مفهوم علم اللغة الحاسوبي

- علم اللغة الحاسوبي **أحد مجالات علم اللغة التطبيقي، وهو مجال بينيّ تلتقي فيه المعارف والقوانين اللغوية التي يهتم بها علم اللغة مع التطبيقات والقوانين الرياضية المنطقية التي تهتم بها علوم الحاسب.**
- يهتم بمعالجة الظواهر **اللغوية حاسوبيًا، وبتصميم برمجيات حاسوبية لتحليل النظام اللغوي بمستوياته المختلفة،** وتطبيقها في مجالات البحث اللغوي، وتعليم اللغة، والترجمة الآلية، والتحقيق الجنائي، وصناعة المعاجم الإلكترونية، والتلخيص الآلي، وتصحيح الأخطاء الإملائية.
- عند براون وميلر - في معجم كيمردج لعلم اللغة - : هو توظيف الحاسب الآلي في **دراسة علم اللغة وتطوير برمجيات تؤدي المهام اللغوية** مثل الإعراب ومعالجة المادة اللغوية.
- ديفيد كريستال - في معجم علم اللغة والصوتيات - فيعرّف علم اللغة الحاسوبي بأنه فرع من علم اللغة **تُطبَّق فيه المفاهيم والتقنيات الحاسوبية في معالجة الوظائف اللغوية.**
- وفي موسوعة اللغة العربية وعلم اللغة، يؤسس إفرهار ديترز المقصود بعلم اللغة الحاسوبي العربي على مفهوم علم اللغة مع إضافة متغيرين أساسيين هما "الحاسوبي" و"العربي"، فيقول: إن علم اللغة الحاسوبي العربي يعني ١) **دراسة الكيفية التي يعمل بها النظام اللغوي في العربية باستخدام الوسائل والتقنيات الحاسوبية؛ ٢) البحث في كيفية توظّف التقنيات الحاسوبية في معالجة اللغة العربية.**
- لعلم اللغة الحاسوبي جانبان: نظري وعملي؛ **يهتم الجانب النظري بوضع نظريات وقواعد تحقق غاياته في حوسبة اللغة، في حين يركز الجانب العملي على تطبيق هذه النظريات على اللغة باستخدام التقنيات الحاسوبية.**
- **اختراع الحاسب الآلي عام ١٩٤٨** – ليمثّل توجها عالميا متسارعا إلى التحول من **عصر التدوين الورقي والعمل اليدوي إلى عصر الرقمنة والحوسبة (الأتمتة) في كل مناحي الحياة** المعلوماتية والتجارية والعلمية والخدمية والاتصالية والأرشيفية وغيرها من الاهتمامات الإنسانية.
- **ويرجع استخدام الحاسوب في حقل الدراسات اللغوية إلى أوائل الخمسينات** من القرن العشرين حينما **ظهرت بدايات الترجمة الآلية في أمريكا** وبخاصة ترجمة الوثائق الروسية لأغراض عسكرية (فيما عُرف بالترجمة الميكانيكية). ثم **انتقلت التجربة في الستينات إلى أوربا وتحديدًا في السويد عام ١٩٦١.**
- **يؤرّخ للبداية الفعلية لحوسبة اللغة أوروبيا في أوائل الستينات بإنشاء عدد من المراكز الحاسوبية للغة في أوربا والاتحاد السوفيتي** منها المركز الحسابي لدراسة الأدب واللغة في جامعة كيمبردج.
- **أما عربيًا، فيمكن التأريخ لأول محاولة عربية لمعالجة اللغة حاسوبيًا باقتراح قدمه الطبيب محمد كامل حسين في أوائل السبعينيات إلى اللغوي المصري الدكتور إبراهيم أنيس** عن إمكانية توظيف تقنيات الحاسوب

	(Tagger	(Swift)		Arab) (Diac
الكلمات المفتاحية (keywords) من الوظائف البحثية التي تتيحها:	المدونات اللغوية	المحللات الصرفية	برنامج القارئ الآلي للنصوص	برامج الترجمة الآلية
ظهرت أول محاولة عربية في معالجة اللغة حاسوبيا في:	أوائل الخمسينيات	أوائل السبعينيات	أواخر الثمانينيات	أواخر الستينيات
يمثل .....دليلا على نجاح اللغة العربية في التحدي التقني.	التحليل الإحصائي الحاسوبي للجذور العربية	اعتماد مواصفة موحَّدة لتشفير محارف الكيبورد العربية	تعريب نظم التشغيل والتطبيقات الحاسوبية	استخدام الحواسيب اللوحية في العالم العربي
يُفسّر اختلاف مفاهيم الوحدات اللغوية بين الحاسوبيين واللغويين بـ:	غياب الخبرة اللسانية لدى الحاسوبيين والخبرة الحاسوبية لدى اللغويين	اختلاف اللغة البشرية عن اللغة الصورية من حيث الطبيعة والخصائص	إشكالية المصطلحية الحاسوبية شأنها شأن المصطلحات العلمية	تشتُّت الجهود المبذولة في مجال حوسبة اللغة لحدّ يبلغ التعارض
اقترح العراق معيارًا موحدًا لترميز الأبجدية العربية، عُرف بـ:	CODAR-Z	CODAR1	CODAR-U	CODAR-UF
يقوم برنامج.....بالبحث على مستوى المفاهيم المعرفية في القرآن.	إبصار	قرآني	(Swift)	Google

	اللغة الطبيعية	الاصطناعي	لغة الآتمتة	اللغة الصورية
ساعد .....مهندس اللغات الطبيعية (في بادئ الأمر) في صياغة القواعد النحوية في صورة قواعد رياضية.	النموذج الرياضي ذو المرجعية الصورية	النموذج البنيوي ذو المرجعية الأوروبية	النموذج المنطقي ذو المرجعية الدلالية	النموذج التشومسكي ذو المرجعية الأمريكية
تبنَّت..... إنشاء بنك (BASIM) للمصطلحات الحاسوبية.	جامعة ليدز بإنجلترا	مدينة الملك عبدالعزيز للعلوم والتكنولوجيا	مكتبة الإسكندرية	شركة آلاس بكندا
مدونة القرآن الكريم وترجمة معانيه إلى الإنجليزية- مثال لـ:	المدونة الخاصة	المدونة العامة	المدونة المقارنة	المدونة المتوازية
تُصمَّم ..... بحيث تكون متنامية الحجم، وتُضاف إليها النصوص الجديدة بنسب متوازنة بصورة دورية؛ لرصد التغيرات الطارئة على الاستعمال اللغوي ومقارنتها.	المدونة التاريخية	المدونة المقارنة	المدونة المحدَّثة	المدونة العامة
تتضمن .....نصوصًا من فترات زمنية متعاقبة؛ بغرض رصد التغيُّر الطارئ على اللغة، مثل: مدونة (Helsinki)	المدونة المقارنة	المدونة العامة	المدونة المحدَّثة	المدونة التاريخية
تمثل صور رسم الهمزة (أ إ ؤ ئ ء) في مواصفة الكتابة العربية:	محارف المد والحركات	محارف الصوامت	محارف خاصة	محارف الضبط
قدَّم د. مرياتي دراسة إحصائية حاسوبية	لجذور معجم الصحاح	لأبواب الفعل الثلاثي في المعجم الوسيط	لدوران الحروف في الجذور	لجذور معجم لسان العرب

شركة RDI

- ✓ صممت المشكل الآلي Arab Diac، يقوم بتشكيل النص العربى
- ✓ برنامج تحشية الفئات الكلامية Arab Tagger: لعنونة كل فئة كلامية بإضافة رمز مميز لها عن الفئات الأخرى، كتمييز الفعل برمز V و الاسم N والصفة ADJ وغيره.
- ✓ برنامج محرك البحث النصي العربي Swift: وهو برنامج يتيح إمكانية البحث الاشتقاقي فى النصوص العربية بالجذر
- ✓ برنامج Arab Dictions: وهو برنامج يقوم هذا البرنامج بتحليل الكلمات العربية إلى صيغها الصرفية وربطها بالمادة المعجمية التي تنتمي إليها
- ✓ برنامج ARAB TALK وهو برنامج يحول النصوص المكتوبة إلى كلام منطوق

شركة صخر Sakhr

- ✓ صخر للتدقيق الإملائى Sakhr Corrector: وهو برنامج يقوم باكتشاف الأخطاء الإملائية
- ✓ صخر للتشكيل الآلى Automatic Diacritizer: ويقوم هذا البرنامج بوضع علامات التشكيل في النص العربي المكتوب
- ✓ المحلل الصرفي Morphological Analyzer: يقوم هذا البرنامج بتجذير الكلمة العربية بإزالة الزوائد
- ✓ صخر لاستخراج الكلمات المفتاحية Sakhr Keywords Extractor: يحلل هذا البرنامج النصوص العربية ويحدد العبارات والأفكار الأساسية فيها بشكل آلي
- ✓ صخر للتلخيص العربي آليا Arabic Summarizer: يقوم هذا المحرك بتحديد الجمل والأفكار المهمة داخل النص ويعرضها في شكل ملخص
- ✓ إبصار: وهو برنامج مصمم للمعاقين بصريا والمكفوفين يمكنهم من قراءة الكتب
- ✓ صخر للترجمة الآلية: وهو محرك للترجمة الآلية من العربية إلى الإنجليزية والعكس

جامعة ليدز Leeds بإنجلترا:

- ✓ برنامج محاورة القرآن "قرآني" - وهو برنامج للبحث على مستوى المفاهيم المعرفية في القرآن
- ✓ مدونة القرآن الكريم Coranic Arabic Corpus مفتوحة المصدر للبحث المجاني، ومُحشّاة نحويًا وصرفيًا
- ✓ تصميم المفهرس السياقي للمدونات العربية aConCorde وهو متصفح يتيح عرض السياقات النصية للكلمة المبحوث عنها في نصوص المدونة
- ✓ نظام التحشية الصرفية morphological annotation يقوم على تصريف كل كلمة إلى مكوناتها الصرفية segments التي تتكون منها
- ✓

- دراسات إحصائية لدوران الحروف في الجذور العربية.

- المعجم الحاسوبي في نظام خبير للغة العربية.

- له دراسة إحصائية للجذور العربية في المعاجم القديمة باستخدام الحاسبات الإلكترونية في حساب النسبة المئوية للجذور الثنائية والثلاثية والرباعية والخماسية، وإمكانية الإفادة من نتائج هذا الإحصاء في الترجمة الآلية من العربية وإليها. (**د. محمد مرياتي "من سوريا"**).

- قدم دراسات إحصائية لدوران الحروف في الجذور العربية، وللمعجم العربي، ولدوران الحروف العربية المشكولة ولحروف اللغة العربية. (**د. محمد مرياتي "من سوريا"**).

- له بحث "المعجم الحاسوبي في نظام خبير للغة العربية". (**د. محمد مرياتي"من سوريا"**).

- قدم د. محمد مرياتي دراسات إحصائية حاسوبية لـ(**المعاجم العربية القديمة**).

- من الدراسات المهمة في حوسبة اللغة العربية لـ د. محمد الحناش:

- دراسات حول المعجم الحاسوبي للغة العربية.

- المعجم التركيبي للغة العربية.

- بناء المعاجم الإلكترونية للعربية.

