



**HAL**  
open science

## Corpus, base de données, cartes mentales pour l'enseignement

Cristelle Cavalla, Mathieu Loiseau, Valérie Lascombe, Joanna Socha

### ► To cite this version:

Cristelle Cavalla, Mathieu Loiseau, Valérie Lascombe, Joanna Socha. Corpus, base de données, cartes mentales pour l'enseignement. Peter Blumenthal, Iva Novakova, Dirk Siepmann. Les émotions dans le discours. Emotions in discourse, Peter Lang, pp.327-341, 2014. hal-01375968

**HAL Id: hal-01375968**

**<https://hal.science/hal-01375968>**

Submitted on 3 Oct 2016

**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.

Version -1 de/of

Cavalla, C., Loiseau, M., Lascombe, V., & Socha, J. (2014). Corpus, base de données, cartes mentales pour l'enseignement. In P. Blumenthal, I. Novakova & D. Siepmann (Eds.), *Les émotions dans le discours. Emotions in discourse* (pp. 327-341). Francfort: Peter Lang.

## **Corpus, base de données, cartes mentales pour l'enseignement**

Cristelle Cavalla\*/Mathieu Loiseau\*\*/Valérie Lascombe\*/Joanna Socha\*

### **Abstract**

As part of an ANR<sup>1</sup> project aimed at the linguistic description of the lexicon of emotions in five European languages, we developed an interface for querying a database (EmoProf) and proposed a sequence for teaching collocations of emotions in an FFL classroom. In this chapter, we will describe the tools and justify our use of “mind maps” to teach FFL.

### **Résumé**

Dans le cadre d'une ANR<sup>1</sup> visant la description linguistique du lexique des émotions dans cinq langues européennes, nous avons développé une interface (EmoProf) pour interroger une base de données et proposer des séquences didactiques pour l'enseignement des collocations d'affect auprès d'un public FLE. Nous décrivons ici les outils et les possibilités d'interrogations ainsi que le choix didactique d'insertion de cartes mentales dans une classe de langue.

### **1. Introduction**

Ce travail s'appuie sur des constats établis dans le cadre de différents projets de recherche (EMOLEX,<sup>3</sup> Scientext, FULS<sup>2</sup>) en description linguistique et développement de séquences didactiques pour un public FLE.<sup>3</sup> Les choix didactiques évoquent l'utilisation des corpus en classe de langue, la description linguistique des collocations et l'aide à la mémorisation. Nous avons opté pour l'élaboration de cartes mentales (CM) par les apprenants sans toutefois arrêter définitivement ce choix pour l'enseignement du lexique à des LNN<sup>4</sup> du français.

---

\* Université Grenoble-Alpes, LIDILEM. Cet article existe grâce à la collaboration fructueuse avec Sascha Diwersy (Cologne).

\*\* LIRIS UMR 5205 (équipe SILEX), INSA, Lyon 1.

1 EMOLEX : [www.emolex.eu/](http://www.emolex.eu/), ANR-09-FASHS-017.

2 Scientext : <http://scientext.msh-alpes.fr> ; FULS : <http://webtek-66.iut2.upmf-grenoble.fr>

3 Français Langue Étrangère.

4 Locuteur Non Natif / LN = Locuteur Natif.

Nous expliquerons et justifierons ces choix : un corpus pour la description linguistique et l'enseignement, une base de données (BD) et des CM pour l'enseignement des collocations. Quelques questions quant à l'aide à la mémorisation et à la production écrite des apprenants concluront cet article.

## 2. Une liste de mots extraite du corpus

Le corpus du projet Emolex est constitué de textes littéraires et journalistiques contemporains. Plusieurs phénomènes linguistiques ont été analysés, notamment les collocations des lexies des émotions. Ces descriptions sont stockées (et codées) dans des tableaux (*cf.* tableau 1) qui s'apparentent à des listes de mots augmentées de descripteurs linguistiques. La question qui s'offrait à nous était celle de la réutilisation des ressources produites dans un contexte d'apprentissage. Pour des raisons méthodologiques, cette interrogation conduit à trois sous-questions : comment gérer et didactiser une liste de mots ? Quelles informations (issues des tableaux) sont réellement pertinentes pour l'enseignement/apprentissage des langues ? Comment les rendre disponibles aux apprenants/enseignants ?

emoclass	pivot	cat_pivot	collocatif	cat_collocatif	construction	rel_synt	dimension	fréquence
COLÈRE	indigné	A	(s')être	V	être ~	a4	expérenciation	391
COLÈRE	colère	N	accès	N	accès de ~	n31	aspect	30
COLÈRE	fureur	N	accès	N	un/des accès de ~	n31	aspect	9
COLÈRE	rage	N	accès	N	accès de ~	n31	aspect	10
COLÈRE	agacement	N	accroître	V	accroître ~	n71	causativité	3
COLÈRE	colère	N	affichage	N	affichage de la ~	n31	manifestation	4

*Tableau 1 : exemple de classement des extractions du corpus EmoConc en français*

La liste de mots est souvent controversée dans l'enseignement malgré sa présence récurrente dans les méthodes de FLE. L'utilité des listes est reconnue lorsqu'elles sont organisées et liées à une situation de communication. Il est donc question d'une hypothèse maintes fois vérifiée : classer aide à mémoriser (Goody 1979). Le choix pour ce projet est de classer trois fois : en tableaux, en tables informatiques pour une BD et, enfin, en CM pour les apprenants. Comme la liste participe à la mémorisation, rappelons quelques facteurs d'aide à mémoriser (Boulton 2000 ; Ellis 1997) :

- 1) l'apprentissage par cœur de listes, surtout au niveau A1 et pour les jeunes apprenants qui auraient davantage de facilités pour mémoriser des listes que les adultes ;

- 2) les mises en contexte à tous les niveaux de l'apprentissage ;
- 3) les réflexions métalinguistiques pour classer et organiser dès le niveau A1 ; ce biais didactique est à favoriser auprès des adultes qui possèdent les procédures nécessaires via leur langue première.

