المحاضرة الخامسة

الحاسوب والنحو العربي

الدكتور نبيل علي

نائب رئيس مجلس الإدارة ومدير البحوث

في مؤسسة صخر – مصر

السبت 29 محرم 1417هـ - 15 حزيران 1996م

مقدمة:

تعد معالجة النحو العربي آلياً بوساطة الحاسوب أحد المقومات الرئيسية لإعداد اللغة العربية لكي تلحق بعصر المعلومات، حيث يتوقف على هذا المعالجة مدى نجاحنا في تطوير برامج عربية في العديد من التطبيقات التعليمية والمكتبية والثقافية، وكذلك في تطوير أساليب البحث اللغوي لدينا وتحديث البنية الأساسية للغتنا الأم، وتحصيئة هذه اللغة العظيمة لتصبح عضواٍ بارزاً في نادي "تعدد اللغات" الذي يتشكل حالياً خلال شبكة "الإنترنيت" التي ستكون – بلا شك – ساحة للاحتكاك اللغوي لم يسبق لها مثيل.

يتسم النحو العربي بخصائص عديدة تجعل من أمور معالجته آلياً بوساطة الحاسوب موضوعاً مثيراً سواء على المستوى اللغوي أو المستوي الحاسوبي، هذا من جانب ومن جانب آخر فإن العلاقة الوثيقة التي تربط بين نحو العربية وصرفها ومبناها ومعناها، تفرض على الباحث ضرورة تناوله لهذه القضية واضعاً نصب عينيه علاقات التداخل المتشابكة التي تربط بين الفروع اللغوية المختلفة داخل منظومة اللغة العربية.

تبدأ الدراسة باستعراض سريع للعوامل التي أدت إلى زيادة أهمية اللغة فيعصر المعلومة لتنتقل بعده إلى حديث مختصر عن موقع النحو في منظومة اللغة ثم إبراز خصائص النحو العربي ذات الصلة بالمعالجة الآلية، بعدها يتم سرد النماذج النحوية المختلفة grammar models التي يمكن من خلالها توصيف منظمة النحو العربي بصورة رسمية formal وذلك تحصيئة لدخوله مصاف العلوم المنضبطة والذي بدوره يعد شرطاً أساسياً لتعامل النحو مع الحاسوب من خلال الرياضيات والمنطق والإحصاء، بعد ذلك يقوم الباحث باستعراض نتائج بحثه في تطوير نظام آلي لإعراب اللغة العربية المكتوبة، وقد تمكن من خلاله من تطوير أداة برمجية ذكية لتشكيل الجمل العربية الخالية تماماً من عناصر التشكيل بصورة تلقائية، تنتهي الدراسة بمناقشة دور المعجم العربي في مساندة نظم معالجة اللغة العربية آلياً وكيفية تطوير هذا المعجم لتلبية مطالب هذه النظم.

أولاً : تزايد أهمية اللغة في عصر المعلومات

هناك علاقة حميمة تربط بين اللغة وتكنولوجيا المعلومات سواء بصفتها – أي اللغة– مقوماً رئيسياً لكثير من التطبيقات المعلوماتية الحديثة، أو من حيث كون هذه التكنولوجيا أداة لمساندة البحث اللغوي، بل وإعادة النظر على المنظومة اللغوية ككل، وهو الأمر الذي دفع كثيراً من الأمم أن تراجع بصورة شاملة علاقة لغاتها القومية بتكنولوجيا المعلومات، لقد أدرك ذوو البصيرة والحمية، بل وأصحاب المنفعة العملية أيضاً أن التقاعس في هذا الصدد يمثل تهديداً حقيقياً بظهور فجوة لغوية تفصل بين أمم العالم في عصر المعلومات، ويقصد بالفجوة هنا تخلف اللغة تنظيراً وتعجيماً، استخداماً وتوثيقاً، تعليماً وتعلماً، وليس من قبيل المبالغة أن نقول إن مثل هذا التخلف إن حدث سيؤدي حتماً على انزواء لغات من يتخلفون إلى مصاف الدرجة الثانية وما دونها، أو بقول آخر هناك طبقية عالمية جديدة على وشك أن تظهر في عصر المعلومات ألا وهي "الطبقة اللغوية".

تتضح أهمية اللغة من خلال تتبعنا لمراحل تطور نظم المعلوماتية (شكل رقم 1)، فخلال نصف القرن المنقضي منذ ظهوره تطور الحاسوب خلال عدة نقلات نوعية أدت به في النهاية إلى مواجهة حاسمة مع منظومة اللغة على اتساعها، ففي البداية استخدم الحاسوب كآلة لسحق الأرقام number crunching **واقتصرت تطبيقاته على النواحي التجارية ذات الطابع الرقمي والعمليات الحسابية والمنطقية المحدودة، ولذلك بهدف إصدار الفواتير وكشوف الحساب وقوائم المرتبات وما شابه ذلك.**

في السبعينيات تطور الحاسوب ليصبح آلة لمعالجة المعلومات information processing **وانتشرت نظم معلومات الإدارة** MIS **وتصور البعض خطأ أنها الوسيلة الفعالة في حل المشاكل التي تواجه الإدارة على اختلاف مستوياتها، إلا أن المؤشرات والتحليلات والإحصائيات التي أفرزتها نظم المعلومات تلك ما كانت لتكفي وحدها لحل كثير من المشاكل، قديمها وحديثها، التي تحتاج إلى طرائق أخرى من الحدس والخبرة المكتسبة، أو بقول آخر تحتاج إلى نوع جديد من المعرفة يمتزج فيها ثالوث المعلومة والحكمة والقدرة على**