- بناء قاعدة بيانات حاسوبية لتراكيب القاموس المحيط.

- تدور مجمل أعمال د. محمد الحناش حول (**بناء المعاجم الإلكترونية الحديثة للغة العربية**).

- قدّم عددًا من الدراسات حول المعجم الحاسوبي للغة العربية ضمنها كتابه "المعجم التركيبي للغة العربية".(**د. محمد الحناش "من المغرب"**).

- مجمل أعماله تدور حول بناء المعاجم الإلكترونية للعربية تأتي ضمن "مشروع علمي كبير عمل عليه لسنوات طويلة من البحث اللساني الحاسوبي". (**د. محمد الحناش "من المغرب"**).

- أسهم في بناء قاعدة لبيانات حاسوبية لتراكيب القاموس المحيط، وبناء محلل صرفي للغة العربية. (**د. محمد الحناش "من المغرب"**).

- أسهم د. محمد الحناش في بناء قاعدة لبيانات حاسوبية لتراكيب (**القاموس المحيط**)

- أسهم د. محمد الحناش في بناء قاعدة لبيانات حاسوبية لـ(**تراكيب القاموس المحيط**)

- أسهم في تشفير مداخل المعجم العربي البسيطة والمعربة: الأفعال الخمسة والأسماء والمصادر والمشتقات. (**د. محمد الحناش "من المغرب"**).

- أسهم في مجال حوسبة النحو من خلال بحثه "منطق النحو العربي والعلاج الحاسوبي" (**د. عبدالرحمن الحاج صالح "من الجزائر"**).

- من أوائل من دعوا إلى مشروع "الذخيرة اللغوية"- المعروفة حديثا بالمدونة اللغوية (**د. عبدالرحمن الحاج صالح "من الجزائر"**)

- له كتاب "بحوث ودراسات في اللسانيات العربية" (**د. عبدالرحمن الحاج صالح "من الجزائر"**)

- يرجع الفضل في "النظرية الخليلية الحديثة" إلى (**د. عبدالرحمن الحاج صالح "من الجزائر"**)

- **الذكاء الاصطناعي** هو هندسة تصميم الآلات الذكية التي تحاكي الإنسان في إنجاز العمليات المعقدة، مثل حل المسائل، والإجابة عن الأسئلة.
- يُقارب الذكاء الاصطناعي ما اصطلح عليه حديثًا **بالأتمتة** أي جعل العمليات والأوامر تُنجز أوتوماتيكيًا.
- ظهرت تطبيقات الترجمة الآلية في الخمسينات من القرن العشرين - وكانت تعتمد فقط على القواميس الإلكترونية وبعض القواعد التركيبية البسيطة لبناء الجملة في اللغتين المترجم منها والمترجم غليها؛ لكن سرعان ما اكتُشف أن نتاج هذه المنهجية نتاج ضعيف من حيث البنية والدلالة.
- وقد وجد المهتمون بحوسبة اللغة ضالتهم في **النموذج التوليدي التحويلي الذي اقترحه تشومسكي** والمعروف "بالنظرية النموذجية" في دراسة اللغة الطبيعية.
- وهو **نموذج ذو مرجعية أمريكية** يقوم على تحليل اللغة إلى مكونات تحكمها قواعد أو أنماط رياضية قادرة على توليد عدد لا نهائي من التراكيب اللغوية في اللغة المعينة.
- إلا أن هذا النموذج التوليدي **يتعامل مع المكونات اللغوية بصورة منعزلة**؛ مما دفع مهندسي اللغات الطبيعية إلى إيجاد بديل يتجاوز المعالجة الرياضية للمكونات اللغوية **إلى معالجة العلاقات التركيبية والمنطقية من ناحية والعلاقات الدلالية السياقية من ناحية أخرى؛ فلجأوا إلى النموذج البنيوي ذي المرجعية الأوربية للتغلب على مشكلات اللبس الدلالي.**
- **تعددت مجالات علم اللغة الحاسوبي وتحققت نجاحات كبيرة في:**

١) **التحليل الآلي للصوت البشري** وبحث الخصائص الصوتية كالجهر والهمس والترقيق والتفخيم والنبر والتنغيم والوقفات الصوتية بصورةٍ علميةٍ دقيقةٍ، ما ساعد في تعليم الأصوات وعلاج أمراض الكلام وإجراء الدراسات الصوتية على المستويات اللغوية المختلفة وإعداد الأطالس المعجمية وغيرها.

٢) **التعرف على الكلام المنطوق وتحويله إلى كلام مكتوب والعكس**، والإفادة منه في مساعدة المكفوفين في القراءة والكتابة والتعلُّم الذاتي.

٣) **التعرُّف على البصمة الصوتية** والإفادة منها في التحقيقات الجنائية، وتطوير نظم التشفير الصوتي للحسابات البنكية والهواتف الذكية ونحوها.

٤) **التحليل الآلي للصرف**، بالتعرف على صيغ الكلمات واشتقاقاتها وأوزانها وتحليلها إلى جذور ومورفيمات وصيغ تؤدي معاني وظيفية مختلفة.

٥) **الترجمة الآلية** من لغة إلى لغة.

٦) **التدقيق الإملائي وتصحيح الأخطاء النحوية**.

٧) **التنقيب في النصوص وتلخيصها** بالاكتفاء بالمعلومات الأساسية وحذف المعلومات الهامشية التي تحمل إضافات أو أمثلة أو شروح ، إلخ.

٨) **الإحصاء اللغوي** لجذور الأفعال، ولمداخل المعاجم، وللمفردات الأكثر شيوعًا في الاستعمال اللغوي، وللمتصاحبات اللفظية، وللأدوات النحوية وللتراكيب العبارية والأسلوبية المختلفة.

بحث المهتمون بحوسبة اللغة والترجمة الآلية عن نموذج يعالج مشكلات:	اللبس الدلالي للمكونات اللغوية	توليد الأنماط والأمثلة اللانهائية	العلاقات التركيبية والمنطقية للغة	الثابت والمتغير في الأبنية الصرفية
نجاح اللغة العربية في اجتياز التحدي التقني دليل على:	توحُّد الجهود العربية في حوسبة اللغة	إسهام رابطة مصنعي الحواسيب الأوربية في حوسبة العربية	نجاح المبرمجين العرب في وضع لغات برمجة بالعربية	قابلية النظام اللغوي العربي للتطويع والحوسبة
التعاون المشترك بين د. إبراهيم أنيس ود. علي حلمي موسى في دراسة الجذور العربية إحصائيًّا باستخدام الحاسوب ــ يمثل حلا للتغلب على تحدِّي	تنوع التجارب الإحصائية على أساس حاسوبي	المصطلحات الحاسوبية	غياب الخبرة اللسانية والصورية	تشتُّت الجهود المبذولة في حوسبة اللغة
تُقارب فكرة (الثابت والمتغير) في الميزان الصرفي العربي فكرة توليد الأنماط والأمثلة اللانهائية في الرياضيات؛ حيث:	تمثل الزوائد والحركات خانات المتغيرات	تمثل الزوائد والحركات خانات الثوابت	تمثل الأصول (ف ع ل) خانات الثوابت	تمثل الزوائد والأصول (ف ع ل) خانات المتغيرات
نجحت شركة (ALIS) بكندا في	تصميم نسخة عربية من نظام التشغيل ARABIC/DOS	وضع أول مواصفة لتشفير المحارف العربية	صناعة أول لوحة مفاتيح عربية- عربية	تعريب قوائم نظم التشغيل في الهواتف الذكية
يدلُّ تنوُّع المحاولات لصناعة المعجم	التعاون	تشتُّت الجهود	الإسهام	تكامل الجهود

- رأى الدكتور عبدالرحمن الحاج صالح أنه يمكن توظيف "النظرية الخليلية الحديثة" في:

- مجالات المعالجة الآلية للنصوص العربية.
- تعليم العربية.
- اصطناع الكلام المنطوق الآلي.
- معالجة أمراض الكلام.

- له دراسة إحصائية لجذور الأفعال الثلاثية في المعجم الوسيط، في ثلاثة مجلدات (د. محمد جواد النوري).
- للدكتور محمد جواد النوري دراسة إحصائية لجذور الأفعال الثلاثية في (المعجم الوسيط)
- للدكتور محمد جواد النوري دراسة إحصائية في المعجم الوسيط لـ (جذور الأفعال الثلاثية).
- دراسة الدكتور محمد جواد النوري للأفعال الثلاثية في المعجم الوسيط دراسة (صوتية صرفية).

٤,٢- الإسهامات المؤسسية في حوسبة اللغة العربية

- يمكن تصنيف الجهود المؤسسية في حوسبة اللغة العربية في اتجاهين:

الأول: يعمل على تصميم تطبيقات حاسوبية وبرمجيات باللغة العربية.

- الثاني: يعمل على تطوير نظم التشغيل بحيث تتوافق مع العربية إلى جانب اللغات الأخرى.

- من المؤسسات التي تخصصت في تطوير نظم التشغيل باللغة العربية:

- (شركة ALIS) لمؤسسها (بشير الحليمي): نجحت في تصميم نسخة عربية من نظام التشغيل ARABIC / DOS تمكن الحاسب الآلي من التعامل مع الحروف والرموز الكتابية العربية بصورة تُوائم نظام التشغيل MS/DOS في التعامل مع الرموز والحروف اللاتينية.
- (المنظمة العربية للعلوم): يرجع إليها الفضل في وضع "المعجم العربي الموحد لمصطلحات الحاسبات الإلكترونية".
- (مكتب تنسيق التعريب بالمغرب): يرجع إليه الفضل في وضع "المعجم المُوحّد لمصطلحات المعلوماتية".
- (مدينة الملك عبدالعزيز للعلوم والتكنولوجيا): يرجع إليها الفضل في إنشاء بنك "باسم" BASIM "البنك الباسمي" للمصطلحات الحاسوبية.
- (جامعة الإسكندرية): يرجع إليها الفضل في مشروع "المدونة الدولية للغة العربية".

- من سمات "المدونة الدولية للغة العربية" لجامعة الإسكندرية:

- احتواؤها على ١٠٠ مليون كلمة.
- جمعها من العربية المكتوبة.
- التمثيل الجغرافي والتنوع الموضوعي والتوازن في حجم المادة.

- يرجع الفضل في تصميم برنامج المشكل الآلي إلى (شركة RDI).
- يرجع الفضل في تصميم برنامج تحشية الفئات الكلامية إلى (شركة RDI).
- البرنامج الذي يقوم بتشكيل النص العربي الخام بنسبة دقة تصل إلى ٩٦% (برنامج المُشكِّل الآلي)
- صُمِّم للتوافق مع الصرف والتركيب العربي؛ إذ يقوم بعنونة كل فئة كلامية بإضافة رمز مميز لها عن الفئات الأخرى، كتمييز الفعل عن الاسم والصفة والحرف وغيره. (برنامج تحشية الفئات الكلامية).

- برنامج مصمم للمعاقين بصريًا والمكفوفين، يمكنهم من قراءة الكتب والمستندات المطبوعة، بالإضافة إلى الملفات الإلكترونية. (برنامج إبصار)

- من الوظائف اللغوية الحاسوبية التي يؤديها برنامج إبصار لشركة صخر Sakhr :

- قراءة الكتب والمستندات المطبوعة والملفات الإلكترونية.
- التعرف الضوئي على الحروف.
- تحويل النصوص إلى كلام منطوق.
- كتابة النصوص باللغة العربية والإنجليزية.
- حفظ النصوص وطباعتها بلغة بريل.