Ainsi, est-il pertinent de classer les éléments de la liste, trouver des contextes d'utilisation et faire réfléchir les apprenants à leur sujet.

### **3. Une BD pour linguistes et enseignants**

Le tableau 1, bien que recensant un grand nombre d'observations, est un document de travail plus qu'une ressource à part entière. L'inclusion des informations dans une BD nécessite une formalisation informatique des phénomènes décrits, qui à son tour permettra aux utilisateurs d'aborder son contenu différemment. Cette formalisation consiste à rendre explicite pour la machine des informations accessibles à l'humain, et ainsi offrir la puissance de calcul d'un ordinateur pour manipuler des objets ayant un sens pour l'utilisateur. Un exemple de ce travail de formalisation est la définition des lexies de telle manière que la colère de *l'accès de colère* soit la même que la colère de *l'affichage de la colère*, ce qui n'est pas le cas (du point de vue de l'ordinateur) dans le tableau 1, puisque les deux lexies appartiennent à deux lignes différentes. En outre, une BD permet de modéliser les relations, comme le fait que cette lexie *colère* puisse avoir une relation de type « collocation » avec les collocatifs *affichage* et *accès*, ou encore de lier cette collocation avec un exemple attesté du corpus.

L'un des atouts de cette base n'est pas seulement l'interrogation qu'elle propose – simple, claire et puissante car beaucoup de données sont accessibles –, mais le fait d'avoir accès à des informations rares pour le français :<sup>5</sup> la description linguistique des collocations accompagnées d'exemples du corpus.

Une fois implantée, la base est une source d'informations pour des utilisateurs avertis, mais dont l'accès peut être simplifié (souvent au détriment de la puissance expressive) par le développement d'une interface. Chaque interface vers la base de données doit prendre en compte les besoins des utilisateurs concernés : ce qui revient à l'identification, parmi les questions auxquelles la base peut répondre, de celles que peuvent poser les utilisateurs. Notre propos étant

---

<sup>5</sup> Il existe Antidote, qui propose notamment de grandes listes de collocations, cependant ce logiciel est payant (<http://www.druide.com/antidote.html>).

l'enseignement du lexique en classe de FLE, certaines questions s'imposent d'elles-mêmes : Comment connaître la collocation la plus fréquente parmi tous les champs ? Les bases (ou pivots ; le mot qui conserve son sens propre) qui apparaissent avec le même collocatif, ou, inversement, le collocatif qui apparaît avec plusieurs bases ? Quelle valeur sémantique est représentative de tel champ ?

Ces croisements de données sont autant d'informations que d'exercices sur divers points linguistiques. La collocation n'est que rarement transparente pour un LNN,<sup>6</sup> car il reste difficile à trouver ou comprendre un collocatif dont la déviance sémantique est souvent imprévisible (Cavalla *et al.* 2009 ; Hill 1993 ; Nesselhauf 2005 ; Verlinde *et al.* 2006) – ce qui n'est pas le cas pour des natifs (Hausmann 1979 ; Tutin 2010) :

[La] transparence n'empêche nullement la collocation d'être imprédictible. L'apprenant étranger, tout en la comprenant (s'il comprend les mots combinés), ne saurait automatiquement la reproduire. (Hausmann 1989, 1010)

Donc, faire manipuler les données permet souvent de mieux les comprendre. L'apprenant sera guidé dans sa démarche de classement, mais nous guiderons d'abord l'enseignant.

Ainsi, l'interface de la BD, nommée EmoProf, doit-elle être le plus accessible possible. Deux interfaces existent : l'une pour le public et l'autre pour ceux qui désirent peupler la base afin d'augmenter le nombre de sorties possibles. La structure interne de la base de données est d'ores et déjà apte à traiter des collocations de tout champ sémantique (pas seulement le lexique des émotions).

---

<sup>6</sup> « The main difference between native and non-native speakers is that the former have met far more English and so can recognize ready-made chunks, which enable them to process and produce language at a much faster rate. » (Hill 1993, 32).

**EmoProf**  
- Support didactique -

Guide

**Bienvenue** **Vocabulaire présélectionné** **Vocabulaire complet**

Champs sémantiques :  Admiration  Colère  Déception  Jalousie  Joie  TOUS ?  
 Mépris  Respect  Surprise  Tristesse

Mot(s) de base (Pivot) :  C ?

Catégorie(s) grammaticale(s) (Cat. Gr.) :  Vers la séquence ...

Mot(s) associé(s) (Collocatif) :  ?

Catégorie(s) grammaticale(s) (Cat. Gr.) :  ?

