المبحث الثاني

قياس خاصية تنوع المفردات

في الأسلوب

عند العقاد والرافعي وطه حسين

المبحث الثاني
قياس خاصية تنوع المفردات في الأسلوب
عند العقاد والرافعي وطه حسين

الفاتحة

المعجم الذي يستخدمه الكاتب أو الشاعر هو من أبرز الخواص الأسلوبية الدالة عليه، والمبينة عن سر صناعة الإنشاء عنده، لذلك يؤدي فحص الثروة اللفظية *vocabulary richness* كما تظهر في النصوص إلى استبانة واحد من أهم الملامح المميزة للأسلوب، فما المفردات إلا الخلايا الحية التي يتحكم المنشىء في تخليقها وتنشيط تفاعلاتها على نحو يتحقق به للنص كينونته المتميزة في سياق النصوص وللمنشىء تفرده بين المنشئين.

وتختلف الثروة اللفظية بين المنشئين من جهتين:

الأولى : حجم الثروة اللفظية ***vocabulary size*.**

والثانية: طرق استخدام هذه الثروة اللفظية والتصرف فيها عند صياغة النص.

ومن المتوقع عند الموازنة بين عدد من الأساليب أن يمتاز بعضها من بعض من إحدى الجهتين أو كلتيهما. وينبغي عند قياس حجم الثروة اللفظية لدى المنشىء أن نميز بين نوعين مختلفين منها:

أولهما: قياس حجم ما يعرفه المنشىء من ألفاظ، و - بعبارة أخرى - كَمّ الألفاظ الذي يمكنه أن يتعرف إليه فيما يقرأ أو يستمع إليه من نصوص اللغة.

وثانيهما: قياس كَمّ ما يستخدمه من الألفاظ بالفعل في صياغة نصوصه.

وبدهي أن كَمّ ما يعرفه من ألفاظ سيكون أكبر بكثير من الكَمّ المستخدم منها، فالدارس المختص بالأدب الجاهلي والقارىء الشغوف به كلاهما يمكنه التعرف إلى كثير مما تحفل به نصوص الجاهليين من غريب، لكنا نشك كثيراً في أن أحدهما سيستخدم أكثر هذا الغريب إذا أراد أن يصوغ شعراً أو نثراً. وثمة طرق إحصائية تصطنع لقياس النوع الأول تقوم على استنطاق المنشىء نفسه مباشرةً بأن يطلب إليه الباحث الإجابة على عدد من الأسئلة توضع بحيث يمكن من خلالها تقويم الثروة اللفظية لديه، وأما النوع الثاني فطريقنا إليه هو فحص النص الأدبي بعد أن يفرغ منه المنشىء. وهذا الفحص مفيد من جهتين:

الأولى : أنه يُعِينُ على معرفة جانب من أهم جوانب صناعة الإنشاء عند المنشىء، والكيفيات التي يتصرف بها في ثروته اللفظية.

الثانية: أنه يصلح - عند الموازنة بين أكثر من نص لأكثر من منشىء - مؤشراً دالاً على تمايز أساليب منشئيها من حيث الزيادة والنقص في كَمّ الثراء المعجمي بوجه عام، ومن البدهي أن الحكم الذي نصل إليه في هذه القضية سيكون حكماً نسبيّاً لا مطلقاً.

وتنوع المفردات ***vocabulary diversification*** هو من الخصائص الأسلوبية التي يمكن التوصل بقياسها في عدد من النصوص إلى إجابة مدعومة بالدليل الإحصائي على سؤالين مهمين:

الأول : أي هذه النصوص يعبر عن ثراء معجمي نسبي إذا ما قورن بغيره؟

الثاني: كيف يستخدم المنشىء خاصية التنويع بين مفرداته عند صياغة النص؟

ومقتضى وصفنا لتنوع المفردات بأنه من بين الخصائص الأسلوبية يعني أنه ليس بالخاصية الوحيدة أو الحاسمة في مجال التمييز بين الأساليب. كما

يعني أيضاً أن الخصائص الأسلوبية من الكثرة والتعدد بحيث ينبغي اعتبار هذا التعدد إذا ما أريد تحديد العلاقات بين الأساليب المتنوعة على نحو أوسع وأشمل. وستتضح هذه المقولة للقارىء فيما يأتي من مناقشة في الفقرة الخامسة من هذا المبحث.

ويهدف هذا البحث إلى تقديم عرض نظري لإحدى الطرق الإحصائية المستخدمة في قياس خاصية تنوع المفردات مع دراسة نصية لنماذج من الكتابة العربية، وستعالج الدراسة المسائل الآتية على الترتيب:

١ - تحديد العيِّنات التي أجريت عليها البحث

٢ - عرض للمقياس وطريقة تطبيقه على العينات

٣ - طرق حساب نسبة التنوع

٤ - نتائج القياس

٥ - ملاحظات على نتائج القياس

٦ - العلاقة بين خاصية التنوع وصعوبة الأسلوب

١ - العيِّنات

يتناول هذا البحث بالدراسة ثلاثة نماذج لثلاثة من أعلام الأدب في العصر الحديث هم: عباس محمود العقاد ومصطفى صادق الرافعي وطه حسين، وقد آثرنا هؤلاء الأعلام بالدراسة لأسباب منها:

اولاً: أن الثلاثة هم من أبرز الأعلام العرب في صناعة النثر، ومن ثم كان تأثيرهم في مجال الفكر والثقافة من جهة، وفي فن الكتابة والأسلوب من جهة أخرى عظيماً، وكان لأدبهم نفوذه القوي وانتشاره الواسع - على تفاوت - بين قراء العربية والمختصين بآدابه.

ثانياً: أن أدب هؤلاء الأعلام الثلاثة قد حظي في أبعاده الفكرية والحضارية من الدارسين والمختصين بأكبر اهتمام، على حين ظل الجانب

الأسلوبي مهملاً أو شبه مهمل، فلم يقع لنا - في حدود ما قرأنا - دراسة أخلصت نفسها لهذا الجانب لديهم، إلا من نتف مبثوثة هنا وهناك هي في الغالب أحكام ذاتية صيغت في عبارات مرنة توهم دون أن تحدد، وتُلمِحُ دون أن تبين.

ثالثاً: أن بعض الأحكام الذاتية التي اشتملت عليها دراسات سابقة قد تكون ذات قيمة نقدية عالية ولا شك، وذلك لصدورها عن أدباء ودارسين أكثرهم ذو أرومة عريقة في صناعة الأدب، بيد أن صياغتها في عبارات ذوقية وجدانية تجعل من الصعب على القراء والدارسين تحديد المراد منها، وعلى كتابها أن يجيبوا إذا سألَتَهم على صدقها البرهان، ومن ثم كانت المعالجة الموضوعية ضرورة يمكن بها تفسير هذه الأحكام، أو نقض ما لا يقوم عليه منها دليل.