اتخاذ القرار، وهو الأمر الذي انتقل بآلتنا الفريدة من كونها آلة لمعالجة المعلومات إلى آلة لمعالجة المعرفة knowledge processing**، وعندها حدثت**

أمام تحد حقيقي خلاصته: كيف يمكن لهذه الآلة الصماء أن تتعامل مع إشكالية اللغة بكل ما يكتنفها من لبس واستعارة وحذف واقتضاب وإطناب؟!! لقد دفع هذا التحدي بالبحث اللساني إلى مشارف جديدة، وفرض على أهل اللغة إعادة النظر إلى منظومة بصورة شاملة، بل وأيضاً العلاقات الوثيقة التي تربط هذه المنظومة بخارجها، ونقصد بذلك علاقاتها بالمنظومات التربوية والثقافية والتكنولوجية بل بالاقتصادية والسياسية أيضاً.

ولم يكن الأمر على صعيد الحاسوب أقل إثارة وتحدياً، فقد فرضت عليه المواجهة مع اللغة ضرورة الارتقاء بكثير من خصائصه وقدراته حتى يتهيأ لهذا اللقاء المثير، وقد شملت نواحي ارتقائه زيادة سرعته الحسابية وسعة ذاكرته وطاقة تخزين وسائطه المغناطيسية والضوئية، والأهم من ذلك الارتقاء بأساليب برمجته التي سعت على التخلص من طابعها القطعي الخورازمي algorithmic & deterministic **الذي لا يستطيع إلا التعامل مع المدخلات الصريحة الظاهرة، وذلك حتى تصبح الآلة أكثر قدرة في التعامل مع غير المحدد وغير القطعي وغير السافر، فدون هذه القدرة يتعذر التعامل مع غموض التعبير اللغوي ومرونة تراكيبه، وتعدد معاني ألفاظه وترادفها، ويستحيل دونها استنباط المحذوف عند الاقتضاب وتجنيب الفائض عند الإطناب.**

وكدليل على ما أحدثه اللغة في تطوير منظومة الحاسوب نقدم هنا تلخيصاً للمقومات الرئيسية لمنظومة الجيل السادس للحاسوب (شكل رقم 2) ممثلة في العلوم الأساسية التي قامت عليها تكنولوجياته المختلفة من جانب، ونوعية التطبيقات التي انبثقت من هذه التكنولوجيات من جانب آخر.

تشمل قائمة العلوم الأساسية:

– علم وظائف الأعضاء.

- **علم النفس .**
- **اللسانيات.**
- **المنطق.**

ولسنا بحاجة هنا لأن نؤكد العلاقة الوثيقة بين اللغة وكل من هذه العلوم الأساسية، أما التكنولوجيات الرئيسية فمعظمها ذو أبعاد لغوية مختلفة تتراوح ما بين تلك التي على مستوى الحرف والكلمة (مثل نظم القراءة الآلية) إلى تلك التي تغوص إلى مستوى المعاني والمفاهيم كنظم الفهم الأتوماتي automatic understanders وتمييز الكلام آلياً automatic speech recognition.

وأما على صعيد التطبيقات فهناك أربعة تطبيقات أساسية للجيل السادس وهي:

- النظم الخبيرة Expert system
- الترجمة الآليةMachine translation
- النظم الذكية للإنسان الآلي.Interlligent robotics
- تطبيقات مساندة الحاسوب لتصميم والتصنيع.CAD/ CAM

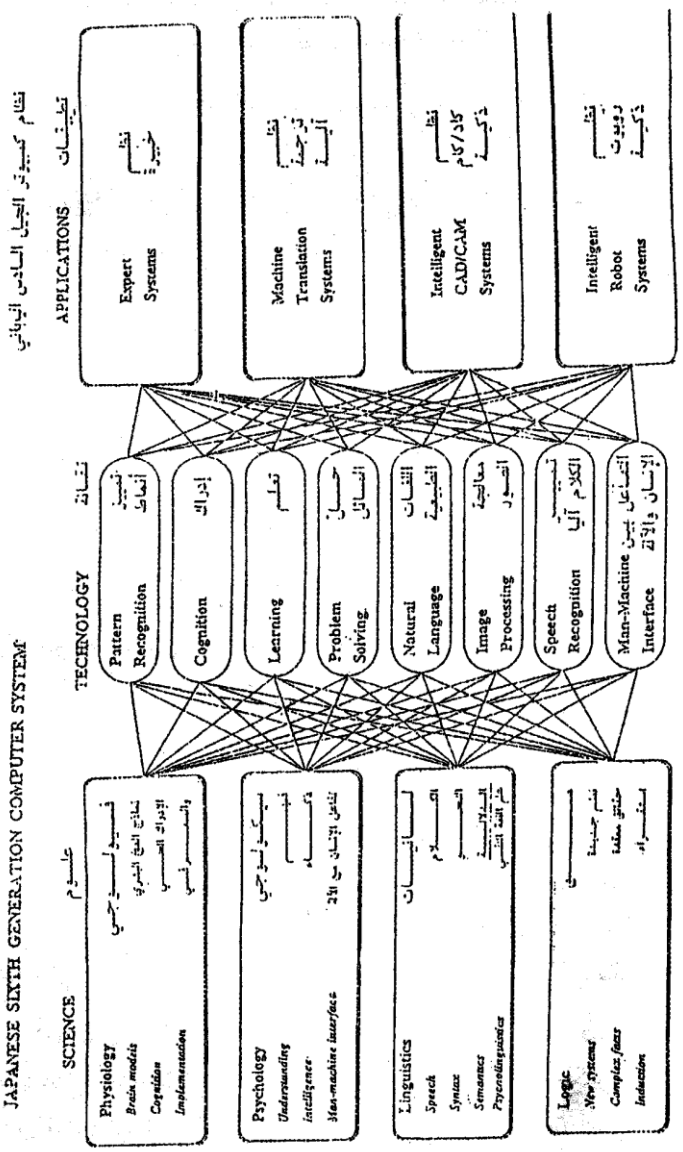
باستثناء النوعية الأخيرة ترتبط هذه التطبيقات مع اللغة بصلات وثيقة بدرجات متفاوتة، فالنظم الخبيرة تحتاج لتخزين المعارف التي تؤسس عليها خبرتها وبالتالي تحتاج إلى اللغة كأهم وسائل تمثل المعرفة ونقلها، ونظم الإنسان الآلي الذكية تحتاج إلى قدرات لغوية لكي تستوعب الأوامر وتتواصل مع الإنسان البشري، أما الترجمة الآلية فهي بحكم طبيعتها تطبيق لغوي صرف.

