



كلية الآداب
قسم اللغة العربية وآدابها



جامعة عين شمس

المدونات الحاسوبية لأرصدة اللغة العربية

دراسة وصفية تحليلية

"أطروحة مقدمة للحصول على درجة الماجستير في علم اللغة"

إعداد الباحث

محمد رمضان علي علي الغليض

إشراف

الأستاذ الدكتور
محسن عبد الرازق رشوان
أستاذ هندسة الإلكترونيات
والاتصالات
كلية الهندسة- جامعة القاهرة

الأستاذ الدكتور
علي محمد هنداوى
الأستاذ المتفرغ بقسم اللغة العربية
وآدابها
كلية الآداب- جامعة عين شمس

اهداء

أبي العظيم وأمي الغالية إخوانى وأخواتي، زوجتي
الحبيبة والأميرات، وإلى أرواح الشهداء،،،
إليكم جميعاً أهدي هذه الثمرة - نرجوها طيبة -
تعبرنا عن عرفان بالجميل، وشفيعي في ذلك نية
حسنة أرادت الخير والإحسان.

شُكْرٌ وَتَقْدِيرٌ

قال رسول الله ﷺ

"مَنْ لَا يَشْكُرُ النَّاسَ لَا يَشْكُرُ اللَّهَ"

صدق رسول الله ﷺ

أتقدم بخالص الشُّكْر والتَّقْدِير إلى أستاذِي الكريمين:

الأستاذ الدكتور / علي محمد هنداوي

الذى غمرني بكرمه وحسن توجيهه بقبوله الإشراف على هذه الرسالة، فله خالص الشُّكْر وعظيم التَّقدِير.

والأستاذ الدكتور / محسن رشوان

الذى أشرف على الجانب الحاسوبى من الرسالة، وأفاض على صاحبها بالمساندة العملية والعلمية؛ فله جزيل الشُّكْر واللُّود.

كما أتقدم بالشُّكْر الجزيل إلى أستاذِي الكريمين:

الأستاذ الدكتور / حسن محمود نصر السيد

و

الأستاذ الدكتور / خالد فهمي إبراهيم

على تفضيلهما بقبول مناقشة هذه الرسالة، وتحمل مشقة قراءتها وتخصص ما فيها لمناقشتها وتصويب ما فيها.

وأنى لأتقدُّم بخالص الشُّكْر للأستاذ الدكتور / فرانتشيك أوندراش

والسيدة الدكتورة / حسنة عبد السميع

الذين تعهدا صاحب الرسالة في مهده الأكاديمي وتبنيا فكرة الدراسة في مهدها؛ فلهما أجزل الشُّكْر وأوفاه.

كما أتقدُّم بالشُّكْر الجزيل للأستاذ الدكتور / بيتر زيمانك على ما بذل من وقته الثمين في توجيه صاحب الرسالة.

كما أتوجه بخالص الشُّكْر والامتنان والتَّقدِير

للأستاذ الدكتور / سامح الانصارى

والأستاذ الدكتور / سلطان بن ناصر المجبول

والأستاذ الدكتور / عبدالمحسن الثبيتي

على مساعدتهم الإيجابية والفعالة في تعزيز مدونة الصومات العربية

فهرس الموضوعات

الصفحة	الموضوعات
١	المقدمة.
٣	طبيعة الدراسة.
٣	أهمية الدراسة.
٤	مشكلة الدراسة.
٤	أهداف الدراسة.
٤	منهج الدراسة.
٤	الدراسات السابقة.
٦	حدود الدراسة.
٧	هيكلة الرسالة.
٩	الفصل الأول: أنواع المدونات اللغوية الحاسوبية.
٩١	الفصل الثاني: دراسة نظرية لبعض نماذج توظيف التقنيات الحديثة في إنتاج بعض التطبيقات اللغوية المنبثقة من المدونات الحاسوبية.
١٠٩	الفصل الثالث: بناء مدونة للصوات الصعبة في اللغة العربية.
١٥٦	الخاتمة.
١٥٧	١. النتائج.
١٥٩	- جوانب القصور.
١٦٠	٢. التوصيات.
١٦١	قائمة الملاحق.
١٧٦	قائمة المراجع.
١٧٧	أولاً- المراجع العربية.
١٨٢	ثانياً- المراجع الأجنبية.
١٨٤	ثالثاً- الواقع الإلكترونية والبرمجيات.
١٨٨	ABSTRACT