رابعاً: أن الشعر قد استأثر باهتمام الجمهرة من دارسي الأدب العربي ونقاده، لذلك رأينا أن نسهم بجهد متواضع في رفع الغبن الذي حاق بصناعة النثر في العربية، حين لحظنا أن مكتبة الدراسات الأسلوبية تشكو ندرة الأعمال التي اتخذت من نصوص النثر موضوعاً لها، على أن كثيراً مما كتب في هذا المجال إنما يقوم دليلاً واضحاً على أزمة المنهج التي يعاني منها درس الأدب العربي، والتي لا مخرج له منها - فيما نحسب - إلا بالإفادة من مناهج اللسانيات الحديثة عامة، والدراسات الأسلوبية خاصة على ما أسلفنا بيانه في غير هذا المكان[1] وقد شملت العينات الثلاث:

(١) الجزء الثاني من وحي القلم للرافعي[2]

وقد اخترنا منه مقالاً بعنوان: «سمو الفقر في المصطلح الاجتماعي الأعظم»، وهو مقال كتبه الرافعي في ثلاث حلقات متتابعات، لهذا كان اختيارنا له، فهو أول مقال في هذا الجزء يمكنه أن يستغرق العينة المطلوبة، تلك التي حددناها بثلاثة آلاف كلمة استغرقتها - على وجه التقريب - الحلقتان الأولى والثانية من المقال.

(٢) عبقرية محمد للعقاد[٣]

وقد اخترنا منه ثلاثة الآلاف الأولى من كلمات الكتاب.

(٣) «الفتنة الكبرى»: «عثمان» لطه حسين[٤]

هنا أيضاً تم اختيار ثلاثة الآلاف الأولى من كلمات الكتاب، وبذلك بلغ مجموع العينات الثلاث تسعة آلاف كلمة، وهو كَمٌّ لا بأس به في مجال دراسة تنوع المفردات.

ويُلْحَظُ أن النماذج المختارة تنتمي جميعها إلى مجال الأدب الإسلامي. وفيها يعالج الكاتب أحداثاً وشخصيات من التاريخ الإسلامي على طريقته الخاصة التي ينفرد بها، وعلى أي حال فإن تشابه الموضوع العام للعينات هو شرط تحسيني وليس شرطاً من شروط الصحة، ذلك لأننا لا ندرس دلالة كلمات بعينها ولا نعالج شكلها اللغوي في النصوص، ولكنا نفحص خاصية التنوع في المفردات التي يستخدمها الكاتب في نصوصه أيّاً ما كان الغرض الموضوعي الذي تنتمي إليه هذه المفردات. ومن المتوقع أن تتماثل نسب التنوع في أسلوب الكاتب بقطع النظر عن تغير الموضوعات التي يتناولها.

وتبقى لنا مسألة مهمة في هذه الفقرة تختص ببيان ما نعنيه بمصطلح «الكلمة» الذي اعتمدناه في تحديد كمِّ العينات المختارة، والحق أن تحديد هذا المصطلح استنفد جهداً كبيراً من اللسانين المحدثين حتى استيأس بعضهم فحمله اليأس على الشك في وجودها، وعدّها «خرافة علم اللغة»[٥].

ويمكن أن نرد هذا الخلاف - فيما نظن - إلى أنهم يحاولون تعريفاً لمفهوم الكلمة في اللغة المنطوقة، حيث تتحقق فيزيائياً في هيئة كم متصل *continuum* من المؤثرات الصوتية المتنوعة. وتجزئة هذا الكم المتصل إلى الوحدات التي يتكون منها تكون غالباً موضوعاً للاجتهاد والاختلاف. أما على مستوى اللغة المكتوبة فقد تولى العرف وتقاليد الرسم الكتابي تقديم تحديد

للكلمة صادر عن منطق اللغة الخاص. وفي الكتابة تظهر الكلمة على هيئة مجموعة من الحروف تتصل خطّاً، ويفصل بينها وبين ما سواها فراغ أوسع نسبيّاً من كلتا الجهتين[٦]. وقد اعتمدنا في هذه الدراسة هذا التعريف لأن الاختلاف عليه قليل من ناحية، ولأننا ندرس بالفعل نصّاً مكتوباً لا مقروءاً من ناحية أخرى. وإذن فليس المعتمد هنا على مصطلح الصوتيم *phoneme* وحدةً لتعيين الكلمة، بل على الحرفيم *grapheme*[٧]. وتكون الكلمة بهذا المفهوم سلسلةً من الحرفيمات المتصلة في الرسم، يفصل بينها وبين ما يسبقها وما يلحقها فراغ أوسع نسبيّاً، ولهذا الفراغ دلالته على استقلال هذه المجموعة في تقاليد الكتابة، وهذا الفراغ بهذه الدلالة يشكل حرفيماً من حرفيمات اللغة العربية المكتوبة يمكن تسميته حرفيم المفصل *juncture grapheme*.

ذلكم أهم ما يتعلق بالعينات التي اخترت للدراسة، أما طريقة قياسها فهي موضوع الفقرة الآتية.

٢ - القياس:

هناك عدة مقاييس اقترحت خاصية تنوع المفردات[٨] ومن أهمها ما اقترحه و. جونسون في دراسة بعنوان «**اللغة والعادات السليمة في الكلام**»[٩] وكتاب «**الناس في المآزق**»[١٠]، وفيهما يرى جونسون أن في الإمكان إيجاد نسبة لتنوع المفردات في النص أو في جزء منه إذا ما حسبنا فيه النسبة بين الكلمات المتنوعة (أي المختلفة بعضها عن بعض) والمجموع الكلي للكلمات المكوِّنة له.

ويطلق جونسون على الكلمات المتنوعة مصطلح «**الأنماط**» *types* وعلى المجموع الكلي للكلمات مصطلح «**التحققات**» *tokens*، ومن ثم يُطلَق على نسبة التنوع ***type-token ratio*** (وتختصر عادة إلى ***(TTR)***).

ويقتضي هذا المقياس أن ندخل في دائرة الكلمات المتنوعة كل كلمة

جديدة ترد في النص - أو في بعض أجزائه - لأول مرة، مع احتسابها مرة واحدة في العدد مهما تعددت مرات ورودها في الجزء الذي نفحصه من النص. وتُعَد مثل هذه الكلمة «نمطاً» *type*، وبعد إحصاء عدد الكلمات المتنوعة «الأنماط» يتم إيجاد نسبة التنوع بقسمة عددها على حاصل الجمع الكلي للكلمات «التحققات» *tokens*.

وواضح أن التوصل إلى عدد الأنماط في نص ما ليس أمراً بالغ السهولة، فقد اقتضانا تقريب هذه العملية وبيان مراحلها للقارىء - ذلك بالنسبة لكل عيِّنة - أن نقوم بما يأتي:

(١) عمل نموذج لجدول تكون عدد خاناته حاصل ضرب ١٠ × ١٠ وبذلك يصل مجموع الخانات في الجدول الواحد ١٠٠ خانة (انظر النموذج في الجدول رقم (١))[١١].

(٢) استنساخ العيِّنة كلها في هذه الجداول، بحيث تكتب كل كلمة في خانة مستقلة، وبذلك استغرقت العيِّنة الواحدة (وهي التي تتكون من ثلاثة آلاف كلمة) ٣٠ جدولاً.

