

Adaptation de la plateforme corporelle ScienQuest pour l'aide à la rédaction en langue seconde

Achille Falaise

Université Grenoble Alpes, LIG-GETALP, F-38040 Grenoble

achille.falaise@imag.fr

RÉSUMÉ

La plateforme ScienQuest fut initialement créée pour l'étude linguistique du positionnement et du raisonnement dans le corpus Scientext. Cette démonstration présente les modifications apportées à cette plateforme, pour en faire une base phraséologique adaptée à l'aide à la rédaction en langue seconde. Cette adaptation est utilisée dans le cadre de deux expérimentations en cours : l'aide à la rédaction en anglais pour les scientifiques, et l'aide à la rédaction académique en français pour les apprenants.

ABSTRACT

Adaptation of the corpus platform ScienQuest for assistance to writing in a second language

The ScienQuest platform was initially created for the linguistic study of positioning and reasoning in the Scientext corpus. This demonstration introduces modifications to this platform, transforming it into a phraseological database adapted for assistance to writing in a second language. This adaptation is used as part of two ongoing experiments: an assistance to writing in English for scientists, and an assistance to academic writing in French for learners.

MOTS-CLÉS : Aide à la rédaction, langue seconde, ScienQuest, Scientext.

KEYWORDS : Writing assistance, second language, ScienQuest, Scientext.

1 Introduction

De plus en plus de personnes sont amenées à écrire dans une langue seconde, c'est à dire une langue qu'ils maîtrisent partiellement, mais qui n'est pas leur langue maternelle. Ces personnes, qui disposent déjà d'un bagage lexical, peuvent néanmoins avoir des difficultés à s'en servir, du fait d'une connaissance imparfaite de l'usage de la langue. Un corpus (Boulton et al., 2006), ou une base phraséologique, alimentée d'exemples choisis pour éclairer certaines notions clés (introduire une idée, se positionner, etc.), peut les aider.

2 La plateforme corporelle ScienQuest

La plateforme ScienQuest (Falaise *et al.*, 2011) fut initialement créée pour l'étude linguistique du positionnement et du raisonnement dans le corpus de textes scientifiques Scientext (Tutin *et al.*, 2009). Cet outil permet de rechercher des concordances dans ce corpus, en fonction de critères linguistiques. Il est destiné à des linguistes : les critères sont formulés en termes linguistiques, les concordances peuvent comporter des erreurs (notamment dues à des erreurs d'annotation dans le corpus) et des résultats peu

pertinents (les résultats peuvent être filtrés mais ce filtrage ne peut pas être sauvegardé) ; enfin, certaines fonctionnalités sont superflues pour des utilisateurs non-linguistes (par exemple, les statistiques lexicométriques).

3 Adaptation de la plateforme

L'adaptation de ScienQuest a consisté en l'ajout de trois fonctionnalités à la plateforme :

1. la possibilité d'exporter et de sauvegarder sur le serveur des concordances filtrées ;
2. la possibilité d'importer ou de restaurer ces concordances filtrées ;
3. la création d'un nouveau mode de visualisation simplifié, limité à la visualisation des concordances.

4 Conclusion

L'adaptation présentée dans cette démonstration sert de base à deux expérimentations en cours, concernant l'aide à la rédaction en anglais pour les scientifiques (Jacques *et al.*, 2013) et l'aide à la rédaction académique en français pour les apprenants (Tutin et Falaise, 2013).

Références

BOULTON, A., WILHELM, S. (2006). Habeant Corpus-they should have the body. Tools learners have the right to use. In *ASp*, 49-50, pages 155-170.

FALAISE, A., TUTIN, A., KRAIF, O. (2011). Exploitation d'un corpus arboré pour non spécialistes par des requêtes guidées et des requêtes sémantiques. In *Actes de TALN 2011*, Montpellier, pages 187-215.

JACQUES, M.-P., HARTWELL L., FALAISE A. (2013). TAL et linguistique de corpus pour aider la rédaction scientifique en anglais. In *Actes de TALN 2013*, Les Sables d'Olonne.

TUTIN A., FALAISE A. (2013). Multiword expressions in scientific discourse: a corpus-driven database. In *eLex 2013*, Tallinn, Estonie.

TUTIN A., GOSSMANN F., FALAISE A., KRAIF O. (2009). Autour du projet Scientext : étude des marques linguistiques du positionnement de l'auteur dans les écrits scientifiques. In *Journées Linguistique de Corpus*, Lorient.