

اللغة العربية وتكنولوجيا المعلومات

9: ا نحو نظرة أوسع للغة

اللغة، بلا منازع، هي أوضح خصائص الجنس البشري تمييزاً له، ودلالة على طبيعته الفريدة، وتأكيد لحقيقة تسنمه الذروة العليا لمرتقى الكائنات الحية، وهي-أي اللغة-ليست مجرد نظام لتوليد الأصوات الناقلة للمعنى، فهي كما قالوا عنها مرآة العقل، وأداة الفكر، ووعاء المعرفة، والهيكـل الحديدي الذي يقيم صلب المجتمعات الإنسانية.

واللغة، بجانب كونها ظاهرة نفسية فسيولوجية، نشاط جماعي ناتج عن التفاعلات الاجتماعية، وهي ترتبط ارتباطاً وثيقاً بحاضر جماعتها وتاريخها، فتاريخ اللغة يكاد أن يكون هو تاريخ شعوبها، ونحن نتاج لغتنا، فكما يقول الفيلسوف الألماني ولهام همبولدت صاحب نظرية الحتمية اللغوية (127): «إن الناس هم تبع في تفكيرهم وإحساسهم ومشاعرهم ونظرتهم للكون للعادات التي اكتسبوها من خلال ممارستهم للغة»، لذا فاللغة تسمو بأهلها مع سموهم، وتتخطى الشعوب مع لغاتها

وبلغاتها، واعتزاز الأمم بلغاتها أمر معروف للجميع فلسنا وحدنا الذين وصفنا لغتنا العربية بأنها «أشرف اللغات وأوسعها وأفصحها وأولها وآخر ما بقي منها»، و«إن لسان العرب فوق كل لسان ولا تدانيها لسان أخرى من أسنة العالم جمالا ولا تركيا ولا أصولا» (55)، فقد سبقنا اليونانيون في ذلك، وهم يصفون لغات غيرهم بنباح الكلاب، ونقيق الضفادع، ويستحل اليهود الكذب بغير العبرية (14).

وبين جميع أنواع السلوك الذهني تبرز اللغة كأكثرها تجليا حيث يمكننا أن نمسك بطرفي دخلها وخرجها فكل ما ننطق به ونسمعه محسوس ومدرك بصورة مباشرة، ولنقارن ذلك لتوضيح ما نقصده بغموض وتميع العمليات الذهنية المصاحبة لإدراكنا البصري وردود فعلنا الذهنية والنفسية تجاه ما نشاهده، أو نتعرض له، يفسر لنا ذلك أهمية اللغة كمدخل لسبر أغوار الآليات الخفية للذهن البشري عموما، وتلك التي تتعامل مع الرموز بشكل خاص، بل ليس من قبيل الإسراف القول إن اللغة ركيزة أساسية لوحدة العلوم ووحدة الفكر، وتمثل مناهجها نموذجا معرفيا إرشاديا paradigm يمكن تطبيقه على ما هو خارج نطاق اللغة. وهاهي علوم الوراثة تقتض من اللغة مصطلحات من قبيل الشفرة الوراثية والجملة الوراثية والنص الوراثي والفائض redundancy.

وهذا شأنها لا يبدو من المستغرب، أن تتجاوز آثار اللغة مجالات التربية والثقافة إلى المجالات السياسية والاقتصادية والتكنولوجية بل والعسكرية أيضا، ولم يكن انتشار آثار اللغة وتغلغلها في الكيان المجتمعي في يوم من الأيام كما نشهده الآن مع التقدم الهائل في تكنولوجيا المعلومات، فاللغة في مجتمع المعلومات موضع الصدارة، وكيف لا واللغة أهم مقومات ذكاء الإنسان محور هذا المجتمع، ومصدر الذكاء الاصطناعي للكمبيوتر أداة هذا المجتمع الرئيسية.

أشرنا في الفصل السابق عن علاقة السياسة باللغة، وهي العلاقة التي تحظى بمزيد من الاهتمام في أيامنا هذه، لقد كان فن الخطابة شرطا من شروط الحاكم في الماضي، والآن أصبح الخطاب الإعلامي ومناوراته اللغوية من أخطر أسلحة السلطة في العصر الحديث، ولقد أدت التغيرات السياسية التي شهدتها العالم في السنوات الأخيرة إلى ما يمكن أن نسميه بظاهرة

«الانفجار اللغوي»، فما نحن نسمع عن إحياء الشعوب للغاتها المندثرة أو المحرمة من لغة كاتلان في إسبانيا ولغة أهل ويلز، وإحياء التحدث بلهجة مندرين الصينية في سنغافورة، بعد أن سادتها الإنجليزية، وأخذت اللغة بعدا سياسيا-اقتصاديا-ثقافيا مع ظهور شعار أوروبا الموحدة حيث ينظر إلى تعدد لغات المجموعة الأوروبية تارة كعقبة أساسية تحول دون انصهارها في كيان موحد، وتارة أخرى كمصدر لطاقة ثقافية كامنة يمكن أن تشري الحضارة الأوروبية بل حضارة العالم أجمع.

أما أهمية اللغة اقتصاديا فيكفي هنا أن نشير إلى ما سبق أن ذكرناه بشأن الميزة النسبية الهائلة التي تتمتع بها الولايات المتحدة في تنافسها الشديد مع اليابان للسيطرة على سوق البرمجيات العالمي، وذلك نظرا لشيوع اللغة الإنجليزية وبساطتها وانغلاق اللغة اليابانية وصعوبتها، ولا يخفى على أحد ما للغة من أهمية في صناعة البرمجيات خاصة بعد دخولها في مجال الإنسانيات (انظر الفقرة 4: 2: 3 من الفصل الرابع).

أما أهميتها التكنولوجية فلا تتبع فقط من علاقة اللغة بتكنولوجيا الطباعة والاتصالات والبرمجيات بل أيضا من الدور الخطير الذي تلعبه اللغة حاليا في تشوير معمارية الكمبيوتر إلى درجة اعتبار كمبيوتر الجيل الخامس حاسبا لغويا في المقام الأول، حيث الهدف منه-كما خلص البعض- هو كسر حاجز العزلة اللغوية التي تعاني منها اليابان (141: 5) أملا في السيطرة على سوق المعلوماتية العالمي والذي يعد فيه تعامل تكنولوجيا المعلومات مع لغات العالم المتعددة عاملا حاسما في تحقيق هذه السيطرة. على الرغم من كل هذه الأهمية لم تحظ اللغة بقدر الاهتمام الجديرة به، هل يمكن أن يرجع ذلك إلى كونها لصيقة بنا لا نستطيع أن ننسلخ عنها لنتمتعنا عن بعد؟ أو لكونها شاغل الجميع إلى حد أن باتت شريفة يتنازع عليها علماء النفس وعلماء الفسيولوجي، وعلماء الاجتماع، وعلماء التربية، ومن قبلهم الفلاسفة ومن بعدهم علماء الكمبيوتر. ويزعم هذا الفريق الأخير بأنه لا حل لمعضلة اللغة دون اللجوء إلى أساليب الذكاء الاصطناعي وهندسة المعرفة، في الوقت نفسه الذي يرددون فيه أنه لا أمل في ارتقاء حقيقي لآلاتهم الحاسبة ما لم تستطع تلك الآلات محاكاة وظائف الذهن اللغوية.

ولا شك أن وراء تأخر المعالجة العلمية الجادة للغة، وعدم دخولها مصاف العلوم المنضبطة إلا أخيراً، سببا رئيسيا، وكافيا هو: تعقدها، وما أعقدها تلك اللغة من ظاهرة، اللغة التي تبين رغم لبسها وغموضها، وتلتبس وتغمض مرتدية أقتعة السفور، وتطنب لتقتص من المعني وتقتضب لإبانته، اللغة التي قال عنها الجرجاني في أسرار بلاغته: «ترك الذكر أفصح من الذكر، والصمت عن الإفادة أزيد للإفاد»، والتي قال عن ألفاظها ومعانيها صاحب البيان والتبيين: «إن الألفاظ دائما ليست على قياس المعاني، وللمعاني أقدار ينبغي أن يدركها ويعرفها الإنسان فهي حسب أقدار المستمعين ومستوياتهم الفكرية». مصدر الإشكالية إذن، في كيفية التصدي مرونة اللغة واتساع مفرداتها، وتعدد معانيها، وتباين استخداماتها، وفي كيفية التعامل مع لبسها، وغموضها وحذفها وتفشي مجازها، والكشف عن مضامين كنايتها وبنية جملها ونصوصها.

لقد فجرت تكنولوجيا المعلومات إشكالية اللغة كما لم يحدث لها من قبل، بعد أن أظهرت المواجهة بينهما الحاجة الماسة إلى المراجعة الشاملة للمنظومة اللغوية ككل، وذلك حتى تنهياً للغة للقاء هذه الآلة المثيرة المتحدية، هذا على جبهة اللغة، أما على جبهة الكمبيوتر فقد كان لزاما عليه أن يتخلص من معماريته التقليدية، آلة فون نيومان، ذات الطابع المركزي الصارم (انظر الفقرة 3: 2: 2 من الفصل الثالث)، وذلك حتى يتأهل هو الآخر للقاءه الحاسم مع اللغة، وهكذا وجد علماء اللغة وعلماء الكمبيوتر أنفسهم أمام تحد حقيقي، مصدره أن ما هو متاح حتى الآن من علم وتقنية لا يكفي لمواجهة إشكالية «اللغة-الكمبيوتر»، ولا بديل عن استخدام مناهج مبتكرة وشق دروب علمية جديدة في مناطق لم يتطرق لها العالم من قبل، وهو الأمر الذي أدى إلى إنشاء مراكز بحثية متخصصة في علاقة اللغة بتكنولوجيا المعلومات في الولايات المتحدة والاتحاد السوفييتي وألمانيا ومناطق أخرى من العالم.

ومع ظهور الحاجة للمراجعة الشاملة عادت الأسئلة ذات الطابع الفلسفي تطرح نفسها من جديد، هل اللغة ظاهرة عشوائية اعتباطية أو تحت تجلياتها الظاهرة هناك نظام متسق تحكمه القواعد والمبادئ؟ هل ندرس سلوكها الظاهر المحسوس أو نحاول استجلاء المعرفة اللاواعية التي تحكم آليات

نطقنا وفهمنا؟ هل اللغة سلسلة من الرموز الخطية، أو شبكة من العلاقات المتداخلة، أو هيكلية من مستويات متوازية متتالية؟ وما الوحدة اللغوية الأساسية، التي يرتكن إليها الدرس اللغوي: أهي اللفظ أو الجملة أو السياق؟ وهل المدخل لدراسة بنيتها الداخلية مباني تراكيبها، وأنماط نظمها، أو دلالة معانيها وتجاوز استعارتها، وما تبطنه كنايتها؟ ودعنا نواصل الأسئلة، وما أكثرها: هل ندرس اللغة المنطوقة أو مآثرها المكتوبة؟ وهل شاعنا هو اللغة كما يجب أن تكون، أو اللغة كما تمارس في واقع الحياة اليومية؟

أمام هذه الحيرة ذهب البعض إلى أن إخضاع اللغة لضوابط العلم المنضبط ما هو إلا وهم خادع، فمن أين للرياضيات والمنطق والإحصاء أن تحاصر هذا الكم الهائل من ظواهرها المعقدة؟ وكيف يمكن لهذه الآلة الصماء أن تحاكي ملكة اللغة بمرونتها ومرادفاتها وشحنة الانفعالات الكامنة وراء تعابيرها؟ في مقابل هذا الرأي، هناك من يؤكد إمكان تحقيق ذلك شريطة أن تتوافر لدينا وسائل علمية جديدة، منطلق أرقى من منطلق الدرجة الأولى لأرسطو، ورياضيات حديثة، وإحصاء متقدم، وفوق ذلك معالجة آلية مغايرة لأساليب البرمجة التقليدية، وهكذا ظهر إلى الوجود علم اللسانيات الحاسوبية computational linguistics، وهندسة اللغة language engineering، وقد صاحبهما ثورة علمية حقيقية في معظم فروع اللسانيات مازال صداها يتردد إلى الآن، وربما لفترة قادمة ستطول، لقد أيقن الجميع أن دخول اللغة مضاف العلوم المضبوطة، شرط أساسي لكي تتبعها في ذلك علوم الاجتماع والأدب والنقد، بل وعلم استرجاع المعلومات information retrieval الذي مازال قيد البرمجيات الهندسية العملية.