(٣) حصر الأنماط في كل جدول على حدة، وذلك بمراجعة أول كلمة من كلماته على سائر الكلمات الباقية فيه وعددها ٩٩ كلمة، ثم شطب أي تكرار لهذه الكلمة يمكن أن يوجد في حدود الجدول الواحد. ثم نبدأ بعد ذلك بمراجعة الكلمة الثانية فيه بالطريقة السابقة على الكلمات الباقية (وسيصير عددها ٩٨ كلمة) حتى تنتهي جميع المئة.

ثم نقوم بمثل ذلك في سائر الجداول الأخرى وعددها بالنسبة للعينات الثلاث ٩٠ جدولاً.

(٤) الكلمات التي بقيت دون شطب تمثل ما نعنيه بالأنماط، وهذه يتم حصرها وكتابة عددها أسفل كل جدول.

بيد أن الخطوات الأربع السابقة تؤدي إلى حصر الأنماط في كل جدول على حدة. وهذا أمر مطلوب كما سنرى بعد؛ ولكنها لا تحصر الأنماط بالنسبة إلى العيِّنة كلها؛ فقصارى ما نصل إليه من تطبيق هذه الخطوات ألا تتكرر الكلمة الواحدة في كل مئة، وهذا لا يمنع من تعدد مرات ورودها فيما يلحق من جداول. ومن ثم يتطلب الأمر القيام بخطوات أخرى لحصر الأنماط على مستوى العيِّنة كلها، وهذه هي:

(١) مراجعة كل كلمة لم تشطب في الجدول الأول على جميع الكلمات التي لم تشطب في الجداول التسعة والعشرين اللاحقة، بحيث يتم شطب جميع تكرارات الكلمة على مستوى النص كله.

وقد أجرينا الشطب في هذه المرة بقلم ذي لون مخالف، أو بإشارة مخالفة حتى يمتاز للباحث ما تم شطبه على مستوى الجدول الواحد مما تم شطبه على مستوى العيِّنة كلها).

(٢) مراجعة كل كلمة لم تشطب في الجدول الثاني على جميع الكلمات التي لم تشطب في الجداول اللاحقة (وعددها ٢٨ جدولاً). وهكذا حتى نفرغ من جميع الجداول الثلاثين التي تتكون منها العينة. وقل مثل ذلك في العينتين الأخريين بجداولهما الستين.

(٣) لكي نضمن دقة الحصر قمنا بعد إجراء هاتين المجموعتين من الخطوات باستنساخ الكلمات الباقية دون شطب في جداول مماثلة تحمل سلسلة الأرقام نفسها من ١ إلى ٣٠ في كل عينة من العينات الثلاث. ومن المتوقع أن يكون عدد الكلمات المستنسخة في الجداول الأخيرة أقل بكثير من عددها في الجداول الأصلية، وهذا ما يسهل عملية حصرها بالطريقة السابقة نفسها وباتباع المراحل نفسها، وذلك حتى نستدرك ما عسى أن يكون قد تفلت من نظرنا أثناء الحصر الأول، وحتى يطمئن الباحث تماماً إلى دقة مراحل القياس.

جدول رقم (١)

نموذج جدول الاستنساخ

مقياس جونسون لاختبار تنوع المفردات في النص

تعود	بنا	هذه	المقدمة	ثلاثين	سنة	إلى	اليوم	الذي	سمعت
فيه	أول	اقتراح	بتأليف	كتاب	عن	محمد	عليه	السلام	وكنت
أقيم	يومئذ	في	ضاحية	العباسية	البحرية	على	مقربة	من	الساحة
التي	كانت	معدة	للاحتفال	بالمولد	النبوي	في	كل	عام	ولنا
رهط	من	الأصدقاء	المشتغلين	بالأدب	يشتركون	في	قراءة	كتبه	العربية
والإفرنجية	ويترددون	معاً	على	الأحياء	الوطنية	وقلما	ويترددون	على	غيرها
فلا	يزالون	متنقلين	فترة	بعد	فترة	بين	الحي	الزينبي	والحي
الحسيني	أو	بين	منشية	القلة	وضاحية	العباسية	أو	بين	الروضة
والخليج	على	حسب	المناسبات	وعلى	غير	مناسبة	في	كثير	من
الأوقات	وكان	رهطاً	له	نقائض	الدنيا	مجتمعات	نقائض	الشباب	نقائض

مصدر النص: عبقرية محمد صفحة ١١ - المؤلف العقاد - الجدول (١).

No. of types (عدد الأنمـــاط) ٧٠

No. of tokens (عدد التحققات) ١٠٠

TTR (نسبـة التنـــوع) ٠,٧

(٤) راجعنا بعد ذلك جداول التصفية على الجداول الأصلية لشطب ما تم اكتشافه من تكرارات.

(٥) تم حصر عدد «الأنماط» في هذه المرحلة في كل جدول من جداول التصفية، مع مراجعة حاصل الجمع على الجدول الأصلي المقابل فإذا توافق الرقمان كان ذلك دليلاً على دقة الإحصاء، وإلا فلابد من إعادة التدقيق لاستكشاف أسباب التخالف واستدراكها.

(٦) يكتب عدد الكلمات المستخرج من المرحلة السابقة تحت الجدول الخاص به.

ومن الواضح أننا بذلك نكون قد استخرجنا رقمين من كل جدول: الأول «للأنماط» على مستوى الجدول، والثاني «للأنماط» على مستوى العينة كلها ومن ثم يجب تمييز كل رقم بعلامة مميزة. (يمكن - على سبيل المثال - كتابة كل رقم باللون الذي استعمل في شطب الكلمات ذات العلاقة بهذا الرقم).

(٧) نتبع الخطوات السابقة نفسها مع العينتين الأخريين؛ كل على حدة.

بهذه المجموعة من الخطوات يمكن التوصل إلى عدد «الأنماط» (الأنواع) على المستويين:

الأول: عددها بين كل مئة كلمة من كلمات العينة.

الثاني: عددها في العينة المدروسة كلها.

وسنرى أهمية هذين المستويين عند الكلام على طرق استخراج النسبة المطلوبة من المعلومات المتوافرة لدينا نتيجة إجراء العمليات السابق ذكرها.

وقبل أن نأخذ في بيان هذه الطرق نود أن نوضح الشروط التي أجرى

تحتها الحصر السابق، فإلى أي مدى وعلى أي أساس يمكن أن تعد كلمة ما مختلفة (أي «نمطاً») بالنسبة للأخرى؟ وبين يدي الإجابة عن هذا السؤال نود أن نقدم هذا التنبيه، إذ الشروط التي سنذكرها إجرائية محض، وهي لا تلزم غيرنا ممن يريد معالجة عينة أسلوبية ما على هذا المنهج، فمن حقه أن يحدد شروطه على النحو الذي يراه شريطة أن يلتزمها في جميع ما يدرس التزاماً صارماً.

