

الاستقراء اللغوي والمعالجة الآلية للغة

سليمان بن علي

قسم اللغة العربية وآدابها

جامعة الأغواط - الجزائر

يتعلق موضوع هذه الدراسة بمسألتين جد مهمتين، هما الاستقراء والمعالجة الآلية للغة، إذ نجد استقراء القدماء للظواهر اللغوية في جميع المستويات تقريبا استقراء غير تام على الرغم مما بذلوا من جهد عظيم في ذلك؛ مما قد يعني إغفال بعض الحقائق في وصف بنية العربية ونظامها الصوتي أو الصرفي أو التركيبي؛ ومن هذا المنطلق يجب أن نسعى جاهدين، ومن خلال المعالجة الآلية للغة وما تقدمه من دقة في البحث وسرعة في الإنجاز، إلى استدراك ما فات بعض القدماء في استقراءهم لبعض تلك الظواهر، وإعادة النظر في بعض أحكامهم الصادرة عن ذلك الاستقراء. ولكي لا نظلمهم فإن علينا أن نشير إلى مواطن الدقة في أعمالهم - وهي كثيرة - بعد أن أثبتت الإحصاءات الدقيقة التي يقدمها الحاسوب في كثير من الأحيان صحة ما ذهبوا إليه.

اتخذ اللغويون والنحاة العرب من ثنائية الزمان والمكان مقياسا هاما من مقاييس الاحتجاج(*) باللغة وأخذها عن أصحابها وتسجيلها في مدوناتهم الصوتية والصرفية والمعجمية والنحوية، ثم راحوا بعد ذلك يدرسون ما جاء في تلك المدونات ليقعدوا القواعد ويضعوا الضوابط والمقاييس التي ينبغي اتباعها في المحافظة على نقاء العربية على جميع المستويات المذكورة آنفا.

(*) تساءل الدكتور حلمي خليل عن مدى تأثير نظرية الاحتجاج من خلال تلك القيود الزمانية والمكانية والجنسية على المادة اللغوية للمعجم العربي، ثم ذكر أن الإجابة عن هذا التساؤل تنحصر في مصطلح واحد هو (المولد)، لأن القدماء عدوا المولد حارج حرم الفصاحة، وأغلقوا دونه أبواب الاستعمال والمعاجم على حد سواء، رغم أنه يجري على القياس والفصيح من جهة كونه ألفاظا عربية الأصل والصيغة أعطيت دلالة جديدة، وما ذلك إلا لأنهم لم يجدوا لهذه الاستعمالات شواهد فيما جمعه من أفواه عرب البادية (انظر: مقدمة لدراسة التراث المعجمي العربي ص 116-117)، والحق كذلك أن من ألفاظ المولد ما تم توليده صيغة ومعنى أيضا، لا معنى فحسب، والأمثلة على ذلك كثيرة نذكر منها كلمة (فطرة) بضم الفاء، بمعنى صدقة الفطر، فهذا المعنى موجود منذ شرعت تلك الصدقة، ولكن التعبير عنه كان بعبارة: صدقة الفطر، فولدت له صيغة (فعله) المذكورة للدلالة عليه، وهي لم تكن مستعملة فيه قبل ذلك.

ومن هنا ظهر التشدد في وصف كل كلام مخالف لكلام القبائل الفصيحة في نظرهم (باعتبار الزمان والمكان المذكورين على الأقل) بالخطأ، فوسم كثير من اللغويين والنحاة بالخطأ واللحن الكثير من الأداءات والصيغ والألفاظ، لا لشيء إلا لأن في اللغة ما هو أفصح منها وأعرف، أو لأن تلك الأداءات والصيغ والألفاظ مأخوذة عن قبائل متهمه في فصاحتها^(١). والأمثلة على ذلك كثيرة، حسبنا منها ما سنورده في عدة مواضع من هذه الدراسة.

