

Constitution de ressources pédagogiques numériques : le lexique des affects

Magdalena Augustyn, Sabrina Ben Hamou, Gwendoline Bloquet, Vannina Goossens, Mathieu Loiseau, Fanny Rinck

▶ **To cite this version:**

Magdalena Augustyn, Sabrina Ben Hamou, Gwendoline Bloquet, Vannina Goossens, Mathieu Loiseau, et al.. Constitution de ressources pédagogiques numériques : le lexique des affects. Colloque International des étudiants-chercheurs en didactique des langues et linguistique., Jul 2006, Grenoble, France. halshs-00418143

HAL Id: halshs-00418143
https://shs.hal.science/halshs-00418143

Submitted on 14 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CONSTITUTION DE RESSOURCES PÉDAGOGIQUES NUMÉRIQUES : LE LEXIQUE DES AFFECTS

Magdalena AUGUSTYN, Sabrina BEN HAMOU, Gwendoline BLOQUET, Vannina GOOSSENS, Mathieu LOISEAU, Fanny RINCK
magdalena.augustyn@bvra.e.u-grenoble3.fr, *sabrina_ben_hamou@yahoo.fr*, *bloquetgwen@yahoo.fr*, *vannina.goossens@u-grenoble3.fr*, *mathieu.loiseau@u-grenoble3.fr*, *fanny.rinck@u-grenoble3.fr*
LIDILEM, Université Stendhal Grenoble3
BP 25 - 38040 Grenoble cedex 9

Abstract : The work presented in this article is part of a project aiming at the development of French teaching pedagogical resources and their NLP[1]-based exploitation. The article focuses on the "affect lexicon" to exemplify the process of computer modelling of textual resources (the description of the lexicon and its semantic annotation within the corpus) on which the extraction of examples and the generation of exercises are based.

1. Introduction

Le travail présenté s'inscrit dans un projet[2] qui vise la constitution et l'exploitation de ressources linguistiques pour la didactique du français, à destination d'un public de français langue étrangère (niveau intermédiaire à avancé) et d'un public de français langue maternelle (collège et lycée). Les ressources, textuelles, sont traitées à l'aide d'outils de traitement automatique des langues (TAL), et servent l'enseignement et l'apprentissage de phénomènes énonciatifs, liés à l'expression du point de vue et de la subjectivité dans les textes.

Nous nous centrerons ici sur le travail préalable à l'exploitation pédagogique, et relatif au lexique des affects : le ressenti ou l'attitude du narrateur et/ou de ses personnages sont caractérisés lexicalement, comme relevant de la « joie », de la « tristesse », etc.

Le projet réunit des linguistes, des spécialistes du TAL et des didacticiens, et prend comme point de départ l'idée qu'une approche en corpus de ces phénomènes doit favoriser leur enseignement/apprentissage. Il s'agit de recourir à des données langagières attestées, qui permettent aux apprenants de repérer, d'analyser et de manipuler ces phénomènes, leurs co-occurrences et leurs effets en termes d'expression de la subjectivité. Une telle approche en corpus doit ainsi favoriser la réflexion métalinguistique des apprenants, dans une perspective en particulier de réinvestissement des apprentissages lexicaux dans la production de textes. Cet objectif requiert d'une part une modélisation informatique des ressources textuelles : il s'agit de faire en sorte qu'elles soient utilisables en tant que ressources pédagogiques, que ce soit comme réservoir d'exemples pour l'enseignant, qui doit pouvoir formuler des requêtes dans le corpus, ou comme réservoir d'exercices pour l'apprenant. D'autre part, une description linguistique du lexique des affects doit être faite de manière à en faciliter l'enseignement et l'apprentissage. Elle doit permettre d'annoter le phénomène dans les textes du corpus, ainsi enrichi d'informations sémantiques qui sont exploitées automatiquement pour la formulation de requêtes dans le corpus et la génération d'exercices.

[1] Natural Language Processing

[2] Plan pluri-formation piloté par le Lidilem (Université Stendhal Grenoble 3).

Ce sont ces deux aspects, la modélisation informatique des ressources et la description du lexique des affects à des fins d'annotation sémantique que nous présentons ici.

