

المقطع السادس: برمجيات تحليل الخطاب حاسوبيا-غواص Ghawwas_V4.6 نموذجاً-

الأهداف العامة للمقطع:

يهدف هذا المقطع إلى :

-التعريف بتحليل الخطاب حاسوبيا

-فهم طريقة عمل برمجية غواص

مقدمة:

تحليل الخطاب ليس حقلاً ذا علم صارم وبمتصورات تأسيسية، ومرحلية، ومن ثم كيانية؛ كونه قد بدأ قديماً في سياق دراسات النصوص البلاغية والفقهاء لغوية إلى أن تضافرت مدارس وحقول أخرى في تكوين شعبه، وسبقه تحليل المحتوى analysis content الذي ظهر في مطلع القرن العشرين في الولايات الأمريكية المتحدة، ثم تشعبت بتشعب المدارس اللسانية والوظيفية والإثنوجرافية، بدءاً من عام 1952م عند هاريس Harris ثم نال أبعاداً عديدة متضافرة التداخل التداولي والنفسي والاجتماعي والسلوكي غير أن حجم الخطاب المحلل في المجمل لا يتجاوز صفحات معدودة سواء كانت مكتوبة أو منقولة نسخاً من المنطوق وبعد تطور أدوات معالجة اللغات ظهرت أبحاث في تحليل الخطاب تقوم على جمع نصوص رقمية تبلغ آلاف الكلمات، واعتمدت هذه الأبحاث على تحليل التكرارات والكشافات السياقية بوصفهما المنطلق لأي منهج من مناهج تحليل الخطاب

1-تحليل الخطاب حاسوبيا:

تتنوع اتجاهات تحليل الخطاب وتحليل الخطاب النقدي بتنوع مصادر التحليل اللغوي المنطوق أو لمكتوب، فمنها ما يكون التركيز فيه على الكلمة أو العبارة أو الجملة التي تحقق عنصر التواصل الفعال، وذلك كله بالتركيز على المعرفة حول اللغة، وكيفية تشكلها عبر النصوص والكلام، وبالتركيز على لعلاقة بين اللغة وبين مختلف السياقات الاجتماعية والثقافية المنتجة فيها، كما يدرس الاتجاهات التي تمثل اللغة بها تصورات العالم المختلفة حول العالم وحول اختلاف الوفاق بين منتجي اللغة وبين السياق الاجتماعي العام والخاص.

ويجري ذلك عن طريق توظيف أدوات معالجة اللغة من التكرارات ، frequencies والكشافات السياقية الممتدة ، expanded concordances والبيانات الإضافية ، metadata ومواقع expanded concordancers في الكشافات السياقية الممتدة span positions المدى وحسابات المقارنات الإحصائية بين المدونة الرئيسة والمدونة المرجعية الأولية.

تعد الأسلوبية الإحصائية من أهم و أبرز اتجاهات الدرس الأسلوبي المعاصر، و السبب في ذلك يعود، دون أدنى شك، إلى إفادة هذه الأخيرة من مناهج العلوم الدقيقة كالرياضيات و الإحصاء، و اعتمادها على أحدث برمجيات الإعلام الآلي و تقنيات الحساب الإلكتروني، و هي لا محالة، تسعى من خلال هذا و ذاك إلى خلق بديل علمي موضوعي في مجال الدراسات النقدية الأدبية، و تهدف إلى توطيد العلاقة بين العلوم الإنسانية و العلوم التجريبية قصد تحقيق نتائج موثوق بها، بعيدة كل البعد عن الأحكام الذاتية و التقييمات الحدسية. تتبنى الأسلوبية الإحصائية، في معالجة النصوص الأدبية، أدوات إحصائية تقوم على تعيين خصائص الخطاب الأدبي و تحديد سمائه البلاغية و الجمالية، من خلال حصر و قياس و وحدات بنيوية ذات وظيفة معينة.

و تقوم الأسلوبية الإحصائية على الوصف الموضوعي و القياس الكمي الذي يستخدم إجراءات التحليل الإحصائي و الرياضي، و يرى أصحاب هذا الاتجاه أن الأسلوب هو المجموع الشامل للبيانات القابلة الالتقاط و التحديد الكمي في بنية النص الشكلية، و لقد اتجهت الكثير من البحوث إلى تحليل العلاقة بين المفردات و معدلات تكرارها و إلى الدراسة الكمية لأطول الكلمات و الجمل، فيقيس بعضهم متوسط طول الجمل و معدل الكلمات فيها و متوسط طول الكلمات و معدل المقاطع و الحروف المكونة لها.