**Afficher** **Télécharger le tableau**

Filter:

Champ sémantique	Pivot	Cat.Gr.	Collocatif	Cat.Gr.	Construction	Valeur sémantique	Fréquence	Exemples
DECEPTION	<u>déçu</u>	Adj	être	V	être <u>déçu</u>	EXPERIENCIATION PRESENCE EXPERIENCIATION NEUTRE	248	<a href="#">afficher</a>
DECEPTION	<u>déçu</u>	Adj	peu	Adv	peu <u>déçu</u>	INTENSITE FAIBLE	129	<a href="#">afficher</a>
RESPECT	<u>respect</u>	N	grand	Adj	avoir un <u>grand respect</u> pour qqn	INTENSITE FORTE	101	<a href="#">afficher</a>

Figure 1 : écran 1 – interface d’interrogation

Pour l’interface, nous avons associé à la terminologie linguistique des lexies dont le sens semble transparent. L’usage des termes linguistiques en parallèle, outre le fait d’économiser de l’espace dans le tableau de résultat, vise à avoir une valeur formative. L’interrogation s’effectue à l’aide des étiquettes grammaticales connues des enseignants de FLE (adjectif = Adj, Nom = N, etc.). L’interface offre deux possibilités d’extraction des données :

- Le « vocabulaire présélectionné », qui rassemble les lexies utilisées pour l’élaboration des séquences didactiques. Ceci permet aux enseignants de travailler exactement comme nous l’avons fait et de s’approprier les exemples en manipulant la base.
- Le « vocabulaire complet », qui rassemble toutes les données de la BD.

Quelle que soit l’entrée choisie, l’interrogation reste identique ; les tableaux indiquent la fréquence et un lien vers des exemples dans le corpus. Pour remplir le formulaire, une aide est proposée et des exemples sont donnés.

L’interface propose une recherche booléenne (chaque propriété est présente ou absente dans une collocation). Selon la terminologie en logique, les propriétés indiquées dans chaque ligne sont traitées comme des *disjonctions* (dans l’écran 1, le mot de base est « déçu » *OU* « respect », etc.) alors que les lignes entre elles sont traitées comme des *conjonctions* : pour faire partie des résultats, la collocation doit avoir comme pivot l’un des lemmes de la liste *ET* comme

collocatif ceux qui ont été choisis (*être, grand, peu*). Tous les champs sont optionnels, ce qui signifie que la requête ne porte que sur le champ sémantique sélectionné ou tous par défaut. Voici quelques exemples de requêtes :

1) extraire toutes les collocations dont le pivot est un nom associé à un adjectif,

- ne déterminer aucun champ,
- catégorie grammaticale du mot de base : N,
- catégorie grammaticale du mot associé : Adj.

2) extraire toutes les collocations d’une liste fermée de mots de base de plusieurs champs et de catégories grammaticales différentes, avec une liste fermée de collocatifs ayant aussi des catégories grammaticales différentes (écran 1).

- champs : RESPECT, DECEPTION
- mots de base : *estime, respect, amertume, déçu*
- mots associés (collocatifs) : *grand, être, peu*

Filter:

Champ sémantique ▲	Pivot ⚡	Cat.Gr. ⚡	Collocatif ⚡	Cat.Gr. ⚡	Construction ⚡	Valeur sémantique ⚡	Fréquence ⚡	Exemples ⚡
DECEPTION	<u>déçu</u>	Adj	être	V	être <u>déçu</u>	EXPERIENCIATION PRESENCE EXPERIENCIATION NEUTRE	248	afficher
DECEPTION	<u>déçu</u>	Adj	peu	Adv	peu <u>déçu</u>	INTENSITE FAIBLE	129	afficher
RESPECT	<u>respect</u>	N	grand	Adj	avoir un <u>grand respect</u> pour qqn	INTENSITE FORTE	101	afficher
RESPECT	<u>estime</u>	N	grand	Adj	<u>grande estime</u>	INTENSITE FORTE	25	afficher
RESPECT	<u>estime</u>	N	peu	Adv	avoir <u>peu d'estime</u> pour/de qqn	INTENSITE FAIBLE	20	afficher

Nombre de résultats: 5

Figure 2 : écran 2 – exemple d’extraction de croisements de données

Les lexies sont classées, par défaut, par la classe sémantique (ordre alphabétique) mais le tableau est interactif et l’en-tête de chaque colonne permet de s’en servir comme critère de tri. En outre, il est possible de classer les sorties, une fois affichées, en utilisant la case « filtrer » ; exemple : dans l’écran 2, entrer « forte », permettra de n’afficher que les éléments à valeur sémantique d’intensité forte (ex : « *grande surprise* »).

Pour les enseignants de FLE, nous proposons des exemples d'exploitation de ces données à partir d'extractions à classer sous forme de CM. Les exemples proposés ne sont pas exhaustifs et peuvent être adaptés à d'autres besoins. L'aspect métalinguistique d'un tel classement – quels que soient le niveau en langue et l'âge des apprenants – nous est apparu pertinent pour l'aide à la mémorisation des éléments phraséologiques.

#### **4. Des séquences pour le FLE**

Le site EmoProf contient des pistes didactiques guidant l'enseignant vers l'utilisation des extractions et de la base. Les séquences sont téléchargeables. Les « pistes didactiques » proposées ne sont pas des séquences complètes car chaque enseignant peut intégrer ce lexique et ces outils quand il le souhaite au fil de sa progression. Dans chaque proposition, on élabore une CM qui est contextualisée au sein des pistes proposées.

Ainsi avons-nous deux objets d'études : les collocations et les cartes mentales. Les collocations relèvent d'une partie du lexique et les cartes sont un outil pour l'enseignement d'un objet. Notons que les collocations sont peu enseignées en FLE et que les CM paraissent répondre à plusieurs de nos attentes en termes d'aide à la mémorisation du lexique.

