيز على بداية رسم النص، كذلك قد تنتج عند رفع القلم عن السطح حيث تسمى كلابات Hook،"[58] ولأجل هذه النقائص على مستوى المسح الضوئي جاء دور المعالجة الأولية، أو كما أشرنا إليها في عنوان الفقرة بالمرحلة التمهيدية، لـ " تنقية الصورة من الشوائب غير النصية (مثل الرسوم والاطراف السوداء)، ويعتمد وضوح الصورة على عوامل عدة، منها تاريخ الأصل وطريقة الطباعة (ليزر أو نقطية)، وضوح الخط، جودة الورق، سلامة الأصل من التظليل والتخطيط،"[59]

مرحلة التقطيع segmentation

تؤدي هذه المرحلة دورا محوريا في عملية المعالجة الآلية للمكتوب اللغوي، حيث أنها تقوم بعملية تحليل الكل إلى أجزاء غير قابلة للتقسيم ومحافظة في الآن ذاته على شكل المادة اللغوية، وهذا العمل فيها إشارة إلى الدقة المتناهية في معالجة الانتاج اللغوي المكتوب للوصول إلى نتائج متماثلة بين الممارسة البشرية والممارسة الانسانية للمكتوب اللغوي، فهي تقوم "بتفريق النص المكتوب إلى حروف منفصلة بعضها عن بعض، ووضع كل حرف في مصفوفة بشكل منفرد تمهيدا لاستخراج السمات لكل حرف، وتعتبر هذه المرحلة من أصعب المراحل وأكثرها اهمية في تحديد نسبة التعرف."[60]

وهي المرحلة التي أشرنا إليها في الفقرة السابقة، والتي تعمل على الحصول على " أجزاء أساسية من النصوص كالحرف بالنسبة للكلمة، وقد يجري التقطيع على عدة مستويات، كتقطيع صور النصوص إلى أسطر، وتقطيع الأسطر إلى كلمات أو دون الكلمات Subwords (دون الكلمات) هو ما يكتب

متصلا في العربية لعدم انتهاء الكلمة وعدم ورود حرف لا يتصل بما بعده ضمنها، وتسمى أيضا Pieces of arabic words أو Connected Components في أحد معانيها."[61]

" أشهر مستويات التقطيع وأهمها على الاطلاق .... هو تقطيع صور النصوص إلى محارف Character Segmentation حيث إنه ... يسهل عمليات التعرف الآلي اللاحقة."[62]

مرحلة استخلاص الملامح Feature Extraction

من مهام هذه المرحلة " التعبير المختصر والمركز عن الصور المراد التعرف عليها بأهم ملامحها *Features* وذلك تصغيرا لحجم البيانات وتسريعها لوقت المعالجة من جانب، وتركيزا على ما يهم القارئ من المحارف وإهمالا لما لا يهم القراءة كفروقات الخطوط الفردية، من الجانب الآخر،"[63] ثم تأتي مرحلة التصنيف Classification، هذه المرحلة التي " يطلق عليها مجازا (التعرف) تهدف لمعرفة رمز النص من ملامحه بعد تعلمه من أمثلة. تمر المصنفات بمرحلتين على الأقل: مرحلة التدريب والنمذجة Training and Modelling، ثم مرحلة التعرف والتصنيف الفعلي Recognition and Classification. كما تمر بعض المصنفات بمرحلة تحقق Validation لتحسين تدريبها ونمذجتها، وبمرحلة اختبار *Testing* لتقرير نسب نجاحها في البحوث العلمية والمسابقات،"[64] ثم تأتي المرحلة الأخيرة والتي يطلق عليها مرحلة المعالجة اللاحقة Postprocesing التي يستعان فيها " للتعرف الآلي بمعاجم Lexicons وقواعد لغوية Linguistic Rules لما تقبله اللغة أو ترفضه، أو بنماذج احصائية Statistical *Models* للشائع لغويا كالورودات الأقرب N-*Grams* ، لترجيح أو استبعاد نتائج التعرف، لا سيما عندما تكون الكلمات المراد التعرف عليها محصورة في مجال محدد كالطب أو الهندسة أو أسماء مدن Domain-Specific."[65]