سنسعى في هذا الفصل إلى تناول خصائص منظومة اللغة العربية ومظاهر أزمتهما الراهنة، وذلك من منظور معلوماتي، بعدها نتناول مراحل تطور دراسة اللغة، لتنتقل بعد ذلك إلى استخدام المعلوماتية كأداة للغة العربية، ويعتبر هذا الفصل تلخيصا لكثير من الأفكار التي سبق للكاتب طرحها في كتابه عن «اللغة العربية والحاسوب» (45).

9: 2 خصائص منظومة اللغة العربية من منظور معلوماتي

اللغة العربية أعقد اللغات «السامية» وأغناها صوتا وصرفا ومعجما،

وقد تحدث عنها كثيرون، عرب ومستشرقون، بعضهم يمجّد عبقريتها وسحرها، والبعض الآخر يرميها بالقصور والعجز عن ملاحقة التطور والصعوبة التي لا مبرر لها، والتي تستدعي ضرورة تبسيط كتابتها وصرفها ونحوها. سنحاول هنا أن نتجاوز حدود الأحكام القيمية مركزين الحديث عن خصائص اللغة العربية من منظور معلوماتي طارحين جانباً تلك الخصائص التي اعتبرناها ليست ذات أهمية مباشرة في سياقنا الراهن، من أمثال ما قيل عن قدم العربية وعراقتها، ووحدتها وتواصلها وثبات أصوات الحروف فيها، وأنها ليست لغة قبيلة أو جماعة بعينها وإنما هي لغة الشعب العربي كله (2: 959).

في رأي الكاتب أن أهم خصائص منظومة «اللغة العربية» هي:

- التوسط اللغوي.
- حدة الخاصية الصرفية.
- المرونة النحوية.
- الانتظام الصوتي.
- ظاهرة الإعراب.
- الحساسية السياقية.
- تعدد طرق الكتابة وغياب عناصر التشكيل.
- ثراء المعجم واعتماده على الجذور.
- شدة التماسك بين عناصر المنظومة اللغوية.

ولا يتسع المجال هنا إلى عرض مفصل لهذه الخصائص وسنكتفي هنا بتناول كل منها في إيجاز شديد، نشرح في البداية المقصود بالخاصية اللغوية المعينة لتعيد طرحها بعد ذلك من منظور معلوماتي.

أ- التوسط اللغوي

المقصود بالخاصية: من وجهة نظر طوبوغرافيا اللغات: تتسم العربية بتوازن في معظم خصائصها اللغوية يضعها في منطقة الوسط-(وكذلك جعلناكم أمة وسطاً) صدق الله العظيم-بين أطراف كثير من المحاور التي تحدد مجالات التنوع اللغوي، فهي تتحاز نحو الشائع اللغوي وتكره الشاذ والشارد وتجمع بين كثير من الخصائص اللغوية المشتركة مع لغات أخرى وسنكتفي هنا ببعض الأمثلة:

- ليست أبجدية الكتابة العربية فونيمية صرفة كـ«الإسبانية» و«الفلندية»، حيث يناظر كل حرف فونيميا (صوتا) واحدا، كما أنها ليست مقطعية، syllabic كـ«اليابانية» حيث رموز الأبجدية عبارة عن مقاطع يتكون كل منها من صامت consonant يتبعه حركة (صائت vowel) مثل «ما»، «كي» و«فو»، فالأبجدية العربية رغم كونها فونيمية أساسا، فإنها تتضمن حروفا ذات طبيعة مقطعية وهي: لا، لأ، لاء، لأ، إ، آ، ؤ، ث..

- بالنسبة لحالات الإعراب تمثل العربية (ولها 3 حالات للإعراب) موضعا وسطا بين قصور الحالات الإعرابية كما في الإنجليزية وكثرتها النسبية كما في الروسية (6 حالات إعراب).

- وبالنسبة لترتيب الكلمات داخل الجمل يجمع نحو العربية بين الجملة الإسمية والجملة الفعلية.

- وبالنسبة لاستخدام الأسماء الموصولة (الذي، اللتين،)، تصل العربية المعرف definite، ولا تصل النكرة indefinite، بينما تصل اللغات الجرمانية كالإنجليزية المعرفة والنكرة، في حين تشرّد الصينية فتسقط ظاهرة «الصلة» تماما.

- تطابق العربية بين الفعل والفاعل وهي في ذلك وسط بين الإنجليزية الخالية من هذا التطابق-باستثناء حالة الفعل المضارع مع الفاعل المفرد الغائب (he goes)، واللغات التي تطابق بين الفعل وفاعله ومفعوله كما في لغة أهل «جورجيا».

- وفيما يخص صيغ الأفعال المزيدة تعد العربية (15 صيغة مزيدة) وسطا بين اللغات ذات صيغ الأفعال المحددة كالإنجليزية، والأخرى ذات العديد منها كالإسبانية (يتجاوز عدد صيغ أفعالها ثلاثين صيغة).

المنظور المعلوماتي: يفرض التوسط اللغوي ضرورة الاهتمام بالدراسات اللغوية المقارنة comparative والتقابلية contrastive حيث تمثل مصدرا غنيا لتفهم خصائص لغتنا في إطار أشمل وأعمق، وقد أدهشني أن تزعم كثير من الدراسات التي تناولت خصائص اللغة العربية انفرادها بخصائص لا تحتكرها مثل ما يقوله مهدي المخزومي «للعربية سمة تميزها عن اللغات الأخرى تلك هي أن الكلمة في أثناء الجملة تحمل معها ما يدل على صفتها الإعرابية»، وما زعمه المستشرق براغستراسر في انفرادها بضمير الشأن

وما زعمه الكرملبي بقوله «مما وسع كلام الناطقين بالضاد توسعا لا يقابله شيء في سائر اللغات المعروفة ما وقع فيها من القلب والإبدال والتصحييف والتعريف» (55)، وقد أثبتت الدراسات الحديثة في طوبوغرافيا اللغات أن جميع هذه الخصائص ليست حكرا على العربية. ألا يظهر ذلك شدة حاجتنا لبنك معلومات لدعم البحث اللغوي الحديث ٥.

علاوة على ذلك يعني توسط اللغة العربية أن أمور معالجتها آليا ستتعرض للعديد من القضاب حيث على النظم الآلية أن تتعامل مع الظواهر سواء كثرت أو قلت، شاعت أو ندرت مادامت قد تعرضت لها اللغة العربية، في الوقت نفسه يمكن لخبراء العرب الاستفادة من الحصاد الهائل المتوافر حاليا من نظم معالجة اللغات الطبيعية لينتقوا ويوفقوا منه ما يهم شؤون لغتنا «الوسطى».

ب - حدة الخاصية الصرفية

المقصود بالخاصية: تتسم اللغات السامية بخاصية الاشتقاق الصرفي المبني على، أنماط الصيغ، لا يباري اللغة العربية أي لغة أخرى، سامية أو غير سامية، في حدة خاصيتها الاشتقاقية، وربما يعد هذا هو المظهر الوحيد لتطرفها اللغوي، إذ إنها تتميز بالاطراد الصرفي شبه المنتظم الذي أدى بالبعض إلى وصفها بالجبرية (نسبة إلى علم الجبر)، بدرجة تقترب من حد الاصطناع (147: 4). وبجانب اطراد الاشتقاق هناك أيضا ظاهرة التعدد الصرفي، كتعدد صيغ الجمع (مثال: كاتبون، كتبة، كُتاب).

المنظور المعلوماتي: إن انتظام الصرف العربي واطراده يزيد من قابليته للمعالجة الآلية computationality، وتعد هذه المعالجة هي المدخل الطبيعي لمعالجة المنظومة الشاملة للغة العربية، حيث يعد معالج الصرف الآلي morphological processor مقوما أساسيا في ميكنة المعجم العربي، وتطوير نظم آلية للإعراب الآلي والتشكيل التلقائي، كما سنوضح في الفقرة 9: 6: 4 من هذا الفصل.

ج- المرونة النحوية

المقصود بالخاصية: تقصد بالمرونة النحوية هنا تلك الحرية النسبية التي نلاحظها في ترتيب الكلمات داخل الجمل العربية، تظهر هذه المرونة في توسعات تراكيب الجمل بفعل آليات التحويل النحوي المختلفة، كالتقديم

والتأخير (مثال: إذا بلغ الفطام لنا رضيع. والأصل فيها إذا بلغ رضيع لنا الفطام)، والحذف والإبدال النحوي (من أمثلة ذلك استخدام اسم الفاعل بدلاً من الفعل «ضاربا أخاه» والمصدر بدلاً من الظرف «استيقظت طلوع الشمس» وما شابه).

المنظور المعلوماتي: تمثل هذه المرونة تحدياً حقيقياً للتظهير العربي ومعالجة النحو العربي آلياً، مصدر هذه الصعوبة أن النماذج اللغوية المتاحة لصياغة قواعد النحو لأغراض المعالجة الآلية قد صممت أصلاً لتلائم مطالب اللغة الإنجليزية التي تتسم بالصرامة النسبية لرتبة الكلمات strict word order داخل جملها، يعني ذلك احتياجنا إلى بحوث أساسية، وتطبيقية لكتابة قواعد النحو الصوري formal grammar اللازم لمعالجة النحو العربي آلياً، حيث يجب أن يراعى في كتابته جميع البدائل الممكنة لأنماط الجملة العربية بفعل عمليات التقديم والتأخير، والحذف والإبدال والإضمار. ولتوضيح الفرق يحتاج نحو الإنجليزية إلى ما يقرب من 1000 قاعدة رياضية في حين وصل عدد القواعد لنحو اللغة العربية غير المشكولة الذي قام بصياغته الكاتب إلى ما يزيد على 12 ألف قاعدة.

د- الانتظام الصوتي

المقصود بالخاصية: تتميز القواعد الصوتية للعربية بالاطراد، ويتسم نظام مقاطعها الصوتية ونبرها بالبساطة، إذ إن جميع هذه المقاطع لا بد وأن تبدأ بحرف صامت (غير متحرك consonant)، ولا تتضمن أكثر من صامتين، وتحديد مواضع النبر في الكلمة العربية يتوقف على سلسلة حروف الكلمة دون اعتماده على عوامل خارجية، مثل اعتماد موضع النبر في بعض الكلمات الإنجليزية على قسم الكلم part-of-speech (مثال: الفرق في نطق progress في حالة الفعل وحالة المصدر).