والذي يعيننا في هذا المجال أن هناك من اللغويين من رأى أن من تكلم بلهجة من لهجات العرب، أو قاس عليها - ولو كانت نادرة أو رديئة - فهو مصيب لكلام العرب غير مخطئ له، ومن هؤلاء ابن جنبي الذي يقول: (فالناطق على قياس لغة من لغات العرب مصيب غير مخطئ، وإن كان غير ما جاء به خيرا منه)^(٢). وهو بذلك يرى أن كل ما جاء عن العرب فهو من كلامها، وأنك إذا جئت بشيء من ذلك مع وجود ما هو أفصح منه وأكثر شيوعا تعتبر مخطئا في اختيار الأحسن والأجود لا مخطئا لكلام العرب، يدلنا على ذلك قوله - في نفس السياق - عند حديثه عن الكشكشة والكسكسة وغيرها من الظواهر اللغوية التي عرفت بها بعض اللهجات العربية: (فإذا كان الأمر في اللغة المعول عليها هكذا وعلى هذا، فيجب أن يقل استعمالها وأن يتخير ما هو أقوى وأشيع منها. إلا أن إنسانا لو استعملها لم يكن مخطئا لكلام العرب لكنه كان يكون مخطئا لأجود اللغتين، فأما إن احتاج إلى ذلك في شعر أو سجع فإنه مقبول منه غير منعي عليه، وكذلك إن قال: يقول على قياس من لغته كذا وكذا، ويقول على مذهب من قال كذا وكذا)^(٣).

(١) انظر: نعمة رحيم، مناهج التصويب اللغوي. مجلة المورد العراقية. مج ٦، ع ١، ١٩٧٧، ص ١٤.

(٢) ابن جنبي، الخصائص. ت: محمد علي النجار. المكتبة العلمية. دون ط. ١٢/٢.

وهكذا تفتن ابن جنبي إلى الكثير من المسائل المتعلقة بجانب اختلاف اللهجات وضرورة مدارستها، فالف (المحتسب في شواذ القراءات والإيضاح عنها) باعتبارها من كلام العرب، كما درس ما سماه تداخل اللغات أو تركيبها، وذلك بأن يسمع الفصيح لغتين فصاعدا فيركب لغة ثالثة أو رابعة.. وهكذا، وقصة الصقر والسقر والزقر معروفة فلا حاجة إلى ذكرها، وقد ردّ من خلال ذلك كله على كثير من اللغويين والنحاة الذين حكموا على مثل تلك الاستعمالات وغيرها بالشذوذ، آخذين في ذلك بمعيار الأفصح، ويكفي - في هذا الصدد - أن نورد له النص التالي المقتطع من (باب في تركيب اللغات)، الذي يقول فيه: (اعلم أن هذا موضع قد دعا أقواما ضعف نظرهم وخفت إلى تلقي ظاهر هذه اللغة أفهامهم أن جمعوا أشياء على وجه الشذوذ عندهم، وادعوا أنها موضوعة في أصل اللغة على ما سمعوه بأخرة من أصحابها، وأنسوا ما كان ينبغي أن يذكره، وأضاعوا ما كان واجبا أن يحفظوه. ألا تراهم كيف ذكروا في الشذوذ ما جاء على (فعل يفعل) نحو: نعم ينعم، ودمت تدوم، ومت تموت. وقالوا أيضا فيما جاء من (فعل يفعل) وليس عينه ولا لامه حرفا حلقيا، نحو: قلى يقلى، وسلا يسلى، وجبى يجبى، وركن يركن، وقنط يقنط، ومما عدوه شاذا ما ذكروه من (فعل فهو فاعل) نحو: طهر فهو طاهر، وشعر فهو شاعر، وحمض فهو حامض، وعقرت المرأة؛ فهي عاقر، ولذلك نظائر كثيرة. واعلم أن أكثر ذلك وعامته إنما هو لغات تداخلت فتركبت على ما قدمناه في الباب الذي هذا الباب يليه، هكذا ينبغي أن يعتقد وهو أشبه بحكمة العرب) (١).

لقد ذكرت هذه الأمثلة والنصوص البسيطة لأبين مدى الاختلاف بين من تشدد في رفض كل كلام نطقت به العرب مما لا يتفق مع ما قعد من قواعد، وبين

(١) الخصائص ج: ١ ص: ٣٧٤.