إن علاقة الحروف العربية فيما بينها علاقة ثنائية؛ إذ أن بعضها يقبل الاتصال بغيره والبعض الآخر لا يقبل، ومسألة معالجة هذه القضية آليا ليست بالأمر الهين " ففي أنظمة تجارية لبرامج التعرف على الحروف العربية نجد أن نسبة تقارب 10 % من الأخطاء على مستوى الكلمات، يكون أصل الخطأ فيها خطأ في فصل الكلمات بعضها عن بعض، هذا إذا كانت الوثيقة واضحة،"[66] وهذه الإشكالية المخالفة لطبيعة الحرف اللاتيني هي التي ساعدت في عرقلة تطور الحرف العربي بوتيرة متسارعة حيث أن البرامج الحاسوبية مكيفة تبعا للحرف اللاتيني.

وقد جاءت هذه البرامج لتمكن الإنسان " التعرف الآلي على الكتابة ....لمحاكاة بعض قدرات البشر،"[67] من خلال التعرف على الأنماط وتمييز الحروف، غير أن أدوات القراءة الآلية أصبحت تتضمن أيضا عمليات أخرى " كمعالجة الصور .....و تحديد مواضع الكتابة في الصور، وتحسين جودة الصور لتسهيل التعرف على كتابتها، وتصحيح نتائج التعرف الآلي على الكتابة لغويا،"[68] وهذه إشكالية أخرى تضاف إلى صعوبة تناول الخط العربي وفق البرامج الحاسوبية المكيفة تبعا للحرف اللاتيني.

– خاتمة:

أهم ما توصلنا إليه في هذا البحث هو أن من أسباب تأخر اللغة العربية في مجال اللسانيات الحاسوبية – على غرار اللغات الأخرى التي قطعت شوطا كبيرا في هذا المجال – هو أن الأبحاث التي تعتني بالمعالجة

الآلية للغة على جميع مستوياتها لاسيما معالجة الخط العربي تجرى داخل قوقعة التخصص الواحد، فاللسانيون يعملون لوحدهم و الحاسوبيون كذلك، في حين يستدعي الامر تضافر الجهود في إطار التخصصات البينية و العمل تحت مظلة واحدة يجتمع فيها اللغوي والرياضي والحاسوبي وعالم الاجتماع وعلماء النفس من خلال إنشاء مخابر بحث بينية تكون فرقها هذه التخصصات التي ذكرناها سابقا ..

–. قائمة الإحالات:

1 ميشال زكريا، 1993، قضايا ألسنية تطبيقية – دراسات لغوية اجتماعية نفسية مع مقارنة تراثية –، دار العالم للملايين، بيروت، ط1، ص 9

2 مازن الوعر، 1989، دراسات لسانية تطبيقية، دار طلاس للدراسات والترجمة والنشر،ط1،.ص 23

3 ميشال زكريا، قضايا ألسنية تطبيقية، ص 9 – 10

4 عبده الراجحي، 1995،علم اللغة التطبيقي وتعليم العربية، دار المعرفة الجامعية، الإسكندرية، دط، ص 13

5 شارل بوتون، دت، اللسانيات التطبيقية، ت: قاسم مقداد وآخر،دار الوسيم للخدمات الطباعية، دمشق،دط، ص 7

6 عبده الراجحي، علم اللغة التطبيقي وتعليم العربية، ص 7

7 الحبيب النصراوي، (دت)، الاحتكاك اللغوي وأثره في العربية الحديثة، مجلة المجلس الأعلى للغة العربية، ج1، ص 105

8 ينظر:إبراهيم السامرائي، 1981، التطور اللغوي التاريخي، دار الاندلس للطباعة ، بيروت، ط2،ص 146 – 150