أما في هذا المبحث فقد رأينا أن تحقيق قياس أدق لخاصية تنوع المفردات يتطلب الالتزام بما يأتي:

(١) يُحتسب الفِعْلُ كلمة واحدة مهما اختلفت صيغته بين مضي ومضارعة وأمر، ومهما اختلفت كذلك جهات إسناده إلى المفرد والمثنى والجمع تذكيراً وتأنيثاً.

(٢) لا يعتد باختلاف صيغ الأسماء إفراداً وتثنية وجمعاً بوصفها أنماطاً إلا إذا كان المثنى أو الجمع من غير لفظ المفرد.

(٣) لا يعتد باختلاف الاسم تذكيراً وتأنيثاً بوصفها أنماطاً إلا إذا كان المؤنث من غير لفظ المذكر.

(٤) إذا تعددت صيغ الجموع احتسبت (أنماطاً)؛ أي كلمات مختلفة.

(٥) إذا اتصلت بالاسم اللاحقةُ الدالة على النسب أو لاحقة المصدر الصناعي فإن الصور الثلاث تعد (أنماطاً)، وعلى ذلك فمثل «إنسان - إنساني - إنسانية (مصدر صناعي)» تعد ثلاث كلمات مختلفة.

(٦) إذا دلت الكلمة على أكثر من معنى معجمي على جهة الاشتراك عُدّت كلمات مختلفة (أي أنماطاً).

(٧) يعتد بالكلمة الرئيسة فقط مهما تعددت السوابق واللواحق، فكلمات مثل: «محمد - لمحمد - ، هذا - بهذا - لهذا، ما (موصولة) - بما - كما -

فيما، له - لنا - لكم» تحتسب كل مجموعة منها كلمة واحدة.

(٨) إذا اختلفت صيغ الأفعال بين ثلاثية ورباعية وخماسية وسداسية، وكذلك المصادر والمشتقات فإن وحدة الجذر لا تحول دون احتسابها أنماطاً.

تلكم هي أهم الشروط التي التزمناها في الإحصاء. هذا، وإن هذه الشروط في تطبيق المقياس هي أصلح شيء في ظننا لمقاربة النصوص ذات الطول المناسب. أما إذا أريد إخضاع مدونة كبيرة في حجمها للمعالجة الحاسوبية، فربما كان من الأنسب تعديل الشروط أو تكييفها بحيث يكون مطلق الاختلاف في الصورة الكتابية البصرية مُثْبِتاً لوجود التنوع؛ ذلك أن اعتماد هذا المبدأ سييسر مهمة المبرمج إلى حد كبير. وصحيح أن الطريقة اليدوية أدل وأدق في مقام قياس التنوع، بيد أن الطريقة الثانية أسرع وأنجع عند معالجة المدونات الكبرى. والآن نعرض للطرق التي يتم بها حساب نسبة التنوع.

٣ - طرق حساب النسبة

اقترح جونسون أربع طرق يمكن حساب نسبة تنوع المفردات باستخدام واحدة منها أو أكثر حسبما يراه الباحث مفيداً ومحققاً لهدفه من الدراسة. وهذا عرض مع التمثيل للطرق الأربع[١٢] قبل تطبيقها على العينات الثلاث التي اخترناها للدراسة:

الطريقة الأولى: إيجاد النسبة الكلية للتنوع *Over-All TTR*:

وفيها تحتسب نسبة التنوع على مستوى النص أو العينة كاملة. ويتطلب حساب النسبة بهذه الطريقة حصر الأنماط في النص كله، وقسمة عددها على الطول الكلي «التحققات» مُقَدّراً بعدد الكلمات المكونة للنص.

مثال:

إذا كان لدينا نص يتكون من ١٠٠٠ كلمة، وكان عدد الأنماط فيه ٢٥٠ كلمة فإن النسبة الكلية للتنوع تحسب بقسمة $\frac{٢٥٠}{١٠٠٠}$ وتساوي بذلك ٠,٢٥.

الطريقة الثانية: إيجاد القيمة الوسيطة التنوع *the mean segmental*

ويتطلب استخدام هذه الطريقة اتباع الخطوات الآتية:

(١) تقسيم النص أو العينة إلى أجزاء متساوية الطول.

(٢) حساب نسبة الأنماط إلى التحققات في كل جزء على حدة.

(٣) استخراج القيمة الوسيطة لقيم نسبة التنوع في الأجزاء المختلفة. وذلك بجمع هذه القيم ثم قسمتها على عدد الأجزاء المكونة للنص.

مثال:

لنفترض أن لدينا نصاً يتكون من ٢٠٠ كلمة، وقسمناه إلى ثلاثة أجزاء بحيث يتكون كل جزء من ١٠٠ كلمة.

فإذا كان عدد الأنماط في الأجزاء الثلاثة على التوالي ٦٠، ٥٠، ٤٠، فإن النسب ستكون على الترتيب ٠,٦، ٠,٥، ٠,٤، كما سيكون مجموعها ١,٥. وبقسمة هذا العدد على ٣ (وهو عدد الأجزاء) تصير القيمة الوسيطة للتنوع في هذا النص ٠,٥.

الطريقة الثالثة: إيجاد منحنى تناقص نسبة التنوع

the decremental TTR curve

ويتطلب ذلك:

(١) تقسيم النص إلى أجزاء متساوية الطول.

(٢) حساب النسبة في الجزء الأول من النص، وذلك بحصر الأنماط وقسمة عددها على المجموع الكلي للتحققات في هذا الجزء.

(٣) حصر الأنماط في الجزء الثاني من النص دون أن ندخل فيها أي كلمة سبق ورودها في الجزء الأول.

(٤) إيجاد النسبة في الجزء الثاني بقسمة عدد الأنماط التي تم حصرها على المجموع الكلي لتحققات الجزء الثاني فقط.

(٥) تتبع الطريقة نفسها مع الجزء الثالث وكذلك سائر الأجزاء إلى أن تنتهي جميع الأجزاء المكونة للعينة.

مثال:

لنفترض أنه عند فحص النص الذي يتكون من ٣٠٠ كلمة مقسماً على ثلاثة أجزاء قد تبين لنا أن عدد الأنماط في الجزء الأول ٦٠ كلمة، وأن عددها في الجزء الثاني (وهي التي لم تظهر من قبل في الجزء الأول) هو ٤٠ كلمة، وعددها في الجزء الثالث (بشرط عدم ورود أي منها في الجزأين السابقين) ٢٠ كلمة - فإن حساب منحنى تناقص النسبة يتم بالطريقة الآتية:

النسبة في الجزء الأول = $\frac{٦٠}{١٠٠}$ = ٠,٦

النسبة في الجزء الثاني = $\frac{٤٠}{١٠٠}$ = ٠,٤

النسبة في الجزء الثالث = $\frac{٢٠}{١٠٠}$ = ٠,٢

ومعنى ذلك أن خاصية التنوع تتناقص على الموالاة تبعاً للنسب ٠,٦، ٠,٤، ٠,٢.