من تساهل واعتبر لغة العرب أكثر من أن يلحن فيها أحد، أو كما أثر عن بعض القدماء قوله: (على ما سمعت من كلام العرب ليس أحد يلحن إلا القليل).
ومما لا شك فيه أن كل تلك الجهود قد قامت على استقراء لما تكلمت به العرب بتتبعها في حلها وترحالها، وأخذ الكلام عنها مشافهة، وربما مساءلتها عن كثير من قضايا اللغة. ولكننا - مع كل ما نكنه لعلمائنا القدامى من تقدير على ما قدموا للعربية - نجد بعض الثغرات التي مكّنت لإغفال شيء من الأداءات والصيغ والألفاظ في مدوناتهم التي جمعوها. وقد تمثلت أهم تلك الثغرات في عدم دقة حكمهم - في أحيان كثيرة - على إهمال العرب لبعض تلك الأداءات والصيغ والألفاظ، فغدا استقراؤهم بذلك استقراء ناقصا، حكمه بعضهم بصرامة في قبول اللغة أو رفضها من الناطقين بها خلال عصر الاحتجاج أو بعده، أو في مناطق خارجة عن نطاق من يحتج بلغتهم.

إن الحكم باستعمال العرب لهذا الأداء أو ذاك أو لتلك الصيغة دون غيرها أو إهمالهم إياها يقتضي تتبع القبائل التي يحتج بكلامها تتبعاً دقيقاً، بل ومعايشتهم دهرًا يمكن فيه تسجيل كل ما يتكلمون به في شتى الظروف، وأنى للغويين أو النحاة ذلك في ذلك الوقت، بل وأنى لنا ذلك في عصرنا هذا مع كل ما يتاح لنا من وسائل وأجهزة لم تخطر على بال المتقدمين؟!^(١). وقد بما قال الإمام الشافعي: (لسان العرب أوسع الألسنة مذهبا وأكثرها ألفاظا، ولا نعلمه يحيط بجميع علمه إنسان غير نبي، ولكنه لا يذهب منه شيء على عامته، حتى لا يكون موجودا فيها من يعرفه. والعلم به عند العرب كالعلم بالسنة عند أهل الفقه، لا نعلم رجلا جمع السنن فلم يذهب منها عليه شيء. فإذا جُمع علم عامة

(١) انظر: محمد حسن حنبل، الاستدراك على المعاجم العربية - في ضوء مئتين من المستدركات

الجديدة على لسان العرب وتاج العروس - دار الفكر العربي، القاهرة. ١٩٨٦، ص ١٧ بتصرف.

أهل العلم بها أتى على السنن، وإذا فُرِّقَ علمُ كل واحد منهم: ذهب عليه الشيء منها، ثم كان ما ذهب عليه منها موجودا عند غيره^(١)، وهكذا يستدرك اللاحق على السابق أيضا. وليس هذا الحكم قيميا لا يتصل بالحقائق اللغوية، بل هناك الكثير من الأدلة - التي سنسوق بعضها - لإثباته.

ومما يدل على ذلك استدراك بعض القدماء على بعض من تقدمهم الكثير من صيغ الكلمات العربية، فقد استدرك الزبيدي ٨٠ بناء من أبنية الأسماء التي بلغت عند سيبويه ٣٠٨ بناء، كما استدرك ٦ أبنية من أبنية الأفعال التي وصلت عند سيبويه إلى ٣٤ بناء. هذا على صعيد الأبنية أو الصيغ، أما على صعيد الجذور - التي تعتبر أمرا مهما في تبويب المعاجم بجعلها مداخل للكلمات المشتقة منها - فقد نبه الأزهري على كثير من المفردات التي عدها الخليل مهملة في معجمه، من ذلك: عه - عهك - عجه - عثق - عكش - كعت - جعز - عجب . . وغيرها، وهذا قد يؤثر بدوره على عدد الجذور والمداخل^(*) في المعاجم العربية، فقد جاء في الصحاح ٥٩١٨ جذرا، بينما بلغ مجموع ما تناوله لسان العرب (وهو يضم محتويات: التهذيب والحكم والصحاح والنهاية وتنبيهات ابن بري) ٩٢٧٣ جذرا، أي ما يقارب الضعف، وبلغ ما احتواه تاج العروس من الجذور ١١٩٧٨، ناهيك عما جاء في المعاجم التي أُلِّفت في الاستدراك على غيرها كفئات الجمهرة، والتكملة على الصحاح والصلة والقراح بتكمل الصحاح . . وغيرها. هذا عدا ما

(١) الشافعي، الرسالة. ت: أحمد محمد شاكر. المكتبة العلمية، بيروت. دون ط. ص ٤٢.