.. من الوظائف اللغوية الحاسوبية التي يؤديها برنامج إبصار لشركة صخر Sakhr:

- تحويل النصوص إلى كلام منطوق.
- كتابة النصوص باللغة العربية والإنجليزية.

- من الوظائف اللغوية الحاسوبية التي يؤديها برنامج إبصار لشركة صخر Sakhr تحويل النصوص المكتوبة إلى كلام منطوق.

- من الوظائف اللغوية الحاسوبية التي يؤديها برنامج إبصار لشركة صخر Sakhr كتابة النصوص بـ:

- اللغة العربية.
- الإنجليزية.

- من الوظائف اللغوية الحاسوبية التي يؤديها برنامج إبصار لشركة صخر Sakhr كتابة النصوص بـ (**اللغة العربية والإنجليزية**)

- من الوظائف اللغوية الحاسوبية التي يؤديها نظام صخر للترجمة الآلية لشركة صخر Sakhr :

- ترجمة الجمل أو الفقرات من العربية إلى الإنجليزية.
- ترجمة الجمل أو الفقرات من الإنجليزية إلى العربية.

- إسهامات جامعة ليدز في حوسبة اللغة العربية: يهتم فريق العمل بكلية الحاسب الآلي بريادة إريك آتويل Atwel Eric في جامعة ليدز Leeds بإنجلترا بإجراء بحوث وتطبيقات حاسوبية لمعالجة اللغة العربية آليًا من خلال:

- بناء مدونات عربية.
- تصميم محللات صرفية عربية.
- تصميم مفهرسات آلية للنص العربي.
- تصميم تحشيات صرفية ونحوية للمدونات اللغوية على غرار تحشيات المدونات الإنجليزية.

- من أكثر إسهامات جامعة ليدز في حوسبة العربية ريادة وأهمية مشاريع حوسبة (**القرآن الكريم**)

- من إسهامات جامعة ليدز في حوسبة اللغة العربية:

- برنامج محاورة القرآن "قرآني".
- مدونة القرآن الكريم.

- أول من اقترح معالجة اللغة العربية معالجة حاسوبية (**د. محمد كامل حسين**).

- يطلق على المحاولات الأولى في حوسبة اللغة العربية (**مرحلة الإحصاء اللغوي**).

- شهدت مرحلة الإحصاء اللغوي تعاونًا بين (**د. إبراهيم أنيس ود. علي حلمي موسى**).

- أول محاولة إحصائية لأصول المواد اللغوية في اللغة العربية في سبعينيات القرن الماضي كانت في (**الكويت**)

- أول محاولة إحصائية لأصول المواد اللغوية في اللغة العربية كانت لإحصاء جذور معجم (**الصحاح للجوهري "تاج اللغة وصحاح العربية"**).

- من سمات مرحلة الإحصاء اللغوي في الحوسبة العربية:

- التعاون بين اللغويين "إبراهيم أنيس" والفيزيائيين "علي حلمي موسى".

- إحصاء جذور معجم الصحاح للجوهري، وفتح المجال لإحصاء جذور أشهر المعاجم العربية.

- التغلب على غياب الخبرة اللسانية "نقص الخبرة اللسانية في مقابل نقص الخبرة الحاسوبية.

- فتح مجال بحثي جديد في حقل الدراسات اللغوية وفق منهج حاسوبي وأسس علمية دقيقة.

- شهدت السبعينيات عددًا من الدراسات الإحصائية لجذور بعض المعاجم اللغوية منها:

- الصحاح للجوهري.

- لسان العرب لابن منظور.

- تاج العروس للزبيدي.

- تعد تجربة الدكتور إبراهيم أنيس والدكتور علي حلمي موسى في الحوسبة الإحصائية للجذور العربية مثالاً على (**التغلب على غياب الخبرة اللسانية "نقص الخبرة اللسانية في مقابل نقص الخبرة الحاسوبية"**).

- تمثلت المحاولة الأولى في حوسبة اللغة العربية بمصر عند الدكتور علي حلمي موسى في (**ألفاظ القرآن الكريم ومقارنتها بألفاظ معجم الصحاح**).

- كان لمرحلة الإحصاء اللغوي كبير الفضل في فتح مجال بحثي جديد لمعالجة اللغة معالجة حاسوبية قائمة على:

- القوانين الرياضية.

- المنطق الصوري.

- الذكاء الاصطناعي.

- شهد علم اللغة الحاسوبي إسهامات متزايدة نظريًا وتطبيقيًا في (**الثمانينيات**).

- شهدَ علم اللغة الحاسوبي في العالم العربي -منذ الثمانينيات وحتى الآن- إسهامات متزايدة على مستوى:

- البحوث والمؤلفات باللغة العربية والإنجليزية.

- التطبيقات الحاسوبية على اللغة العربية.

- الهيئات والمنظمات والأقسام العلمية المتخصصة.

٣,١- ملامح علم اللغة الحاسوبي في التراث العربي

- تتطابق تقليبات الخليل بن أحمد مع (**الخوارزميات الرياضية "خوارزميات البرامج الحاسوبية**)

حوسبة الحرف العربي

- ✓ في أواخر الستينيات من القرن العشرين، عُقد أول مؤتمر عربي بالمغرب عام (١٩٧٦) للخبراء العرب لوضع مواصفة موحدة لإدخال العربية على النظام الحاسوبي. وانتهى المؤتمر إلى اعتماد أول معيار موحّد لتشفير الرموز الكتابية العربية (CODAR1)
- ✓ اقترح العراق معيارًا موازيًا لمعيار (CODAR1) في المؤتمر العربي الذي عُقد بتونس، وانتهت المداولات اعتماد مشروع مؤقت تحت اسم CODARZ، ثم عُرض على المؤتمر العام لليونسكو في ديسمبر (١٩٧٦) بكينيا ليستقر الرأي على تبني المشروع.
- ✓ وفي عام (١٩٧٧) عُقد مؤتمر بروما لوضع مواصفة مشتركة لتشفير الرموز العربية حاسوبيًا، ونجح المشاركون في الاتفاق على معيار جديد موحّد لترميز الكتابة العربية هو CODAR-U، ثم سُجل لدى المنظمة الدولية للتقييس ISO،
- ✓ وفي عام ١٩٨٢ اقترحت الرابطة الأوربية لمصنعي الحواسيب (ECMA) معيارًا متكاملًا لرقمنة الحرف العربي، نوقش الاقتراح في مؤتمر طارئ بالرباط ، انتهى إلى اعتماد معيار CODAR – UF/ DF ليكون مواصفة معيارية موحدة.
- ✓ نظرًا للتعدية اللغوية والهيمنة اللاتينية في عمليات الرقمنة والتشفير الحاسوبي، لم تُقبل مبادرة تصنيع حاسوب عربي أحادي اللغة، فكان لا بُد من اعتماد الرموز اللاتينية إلى جانب الرموز العربية في إعداد أية مواصفة لرقمنة الكتابة العربية، وعلى هذا وُضع المعيار العربي ASMO708 القائم على الثنائية اللاتينية العربية،
- ✓ يتكون المقياس العربي لمجموعة المحارف ٤٥، بواقع ٢٨ صوتًا صامتًا هي: (ء ب ت ث ج ح خ د ذ ر ز س ش ص ض ط ظ ع غ ف ق ك ل م ن هـ و ي)، بالإضافة إلى الحركات وعددها ٦ (الفتحة والضمة والكسرة إضافة إلى المد بالألف وبالواو وبالياء) ومحارف الضبط وعددها ٤ (التنوين والشدة والسكون وهمزة الوصل، وبتركيب هذه الرموز مع بعضها يبلغ عددها ١٠) والمحارف الخاصة وعددها ٧ (الهمزة بأشكالها المختلفة: إ أ ؤ ئ والألف المقصورة: ى والمد: آ وأخيرا التاء المربوطة: ة).

برامج حاسوبية للتحليل الصوتي

١) برنامج Audacity Beta: لعرض الصوت على شكل موجة مرئية، ثم يقابلها بالرسم الطيفي.

٢) برنامج Speech Filing System: لعرض صورة دقيقة لمستويات الصوت من حيث القوة والشدة، وشكل الموجة، وصورتها الطيفية... إلخ.

- بناء الثاني لدى الخليل بن أحمد يقابل (types و tokens)، فحساب عدد الكلمات داخل المدونة دون احتساب عدد المرات types.

- تتطابق فكرة الثابت والمتغير في الصرف العربي مع (**حدود الرياضيات والأشكال الهندسية في البرمجة**)، فالعرب القدامى وضعوا نموذجًا للوزن الصرفي هو (فعل) فكل كلمة لها ميزان صرفي، والفاء والعين واللام هي المتغير، والثابت هي الزوائد التي معها والحركات، فما يسكن في خانة الفاء والعين واللام هو الذي يتغير. أما الزوائد فهي التي تثبت.

م	ف	ع	و	ل
م	ك	ت	و	ب
م	ق	ت	و	ل

- تسعى الشركات الكبرى إلى ابتكار لغات مبرمجة متطورة في برمجة (**اللغات الطبيعية**).

- تتقارب لغة البرمجة مع اللغة الإنسانية في (**منطقيتها ورياضيتها**).

- المنهج: يتأسس علم اللغة الحاسوبي على:

- النظرية اللغوية من خلال فهم خصائص النظام اللغوي المراد حوسبته "أساس معرفي".

- تصميم التطبيقات أو برمجيات المعالجة الآلية ثانيًا "أساس هندسي حاسوبي".

- يعمل المختصون في حوسبة اللغة على تحويل خصائص (**اللغات الطبيعية إلى لغات صورية وقواعد رياضية**)

- يعمل المختصون في حوسبة اللغة على تحويل خصائص (**اللغات الطبيعية إلى لغات صورية**)

- النموذج التوليدي التحويلي لدى تشومسكي "النظرية النموذجية":

- يدرس اللغة الطبيعية.

- ذو مرجعية أمريكية.

- يحلل اللغة إلى مكوناتها التركيبية

- ساعد هذا المنهج مهندسي اللغات الطبيعية "الحاسوبيين في صياغة القواعد النحوية في صورة قواعد رياضية يتولد عنها عدد لا نهائي من التراكيب اللغوية

- إذن ساعد المنهج التوليدي التحويلي "النظرية النموذجية" لدى تشومسكي "الحاسوبيين في (**صياغة القواعد النحوية في صورة قواعد رياضية**)

- المنهج البنيوي لدى سوسير:

- "منهج أروبي.

- يراعي جانب التركيب "النحو" والمعنى"، أي العلاقات الأفقية "التركيبية النحوية الموقعية" والرأسية "الاستبدالية الدلالية المعجمية".

- عدل الحاسوبيون عن نموذج تشومسكي التوليدي التحويلي إلى نموذج دي سوسير البنيوي لـ:

- تعامل تشومسكي مع المكونات اللغوية بصورة منعزلة.

- ركز إفرهار ديترز في تعريفه علم اللغة الحاسوبي العربي على:

- الكيفية التي يعمل بها النظام اللغوي في العربية باستخدام الوسائل التقنيات الحاسوبية.
- البحث في كيفية توظيف التقنيات الحاسوبية في معالجة اللغة العربية.