Le corpus utilisé représente 850.000 mots. Il comprend des textes littéraires principalement, en raison du public visé ainsi que des contraintes imposées par la diffusion des textes. L'intérêt qu'il y aurait à disposer d'autres types de textes, notamment des textes journalistiques et des textes en usage sur Internet se heurte à des problèmes de droit. On peut également se demander dans quelle mesure le contenu du web peut servir de corpus, a fortiori à des fins d'exploitation pédagogique.

2. Modélisation informatique des ressources

La mise à disposition pour des enseignants de français de ressources pédagogiques numériques fait intervenir diverses compétences : didactiques (exploitation des ressources constituées), linguistiques (description des phénomènes linguistiques) et informatiques (numérisation et mise à disposition des ressources). Tous les aspects de ce projet sont interdépendants. La description des phénomènes linguistiques nécessite d'avoir des informations d'ordre didactique (comment les enseignants voudront-ils exploiter les informations linguistiques mises à leur disposition, pour quel type d'activité ?) et relevant du TAL[3] (dans quelle mesure l'annotation des phénomènes pourra-t-elle être automatisée ?). De leur côté, les didacticiens ne peuvent préparer l'exploitation pédagogique sans une connaissance préalable approfondie des ressources linguistiques disponibles (quelles informations seront disponibles dans les ressources et comment elles seront présentées ?).

D'un point de vue informatique, le processus de constitution du corpus est lui aussi tributaire des choix de modélisation linguistique des phénomènes et sera détaillé dans la partie suivante à travers l'exemple du lexique des affects. De la même manière, la création du système *via* lequel les utilisateurs exploiteront les ressources doit prendre en compte leurs attentes et dépend des informations contenues dans ces ressources.

Toutes ces contraintes nécessitent un développement incrémental pour chaque notion incluse dans les ressources développées. Ce type de fonctionnement requiert une architecture globale du système la plus générique possible. Nous utilisons donc l'architecture de la plateforme d'apprentissage des langues MIRTO[4], développée à l'université Stendhal. La modularité de la plateforme provient de la dissociation des données (corpus) et des traitements (Antoniadis *et al.*, 2004a) et permet ainsi d'éviter les principaux défauts des logiciels d'apprentissage des langues : la pauvreté des informations de sens associées aux données langagières (Antoniadis *et al.*, 2004b) et le manque de diversité des exemples disponibles.

Dans cette architecture (cf. Figure 1), un script (S_i) est un programme informatique qui crée une activité à partir d'un texte donné. Cette activité peut aussi bien être un exercice lacunaire géré entièrement par l'ordinateur (ex : exercice lacunaire sur les noms d'affect) que la présentation de certaines facettes du texte pour un travail dirigé par l'enseignant en présentiel. On peut, par exemple imaginer un script qui, dans un texte donné, mettrait en évidence toutes les occurrences des noms d'affect de la même manière afin de générer, après un travail d'exploration, une discussion aboutissant à la description des différentes catégories de noms d'affect. Dans la première phase du travail, les apprenants surligneraient d'une manière différente chacun des noms d'affect en fonction des types qu'ils remarquent, avant de confronter, dans une seconde phase, la typologie ainsi créée avec la typologie existante. Pour un tel script, les différentes phases du travail seraient configurables par l'enseignant, qui fournirait ainsi le contexte didactique de l'application du script :

[3] Traitement Automatique des Langues

[4] Multi-apprentissages Interactifs par des Recherches sur des Textes et l'Oral

- choix du (ou des) texte(s) au(x)quel(s) appliquer le script
- choix des phénomènes à mettre en valeur (noms d'affect dans l'exemple ci-dessus)
- avec quelle précision (domaine, sous-domaine ou distinction plus fine encore, *cf.* partie 3)
- interaction avec l'utilisateur (visualisation, regroupement en catégories)
- et enfin la proposition ou non d'une « correction » (comparaison de la catégorisation produite par l'apprenant avec celle effectuée au préalable dans le corpus).