(*) المدخل Entry عبارة عن الوحدة التي ستوضع تحتها بقية الوحدات المعجمية الأخرى، أو المادة المعجمية التي تتألف عادة في المعاجم اللغوية من الكلمات المشتقة وغير المشتقة. وغالبا ما يتكون هذا المدخل في مثل هذا النوع من المعاجم من الجذر Root الذي يمثل البنية الأساسية للكلمات والمشتقات، وقد يرمز له في غير المعاجم بعلامة الجذر التربيعي في الرياضيات. وقد عرفت المعاجم العربية أنواعا وطرقا كثيرة في ترتيب المداخل (مقدمة لدراسة التراث المعجمي ص ٢١-٢٢).

استدركه المحدثون في ضوء التطور اللغوي الذي أصاب العربية مبني ومعنى^(١). إن كل ما سبق يعني ضرورة تضافر جهود جميع الأفراد والهيئات والمجامع للاقتراب من وصف شامل ودقيق لبنية اللغة العربية على جميع مستوياتها الصوتية والصرفية والتركيبية والدلالية، كما يعني إعادة النظر في بعض ما وصل إلينا من أحكام وقواعد حول بنية العربية ونظامها.

وقد أتيح لنا من أسباب هذا التضافر وتلك الدقة الشيء الكثير، فعلينا أن نستغل ما تمدنا به المناهج اللغوية الحديثة، وما تشهده من تطور في وسائل الدراسة، وعلى رأسها المعالجة الآلية للغة في جميع مستوياتها.

ولعل في دراسة الدكتور حلمي موسى، التي سنعرض بعض منجزاتها، خير دليل على ضرورة التماس المعالجة الآلية للغة كبديل أفضل للرحلة والمشاهدة التي عرفت قديماً، ليس على مستوى الاستدراك على القدماء فيما فاتهم فحسب، بل على مستوى تأكيد بعض ما وصلوا إليه من نتائج دقيقة في وصف العربية واستقراء خصائصها.

وحديثنا عن المعالجة الآلية للغة في العصر الحديث لا يعني بحال عدم تفتن القدماء لمعالجة اللغة معالجة رياضية إحصائية - والتي تقوم عليها تلك المعالجة الآلية في غالب الأحيان - فقد عُرِفَ الخليل ابن أحمد بإحصاء مفردات اللغة العربية بطريقة رياضية دقيقة، وكانت محاولته تلك - فيما نعلم - أول محاولة من نوعها في تاريخ العربية. وذلك باستخدامه نظريتي (التوافق والتبادل) لجميع الأحرف الهجائية بالنسبة لأحجام الكلمات المعروفة في اللغة العربية، وهي الثنائي والثلاثي والرباعي والخماسي؛ مما أنتج من الثنائي ٧٥٦ احتمالاً (أي حاصل ضرب ٢٧×٢٨)، وفي الثلاثي ١٩٦٥٦ احتمالاً (أي حاصل ضرب ٢٦×٢٧×٢٨)،

(١) انظر: الاستدراك ص ١٩-٢٠.

وفي الرباعي ٤٩١٤٠٠ احتمالاً (أي حاصل ضرب ٢٨×٢٧×٢٦×٢٥)، وفي الخماسي ١١،٠٠٠٠٠٠٠٠٠٠ (أي حاصل ضرب ٢٨×٢٧×٢٦×٢٥×٢٤). كما أن تجمع أحرف أي كلمة في الثلاثي أنتج عنده ست ٦ صور، وفي الرباعي أربعاً وعشرين ٢٤ صورة، وفي الخماسي مائة وعشرين ١٢٠ صورة^(١). وهكذا نجد القدماء قد تفتنوا إلى أهمية المنهج الرياضي - إن صح التعبير - في عملية الإحصاء اللغوي وفي غيرها من المسائل المتعلقة بذلك، والتي سنذكر بعضها فيما يستقبل من هذه الدراسة. غير أن الخليل صادفته هنا مشكلة دقيقة، وهي أن هذه الحصيلة بعضها مستعمل وبعضها مهمل، ولكي يسقط المهمل من حسابه اعتمد على خبرته الصوتية واللغوية بقوانين اللغة العربية، والحدود التي تسمح بها هذه القوانين^(٢). وربما هذا ما أدى به إلى إسقاط بعض ما تكلمت به العرب، مما استدرك عليه بعد.