الطريقة الرابعة: إيجاد منحنى تراكم نسبة التنوع

the cumulative TTR curve

ويتم حسابه على الوجه الآتي:

(١) تقسيم النص إلى أجزاء متساوية الطول.

(٢) إيجاد النسبة بين الأنماط والمجموع الكلي لتحققات الجزء الأول.

(٣) بالنسبة للجزء الثاني يتم إيجاد النسبة بين الأنماط (التي لم يسبق لها أن ظهرت في الجزء الأول) وبين المجموع الكلي لتحققات هذا الجزء فقط.

(٤) نقوم بجمع عدد الأنماط في الجزء الأول إلى عدد الأنماط في الجزء الثاني، ثم نحصل على نسبة التراكم بقسمة حاصل جمعها على المجموع الكلي لتحققات الجزأين معاً.

(٥) نسبة التراكم في الجزء الثالث تساوي حاصل جمع عدد الأنماط في الأجزاء الثلاثة، مقسوماً على الطول الكلي للنص «مقدراً بعدد التحققات المكونة للأجزاء الثلاثة». وهكذا حتى تنتهي جميع الأجزاء المكونة للنص أو العينة.

مثال:

يمر إيجاد منحنى التراكم للعينة المذكورة في المثال السابق بالخطوات الآتية:

النسبة في الجزء الأول = $\frac{٦٠}{١٠٠}$ = ٠,٦

النسبة في الجزء الثاني = $\frac{٤٠}{١٠٠}$ = ٠,٤

نسبة تراكم التنوع حتى نهاية الجزء الثاني = $\frac{٤٠ + ٦٠}{٢٠٠}$ = ٠,٥

نسبة التنوع في الجزء الثالث = $\frac{٢٠}{١٠٠}$ = ٠,٢

نسبة تراكم التنوع حتى نهاية الجزء الثالث = $\frac{٢٠ + ٤٠ + ٦٠}{٣٠٠}$ = ٠,٤

ويقتضي ذلك أن التنوع في هذه العينة يتراكم على الموالاة بالنسب ٠,٦، ٠,٥، ٠,٤ وهكذا يستبين لنا الفرق بين إيجاد منحنى التناقص (المبين في الطريقة الثالثة) وإيجاد منحنى التراكم (الطريقة الرابعة). ويَبيِّنُ هنا أن التناسب بينهما عكسي لا طردي.

تلكم هي الطرق الأربع التي يمكن باستخدامها حساب نسبة تنوع المفردات في الأسلوب[١٣]. وليس من بينها واحدة هي أولى بالاتباع من الأخريات فجميعها صالح لقياس هذه الخاصية، ولإعطائنا مؤشراً دالاً على حجم الثروة اللفظية وتنوعها في أسلوب الكاتب أو الشاعر. واختيار طريقة دون أخرى هو أمر مرجعه إلى الباحث نفسه. بيد أن في الإمكان بوجه عام أن نقول إن الطريقة الأولى الخاصة بإيجاد النسبة الكلية للتنوع (بشرط معرفتنا بالطول الكلي للنص) والطريقة الثانية المتعلقة بإيجاد القيمة الوسيطة لنسبة التنوع بين أجزاء النص يمكن أن تمدا الباحث بمؤشر أكثر وضوحاً إذا كان هدفه التمييز بين أسلوب منشىء ومنشىء من غير تعرض تفصيلي لنقد النص واستخراج المعدلات التي تدخل بها الكلمات الجديدة فيه، إذ إنهما لا تغنيان في هذا المجال غناء الطريقتين الثالثة والرابعة اللتين يمكن باستخدامهما إيجاد نسبة تناقص التنوع ونسبة التراكم على التوالي، وهذا لا ينفي إمكان الاعتماد على الطرق الأربع جميعاً في تمييز أساليب المنشئين.

ولعل تتبع هذا العرض لطرق حساب نسبة التنوع يعين على توضيح الأسباب التي من أجلها قمنا في العيِّنات الثلاث بحصر الأنماط بالنسبة لكل جدول على حدة ثم بالنسبة للعيِّنة كلها.

ولقد استخدمنا في معالجة العيِّنات الثلاث الطرق الأربع التي أسلفنا شرحها، وذلك بهدف تمييز أساليب الأعلام الثلاثة، وإن كنا في بحثنا هذا لم نستنفد جميع الإمكانات التي يمنحنا إياها استخدام الطريقتين الثالثة والرابعة، فذلك ما لم نقصد إليه حتى لا نخرج بالبحث إلى إطالة تتجاوز الحدود المرسومة له سلفاً، ولنا إن شاء اللّه عود في قابل لدراسة أكثر تفصيلاً نعالج فيها مسألة تنوع المفردات، ونضيف بها مقاييس أخرى معتمدة في هذا المقام.

٤ - نتائج القياس

نسجل في مجموعة الجداول والرسوم البيانية الآتية النتائج التي توصلنا إليها باستخدام هذا المقياس لفحص النماذج المختارة من كتابات العقاد والرافعي وطه حسين.

جدول رقم (٢)

النسبة الكلية للتنوع في العينات الثلاث

الكاتب	النسبة الكلية للتنوع
العقاد	٠,٣٩
الرافعي	٠,٣٣
طه حسين	٠,٢٥

جدول رقم (٣)

نسبة التنوع باستخدام القيمة الوسيطة في العينات الثلاث

(كل عينة مقسمة إلى ٣٠ جزءاً في ٦ مجموعات، وتتكون المجموعة من ٥٠٠ كلمة)

الكاتب	النسبة الكلية للتنوع						القيمة الوسيطة
	١	٢	٣	٤	٥	٦	
العقاد	٠,٧٣	٠,٧٥	٠,٦٨	٠,٧٢	٠,٧٤	٠,٧٠	٠,٧٢
الرافعي	٠,٧٢	٠,٦٦	٠,٧٢	٠,٧٠	٠,٧٣	٠,٧٢	٠,٧١
طه حسين	٠,٦٩	٠,٥٩	٠,٦٧	٠,٦٦	٠,٦٤	٠,٦٣	٠,٦٥

جدول رقم (٤)

نسبة تناقص التنوع

(كل عينة مقسمة إلى ستة أجزاء، والجزء يتكون من ٥٠٠ كلمة)