ثانياً: موقع النحو في منظومة اللغة

يمثل النحو والصرف العربيات توأمين يصعب الفصل بينهما سواء على مستوى التنظير أو المعالجة الآلية، وتأثير النحو على الصرف يكمن في الدور الذي تلعبه الآليات النحوية المختلفة في تحديد بنية الكلمة، ونقصد بها آليات الإعراب والتصريف والإلصاق والتقديم والتأخير، أما تأثير الصرف على النحو فيتخلص أساساً في الدور الذي تلهبه الصيغة الصرفية للكلمة، ومعاني مبانيها – بالتالي-، في تحديد صرفها النحوي من حيث طبيعة المكملات وعددها Complements والفضلات adjuncts التي يمكن أن تتعلق بالكلمة تركيبياً.

أما عن علاقة النحو بمنظومة الصوتيات فتتركز في أثر البنية النحوية للجمل على أنماطها التنغيمية intonational patterns والعكس، والدور الذي يلعبه

التلوين الصوتي في كشف بنية السياق المتصل من خلال إبراز مناطق التبئير focusing والفواصل بين الجمل والفقرات.



الكمبيوتر : الكمبيوتر كآلة ذكية لا آلة حاسبة ـــ الكمبيوتر آلة لمعالجة الرموز لا معالجة الأرقام

الشكل رقم ٢

اقتصر الحديث عن علاقة النحو بالمعجم في تلك الأمـور المتعلقـة بـاللزوم والتعدي وعلاقة أبواب الفعل بهما، وبينما تركز المعاجم العربية علـى المعطيـات الصرفية فقد أهملت المعطيات النحوية التي يصعب دونها توصيف كيفية استخدام

الكلمات داخل السياق، وأثر هذا السياق على تحديد معاني هذه الكلمة، لقد طغت الكلمة العربية على سياقها، أو بقول آخر لقد طغى الصرف على النحو والمعنى وربما يرجع ذلك على أن الكلمة العربية ذات الطابع الاشتقاقي تتسم بقدرة عالية على التضمين أو احتواء المعاني.

تشمل العلاقة النحوية – المعجمية تحديد القيود الدلالية والنحوية solectional restrictions لمتعلقات الفعل من فواعل ومفاعيل وحروف وظروف، وكذلك بمتعلقات الاسم من حيث طبيعة ما يضاف إليه وما يمكن أن يرتبط به من حروف وظروف وأدوات، إضافة إلى ما أسسه عبدالقاهر الجرجاني في أسرار بلاغته ودلائل إعجازه لم تحظ العلاقة النحوية – الدلالية بما تستحقه من قبل اللغويين والبلاغيين، ربما يكون وراء ذلك صعوبة التعامل مع شق المعنى اللغوي عموماً، وعجز المنطق الصوري ذي الرتبة الأولي first order logic عن تذليل هذه الصعوبة، لقد حير زيغ المعنى الفلاسفة الذين فشلوا –عبر القرون– في تحديد ماهيته، وقد تحاشي الخوض فيه كثير من اللغويين (بلومفيلد على سبيل المثال) ويعتبره علماء الحاسوب المعضلة الأساسية في إكساب آلتهم قدراً من الذكاء الاصطناعي.

إن العلاقة بين ظاهرة التركيب النحوي – كما يتبدى في رتبة سلاسل الكلمات وما يطرأ على هذه الكلمات من تغيرات وتبديل وتأخير– والمعنى الذي يبطنه هذا التركيب تمثل أهم القضايا التي يتناولها الدرس اللساني الحديث، وفي حديثي هذا سوف أتجنب مواجهة السؤال المحوري الذي أثار – وما زال يثير– جدلاً هائلاً ألا وهو: أي من الشقين النحوي أو الدلالي له الأولوية في توصيف الأداء اللغوي وتفسيره بصفة عامة، أو بقول آخر هل معنى الجملة تابع لبنيتها التركيبية، أو يتولد التركيب كتابع للمعنى تجسيداً أو تحقيقاً له؟ لا شك أن العلاقة بينهما – ككل العلاقات البينية بين فروع اللغة المختلفة– هي علاقة تبادلية، وأحيل القارئ هنا إلى مقولة عالم اللغة الإنجليزية المعاصر جون سينكلير إذ قال: "إن كل تغيير في الشكل لا بد أن يتبعه تغيير في المعنى".