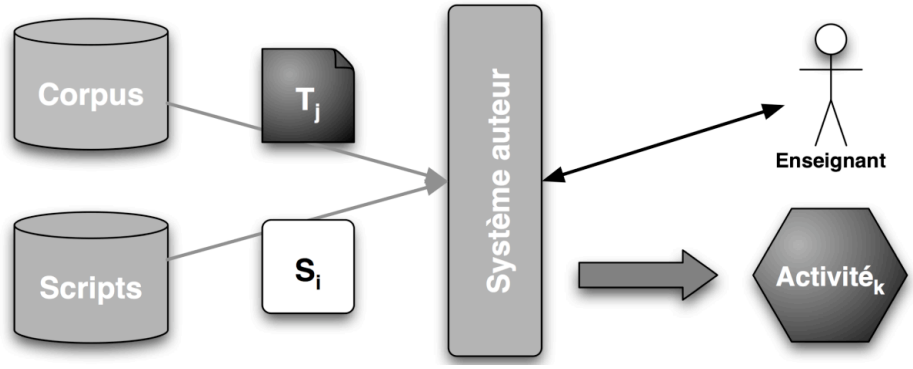


Figure 1 - Génération d'une activité

La création de scripts demandant des compétences pointues en informatique est réservée à un concepteur de scripts, qui proposera, comme nous l'avons vu à travers l'exemple ci-dessus, la possibilité à l'enseignant de définir le contexte didactique de son activité.

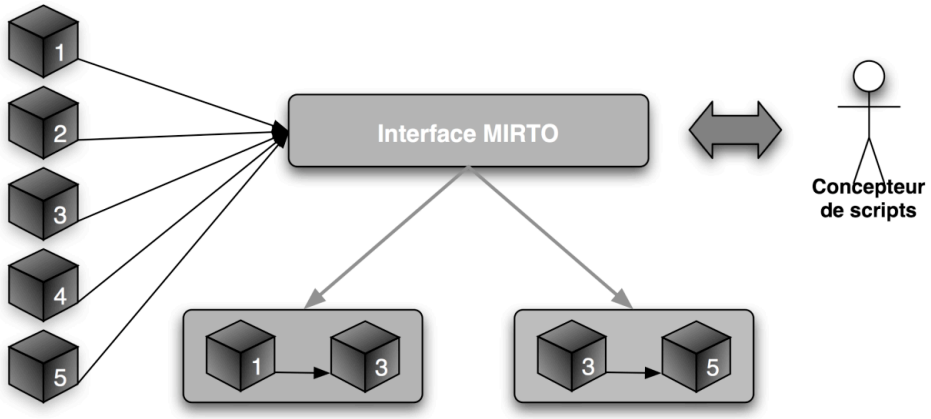


Figure 2 - Conception de scripts dans MIRTO

Via l'interface MIRTO, le concepteur de scripts a accès à un ensemble de traitements de bas niveaux (fonctions numérotées dans la Figure 2). Ces fonctions peuvent relever du TAL (ex : analyse morphologique) ou non (ex : *parser* XML). Le concepteur est alors libre de combiner ces diverses fonctions pour créer des scripts qui seront ensuite accessibles aux enseignants.

Un script de création d'exercices lacunaires portant sur le lexique des affects fera donc intervenir :

- une fonction permettant d'interpréter du XML (*parser* XML),
- une fonction permettant de sélectionner les éléments balisés comme faisant partie du lexique des affects en fonction de paramètres donnés par l'utilisateur enseignant (ex : la catégorie syntaxique des mots, le type, le domaine, *cf.* partie 3),
- une fonction permettant de créer les trous, qui pourra elle même faire appel à
- une fonction de lemmatisation, selon que l'enseignant (utilisateur du script) voudra ou non afficher les lemmes correspondants aux mots escamotés (en fonction du

niveau des apprenants, l'enseignant choisira entre un exercice lacunaire sans indication ou en affichant sous l'énoncé la liste des formes lemmatisées ou encore la liste des mots tels qu'ils apparaissent dans le texte),

- et enfin une fonction de correction automatique des exercices.

A partir d'un tel script, l'enseignant pourra créer autant d'exercices qu'il y aura de textes dans le corpus, palliant ainsi le manque de diversité des exemples disponibles, récurrent dans les plateformes d'enseignement des langues. Les outils TAL présents sous forme de fonction dans les scripts, ainsi que les annotations préalables du texte enrichissent le sens associé au texte dans le cadre de son utilisation pour l'enseignement du français.