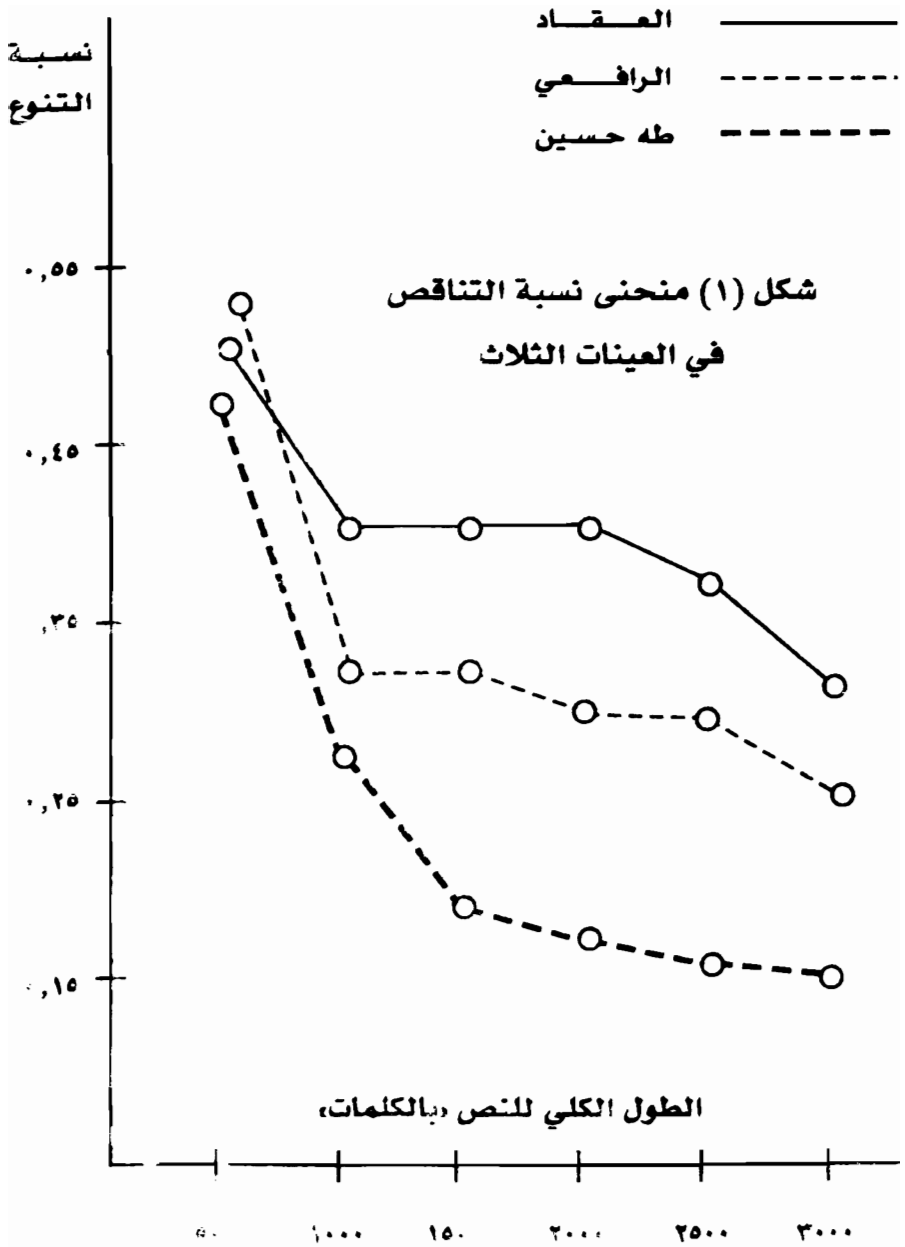
الكاتب	نسبة تناقص التنوع بين الأجزاء					
	١	٢	٣	٤	٥	٦
العقاد	٠,٥١	٠,٤٠	٠,٤٠	٠,٤٠	٠,٣٦	٠,٢٧
الرافعي	٠,٥٣	٠,٣٣	٠,٣٣	٠,٣٠	٠,٢٩	٠,٢٤
طه حسين	٧٤,٠	٠,٢٨	٠,٢٢	٠,٢٠	٠,١٨	٠,١٧

جدول رقم (٥)

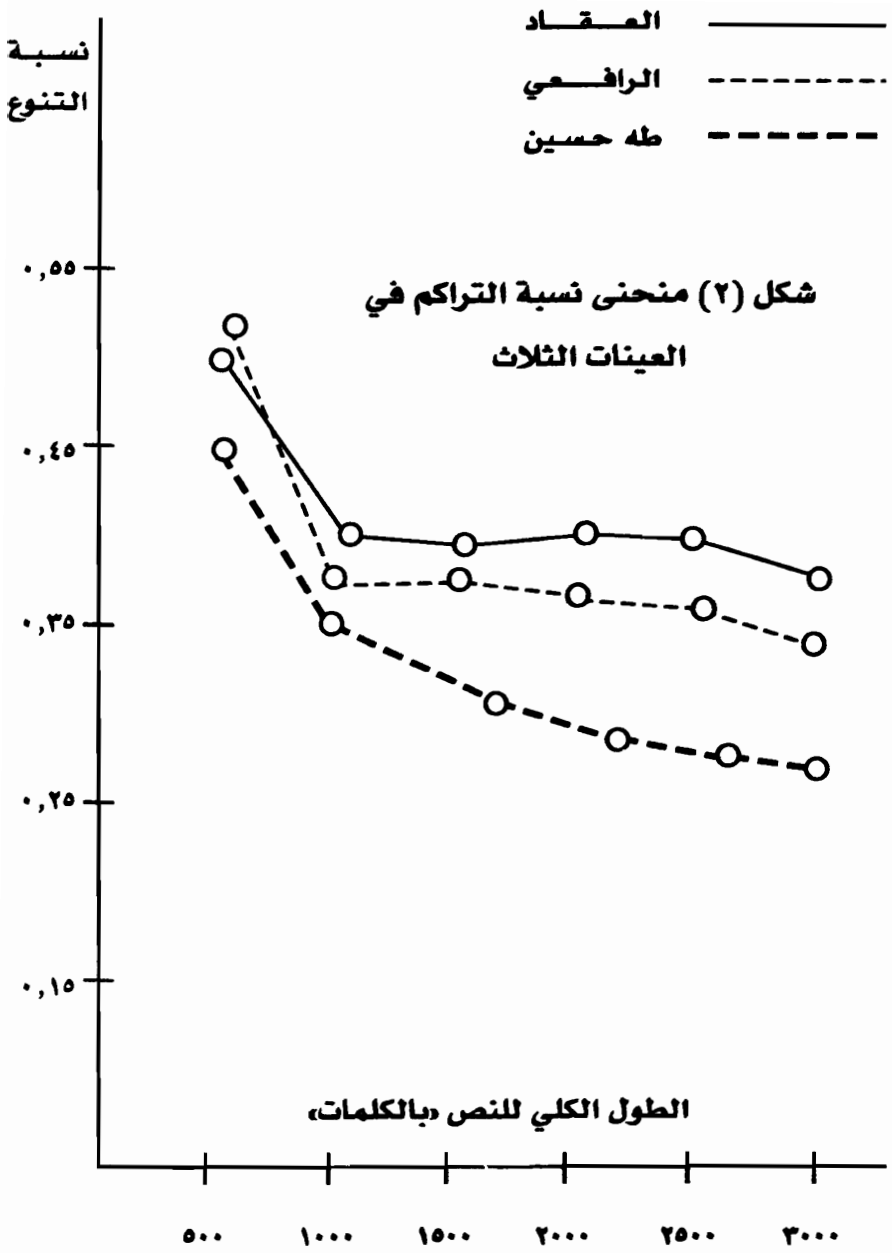
النسبة التراكمية للتنوع في العينات الثلاث

(كل عينة مقسمة إلى ستة أجزاء، والجزء يتكون من ٥٠٠ كلمة).

