

Exemples d'application de l'AFC aux structures du vocabulaire :

2010 Viprey J.-M. & Schepens P. « Dérivation lexicale et dérive du discours : « mutualisation, mutualiser » in JADT 2010. Roma : LED.

http://lexicometrica.univ-paris3.fr/jadt/jadt2010/allegati/JADT-2010-0489-0498_119-Viprey.pdf

2006 Viprey J.-M. « ...un de ces syntagmes qui... » in *Corpus* n°5, *Corpus et stylistique*, 2006, CNRS-UNSA. (28 p.)

<http://corpus.revues.org/index713.html>

2005 Viprey J.-M. « Structure non séquentielle du texte » in *Langages* n° 161, *Unité(s) du texte* dir. Dominique Legallois. – Paris : Larousse. (pp. 65-82) [joint]

2005 Viprey J.-M. « Méthodes pour la lecture des corpus » in *Sémantique et corpus* dir. Anne Condamines. – Hermès (pp.245-276) [joint]

L'article où a été présentée la projection géodésique :

2006 Viprey J.-M. « Ergonomiser la visualisation AFC dans un environnement d'exploration textuelle : une projection 'géodésique' » in *JADT'06, 8èmes Journées internationales d'Analyse statistique des Données Textuelles*. – Besançon, Presses Universitaires de Franche-Comté. (12 p.)

<http://lexicometrica.univ-paris3.fr/jadt/jadt2006/PDF/II-087.pdf>

Le site d'Astadiag :

http://laseldi.univ-fcomte.fr/document/viprey/page_JMV.htm

► **DiaTag** est un ensemble modulaire destiné à répondre aux besoins immédiats et approfondis d'élaboration et de préparation de bases textuelles et d'extraction de corpus textuels en français des 19-20èmes siècles. Révisé pour être compatible avec les prescriptions de la TEI et les contraintes de l'interopérabilité, il intègre des modules de correction d'OCR, d'extraction de sous-corpus, d'annotation linguistique poussée avec levée d'ambiguïté semi-automatique. Il recourt à des dictionnaires de corpus, à des scripts de levée d'ambiguïté, et prête une attention particulière aux convivialités de décision et à leur ergonomisation. **DiaTag** fait l'objet d'études spécifiques, co-financées par la Région Franche-Comté, en vue de tester la meilleure faisabilité de l'acquisition de grandes bases numériques plein texte à partir des collections de presse du 19^{ème} siècle ou antérieures sans saisie clavier intégrale.

► **Astartex** est un environnement entièrement original, modulaire (de l'utilisation en 1^{er} cycle à l'hypertexte avancé), à forte consistance statistique multidimensionnelle (avec recours à une heuristique à base de cartographie statistique), et fondé sur les outils d'un *dictionnaire dynamique multimodal* (contextes « bruts », histogrammes, cartes, sensibilité au clic vers diverses fonctions de saisie et de relance...).

Ces deux développements, intégrés dans le logiciel AstaDiag, font l'objet d'un enrichissement coopératif de ressources et sont associés à la construction de la plate-forme de fonctionnalités *Textométrie* en cours de développement dans le cadre d'un projet ANR. Ils font l'objet de perspectives d'intégration plus poussée avec les logiciels SATO (Daoust-UQAM : thèse Daoust, Ecole d'Été MCD, visés ci-dessous) et Hyperbase (Brunet-UNSA-CNRS : délégation Viprey auprès de l'UMR BCL-CNRS).