En effet, dans le cadre de ce projet, il ne s'agit pas de travailler avec du texte tout venant mais avec le texte issu du corpus, dont nous allons présenter l'annotation à travers l'exemple du lexique des affects.

3. Le lexique des affects

3.1. Description et organisation du lexique

Le lexique des affects a donné lieu à de nombreuses recherches (voir notamment les numéros 105 de *Langue Française,* Balibar-Mrabti, 1995, et 32 de *Lidil*, Grossmann et Tutin, 2005), sur lesquelles nous nous sommes appuyés pour réaliser des travaux de description linguistique (Tutin *et al*., 2005, Goossens, 2005) et des travaux à visée didactique (Grossmann et Boch, 2003, Boch *et al*., 2005, etc.). Tous ces travaux ont mis en évidence d'une part la difficulté qu'il y a à établir quelles unités lexicales font partie du lexique des affects et d'autre part, la complexité sémantique et le caractère polysémique de ces unités lexicales. Pour ce projet, nous avons choisi de retenir une définition assez large de ce qui constitue le lexique des affects afin d'analyser au mieux l'expression de la subjectivité et la polysémie. Nous avons ainsi considéré comme appartenant au lexique des affects les unités lexicales qui peuvent renvoyer à un processus psychologique de type affectif – même s'ils ne renvoient pas obligatoirement à un affect dans tous leurs emplois, comme nous le verrons par la suite. Ceci explique que nous ayons pris en compte des noms comme *bonté* ou *orgueil*, qui s'apparentent en règle générale à des qualités mais qui peuvent également prendre un sens affectif, comme par exemple *bonté* dans le passage suivant, issu de notre corpus :

> *« Et plus leurs cœurs se rapprochaient, plus ils s'appelaient avec cérémonie "Monsieur et Mademoiselle", plus aussi leurs regards se souriaient, se mêlaient ; et il leur semblait qu'une bonté nouvelle entrait en eux »* (Maupassant, *Une vie)*.

Afin de pouvoir annoter le lexique des affects dans notre corpus nous avons élaboré une liste des unités lexicales que nous considérons comme appartenant au lexique des affects. Nous avons pris comme point de départ des listes existantes, notamment celles de Y. Y. Mathieu (2000) pour les verbes. Dans un souci d'exhaustivité, nous avons alors cherché à élargir ces listes. Nous sommes partis de dictionnaires de référence dans lesquels nous avons navigué par synonymes à partir des unités lexicales de la liste initiale. Nous avons ainsi abouti à l'élaboration de quatre listes distinctes, une pour chaque catégorie grammaticale du lexique des affects : noms, verbes, adjectifs et adverbes. Nous allons nous centrer dans la suite de cet article sur le traitement réservé aux noms.

Afin de décrire et d'organiser notre liste de noms d'affect dans une perspective d'exploitation pédagogique, nous avons élaboré une typologie de ces noms. Nous avons fait le choix d'une typologie notionnelle, ce qui nous a semblé le plus adapté pour l'accès sémantique aux éléments lexicaux par les apprenants. Pour cette même raison, nous avons également fait le choix de nous limiter à deux niveaux hiérarchiques : des domaines, qui sont

au nombre de 27, divisés pour certains en sous-domaines. Il s'agit d'étiquettes conventionnelles regroupant une même « famille » sémantique de noms d'affect. Par exemple, le domaine *joie* se divise en 4 sous-domaines :

Domaine 'joie'			
'Joie'			
'bonheur-satisfaction'	**'enthousiasme-entrain'**	**'joie-gaieté'**	**'plaisir-jouissance'**
Bonheur, béatitude, contentement, félicité, ivresse…	***Alacrité, enthousiasme, optimisme, entrain***	***Euphorie, liesse, enchantement, joie, gaieté…***	***Goût, plaisir, jubilation, volupté, jouissance…***

Tableau 1 - Domaine 'joie'

Un certain nombre de noms se trouvent dans plusieurs domaines ou sous-domaines, du fait de la polysémie propre à chaque nom. *Horreur*, par exemple, se trouve à la fois dans le domaine « peur » et dans le domaine « dégoût ». Mais ce n'est pas le seul type de phénomènes polysémiques auquel donnent lieu les noms d'affect et que nous ayons décrit dans ce projet.