ثالثاً: خصائص النحو العربي

بداية، دعنا نقر هنا بأن نحو "العربية" هو حالة خاصة من النحو العام الذي، تسعى النظريات الحديثة إلى وضع أسسه لجميع اللغات ، ولا يعني ذلك طمس خصوصيته، بل على العكس، فإن ذلك يؤكد أهمية معرفة ما يميزه من خصائص حتى يمكن إدراجه، وتحديد موقعه.

من خلال فهمنا لخصائص النحو العربي ستبرز المطالب الأساسية لمعالجته آلياً، ونوعية المشاكل التي يجب أن نتصدى لها، إن معظم النظم النحوية الآلية قد أقيمت على أساس اللغة الإنجليزية، لذا فإن تقييم مدى ملاءمة هذه النظم لمعالجة النحو العربي لا بد أن يبنى على أساس الفهم الدقيق لخصائصه، وأهم الفروق التي تفصل بينه وبين نحو "الإنجليزية".

وفي رأي المؤلف أن أهم خصائص النحو العربي التي تعنينا هنا بالدرجة الأولى هي:

– العلاقة العضوية بين النحو والصرف.

- رتبة الكلمات داخل الجملة العربية.

- المرونة النحوية.

- التوسط النحوي للغة العربية.

بغرض الإيجاز سيقتصر الحديث هنا على خاصية التوسط النحوي للغة العربية حيث سيتكشف من خلاله بعض الأمور المتعلقة بالخصائص النحوية الأخرى.

في رأي المؤلف، يأخذ نحو "العربية"، موقعاً وسطاً بين اللغات الإنسانية، وهو ما سنسميه هنا بخاصية "التوسط النحوي"، ولا نتحدث هنا تحت وقع الغيبيات اللغوية بل على ضوء معطيات علم طوبوغرافيا اللغات، ونعني بالتوسط النحوي مثل نحو "العربية" إلى الشائع، وكرهه للشاذ والشارد، بجانب جمعه بين عديد من الخصائص المشتركة مع أنحاء لغات أخرى.

لتأسيس ظاهرة التوسط النحوي التي نزعمها هنا نورد فيما يلي بعض الشواهد التي تساعد على تحديد موقع النحو العربي في خريطة أنحاء اللغات الإنسانية، وذلك باعتبار الخواص اللغوية التالية:

- رتبة الكلمات.
- المرونة النحوية.
- استخدام عناصر الكينونة.
- المطابقة.
- ظاهرة الوصل.
- النفي.
- حالات الفعل (أو درجة يقين المتكلم).
- أداة التعريف.
- أدوات المصدرية.
- استخدام الأداة كفاعل.
- درجة الإضمار.

والتالي ملخص بالشواهد التي تؤكد توسط نحو العربية في كل من هذه الخواص:

الظاهرة	موقع النحو العربي
1- رتبة الكلمات في الجملة.	تجمع العربية بين الجملة الاسمية والجملة الفعلية.
2- المرونة النحوية.	العربية وسط ما بين صرامة قيود الرتبة كما في "الإنجليزية"، والمرونة المفرطة كما في "الهنغارية"، و"اليابانية".
3- الخاصية الإعرابية.	العربية (3 حالات إعراب) وسط بين اللغات "الإنجليزية"، واللغات المسرفة فيه كـ"الروسية" (6 حالات إعراب) و"اللاتينية" (5 حالات إعراب).
4- استخدام عناصر	اللغة العربية وسط في استخدام فعل الكينونة، فليس استخدامه ملزماً

الكينونة	في جميع الحالات كما في اللغات الجرمانية، وليس مغفلاً تماماً كما في بعض اللغات الإفريقية، حيث تستخدم العربية ماضي الكينونة وتترخص في استخدام مضارعه. وتجمع العربية أيضاً بين كينونة الفعل وكينونة الاسم (الضمير هنا) (مثال: "الله هو المنتقم") وكينونة الأداة (مثال: "ما محمد بغائب")
5- المطابقة	ليست ظاهرة المطابقة في العربية محدودة للغاية كما هي في "الإنجليزية" حيث تقتصر على حال الفعل المضارع مع الفاعل المفرد الغائب، ولا مفرطة كما في "الألمانية"، "والفرنسية" حيث تمتد المطابقة لتشمل أدوات التعريف مع الأسماء ، وتخلو "العربية" كذلك من ظواهر المطابقة المتطرفة الفعل كما في بعض لغات تتم فيها مطابقة الفعل مع أكثر من عنصر من عناصر إسناده، كمطابقة الفعل مع الفاعل والمفعول به كليهما، كما في لغة أهل "جورجيا".
6- ظاهرة الوصل	تصل العربية المعرفة ولا تصل النكرة، وهي في ذلك وسط ما بين اللغات الجرمانية على سبيل المثال، التي تصل المعرفة والنكرة، ولغات شاردة في هذا الخصوص مثل"الصينية"، وعدد من لغات القبائل الاسترالية، التي لا تعرف التركيبات الموصولة أصلاً.
7- النفي	وهي في ذلك وسط بين اللغات الموغلة في ظاهرة النفي، كاللغة الفرنسية، واللغات التي تكتفي في نفيها بأداة نفي واحدة كـ"الروسية" وتجمع العربية في نفيها أيضاً بين نفي الأداة ونفي الفعل (الفعل الناقص "ليس").
8- البناء للمجهول	البناء للمجهول في "العربية" بلا جملة فاعل، وهو الوضع الأصلي كما دلل على ذلك البعض، في نفس الوقت الذي يجوز فيه إضافة جملة الفاعل باستخدام كلمات مثل "بواسطة"، "على يد" مثال (حسم الأمر من قبل الحاكم). تخلو العربية كذلك من الشارد اللغوي فيما يخص البناء للمجهول، فليست "العربية" كـ "اللاتينية" ذات صيغ متعددة للبناء للمجهول،