فهرس الجداول

الصفحة	عنوان الجدول	رقم الجدول
٦٩	مقدمة كلارا CLARA	١-١
٩٧	اختبارات قياس الكفاءة اللغوية في مجال اللغة العربية للناطقين بغيرها.	١-٢
١٠٤	المقاييس المترتبة العام لاختبار اللسان العربي (اسع).	٢-٢
١٠٦	التوصيفات التي وضعها ووضعوا الإطار المرجعي الأوروبي المشترك للمهارات.	٣-٢
١٠٧	ارتباط درجات امتحان (اسع) بالإطار الأوروبي المرجعي الموحد للغات من واقع درجات اختبار (IELTS).	٤-٢
١١٠	بيان الأصوات الصامتة.	١-٣
١١٢	بيان أكثر الألسنة من حيث عدد الناطقين بها غير اللغة العربية.	٢-٣
١١٣	البلاد التي تتحدث تلك اللغة غالباً.	٣-٣
١١٥	الصومامت العربية ولغة الماندرين الصينية.	٤-٣
١١٦	صومامت لغة الماندرين الصينية بالمقارنة بالصومامت العربية.	٥-٣
١١٧	الصومامت العربية ولللغة الأسبانية.	٦-٣
١١٧	صومامت اللغة الأسبانية بالمقارنة بالصومامت العربية.	٧-٣
١١٨	الصومامت العربية ولللغة الإنجليزية.	٨-٣
١١٩	صومامت اللغة الإنجليزية بالمقارنة بالصومامت العربية.	٩-٣
١٢٢	الفارق الصوتي بين اللغتين العربية والتركية.	١٠-٤
١٢٣	صومامت الألقيائية الصوتية العربية الدولية.	١١-٣
١٢٤	صومامت الألقيائية الدولية.	١٢-٣
١٢٧	مقارنة بين رموز IPA للغات المختلفة ولللغة الإنجليزية.	١٣-٣
١٢٨	الأصوات الصامتة والحركات كما ذكرها د. كمال بشر في كتاب "علم الأصوات".	١٤-٣
١٤٣	صفات الصومامت العربية (١).	١٥-٣
١٤٤	صفات الصومامت العربية (٢).	

فهرس الأشكال

الصفحة	عنوان الشكل	رقم الشكل
١٧	مثالاً لخطأ شائع من "معجم لونجمان للاختبارات".	١-١
١٧	تحليل الأخطاء اللغوية لدارسي اللغة العربية - طلبة كوربين - .	٢-١
٢٠	أربعة مصادر أساسية للمدونة اللغوية العالمية العالمية.	٣-١
٢٠	فئات المدونة اللغوية العالمية الإحدى عشرة.	٤-١
٢١	مصادر المدونة من داخل وخارج الوطن العربي.	٥-١
٢٢	صفحة تسجيل الدخول للمدونة اللغوية العالمية.	٦-١
٢٤	نموذج من نتائج استخدام محركات بحث المدونة.	٧-١
٣١	واجهة موقع مدونة فيينا للغة العربية المنطقية.	٨-١
٣١	واجهة صفحة بحث موقع مدونة فيينا للغة العربية المنطقية.	٩-١
٣٢	نموذج من نتائج استخدام محركات بحث المدونة.	١٠-١
٣٦	نموذج من شروحاً مبتكرة لبناء الجملة القرآنية باستخدام الرسوم البيانية.	١١-١
٣٧	ترجمة رموز الوظائف النحوية في مخطط وصف المدونة العربية للغة القرآن الكريم.	١٢-١
٣٨	صفحة التسجيل في موقع المدونة العربية للغة القرآن الكريم.	١٣-١
٣٩	واجهة صفحة البحث في موقع المدونة العربية للغة القرآن الكريم.	١٤-١
٤١	نموذج من شاشة لشرح الكلمة على مستوى قواعد اللغة العربية نحواً وصرفًا، وكذلك الترجمة الإنجليزية للآيات من موقع المدونة العربية للغة القرآن الكريم.	١٥-١
٤٥	صفحة تسجيل الدخول للمدونة اللغوية لمتعلمي اللغة العربية.	١٦-١
٤٧	نموذج من نتائج استخدام محركات بحث المدونة اللغوية لمتعلممي اللغة العربية.	١٧-١
٥١	صفحة تسجيل موقع الكوربوس العربي.	١٨-١
٥٣	نموذج من نتائج استخدام محركات بحث موقع الكوربوس العربي.	١٩-١
٦٤	واجهة صفحة البحث في موقع المدونة اللغوية العربية لمدينة الملك عبدالعزيز.	٢٠-١
٦٥	نموذج من نتائج استخدام محركات بحث المدونة.	٢١-١
٧٢	صفحة تسجيل موقع المدونات المتوازية لوثائق الأمم المتحدة.	٢٢-١
٧٥	نموذج من نتائج استخدام محركات بحث مدونة نظام الأمم المتحدة للمعلومات البليوغرافية.	٢٣-١
٨١	نموذج من نتائج استخدام محركات بحث المدونة اللغوية المتوازية عربي - إنجليزي.	٢٤-١

٨٦	نموذج من نتائج استخدام محركات بحث المدونات النصية العربية للجامعة الإسلامية العالمية بماليزيا (المُنْقَبُ العربي).	٢٥-١
١٠٧	ارتباط درجات امتحان (IELTS) بالإطار المرجعي الموحد للغات.	١-٢
١٢٤	صوائب الألفائية الولية.	١-٣
١٢٥	حركة اللسان مع الصوائب الأساسية.	٢-٣
١٢٨	الحركات العربية (الصوائب) والحركات المعيارية.	٣-٣
١٢٩	بيان الصوامت العربية من حيث طريقة النطق.	٤-٣
١٤٧	أداة خافض اللسان.	٥-٣