ولكن يجدر بنا أن نشير هنا إلى أن الجديد في العصر الحديث تمثل في التوسع في القيام بالإحصاء اللغوي؛ نتيجة الاستعانة بالحاسوب، مما أصبح يدخل تحت ما يسمى بـ (المعالجة الآلية للغة)، وصار أحد المجالات التطبيقية المهمة لما يسمى بـ (علم اللغة الحاسوبي) الذي يمس عدة جوانب من اللغة كنظام اشتقاق الكلمة العربية، ومعالجة الكلمات والنصوص وأوزان الشعر، يقول حلمي خليل: (وقد أسفر اللقاء بين علم اللغة ونظرية المعلومات والحاسبات الآلية عن ظهور فرع من فروع علم اللغة المعاصر، هو علم اللغة الآلي Computational Linguistics، الذي يستخدم مفاهيم وتقنيات الحاسب الآلي Computer في التحليل اللغوي، مثل

(١) انظر: أحمد مختار عمر، البحث اللغوي عند العرب. عالم الكتب، ط٧، ١٩٩٧، ص ١٧٨ وما بعدها.

ويسري عبد الغني عبد الله، معجم المعاجم العربية. دار الجيل، بيروت. ط ١، ١٩٩١، ص ٩٠ وما بعدها.

(٢) معجم المعاجم العربية ص ١٠٣.

تخليق الكلام Speech synthesis، وتعلم اللغة سواء لأبنائها أو لغير الناطقين بها، كما أحدث هذا العلم نقلة نوعية ضخمة في وضع المعاجم وتأليفها وكذا ما يسمى بالبنوك اللغوية (Language Bank)^(١)، وسنقتصر في عرض عمل الدكتور علي حلمي موسى - أستاذ الفيزياء النظرية سابقا بجامعة الكويت - على بعض الجوانب الهامة التي تتصل بإحصاءات الحاسوب لجذور اللغة العربية، ومقارنتها ببعض ما توصل إليه القدماء في هذا الشأن.

لقد قسم الدكتور علي حلمي دراسته إلى أربعة أقسام، تناول في القسم الأول الجذور الثلاثية في صحاح الجوهري، وفي القسم الثاني الجذور غير الثلاثية فيه، وفي الثالث جذور لسان العرب، أما الرابع فجعله جذور تاج العروس. وحسبنا هنا أن نتطرق إلى بعض ما جاء في الجزء الثالث من هذه الدراسة، وهو الخاص بالمعالجة الآلية لجذور لسان العرب لابن منظور، ولعل أول ملاحظة يمكن إبدائها - قبل ذكر بعض عمله وطريقته فيه - هو أنه استبعد بعض الجذور لقلتها، على حد تعبيره، وهو عمل ربما لا يقبل به بعض الدارسين لما سينجم عنه من نتائج، لأننا عند دراسة ترددات الحروف - كما فعل هو - سنجد بعض الحروف لا يتردد مع بعضها الآخر فيما وضعناه من جذور قيد الدراسة، ولعلنا نجد نسبة - ولو ضئيلة - من هذا التردد فيما أبعدناه من جذور!، خاصة وأن الدكتور حلمي قد أبعد بعض الجذور على أكثر من مستوى (مستوى الجذور الثنائية، ومستوى ما زاد عن خمسة أحرف)، يقول في ذلك كله: (أورد ابن منظور في معجمه جذورا تختلف من جهة عدد حروفها، فمنها الثنائية، وهي قليلة جدا، ومنها الثلاثية، وتمثل الغالبة، وتليها الرباعية ثم الخماسية، ويندر وجود السداسية وما فوقها، والجذور الثنائية يقل عددها عن العشرين، أي أنها تمثل جزءا من ٥٠٠ جزء من جذور

(١) حلمي خليل، مقدمة لدراسة التراث المعجمي العربي. دار المعرفة الجامعية، الإسكندرية، ٢٠٠٣، ص ٢٩.