#### **5. L'enseignement des collocations**

Chercheurs et didacticiens du FLE ont développé des outils pour l'aide à l'enseignement/apprentissage des éléments phraséologiques : la BLF<sup>7</sup> associée au site d'exercices Alfalex<sup>8</sup> (Verlinde *et al.* 2006) ; le site « La peur »<sup>9</sup> et le développement de séquences autour des expressions des émotions (Cavalla *et al.* 2005, Cavalla *et al.* 2009) ; l'ouvrage pour l'enseignement de la phraséologie dans son ensemble (Gonzalez-Rey 2008). Nous espérons que cet engouement touchera bientôt les auteurs de manuels de FLE qui intègrent avec parcimonie ces éléments linguistiques fréquents. Les chercheurs précédemment cités proposent diverses solutions quant à leur abord en classe (Binon *et al.* 2003, Cavalla 2009, Gonzalez-Rey 2010) et posent désormais la question de l'aide à la mémorisation d'éléments aux constructions qui paraissent parfois complexes (Tutin 2007).

---

7 Base Lexicale du Français, URL : <http://ilt.kuleuven.be/blf/> [consulté le 26 février 2013].

8 URL : <http://www.kuleuven.be/alfalex> [consulté le 26 février 2013].

9 LIDILEM, URL : <http://w3.u-grenoble3.fr/lidilem/projets/ESC/> [consulté le 26 février 2013].



L'utilisation des corpus est un atout pour l'enseignement des collocations car l'extraction sous forme de concordances permet à l'enseignant de repérer rapidement les formes figées associées à une lexie (Chambers *et al.* 2011 ; Johns 2002). Ceci soulève la question de l'intégration des corpus en classe de FLE, sujet que nous n'aborderons pas ici (Aston 2002 ; Boulton 2008 ; Cavalla *et al.* à paraître ; Chambers 2007 ; Fligelstone 1993). Notons qu'ils sont utilisés en classe de langue par un nombre croissant d'enseignants. Le concordancier, parfois gratuit,<sup>10</sup> parfois utilisable en ligne, est probablement l'outil le plus représenté parmi les dispositifs pédagogiques visant à intégrer la notion de corpus en classe de langue. Peu d'approches pédagogiques existent autour des corpus, dont la profusion lexicale pourrait engendrer une abondance de type d'exploitation (Boulton 2008 ; Chambers *et al.* 2011).

Nous proposons ici un autre point d'accès vers les corpus, par le biais des CM. Celles-ci imposent un recours au corpus pour vérifier une intuition ou pour mener une réflexion métalinguistique fondée sur les tableaux fournis par la base de données (qui renvoient eux-mêmes vers des concordances).

## **6. Choix de cartes mentales**

Les CM sont des schémas de type arborescent développés avec l'ambition de suivre un modèle du fonctionnement cérébral (pensée par association). Le type de relations représentées étant variable (lien hiérarchique, lien sémantique), les CM apparaissent dans différents contextes. Elles constituent un intermédiaire entre la liste classique et la production écrite (Barna 2009) ; ceci pourrait satisfaire les trois critères indiqués par Ellis (1997) et Boulton (2000) : apprentissage de listes, mise en contexte, réflexion métalinguistique pour aider à la production écrite par exemple. Certains y voient un « [o]util pédagogique puissant et versatile » (Peters *et al.* 2005) qui s'adapte à tous les niveaux de l'apprentissage, à tous les concepts et à tous les âges.

Les concepts principaux sont organisés en fonction de liens qui explicitent la relation entre ces concepts : des exemples, des composantes, des caractéristiques, etc. (Van Zele/Lenaerts/Wieme, 2004, cité par Peters *et al.* 2005)

---

<sup>10</sup> Par exemple : <http://www.ddc-concordance.org/>

De fait, le classement satisferait nos attentes en termes de représentation de la syntagmatique des collocations. Nous savons (Boulton 2000 ; Goody 1979) qu'organiser des listes, se concentrer sur un élément sont des biais pertinents pour l'apprentissage et donc la mémorisation. La manipulation du lexique par classement ou découpage (morphologique par exemple) permet à l'apprenant de se concentrer sur l'objet, voire de trouver sa place d'acteur « linguiste ».

## 7. Exemple de pistes didactiques

Les CM apparaissent soit en début, en milieu ou en fin de proposition didactique. L'objectif est d'aboutir à une CM qui pourrait représenter un état du savoir des apprenants à un moment  $t$ . Il est donc prévu que les cartes soient extensibles (à l'infini en théorie) aux autres savoirs des apprenants dans la pratique.

La progression proposée est inductive (aspect renforcé par l'élaboration des cartes mentales) et nous permet d'avoir un cadre didactique autour des CM. Plusieurs champs sémantiques sont pris comme exemples ; chaque séquence a un déroulement différent en fonction des objectifs visés. L'enseignant est guidé vers l'utilisation de la base et l'élaboration de la CM. Ces aspects se retrouvent dans les rubriques « Cadre de la séquence », « Déroulement » et « Réinvestissement », rubriques communes à tous les exemples.