3.2. La polysémie des noms d'affect

Comme nous l'avons évoqué précédemment, les noms sélectionnés ne renvoient pas toujours à des affects, ou pas uniquement à des affects. Nous nous trouvons ainsi, en fonction du contexte, en présence de cas de polysémie régulière ou systématique (lorsqu'une dérivation sémantique existe pour tous, ou en tout cas une grande partie, des éléments d'une classe sémantique, en l'occurrence les noms d'affect), ou de polysémie que l'on pourrait appeler irrégulière (lorsque cette dérivation est spécifique à une unité lexicale particulière). Nous avons ainsi distingué quatre cas de figure différents, que nous allons expliquer et exemplifier.

Tout d'abord, et ceci représente évidemment les cas les plus courants, le nom peut exprimer un affect. Par exemple, dans :

> *« Parfois, si les préjugés ou les instincts de son pays revenaient l'assaillir et lui montraient une vengeance facile au détour d'un sentier, il les écartait avec horreur en pensant à ses camarades de régiment, aux salons de Paris, surtout à miss Nevil » (*Mérimée*, Colomba)*

horreur renvoie a un affect du domaine « dégoût ». Dans :

> *« Le berger s'était emparé du cheval, mais il avait laissé tomber selle et bride, et paraissait saisi d'horreur » (*Mérimée*, Colomba)*

il appartient dans ce cas au domaine « peur ». C'est également le cas pour des noms comme *orgueil*, *bonté* ou encore *tendresse*[5], comme dans :

> *« entre eux flottait déjà cette subtile et vague tendresse qui naît si vite entre deux jeunes gens, lorsque le garçon n'est pas laid et que la jeune fille est jolie » (*Maupassant*, Une Vie)*.

Ces noms peuvent également exprimer un affect couplé à une autre dimension sémantique spécifique. Dans de très nombreux cas, par exemple, le nom renvoie à la fois à un affect et à une manifestation de cet affect (ou une attitude exprimant cet affect), comme dans :

> *« Colomba parlait avec enthousiasme de la beauté de miss Nevil » (Mérimée,* Colomba*) ou encore « Il y eut parmi toute l'étude, un murmure d'admiration » (*Daudet*, Le petit chose)*.

[5] Qui s'apparentent en règle générale à des qualités.

Parfois, ces noms n'expriment pas un affect mais une autre acception, dérivée par une relation régulière. Par exemple dans :

> *« sous sa rude moustache, effroi des chiens et des traîneurs de bancs, se cachait un bon sourire attendri » (*Daudet*, Les contes du lundi)*

effroi ne renvoie pas véritablement à un affect mais plus précisément à ce qui cause cet affect, en l'occurrence la moustache. C'est également le cas pour *peine* et *embarras* dans :

> *« je suis une peine et un embarras à ma famille » (*Sand*, La petite Fadette)*

Pour donner un autre exemple de ces acceptions polysémiques régulières, le nom peut exprimer une qualité d'une personne. C'est le cas bien entendu de noms comme *orgueil, bonté* ou *tendresse* que nous avons évoqués précédemment et qui s'apparentent habituellement à des qualités, mais également de noms qui désignent généralement des affects comme *passion* dans :

> *« il est impossible de parler des Corses sans attaquer ou sans justifier leur passion proverbiale » (*Mérimée*, Colomba).*

Enfin, dans certains cas, ces noms n'expriment ni un affect ni une acception polysémique régulière, comme par exemple *blessure* dans :

> *« La blessure à la tempe est grave, n'est-ce pas ? » (*Leroux*, Le mystère de la chambre jaune)*

ou encore *bonheur* dans :

> *« Elle cherchait le trèfle à quatre feuilles, qui se trouve bien rarement et qui porte bonheur à ceux qui peuvent mettre la main dessus » (*Sand*, La petite Fadette).*

Ces différents cas de figure ainsi que les domaines et/ou sous-domaines sont déterminés et annotés en contexte pour chaque nom, de même qu'un certain nombre d'autres traits que nous allons maintenant présenter.