س/ اقتصر إفرهار ديترز في تعريفه لعلم اللغة الحاسوبي على دراسة الكيفية التي يعمل بها النظام اللغوي في العربية باستخدام الوسائل والتقنيات الحاسوبية. (خطأ)

- يهتم الجانب التطبيقي في علم اللغة الحاسوبي بـ:

- التطبيقات الحاسوبية وتطويرها.
- التحقيق الجنائي والتلخيص الآلي وتصحيح الأخطاء الإملائية.

- يُؤَرِّخ لعصر الأتمتة "تحوّل الشيء ليعمل أوتوماتيكيًا" بـ (**الأربعينيات**).
- يُؤَرِّخ لاستخدام الحاسوب "الحاسب الآلي" في حقل الدراسات اللغوية بـ (**أوائل الخمسينيات**).
- يُؤَرِّخ لاستخدام الحاسوب "الحاسب الآلي" في حقل الدراسات اللغوية "البداية الأولية لحوسبة اللغة" بـ (**الخمسينيات**).
- يرجع لاستخدام الحاسوب في حقل الدراسات اللغوية إلى أوائل الخمسينات من القرن العشرين حينما ظهرت بدايات الترجمة الآلية في (**أمريكا**).

- يرجع استخدام الحاسوب في حقل الدراسات اللغوية إلى أوائل الخمسينات من القرن العشرين في أمريكا من خلال:

(ترجمة الوثائق الروسية لأغراض عسكرية "الترجمة الميكانيكية، الترجمة الآلية").

- يُؤَرِّخ لاستخدام الحاسوب "الحاسب الآلي" في حقل الدراسات اللغوية في أوربا بـ(**الستينيات**).
- انتقلت التجربة الأمريكية لاستخدام الحاسوب في حقل الدراسات اللغوية إلى أوربا في (**السويد**)
- أولى الجامعات في تقديم أولى محاولات المعالجة الآلية للغة الطبيعية وتحديدًا في مجال الترجمة الآلية إلى اللغة الإنجليزية جامعة (**جورج تاون الأمريكية**).
- يؤرخ للبداية الفعلية "الحقيقية" لحوسبة اللغة بـ(**مركز التحليل الآلي للغة بمدينة gallarat الإيطالية**).
- البداية الفعلية "الحقيقية لحوسبة اللغة (**إيطالية**).
- يُؤَرِّخ للبداية الحقيقية "الفعلية" لحوسبة اللغة بـ (**الستينيات**).

- من ملامح البداية الفعلية "الحقيقية" لحوسبة اللغة في الستينيات:

- وضع أسس استخدام الحاسب الآلي في معالجة اللغة.
- إنشاء عدد من المراكز الحاسوبية المتخصصة في أوربا والاتحاد السوفيتي.
- ظهر المركز الحسابي لدراسة الأدب واللغة في جامعة كيمبردج في (**الستينيات**).
- ظهر المركز المعجمي بمجمع دالال كرومسكا بإيطاليا في (**الستينيات**).
- ظهر معهد الألسنية التابع لمجمع العلوم بكييف في أوكرانيا في (**الستينيات**).

- المحاولات العربية "مرحلة الإحصاء اللغوي":

- يؤرخ لدخول علم اللغة الحاسوبي إلى اللغة العربية أو الحوسبة العربية بـ(**السبعينيات**).

التاريخي للغة العربية على:	المشترك بين اللغويين والحاسوبيين في حوسبة المعجم العربي	والتنافسية الإقليمية والرغبة في الريادة والسَّبق	المؤسسي في حوسبة اللغة العربية	لرصد التطور التاريخي للألفاظ العربية
يُؤرِّخ لأولى مراحل حوسبة اللغة العربية التي شهدت تعاونًا بين اللغويين والحاسوبيين بمرحلة:	ترميز المحارف العربية	الترجمة الآلية وفق النموذج التشومسكي	المدونات التقليدية في صورتها الورقية	الإحصاء اللغوي لجذور المعاجم العربية
اقترن استخدام الحاسوب في حقل الدراسات اللغوية في أوائل الخمسينيّات بـ:	إنشاء عدد من المراكز الحاسوبية للغات في الاتحاد السوفيتي	التأسيس لاستخدام الحاسب الآلي في معالجة اللغة بأوروبا	الترجمة الميكانيكية للوثائق الروسية بأمريكا	التحوُّل من التدوين الورقي إلى الرقمنة في المجالات المختلفة عالميًا
يُعرف وَسْم الفئات الكلامية بالمختصرات، مثل: [صفة adj] في المدونات اللغوية بـ:	التحشية الصرفية Tagging	الاختصارات Abbreviation	الاقتطاع الهجائي Archeology	الفهرسة الآلية Concordance
اعتماد المعيار العربي (ASMO708) القائم على الثنائية اللاتينية- العربية في تشفير اللغة العربية دليل على	الهيمنة اللغوية والصراع التقني	التكامل بين اللغات في الحوسبة اللغوية	الرغبة في إصدار جيل من الحواسيب أحادية اللغة	تحدي اختلاف اللغة الصورية التي تواجهه اللغات
من الإسهامات الحاسوبية لجامعة ليدز:	برنامج تحشية الفئات الكلامية Arabic)	محرك البحث النصي العربي	مدونة القرآن الكريم	برنامج المشكل الآلي

٩) **صناعة المعاجم الإلكترونية** وتطبيقاتها المختلفة على مواقع الإنترنت، أو الحواسيب المكتبية والمحمولة واللوحية، أو الهواتف الذكية.

١٠) **بناء المدونات اللغوية** بأنواعها المختلفة، وتطبيقاتها البحثية والتعليمية.

١١) **تعليم اللغة** لأبنائها ولغير الناطقين بها وتوظيفها في تنمية مهارات الاستماع والقراءة والكتابة والتقييم وتحليل الأخطاء.

الإسهامات العلمية في حوسبة العربية

- يقوم علم اللغة الحاسوبي على أساسين: نظري وتطبيقي، وهو ما يستلزم **تكاملا وتعاونا بين اللغويين والحاسوبيين** للإسهام في معالجة الخصائص اللغوية وكيفية عمل النظام اللغوي.
- وقد نجحت الجهود العربية حديثًا في تطبيق الحوسبة اللغوية في هذه المجالات على اللغة العربية، وتنوعت هذه الجهود ما بين إسهامات فردية، وبحوث أكاديمية، ومشروعات مؤسسية، منها:
- ففي الجزائر يبرز د. عبد الرحمن الحاج صالح وتصوره حول وضع "أنموذج لساني للعلاج الآلي للغة العربية"،
- و في مصر اشتُهر د. **نبيل علي بما طرحه من مشروع متكامل لحوسبة اللغة العربية بمستوياتها المختلفة ضمنها كتابه "اللغة العربية والحاسوب "**مقارنة بحوسبة اللغة الإنجليزية،
- وفي مصر أيضًا صُمِّمت مدونات لغوية عامة للغة العربية، وأنشئت عشرات الشركات والمعاهد والأقسام العلمية المهتمة بحوسبة العربية،
- وفي المملكة العربية السعودية **تبنَّت مدينة الملك عبدالعزيز للعلوم والتقنية إنشاء بنك "باسم" وهو بنك آلي للمصطلحات،**
- وفي المغرب صمَّم الأستاذ **أحمد الأخضر طريقة تكنولوجية آلية لتعريب الحاسوب،**
- وفي إنجلترا أطلقت **جامعة ليدز عددًا من التطبيقات الحاسوبية المتخصصة في تحليل اللغة العربية.**

تحديات الحوسبة اللغوية

- أبرز التحديات التي تواجه حوسبة اللغة بشكل عام، **والعربية بشكل خاص هو هيمنة اللغات الوسيطة (خاصة اللاتينية والإنجليزية)**
- قد نجحت العربية في تحدي الوجود التقني باعتماد مواصفة موحّدة لإدخال النظام الكتابي إلى الحاسب وصناعة لوحة مفاتيح عربية
- **اختلاف طبيعة اللغة البشرية من حيث قواعدها وخصائصها عن طبيعة اللغة الصورية بقواعدها الرياضية**

			العربية	
شهدت أوائل الثمانينيات	تصميم أول مدونة إلكترونية (Brown)	ظهور فكرة المدونات الورقية	الانتقاد التشومسكي لفكرة المدونات اللغوية	ظهور مصطلح المدونات اللغوية (Corpus)
تقوم حوسبة اللغة على ثنائية (اللغة والحاسوب)، وثنائية (الهيمنة والهوية) اللغوية.	صح	خطأ		
يُستخدم برنامج (Acapela Multimedia) في:	ضبط النص العربي بالشكل	تحويل النص المنطوق إلى مكتوب	ترجمة النص المكتوب من (ع) إلى (E) والعكس	تحويل النص المكتوب إلى منطوق
يؤسّس..... لمشروع متكامل نظريًا و تطبيقيًا في حوسبة اللغة العربية على المستويات المختلفة.	د. نبيل علي في كتابه "اللغة العربية والحاسوب"	د. محمد الخناش في كتابه "مقدمات في المعالجة الحاسوبية للغات الطبيعية"	د. عبدالرحمن الحاج صالح في كتابه "منطق النحو العربي والعلاج الحاسوبي"	د. عبده ذياب في كتابه "الحاسوب واللغة العربية"
يُعد....من أوائل من دعوا إلى مشروع "الذخيرة اللغوية" المعروفة حديثًا بالمدونة اللغوية (Corpus).	د. نبيل علي من مصر	د. مازن الوعر من سوريا	د. عبدالرحمن الحاج صالح من الجزائر	د. محمد الخناش من الغرب

- المفهرس السياقي للمدونات العربية.
- التحليل الصرفي "نظام التحشية الصرفية".
- يرجع الفضل في تصميم برنامج محاورة القرآن "قرآني" إلى (**جامعة ليدز**).
- برنامج للبحث على مستوى المفاهيم المعرفية في القرآن يقوم على تمثيل الإطار المعرفي لتحديد المفاهيم الأساسية "المفتاحية" في القرآن الكريم وعرض العلاقات المنطقية بين هذه المفاهيم. (**برنامج محاورة القرآن "قرآني"**)

- من الوظائف اللغوية الحاسوبية التي يؤديها برنامج محاورة القرآن "قرآني" لجامعة ليدز:

- تحديد المفاهيم الأساسية "المفتاحية" في القرآن الكريم.
- عرض العلاقات المنطقية والدلالية بين هذه المفاهيم.

- من الوظائف اللغوية الحاسوبية التي يؤديها برنامج محاورة القرآن "قرآني" لجامعة ليدز (**تحديد المفاهيم الأساسية "المفتاحية" في القرآن الكريم وعلاقاتها المنطقية والدلالية**)
- يرجع الفضل في تصميم مدونة القرآن الكريم إلى (**جامعة ليدز**).
- مُدَوَّنة مفتوحة المصدر للبحث المجاني ومحشاة نحويًا وصرفيًا لكل كلمة من كلمات القرآن الكريم (**مدونة القرآن الكريم**).
- من الوظائف التي تؤديها مدونة القرآن الكريم لجامعة ليدز وظائف (**نحوية وصرفية**)
- يرجع الفضل في تصميم المفهرس السياقي للمدونات العربية إلى (**جامعة ليدز**).
- يتيح عرض السياقات النصية للكلمة المبحوث عنها في نصوص المدونة مع إمكانية إجراء العلميات الإحصائية (**المفهرس السياقي للمدونات اللغوية بجامعة ليدز**)

- من الإمكانات التي يتيحها المفهرس السياقي للمدونات اللغوية بجامعة ليدز:

- عرض السياقات النصية للكلمة.
- إجراء العمليات الإحصائية.