المنظور المعلوماتي: تعد هذه الخاصية ذات أهمية كبيرة في توليد الكلام العربي speech synthesis، وتمييزه speech recognition آلياً، حيث يسهل «تطريز» أنماط النبر في الكلمات المولدة آلياً بحيث يبدو الصوت المولد طبيعياً لا ميكانيكياً robotic، في حين يمكن استغلال انتظام قواعد النبر في تمييز الكلمات المنطوقة آلياً حيث يسهل نسبياً من عملية التعرف على مواضعه، والتي تساعد بدورها على التعرف على بنية الكلمة المنطوقة.

هـ- الحساسية السياقية

المقصود بالخاصية: يقصد بالحساسية السياقية contextual sensitivity تأخي العناصر اللغوية مع ما يحيطها، أو يرد معها، من عناصر، تتسم اللغة العربية بحساسية سياقية عالية مقارنة باللغة الإنجليزية مثلا، ويمكن اقتفاء مظاهر هذه الحساسية على مستويات متعددة، فعلى مستوى الكتابة يتوقف شكل الحرف العربي على الحرفين السابق واللاحق له (لاحظ تغير شكل العين في مرعى، مزارع، معدة، مصنع)، وعلى مستوى النحو تتبدى مظاهر هذه الحساسية في علاقات المطابقة agreement، كتطابق الصفة مع الموصوف، وتطابق الفعل والفاعل والمبتدأ مع خبره المشتق، وإعراب الفعل مع أداة النفي السابقة له (لم يقل، لن يقول).

المنظور المعلوماتي: تمثل هذه الحساسية السياقية أحد مواضع الاختلاف الرئيسية بين اللغة الإنجليزية واللغة العربية، وهو ما يجعل معالجة اللغة العربية أليا أعقد، وربما بكثير، من معالجة اللغة الإنجليزية، والأمثلة على ذلك كثيرة، نذكر منها أن تعامل الكمبيوتر مع نسق الكتابة الإنجليزية ذات الحروف المنفصلة، يعد بسيطا إذا ما قورن بذلك للكتابة العربية ذات التشبيك والتشكيل، فيكفي أن الأخيرة تحتاج-خلافًا عن الإنجليزية- إلى وسيلة برمجية آلية لانتقاء شكل الحرف المناسب تلقائيا وفقا للحرفين المحيطين به، ومثالنا الآخر يتعلق بمعالجة النحو العربي أليا حيث نجم عن الحساسية السياقية صعوبات جمة في أسلوب كتابة النحو الصوري formal programmer الذي سبق الإشارة إليه هنا، والذي تعمل على أساسه نظم الإعراب الآلي.

و- تعدد طرق الكتابة وغياب عناصر التشكيل

المقصود بالخاصية: فيما يخص عناصر التشكيل يمكن كتابة العربية

بثلاث طرق:

- كتابة تامة التشكيل.

- كتابة مشكولة جزئيا.

- كتابة خالية من التشكيل.

لقد نشأت الكتابة العربية أصلا دون تشكيل، وقد شاع التغاضي عن استخدام حركات التشكيل في «العربي» الحديثة إلى أن أصبح عدم التشكيل

عادة مترسخة في قراءة العربية وكتابتها، بل وأكد أزعم في سك مصطلحاتها الجديدة أيضا حيث يراعي واضعو المصطلح الحديث بقدر الإمكان عدم التباس قراءته خاليا من التشكيل، مثال ذلك رفض مصطلح حساب بالسين المشددة كمكافئ للكمبيوتر نظرا لاختلاطه مع علم الحساب. وكل الدعاوى المنادية بضرورة الالتزام بالتشكيل هي نوع من التفاؤل غير الواقعي، وأميل إلى القول بأنها تغفل أحد الأسس العميقة التي ترتكز عليها كتابة العربية ونطقها، ولم أجد ما يقنعني بما خلص إليه عباس حسن من «أن ترك الشكل في الكتابة العربية ليس أصلا من أصولها، ولا ضرورة محتومة فيها، بل ربما كان العكس هو الصحيح، كما ينطق بذلك الخط العربي» (14)، إن التشكيل على أهميته سيظل في رأيي مجرد أداة تربوية لتعليم صغارنا مبادئ نطق العربية وكتابتها.

المنظور المعلوماتي: يؤدي إسقاط علامات التشكيل إلى ظهور حالات معقدة من اللبس حيث يمكن أن يمثل هيكل skeleton الكلمات الخالية من عناصر التشكيل عدة قراءات محتملة له (ولنأخذ مثلا كلمة «وجد» وأستحث القارئ على تمثل قراءاتها الممكنة مرة باعتبار الواو من أصل الكلمة ومرة باعتبارها أداة عطف)، حتى أبسط الكلمات مثل «لن»، و«أنتم» التي تبدو من أول وهلة خالية من اللبس تلتبس هي الأخرى مع غياب التشكيل فيمكن بجانب قراءتها الدارجة أن تكون صيغ فعل الأمر أو الجزم من «لان» و«أنمي». تتفاعل حالات اللبس الناجمة عن غياب التشكيل مع حالات اللبس الأخرى التي تشترك فيها العربية مع باقي اللغات، كاللبس المعجمي في كلمة «عين» (بمعنى البئر أو الجاسوس أو الرأس أو ذات الشيء) واللبس التركيبي في شبه جملة شاعر النيل العظيم (باحتمال كون العظيم صفة للنيل أو شاعره)، يؤدي ذلك إلى أنواع معقدة للغاية من اللبس المركب متعدد المستويات multi-level ambiguity.

تمثل مشكلة غياب التشكيل عقبة أساسية أمام معالجة النصوص العربية آليا، ولتخيل القارئ معي مدى الإحباط الذي يشعر به الباحث إن أراد أن يبحث عن كلمة «علم بمعنى مفرد علوم» من نصوص مخزنة على الكمبيوتر ليظهر لنا نظام استرجاع المعلومات، جميع حالاتها الاسمية والمصدرية والفعلية كعلم بمعنى «راية»، وبمعنى الفعل في «علمه الدرس»، وكذلك

الفعل والمصدر في «علم الخبر»، ويستحيل بالطبع أن نطالب المستخدم، حلا لهذه المشكلة، بإدخال بياناته كاملة التشكيل، حيث لا يتطلب ذلك جهدا مضاعفا لإدخالها من خلال لوحات المفاتيح فقط بل جهدا مضنيا للغاية لا يقدر عليه كثيرون لضبط تشكيلها يدويا.

إن معالجة النصوص العربية آليا تحتاج إلى وسيلة برمجة آلية لتشكيل النصوص تلقائيا، وقد أسفر البحث الذي قام به الكاتب على مدى خمس سنوات متواصلة منذ 1988، عن تطوير نظام آلي لإعراب النصوص العربية automatic parsing، وكذلك تشكيلها آليا automatic diacritization.

ز- ثراء المعجم واعتماده على الجذر

المقصود بالخاصية: يرتبط تنظيم المعجم في أي لغة بصلات وثيقة مع طبيعة عمليات تكوين الكلمات بها، لذا فقد كان من الطبيعي أن يأتي تنظيم معجمنا العربي على أساس الجذور، أصل تكوين الكلمات، لا على الترتيب الأبجدي كما في «الإنجليزية» مثلا. تتميز شجرة المفردات العربية، شأنها في ذلك شأن جميع اللغات السامية، بقلّة الجذور وتعدد الأوراق تعددا مضاعفا، لذا فهي توصف بأنها شجرة «ثقيلة القاع»، فرغم صغر نواة المعجم العربي (أقل من عشرة آلاف جذر)، تتعدد المفردات بصورة هائلة وذلك بفعل خاصية الاشتقاق أو «الإنتاجية الصرفية» العالية بلغة الصرفيين. فمن أساس الاشتقاق للجذر النمطي «ف.ع.ل» يمكن «إنتاج» 15 صيغة من مزيادات الأفعال تتيح كل صيغة منها صيغا مطردة وغير مطردة للمصادر وأسماء الأفعال والمفعول وأسماء المكان والزمان وصيغ التفضيل والمبالغة وصيغ الجموع، علاوة على ذلك تتميز العربية بثراء مفرداتها وكثرة مترادفاتها.

المنظور المعلوماتي: مما سبق يتضح أن المعجم العربي ليس مجرد قائمة مفردات، كما يجوز القول بالنسبة للإنجليزية، بل بنية معقدة من العلاقات التي تربط بين مشتقات الجذور، وصيغ الأفراد والجمع والمترادفات وما شابه، ناهيك عن علاقات الاشتقاق الأعظم كما أورده ابن جني في «الخصائص» ليمرر ظاهرة اتصال المعنى وإن اختلف ترتيب حروف الأصل (مثال ذلك ك. ل. م فهي حيث تقلبت تدل على القوة والشدة)، وأكاد أزعم أن فهم بنية المعجم العربي المعقدة مازالت دون تناول معظم متخصصينا

وباحثينا في ظل القيود التي تفرضها عليهم الوسائل اليدوية التقليدية لبناء المعاجم وتحديثها واستخدامها، واعتقادي الراسخ أنه لا أمل في سبر أغوار بنية المعجم العربي إلا باللجوء إلى نظم المعلومات باستخدام الكمبيوتر، ولم تعد ميكنة المعجم أمرا من قبيل الرفاهية الفنية بل مطلباً أساسياً نابعا من طبيعة معجمنا، وضرورة تحديثه، وإعادة تنظيمه بما يليبي مطالب البشر ومطالب النظم الآلية على حد سواء.

ح - شدة التماسك بين عناصر منظومة «العربية» المقصود بالخاصية: تتميز منظومة اللغة العربية بشدة التماسك بين عناصرها، ومظاهر ذلك عديدة نكتفي هنا ببعض منها:

- التداخل الشديد بين منظومتي الصرف والمعجم، إذ انعكست الخاصية الاشتقاقية بشكل واضح على تنظيم المعجم وأسلوب استخدامه.

- العلاقة العضوية بين نحو العربية وصرفها، فهما توأمان يصعب الفصل بينهما.

- التداخل الشديد بين منظومتي الصرف (المورفولوجي) والصوتيات (الفونولوجي)، والذي يبدو واضحا في الدور الذي يلعبه الإبدال والإعلال (العمليات المورفولوجية) في تحديد بنية الكلمة العربية.