3.3. L'annotation du lexique des affects

L'annotation a été réalisée en deux étapes et sous la forme de balises XML à l'aide d'outils de traitement de corpus : Nooj (Silberztein, 2002) dans un premier temps, pour l'annotation automatique sans attributs puis le filtrage (par exemple suppression d'erreurs de catégorisation comme *je doute, sans doute…*) ; puis Intex (Silberztein, 1992) dans un deuxième temps, pour l'insertion automatique des attributs. A ce stade toutes les valeurs possibles sont attribuées à chaque nom (tous les sous-domaines par exemple) qui sera enfin vérifié et désambiguïsé manuellement en fonction du contexte. Au final, nous aboutissons à un balisage de ce type :

> *« Jean Peyrol devient rouge de* `<LEXIQUE TYPE ="affect+manif_attitude" CAT ="N" DOMAINE ="plaisir" NV_LANGUE ="courant" INTENSITE = "moyen" POLARITE ="positif" FIGE ="" FIGURE_STYLE ="">`*plaisir* `</LEXIQUE>` *et s'empresse autour de son illustre visiteur. » (*Daudet*, Le petit chose).*

Le choix des différents attributs annotés pour le lexique des affects et leur description a été déterminé par la perspective de l'exploitation pédagogique du corpus. Certains sont communs à toutes les catégories syntaxiques comme le domaine, le niveau de langue (courant, littéraire, etc.), l'intensité (haut ou moyen) et la polarité (négatif, positif ou neutre), alors que d'autres sont spécifiques à une catégorie comme le figement pour les noms, le type agentif ou causatif pour les verbes.

En ce qui concerne les noms, nous avons tout d'abord un attribut TYPE, qui correspond au traitement de la polysémie que nous avons abordé précédemment. Quatre configurations sont possibles : affect, affect couplé à une dimension sémantique particulière (affect+manif_attitude, etc.), non affect mais acception polysémique régulière (non affect+cause_objet, non affect+qualité, etc.), ou non affect. L'attribut DOMAINE comprend le domaine du nom ou le sous-domaine lorsqu'il y en a un, tels qu'établis dans notre typologie.

Dans le cas de contraintes syntaxiques propres à certaines constructions (comme par exemple *c'était pitié de*, *faire peine, de peur de,* etc.), nous annotons cet emploi discursif particulier à l'aide de l'attribut FIGEMENT, qui comporte deux valeurs : figé ou non. Cette description sera par la suite affinée par des marqueurs de niveau de figement (collocations, expressions figées, etc. cf. Tutin et Grossmann, 2003).

Enfin, nous annotons également l'emploi stylistique des noms d'affect par l'attribut FIGURE_STYLE avec la désignation du type de procédé employé (métaphore, métonymie, personnification, etc.).

Les verbes, adjectifs et adverbes d'affect ont été annotés selon la même méthode. De même, le corpus contient des annotations pour d'autres phénomènes (discours rapporté, modalisation, etc.) et ce sont ces informations qui sont ensuite interprétées par les scripts de la plateforme MIRTO.

4. Conclusion

Il reste pour finir à dire quelques mots de l'exploitation didactique des annotations faites dans le corpus. La perspective didactique détermine d'abord comme nous l'avons vu le choix des annotations. Des entrées notionnelles ont été privilégiées, pour décrire le lexique des affects à partir de domaines tels que la joie, la tristesse, etc., et il faut gérer un équilibre entre la finesse de la description et la nécessité que les annotations fassent sens pour les enseignants. Un des enjeux de l'annotation a par ailleurs consisté à mettre l'accent sur le problème de la polysémie : le but n'est pas de proposer une interprétation univoque des phénomènes, mais au contraire de défendre l'idée que la polysémie peut et doit faire l'objet d'un traitement en classe de langue.

Outre leur utilisation *via* les scripts (évoquée dans la partie 2), les annotations sont exploitées dans les requêtes que pourra formuler l'enseignant, pour tirer des exemples du corpus, ou même les apprenants, dans le cadre d'une activité d'exploration du corpus encadrée par l'enseignant pour examiner les phénomènes qui constitueront l'objet du cours. Un script pourra s'appliquer ensuite aux textes choisis. D'autre part ces annotations permettent aux apprenants de bénéficier d'exercices générables automatiquement. Dans les deux cas, il s'agit de privilégier une analyse et une manipulation du phénomène en contexte, et de sensibiliser les apprenants aux associations lexicales privilégiées (par ex. *rouge de colère*, etc.). Les ressources doivent permettre un travail sur les effets de sens induits par le contexte, leur diversité et les éventuelles formes d'ambiguïté et de polysémie. Le rôle de l'enseignant apparaît alors comme central dans le déroulement des séances : les outils proposés ne sont pas une fin en soi, mais peuvent par contre être le moteur de nouvelles activités non automatisables, et en premier lieu les échanges au sein de la classe que l'observation de ces phénomènes linguistiques doit susciter.