تتم عملية ترميز الكلمات في نظام التحشية الصرفية بجامعة ليدز بـ

- دمج كل الرموز الدالة على المكونات الصرفية في الكلمة الواحدة.
- إدراج الكلمة في سياقها المجاور للكلمة الأخرى.
- الإشارة إلى المعلومات الصرفية.

- نجحت اللغة العربية في تحدي الوجود التقني بالتصديق على ترميز المحارف العربية وصناعة الكيبورد العربي في (**السبعينيات**)
- يُؤرِّخ لِعقدِ أول مؤتمر عربي لوضع مواصفة مُوحدة لإدخال العربية على النظام الحاسوبي بـ(**الستينيات**).
- اعتمدت أبجدية عربية للرموز الوظيفية وعلامات التشكيل، متفقة مع الأبجدية الدولية، متطابقة مع المواصفة الدولية (ISO/IEC646).

في البحث اللغوي، فاستعان د. إبراهيم أنيس بالدكتور علي حلمي موسى في **دراسة إحصائية لجذور معجم الصحاح باستخدام الكمبيوتر، فيما عُرف بمرحلة الإحصاء اللغوي.**

- **منذ الثمانينيَّات وحتى الآن، يشهد علم اللغة الحاسوبي في العالم العربي إسهامات متزايدة على الجانبين النظري والتطبيقي.** وقد تنوعت هذه الجهود لتشمل ١) بحوثا ومؤلفات علمية، ٢) تصميمات للبرامج والتطبيقات الحاسوبية، ٣) منظمات وأقسام علمية متخصصة في علم اللغة الحاسوبي.
- المدقق في الموروث العربي تستوقفه بعض المنهجيات التي تدل على عبقرية العقلية العربية في التعامل مع المادة اللغوية وفق نظام رياضي منطقي.
- يُستدل على هذا بمنهج الخليل بن أحمد الفراهيدي في جمع مداخل معجم حيث قدّم الخليل مجموعة من **الأسس والمعايير لا تقل في منطقيَّتها ورياضيتها عن خوارزميات البرمجيات الحاسوبية بمفهومها الحديث.**
- **ومن هذه الأسس نظام التقليبات** – أي تقليب الجذور العربية – بحيث يكون الحاصل جميع الصور الممكن توليدها من الجذر الواحد، **ومن ثمّ فحاصل تقليب الجذر الثنائي صورتان، والثلاثي تسع صور، والرباعي اثنتا عشرة، والخماسي أربع وعشرون. ثمّ يُقسم حاصل الصور الناتجة عن عملية التقليب إلى مهمل ومستعمل، ومن خوارزمية صوتية واستعمالية.**
- اعتماد الخليل على الأبنية الصرفية في الترتيب الداخلي لمواد المعجم؛ **إذ اعتمد على ستة أبنية** هي الثنائي والثلاثي الصحيح، والثلاثي المعتل واللفيف المفروق والمقرون، والرباعي، **وعدّ الرباعي الذي يتألف من أصلين مكررين مثل (عسعس) من بناء الثنائي، وهو ما يتفق مع ما استقرت عليه برمجيات المدونات اللغوية حديثا في التفريق بين Types و Tokens؛**
- وإذا دققنا النظر في **نظام الدوائر العروضية في التراث العربي**، لالتمسنا دليلًا آخر على عبقرية العقلية العربية في تأليف نظام منطقي يُحاكي الأنظمة الحاسوبية في دقته ومنطقيته في توزيع أوزان الإيقاع الشعري وتقسيمه إلى بحور شعرية لكل منها وزنه وتفعيلاته وإيقاعه الموسيقي.
- **فكرة "الثابت والمتغير" في الصرف العربي تقف دليلا قويًا على عبقرية اللغويين العرب القدامى؛ إذ وضعوا رموزًا للأوزان الصرفية (هي الفاء والعين واللام)** التي تمثل **البنية الرمزية** لأبواب الفعل وصيغ المشتقات والأوزان الصرفية بشكل عام. **فصيغة مِفعال على سبيل المثال تقوم على ثوابت هي الحركات والسكنات والزوائد، ومتغيرات هي الحروف الأصول (ف، ع، ل)**، وبهذا التكوين يمكن توليد عدد لانهائي من أمثلة هذه الصيغة. **هذه المنهجية في التجريد التمثيلي تقارب ما يُعرف في الرياضيات بحدود الرياضيات والأشكال الهندسية في تخليق الأنماط والأمثلة التي تتولد عنها.**
- **اللغة الطبيعية** هي منظومة من العلاقات أو الرموز اللغوية التي يستخدمها الإنسان في التواصل والأغراض المختلفة، ومن أمثلتها العربية والإنجليزية وغيرها.
- **اللغة الصورية** هي منظومة من الرموز تحكمها قواعد صورية و قوانين رياضية وعلاقات منطقية، ومن أمثلتها لغات البرمجة، والتكويد، والتشفير.

- البرنامج الذي يقوم بعنونة كل فئة كلامية بإضافة رمز مميز لها عن الفئات الأخرى، كتمييز الفعل عن الاسم والصفة والحرف وغيره. (**برنامج تحشية الفئات الكلامية**)

- يقوم برنامج تحشية الفئات الكلامية لشركة RDI بـ (**عنونة الفئات الكلامية وتمييزها**)

- برنامج يتيح إمكانية البحث الاشتقاقي في النصوص العربية بالجذر اللغوي أو التركيب أو فئة الكلمة.(**برنامج مُحَرِّك البحث النصي العربي**)

- يقوم برنامج مُحَرِّك البحث النصي العربي لشركة RDI بـ(**إمكانية البحث الاشتقاقي في النصوص العربية بالجذر اللغوي أو التركيب أو فئة الكلمة**).

- يتيح برنامج مُحَرِّك البحث النصي العربي لشركة RDI البحث الاشتقاقي في النصوص العربية بـ:

- الجذر اللغوي.
- التركيب.
- فئة الكلمة.

- برنامج يقوم بتحليل الكلمات العربية إلى صيغها الصرفية وربطها بالمادة المعجمية التي تنتمي إليها والمرتبة وفق الجذر اللغوي. (**برنامج** Arab Dictions)

- يقوم برنامج Arab Dictions لشركة RDI بـ(**تحليل الكلمات العربية إلى صيغها الصرفية وربطها بالمادة المعجمية**)

- يقوم برنامج Arab Dictions لشركة RDI بـ:

- تحليل الكلمات العربية إلى صيغها الصرفية.
- الربط بالمادة المعجمية.

- برنامج يحول النصوص المكتوبة إلى كلام منطوق، ويعتمد على برنامج المشكل الآلي (**برنامج** TALK ARAB)

- برنامج ARAB TALK لشركة RDI يقوم بـ(**تحويل النصوص المكتوبة إلى كلام منطوق**)

- برنامج ARAB TALK لشركة RDI يعتمد على (**برنامج المُشَكِّل الآلي**)

- برنامج ARAB TALK:

يُحَوِّل النصوص المكتوبة إلى كلام منطوق.

- يعتمد على برنامج المُشَكِّل الآلي.
- برنامج ARAB TALK يُحَوِّل النصوص المكتوبة إلى كلام منطوق.

- البرمجيات التي أنتجتها شركة صخر Sakhr

- من الإسهامات المؤسسية في حوسبة اللغة العربية شركة صخر Sakhr:

- برنامج التدقيق الإملائي وبرنامج التشكيل الآلي.
- برنامج المحلل الصرفي وبرنامج استخراج الكلمات المفتاحية.
- محرك صخر للتلخيص العربي.
- برنامج إبصار ونظام صخر للترجمة الآلية.

٤- المدونة المتوازية:

- **تتكون من مدونتين فرعيتين أو أكثر تتضمن نصوصًا بعينها** مترجمة من لغة ما إلى لغات أخرى كترجمات القرآن الكريم الإنجليزية مثلا أو ترجمة الأعمال الأدبية ككتاب ألف ليلة وليلة من العربية إلى الإنجليزية. **تستخدم المدونات المتوازية في تعليم اللغات، ودراسات الترجمة التي تبحث في المكافئات اللغوية** لبعض التعبيرات والتراكيب العبارية في اللغات المختلفة، وكذلك الفوارق والخصائص اللغوية بين اللغات.

٥- المدونة التعليمية:

- **تتكون من كتابات متعلمي اللغة الأجنبية،** بغرض إجراء مقارنات بينها وبين النظام اللغوي المعياري في اللغة المتعلمة، للإفادة منها في رصد الخصائص اللغوية لدى متعلمي اللغة من الأجانب، ومعالجة الأخطاء، وتطوير المقررات وطرق التدريس.. إلخ. **من أمثلة المدونات التعليمية في اللغة العربية "المدونة اللغوية لمتعلمي اللغة العربية"** Arabic learner Corpus – ALC، ويمكن توظيف بياناتها في:

٦- المدونة المنهجية: وتحتوي على المواد أو المقررات التي تعليمية التي يدرسها التلاميذ ومتعلمو اللغة؛ بغرض مقارنة السياقات التعليمية المختلفة التي ترد فيها كلمة أو عبارة ما، وكذلك الوقوف على مدى واقعية اللغة التي تُقدم بها المناهج التعليمية من خلال مقارنة المدونة المنهجية بمدونات حديثة تعكس اللغة المستخدمة في المجتمع.

٧- المدونة التاريخية: **تحتوي على نصوص من فترات زمنية متعاقبة**؛ بغرض تتبع مظاهر التغير الطارئ على اللغة عبر الزمن، ورصدها والإفادة منها في التأليف المعجمي والدراسات اللغوية التاريخية. **من أمثلة المدونات التاريخية مدونة (Helslnkl Corpus)** التي تتكون من ١,٥ مليون كلمة.

٨- المدونة المحدَّثة: وهي مدونة تصمم بحيث يتم تحديثها شهرًا أو سنويًا، بغرض رصد التغيرات اللغوية الحديثة الطارئة على الاستعمال اللغوي، ومن هنا فهي متنامية الحجم بشكل كبير، وتضاف إليها النصوص الجديدة بنفس النسبة الثابتة لكل جنس لغوي.

أسئلة المحاضرات للمراجعة

-اقتصر تعريف براون وميلر لعلم اللغة الحاسوبي على توظيف الحاسب الآلي في دراسة علم اللغة.(خطأ).

- ركز ديفيد كريستال في تعريفه لعلم اللغة الحاسوبي على كيفية معالجة (**اللغة الطبيعية**).