- شدة الصلة بين مباني الصيغ الصرفية ومعانيها (كارتباط صيغة «انفعل» بمعنى المطاوعة وصيغة «تفاعل» بمعنى المبادلة)، وارتباط النحو العربي بالمنطق، وفي رأيي أنه لا سند للرأي القائل إن هناك ثمة انفصالا لفظ عن المعنى في العربية، فهو أمر يتناقض في جوهره مع وظيفة اللغة عموما، ناهيك عن لغتنا العربية ذات القدرة العالية على الإبانة والفصاحة، والتعبير بموجز اللفظ عن عمق المعاني، ويجلي لبسها وغموضها رغم كتابتها المختزلة غير المشكلة التي تسقط الحركات القصار، وربما تكون المشكلة في محاولة فهمنا ظاهرة الصلة بين ألفاظ العربية ومعانيها من منظور أرسطي عفي عليه الزمن، ولا يتسع المقام هنا إلى تفنيد هذا الزعم في ضوء ما أورده الجرجاني في أسرار البلاغة ودلائل الإعجاز، أو ما خرج به عالم الدلالة الحديث سواء الدلالة المعجمية lexical semantics، أو الدلالة الصورية formal semantics من أسس لتناول علاقة الألفاظ بمعانيها، وسأكتفي هنا بما قاله ابن رشيق القيرواني سنة 456 للهجرة: «اللفظ جسم روحه

المعنى، وارتباطه به كارتباط الروح بالجسم»، وهو يقترب بشدة إلى ما خلص إليه دي-سوسير بعد تسعة قرون من صعوبة فصل الرمز عن معناه، فهما على حد تعبيره الشهير كصفحتي الورقة الواحدة.

المنظور المعلوماتي: يمثل هذا التماسك لمنظومة اللغة العربية سلاحا ذا حدين، فهو من جانب يزيد من صعوبة النظم الآلية حيث يصعب فصل المعالجات الآلية للنحو الآلي عن تلك الخاصة بمعالجة صرفها ومعجمها، ويتطلب ذلك حاسبات ذات إمكانيات وقدرة أكبر لمعالجة اللغة العربية آليا (انظر الفقرة 3:2:4 من الفصل الثالث)، ومن جانب آخر يعمل هذا التماسك لصالح نظم الفهم الأتوماتي وإجلاء اللبس الناجم عن غياب التشكيل حيث تعتمد هذه النظم على مبدأ تضافر القرائن اللغوية على اختلاف مصادرها.

9: 3 مراحل تطور دراسة اللغة

رأينا أن نوجز هنا مراحل تطور دراسة اللغة لغرضين أساسيين:

- إبراز شدة التخلف الذي يعاني منه حاليا التنظير للغة العربية الذي يكاد يكون قد توقف منذ القرن الخامس الهجري تقريبا، ومما يثير الحسرة أن كثيرا من أسس اللسانيات الحديثة لها أصولها الراسخة في تراثنا اللغوي الذي خلفه لنا سلفنا العظيم.

- إبراز الدور الذي لعبته وتلعبه تكنولوجيا المعلومات التي دفعت بالدرس اللغوي إلى آفاق جديدة.

يمكن تلخيص مراحل تطور دراسة اللغة في عدة مراحل هي:

أ- مرحلة ضبط السلوك اللغوي: الأساس هنا هو النظر إلى اللغة بصفتها سلوكا اجتماعيا لا بد من إخضاعه لـ «تشريع» لغوي، فالنحو كما عرفه «صبح الأعشى» هو «ميزان العربية» والقانون الذي تحكم به في كل صورة من صورها، والنحو قياس يتبع والاطراد مناط القياس، وهدف النحو هو منع اللحن وحفظ النص القرآني كتابة وتلاوة، لقد استتبع ذلك تزايد نفوذ النحاة حتى ساغ أن يطلق على هذه المرحلة عصر «ديكتاتورية النحاة».

ب- مرحلة دراسة التباين اللغوي: سادت نظرية التطور لداروين الفكر العلمي طوال القرنين الأخيرين-على الأقل-من القرن التاسع عشر والعقود الثلاثة الأولى من القرن العشرين، وقد صاحب ذلك اهتمام الدول المستعمرة

بدراسة لغات أهالي المستعمرات (127: 9)، وظهر علم فقه اللغة ليدرس أصل اللغات ويقارن بينها ويوصف ويصنف خصائصها وفصائلها، لقد ذهب إلى الأبد عصر الغيبيات اللغوية والأحكام القيمية لتقويم اللغات بين لغات راقية ولغات بدائية أو منحطة، وذلك بعد أن ضاع وهم الرقي الذي تصوره الأوروبيون عن لغاتهم إثر اكتشافهم أن اللغة الهندية القديمة (السنسكريتية) هي أصل اللغات الأوروبية التي أصبحت تعرف بعده باللغات الهندوأوروبية، بسقوط الغيبيات وتخلص الفكر اللغوي من نظراته الضيقة اتجه من دراسة لغة أو عدة لغات إلى دراسة ظاهرة اللغة ككل وتهيأت اللغة بذلك لدخول مرحلة الضبط العلمي.

ج- مرحلة وضع النواة التنظيرية: يرجع الفضل في إطلاق الشرارة الأولى لعلم اللغة المحض إلى «فرديناند دي-سوسير» الذي جعل من ثنائية الرمز والمدلول والورقة بينهما بمثابة الإشكالية المحورية للمدرس اللغوي، وفرق بين اللسانيات وبين مما ارتبط بها فيما سبق من علوم أخرى لدراسة تطور اللغات وتاريخها وتعليمها وتحليل النتائج اللغوي، ومن جانب آخر وأهم فقد فصل دي-سوسير بين نظام اللغة الذي يتحكم في أدائها وبين الكلام الفعلي أو السلوك اللغوي الذي يمارسه الناطقون بها.

د- مرحلة تحليل المركب اللغوي: ظهر الاتجاه التحليلي في الولايات المتحدة متأثراً بعلم الكيمياء في تحليل المركبات العضوية وغير العضوية إلى عناصرها الأولية، وهكذا تم تحليل الإشارة الكلامية إلى فونيمات (وحدة الصوت اللغوي)، والكلمات المركبة إلى صرفيمات أو مورفيمات (morphemes mismanagement mis+manage+ment)، ليصبح التعامل مع العنصر «الذري» اللغوي، وفهم طبيعة مركباته مدخل التحليل اللغوي، ورغم نجاح هذا المنحى التحليلي على صيد الصوتيات (الفونولوجي) وبدرجة أقل على صعيد الصرف (المورفولوجي) فقد ظهر قصوره في التعامل مع النحو بتركيباته اللانهاية. واكب ظهور الاتجاه التحليلي للمعطيات اللغوية الفعلية ظهور المدرسة السلوكية في الدرس اللغوي، وهي المدرسة التي تعتمد أساساً على ملاحظة الظواهر (المؤثرات وردود الأفعال) وتحليل المعطيات دون غيرها وترفض اللجوء إلى المبادئ الغائية، أو النماذج الذهنية كأساس لتفسير ظواهر النشاط اللغوي.

د - وضع النموذج الإحصائي للغة: كعادته، يمهّد الإحصاء الأرض للحرث العلمي الدقيق، وقد ظهرت بوادر الإحصاء اللغوي في نهاية القرن الماضي عندما استخدم لأغراض ذات طابع عملي أكثر منه نظرياً، مثل تحقيق التراث، والتحليل الكمي لأساليب الأدباء والشعراء، وبعض الأمور المتعلقة بالفهرسة وتنظيم المعاجم وتصميم شفرات الاتصال للأغراض العسكرية، تم وضع النموذج الإحصائي للغة من خلال نظرية المعلومات د «شانون» الأمريكي (انظر الفقرة 2: 3 من الفصل الثاني) والسلاسل الإحصائية time series لـ «ماركوف» الروسي، وصار بالإمكان قياس كمية المعلومات التي تتضمنها سلاسل الرموز اللغوية، كما دان الفائض اللغوي linguistic redundancy للتحليل الكمي مما أبرز أهمية هذه الظاهرة اللغوية الأصيلة وتفضيها في منظومة اللغة. وتصور البعض خطأ أن الإحصاء هو الحل، وأن اللغة قد خضعت أخيراً لسطوة التحليل الكمي، ولكنهم سرعان ما اكتشفوا خطأ تصورهم عندما أدركوا أن نظرية المعلومات تنظر إلى اللغة نظرتها إلى الإشارات الهندسية، نظرة تختزلها إلى إشارة خالية من المعنى لا تأخذ من اللغة إلا تجلياتها السطحية دون النفاذ إلى معانيها وبنيتها الداخلية والياتها الدفينة المولدة لهذه التجليات.

هـ - بدء استخدام الكمبيوتر في مجال اللغة: صاحب ظهور الكمبيوتر في أواخر الأربعينيات، تهاوّل شديد عن استخداماته المحتملة، في مجالات التحليل اللغوي والترجمة الآلية، وكما هو متوقع باءت المحاولات الأولى بفشل ذريع، فكيف يمكن أن يتعامل الكمبيوتر في مراحلها الأولى، والذي تعمل آلياته على أساس من القطع والاطراد البحث، مع معضل اللغة بلبسها وترادفها ومجازها وفائضها، وذلك قبل أن تدين للتجريد الرياضي والصياغة النظرية الدقيقة، وهما بمثابة تصريح الدخول لـ «ساحة» المعالجة الآلية بواسطة الكمبيوتر.

و- وضع النموذج الرياضي للغة: وضع برتراند راسل الأسس الرياضية لنظرية صورية للغات الرمزية formal theory of language، وفي نهاية الخمسينيات ظهرت نظرية النحو التوليدي generative grammar على يد نعوم تشومسكي، لتضع أسس النموذج الرياضي للغات الإنسانية، وهو ما يعتبر بداية الانطلاق الحقيقي للسانيات الحديثة، وقد أقيمت النظرية

التوليدية على عدة ركائز أهمها:

- هدف الدرس اللغوي ليس هو تحليل عينات الأمثلة أو توصيف حالات الاطراد والشذوذ أو وضع قائمة بمعايير الحكم على صحة الجمل، بل الهدف هو وضع نحو توليدي رياضي صريح explicit للغة ما، بحيث يمكنه توليد جميع الجمل المسموح بها من قبل هذه اللغة بمجموعة من القواعد الرياضية، تماما كما تولد المعادلة العامة $أ س + ب ص + ج = صفر$ على سبيل المثال جميع الخطوط المستقيمة المحتملة. أو توليد جميع المتواليات العددية أو الهندسية الممكنة بصيغة رياضية تعبر عن الحالات اللانهائية لسلاسل هذه المتواليات.

- على حين كان شاغل اللسانيين في الماضي هو البحث عن التباين اللغوي-linguistic difference، ركز تشومسكي في بحثه عن العموم اللغوي (linguistic uni-versality)(127- 10)، وذلك بتركيزه على القاسم المشترك بين اللغات بغية الوصول إلى نحو عام يفسر ظواهر التشابه والتباين بينها، في الوقت نفسه والذي في استخدامه كأساس لتفسير ظاهرة اكتساب الأطفال للغتهم الأم بصورة تلقائية.

- دراسة المعرفة اللغوية اللاواعية في الذهن البشري، فقد تبنى تشومسكي النموذج الذهني Mentalistic model للغة الذي يفترض أن الأطفال يولدون بغريزة لغوية عامة يقومون بتنميتها وتوجيهها من خلال تفاعلهم مع بيئتهم اللغوية الخالصة، إلى المطالب المعينة للغتهم الأم. ويمثل ذلك اختلافا جوهريا عن نظرة السلوكيين التي كانت سائدة عندئذ والتي ركزت على مظاهر السلوك اللغوي المحسوسة مفترضة أن الإنسان يولد ومخه صفحة بيضاء، وأن اكتساب اللغة يتم من خلال الممارسة الفعلية واكتساب العادات وترسيخها.