Une première expérimentation (Cavalla et Tutin, 2006) a été menée dans un cours de FLE de niveau intermédiaire à avancé ; elle valide l'intérêt d'un travail d'analyse en corpus et de mémorisation par des exercices pour l'acquisition des collocations « verbe+nom » et leur réemploi dans une activité de production de textes.

Références

Références bibliographiques

ANTONIADIS, Georges, ECHINARD, Sandra, KRAIF, Olivier, LEBARBE, Thomas, LOISEAU, Mathieu & PONTON, Claude (2004a). NLP-based scripting for CALL activities, *Actes du workshop de Coling '04 sur eLearning for Computational Linguistics and Computational Linguistics for eLearning*, Genève, Suisse, août 2004.

ANTONIADIS, Georges, ECHINARD, Sandra, KRAIF, Olivier, LEBARBE, Thomas, LOISEAU, Mathieu & PONTON, Claude (2004b). CALL: from current problems to NLP solutions, *Actes de EUROCALL*, Vienne, Autriche, 1-4 septembre 2004.

BALIBAR-MRABTI, Antoinette (Ed) (2005). Grammaire des sentiments, *Langue Française*, *105*.

BOCH, Françoise, CAVALLA, Cristelle & GROSSMANN, Francis (2005). Apprendre à intégrer le lexique des sentiments au cours de l'écriture de textes narratifs, *Colloque international Lexique et production verbale : vers une meilleure intégration des apprentissages lexicaux*, 9-10 mars 2005, Lyon, France : Institut National de Recherche Pédagogique.

CAVALLA, Cristelle & TUTIN, Agnès (2006). Lexique de l'émotion en classe de FLE : une expérimentation basée sur corpus, *7eConférence Teaching and Language Corpora : TaLC2006*, 1-4 juillet 2006, Paris 7, France.

CRUSE, D. Alan (1986). *Lexical Semantics*. Cambridge : University Press.

GOOSSENS, Vannina (2005). Les noms de sentiment : esquisse de typologie sémantique fondée sur les collocations verbales, *Lidil*, *32*:103-121.

GROSSMANN, Francis & BOCH, Françoise (2003). Production de textes et apprentissage lexical : l'exemple du lexique de l'émotion et des sentiments, *Repères*, *28*: 117-135.

GROSSMANN, Francis & TUTIN, Agnès (Eds) (2005). Sémantique des noms et adjectifs d'émotion, *Lidil*, *32*.

MATHIEU, Yvette Yannick (2000). *Les verbes de sentiment : de l'analyse linguistique au traitement automatique*. Paris : CNRS.

NUNBERG, Geoffrey & ZAENEN, Annie (1997). La polysémie systématique dans la description lexicale, *Langue Française*, *113*: 12-23.

PUSTEJOVSKY, James (1991). The Generative Lexicon, *Computationnal Linguistics*, *17*, 4: 409-441.

PUSTEJOVSKY, James (1995). *The Generative Lexicon*. Cambridge : MIT Press.

TUTIN, Agnès & GROSSMANN, Francis (Eds) (2003). *Les collocations, analyses et traitement.* Amsterdam : De Werelt.

TUTIN, Agnès, NOVAKOVA, Iva, GROSSMANN, Francis & CAVALLA, Cristelle (2006). Esquisse de typologie des noms d'affect à partir de leurs propriétés combinatoires, *Langue Française*, *150*: 32-49.

Références aux logiciels

SILBERZTEIN, Max (1992). *Intex*, http://intex.univ-fcomte.fr, (nd). Consulté en septembre 2006.

SILBERZTEIN Max (2002). *Nooj*, http://perso.wanadoo.fr/rosavram/pages/noojpag.html, (nd). Consulté en septembre 2006.