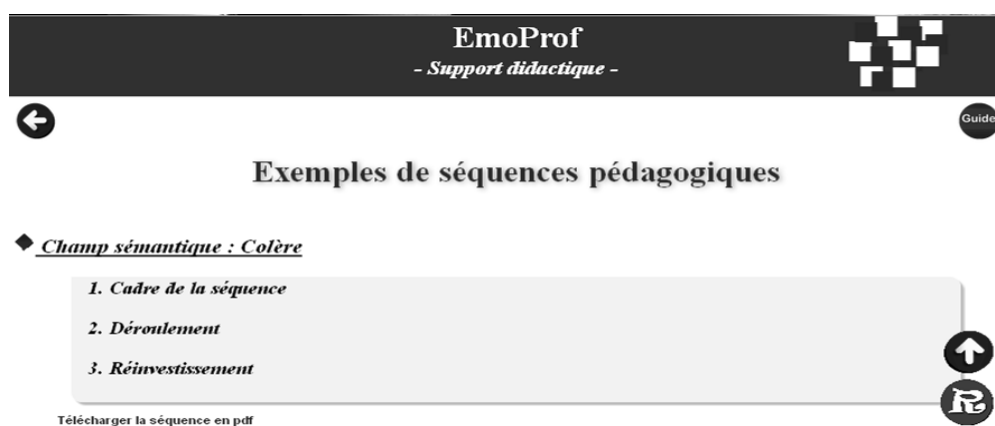


Figure 3 : écran 3 – interface d'accès aux séquences

Différents niveaux sont abordés afin de tester les CM et trier le lexique. Par exemple, l'un des critères de tri des lexies du champ COLÈRE est celui de « constructions », où les lexies du champ sont « objet » du verbe associé (*apaiser sa colère, cacher son exaspération*, etc.). Le cadre de la

séquence varie en fonction des objectifs visés, mais préconise toujours une contextualisation du lexique dans des documents authentiques. Répertoire avec les apprenants leurs connaissances sur ce lexique est préconisé quasi systématiquement. Ainsi, toutes les propositions guident-elles les apprenants vers une réflexion sur les lexies et les encouragent à trouver des liens entre elles et à créer leurs propres données lexicales. Il faut alors envisager un temps de création de la CM par les apprenants. Des consignes sont proposées dans la rubrique « Déroulement » afin de guider les enseignants et engendrer d'autres idées pour l'exploitation des CM en classe.

La démarche pour élaborer la carte préconise d'avancer par étapes. En effet, il n'est pas nécessaire d'envisager la CM dans son entier dans un premier temps ; au contraire, plusieurs temps pédagogiques sont nécessaires à la mémorisation. De fait, l'élaboration de parties de CM en fonction d'une logique syntagmatique (ou autre) aide à la mémorisation de ces liens parmi tous ceux rencontrés autour d'une lexie. Une fois plusieurs CM élaborées, il semble pertinent de les rassembler. Pour associer plusieurs CM, il suffit parfois d'un seul lien entre deux éléments. Ci-dessous, dans la figure 4 (CM1), le classement sémantique prend en compte les verbes permettant de désigner le stade évolutif de l'affect : le début, le milieu, la fin. La combinatoire syntaxique des éléments apparaît dans le sens des flèches entre les éléments ; par exemple *se mettre en* est un nœud qui se combine avec *colère* ou *rage*, les flèches partent donc du nœud pour se diriger vers l'objet qui est postposé au verbe. Dans la CM2 (cf. figure 5), les flèches jouent leur rôle syntagmatique, et d'autres nœuds ajoutent un lien sémantique entre les éléments ; « visible » indique que l'émotion se « voit » par différents biais exprimés par le verbe.

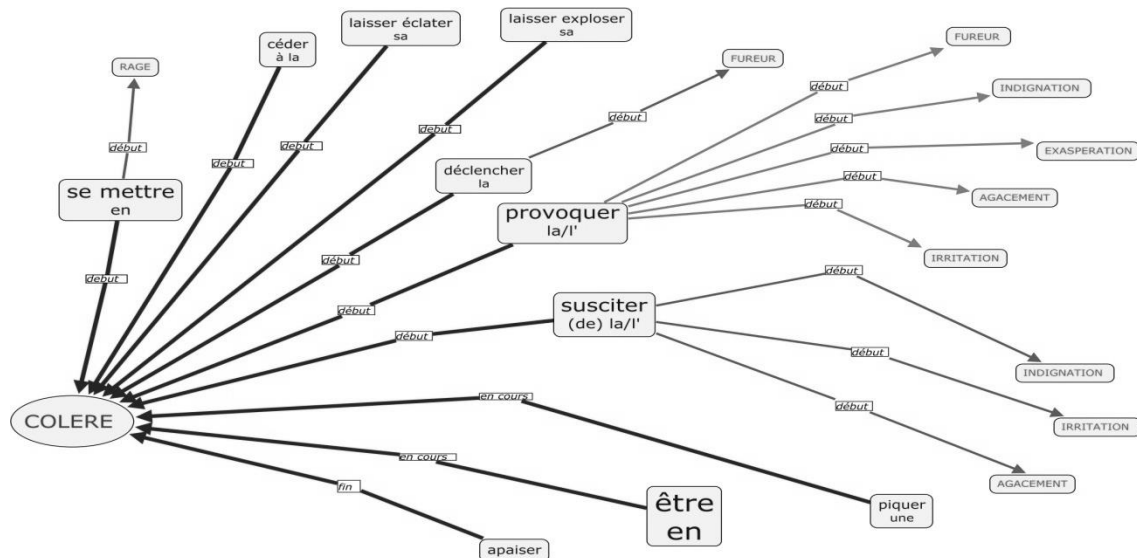


Figure 4 : CM1 – évolution de la colère (début, milieu, fin)