معجم لسان العرب . وقد استبعدت هذه الجذور من الإحصائيات لهذا السبب . . .
 أما ما زاد على خمسة أحرف، وهو يبلغ حوالي ست عشرة كلمة، فهو في أغلب
 الأحوال من أصل غير عربي، كما أنه يمثل نسبة ضئيلة جدا من مجموع كلمات
 المعجم، لذا فقد رأيت استبعاد ذلك من الإحصائيات، وعلى هذا أصبحت
 الدراسة مقتصرة على الجذور الثلاثية والرباعية والخماسية الواردة في معجم لسان
 العرب، وهي تمثل أكثر من ٩٩,٦٪ من جذور المعجم^(١).

وما يمكن ملاحظته عليه في مقولته هذه - إضافة إلى ما سبق - خطأه في استبعاد
 الكلمات التي رأى أن أغلبها من أصل غير عربي، وهو خطأ منهجي، إذ يجوز أن
 تغير مراعاة مثل تلك الكلمات من نتائج إحصاء تردد بعض الحروف، خاصة إذا
 علمنا أن العرب غيرت بعض تلك الكلمات بالنقص من حروفها أو استبدالها
 بغيرها مما تنطق به، وجرت إلى تخفيف ثقل العجمة، واستعملتها في أشعارها
 ومحاوراتها، حتى جرت مجرى الصحيح، وحتى وقع بها البيان، فنزل بها
 القرآن^(٢)، وبهذا الاعتبار عالج الجواليقي اختلاف مذهبي العلماء في ورود
 كلمات أعجمية في القرآن الكريم فقال: (وكلاهما مصيب إن شاء الله، وذلك أن
 هذه الحروف بغير لسان العرب في الأصل، فقال أولئك على الأصل، ثم لفظت به
 العرب بالسنتها، فعربته، فصار عربيا بتعريبها إياه، فهي عربية في هذه الحال،
 أعجمية الأصل، فهذا القول يصدق الفريقين جميعا)^(٣)، وقد عالجت هذه المسألة
 بما يغني عن إعادة جزئياتها هنا، وذلك في مقالة لي بمجلة الآداب واللغات الصادرة
 عن قسم اللغة العربية وآدابها بجامعة الأغواط، وذلك ضمن عددها الأول فليرجع
 إليه .

(١) علي حلمي موسى، إحصائيات جذور معجم لسان العرب، مطبوعات جامعة الكويت، ١٩٧٢، ص ٩٠.

(٢) انظر: القرطبي، الجامع لأحكام القرآن. دار إحياء التراث، بيروت. ١٩٨٥، ١/٦٨-٦٩.

(٣) الجواليقي، المعرب. ت: خليل عمران المنصور. دار الكتب العلمية، بيروت. ط ١، ١٩٨٩، ص ٦.

وقد قدم الدكتور حلمي بين يدي عمله هذا عدة جداول لترددات الحروف كل على حدة، وترددات بعضها مع بعض، وغير ذلك مما يمكن تلخيصه فيما يلي^(١):

١- ترددات الحروف بحسب مواقعها:

بيّن في الجدول الخاص بذلك تردد حروف الجذور الثلاثية في الموقع الأول والموقع الثاني والموقع الثالث، ثم عيّن مجموع مرات التردد لكل حرف والنسبة المئوية لهذا التردد، فنجد مثلاً أن تردد حرف الراء هو ٣٤٢ مرة في الموقع الأول، و٤٢٣ مرة في الموقع الثاني، و٤٤٠ في الموقع الثالث، ويكون المجموع حينئذ ١٢٠٥، والنسبة المئوية هي ٦٤,١٤٤٪.

٢- الترتيب التنازلي لتردد الحروف في كل موقع على حدة:

ظهر في الجدول الخاص بهذا العمل أن حرف النون في الموقع الأول هو الأكثر تردداً، إذ تردد ٣٩٧ مرة، يليه حرف الواو الذي تردد ٣٥٦ مرة، ثم الراء بـ ٣٤٢ مرة... وهكذا.

وظهر أن حرف الواو في الموقع الثاني هو الأكثر تردداً بـ ٤٧٣ مرة، يليه الراء بـ ٤٢٣ مرة، ثم الياء بـ ٣٨٥ مرة... وهكذا.

كما ظهر حرف الألف في الموقع الثالث في أعلى القائمة بـ ٥٢٦ مرة، يليه الراء بـ ٤٤٠ مرة، ثم الميم بـ ٤٢٩ مرة... وهكذا مع بقية الحروف.