9 جوليت غارمادي، 1990، اللسانة الاجتماعية،ت:د.خليل أحمد خليل، دار الطليعة للطباعة، بيروت، ط1،ص 17

10المرجع نفسه، ص 16

11بيتر ترادجل، السوسيولسانيات – مدخل إلى دراسة اللغة في علاقتها بالمجتمع –، ت: محمد كرم الدكالي، منشورات إفريقيا الشرق، الدار البيضاء

12استبدال عبارة (منتصف هذا القرن) بعبارة (منتصف القرن الماضي) باعتبار أن إشارة الكاتب لهذه الفترة الزمنية كانت في نفس القرن المشار إليه.

13بول كوبلي،2016،دليل راوتليدج لعلم السيمياء واللغويات،ت:هبة شندب،المنظمة العربية للترجمة،بيروت،ط1،ص 315

14المرجع نفسه، ص 315

15 نهاد الموسى، 2007، التركيب الإضافي في العربية ـ نحو توصيف جديد في ضوء اللسانيات الحاسوبية ـ ، رسالة ماجستير، الجامعة الأردنية، ص: 53-54.

16 سناء منعم، 2015، اللسانيات الحاسوبية والترجمة الآلية ـ بعض الثوابت النظرية والإجرائية ـ منشورات مختبر العلوم المعرفية، عالم الكتب الحديث، الأردن، ط1،ص 64

17 نهاد الموسى، التركيب الإضافي في العربية ـ نحو توصيف جديد في ضوء اللسانيات الحاسوبية ـ ، ص: 54

18 روسلان ميتكوف وآخرون، 2018، دليل أكسفورد في اللسانيات الحاسوبية ج1 أساسيات علم اللسانيات الحاسوبية، ت: طارق إبراهيم، المنظمة العربية للترجمة، بيروت، ط1، ص 9

19 المرجع نفسه، ص 11

20 A Dictionary of Linguistics and Phonetics 6th Edition. David Crystal, Sixth Edition, Blackwell Publishing Ltd. 2008, p97

21 أحمد فؤاد باشا، 2018، تنمية اللغة العلمية العربية وتحديات التعريب والحوسبة والتجدد الحضاري، دار الكتب والوثائق القومية، القاهرة، دط، ص 163-164

22المرجع نفسه، ص 164

23المرجع نفسه، ص 164

24 منصور بن محمد الغامدي وآخرون، 2017، مدخل إلى اللسانيات الحاسوبية، تح : د. عبد الله بن يحيى الفيفي، دار وجوه للنشر والتوزيع، الرياض ط1،ص 6