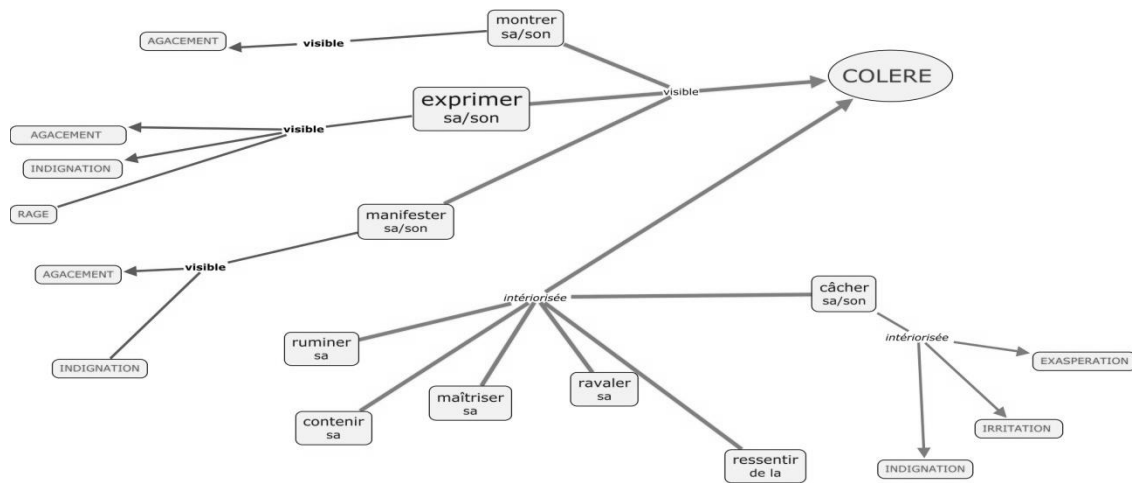


Figure 5 : CM2 – présence visible ou non de la colère

Nous pouvons davantage scinder ces CM si cela convient à la progression prévue par l’enseignant ; la CM1 se partage en trois : le début, le milieu et la fin de l’émotion. La CM3 (cf. figure 6) ci-dessous rassemble les deux premières afin d’avoir une vue d’ensemble des liens possibles entre toutes les lexies. Notons qu’il est possible d’ajouter des couleurs indiquant, par exemple, des sous-groupes sémantiques.

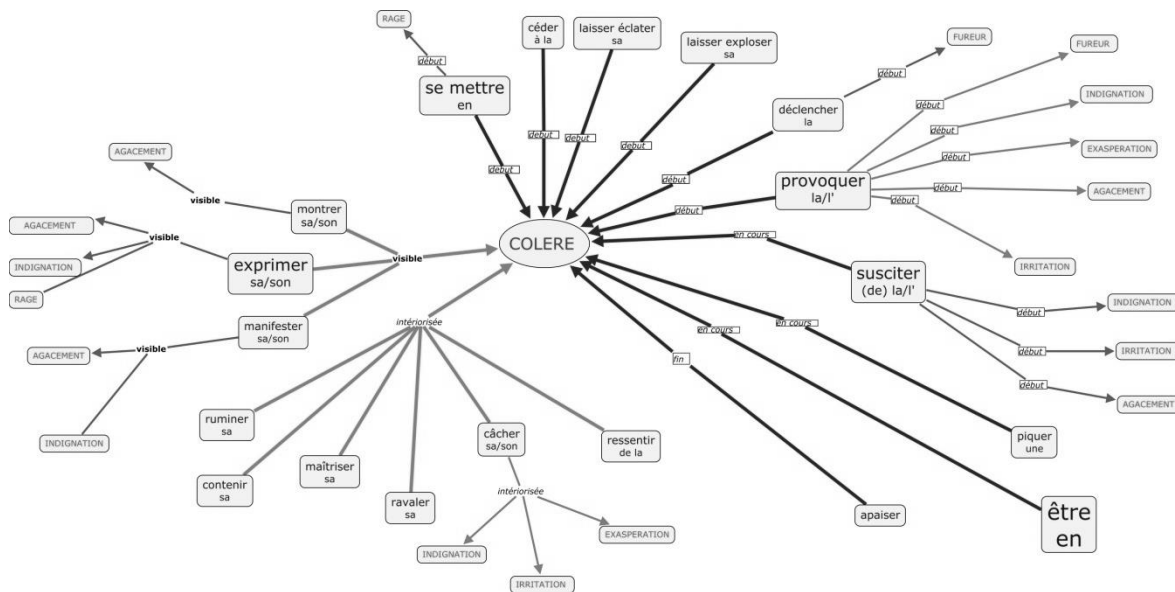


Figure 6 : CM3 – CM1 et CM2 rassemblées

Ces cartes sont élaborées à l'aide d'un logiciel gratuit « CmapTools » (Cañas *et al.* 2004) ; d'autres logiciels existent et l'utilisation du papier fonctionne également. L'important n'est pas tant l'outil pour la fabriquer que le processus de fabrication lui-même. L'objectif, dans le choix de ce dispositif pédagogique, est que la manipulation des lexies, pour faire la carte, aide les apprenants à conceptualiser et retenir différents types d'informations essentielles pour l'utilisation à bon escient, en contexte, de la lexie en question (son sens, sa syntagmatique pour les collocations, etc.). En d'autres termes, nous faisons l'hypothèse que la discussion métalinguistique, inhérente à la tâche de classification, peut avoir un effet analogue à celui de la discussion métalinguistique dans le cadre d'une tâche de production (Swainn 2000). À défaut, se posera la question de savoir si les compétences méthodologiques acquises (discussion métalinguistique à partir d'exemples, introduction au corpus, outils pour la conceptualisation) peuvent favoriser la prise de conscience métalinguistique et l'autonomie des apprenants dans le cadre d'activités de production.