- اقتصر تعريف ديفيد كريستال لعلم اللغة الحاسوبي على المفاهيم والتقنيات الحاسوبية في معالجة الوظائف اللغوية والصوتية.(خطأ)

- أسَّس علم اللغة الحاسوبي العربي على مفهوم علم اللغة مع إضافة متغيرين أساسيين هما "الحاسوبي" و"العربي".(**إفرهار ديترز في "موسوعة اللغة العربية وعلم اللغة"**)

- البرنامج الذي يقوم بعنونة كل فئة كلامية بإضافة رمز مميز لها عن الفئات الأخرى، كتمييز الفعل عن الاسم والصفة والحرف وغيره. (**برنامج تحشية الفئات الكلامية**)

- يقوم برنامج تحشية الفئات الكلامية لشركة RDI بـ (**عنونة الفئات الكلامية وتمييزها**)

- برنامج يتيح إمكانية البحث الاشتقاقي في النصوص العربية بالجذر اللغوي أو التركيب أو فئة الكلمة.(**برنامج مُحَرِّك البحث النصي العربي**)

- يقوم برنامج مُحَرِّك البحث النصي العربي لشركة RDI بـ(**إمكانية البحث الاشتقاقي في النصوص العربية بالجذر اللغوي أو التركيب أو فئة الكلمة**).

- يتيح برنامج مُحَرِّك البحث النصي العربي لشركة RDI البحث الاشتقاقي في النصوص العربية بـ:

- الجذر اللغوي.
- التركيب.
- فئة الكلمة.

- برنامج يقوم بتحليل الكلمات العربية إلى صيغها الصرفية وربطها بالمادة المعجمية التي تنتمي إليها والمرتبة وفق الجذر اللغوي. (**برنامج Arab Dictions**)

- يقوم برنامج Arab Dictions لشركة RDI بـ(**تحليل الكلمات العربية إلى صيغها الصرفية وربطها بالمادة المعجمية**)

- يقوم برنامج Arab Dictions لشركة RDI بـ:

- تحليل الكلمات العربية إلى صيغها الصرفية.
- الربط بالمادة المعجمية.

- برنامج يحول النصوص المكتوبة إلى كلام منطوق، ويعتمد على برنامج المشكل الآلي (**برنامج TALK ARAB**)

- برنامج ARAB TALK لشركة RDI يقوم بـ(**تحويل النصوص المكتوبة إلى كلام منطوق**)

- برنامج ARAB TALK لشركة RDI يعتمد على (**برنامج المُشَكِّل الآلي**)

- برنامج ARAB TALK:

يُحَوِّل النصوص المكتوبة إلى كلام منطوق.

- يعتمد على برنامج المُشَكِّل الآلي.
- برنامج ARAB TALK يُحَوِّل النصوص المكتوبة إلى كلام منطوق.

- البرمجيات التي أنتجتها شركة صخر Sakhr

- من الإسهامات المؤسسية في حوسبة اللغة العربية شركة صخر Sakhr:

- برنامج التدقيق الإملائي وبرنامج التشكيل الآلي.
- برنامج المحلل الصرفي وبرنامج استخراج الكلمات المفتاحية.
- محرك صخر للتلخيص العربي.
- برنامج إيصار ونظام صخر للترجمة الآلية.

مفهوم علم اللغة الحاسوبي

- علم اللغة الحاسوبي **أحد مجالات علم اللغة التطبيقي، وهو مجال بيني تلتقي فيه المعارف والقوانين اللغوية التي يهتم بها علم اللغة مع التطبيقات والقوانين الرياضية المنطقية التي تهتم بها علوم الحاسب.**
- يهتم بمعالجة الظواهر **اللغوية حاسوبيًا، وبتصميم برمجيات حاسوبية لتحليل النظام اللغوي بمستوياته المختلفة**، وتطبيقها في مجالات البحث اللغوي، وتعليم اللغة، والترجمة الآلية، والتحقيق الجنائي، وصناعة المعاجم الإلكترونية، والتلخيص الآلي، وتصحيح الأخطاء الإملائية.
- عند براون وميلر - في معجم كيمردج لعلم اللغة - : هو توظيف الحاسب الآلي في **دراسة علم اللغة وتطوير برمجيات تؤدي المهام اللغوية** مثل الإعراب ومعالجة المادة اللغوية.
- ديفيد كريستال - في معجم علم اللغة والصوتيات - فيعرّف علم اللغة الحاسوبي بأنه فرع من علم اللغة **تُطبَّق فيه المفاهيم والتقنيات الحاسوبية في معالجة الوظائف اللغوية.**
- وفي موسوعة اللغة العربية وعلم اللغة، يؤسس إفرهار ديترز المقصود بعلم اللغة الحاسوبي العربي على مفهوم علم اللغة مع إضافة متغيرين أساسيين هما "الحاسوبي" و"العربي"، فيقول: إن علم اللغة الحاسوبي العربي يعني ١) **دراسة الكيفية التي يعمل بها النظام اللغوي في العربية باستخدام الوسائل والتقنيات الحاسوبية؛ ٢) البحث في كيفية توظف التقنيات الحاسوبية في معالجة اللغة العربية.**
- لعلم اللغة الحاسوبي جانبان: نظري وعملي؛ **يهتم الجانب النظري بوضع نظريات وقواعد تحقق غاياته في حوسبة اللغة، في حين يركز الجانب العملي على تطبيق هذه النظريات على اللغة باستخدام التقنيات الحاسوبية.**
- **اختراع الحاسب الآلي عام ١٩٤٨** – ليمثل توجها عالميا متسارعا إلى التحول من **عصر التدوين الورقي والعمل اليدوي إلى عصر الرقمنة والحوسبة (الأتمتة) في كل مناحي الحياة** المعلوماتية والتجارية والعلمية والخدمية والاتصالية والأرشيفية وغيرها من الاهتمامات الإنسانية.
- **ويرجع استخدام الحاسوب في حقل الدراسات اللغوية إلى أوائل الخمسينات** من القرن العشرين حينما **ظهرت بدايات الترجمة الآلية في أمريكا** وبخاصة ترجمة الوثائق الروسية لأغراض عسكرية (فيما عُرف بالترجمة الميكانيكية). ثم **انتقلت التجربة في الستينات إلى أوربا وتحديدًا في السويد عام ١٩٦١.**
- **يؤرَّخ للبداية الفعلية لحوسبة اللغة أوروبيا في أوائل الستينات بإنشاء عدد من المراكز الحاسوبية للغة في أوربا والاتحاد السوفيتي** منها المركز الحسابي لدراسة الأدب واللغة في جامعة كيمبردج.
- **أما عربيًا، فيمكن التأريخ لأول محاولة عربية لمعالجة اللغة حاسوبيًا باقتراح قدمه الطبيب محمد كامل حسين في أوائل السبعينيات إلى اللغوي المصري الدكتور إبراهيم أنيس** عن إمكانية توظيف تقنيات الحاسوب

- صاحب كتاب "اللغة العربية والحاسوب" هو (**د. نبيل علي**)

- أول من تناول مفهوم الهندسة اللسانية العربية بجميع مستوياتها. (**د. نبيل علي**)

- أول من حاول إمكانية تطويع الحاسوب للنظام اللغوي وليس العكس. (**د. نبيل علي**)

- رأى أن العلاقة بين اللغة والهندسة الحاسوبية علاقة متبادلة يستخدم فيها الحاسوب بشكل أساسي في تصميم نماذج تحليلية للمستويات اللغوية المختلفة. (**د. نبيل علي**)

- أولى د. نبيل علي أهمية خاصة لـ (**الصرف الحاسوبي**)

- من أهم أعمال د. نبيل علي في حوسبة اللغة العربية:

- أسهم في تصميم أول محرك بحث عربي على أساس صرفي. (**د. نبيل علي**)

- أسهم في أول قاعدة بيانات معجمية للغة العربية. (**د. نبيل علي**)

- أسهم في برنامج صخر للقرآن الكريم. (**د. نبيل علي**)

- أسهم في بناء قاعدة بيانات للشعر العربي. (**د. نبيل علي**)

- أسهم في برنامج الإعراب الآلي. (**د. نبيل علي**)

- أسهم برنامج التشكيل الآلي. (**د. نبيل علي**)

- أسهم في نموذج العمل المتقدم لتعليم العربية وتعلمها. (**د. نبيل علي**)

- مؤلف كتاب "العرب وعصر المعلومات". (**د. نبيل علي**)

- مؤلف كتاب "العقل العربي وعالم المعرفة". (**د. نبيل علي**)

- أجرى دراسة لمنظمة إسكوا عام ٢٠١٠م عن "مسح للمحتوى الرقمي العربي". (**د. نبيل علي**)

- مؤلف كتاب "اللغة العربية والحاسوب" (**د. نبيل علي**)

- مؤلف كتاب "الحاسوب واللغة العربية" (**د. عبده ذياب العجيلي**)

- مؤلف كتاب "العربية – نحو توصيف جديد في ضوء اللسانيات الحاسوبية" (**د. نهاد الموسى**)

- اشتمل كتاب "العربية – نحو توصيف جديد في ضوء اللسانيات الحاسوبية" للدكتور نهاد الموسى على رؤى حاسوبية حاول إسقاطها على أنظمة العربية خاصة:

- النحو "الإعراب".

- الصرف "البنية".

- المعجم.

- التصويب اللغوي "الأخطاء النحوية، والصرفية، والإملائية.

- حاول د. نهاد الموسى في كتابه "العربية – نحو توصيف جديد في ضوء اللسانيات الحاسوبية" الانتقال **(من وصف العربية إلى توصيفها)**.

- من الدراسات المهمة في حوسبة اللغة العربية لد/ محمد مرياتي:

- دراسة إحصائية للجذور العربية في المعاجم القديمة.

وعلاقاتها المنطقية. وهي قضية متصلة بتحدي الترميز أو التكوين والتشفير – أي **تحويل القواعد التي تحكم عمل النظام اللغوي إلى قواعد رياضية بحيث يفهمها الحاسوب ويعالجها**، وقد ترتب على هذا الاختلاف اللغوي الحاسوبي **اختلاف مفهومي؛ إذ إن تعريف الحاسوبيين للوحدات اللغوية (ومنها الكلمة، والجملة، والمركب اللغوي)**

- السبب الرئيسي الذي دعا تشومسكي وضع أسس منهجه النموذجي في معاجلة اللغة الطبيعية – هو **سعيه الدائم إلى التأقلم مع متطلبات المعالجة الآلية للغات الطبيعية**
- **غياب الخبرة اللسانية** لدى الحاسوبيين وغياب الخبرة الحاسوبية لدى اللغويين **مما يستلزم تعاونا مشتركا كما حدث بين د إبراهيم أنيس ود علي حلمي موسى في دراسة الجذور العربية.**
- **تحدي المصطلحية الحاسوبية**، فلا تزال العربية تُعاني من قصور في استيعاب المصطلح الحاسوبي ودلالاته المستحدثة والمتطورة بشكل متسارع، وإن كانت هناك جهود محمودة في تعريب قوائم نظم التشغيل
- ظهرت مناهج ما بعد التوليدية، منها منهج النحو التأليفي الذي أصبح متخصصًا في الهندسة اللسانية ذات التوجه الاستقرائي خلافًا للنظرية التوليدية التي باتت تُصنف في إطار التوجه المفهومي بمعناه المنطقي.
- التحدي الأبرز الذي يواجه حوسبة العربية وهو السبب لكثير من التحديات والمشكلات والمعوقات- هو **تشتت الجهود المبذولة في محال علم اللغة الحاسوبي في العالم العربي وتنوعها لحد يبلغ التشابه والتعارض في الوقت نفسه**؛ لأنها تقف وراءها غايات تنافسية وإقليمية بغرض إحراز السبق أو الرغبة في الزيادة، أو الأغراض الربحية، ومثالها برامج التحليل اللغوي مغلقة المصدر، وكذلك محاولات بناء مدونات لغوية عامة ممثلة للغة العربية ، والمحاولات المتفرقة في سبيل صناعة معجم تاريخي للغة العربية.