25 محسن رشوان وآخرون، 2019، مقدمة في حوسبة اللغة العربية، تح : المعتز بالله السعيد وآخر، دار وجوه للنشر والتوزيع، الرياض ط1، ص 17
26 المرجع نفسه، ص 17
27 A Dictionary of Linguistics and Phonetics 6th Edition. David Crystal, p 97
28محسن رشوان وآخرون، مقدمة في حوسبة اللغة العربية، ص 17
29غانم قدوري، 1982،رسم المصحف دراسة لغوية تاريخية، اللجنة الوطنية للاحتفال بمطلع القرن 1515، العراق، ط1، ص21
30 ابراهيم أنيس، 2003، في اللهجات العربية، مكتبة الانجلو المصرية، القاهرة، (د. ط)، ص 33
31ابن فارس أحمد، الصاحبي في فقه اللغة وسنن العرب في كلامها، المكتبة السلفية، مطبعة المؤيد، القاهرة، (دط)، 1910، ص 8 -9
32النمل : 28-29-30-31
33البلاذري احمد بن جابر ، 1901م، فتوح البلدان، شركة بيع الكتب العربية، القاهرة، ط1، ص 476
34ينظر : ابن عبد ربه، (د. ت)، العقد الفريد، لجنة التأليف والترجمة والنشر، القاهرة، (د. ط)، ج4، ص 157.
35ينظر : الطاهر أحمد مكي، 1970م، دراسة في مصادر الأدب، دار المعارف، ط2، ج1، ص 38.
(36) ابن دريد، جمهرة اللغة، 1345 هـ. دائرة المعارف العثمانية، حيدر أباد، ط1، ، ج2، ص91.
(37) ابن جني، 1954، سر صناعة الاعراب، مصطفى البابي الحلبي، القاهرة، ط1، ج1، ص 45.
(38) ابن دريد، جمهرة اللغة، ج2، ص 91.
(39) ابن النديم، 1881م، الفهرست، ليبسك، ألمانيا، (د. ط)، ص 4-5.
(40) ابن خلدون، 1956،كتاب العبر وديوان المبتدأ والخبر، دار الكتاب اللبناني، بيروت، دط، ج1، ص 756/755
(41) ينظر : عبد الصبور شاهين، 1426، تاريخ القرآن، دار نهضة مصر، دط، ص 64.
(42) الجوهري، 1956، تاج اللغة وصحاح العربية، دار الكتاب العربي، القاهرة، دط، ج1، ص 487.
(43) ابن النديم، الفهرست، ص 5.
44 المفصل في تاريخ العرب قبل الإسلام، د. جواد علي، جامعة بغداد، ط2، 1993، ج8، ص 220
45 المرجع نفسه ج8، ص 181
46 تطوير الخط العربي وفنونه، د. عصام العسيري، مجلة الرياض الرقمية https://www.alriyadh.com/1830855
47 يوهانس فريدريش،2013،تاريخ الكتابة،ت:د.سليمان أحمد الضاهر،الهيئة العامة السورية للكتاب،دط،ص15-16
48 مُحَمَّد عطية وآخرون، 2019، تطبيقات أساسية في المعالجة الآلية للغة العربية، تح : : د. محسن رشوان و المعتز بالله السعيد، دار وجوه للنشر والتوزيع، الرياض ط1، ص 24
49 تطبيقات أساسية في المعالجة الآلية للغة العربية، مُحَمَّد عطية وآخرون، مركز الملك عبدالله لخدمة اللغة العربية، الرياض ط1، 2019، ص 35
50 مُحَمَّد عطية وآخرون، 2019، العربية والذكاء الاصطناعي، ت : المعتز بالله السعيد، دار وجوه للنشر والتوزيع، الرياض، ط1، ص 96
51 المرجع نفسه، ص 97
52 مُحَمَّد عطية وآخرون، تطبيقات أساسية في المعالجة الآلية للغة العربية، ص 24
53 المرجع نفسه، ص 24
54 روعة أبو الشامات، 2014، التعرف الآلي على الأحرف العربية باستخدام سماتها البنيوية، مجلة جامعة تشرين للبحوث والدراسات العلمية_سلسلة العلوم الهندسية، المجلد 36 العدد6،ص 514
55 المرجع نفسه، ص 514
56 مصطفى أبو زريدة وآخرون، 2013، المراحل التمهيدية وأهميتها في أنظمة التعرف على الكتابة العربية، المجلة الدولية للتطبيقات الاسلامية في علم الحاسب والتقنية، المجلد 1، العدد 2، سبتمبر / أيلول 14-23، ص 18

57 المرجع نفسه، ص 18
58 ينظر : المرجع نفسه، ص 18
59 التعرف الآلي على الأحرف العربية باستخدام سماتها البنيوية، روعة أبو الشامات، مجلة جامعة تشرين للبحوث والدراسات العلمية_سلسلة العلوم الهندسية، المجلد 36 العدد6، 2014، ص 518
60 روعة أبو الشامات، التعرف الآلي على الأحرف العربية باستخدام سماتها البنيوية، ص 518
61 يوسف سالم العريان وآخرون، 2019، تطبيقات الذكاء الاصطناعي في خدمة اللغة العربية، تح : د. يوسف سالم العريان، دار وجوه للنشر والتوزيع، الرياض ط1، ص 24
62 المرجع نفسه، ص 24
63 المرجع نفسه، ص 26
64 المرجع نفسه، ص 27
65 المرجع نفسه، ص 28
66 محسن رشوان وآخرون، مقدمة في حوسبة اللغة العربية، ص 24
67 يوسف سالم العريان وآخرون، تطبيقات الذكاء الاصطناعي في خدمة اللغة العربية، ص 17
68 المرجع نفسه، ص 17