الكاتب	النسبة التراكمية للتنوع في العينات الثلاث					
	١	٢	٣	٤	٥	٦
العقاد	٠,٥١	٠,٤٠	٠,٤١	٠,٤١	٠,٤١	٠,٣٩
الرافعي	٠,٥٣	٠,٤٢	٠,٣٩	٠,٣٧	٠,٣٥	٠,٣٣
طه حسين	٧٤,٠	٠,٣٨	٠,٣٢	٠,٢٩	٠,٢٧	٠,٢٥



شكل (١) منحنى نسبة التناقص في العينات الثلاث



شكل (٢) منحنى نسبة التراكم في العينات الثلاث

٥ - ملاحظات على النتائج

نلاحظ أن قياس النسبة الكلية للتنوع يرشدنا إلى أن أكثر الأساليب الثلاثة تنوعاً هو أسلوب العقاد (٠,٣٩)، وأقلّها هو أسلوب طه حسين (٠,٢٥) على حين يتوسط أسلوب الرافعي بينهما، وإن كان أقرب إلى الأسلوب الأول منه إلى الثاني (٠,٣٣)، ودلالة النسبة الكلية على التنوع صحيحة إذا ما توافر فيها شرطان:

الأول : أن تكون أطوال العينات التي هي موضوع المقارنة متساوية.

الثاني: أن نعرف تحديداً الطول الكلي للعينة.

وقد توافر لنا الشرطان فيما عالجنا من عينات؛ فحددناها بثلاثة آلاف كلمة لكل عينة، ومن ثم فالحكم الذي توصلنا إليه صحيح في حدود المادة المختارة والشروط التي طبقت عليها.

ويشهد لصحة الحكم أن قياس الخاصية باستخدام الطرق الأخرى يؤدي بنا إلى النتيجة نفسها. فالقيمة الوسيطة للتنوع في أسلوب العقاد (٠,٧٢)، وهي عند الرافعي (٠,٧١) وعند طه حسين (٠,٦٥).

ويفسر لنا الشكلان ١، ٢ الكثير من طبيعة المقياس من جهة، ومن خصائص أساليب الأعلام الثلاثة من جهة أخر. فأما عن طبيعة المقياس نفسه فقد نبه ج. ب. استوب ***G. B. Estoup***[١٤] إلى إحدى الصعوبات التي تواجه من يستخدمه من الباحثين، وفحواها أن معدل الزيادة في عدد الأنماط أقل بكثير من معدل الزيادة في المجموع الكلي للكلمات المكونة للنص، إذ إن احتمال تكرار الكلمات يزيد بزيادة طول النص؛ حتى إن الأجزاء الأخيرة منه قد تتشكل في الأعم الغالب من كلمات سبق ورودها، وتتضاءل الفرصة أمام الكلمات الجديدة للظهور. ولما كانت هذه ظاهرة عامة تحكم العلاقة ما بين الأنماط والمجموع الكلي لتحققات النصوص وجدنا أن الاتجاه العام للمنحنيات في الشكلين ١، ٢ واحداً مع جميع الأساليب، فهي جميعاً تبدأ بقيمة أعلى ثم تتجه إلى الانحدار. بيد أن الكتاب الثلاثة يختلفون اختلافاً مميزاً في درجات

الانحدار: حيث يبدو المنحنى الممثل لأسلوب العقاد أقل الثلاثة انحداراً، يليه المنحنى الخاص بالرافعي، على حين يبدو انحدار المنحنى أكثر وضوحاً في أسلوب طه حسين.

ويرتبط ذلك كله بنتائج قياس نسبة التناقص (ويمثلها الشكل ١، والجدول ٤) وقياس نسبة التراكم (ويمثلها الشكل ٢، والجدول ٥). ومن هذه الزاوية نجد:

١ - أن أسلوب العقاد يتميز بنسبة تراكم أعلى ونسبة تناقص أقل. ومن ثم يتخذ أسلوبه نسبة تنوع ثابتة تقريباً ابتداء من الجزء الأول حتى الجزء الثالث في النص؛ أي على مدى النصف الثاني كله من العينة.

٢ - أن أسلوب الرافعي يتميز بنسبة تراكم أقل نوعاً ما من أسلوب العقاد، كما أن نسبة التناقص في التنوع عنده أعلى من سابقه، ويتناقص التنوع عنده بشكل حاد بعد خمس مئة الكلمات الأولى، ثم يثبت تقريباً في الجزأين الثاني والثالث، وتسجل القيمة نقصاً آخر مع ثباتها تقريباً في الجزأين الرابع والخامس، ثم تتحدر انحداراً واضحاً في الجزء السادس الأخير.

٣ - أن أسلوب طه حسين يتميّز بأنه أقل الأساليب الثلاثة في نسبة التراكم، وأنه أعلاها جميعاً في نسبة التناقص. ويكاد يقف التنوع فيه عند نسبة ثابتة؛ ابتداء من الجزء الرابع (من ١٥٠٠ - ٢٠٠٠ كلمة) حتى الجزء السادس الأخير.

ونخرج مما سبق بأن أسلوب العقاد أعلى الأساليب الثالثة تنوعاً، يليه أسلوب الرافعي، ثم بعد فاصل كبير نوعاً يأتي أسلوب طه حسين، ومن ثم يمكن القول - بطريق الاقتضاء بأن حجم الثروة اللفظية عند الكتاب الثلاثة يسير في خطوط موازية غالباً للمنحنيات التي يسجلها قياس خاصية التنوع.

ولقد لحظنا أن الفارق بين نسبة التنوع عند العقاد والرافعي ليس كبيراً، على حين يفصل بين الكاتبين من جهة وطه حسين من جهة أخرى فارق ملحوظ. ويمكن التماس العلة لهذا الأمر في الفارق ما بين خواص اللغة

المكتوبة *written language* واللغة المنطوقة *spoken language*؛ فالحق أن أسلوب الرافعي والعقّاد أسلوبٌ كتابيٌّ خالص، يخضع للتسويد والتبييض والتنقيح والتحكيك. أما أسلوب طه حسين فنحن نتوقع أنه أسلوب وسط ما بين المكتوب والمنطوق، إذ إنه يملي كتبه على مستملاه، وحينئذ تكون الفرصة لطول التنقيح والمراجعة أقل مواتاة. وإذن فهو فارق ما بين مماناة التجويد المقصود للأسلوب وما يشبه أن يكون تلقائية الأداء.

ونود أن نؤكد أن الوصول إلى هذه النتيجة من الموازنة بين الأساليب الثلاثة لا تعني - في هذه المرحلة من البحث - قدحاً أو مدحاً، بقدر ما تعني التشخيص والتوصيف، وتحديد موقف هذا الأسلوب أو ذلك من المعيار الإحصائي الذي يجري تحكيمه.