وبقدر ما حظيت به نظريات تشومسكي من تأييد وترحيب، بعد أن وجد فيها اللسانيون تميزهم العلمي، واستقلالهم الفكري عن مناهج العلوم الأخرى التي فرضت على اللغة في الماضي نوعا من الرعاية العلمية، فقد واجهت نقدا شديدا من قبل فلاسفة اللغة، وعلماء النفس وعلماء الكمبيوتر أيضا، فلاسفة اللغة يرمونها بعدم الواقعية بعد أن أهملت الوظيفة الاتصالية للغة، وأبعادها الاجتماعية بالتالي، وركزت على نظام اللغة في صورتها

السليمة المفروضة، لا تلك التي تمارس في الحياة العملية بكل ما يعترها من انفعالات وزلل ومؤثرات خارجية، أما النفسيون السلوكيون فيرون في لجوء تشومسكي لافتراض بنى معرفية مجردة وملكة لغوية معينة نولد بها انتكاسة إلى الغائية، وقصورا عما يمكن أن يتحقق من خلال التحليل الدقيق لسلوك اللغوي المحسوس، أما علماء الكمبيوتر فيرونها مسرفة في تجريداتها وعمومياتها بقدر يصعب معه استغلالها عمليا كأساس لتطوير نظم واقعية لمعالجة اللغة آليا.

ز- وضع أسس علوم الدلالة والمعرفة: كما كان التركيب معضلة «المدرسة التحليلية» كانت معضلة المدرسة التوليدية هي-بلا شك-مسألتي الصرف (خاصة جانبه الاشتقائي) والدلالة، وقد أيقن الجميع منذ البداية أن إخضاع اللغة للتمثيل المنطقي بهدف استخلاص العلاقات الدلالية يمثل تحديا هائلا للغويين والمناطقية على حد سواء. وبات واضحا حاجتنا إلى منطق جديد يتجاوز حدود ثنائية الصواب والخطأ لمنطق الأرسطي، إذ كيف يمكن أن يتعامل هذا المنطق القاطع مع العلاقات المعنوية لألفاظ مثل «أحيانا» «ومن المحتمل» و«بقدر كبير» و«في الأيام القليلة القادمة» و«على ما يبدو»، وليت الأمر يقف عند هذا الحد، فقد ظهرت استحالة توصيف عمليات التواصل اللغوي دون تمثيل الخلفية المعرفية المشتركة التي تربط بين المتكلم والمستمع، وهي الخلفية التي تشمل شركة المدركات والخبرات والمعتقدات والمسلمات والاصطلاحات عن العالم خارج نطاق اللغة، بجانب معرفة ظروف المقام التي يتم في ظلها الحدث اللغوي.

ح - مرحلة معالجة اللغة آليا بوساطة الكمبيوتر: بعد أن دانت اللغة للصياغة الإحصائية والرياضية وجزئيا للتحليل المنطقي أصبح الطريق ممهدا لدخولها مرحلة المعالجة الآلية تأكيدا لوصولها لمرحلة متقدمة من النضج العلمي، وهو ما سنتناوله بمزيد من التفصيل فيما بقي من هذا الفصل.

9: 4 العلاقة بين تكنولوجيا المعلومات واللغة العربية

9: 4: 1 عن حتمية الالتقاء وعوامله

منذ ظهور الكمبيوتر في أواخر الأربعينيات، وصلته باللغة تتوثق وتتأصل

في كلا الاتجاهين، وكان من المنطقي، بل من الحتمي أيضا، أن تلتقي اللغة بالكمبيوتر، وذلك لسبب أساسي وبسيط، هو كون اللغة تجسيدا لما هو جوهري في الإنسان، أي نشاطه الذهني بكل تجلياته، في الوقت نفسه الذي يتجه فيه الكمبيوتر نحو محاكاة وظائف الإنسان وقدراته الذهنية، لقد تدرج هذا الالتقاء حتى بلغ درجة عالية من التفاعل العلمي والتقني بصورة لا مثيل لها، ووراء ذلك عدة أسباب متباينة ومتعددة المصادر من أهمها في رأي الكاتب:

أ- التطور الهائل في علوم اللسانيات وخضوع كثير من جوانبها بشكل أو بآخر للمعالجة الرياضية والمنطقية والإحصائية.

ب - الوثبات العلمية التي تحققت في ميادين علوم الكمبيوتر بصفة عامة، وفي مجالات نظرية الأتوماتيات theory of outomata، وتصميم لغات البرمجة، وأساليب الذكاء الاصطناعي خاصة (انظر الفقرة 4: 3 من الفصل الرابع).

ج - التقدم المذهل الذي أحرزته تكنولوجيا المعلومات في مجال العتاد hardware والبرمجيات software، وما صاحب ذلك من تزايد الحاجة إلى لغات برمجة أرقى، لغات تتسم بالقوة والمرونة معا، وهو ما أدى إلى الاستهزاء بكثير من خصائص اللغات الإنسانية في تصميم هذه اللغات الاصطناعية. د- ظاهرة انفجار المعلومات مما يتطلب معها استحداث وسائل آلية ذات كفاءة عالية لتنظيم هذا الفيض المتزايد من المعلومات المتنوعة، وزيادة كفاءة تخزينها واسترجاعها وتوظيفها.

هـ - انتشار الحاسبات الشخصية والمنزلية وما استتبعه ذلك من ضرورة تسهيل التعامل مع الكمبيوتر بحيث يصبح في متناول الشخص العادي غير المتخصص، وبالتالي ضرورة التعامل بلغة طبيعية natural language أو بلغات برمجة شبيهة بها.

و- ظهور الحاسبات فائقة السرعة supercomputer مما مكن من تطوير نظم عملية لمعالجة اللغة آليا في ظل قيود الحيز والزمن والكلفة التي تفرضها الجدوى الاقتصادية والفنية لمثل هذه النظم (انظر الفقرة 3: 2: 4 من الفصل الثالث).

ز- بداية ظهور النظم الآلية الخبيرة، مثل تلك الخاصة بتشخيص

الأمراض، وتقديم الاستشارات الفنية والقانونية وبعض النظم الآلية للتعليم الذاتي (انظر الفقرة 4: 3: 7 من الفصل الرابع). وتتطلب هذه النظم قدرة على الحوار مع المستخدم البشري بلغة سهلة تقترب من لغته الطبيعية.

ح - انتشار الكمبيوتر كوسيلة للتعليم والتعلم بصفة عامة، وتعليم وتعلم اللغات بصفة خاصة.

ط - دخول تطبيقات الكمبيوتر مجالات علوم الإنسانيات، كالتاريخ وعلم الاجتماع والأدب والنقد، وجميعها ذات صلة وثيقة باللغة على خلاف التطبيقات الإدارية والتجارية.

وجملة الأسباب التي أدت إلى تقوية الصلة بين اللغة، وتكنولوجيا المعلومات يخصها التطور في ثلوث «اللغة-الكمبيوتر-التطبيق»، لقد تهيأت اللغة للمعالجة الآلية بدخولها مجالات التحليل الرياضي والمنطقي والإحصائي، وتهيأ الكمبيوتر من جانبه للقائه مع اللغة بالسرعة الفائقة وضخامة الذاكرة وضآلة الحجم وأساليب الذئ الاصطناعي، ولغات البرمجة الراقية، أما التطبيق فقد شق هو الآخر طريقه بنجاح إلى مجالات التعليم والإنسانيات والنظم الخبيرة.

هذا عن لقاء تكنولوجيا المعلومات مع اللغة بصفة عامة فماذا عن طبيعة هذا اللقاء مع اللغة العربية؟ وهو السؤال الذي سنحاول أن نجيب عنه في فقرتنا القادمة باختصار.

9: 4: 2 عن لقاء اللغة العربية بتكنولوجيا المعلومات

علينا أن نقر بداية بحقيقة مهمة هي أن هذا اللقاء ينطوي على مواجهة غير متكافئة بين اللغة العربية واللغة الإنجليزية، وذلك نظرا لطغيان الأساس الإنجليزي على توجهات تكنولوجيا المعلومات، فقد حددت هذه التوجهات أصلا لتبلي المطالب الخاصة باللغة الإنجليزية ودعنا نلخص فيما يلي بعض مظاهر هذه الظاهرة الطاغية:

- تصميم معظم لغات البرمجة باللغة الإنجليزية.
- استخدام شفرات لتبادل البيانات مصممة أصلا للتعامل مع الأبجدية الإنجليزية المحدودة في عدد حروفها وأشكال هذه الحروف.
- تصميم أساليب نظم تخزين المعلومات واسترجاعها على أساس أن اللغة الإنجليزية هي لغة الهدف.

- القسم الأكبر من مكتبة البرامج الجاهزة باللغة الإنجليزية.
- معظم الكتب والمراجع والدوريات والبحوث باللغة الإنجليزية، وربما يبرز حجم المشكلة لو أدركنا أن 90% من سكان العالم لا يتكلمون الإنجليزية أصلاً. ولا شك أن تكنولوجيا المعلومات تتجه نحو مزيد من المرونة اللغوية بتخلصها من قيود الأساس الإنجليزي خاصة بعد ظهور اليابان ولغتها الخاصة كقوة عالمية في هذا المجال. أشرنا فيما سبق إلى بعض مظاهر أزممتا اللغوية الحادة والتي من أبرزها تخلف قصور التنظير لها، وقصور المعاجم العربية، وقصور أساليب تعليم اللغة العربية، علاوة على ثنائية الفصحى والعامية، وسواء منفردة أو مجتمعة، تمثل أوجه القصور تلك عوائق حقيقية أمام جهود معالجة العربية آلياً. لقد أصبحنا في حاجة ماسة إلى إعادة النظر بصورة شاملة في مناهج تنظيرنا اللغوي، وذلك بهدف بلورة أساليب متقدمة لصياغة قواعد النحو والصرف بصورة رسمية formal ومنضبطة exact وسافرة explicit، وذلك تمهيدا لتطويع اللغة العربية لمطالب المعالجة الآلية، ولا نقصد بذلك اختزال قواعد العربية، أو تعديلها بل توافر البنى الأساسية اللازمة لمعالجتها آلياً، إن العالم يشهد منذ فترة تزيد على الأربعين عاما ثورة حقيقية في مجال اللسانيات، كان من نتائجها ظهور العديد من النماذج اللغوية، أو المناهج العامة تناول إشكالية اللغة بصورة أعمق وأشمل، وقد قام بوضع بعض هذه النماذج لغويون، والبعض الآخر حاسوبيون، وكدليل على مدى الثراء العلمي في مجال التنظير اللغوي نكتفي هنا بسرد قائمة هذه النماذج اللغوية:

- نحو تحليلي Analytical Grammar .
- نحوي توليدي Generative Grammar .
- نحو توليدي تحويلي Transformational Generative Grammar .
- نحو وظيفي Functional Grammar .
- نحو معجمي Lexical Grammar .
- نحو وظيفي معجمي (Lexical Functional Grammar) LFG .
- نحو علاقي Relational Grammar .
- نحو طبقي Stratificational Grammar .
- نحو تعميم البنية النحوية Generalized Phrase Structure Grammar

.((GSPG

- نحو بنية الجملة المبني على الرأس Head Phrase Structure Grammar

.((HPSG

- نحو المقولات Categorical Grammar .