الجهود الفردية والمؤسسية والدولية في حوسبة العربية

د. نبيل علي - مصر

- ✓ ألف كتاب: "**اللغة العربية والحاسوب**" الصادر في سنة ١٩٨٨م
- ✓ **أسس لمشروع متكامل نظريًا وتطبيقيًا** لحوسبة اللغة العربية صوتيًا، وصرفيًا، ونحويًا، ومعجميًا مقارنة بحوسبة اللغة الإنجليزية
- ✓ أول من تناول مفهوم الهندسة اللسانية العربية بجميع مستوياتها وإمكانية تطويع الحاسوب للنظام اللغوي وليس العكس.
- ✓ أسهم د. نبيل علي في **تصميم أول محرك بحث عربي على أساس صرفي**، وأول قاعدة بيانات معجمية للغة العربية، وبرنامج صخر للقرآن الكريم، **وقاعدة بيانات للشعر العربي، وبرنامج الإعراب الآلي**، وبرنامج التشكيل الآلي، ونموذج العمل المتقدم لتعليم العربية وتعلمها.

٣) برنامج Ic speech analyzer: لرسم الموجة الصوتية بيانيًا، وتحليل طيفها رقميًا، والتوزيع الطيفي للإشارة الصوتية... إلخ.

٤) برنامج Praat:

- يُحلّل الموجات الصوتية والخصائص الفيزيائية للصوت.
- يجمع Praat بين الدراستين الطيفية والاهتزازية.
- يدعم كتابة التعليقات بالعربية، إلا أن واجهته لم تعرب بعد
- الكتابة العربية تقرأ في هذا البرنامج من اليسار إلى اليمين.
- يقيس طول الموجة، سعة الموجة، التردد أ ، الصورة الطيفية للصوت، التنغيم).

- للتحليل الآلي للصوت البشري مجالات تطبيقية متعددة، منها (علم اللغة الجنائي) في إثبات هوية المتحدث، من خلال تحليل البيانات الصوتية المقترنة بالحدث الجنائي.
- ويستعان في عملية التعرف على هُويَّةِ المتحدث بطريقتين أساسيتين، هما: **١) التعرف السماعيّ:** عن طريق ما يعرف بطابور الشخصية الصوتي لشهود الاستماع؛ حيث يعرض صوت المتهم مع عدة أصوات مشابهة له على شريط مسجل في شكل طابور، ليميز من بينها صوت الجاني و **٢) استخدام التقنيات الحديثة في التحليل الصوتي:** وتشمل هذه الطريقة ما يعرف بالبصمة الصوتي. وقد ساعد في تقدم هذا الفرع تطور تسجيل الأصوات وتحليلها إلكترونيًا، بحيث أصبح بالإمكان تسجيل الأصوات وتحويلها إلى أشكال بيانية.
- ومن البرامج الحاسوبية في معالجة الصوت: برنامج النطق الآلي للنصوص Text to speech: لتحويل أي نص مكتوب إلى نص منطوق بصوت بشري مركب، ولا يدعم اللغة العربية.
- **أما البرامج التي تدعم النصوص العربية فمن أشهرها:**

١) برنامج Acapela Multimedia، للقراءة الآلية للنصوص، صممته شركة Acapela عام ٢٠٠٣، ثم صدرت النسخة العربية منه عام ٢٠٠٤. ويتيح البرنامج خيارات للنطق بالعربية بصوت رجل وسيدة.

٢) برنامج Oddcast 0.5: للقراءة الآلية للصوت، صممته شركة Oddcast في نيويورك عام ١٩٩٩، ويدعم اللغة العربية.

ومن برامج التعرف على الصوت العربي المنطوق وتحويله إلى نص مكتوب بدقة عالية.

٣) برنامج Google Translate 0.6

٤) وبرنامج Talk Typer، وهما من التطبيقات التي تستخدم أون لاين، ويدعمان اللغة العربية.

٥) برنامج ReadSpeaker TTS

التحليل الحاسوبي للمدونات اللغوية

- **في الأربعينيات** ظهرت فكرة المدونات (أي قبل شيوع الاتجاه التشومسكي في التفكير اللغوي).
- في **أواخر الخمسينيات** تعرضت فكرة استخدام المدونة في إصدار الأحكام اللغوية إلى انتقاد شديد – خاصة من قبل تشومسكي.
- **أواخر الستينيات** بناء أول مدونة إلكترونية **للإنجليزية الأمريكية** هي مدونة (Brown).
- **أوائل الثمانينيات** ظهر مصطلح المدونات اللغوية.

- **مصطلحيًا، يفرق اللغويون بين المدونة** corpus، **ولغويات أو لسانيات المدونة** corpus linguistics:

١- **المدونة اللغوية:** قاعدة بيانات إلكترونية تحتوي على عينة كبيرة من النصوص (المكتوبة أو المنطوقة) التي تعكس الاستعمال الحي للغة بغرض تحليلها تحليلا لغويًا باستخدام برامج وأدوات تحليل متخصصة.

٢ - **أما لغويات المدونة فهي** دراسة اللغة على أساس أمثلة وعينات حية تعكس واقع الاستعمال اللغوي.

من أشهر أنواع المدونات ما يلي:

١- المدونة الخاصة: وتتضمن نصوصًا تمثل جنسًا لغويًا معينًا؛ كالصحافة أو الشعر أو الرواية أو القانون أو الكتابات الأكاديمية أو مواد من مواقع التواصل الاجتماعي أو تفريغات لمواد منطوقة.. إلخ.

- من أشهر البرامج مفتوحة المصدر في بناء المدونات العربية الخاصة وتصفحها **(غواص Ghawaas 4.0)**، وهو من تصميم **مدينة الملك عبدالعزيز للعلوم والتكنولوجيا**، ومزود بمقاييس إحصائية، ويتوافق مع العربية والإنجليزية.

٢- المدونة العامة (المدونة المرجعية): تحتوي على نصوص من أجناس لغوية مختلفة، وقد تتضمن نصوصًا مكتوبة وأخرى منطوقة، لتمثل الواقع اللغوي في دولة بعينها أو دول متعددة.

- من أشهر المدونات العربية العامة:

- **المدونة الدولية للغة العربية**، التي قام ببنائها فريق عمل بمكتبة الإسكندرية؛ إذ بدأت مكتبة الإسكندرية في مشروع بناء المدونة الدولية للغة العربية، تتضمن ١٠٠ مليون كلمة،

- **المدونة اللغوية العربية لمدينة الملك عبدالعزيز للعلوم والتقنية (المدونة العربية) التي تحوي مليار كلمة مما دُوِّنَ بالعربية** ابتداءً من العصر الجاهلي وحتى العصر الحديث.

٣- المدونة المقارنة:

- تتضمن مدونتين فرعيتين تمثلان لغتين مختلفتين كالعربية الإنجليزية، أو تنوعين لغويين كالعربية المصرية والعربية المغربية. ويجب أن تتضمن المدونتان نفس الجنس اللغوي وبنفس الحجم. تُستخدم المدونات المقارنة في دراسة خصائص التنوعات اللغوية، وإجراء بحوث تتعلق بالترجمة، أو تعليم اللغات. من أشهر أمثلتها مدونة (ICE)، التي تتضمن مليون كلمة لكل تنوع من تنوعات اللغة الإنجليزية.

- ✓ د.عبده ذياب العجيلي: ألف كتاب "الحاسوب واللغة العربية" ١٩٩٦ عرض فيه أهم التطبيقات الحاسوبية على مستويات اللغة العربية المختلفة، والترجمة الآلية.
- ✓ د. نهاد الموسى: ألف كتاب "**العربية – نحو توصيف جديد في ضوء اللسانيات الحاسوبية**" سنة ٢٠٠٠، يشتمل على رؤى حاسوبية حاول المؤلف إسقاطها على أنظمة العربية، خاصة النحو (الإعراب)، والصرف (البنية)، والمعجم، إضافة إلى التصويب اللغوي (الأخطاء النحوية، والصرفية، والإملائية).
- ✓ د. محمد مرياتي (سوريا) قدّم **دراسة إحصائية للجذور العربية في المعاجم القديمة باستخدام الحاسبات الإلكترونية في حساب النسبة لدوران الحروف في الجذور العربية**، وللمعجم العربي، ولدوران الحروف العربية المشكولة ولحروف اللغة العربية
- ✓ د. محمد الحناش (المغرب) قدم عددًا من الدراسات حول **المعجم الحاسوبي للغة العربية ضمنها كتابه "المعجم التركيبي للغة العربية"** الصادر سنة ١٩٩٢.
- ✓ ومجمل أعماله حول **بناء المعاجم الإلكترونية للعربية**،
- ✓ كما أسهم في بناء قاعدة لبيانات حاسوبية لتراكيب القاموس المحيط، وبناء محلل صرفي للغة العربية،
- ✓ د. مازن الوعر (سوريا) اهتم بمجال المعالجة الآلية لنحو العربي، منها "التوليد الصوتي والنحوي والدلالي لصيغ المبني للمجهول في اللغة العربية – معالجة لسانية حاسوبية" ٢٠٠٢.
- ✓ د. عبد الرحمن الحاج صالح (الجزائر) أسهم في مجال حوسبة النحو من خلال بحثه **"منطق النحو العربي والعلاج الحاسوبي"**.
- ✓ من أوائل من دعوا إلى **مشروع "الذخيرة اللغوية" - المعروفة حديثًا بالمدونة اللغوية** التي توفر للباحثين إمكانية البحث في النصوص اللغوية المختزنة في ذاكرة الحاسوب.
- ✓ **عالج النظرية الخليلية الحديثة التي يُعتمد عليها في مجالات المعالجة الآلية للنصوص العربية**، وتعليم العربية واصطناع الكلام المنطوق الآلي
- ✓ من المؤسسات التي تخصصت في تطوير نظم التشغيل بالعربية شركة "ALIS" لمؤسسها الجزائري بشير الحليمي في كندا، التي نجحت في **تصميم نسخة عربية من نظام التشغيل** ARABIC/DOS تمكّن الحاسب الآلي من التعامل مع الحروف والرموز الكتابية العربية
- ✓ المنظمة العربية للعلوم لعبت دورًا في **تشفير الحرف العربي ووضع المعجم العربي الموحد لمصطلحات الحاسبات** الإلكترونية عام ١٩٨١ والذي يضم ٣٤١٤ مصطلحا،
- ✓ مكتب تنسيق التعريب بالمغرب وضع **"المعجم الموحد لمصطلحات المعلوماتية"** عام ٢٠٠٠.
- ✓ مدينة الملك عبدالعزيز للعلوم والتكنولوجيا بالرياض اهتمت **بإنشاء بنك "باسم"** BASIM **للمصطلحات الحاسوبية عام 1994**، وهو بنك موسوعي لمصطلحات علوم الحاسب
- ✓ جامعة الإسكندرية قامت بتصميم مشروع **"المدونة الدولية للغة العربية"** International Corpus of Arabic تحتوى على ١٠٠ مليون كلمة، جُمعت من العربية المكتوبة

- إغفال تشومسكي المعنى.
- التغلب على مشكلات اللبس الدلالي والترجمة.
- اعتمد الحاسوبيون على منهج دي دي سوسير البنيوي في بناء المدونات اللغوية التي تقرأ (**أفقيًا "في المفهرسات اللغوية" ورأسيًا لتحليل الدلالات السياقية**).