ومن الأهمية بمكان أيضاً أن نؤكد أنّ أسلوب المنشئ الفرد لا يمكن تمييزه بالطرق الإحصائية على نحو متكامل إلا باستخدام منظومة من المقاييس المتنوعة يمكن بها قياس عدد دالٍّ من الخصائص الأسلوبية. وأما عند الموازنة بين أكثر من منشئ فمن المتوقع أن تتقاطع خطوط توزيع الخصائص الأسلوبية على نحو غير منتظم، فقد يتفق الأسلوبان (أ) و (ب) في خاصية يختلفان فيها عن الأسلوب (جـ)، على حين يثبت استخدام مقياس آخر لخاصية أخرى التشابه بين (أ) و (جـ) دون (ب). من ثم يتم التحديد والتمييز بين الأساليب على أساس اعتماد أكبر مجموعة ممكنة من الخصائص التي يتميّز بها أسلوب من أسلوب، مع وجود الفرصة للتشابه بين هذا الأسلوب أو ذلك في خاصية أو أكثر.

والعينات الثلاث التي نقوم الآن بفحصها تقدم لنا دليلاً جيداً على هذه الظاهرة، فالتقارب بين أسلوب الرافعي وأسلوب العقّاد في خاصية تنوع المفردات لا يعني أن الأسلوبين شيء أحد، وإنما الذي يعنيه أن هذه الخاصية على وجه الخصوص لا تصلح وحدها مميزاً حاسماً بين أسلوب الرافعي وأسلوب العقّاد، على حين تصلح مميزاً جيداً وقادراً على النهوض بمهمة التمييز بين أسلوب كل منهما وأسلوب طه حسين. وإذن فهناك خصائص أُخَر

غير هذه الخاصية هي التي ستتكفل بالتمييز الحاسم بين أسلوبي الرافعي والعقّاد فإن قيل: وما هي؟ قلنا: ذلك متروك للتجربة والاختبار وتطبيق المقاييس الأخرى المقترحة. وقد قمنا في غير هذا المبحث بدراسة إحصائية أخرى وازنّا فيها بين أسلوبي العقّاد وطه حسين من حيث خاصية الانفعالية والعقلانية، وتبين لنا صلاحية معامل بوزيمان للقيام بهذه المهمة على وجه جدير بالتعويل عليه، ولعله يتاح لنا أو لغيرنا تطبيق هذا المعامل على عينة من أسلوب الرافعي ليتسنى لنا تحديد ملامحه بهذا الاعتبار(١٤).

٦ - علاقة تنوع المفردات بصعوبة الأسلوب

لاحظ بعض العلماء وجود صلة وثيقة بين صعوبة الأسلوب وارتفاع نسبة التنوع فيه، وقد حدا بهم ذلك إلى القول بأن نسبة التنوع هي أفضل مقياس يمكن به اختبار مدى الصعوبة في الأسلوب(١٥).

وترجع العلاقة بين الخاصتين إلى أمر يمكن توقعه، فالكاتب أو الشاعر الذي يتميز بنسبة تنوع عالية في المفردات، أي بوجود عدد كبير من الأنماط، يلجأ عادة إلى استخدام كلمات غير مألوفة لكي يزيد من تنوع ألفاظه، كما أن القارىء مع مثل هذا الأسلوب يكون أكثر توتراً، لأنه لا يزال يفاجأ بجديد كلما أمعن في القراءة.

وتُصدِّقُ النتائج التي حصلنا عليها من قياس العينات الثلاث حكم الذوق الذي يقضي بأن كتابات العقاد والرافعي تُعَدُّ في باب الصناعة الأسلوبية على درجة ملحوظة من الصعوبة والتعقيد إذا ما قيست إلى كتابات طه حسين.

وها نحن أولاء نجد قياس نسبة التنوع يفتح لنا باباً واسعاً لقياس خاصية أخرى من أهم الخواص الأسلوبية التي يهتم الدارس لتحديدها، ألا وهي خاصية صعوبة الأسلوب.

ونجتزئ هنا بهذه الإشارة، فلقياس صعوبة الأسلوب مقاييس أخر غير هذا المقياس، ومجال آخر غير هذا المجال. ولعلنا نوفق في قابل إن شاء الله إلى تناول هذا الموضوع في دراسة نظرية وتطبيقية أخرى.

المصادر والحواشي

(١) انظر مقدمة الطبعة الرابعة من كتابي: «الأسلوب»، وعنوانها: «عن اللسانيات العربية وقراءة النص الأدبي»، وكذلك الفصل الأول من الكتاب بعنوان: «الحاجة إلى منهج».

(٢) رجعنا إلى طبعة دار الكتاب العربي، بيروت، ص ٤٦ - ٥٨ من الجزء الثاني.

(٣) رجعنا إلى طبعة الكتاب المنشورة في الأعمال الكاملة، المجلد الأول، بيروت، ص ص ١١ - ٢٥.

(٤) رجعنا إلى طبعة الكتاب المنشورة في الأعمال الكاملة، المجلد الرابع، بيروت، ص ص ١٩٩ - ٢٠٩.

(٥) راجع في ذلك كتاب حلمي خليل: «الكلمة: دراسة لغوية ومعجمية»، الهيئة المصرية العامة للكتاب، ١٩٨٠، ص ٣٤، ومواضع متفرقة من الفصل الأول الذي أخلصه لاستعراض محاولات تعريف المصطلح.

(٦) انظر المرجع السابق: ص ١٠٢ - ١٠٤.

(٧) ترجمة المصطلحين لشيخي الأستاذ الدكتور عبدالرحمن أيوب، انظر كتابه: «اللغة والتطور» معهد الدراسات العربية العالية، القاهرة: ١٩٦٩، ص ١٠٤.

(٨) من بين هذه المقاييس: مقياس كارول J. B. Carrol وقد عرضه في مقال له بعنوان:

Diversity of Vocabulary and the Harmonic Series Law of Word Frequency Distribution, ***Psych, Rec.,*** 1838, 2, PP. 379 - 386.

وثمة عرض شامل لعدد من هذه المقاييس تجده في:

G. Miller, ***"Language and Communication"***, New York, Toronto, London, 1963, pp. 122 - 126.

(٩) عنوان الكتاب بالإنجليزية: W. Johnson, "Language and Speech Hygiene", ***Gen. Semantics, Monograph*** No. 1, 2 and ed., Chicago, Institute of General Semantics, 1941.

(١٠) ***"People in Quandaries"*** New York, Haper 1946.

(١١) من الممكن - بطبيعة الحال - اختيار نموذج الجدول بحيث يشتمل على عدد أكبر أو أصغر، ولكننا اخترنا الجدول المكون من مئة خانة قصداً لتيسير حساب النسبة.

(١٢) W. Gohnson, ***"People in Quandaries"***, pp. 502 - 505.

(١٣) Miller, op. cit. pp. 122 - 123.

(١٤) انظر فصلاً بعنوان: «نماذج تطبيقية من الأساليب النثرية» في كتابي: «الأسلوب: دراسة لغوية إحصائية».

(١٥) انظر: M. Vogel and C. Washburne, "An Objective Method of Determining Grade Placement of Children's Reading Material", ***Elementary School Journal***, 1928, 28, PP. 373 - 375.