

تطبيق لسانيات المدونات على اللغة العربية

نوظف في وصف اللغة العربية لسانيات المدونات (corpus linguistics) التي تنزل مليارات الألفاظ العربية في سياقات استعمالها الحقيقية، وتضبط تواترها، وتمكن من استجلاء قيود انتقائها، كما تضبط أمثالها المطردة وعباراتها المتجمدة.

نطبق في عملنا لسانيات المدونة القاعدية (corpus based) التي يشرف على أبحاثها قوفري ليش (Goffrey Leech) من جامعة "لانكستر". وقد استفدنا منها في تدقيق المترادفات معجما وتركيبا، وفي ضبط استعمالات الأفعال التوزيعية (distributional verbs) وأفعال العماد (support verbs)، وكذلك في التعرف على المجموع المتواترة، وتلك الآخذة في الاندثار لاستبعادها. واعتمدنا على "اللسانيات الموجهة بالمدونات" (corpus directed linguistics) التي تبناها جون سنكلر (John Sinclair) من جامعة "برمنغهام" في استقرار الظواهر النحوية، من مثل الوقوف على سمات أفعال الوجوب في اللغة العربية. ونظرنا كذلك في مدونات اللغات المنطوقة والمشفوعة بالإيماءات في مختلف الممارسات اللغوية مما جعل اللسانيات تراجع مفاهيمها ومناهج أبحاثها وتولي المدونات الحاسوبية أهمية أوفى.

نتطلع من دراستنا إلى سنّ منهج لسانيّ حاسوبيّ يُوظف في وصف الاستعمال المتحقق تداول للغة العربية.

This book uses Corpus Linguistics to describe Arabic as exemplified through the use of billions of Arabic expressions in real situational contexts. This is based on their frequency of occurrence and selection restriction rules, ranging from the most frequently occurring expressions to the frozen ones. The framework relied upon is Corpus databank under the supervision of Geoffrey Leech at Lancaster University. This bank has been very useful in detailing the nuances of synonyms at the lexical and syntactic levels, determining the verbs' distributional patterns, with special focus on support verbs in addition to the identification of frequent sets as opposed to those sets on the way to getting obsolete. Reliance has also been made on Corpus-directed Linguistics, as advocated by Sinclair at Birmingham University to assign grammatical features such as verbs of obligation in Arabic. Linguistic theory is derived from Speech theory. This is why there is special focus on spoken language in its paratextual dimensions in various linguistic practices, hence the vital importance of Corpus-based Linguistics.

This study looks forward to establishing a computational linguistic approach to use in describing the real usage of Arabic.