	وليست خالية منه كما في بعض اللغات في غينيا الجديدة، وليست أيضاً كالهندية حيث يتم البناء للمجهول دون تبديل لموضع الفاعل والمفعول به.
9- حالات الفعل (درجة اليقين لدى المتكلم).	3 إعراب الفعل المضارع، أو درجة اليقين لدى المتكلم "بالعربية" حالات إعراب" وسط بين غياب الإعراب كما في "الإنجليزية" والإفراط في إعرابه كما في إعراب الفعل "الأيسلندية " (أكثر من ست حالات للإعراب).
10- أداة التعريف	التعريف في "العربية " باستخدام الأداة، أما التنكير فدونها (حيث يحل التنوين محل الأداة) وهي في ذلك وسط بين لغات كـ"الألمانية " و"الفرنسية" المسرفة في استخدام أدوات التعريف والتنكير المتصرفة، ولغات مثل لغة أهل "توجا لوج" التي لا تعرف الأداة: تعريفاً أو تنكيراً.
11- استخدام أدوات المصدرية للفعل.	في حين أن هناك أداة واحدة لمصدرية الفعل في الإنجليزية (to) وهناك 4 أدوات في الفرنسية (mone, pour, a,…) هناك أداتان فقط في العربية (أن، اللام).
12- استخدام الأداة كفاعل	تنقسم اللغات إلى لغات Do- language حيث يمكن أن تحل الأداة محل الفاعل في مثل The wind destroyed the house، ولغات Be- Lnaguage حيث لا يجوز أن تقع الأداة المسببة محل الفاعل مثل اليابانية حيث تستخدم بديلاً مكافئاً لـ"تحطم المنزل بفعل الرياح" ، أما في العربية فيمكن استخدام الصيغتين (حطمت الرياح المنزل- وتحطم المنزل بفعل الرياح).
13- استخدام الضمير	هناك بعض لغات مسرفة في إضمارها ومنها الإنجليزية وتسمى لغات سخانة Hot وهناك لغات مقلة في إضمارها كالصينية وتعرف بالباردة Cold وتتوازن العربية بين هذين الطرفين لذا يمكن أن نصفها بالدفء Warm.

إن التوسط النحوي الذي نزعمه، إنما قصدنا به لفت الانتباه إلى هذه الظاهرة

الهامة، التي تأخذ بعداً أعمق لو نظرنا إلى نحو "العربية " كحالة خاصة من النحو العام. إن هناك فرصاً متعددة لاستغلال خاصية التوسط هذه، لننهل من حصيلة ما تم على صعيد اللغات الأخرى، ننتقي منها ما يلائم نحونا، ففي التوسط اللغوي هداية لنحاتنا لتحسس أداء اللغة العربية مقارنة بأداء اللغات التي تتوسط بينها، أو التي تجمع بين بعض خصائصها.

ساد التوجه التوليدي Generative trend التنظير النحوي، وأعلن النحو التحليلي هزيمته الساحقة أمام أصالة النحو التوليدي، وقدرته على كشف بواطن "السر اللغوي"، يوضح الشكل 3 عائلة النماذج النحوية، التي تؤكد بما لا يدع مجالاً للشك مدى الغنى العلمي الذي فجره مفهوم التوليد، وكيف تعدد، الرؤى النظرية تهاجم المسألة النحوية من زوايا متعددة وبأهداف مختلفة.

لقد ظهرت، وما زالت تظهر، عدة نظريات نحوية، وهي تمثل النتاج الوفير للتفاعل الشديد بين النحويين والدلاليين من جانب، وبين اللغويين وعلماء الحاسوب من جانب آخر، توضح النماذج النحوية المضللة في الشكل تلك التي قام بوضعها لغويون وحاسوبيون، فمن ضمن أحد عشر نموذجاً نحوياً، هناك خمسة نماذج من وضع الحاسوبيين، ويدل هذا بوضوح على مدى التفاعل الشديد بين أهل اللغة وأهل الحاسوب في إغناء الفكر النحوي.

تتضمن شجرة التنظير النحوي النماذج النحوية التالية:

نحو تحليلي

- نحو اعتماديات Dependency Grammar.
- نحو طبقي Stratificational Grammar.

نحو توليدي

- نحو توليد تحويلي Tranformatioal Generative Grammar.
- نحو الحالات الإعرابية Case Grammar.
- نظرية الربط العاملي GB: Government Binding Theory.
- نحو وظيفي FG: Functional Grammar.

