



Étude lexicale de textes arabes pour l'évaluation automatique de la complexité textuelle.

Youcef Ihab MORSI , Yves BORDET, Iana ATANASSOVA.

7ème Colloque International en Traductologie et TAL TRADETAL 2016, Oran, Algérie, mai 2016.



Résumé

Nous proposons une méthode pour évaluer la complexité du vocabulaire des textes en langue arabe en utilisant les résultats d'une analyse lexicale d'un corpus littéraire composé de 40 textes. Cette approche a été proposée par Yves Bordet dans le projet Doxilog appliqué à plusieurs langues, dont le français, l'anglais, l'espagnol, le russe et le chinois. La présence du vocabulaire arabe pérenne dans un texte en arabe est un élément clé permettant de déterminer la complexité du texte et notamment l'âge et le niveau d'études pour les apprenants natifs et le niveau de maîtrise de la langue pour les apprenants de langues étrangères.

Mots clés

Langue arabe, langue pérenne, vocabulaire pérenne, complexité textuelle, lisibilité, lexique, analyse lexicale, Doxilog.