- نظرية العاملية والترابطية (GB) Government Binding theory

وينطلق علماء اللغة شرقا وغربا محاولين تطبيق نماذج التعميد المختلفة على لغاتهم القومية، بل وتعقد المؤتمرات الوطنية لانتقاء أنسب هذه النماذج للتعامل مع الخصائص الأصيلة لهذه اللغات، أين نحن من هذا كله؟ لا يمكن للباحث الأمين إلا الإقرار بأن تنظيرنا اللغوي مازال في سباته العميق، وربما يكون السبب في ذلك هو قصور العدة المعرفية لكثير في من اللسانيين العرب الذي يعوزهم الإلمام بعلوم الرياضيات الحديثة والمنطق والإحصاء وعلم النفس واللسانيات الحاسوبية.

9:4:3 الاتجاهات الراهنة لتعريب المعلوماتية

لا يخرج الموقف الراهن لتعريب الكمبيوتر، ونظم المعلومات في مجمله عن كونه إحدى النتائج المنطقية لأزمتنا اللغوية الحادة، ولواقع مركزنا التقني والعلمي المتأزم، ويمكن تلخيص هذا الوضع الراهن بالملامح الرئيسية التالية: (أ) محاولة استيعاب العربية في نطاق التقنيات المصممة أصلا للغة الإنجليزية. فتحت ضغط الدوافع العملية واللهفة التجارية لموردي المعدات والبرمجيات، تم اتباع طرق «تعسفية» لإخضاع العربية للقيود التي فرضها النموذج الإنجليزي.

وقد تم ذلك إما بالتحايل حول هذه القيود الفنية على حساب المستخدم، وكفاءة النظام، أو الترخص في بعض خصائص اللغة العربية كتقليل أعداد أشكال الحروف وإغفال حركات التشكيل، في التعامل مع الكتابة العربية، إدخال وإخراجا، وتجنب قواعد الإبدال والإعلال، كما هي الحال في كثير من محاولات تطوير نظم آلية لمعالجة الصرف العربي.

إن استيعاب اللغة العربية آليا في نطاق الإنجليزية هي عملية خاطئة

من أساسها وذلك لسببين رئيسيين:

السبب الأول: أن اللغة العربية بصفاتها أعقد اللغات السامية والإنجليزية

باعتبارها من أبسط لغات الفصيحة الهندوأوروبية يمكن النظر إليهما كطرفي

نقيض على محور معالجة اللغة آليا .

السبب الثاني: أن العربية في معظم أمورها أعقد من الإنجليزية كتابة وصرفا ونحوا .

لهذين السببين تبدو عملية استيعاب العربية في إطار الإنجليزية بمثابة منطق معكوس أو محاولة مستحيلة لاستيعاب الأعقد في نطاق الأيسر .

ب - سطحية التعريب:

حيث انصبت معظم الجهود، حتى وقت قريب على كيفية التعامل مع اللغة العربية، على مستوى الحرف دون المستويات اللغوية الأخرى كالصرف والنحو والسياق، وأصبح مفهوم تعريب نظم المعلومات مرادفا لإدخال النصوص العربية من لوحات المفاتيح وطباعة النصوص العربية أو إظهارها على الشاشات المرئية .

ج - الاعتماد على الأجنبي:

فمن الشواهد البارزة أن معظم جهود تعريب تكنولوجيا المعلومات ظلت تفد إلينا من خارج الوطن العربي وبخامة من الولايات المتحدة وكندا والمملكة المتحدة وفرنسا، وقد شهدت الفترة الأخيرة تحولا واضحا نحو توطين هذه الجهود في الوطن العربي خاصة في المملكة العربية السعودية ومصر والكويت وسوريا وتونس والمغرب .

د- غياب البحوث الأساسية في حقل اللسانيات الحاسوبية

فعلى الرغم من جهود الهيئات والمؤسسات السابقة الذكر مازال اهتمام جامعاتنا ومعاهدنا ومجامعنا اللغوية دون الحد الأدنى المطلوب، (انظر الفقرة 6:3:2 من الفصل السادس) حجت اقترحنا مجموعة من موضوعات البحوث الأساسية المتعلقة بمعالجة اللغة العربية آليا .

5:9 تكنولوجيا المعلومات كأداة للغة العربية

9:5:1 تكنولوجيا المعلومات كأداة للإحصاء اللغوي

يمكن استخدام نظم المعلومات الإحصائية في كثير من المجالات، نذكر منها على سبيل المثال لا الحصر:

أ- التقييم الكمي لبعض خصائص اللغة، كمعدلات استخدام الحروف والكلمات والصيغ الصرفية، في نصوص العربية المختلفة، والموازن الشعرية،

وأنواع الأساليب النحوية، والحالات الإعرابية، أو المفردات المعجمية، من أمثلة هذه التطبيقات التي أنجزت بالفعل إحصائيات تواتر الجذور المستخدمة في المعاجم العربية، وإحصائيات عن تواتر استخدام الحروف والحركات والجذور الثلاثية والرباعية وأسماء الأعلام في النص القرآني الشريف، وذلك باستخدام المعجم المفهرس لألفاظ القرآن.

ب - التوصيف الكمي لبعض العلاقات اللغوية أو علاقات النصوص كالعلاقة بين نوع المبتدأ (عاقل / غير عاقل، مجرد/ محسوس، كائن حي / جماد،..)، ونوعية خبره المفرد (جامد أم مشتق، مجرد أم محسوس،..) أو العلاقة بين طول الجملة ومستوى الصعوبة في الكتب التعليمية للمراحل المختلفة.

ج - استخدام التحليل الإحصائي، كذلك في تفسير بعض الظواهر اللغوية مثل تفسير ظاهرة القلب المكاني لدى الأطفال (مثل قولهم «جذب» بدلا من «جذب»)، والتي فسرها إبراهيم أنيس على ضوء إحصائيات تواتر تتابع الحروف (تواتر زوج الحروف «جذ» أقل بكثير من «جب»). ركزت معظم الإحصائيات حاليا، على اللغة المكتوبة دون المنطوقة، وينقصها كثير من الإحصائيات المتعلقة بالعلاقات داخل المعجم العربي، وكذلك الخاصة بعنصر الدلالة (المعنى).

9:5 2 استخدام تكنولوجيا المعلومات في معالجة الكتابة العربية

من أهم تطبيقات الكمبيوتر في التعامل مع منظومة الكتابة العربية نظم القراءة الآلية للنصوص العربية automatic text reading، سواء المنسوخة typed أو المطبوعة printed، وهناك بعض محاولات لقراءة البيانات المكتوبة بخط اليد.

هذا عن شق القراءة، أما إظهار وطباعة النصوص العربية فقد قطعت تكنولوجيا المعلومات شوطا كبيرا في هذا المضمار، وأصبحت قادرة باستخدام طابعات الليزر على توليد معظم أنماط الحروف العربية (الثلاث، النسخ، الكوفي، الرقعة، الديواني،..) سواء بنظام الكتابة المعتاد أو بالخط العثماني المستخدم في كتابة النص القرآني الشريف.

9:5 3 تكنولوجيا المعلومات كأداة للصرف العربي

تم تطوير عدة معالجات آلية للصرف العربي قادرة على القيام بجميع

عمليات التحليل والتركيب لكلمات اللغة العربية، يقوم الشق التحليلي بتفكيك الكلمة، إلى عناصرها الأولية الاشتقاقية والتصريفية، والإعرابية، والواصق السابقة واللاحقة (مثال: تحليل كلمة «وإيجادهم» إلى حرف العطف «و» وساق الكلمة «إيجا» والضمير المتصل «هم»، ثم تحليل ساق الكلمة «إيجاد» إلى الجذر «وجد» على صيغة «إفعال»، بعد عكس عمليات إبدال «و» إلى «ي») أما التركيب الصرفي فيمثل العملية العكسية لتكوين الكلمات من عناصرها الأولية (كأن يغذى للمعالج الصرفي الآلي الجذر «ق و م» ويطلب منه تركيب فعله المضارع على صيغة «استفعل» لجمع المؤنث الغائب فيقوم المعالج بإخراج الكلمة النهائية «يستقمن»)، يتم ذلك من خلال قيام معالج الصرف الآلي بصهر الجذر في القالب الاشتقاقي المطلوب، والقيام أتوماتيا بجميع عمليات الإبدال والإعلال والحذف. يعد المعالج الصرفي الآلي مقوما أساسيا في تحليل النصوص العربية واسترجاعها، وكذلك في عمليات الإعراب الآلي للجمل العربية، وقد تم تطبيقه من قبل الكاتب بالفعل في تحليل كلمات النص القرآني الشريف واسترجاع مضمون هذا النص آليا.

5:9:4 تكنولوجيا المعلومات كأداة للنحو العربي

يمثل استخدام الكمبيوتر في التحليل النحوي للغة العربية تحديا كبيرا سواء على صعيد اللغة أو الكمبيوتر. يقوم نظام النحو الآلي بتفكيك (إعرابها parsing) إلى عناصرها الأولية من أفعال وأسماء وأشباه جمل وظروف وما شابه، وتحديد الوظائف النحوية لكل عنصر (فاعل، مفعول، خبر، صفة، حال، ..) وربط الضمائر بمراجعها والتعويض عن المحذوف، وذلك تمهيدا لتمثيل بنية الجملة بصورة تفصيلية سافرة.

وقد قام كاتب هذه السطور، بتطوير نظام للتحليل النحوي الآلي للغة العربية المكتوبة المشكلة وغير المشكلة يقوم بإعراب الجمل آليا ويشكلها تلقائيا.

يعد المحلل النحوي الآلي مقوما أساسيا لتحليل مضمون النصوص وفهمها آليا، وعمليات الاستخلاص والتلخيص والفهرسة الآلية، وكذلك نظم الترجمة الآلية من العربية إلى اللغات الأجنبية، علاوة على كونه عنصرا لا غنى عنه في نظم اكتشاف الأخطاء الهجائية والنحوية أتوماتيا.

5:9:5 استخدام تكنولوجيا المعلومات في الفهم الأتوماتي للسياق اللغوي

الهدف الأسمى لمعالجة اللغات الإنسانية آليا هو الوصول إلى نظام أوتوماتي لفهم السياق اللغوي في صورته المنطوقة والمكتوبة، وهناك عدة محاولات بدائية في اتجاه تحقيق هذا الهدف، والذي لا يمكن له أن يتحقق دون التصدي العلمي لمعضلة «المعنى في اللغة» على المستوى المعجمي والمنطقي والسياقي، وكذلك دراسة أثر مقام الحدث أو الحديث في تفسير معناه. يصب في نظم الفهم الأتوماتي نتائج المعالجات اللغوية الفرعية (الصرفية والنحوية والدلالية والمعجمية)، علاوة على ذلك لا بد وأن تكتب هذه النظم المعرفة الدارجة naive التي يدركها الإنسان بحسه الطبيعي (انظر الفقرة 2: 4 من الفصل الثاني).

هناك عدة محاولات لتطبيق أساليب التحليل الدلالي على نصوص عربية قصيرة تغطي موضوعا واحدا من فقرات الدستور اللبناني، وكذلك لفهم تمارين عالم الميكانيكا.