مسرد المصطلحات

Arabic Lingua Test (ALT)	اختبار اللسان العربي "اسع"
International Phonetic Alphabet (IPA)	الألفبائية الصوتية الدولية
Arabic Lexicography	الصناعة المعجمية العربية
Concordancer	الكاف السياقي
Unique Words	الكلمات الفريدة/غير المكررة
Computational Linguistics	اللسانيات الحاسوبية
Spoken language Corpus	المدونات اللغوية المنطقية
Corpus	المدونة اللغوية/متن/مكتن
part-of-speech tagging	توصيم نحو
Language Computing	حوسبة اللغة
Phonetics	علم الأصوات
Phoneme	فونيم "وحدة صوتية"
Lists of common words	قوائم الكلمات الشائعة
Corpus Linguistics	لسانيات المدونة
Search Engine	محرك البحث
Parallel corpus	مدونة متوازية
Frequency Information	معلومات التكرار
Linguistic Data Consortium	مؤسسة البيانات اللغوية
Buckwalter transliteration system	نظام باك والتر للنقل الكتابي
DT transliteration system	نظام دي تي للنقل الكتابي
DARPA	وكالة مشاريع البحوث المتقدمة الدفاعية التابعة لوزارة الدفاع الأمريكية

المقدمة

المقدمة

لا يخفى على أحد الإمكانات المتعددة التي أتاحها الحاسُبُ الآلِيُّ والتي ولدت الرغبة في تسخيرِ علومه لخدمة نشاطاتٍ عَدَة، من أهمها الدرس اللغوي؛ إذ يعُد دخول الحاسُبُ الآلِيُّ في مجال اللغة بما عُرِفَ بـ"حوسبة اللغة" خطوة كبيرةً في مجالات اللغات الإنسانية بصفةٍ عامَة، ولاسيما اللغة العربية و مجالاتها، عن طريق المعالجة الحاسوبية للأرصدة اللغوية المكتوبة والمنطقية؛ لتصبح مدونة حاسوبية يُمكِّن التعامل معها، ومن ثم إنتاج العديد من التطبيقات التي تخدم اللغة وتقتربُها إلى الناطقين بها وبغيرها.

وقد اهتمت جهات عدّة بجمع هذه الأرصدة اللغوية، وحشدت لها الجهات والفرق البحثية للعمل عليها والإفادة منها في الخروج بتطبيقاتٍ لغويةٍ مثلك طفرةً كبيرةً في صناعةِ القواميس، وتعليم اللغاتِ للناطقين بغيرها، والبرامج التي تخدم البحث الأكاديمي وتسير مادته على الباحثين في التخصصات المتعددة مُذللةً لهمحدود الجغرافية وما تمثله من عائق وصعوبات.

ومن تلك الجهاتِ داخل مصر والعالم العربي، على سبيل المثال:

- ١- الجامعة الأمريكية بالقاهرة ومساهمتها في مبادرة "أرشيف المجموعات العربية على شبكة المعلومات العالمية"^(١) عام ٢٠١٥ م.
 - ٢- "المدونات النصية المتوازية من الإنجليزية للعربية" التي قامت برصدها جامعة الكويت عام ٢٠٠٣ م.^(٢)
 - ٣- مشروع "الرصيد اللغوي الوظيفي" الذي تقرر إنجازه عن طريق بلدان ثلاثة هي تونس والجزائر والمغرب عام ١٩٦٧ م. وانتهى العمل في ١٩٧٢ م، وأقر وزراء تعليم هذه الدول إدراجه في التعليم ابتداء من عام ١٩٧٥ م.^(٣)
 - ٤- مشروع بناء "المعجم الكبير"، الذي تبناه مجمع اللغة العربية بالقاهرة، صدرت منه حتى الآن - سبعة أجزاء بين عامي ١٩٥٦ و٢٠١٢ م، لم يكتمل بعد.^(٤)

كما عمل على تكوين مثل هذه الأرصدة الإلكترونية للغة العربية من خارج العالم العربي فرق عمل بحثية وأقسام لغة العربية، على سبيل المثال:

(1) [الجامعة تتضمن إلى مبادرة أرشيف المجموعات العربية على الإنترنت](http://aucegypt.edu/ar/news/stories-العربية-المجموعات-أرشيف-إلى-مبادرة)

(2) E-A Parallel Corpus (2003), University of Kuwait, Publications from Kuwait National Council.

(٣) "الرصيد اللغوي للطفل العربي وأهمية الاهتمام بمدى استجابته لحاجاته في العصر الحاضر، الممارسات اللغوية" عبد الرحمن الحاج صالح، ٢٠١٠م، (١). ٩-٢١.