٣- توزيع الجذور الثلاثية بالنسبة للحرفين الأول والثالث:

نجد في هذا الصدد أن الثاء وردت كأول مع الهمزة كثالث ٠٤ مرات، ومع الباء ٠٩ مرات، مع التاء ٠٥ مرات، ومع الثاء ٠٢ مرتين، ومع الجيم ٠٨ مرات. وأنها لم ترد مع واحد من الأحرف التالية كالثالث: ح- ذ- ز- س- ص- ض- ظ- ك- و- ي.

٤- توزيع الجذور الثلاثية بالنسبة للحرفين الأول والثاني:

(١) انظر: إحصائيات جذور معجم لسان العرب ص ٩٣ وما بعدها.

يظهر في هذا التوزيع وقوع الثاء كأول مع الهمزة كثنان ٠٨ مرات، ومع الباء ١٠ مرات، ومع التاء ٠٥ مرات.. إلخ. ولم ترد معها في الموقع الثاني الحروف التالية: ث - ذ - ز - س - ش - ص - ض - ظ - ا.

هذه باختصار شديد بعض ترددات الحروف التي تظهرها بعض جداول الدراسة، والتي بلغت ١١٤ جدولاً موزعة بحسب مواقع جميع الحروف وتردداتها وكذا صفاتها.. وغير ذلك من المسائل المهمة في دراسة بنية العربية ونظامها الصوتي.

ولعل أهم شيء يمكن أن نقف عليه من خلال هذه الدراسة التي قدمها الدكتور حلمي تناوله لبعض الأصوات اللغوية كمجموعات، يمثل بعضها مجموعة الأصوات الحلقية (أ - ح - خ - ع - غ - هـ)، والفموية (ج - غ - ق - ك)، والشفوية (ب - ف - م)، كما يمثل بعضها الأصوات المتقاربة المخرج مثل (د - ذ - ز - س - ص)، وقد أظهرت الإحصاءات التي قدمها أن كل مجموعة منها نادراً ما تتتالي حروفها بعضها مع بعض، إذ بدأ تتابع المجموعة الأخيرة على النحو التالي:

	ش	س	ز	ذ	
ذ	-	١	-	١٧	
ر	-	-	١٩	١	
س	-	٢٠	-	٤	
ش	١٩	-	١	١	

ومن الواضح أن تردد الحرف مع نفسه كثير، مقارنة مع تتابعه - الذي يكاد يكون نادراً - مع غيره من أفراد المجموعة.

إن النتيجة المتوخاة من مثل هذه المعالجة الآلية لتردد مثل هذه الحروف هي الوقوف بدقة على ما ذكره بعض القدماء حول عدم فصاحة الكلمات المفردة،

الراجع إلى تنافر الحروف، إذ اختلف القدماء حول سبب هذا التنافر فهو القرب الشديد في المخرج أم البعد الشديد أم هما معا؟(*) وقد تبين من خلال الإحصاء - وهو ما لخصه الجدول السابق وغيره من الجداول الأخرى - أن التنافر في قرب المخرج لا بعده، وهو ما أشار إليه الخليل حينما شرع في الخصائص الفونولوجية لبناء الكلمات العربية من جهة دخول بعض الأصوات اللغوية في بناء معين أو عدم دخولها، فوجد أن تقارب المخرج بصورة عامة من العلل المانعة لدخول حرفين من حيز واحد في بناء واحد، وذلك على النحو التالي:

- العين والحاء لا يجتمعان في كلمة واحدة لقرب مخرجيهما، إلا أن يشتق من ذلك فعل من جمع بين كلمتين أو أكثر مثل (حيعل).

- العين مع الغين، والهاء مع المهملات لأنهن من حيز واحد.

- القاف لا تجتمع مع الكاف في كلمة واحدة، وتأليفهما معقوم في بناء العربية لقرب المخرج.

- القاف والجميم كيف قلبتا لم يحسن تأليفهما إلا بفصل لازم...

وغير ذلك من المعايير التي حكم بها على المهمل على مستوى البنية الصوتية للكلمة دون المعنى^(١). ولعل هذا ما جعل بعض من جاء بعده يستدرك عليه بعض المواد التي أهملها لذات السبب ولم يتقص أمر استعمالها عند العرب.