9:5: 6 تكنولوجيا المعلومات في تحليل النتاج الأدبي وأساليب الكتاب من أهم المجالات التي استخدمت فيها أساليب الإحصاء والتحليل اللغويين، تلك الخاصة بتحليل النتاج الأدبي، تراثه وحديثه، منشوره ومنظومه، وذلك للأغراض التالية:

أ- تحقيق التراث، ومن أبرز الأمثلة الدراسة التي استخدم فيها الكمبيوتر في تحليل ملحمة «الإلياذة» للتأكد من تفرد هوميروس بإنشائها (19: 145).
ب - التقييم الكمي لخصائص أساليب الكتاب، وقد أجريت دراسة حديثة لتحليل أساليب بعض الأدباء المصريين من حيث أطوال الجمل، ولا شك أنها تحتاج إلى تعميق وتوسيع من حيث تحديد معجم مفردات الأديب ومعدلات استخدامه لها، وتوزيع الأساليب النحوية المختلفة، ومدى استخدامه لظروف adverbs بأنواعها، وروابط الجمل conjuncts، والجمل الاعتراضية disjuncts وخلافه.

ج - التحديد الموضوعي لمدى تأثير الأدباء، والشعراء بمن سبقهم.
د- فهرسة النصوص آليا، وقد تم استخراج المعجم المفهرس للقرآن الكريم والحديث الشريف (كتاب الشيخين)، وذلك باستخدام المعالج الصرفي الآلي الذي أشرنا إليه في الفقرة 9: 5: 3 من هذا الفصل.

9:5: 7 تكنولوجيا المعلومات كأداة لمكننة المعجم العربي

يعاني المعجم العربي من أزمة حادة ومزمنة، ومظاهر أزمتنا المعجمية عديدة، أبرزها عزوف الناطقين بالعربية، عن استخدام معجم لغتهم الأم، وقصور حداد في المصطلحات، ولا يتسع الحديث هنا لمناقشة أسباب هذه الأزمة تفصيلاً وسأكتفي هنا بسرد سريع لبعض ما أراه من أسبابها:

أ- جمود النظرة إلى وسائل (آليات) تكوين الكلمات word formation في العربية حيث طغى الاشتقاق على الآليات الأخرى كتلك الخاصة بالكلمات المركبة (خط النار، إنسان العي، سفير فوق العادة)، والكلمات المزجية («درعمي» الهندوأوروبية).

ب- إهمال العلاقات بين المفردات والفصائل المعجمية، كعلاقات التضاد والترادف والاشتراك اللفظي.

ج- إغفال البعد التاريخي في البحث المعجمي العربي، حيث لا تفرق معظم المعاجم العربية الحالية بين قديم اللفظ وحديثه، وأهملت الدراسات الخاصة برصد التغيرات التي طرأت على معاني الألفاظ على مر العصور. د- الانفصال الحاد بين المجمع وجماعته فهناك شبه انعزال بين عمليات التحديث المعجمي، والاستخدام الفعلي للمفردات في المجالات المختلفة. هـ - ضمور عنصر الدلالة (المعنى) في الدراسات المعجمية.

و - فوضى لغة تعريف مفردات المعجم، وعدم التزامها بأنماط موحدة (من أمثلة هذه التعريفات القاصرة: ضرب من السمك، نبات معروف، على مسيرة ليلتين من الناقة،..)

ز- إغفال الجهود الجارية للسانيات الحاسوبية في تحليل بنية المعجم واستخدام تكنولوجيا المعلومات في مكننة المعاجم automation، ودعم جهود العمل المصطلحي.

وكما قلنا سابقاً، فإن بنية المعجم العربي وثراء مفرداته وتعدد علاقاته تحتم ضرورة مكننته، وذلك للأهداف التالية:

أ- توفير خدمة أفضل للمستخدم من خلال الأساليب المتطورة لاسترجاع المعلومات، وفرز مفردات المعجم بصور مختلفة، واستخراج قوائم المفردات وفقاً لما يحدده المستخدم من معايير.

ب - حصر التعابير المسكوكة idiomatic (مثل: أجهدش بالبكاء، حجر الزاوية)، والتعابير السياقية (مثل: الغزو الثقافي، الأمن الغذائي، مجلس

التعاون الخليجي، سوق الأوراق المالية،..).

ج - محاصرة ظاهرة الإزاحة الدلالية semantic shift التي تطرأ على المفردات العربية كتحويلها من الوصفية إلى الإسمية (أمثلة: شاحنة، مستد، ناظر).

د- تمييط لغة تعريف معاني المفردات.

هـ - إمكان دمج المعجم الممكن في النظم الآلية الأشمل كنظم الإعراب الآلي والفهم الأتوماتي للنصوص والترجمة الآلية.

من الآثار البارزة لاستخدام الكمبيوتر في المجال المعجمي، التوجه نحو بناء المعاجم باستخراج المعاني المختلفة للمفردات من واقع استخدامها في النصوص الفعلية، فقد ظهر أن معاني المفردات التي تتضمنها المعاجم التقليدية لا تغطي جميع المعاني المستخدمة بالفعل، يرجع ذلك إلى أن الجماعة اللغوية لا تتوقف عن ابتداء معان جديدة للمفردات، إما بتوسيع نطاق المعنى أو تضييقه أو إزاحته عن معناه، من أشهر المعاجم التي بنيت على أساس قاعدة الذخيرة هو ذلك الذي قامت بتطويره دار النشر الإنجليزية المعروفة كولينز Collins، بالتعاون مع جامعة برمنجهام. لتحقيق هذه المهمة يلزم توفير قاعدة كبيرة لذخيرة النصوص اللغوية corpus يتم تجميعها من مصادر مختلفة، بحيث تغطي نطاقا واسعا من الموضوعات، وأساليب الكتاب. وقد قام الكاتب خلال تطويره للمحلل النحوي الآلي للعربية بتجهيز قاعدة لذخيرة النصوص العربية، تم تحليلها صرفيا باستخدام المعالج الصرفي الآلي وقد بلغ حجمها حاليا ما يزيد على 12 مليون كلمة.

9: 5: 6 تكنولوجيا المعلومات كأداة لدعم العمل المصطلحي

بجانب الأزمة المعجمية التي تحدثنا عنها في الفقرة السابقة يواجه العمل المصطلحي عدة تحديات أخرى يلخصها الكاتب هنا من دراسة سابقة له (41):

أ- ندرة التأليف والترجمة باللغة العربية، خاصة في مجال التخصصات العلمية والفنية الجديدة، بما لا يعطي فرصة كافية لـ «تعتيق» المصطلح وزيادة استساغته وشيوعه وتوحيده.

ب - الموقف المعارض لتعريب العلوم من قبل كثيرين من داخل جامعاتنا وخارجها.

ج - ظاهرة العلم الضخم وانفجار المعلومات مما يزيد من معدلات الطلب المصطلحي بصورة أسية exponential تعجز جميع الوسائل اليدوية عن مجاراتها .

د- إهمال الرصيد الهائل للعربية الكلاسيكية والسحيقة، وهو الرصيد الذي يمثل المخزون الاستراتيجي «لاستهلاكنا» المصطلحي، لقد بات علينا إحياء المهجور، وأنقل هنا ما قاله جميل الملائكة في هذا الصدد عن ضرورة «الإفادة من الخزين الكبير من الألفاظ القديمة المماتة.. فهي تكاد تكون الأداة الرئيسية المستعملة اليوم لوضح المصطلحات الجديدة في اللغات الأوروبية وإن اختلفت طبيعة تلك اللغات بعض الشيء عن لغتنا العربية، فهم يرجعون في الأكثر إلى أصول اللغة اللاتينية أو اليونانية القديمة» (17).

هـ - عدم كفاية المعاجم المتخصصة التي تمهد لدخول المصطلح المعجم العام، إلا أن الساحة العربية تشهد منذ فترة جهدا مشكورا في هذا المجال. و - إغفال المجامع العربية ثورة للسانيات والمعلومات، وتجدر الإشارة هنا إلى المبادرات التي قام بها أخيرا مجمع اللغة الأردني والمجمع المصري. يرى الكاتب، أنه لا بديل عن استغلال تكنولوجيا المعلومات، وبشكل مكثف لتعويض تخلفنا في مجال العمل المصطلحي، فقد بات واضحا للجميع عجز الوسائل اليدوية التقليدية أمام الطلب المصطلحي المتدفق، يمكن لتكنولوجيا المعلومات أن تقدم دعما حقيقيا للأنشطة الرئيسية للعمل المصطلحي من اقتناء مصادر المعلومات وتحليل المادة المعجمية، وتوثيق المصطلح وإجراء ونشر قوائم المصطلحات وتقديم الدعم للمعجميين في استخدام المعاجم والمكانز العامة والمتخصصة .

لقد انتشر استخدام تكنولوجيا المعلومات لإقامة بنوك المصطلحات في أوروبا وآسيا معاونة المترجمين، ودعم الترجمة الآلية وبناء المعاجم المختصة، ودعم التوحيد القياسي، وكذلك لخدمة الشركات الصناعية الكبرى في ترجمة وثائقها الفنية تلبية لمطالب السوق الخارجي. من أمثلة ذلك بنك المصطلحات الذي أقامته السوق الأوروبية المشتركة المعروف باسم EURODICAUTON وبنك مصطلحات شركة سيمنس Siemens المعروف باسم .TEAM

وهناك عدة جهود عربية رائدة في مجال بنوك المصطلحات وتوحيدها، من أمثلة ذلك بنك المصطلحات المعروف باسم «باسم» الذي قام بتطويره في السعودية د. محمود الصيني، أحد الرواد العرب البارزين في مكننة المعاجم والترجمة الآلية، وبنوك المصطلحات في بيت الحكمة التونسي ومكتب تنسيق التعريب بالمغرب، ولجنة توحيد المصطلحات العسكرية في سوريا. وأود أن أنهي هذه الفقرة، بحدث عن علاقة المصطلح بالتكنولوجيات الحديثة كتكنولوجيا المعلومات والتكنولوجيا الحيوية على سبيل المثال، وما تقوم عليه من علوم نظرية وتطبيقية، تتعامل معظم هذه الفروع المعرفية المستحدثة، مع مفاهيم مركبة يصعب التعبير عنها بكلمات مفردة في كثير من الأحيان، وهو الأمر الذي يستوجب كما ذكرنا سابقا-ضرورة إعادة النظر في أساليب (آليات) تكوين الكلمات في العربية كتكريب المصطلح من أكثر من لفظ compounding (لوحة المفاتيح، الشفرة الوراثية)، أو باستخدام أسلوب المزج blending كما في «درعمي» و«الهندوأوروبية»، وهو ما سأتناوله بمزيد من التفصيل هنا، نظرا لندرة الدراسات حوله ولأهميته بالنسبة لمقامنا الحالي.

أورد جميل الملائكة، في مقالته عن «المصطلح العلمي ووحدة الفكر» (17) بعض أمثلة النحت المزجي التي اقترحها ساطع الحصري، نذكر منها:

- تحشعوري: للدلالة على المفهوم المركب من «تحت + شعوري».
- قبتاريخي: للدلالة على المفهوم المركب من «قبل + تاريخي».
- خامدرسي: للدلالة على المفهوم المركب من «خارج + مدرسي».