المصادر والمراجع

1. الحبيب النصراوي، دت، الاحتكاك اللغوي وأثره في العربية الحديثة، مجلة المجلس الأعلى للغة العربية،
2. الجوهري، 1956، تاج اللغة وصحاح العربية، دار الكتاب العربي، القاهرة، دط،
3. مُحَمَّد سديد، 1998، تاريخ الحرف العربي المطبوع، مجلة التاريخ العربي، الرباط، المغرب، العدد 08،
4. عبد الصبور شاهين، 1426، تاريخ القرآن، دار نهضة مصر، دط،
5. يوهانس فريدريش، 2013، تاريخ الكتابة، ت : د.سليمان أحمد الضاهر، الهيئة العامة السورية للكتاب، دط،
6. نهاد الموسى، 2007، التركيب الإضافي في العربية ـ نحو توصيف جديد في ضوء اللسانيات الحاسوبية ـ ، رسالة ماجستير، الجامعة الأردنية،
7. مُحَمَّد عطية وآخرون، 2019، تطبيقات أساسية في المعالجة الآلية للغة العربية، تح : : د. محسن رشوان و المعتز بالله السعيد، دار وجوه للنشر والتوزيع، الرياض ط1
8. يوسف سالم العريان وآخرون، 2019، تطبيقات الذكاء الاصطناعي في خدمة اللغة العربية، تح : د. يوسف سالم العريان، دار وجوه للنشر والتوزيع، الرياض ط1،
9. إبراهيم السامرائي، 1981، التطور اللغوي التاريخي، دار الاندلس للطباعة والنشر والتوزيع، بيروت، ط2،
10. روعة أبو الشامات، 2014، التعرف الآلي على الأحرف العربية باستخدام سماتها البنيوية، مجلة جامعة تشرين للبحوث والدراسات العلمية_سلسلة العلوم الهندسية، المجلد 36 العدد6،
11. أحمد فؤاد باشا، 2018، تنمية اللغة العلمية العربية وتحديات التعريب والحوسبة والتجدد الحضاري، دار الكتب والوثائق القومية، القاهرة، دط،
12. ابن دريد، 1345 هـ، جمهرة اللغة، دائرة المعارف العثمانية، حيدر أباد، ط1،
13. مازن الوعر، 1989، دراسات لسانية تطبيقية، دار طلاس للدراسات والترجمة والنشر،ط1،.
14. الطاهر أحمد مكي، دراسة في مصادر الأدب، دار المعارف، ط2، 1970م،
15. روسلان ميتكوف وآخرون، 2018، دليل أكسفورد في اللسانيات الحاسوبية ج1 أساسيات علم اللسانيات الحاسوبية، ت: طارق إبراهيم، المنظمة العربية للترجمة، بيروت، ط1،
16. بول كوبلي، 2016،دليل راوتليدج لعلم السيمياء واللغويات، ت : هبة شندب، المنظمة العربية للترجمة، بيروت، ط1،
17. غانم قدوري، 1982، رسم المصحف دراسة لغوية تاريخية، اللجنة الوطنية للاحتفال بمطلع القرن 1515 هـ، العراق، ط1،
18. ابن جني، 1954، سر صناعة الاعراب، مصطفى البابي الحلبي، القاهرة، ط1،