و يعد استخدام الإعلام الآلي في مجال الأسلوبية الإحصائية من أهم ميزات هذه الدراسة الكمية التي أصبحت أكثر شهرة بفضل الإسهامات الجليلة لتقنيات الفحص الإلكتروني و آليات البرمجة المعلوماتية التي تعمل بدورها على تخصيص و إرساء مناهج حديثة في المعالجة الإحصائية للنصوص الأدبية.

فالإحصاء المعجمي لكلمات و مفردات نص ما تعد مهمة جد صعبة تتطلب الكثير من الدقة اللسانية، كما أنه أمر لا يخلو من الهفوات و الأخطاء، خاصة إذا ما تعلق الأمر بإحصاء آلاف الترددات في نص من 15000 إلى 20000 تردد، مثلا. في هذه الحالة، فإن الأمر يتجاوز القدرة البشرية، لذا فقد تم توكيل هذه المهام الحساسة و الدقيقة إلى الآلة و تتم هذه العملية في مرحلتين أساسيتين :الأولى يقوم فيها الباحث (اللساني) بتحديد واختيار القاعدة المعجمية التي ستم معالجتها، و في مرحلة ثانية يقوم الإعلامي (L'opérateur) باستثمار هذا الخيار، و إدراجه في الحقل العملي، و ذلك بتسجيل النص و برامج المعالجة اللازمة، و كذا وسائل المراقبة الأوتوماتيكية التي تقوم برصد الأخطاء، تصحيحها، أو إزالتها تماما

2- المدونات اللغوية كأداة لتحليل الخطاب حاسوبياً:

إن الدراسات التي تعتمد على تحليل مجموعة نصوص تركز على بناء مدونة corpus حتى يتم ضبط هذه النصوص ومعالجتها آلياً وهذه المدونات تختلف حسب نوع الدراسة والهدف المرجو منها .

كلمة "Corpus" كلمة إنجليزية مشتقة من كلمة لاتينية بمعنى "الجسد" وترجمة الكلمة اللاتينية الأصل "corpse" التي تقابل في الإنجليزية "body" وجمعها في اللاتينية هو "corpora" وهي عبارة عن مجموعة من النصوص المتاحة للاستخدام الحاسوبي، ويُطلق على مجموعة المدونات بالإنجليزية لفظ "corpora". وقد عرفها معجم أكسفورد للغة الإنجليزية بأنها "جسم من المادة المكتوبة أو المنطوقة يبني عليه التحليل اللغوي". أما معجم ديفيد كرستال فعرفها بأنها "مجموعة من البيانات اللغوية المكتوبة أو المَحَوَّلة لمادّة مكتوبة من تسجيلات صوتية، يمكن أن تستخدم كنقطة بداية لوصف اللغة أو طريقة إثبات الفرضيات اللغوية، بالرغم من أن المتون كانت فيما قبل نصوصاً ورقية؛ إلا أن المتون تُعرف الآن على أنها النصوص المقروءة آلياً، وأطلق عليها مدونات حاسوبية computer corpora وعندما نتحدث عن المتون والمدونات فإننا نعني الحاسوبية - النصوص الالكترونية .

3- مكونات برمجية خواص Ghawwas_V4.6 :

الواجهة الرئيسية

الأسنة الجانبية (إضافة المدونات)

- 1- إضافة المدونات: تحتوي هذه النافذة على وظائف إضافة وحذف المدونات المراد تحليلها.
- 2- خيارات المعالجة: تتبع هذه النافذة عدة خيارات لتنقيح ومعالجة المدونات.
- 3- المقارنة: تتضمن هذه النافذة الخيارات الخاصة بمقارنة المدونات وتحديد الكلمات المميزة لها.
- 4- المدونة الرئيسية: يظهر فيها أسماء المجلدات الخاصة بالمدونة الأساسية.
- 5- إضافة مجلد: للمدونة الرئيسية.
- 6- حذف مجلد: يتم حذف المجلد أو المجلدات التي تم تحديدها بواسطة المستخدم والموجودة ضمن المدونة الرئيسية.
- 7- المدونة المرجعية: يظهر فيها أسماء المجلدات الخاصة بالمدونة المرجعية المراد مقارنتها بالمدونة الرئيسية.
- 8- إضافة مجلد: للمدونة المرجعية.
- 9- حذف مجلد: يتم حذف المجلد أو المجلدات التي تم تحديدها بواسطة المستخدم والموجودة ضمن المدونة المرجعية.
- 10- حساب التكرار: عند النقر على هذه الأيقونة يتم حساب التكرار لكلمات المدونات.