## **8. Bilan et discussion**

Dans les pistes didactiques proposées, les enseignants de FLE utilisent à leur gré la BD et les CM. Ils peuvent envisager de faire interroger la base par les apprenants ; ultime phase préconisée (Chambers *et al.* 2004) pour l'accès aux corpus en classe de langue.

L'élaboration d'une CM étant inexistante dans les méthodes de FLE (donc méconnue de nombreux enseignants), il est nécessaire de la présenter de façon précise et pédagogiquement contrainte dans un premier temps. En outre, les recherches sur la manipulation de ces cartes montrent qu'elles sont utilisées auprès d'enfants mais de façon moindre pour les adultes. L'aide à la mémorisation grâce à ces CM paraît efficace pour l'adulte car il possède les cadres cognitifs langagiers absents chez l'enfant. Enfin, reconnaissons que la conception d'une CM prend du temps et qu'un réinvestissement est à prévoir. EmoProf propose un réinvestissement des cartes soit en les complétant au fil de l'apprentissage, soit pour des applications externes comme l'aide à la production. La BD est un outil intéressant pour l'enseignement car elle donne accès à un grand nombre d'informations linguistiques précieuses pour l'enseignant ; il faudra toutefois l'enrichir d'autres lexies.

Une fois ces outils et ces séquences établis, reste en suspens la question du passage du classement à la mémorisation du lexique pour un usage en contexte. La question de l'enseignement explicite des formes pour une communication effective est en débat (Ortega 2009, Huot *et al.* 1996). Cependant, la recherche tend vers une coexistence entre un enseignement explicite de certaines formes, associé à une approche implicite (Doughty 2001). Notre entrée, par la forme des collocations, s'associe à l'implicite lorsque le sens de chaque élément n'est pas directement décliné dans la carte, mais que, appartenant à un champ sémantique, l'apprenant doit déduire. Sans l'avoir testé, nous ne pouvons l'affirmer, cependant, puisque classer, organiser du lexique aide à le fixer dans nos mémoires (Goody 1979), l'utilisation de ces CM nous paraît être une aide intéressante pour l'apprentissage du lexique en classe de langue.

Pour finir, les perspectives envisagées pour le projet sont d'élargir les domaines d'application de la BD en y intégrant de nouvelles données (champs sémantiques) ou via la réutilisation du code par d'autres.<sup>11</sup> Nous souhaiterions également proposer d'autres pistes pour l'enseignement des collocations et voir d'autres biais pédagogiques – hormis les CM – pour l'aide à la mémorisation. Des tests d'utilisation de EmoProf sont en cours et nous souhaiterions comprendre comment l'insertion de ces CM en classe de langue pourrait aider à la production notamment d'écrits universitaires.

Enfin, une autre piste serait de concevoir un prototype d'interface qui intègre la BD à un logiciel de création de CM. En effet, la quantité d'informations disponible (augmentée grâce à des outils TAL) pourrait permettre au système d'inférer la sémantique des relations à partir des données appariées. Le système pourrait, dès lors, être utilisé comme support du script d'Argue Graph (Dillenbourg *et al.* 2007), qui vise à faire débattre les apprenants.

## **Bibliographie**

- Aston, Guy (2002). « The Learner as Corpus Designer », *in* : Bernhard Kettemann/Georg Marko (eds.) : *Teaching and Learning by Doing Corpus Analysis*. New York : Edwin Mellen, 9-26.
- Barna, Judith (2009). « L'article de presse à l'épreuve », *Les Langues modernes* 2, 21-27.

---

<sup>11</sup> Le code sera rendu disponible une fois une version stable atteinte.

- Binon, Jean/Verlinde, Serge (2003). « Les collocations : clef de voûte de l'enseignement et de l'apprentissage du vocabulaire d'une langue étrangère ou seconde », *La Lettre de l'AIRDF* 33, 31-36.
- Boulton, Alex (2000). « Aspects lexicaux de l'acquisition < naturelle > et de l'apprentissage < artificiel > en L2 », *Les Mélanges du CRAPEL* 26, 63-90.
- Boulton, Alex (2008). « Esprit de corpus : promouvoir l'exploitation de corpus en apprentissage des langues », *Texte et Corpus* 3, 37-46.
- Cañas, Alberto J./Hill, Greg/Carff, Roger/Suri, Niranjana/Lott, James/Gómez, Gloria/Eskridge, Thomas C./Arroyo, Mario/Carvajal, Rodrigo (2004). « CmapTools : A Knowledge Modeling and Sharing Environment », in : Alberto J. Cañas *et al.* (eds.) : *Concept Maps : Theory, Methodology, Technology. Proceedings of the First International Conference on Concept Mapping*. Pamplona : Universidad Pública de Navarra, URL: <<<http://cmc.ihmc.us/papers/cmc2004-283.pdf>>> [consulté le 26 septembre 2013], 125-133.
- Cavalla, Cristelle (2009). « La phraséologie en classe de FLE », *Les Langues modernes* 1-2009, URL: <<<http://www.aplv-languesmodernes.org/spip.php?article2292>>> [consulté le 26 septembre 2013].
- Cavalla, Cristelle/Crozier, Elsa (2005). *Émotions-Sentiments*. Grenoble : Presses Universitaires de Grenoble.
- Cavalla, Cristelle/Labre, Virginie (2009). « L'enseignement en FLE de la phraséologie du lexique des affects », in : Agnès Tutin/Iva Novakova (eds.) : *Le lexique des émotions et sa combinatoire lexicale et syntaxique*. Grenoble : Ellug, 297-316.
- Cavalla, Cristelle/Loiseau, Mathieu (à paraître). « Scientext comme corpus pour l'enseignement », in : Agnès Tutin/Francis Grossmann (eds.) : *Autour de Scientext*. Grenoble.
- Chambers, Angela (2007). « Popularising Corpus Consultation by Language Learners and Teachers », in : Encarnación Hidalgo Tenorio/Luis Quereda Rodríguez-Navarro/Juan Santana (eds.) : *Corpora in the Foreign Language Classroom*. Amsterdam/New York : Rodopi, 3-16.
- Chambers, Angela/Farr, Fiona/O'riordan, Stéphanie (2011). « Language Teachers with Corpora in Mind: from Starting Steps to Walking Tall », *The Language Learning Journal* 39(1), 85-104.