19. بيتر ترادجل، (دت) ، السوسيولسانيات – مدخل إلى دراسة اللغة في علاقتها بالمجتمع ---، ت: مُحَمَّد كرم الدكالي، منشورات إفريقيا الشرق، الدار البيضاء
20. جواد علي، المفصل في تاريخ العرب قبل الإسلام، جامعة بغداد، ط2، 1993
21. ابن فارس أحمد، 1910، الصاحبي في فقه اللغة وسنن العرب في كلامها، ت:المكتبة السلفية، مطبعة المؤيد، القاهرة، (د. ط)،
22. مُحَمَّد عطية وآخرون، 2019، العربية والذكاء الاصطناعي، تح : : د. المعتز بالله السعيد، دار وجوه للنشر والتوزيع، الرياض، ط1،
23. ابن عبد ربه، العقد الفريد، (د. ت)، لجنة التأليف والترجمة والنشر، القاهرة، (د. ط)،
24. روجر بنروز، 1998، العقل والحاسوب وقوانين الفيزياء، ت : مُحَمَّد وائل الأتاسي وآخر، دار طلاس للدراسة والترجمة، دمشق، ط1،
25. عبده الراجحي، 1995، علم اللغة التطبيقي وتعليم العربية، دار المعرفة الجامعية، الإسكندرية، دط،
26. البلاذري احمد بن جابر ، 1901، فتوح البلدان، شركة بيع الكتب العربية، القاهرة، ط1،
27. ابن النديم، 1881،الفهرست، ليبسك، ألمانيا، (د. ط)،
28. ابراهيم أنيس، 2003، في اللهجات العربية، مكتبة الانجلو المصرية، القاهرة، (د. ط)،
29. روجر بنروز وآخرون، 2009، فيزياء العقل البشري والعالم من منظورين، ت : عنان علي الشهاوي، هيئة أبو ظبي للثقافة والتراث (كلمة)، أبوظبي، ط1،
30. ميشال زكريا، 1993،دار قضايا ألسنية تطبيقية – دراسات لغوية اجتماعية نفسية مع مقارنة تراثية –، العالم للملايين، بيروت، ط1،
31. ابن خلدون، 1956، كتاب العبر وديوان المبتدأ والخبر، دار الكتاب اللبناني، بيروت، دط،
32. جولييت غارمادي، 1990،اللسانة الاجتماعية، ت: د. خليل أحمد خليل، دار الطليعة للطباعة والنشر، بيروت، ط1،
33. شارل بوتون، (دت)، اللسانيات التطبيقية، ت : د.قاسم مقداد ومُحَمَّد رياض المصري، دار الوسيم للخدمات الطباعية، دمشق، دط،
34. سناء منعم، 2015، اللسانيات الحاسوبية والترجمة الآلية ـ بعض الثوابت النظرية والإجرائية ـ منشورات مختبر العلوم المعرفية، عالم الكتب الحديث، الأردن، ط1،
35. عطية سليمان أحمد،2019،اللسانيات العصبية اللغة في الدماغ،الأكاديمية الحديثة للكتاب الجامعي، دط،
36. منصور بن مُحَمَّد الغامدي وآخرون، 2017، مدخل إلى اللسانيات الحاسوبية، تح : د. عبد الله بن يحيى الفيفي، دار وجوه للنشر والتوزيع، الرياض ط1،
37. مصطفى أبو زريدة وآخرون، 2013، المراحل التمهيدية وأهميتها في أنظمة التعرف على الكتابة العربية، المجلة الدولية للتطبيقات الاسلامية في علم الحاسب والتقنية، المجلد 1، العدد 2، سبتمبر / أيلول 14-23،
38. عز الدين البوشيخي، 2018، المعرفة اللسانية وزوايا النظر، السيمنار الأسبوعي (2018/02/28)، المركز العربي للأبحاث ودراسة السياسات، الدوحة.
39. محسن رشوان وآخرون، 2019، مقدمة في حوسبة اللغة العربية، تح : المعتز بالله السعيد وآخر، دار وجوه للنشر والتوزيع، الرياض ط1،

المصادروالمراجع الأجنبية

1. A Dictionary of Linguistics and Phonetics 6th Edition. David Crystal, Sixth Edition, Blackwell Publishing Ltd. 2008,