الشكل التوضيحي 1: الأسنة الجانبية (إضافة المدونات).

الشكل التوضيحي 4: خيارات العرض والنتائج.

- 1- اختيار المدونة: يتم اختيار المدونة التي يرغب المستخدم في الاطلاع على نتائج تحليلها بحيث يمكن له اختيار المدونة الرئيسية أو المرجعية أو المدونتين معاً.
- 2- عرض المجلد: تعمل هذه الخاصية إلى اختيار المجلد المراد الاطلاع على نتائج تحليله من المدونة الرئيسية أو المرجعية حسب اختيار المستخدم للمدونة.
- 3- عرض الملف: عرض نتائج تحليل ملف من مجلد تم اختياره سابقاً.
- 4- خيارات عرض النتائج:

4- طريقة تحليل الخطاب استخدام برمجية غواص:

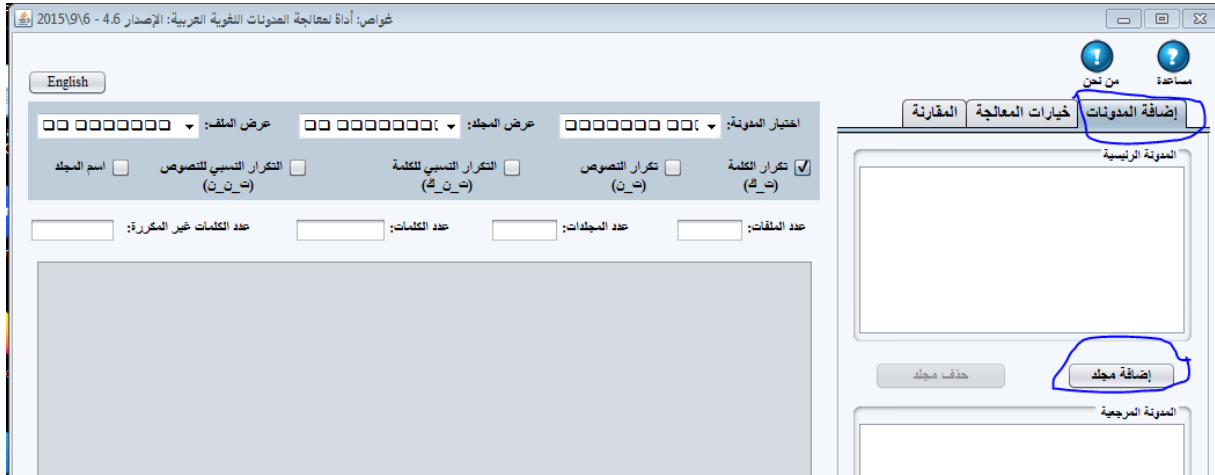
1- إنشاء المدونة:

هناك العديد من المعايير المستخدمة لبناء المدونات من التي تتمثل في حجم المدونة ومدى تنوع اللغة المستخدمة، بالإضافة إلى خصائص النص: كعمر الكاتب وجنسه. وتختلف المعايير بينها، غير أنها تشترك في المجال الذي تستخدم به المدونة، والفترة الزمنية المغطاة، وطريقة استخدام هذه المدونة بشكل منهجي مقنن.

فقد تكون المدونة هنا مجموع مقابلات مع مختصين او افراد العينة او مجموعة نصوص جمعها الباحث او ملاحظات كتبها .

ب-استيراد المدونة الى برمجية غواص:

وذلك بالضغط على اضافة المدونات وبعد ذلك اضافة مجلد



ج- حساب تكرار الكلمات:

باستخدام برمجية غواص نقوم بمعالجة المدونة وذلك بحساب تكرار الكلمات وسياق ورودها واستخداماتها

English

اختيار المدونة: عرض المجلد: عرض الملف: كل الملفات

تكرار الكلمة (ت_ك) تكرار التصوص (ت_ن) التكرار النسبي للكلمة (ت_ن_ك) التكرار النسبي للتصوص (ت_ن_ن) اسم المجلد

عدد الملفات: عدد المجلدات: عدد الكلمات: عدد الكلمات غير المكررة:

الكلمة	ت_ك
من	110
و	74
في	71
لا	39
ان	33
على	32
يا	32
"صباح	25
حتى	18
ما	18
أن	18
الله	17
عن	17
يستقبل	15
لم	15
ولا	14
الذي	13

باركتكشاف السياق للكلمة بضغط على الكلمة مرتين

د-تحليل النتائج