- نحو وظيفي معجمي LFG: Lexical Functional Grammar.
- نحو علاقي RG: Relational Grammar.
- نحو مقولي CG: Categorical Grammar.
- نحو شبكات الانتقال المعززة ATN: augmented Transitio Networks.
- نحو البنية العامة للجمل GPSG: Generalized Phrase Structure Grammar.
- نحو بنية الجملة المتعمد على الرأس HPSG: Head Phrase Structure Grammar.
- نحو ترابطي Unification Grammar.

تختلف هذه النماذج من حيث دقتها في توصيف الظواهر اللغوية وقدرتها على تفسيرها، ولا شك أن اختيار أمثلها لنمذجة اللغة العربية هي أمر يتطلب بحثاً متعمقاً في خصائص تلك النماذج باعتبار خصائص النحو العربي التي أسلفناها ، وأهم الشروط التي يجب أن يوفرها هذا النموذج فيما يخص العربية المكتوبة في رأي الكاتب هي الآتي:

– القدرة على التعامل مع اللبس الناجم عن غياب التشكيل.
- تغطية بدائل رتبة الكلمة الناجمة عن التبديل والتأخير.
- تقليل التوليد المسرف إلى أدني حد overgeneration.

-

- القدرة على كشف بنية الجملة من خلال استغلال ظواهر التآخي النحوي – الصرفي كالمطابقة (بين الصفة والموصوف، الفعل والفاعل) والتآخي النحوي– الدلالي في ارتباط الظروف بمعاني الأفعال، وذلك علاوة على نواحي الارتباط الأخرى مثل ارتباط الفعل أو مشتقه بالمفعول المطلق وارتباط الضمير بمرجعه وارتباط الحروف بمجرورها.

– مساندة عملية التفضيل التركيبي والدلالي preferntial semantice & syntans بهدف اختيار أنسب القرارات في حالة التعدد.

لقد بات الأمر في حاجة إلى مؤتمر قومي عربي لاختيار أنسب هذه النماذج في تناول منظومة اللغة العربية، ولا بد ألا يقتصر الأمر على نموذج واحد، فسبر أغوار منظومة اللغة العربية لن يتأتى إلا من خلال تعدد النظر إليها من مناظير وزوايا مختلفة، وأكاد أزعم أن اللغة العربية أكثر حاجة من غيرها في احتياجها لأكثر من نموذج وذلك بسبب التشابك الشديد بين فروعها المختلفة من نحو و صرف ودلالة وصوتيات، وتشابك هذه الفروع ذاتها مع منظومة المعجم.

رابعاً: نظام الإعراب الآلي للغة العربية المكتوبة

قام الباحث خلال عشر السنوات المنقضية بمشروع بحثي طويل الأجل يعاونه فيه فريق كبير من اللغويين ومهندسي الحاسوب بهدف تطوير سلسلة من المعالجات الآلية تتعامل مع العربية المكتوبة غير المشكولة ابتداء من الحرف حتى السياق المتصل (شكل رقم 4).

على مستوى الحرف تم تطوير أدوات برمجية لقراءة النصوص العربية آلياً Arabic- OCR، وكذلك لتوليد جماليات الخط العربي بأنماطه المختلفة calligraphical generator، وقد تم تطبيق أساليب متطورة كالتعرف على الأنماط pattern receognition وكذلك أساليب التعلم الذاتي automatic learning لتمييز أحجام الحروف وأنماطها تلقائياً، من أصعب المشاكل التي واجهها نظام القراءة الآلية للنصوص العربية آلياً هو الالتباس في تمييز علامات التشكيل والنقاط نظراً لصغر حجمها النسبي، وكذلك صعوبة التعامل مع النصوص ذات جودة الطباعة المنخفضة ، لقد تطلب حل هذه المشكلات اللجوء إلى معالجات صرفية ونحوية ودلالية متطورة بهدف مساندة نظام القراءة الآلية في حسم كثير من حالات الالتباس والخطأ.

بعد نظام القراءة الآلية الذي يتعامل أساساً مع اللغة باعتبار وحدة الحروف انتقل التطوير للتعامل مع عنصر الكلمة، وتم بنجاح تطوير معالج صرفي آلي للغة الحديثة والكلاسيكية، يمكن من خلال هذا المعالج تحليل أي كلمة عربية

بغض النظر عن مستوى تشكيلها (سواء تامة التشكيل أو خالية منه أو مشكلة جزئياً) إلى عناصرها الأولية من سوابق ولواحق وجذر وصيغة صرفية،ويعمل المعالج أيضاَ بصورة عكسية، بغرض تركيب الكلمة من هذه العناصر الأولية، وللتأكد من صحة عمل المعالج الصرفي تم اختباره على نص القرآن الكريم

والكتب التسعة للحديث النبوي الشريف، علاوة على كم هائل من النصوص بلغ عشرة ملايين كلمة تغطي نطاقاً واسعاً من الموضوعات والأساليب ومستوى الصعوبة.

تم استخدام المعالج الصرفي الآلي لتطوير قاعدة البيانات المعجمية للغة العربية حيث الصرف هو قلب المعجم العربي، والصرف بأكمله – كما يقول أندرسون – يكمن داخل منظومة المعجم حيث يمثل قواعد الفائضية المعجمية lexical redundancy، ونقصد بها تلك الخصائص التي يتكرر ظهورها بصورة مطردة أو شبه مطردة عبر كثير من المدخلات المعجمية.