من أبرز التحديات التي تواجه حوسبة اللغة بشكل عام والعربية بشكل خاص- هيمنة (**اللغات الوسيطة "خاصة اللاتينية والإنجليزية**

- من أبرز التحديات التي تواجه حوسبة اللغة العربية هيمنة اللغات الوسيطة في مراحل:

- الترميز.
- الإدخال.
- تصميم نظم التشغيل للتطبيقات الحاسوبية.

- **الاختصارات الكتابية**: يطلق على الاكتفاء بحرف أو أكثر من الكلمة عوضًا عن ذكرها كاملة.
- الحروف الأوائلية: وهو اقتطاع أول حرف من كل كلمة من المصطلح المركب للدلالة عليه بالصيغة المختصرة.
- الرموز الأركرولوجية مثل التقويم والياء اللينة والياء الشامية (ي) والياء اللينة (ى).

- نجحت اللغة العربية في تحدي الوجود التقني باعتماد:

- مواصفة موحدة لإدخال النظام الكتابي إلى الحاسب.
- صناعة لوحة مفاتيح عربية Arabic keyboard .

٢.٢- تحديات الحوسبة اللغوية:

- **اختلاف طبيعة اللغة البشرية عن طبيعة اللغة الصورية**: فقد أدى اختلاف طبيعة اللغة البشرية من حيث قواعدها وخصائصها عن طبيعة اللغة الصورية بقواعدها الرياضية وعلاقاتها المنطقية إلى اختلاف (**مفهومي**)
- **غياب الخبرة اللسانية**: ويمثل ذلك تجربة الدكتور أنيس والدكتور علي حلمي موسى.
- **تحدي المصطلحية الحاسوبية**: ولا تزال العربية تعاني من قصور في استيعاب المصطلح الحاسوبي ودلالاته المستحدثة والمتطورة بشكل متسارع.
- **تشتت الجهود المبذولة**: فهي تعمل في جزر منعزلة، وكثير منها للأسف الشديد – تقف وراءه غايات تنافسية وإقليمية بغرض إحراز السبق أو الرغبة في الزيادة، أو الأغراض الربحية.

- من الذين حاولوا التقريب بين اللغويين والحاسوبيين في المفاهيم (**المدرسة التوليدية التحويلية "تشومسكي"**)

- د. نبيل علي (مصر): تعد جهوده حجر الزاوية في مسيرة البحث اللغوي العربي في اللسانيات الحاسوبية، إذ تؤسس لمشروع متكامل نظريًا وتطبيقيًا لحوسبة اللغة العربية صوتيًا، وصرفيًا، ونحويًا، ومعجميًا مقارنة بحوسبة اللغة الإنجليزية، فله مشروع متكامل لحوسبة اللغة العربية بمستوياتها المختلفة. (**د. نبيل علي في كتابه "اللغة العربية والحاسوب"**)

- له مشروع متكامل لحوسبة اللغة العربية بمستوياتها المختلفة. (**د. نبيل علي في كتابه "اللغة العربية والحاسوب"**)

- **يتكون المقياس العربي لمجموعة المحارف ٤٥، بواقع:**

أ- ٢٨ صوتا صامتا هي: (ء ب ت ث ج ح خ د ذ ر ز س ش ص ض ط ظ ع غ ف ق ك ل م ن هـ و ي).

ب- الحركات وعددها ٦ (الفتحة والضمة والكسرة إضافة إلى المد بالألف وبالواو وبالياء).

جـ- محارف الضبط وعددها ٤ (التنوين والشدة والسكون وهمزة الوصل)

د- المحارف الخاصة وعددها ٧ (الهمزة بأشكالها المختلفة : إ أ ؤ ئ والألف المقصورة: ى والمد : آ وأخيرا التاء المربوطة: ة).

- المقياس العربي لمجموع المحارف **(٤٥)**
- اختلاف رمز الهمزة (أ إ ؤ ئ) يمثل **(محارف خاصة).**
- الهمزة بأشكالها وعلامات الترقيم رموز أكرونولوجية Acronology.
- الكلمة الإنجليزية المكتوبة تحت كل حرف هو مثال اللغة الصورية.
- الحروف منفردة س أو ص أو غيرها مثال اللغة الطبيعية.

أسئلة مجاب عليها

س	أ	ب	ج	د
تُبنى المدونة التعليمية (Learner corpus من المناهج والمقررات التعليمية.	صح	خطأ		
تقوم عملية التنقيب في النصوص وتلخيصها على الاكتفاء بالمعلومات الأساسية، وحذف الإضافات والأمثلة والشروح.	صح	خطأ		
تُقارب فكرة (البناء الصرفي الثنائي) للخليل بن أحمد كما في (شدّ) و (زلزل) فكرة الـ (tokens) و الـ(types) في برمجيات المدونات اللغوية.	صح	خطأ		

- يرجع الفضل في تصميم برنامج صخر للتدقيق الإملائي إلى (**شركة صخر Sakhr**).

- من الوظائف اللغوية الحاسوبية التي يقوم بها برنامج صخر للتدقيق الإملائي:

- اكتشاف الأخطاء الإملائية.
- الأخطاء النحوية.
- الأخطاء الشائعة في الكتابة العربية.

- يرجع الفضل في تصميم برنامج التشكيل الآلي إلى (**شركة صخر Sakhr**).

- من الوظائف اللغوية الحاسوبية التي يقوم بها برنامج صخر للتشكيل الآلي (**وضع علامات التشكيل في النص العربي المكتوب**)

- نسبة دقة برنامج صخر للتشكيل الآلي تبلغ ٩٨%. - يرجع الفضل في تصميم برنامج المحلل الصرفي إلى (**شركة صخر Sakhr**).

- من الوظائف اللغوية الحاسوبية التي يقوم بها برنامج المحلل الصرفي شركة صخر Sakhr (**تجذير الكلمة العربية بإزالة الزوائد والمورفيمات السابقة واللاحقة لتحديد جذر الكلمة**)

- من الوظائف اللغوية الحاسوبية التي يقوم بها برنامج المحلل الصرفي لشركة صخر Sakhr:

- تجذير الكلمة العربية.
- إزالة الزوائد والمورفيمات السابقة واللاحقة لتحديد جذر الكلمة.

- يطلق على البرنامج الذي يقوم بتجذير الكلمة العربية بإزالة الزوائد والمورفيمات السابقة واللاحقة لتحديد جذر الكلمة (**برنامج المُحَلِّل الصرفي**)

- يرجع الفضل في تصميم برنامج استخراج الكلمات المفتاحية إلى (**شركة صخر Sakhr**).

- من الوظائف اللغوية الحاسوبية التي يقوم بها برنامج صخر لاستخراج الكلمات المفتاحية:

- تحليل النصوص العربية.
- تحديد العبارات والأفكار الأساسية.

- يطلق على البرنامج الذي يُحَدِّد العبارات والأفكار الأساسية فيها بشكل آلي في صورة كلمات مفتاحية تساعد في تصنيف النصوص والوثائق في شجرة موضوعات تسهل عملية التصفح والوصول للمعلومات ذات الصلة بالموضوع الواحد. (**برنامج صخر لاستخراج الكلمات المفتاحية**).

- من الوظائف اللغوية الحاسوبية التي يقوم بها مُحَرِّك صخر للتلخيص العربي:

- تحديد الجمل والأفكار المهمة داخل النص.
- عرض الجمل في شكل ملخص نصي قصير.

- يطلق على ما يُحَدِّد الجمل والأفكار المهمة داخل النص وعرضها في شكل ملخص نصي قصير. (**محرك صخر للتلخيص العربي آليًا**)

- يرجع الفضل في تصميم برنامج إيصار إلى (**شركة صخر Sakhr**).

- انتهى المؤتمر إلى اعتماد أول معيار موحد لتشفير الرموز الكتابية العربية، وهو (CODAR1)

- يؤرخ لاعتماد أول مواصفة عربية "أول معيار موحد لتشفير الرموز الكتابة العربية "CODAR1" بـ(**الستينيات**).

المرحلة الثانية: المؤتمر الثاني في تونس:

- قدم العراق معيارًا موازيًا للمعيار السابق في مؤتمر تونس.
- أجري بعض التعديلات على CODAR1.
- اعتمد مشروع مؤقت تحت اسم CODARZ.
- عرض ما توصل إليه على مؤتمر الوزراء العرب المسئولين عن العلوم والتكنولوجيا في المغرب، والمؤتمر العام لليونسكو في ديسمبر بكينيا، ليستقر الرأي على تبني المشروع في **السبعينيات**.

المرحلة الثالثة: المؤتمر الثالث في إيطاليا (روما):

- يؤرخ لعقد مؤتمر روما لوضع مواصفة مشتركة لتشفير الرموز العربية حاسوبيًا بـ(**السبعينيات**)

- انتهت المداولات في مؤتمر روما إلى الاتفاق على المواصفة العربية "معيار جديد موحد لترميز الكتابة العربية" هو (CODAR – U)

- أول مواصفة عربية سجلت لدى الـ ISO كانت (CODAR1/ CODARZ/ CODAR–U/ CODARZ).

- (ISO) ترمز إلى (**المنظمة الدولية للتقييس "منظمة المعايير الدولية"**)

المرحلة الرابعة: مؤتمر الرباط بالمغرب ومناقشة اقتراح الرابطة الأوربية لمصنعي الحواسيب ECMA :

- الجهة التي تملك سلطة تسجيل الرموز الحاسوبية لدى منظمة الأيزو هي (**الرابطة الأوربية لمصنعي الحواسيب**).
- يؤرخ لاقتراح الرابطة الأوربية لمصنعي الحواسيب (ECMA) معيارًا متكاملًا لرقمنة الحرف العربي بـ (**الثمانينيات**).
- نوقش اقتراح الرابطة الأوربية لمصنعي الحواسيب "ECMA" لرقمنة الحرف العربي في (**الرباط بالمغرب**).
- شارك في مناقشة اقتراح الرابطة الأوربية لمصنعي الحواسيب "ECMA" لرقمنة الحرف العربي، والذي نوقش في مؤتمر طارئ بالرباط (**المنظمة العربية للمواصفات والتقييس "الألكسو"**).
- انتهت المداولات في مؤتمر الرباط بالمغرب في الثمانينيات إلى الاتفاق على معيار لترميز الكتابة العربية هو CODAR UF DF

- وقد اعتمدت الجهة الأوربية لمصنفي الحواسيب هذا المعيار باسم (ECMA44)
- واعتمدته منظمة الأيزو ١٩٨٨ باسم ISO/IEC8859-6 الخاص باللغة والكتابة العربية.
- **وتأتي مواصفة الكتابة العربية ISO/IEC8859-6 في المعيار السادس،**