وقد اعترض جميل الملائكة على أسلوب النحت المزجي بصفة عامة، على أساس أنه لا يتفق وطبيعة اللغة العربية، وهو الرأي الذي أعترض عليه هنا، فعدم استساغة هذه المصطلحات المزجية لا يرجع إلى أسلوب «المزج» نفسه بل نوع المزج الذي نهجناه في نحتها، فهو مزج لصقي agglutinative، يلصق مقطعا من لفظ مع مقطع من لفظ آخر ليخرج بلفظ مركب غريب عن قوالب تكوين الكلمة العربية بموازينها الصرفية المعهودة. ولغتنا العربية كما هو معروف ليست لغة لصقية كالتركية والألمانية، بل لغة اشتقاقية انصهارية حيث ينصهر في بنية كلماتها الأصل المعجمي مع صيغة الميزان الصرفي، في وحدة مندمجة ذات إيقاع رصين ومستساغ للغاية،

وذاات قابلية عالية للاشتقاق.

ما أشد حاجتنا إزاء الانفجار المعرفي الذي نعايشه حالياً لتعزيز آليات تكوين الكلمات في العربية، ويأتي أسلوب المزج هنا كإحدى الوسائل المتاحة للدلالة على المفاهيم المركبة التي تسود معظم فروع المعرفة الحديثة، وفي هذا الصدد أطالب بإعادة اكتشاف صيغة الأصل الرباعي (مثل: فعلل وتفعلل وفعللة، وفعلنة) حيث تفوق هذه الصيغة الصيغ الثلاثية في قدرتها على نحت الصيغ المزجية، لتوضيح وجهة نظري أكتفي هنا ببعض الأمثلة: أ- من أنجح أمثلة المزج لفظ «درعمي» المركب من لفظتي (دار + العلوم)، أسر وراء نجاح هذا المصطلح أنه قد خرج بأصل رباعي جديد هو «درعم» ليتوافر بذلك المدخل الطبي السليم، الذي يؤهله لولوج المعجم العربي، وهو المدخل الذي يفجر طاقة الاشتقاق لهذا الأصل المنحوت، حيث يمكن أن يشتق منه الفعل واسم الفاعل واسم المفعول، والمصدر الصناعي (تدرعم، متدرعم، مدرعم، درعمية أو تدرعمية)، ويبدو هذا الأصل في أوج إشعاعه الاشتقاعي عندما تدين له صيغة جمع التكسير «دراعمة» بديلاً أكثر استساغة من الصيغة المطردة لجمع المذكر السالم «درعميون» ليتبوأ بذلك مكانه بشكل نهائي في صلب معجمنا العربي.

(ب) ومثال آخر ناجح من النحت المزجي هو لفظ «تزلج» للدلالة على التزلج على الثلج، فلقد خرج هو الآخر بأصل رباعي انصهاري هو «تزلج» ليوضع بذلك على نقطة البداية السليمة لمسار الاشتقاق الكامل لينبثق منه الفعل «تزلج»، واسم الفاعل «متزلج»، واسم المفعول «متزلج عليه»، واسم المكان «متزلج»، وصيغة الفعل المزيد «تزالج اللاعبان».

(ج) والمثال الأخير من لدى الكاتب عندما حاول أن ينحت مصطلحاً مزجياً انصهارياً لمفهوم التعليم من خلال العمل on-job-learning، وهو أحد المفاهيم الشائعة في الخطاب التربوي الحديث، وقد أغراه على فعل ذلك الحروف المشتركة بين لفظتي «تعلم» و«عمل» والتي أوحى له بأصل رباعي جديد هو «تعلمل»، وأترك للقارئ الحكم على مدى استساغة ما يشتق منه مثل «المتعلمين، علملة، علملي».

9:5 8 تكنولوجيا المعلومات في مجال الترجمة الآلية

ظلت الترجمة الآلية حلماً يراود خيال الكثيرين منذ ظهور الكمبيوتر

في أواخر الأربعينيات، وبعد سلسلة من البدايات الفاشلة أخذت الترجمة الآلية تحقق نجاحا ملموسا في مجال ترجمة الوثائق الفنية والعلمية، وتعد الترجمة الآلية إحدى الغايات النهائية التي لقب في معش روافد نظم التحليل والتركيب اللغويين، لهذا السبب وبجانب كونها تطبيقا قائما بذاته، ينظر البعض إليها كنموذج آلي شامل computerized model لدراسة أداء المنظومة اللغوية.

وسأوجز هنا بعض القضايا الأساسية المتعلقة بالترجمة الآلية:

أ- تمثل الاستعارة والمجاز، والأساليب البلاغية أعقد المشاكل التي تواجه نظم الترجمة الآلية لذا فإن الجهود متجهة حاليا ولفترة قادمة ستطول إلى ترجمة الوثائق العلمية والفنية التي تتسم بالصياغة المنضبطة أو شبه المنضبطة.

ب - من أهم الأمور في تطوير نظم الترجمة الآلية هي نطاق الموضوعات subject domain أو شريحة اللغة sublanguage، التي تتعامل معها هذه النظم، معظم النظم الحالية تركز على موضوع واحد أو نطاق ضيق للغاية من الموضوعات المتقاربة لغويا ومعرفيا، من أشهر هذه النظم هو نظام METEO الكندي، الذي يترجم نشرات الأرصاد الجوية ما بين الإنجليزية والفرنسية. ما أن يتم تحديد نطاق الموضوع حتى يتقلص إلى حد كبير لبس معاني الكلمات word sense ambiguity، وهو أكثر المشاكل صعوبة في عملية الترجمة، فكلمة «قانون» في مجال التشريع ليس لها إلا معنى واحد في الغالب، في حين تتعدد معانيها ما أن يتسع نطاق الموضوعات لتشمل قانون بمعنى الآلة الموسيقية المعروفة، أو بمعنى القاعدة العلمية، أو النظام أحيانا. لهذا السبب يبدو تطوير نظم للترجمة الآلية غير محددة الموضوع domain-unrestricted بعيد المنال في ظل المتاح حاليا من الوسائل اللغوية والمعجمية والآلية.

ج - يمثل التباين بين اللغات، خاصة تلك التي تتدرج تحت فصائل لغوية مختلفة كالإنجليزية في مقابل العربية، أو اليابانية، أو الصينية، مشكلة أساسية لنظم الترجمة الآلية، فما تقوم به لغة ما، على مستوى نظام النحو يمكن أن تقوم به لغة أخرى على مستوى المعجم، وهناك اختلافات جوهرية في رتبة الكلمات داخل الجمل وأشباه الجمل، واستخدام الضمائر، فالعربية مثلا-على عكس الإنجليزية-تقدم الفعل على الفاعل والموصوف على الصفة،

وتستخدم الضمائر المستترة وضمائر الربط.

د- تمثل دقة الترجمة الآلية وأسلوب تقييم نظمها المختلفة معضلة أخرى، وبالتالي حجم التدخل البشري المطلوب لتوضيب النص قبل ترجمته pre-editing أو تهذيبه بعد ترجمته post-editing، وقد حققت نظم الترجمة درجة دقة تراوح ما بين 60-90٪، وفي هذا الصدد يلزم التنويه إلى أن مهمة المترجم، بشرا كان أو آلة، هي نقل المعنى من لغة المصدر source language إلى لغة الهدف target language وليست مهمته فك اللبس أو إجلاء الغموض، بل من أمانة الترجمة أن ينقل اللبس أو الغموض إن وجد فربما قصده صاحب النص الأصلي.

بوجه عام وبشكل نسبي، تعد الترجمة من اللغات الأجنبية إلى العربية، أبسط من الترجمة في الاتجاه العكسي من العربية إلى اللغات الأجنبية، ومن حسن الحظ أن القدر الأكبر من جهد الترجمة يندرج تحت الاتجاه الأول، إلا أن الترجمة من العربية إلى الإنجليزية والفرنسية ذات أهمية خاصة بالنسبة لتعليم صغارنا اللغات الأجنبية، بجانب الأمور المتعلقة بنشر تراثنا خارج حدود الوطن العربي، خاصة للدول التي استقلت حديثا بعد تفكك الاتحاد السوفييتي، المشكلة الأساسية في الترجمة من العربية، هي تلك المتعلقة بغياب التشكيل في معظم النصوص المتوقع ترجمتها، لذا يعد نظام التشكيل التلقائي الذي أشرنا إليه في الفقرة 5:9 من هذا الفصل أحد المقومات الأساسية لنظم الترجمة الآلية من العربية.

أبدت كثير من المؤسسات العربية الأكاديمية والتجارية، اهتماما بنظم الترجمة الآلية، ومنذ منتصف الثمانينيات، وحتى يومنا هذا تقيم السعودية، بصورة شبه سنوية، ندوات وحلقات دراسة للجوانب المختلفة للترجمة الآلية، وقام المركز الإقليمي للمعلوماتية بتونس (IRSIT) بتطوير نموذج أولي لترجمة الجمل الإنجليزية البسيطة إلى العربية (نظام ترجمان)، وقام معهد الإلكترونيات التابع للمركز القومي للبحوث بالقاهرة بتطوير نموذج أولي آخر للترجمة ثنائية الاتجاه ما بين العربية والإنجليزية.

9:5:9 تكنولوجيا المعلومات كأداة لتعليم وتعلم اللغة العربية

تتراوح البرمجة التعليمية في مجال اللغة ما بين تلك الخاصة بألعاب الكلمات بفرض تنمية حصيلة المفردات إلى تلك التي تتبع مناهج متكاملة

لتعليم وتعلم قواعد الصرف والنحو، وإكساب مهارات القراءة والكتابة، هذا ويواجه تعليم اللغات باستخدام الكمبيوتر حاليا عدة مصاعب أهمها:- غياب عنصر الكلام المنطوق.

- صعوبة محاكاة المواقف الطبيعية للاستخدامات اللغوية.
- احتياج معظم البرامج إلى سعة تخزين هائلة لحفظ المادة التعليمية خاصة في مجال تعليم القراءة والإنشاء (انظر الفقرة 3: 2: 4 من الفصل الثالث).

هناك عدة محاولات مشجعة لاستخدام الكمبيوتر التعليمي في تعليم وتعلم قواعد اللغة العربية وتعليم الأبجدية للصغار، وإتقان تهجئة الكلمات وتمييز أقسام الكلم، وإعراب الجمل واستخلاص الجذور وتصريف الكلمات وتكوين جمل قصيرة، والتدريب على استخدام المعجم العربي.

5:9: 15 توليد الكلام العربي وفهمه آليا

يسعى أهل الذكاء الاصطناعي لتطوير نظم آلية قادرة على تمييز الكلام المنطوق وفهمه آليا، وكذلك للتعرف على شخص المتكلم الذي سبق تخزين أنماط صوته، ويمكن تقسيم نظم تمييز الكلام إلى تلك التي تتعامل مع الكلمات المنعزلة والكلام المتصل غير المتدفق connected speech، والكلام المستمر المتدفق continuous speech، وهو أكثرها صعوبة بالطبع. وهناك أيضا محاولات لمحاكاة النطق البشري لتوليد الكلام آليا، وتشير الأدبيات إلى وجود عدة محاولات أولية لتوليد الكلام العربي آليا بتحويل النصوص المدخلة من لوحة المفاتيح إلى مقابلها المنطوق، وهناك بدايات لتمييز الكلام العربي المنطوق على مستوى الكلمات المنفردة.