بإكمال قاعدة البيانات المعجمية ونظام الصرف الآلي أصبح الطريق ممهداً لتطوير نظام الإعراب الآلي لتحقيق الأغراض الآتية:

– فك اللبس الصرفي– النحوي الناجم عن غياب التشكيل وبالتالي إكمال تشكيل الكلمات سواء تلك الخاصة بصلب الكلمة أو نهايتها الإعرابية.

- إعراب فقرات الجمل.
- فك اللبس الدلالي لمعاني الكلمات.
- ربط الضمائر بمراجعها وربط الظروف بأفعالها ومشتقاتها.
- استنباط المحذوف وربطه بما يسبقه من ألفاظ أو استنتاج ما يشير إليه بالعهدية.

ولتوضيح المقصود بفض اللبس الصرفي الناجم عن غياب التشكيل إليكم المثال الآتي وهو مثال مفتعل بغرض التوضيح فقط.

"عاش الرجل الأسود من أهل الجزر يأكل الجزر وتخشى الجزر".

باستثناء الفعل "عاش" جميع الكلمات الأخرى ملتبسة صرفياً ومعجمياً – قام نظام الإعراب الآلي الذي تم تطويره بفض هذا اللبس المركب وبالتالي أصبح ممكناً تشكيل الجملة تلقائياً كما هو موضح في الشكل رقم 5.

خامساً: قاعدة البيانات المعجمية:

لا حاجة بنا أن نؤكد الدور الهام الذي يلعبه المعجم في معالجة اللغة العربية آلياً بصفة عامة، وهو الدور الذي يزداد أهمية بصورة كبيرة فيما يخص

كان من أهم هذه البحوث هو وضع نظام متكامل للسمات الدلالية للغة العربية semantic features بدرجة من التفصيل تكفي لتوصيف جميع معاني الأسماء

والأفعال بدرجة الدقة المطلوبة التي يحتاجها نظام الإعراب الآلي. ويمكن النظر على هذه السمات بصفتها العناصر الذرية أو الأولية التي تكون معنى الكلمة.

لقد نجح فريق المعجميين في توصيف جميع المعاني الواردة في معجم اللغة العربية الحديثة بدلالة هذه السمات، وهو ما أكسبه النظام الآلي قدرة على الفهم الأتوماتي الضحل shallow understanding بالقدر الذي يكفي لإجلاء معظم حالات اللبس بأنواعه المختلفة.

وقد واجه النظام في بداية تطويره عقبة أساسية ألا وهي قصور المعاني التي تتضمنها المعاجم العربية، وعدم التفريق الدقيق بين المعاني المتعددة لنفس الكلمة. للتغلب على هذه العقبة تم بناء قاعدة بيانات ذخيرة النصوص Corpus من عشرة ملايين كلمة، تم تحليل نصوص هذه القاعدة باستخدام المعالج الصرفي الآلي، فتم فك لبس جميع مفرداتها صرفياً وإعرابياً، علاوة على تصنيف هذه المفردات وفقاً نظام دقيق لأقسام الكلم، وكذلك تم ربط الأسماء بمعلقاتها وكذلك الأفعال بمكملاتها وظروفها.

ويوضح (شكل رقم 6) المراحل المختلفة لتطوير المعجمية الحديثة ، ويتضح منه أن التقنيات المبينة على قواعد ذخيرة النصوص تمثل نقلة نوعية ساهمت بصورة أساسية في نقل المعجميات من كونها فناً وحرفة lexicography إلى كونها علماً منضبطاً lexicology، والأهم من ذلك تهيئة المعجم مادة وتنظيماً للمعالجة الآلية، وسينقلنا هذا إلى الحديث عن ضرورة توثيق اللغة العربية لفظاً ومصطلحاً واستخداماً.

في ظل النظم اليدوية يتم توثيق اللغات خلال مجموعة من المراجع المختلفة هي:

- المعاجم العامة.
- المعاجم المتخصصة.
- كتب قواعد النحو.
- مسارد المصطلحات.
- مكانز المترادفات والموضوعات.

- كتب المقتطفات والمأثورت اللغوية.

نتيجة للتطور الهائل في نظم الحاسبات الإلكترونية تمكن الباحثون من تجيمع كم هائل من البيانات اللغوية سواء على مستوى المعجم أو على مستوى الاستخدام الفعلي. إن نقص البيانات في الماضي قد عزل الباحث اللغوي عن لغة الواقع وجعله يميل إلى تقليص إشكالية اللغة لإمكانية استيعابها ذهنياً والإسراف في اللجوء إلى الحدس لدراسة كثير من الظواهر اللغوية المعقدة.

يسعى حالياً علم اللسانيات لاستدراك هذا الخلل المنهجي الرئيسي وذلك من خلال تناول اللغة الواقعية مع اتساعها متحرراً من أسر اللغة المثالية كما يتصور منظروها، لقد ظهر فرع جديد من علم اللسانيات النصية Corpus Linguistics قوامه استخلاص الدليل اللغوي evidence من واقع النصوص الفعلية وليس استناداً إلى مجموعة محدودة من القواعد المطردة وحالات استثنائها.