كما تظهر أهمية المعالجة الآلية للغة في تصحيح بعض الأحكام الخاطئة لعلماء اللغة القدماء، إذ ذكر ابن منظور مثلا تردد بعض الحروف كما يلي: أ - ل - م - ه - و -

(*) ذكر الرماني أن الفصاحة في تعديل الحروف، وأن تعديل الحروف يكمن في الابتعاد عن التنافر، وأن التنافر - حسب رأي الخليل بن أحمد إما في القرب الشديد لمخارج الحروف أو البعد الشديد، غير أن ابن سنان الخفاجي يراه في القرب الشديد، مستدلا بكلمة (الم) غير متنافرة الحروف على الرغم من التباعد الشديد في مخارج حروفها. (انظر: البحث اللغوي عند العرب ص ٩٦-٩٧).

(١) انظر: مقدمة لدراسة التراث المعجمي ص ١٣٥.

ي- ن . في حين أظهرها الحاسوب على النحو التالي : ر- ل - ن - ب - م - ع - ق .

ومن هنا يتضح لنا أن ابن منظور قد أخطأ في تردد حرف الراء .

وهذه المعالجة تعطي على صعيد آخر بعض النظريات القديمة مصداقية وثباتا، وتجعلنا نقف باعتزاز أمام ما أنجزه القدماء من علماء العربية رغم قلة الوسائل إن لم نقل انعدامها التام، فقد أثبتت الإحصاءات التي قدمها الدكتور حلمي ضرورة اشتمال الكلمة الرباعية أو الخماسية الأصل على حرف من حروف الذلاقة (وهي : ر- ل - ن - ف - ب - م) (*)، وأن النون مثلا لا تكون بعدها راء في كلمة عربية، ولا تجتمع الجيم والقاف في كلمة عربية الأصل كذلك .. وغير ذلك من الأحكام التي ذكرها القدماء في دراستهم لبنية الألفاظ في اللغة العربية^(١).

وهناك غير هذا نتائج جد مهمة يمكن استخلاصها من المعالجة الآلية للغة، خاصة ما يتعلق ببعض الأحكام الصادرة عن القدماء فيما يخص بنية العربية صوتيا وصرفيا، بل حتى في دراستهم للعروض وموسيقى الشعر وتردد بعض القوافي دون غيرها، ومدى مطابقة كل ذلك للواقع. ناهيك عن توفيرها جهد سنوات طويلة من العمل الشاق، واستخلاصنا من خلالها الكثير من النتائج في زمن يسير.

والمعالجة الآلية للغة - إضافة إلى ما سبق - تعين بعض فروع علم اللغة بكثير من النتائج التي تحتاجها تلك الفروع في بعض المسائل الإجرائية، كعلم المفردات Lexicology الذي استقر على عدة موضوعات، يهمننا منها هنا:

- حصيلة المفردات التي يستخدمها المتكلم أو الشاعر أو الكاتب بلغة ما .

- مقدار الثروة اللفظية في لغة معينة وإحصاؤها .

(*) وهو ما ذكره الخليل حين رأى أن هذه الحروف أكثر دورانا في بنية الكلمة العربية لسهولة النطق .

(انظر : مقدمة لدراسة التراث المعجمي ص ١٣٦) .

(١) انظر: السيوطي، المهذب فيما وقع في القرآن من المعرب . ت: محمد التونجي . دار الكتاب العربي،

بيروت . ط ١، ١٩٩٥، ص ١٠ وما بعدها .

– مجموعة المصطلحات التي تستعمل في دائرة علمية أو فنية محدودة .
– إحصاء ومقارنة الكلمات المستعملة في عدة لغات طبقا لاحتياجات المتكلمين بها وغالبا ما يستعمل هذا العلم الإحصاء اللغوي الآلي وسيلة من وسائله الهامة^(١).

ونحن في الوقت الراهن في أمس الحاجة إلى استغلال هذا التطور الهائل في وسائل الدراسة لمعالجة قضايا العربية وتطويرها نحو الأفضل والأصلح لنا ولأجيال المتعلمين والباحثين .

(١) انظر: مقدمة لدراسة التراث المعجمي ص ٧٠-٧١ .