- Chambers, Angela/O'sullivan, Ide (2004). « Corpus Consultation and Advanced Learners' Writing Skills in French », *ReCALL* 16(1), 158-172.
- Dillenbourg, Pierre/Tchounikine, Pierre (2007). « Flexibility in Macro-scripts for Computer-supported Collaborative Learning », *Journal of Computer Assisted Learning* 23, 1-13.
- Ellis, Nick C. (1997). « Vocabulary Acquisition : Word Structure, Collocation, Word-class, and Meaning », in : Norbert Schmitt/Michael McCarthy (eds.) : *Vocabulary Description, Acquisition and Pedagogy*. Cambridge : Cambridge University Press, 122-139.
- Fligelstone, Steve (1993). « Some Reflections on the Question of Teaching, from a Corpus Linguistics Perspective », *ICAME journal* 17, 97-109.
- Gonzalez-Rey, Isabel (2008). *La didactique du français idiomatique*. Fernelmont : Éditions Modulaires Européennes.
- Gonzalez-Rey, Isabel (2010). « La phraséodidactique en action : les expressions figées comme objet d'enseignement », *La Clé des Langues* (ENS LYON/DGESCO), URL : <<<http://cle.ens-lyon.fr/espagnol/la-phraseodidactique-en-action-les-expressions-figees-comme-objet-d-enseignement-92012.kjsp>>> [consulté le 26 septembre 2013].
- Goody, Jack (1979). *La raison graphique – la domestication de la pensée sauvage*. Paris : Minuit.
- Hausmann, Franz Josef (1979). « Un dictionnaire des collocations est-il possible ? », *Travaux de littérature et de linguistique de l'université de Strasbourg* XVII(1), 187-195.
- Hausmann, Franz Josef (1989). « Le dictionnaire de collocations », in : Franz Josef Hausmann *et al.* (eds.) : *Wörterbücher : ein internationales Handbuch zur Lexikographie*. Berlin/New York : de Gruyter, 1010-1019.
- Hill, Jimmie (1993). « Revising Priorities : from Grammatical Failure to Collocational Success », in : Michael Lewis (ed.) : *The Lexical Approach, the State of ELT and the Way Forward*. London : Language Teaching Publications, 47-67.
- Johns, Tim (2002). « Data-driven Learning : the Perpetual Challenge », in : Bernhard Kettemann/Georg Marko (eds.) : *Teaching and Learning by Doing Corpus Analysis*. New York : Edwin Mellen, 107-118.
- Nesselhauf, Nadja (2005). *Collocations in a Learner Corpus*. Amsterdam : Benjamins.
- Peters, Martine/Chevrier, Jacques/Leblanc, Raymond/Fortin, Gilles/Malette, Judith (2005). « Compétence réflexive, carte conceptuelle et webfolio à la formation des maîtres », *Canadian*



*Journal of Learning and Technology* 31(3), URL : <<<http://www.cjlt.ca/index.php/cjlt/article/viewArticle/96>>> [consulté le 26 septembre 2013].

Swain, Merrill (2000). « The Output Hypothesis and beyond : Mediating Acquisition through Collaborative Dialogue », in : James P. Lantolf (ed.) : *Sociocultural Theory and Second Language Learning*. Oxford : Oxford University Press, 97-114.

Tutin, Agnès (2007). « Collocations du lexique transdisciplinaire des écrits scientifiques : annotation et extraction des propriétés linguistiques dans la perspective d'une application didactique », *Cahiers de l'Institut de linguistique de Louvain* 31(2-4), 247-262.

Tutin, Agnès (2010). « Le traitement des collocations dans les dictionnaires monolingues de collocations du français et de l'anglais », in : Franck Neveu *et al.* (eds.) : *Congrès Mondial de Linguistique Française (CMLF-2010)*. Paris : Institut de Linguistique Française, 1075-1090, URL: <<[http://www.linguistiquefrancaise.org/articles/cmlf/pdf/2010/01/cmlf2010\\_000141.pdf](http://www.linguistiquefrancaise.org/articles/cmlf/pdf/2010/01/cmlf2010_000141.pdf)>> [consulté le 26 septembre 2013].

Verlinde, Serge/Binon, Jean/Selva, Thierry (2006). « Corpus, collocations et dictionnaires d'apprentissage », *Langue française* 150, 84-98.