ويحتاج ذلك إلى تخزين كم هائل من النصوص اللغوية الفعلية تنتقى وفق أساليب إحصائية دقيقة بحيث تمثل اللغة أو الشريحة اللغوية رهن الدراسة.

تسعى كثير من الأمم حالياً إلى توثيق لغاتها سواء الحديثة أو القديمة من خلال استخدام الإمكانيات الهائلة لتكنولوجيا المعلومات عن طريق بناء قواعد نصوص ضخمة.

تتيح قواعد النصوص اللغوية كثيراً من المهام التي يتعذر على الأساليب اليدوية القيام بها:

– رصد مظاهر دينامية اللغة سواء في تغير معاني ألفاظها أو أساليب تراكيب جملها.

- بناء المعاجم استناداً إلى النصوص الفعلية حيث ثبت تعذر ملاحقة المعاجم التقليدية لما يطرأ على معاني الألفاظ واستحداثها.

- تحليل أساليب الكتاب وعلاقتها بأنواع الخطاب.

- استخلاص الخرائط المعرفية والمناحي الفكرية.

وكما قال قائل إن الصنعة المعجمية منذ نشأتها وحتى الآن قد نهجت مضطرة مساراً غير طبيعي، ويقصد بذلك سعيهم لربط المعنى باللفظ بعد انتزاعه من سياقه، ومحاولتهم المستميتة لاسترداد الطبيعية المفقودة من خلال تصور أسيقة مختلفة للفظ نفسه وذلك لتحديد أثر السياق على اختلاف المعنى الذي هو بحكم طبيعته وليد النص، لقد اكتفت معاجم الماضي بالاستشهاد بمقتطفات النصوص، أما حالياً فالأمر يختلف حيث تستخلص المادة المعجمية من قلب النصوص. تسعى الدراسة لعرض تجربة الكاتب في بناء أول قاعدة نصوص عربية باستخدام الحاسوب، وهي تبرز مدى الاختلاف بين إعداد قواعد النصوص العربية وقواعد النصوص الإنجليزية، مصدر الاختلاف الأساسي هو أن النصوص العربية الواقعية غير مشكلة وبالتالي يلزم إجلاء اللبس الناجم عن غياب التشكيل سواء باستخدام النظم اليدوية أو نظام التشكيل التلقائي الذي قام بتطويره الباحث.

لقد زادت أهمية المعجم في الآونة الأخيرة بصورة كبيرة بعد أن أصبح قاسماً مشتركاً في نظم معالجة اللغات الطبيعية آلياً، من تنسيق الكلمات ونظم استرجاع المعلومات وتحليل النصوص والفهرسة والاستخلاص الآلي وقد، حدث كثيرون عن أوجه قصور المعاجم العربية الحالية التي يلخصها شكل رقم 7.

كلمات ختام:

أترك القارئ ببيانين تاركاً له حرية الربط بينهما، أولهما مخطط زمني يوضح التطور التاريخي للمعجم العربية مقروناً بتطور نظيره الإنجليزي، وهو يوضح بصورة سافرة ريادتنا لعلم المعاجم وصيغتها وقد كنا نسبق أهل الإنجليزية بما يزيد على عشرة قرون، لقد أهدرنا هذا السبق الزمني الكبير وما خلفه لنا سلفنا العظيم، ويكفي أن أربعة قرون مضت بين ظهور الطباعة في منتصف القرن الخامس عشر وظهور أول معجم عربي مطبوع عام 1865–1948 وبداية استغلالنا له في تطوير معجم العربية مادة وتنظيماً واستخداماً، ولا خلاف على ضرورة أن نحافظ على لغتنا بالاحتفاء بتراثنا وتمحصيه واستخداماً وتوظيفه، في الوقت نفسه الذي يجب أن نتابع فيه بكل دقة ما يجري حالياً في مجال اللسانيات والمعجميات الحديثة، أما البيان الثاني فهو نص ورد في كتاب "علم اللغة التقابلي" للدكتور "أحمد سليمان ياقوت عميد كلية الآداب – جامعة بيروت العربية حيث يقول: لا يسعنا ونحن ندرس مرحلة

التوليد هذه إلا أن نتذكر ما كنا نسأل فيه ونحن في المدارس الابتدائية أو الثانوية "اجعل العبارة الآتية للمثنى المذكر ثم للجمع بنوعيه" هذا هو الرجل الذي أكل التفاحة، هؤلاء هن النسوة اللاتي أكلن التفاحة... إلخ".

أليست هذه الإجابة هي الآلة نفسها التي أشار إليها تشومسكي بأن تصور أن

هناك آلية لتوليد الجمل، لها نقطة بداية، وتتوقف عند نقطة نهاية، وتحتوي بين الاثنتين جملاً عديدة أليس إنتاجها هو إجابة ذلك السؤال الذي طرحناه منذ قليل؟ بلى هو كذلك ولكننا مغرمون بأشكال من القول زائفة وزخارف من الرسوم عديمة النفع، نوليها عنايتنا ونهتم بها حتى لا نهتم بأننا مقصرون عن اللحاق بالركب اللغوي في الغرب (ص48).

إلا يجب علينا أن نتصدي لمثل هذه القطيعة المعرفية، فما أحوجنا وقد عجزت عقولنا عن توليد علم جديد أن نعوض ما ينقصنا بالاقتراض الرشيد لما أنتجه غيرنا واستيعابه وتطويعه لمطالب لغتنا وتراثنا.