

Du dictionnaire au site lexical pour l'enseignement/apprentissage du vocabulaire



Anne-Laure Jousse, Alain Polguère, Ophélie Tremblay

1. Introduction

Le dictionnaire pédagogique pourrait être l'outil privilégié de l'enseignement et apprentissage des connaissances lexicales. Cependant, dans sa forme actuelle, il ne contient pas nécessairement toutes les informations dont les enseignants ou les élèves ont besoin, les premiers, dans leur préparation d'activités pédagogiques, les seconds, dans leurs tâches scolaires. Cet article vise à proposer des pistes pour améliorer l'accès aux données sur le lexique par le biais de l'utilisation d'un type de ressource nouvellement apparu, les sites lexicaux, et par le recours à une formation linguistique visant le développement de la compétence lexicale des utilisateurs.

Dans la section 2, le dictionnaire est examiné de façon critique, du point de vue du type d'informations qu'on y retrouve et de la façon dont celles-ci sont encodées. Nous introduisons ensuite, dans la section 3, le site lexical comme alternative au dictionnaire. Nous faisons la présentation d'une approche relationnelle de la modélisation du lexique — celle de la Lexicologie Explicative et Combinatoire — et décrivons une base de données lexicale du français fondée sur cette approche : le DiCo et sa version en ligne DiCouèbe. La section 4 explicite les modes de navigation qui peuvent être intégrés dans un site lexical conçu à partir d'une telle base (approche thématique et approche structurante). Finalement, la section 5 est consacrée aux compétences mises en jeu dans l'utilisation d'une ressource lexicale pour l'enseignement et l'apprentissage du lexique, notamment dans un contexte de production de texte.

2. Limitations des dictionnaires conventionnels

2.1. Dictionnaires comme modèles du lexique

Les dictionnaires de langue se veulent des descriptions du lexique, des « modèles lexicaux », conçus pour être utilisés par le grand public comme ouvrages de référence et de consultation. Ils devraient donc être en mesure de répondre aux questions les plus évidentes que l'on peut se poser sur le fonctionnement des unités lexicales de la langue. Cependant, les dictionnaires remplissent rarement ce rôle, notamment lorsqu'ils sont utilisés dans un contexte de production langagière (et pas simplement pour vérifier l'orthographe ou le sens d'un mot peu fréquent).

Pour bien démontrer la nature des limitations du modèle lexical offert par les dictionnaires conventionnels, nous examinerons plusieurs cas d'échecs dans l'accès aux unités lexicales à partir de la nomenclature des dictionnaires : échecs au niveau de la navigation dans la **macrostructure** des dictionnaires. Ensuite, nous considérerons des cas d'échecs dans l'accès à l'information à l'intérieur des articles de dictionnaires eux-mêmes : échecs au niveau de la navigation dans la **microstructure**. Dans chaque cas, nous prendrons comme point de départ les dictionnaires papier, tout en faisant aussi référence aux dictionnaires électroniques.

2.2. Navigation dans la macrostructure des dictionnaires

La recherche dans les dictionnaires papier est contrainte par l'ordre alphabétique et se caractérise, de ce fait, par une navigation fastidieuse (souvent décourageante, spécialement pour un jeune public). Les dictionnaires électroniques ont l'avantage de donner accès directement aux articles recherchés, puisqu'il suffit de taper le mot désiré. Certains dictionnaires électroniques, comme le dictionnaire du logiciel de correction Antidote, offrent même des alternatives pour les requêtes mal orthographiées. Cependant, que ce soit pour les dictionnaires papier ou les dictionnaires électroniques, l'accès direct par mot identifiant une entrée est souvent insuffisant lorsque l'utilisateur cherche certains moyens linguistiques d'encoder sa pensée.

Imaginons un cas concret où un rédacteur cherche des verbes sémantiquement liés au nom peur. Les dictionnaires papier traditionnels ne proposant pas de structuration thématique, le rédacteur est contraint de chercher à l'intérieur de l'article de peur, ou de mots sémantiquement proches, dans l'espoir d'y trouver mention des verbes recherchés. Par exemple, frayer sera peut-être donné comme synonyme dans l'article de peur, et l'article de frayer lui-même contiendra peut-être la mention du verbe effrayer (dont l'emploi est beaucoup plus courant que celui du dérivé morphologique apeurer). Les dictionnaires électroniques, bien que leur structure soit moins soumise à une nomenclature rigide comme l'est la nomenclature alphabétique, ne proposent généralement pas d'accès au lexique par des thèmes. Cependant, certains modes de recherche permettent de plus ou moins simuler un tel accès. Par exemple, le dictionnaire de la version 5 du correcteur Antidote présente un onglet « Analogies », qui donne accès, à partir de l'article d'un mot donné, aux éléments de la nomenclature dont l'article contient le mot en question. Il ne s'agit cependant pas d'un véritable mode d'accès par grandes catégories thématiques.

Par ailleurs, si la navigation dans le « réseau lexical de la langue » est facilitée par le recours aux dictionnaires électroniques, elle peut devenir erratique dans le cas où l'utilisateur n'a pas une idée précise de l'information recherchée. Les informations visuelles et sonores que fournissent les dictionnaires électroniques destinés aux jeunes, comme le *Petit Robert Junior*, sont un bon complément à une recherche linguistique, mais peuvent peut-être détourner l'utilisateur de son objectif de recherche initial. Bref, dans bien des cas, aussi bien le dictionnaire papier que le dictionnaire électronique échouent quand il s'agit d'accéder de façon simple et systématique aux unités lexicales en tant qu'outils d'encodage de la pensée.

2.3. Accès aux données dans la microstructure des dictionnaires

Les informations que l'on retrouve dans un article de dictionnaire sont de plusieurs ordres : orthographe, prononciation, étymologie, définition, polysémie, synonymes et mots reliés, exemples, etc. Dans un contexte de rédaction de texte, ces informations ne sont pas toutes pertinentes, puisque l'utilisateur cherche avant tout des moyens linguistiques d'encoder sa pensée. Par exemple, il peut chercher une expression précise évoquant le fait qu'un sentiment de peur naît chez une personne. Où trouvera-t-il les expressions *la peur s'insinue en X*, *la peur s'immisce en X*, *la peur naît chez X*, *la peur saisit X...*, qui sont des collocations types contrôlées par peur ? La zone d'exemples de chaque article de dictionnaire contient souvent de telles collocations ; cependant, le sens véhiculé par chacune n'est pas spécifié, ce qui oblige l'utilisateur, dans l'exemple présenté, à chercher le sens des différents verbes employés avec peur pour trouver celui qui correspond le mieux à son intention de communication. Cet exemple montre qu'un tel type de recherche dans le dictionnaire a de fortes chances d'être voué à l'échec, les dictionnaires traditionnels ne rendant pas compte de façon explicite et systématique du phénomène particulier qu'est la collocation.

Cet échec dans la recherche de l'information est-il inéluctable ? Nous pensons qu'il peut être résolu en faisant appel à d'autres modélisations du lexique. De fait, le dictionnaire est moins caractérisé par un problème d'accès aux données que par une modélisation inadéquate de certains phénomènes lexicaux. Il faut donc faire appel à des modèles qui illustrent mieux la façon dont s'encode l'information lexicale et qui reproduisent davantage le processus d'accès aux unités lexicales et aux combinaisons lexicales.

3. Le modèle relationnel du lexique sous-jacent à un site lexical

3.1. Du dictionnaire au site lexical

Le modèle classique du dictionnaire est de plus en plus remis en question par le développement de ressources lexicales « en ligne » qui ne sont pas de simples versions sur support électronique de dictionnaires papiers, mais de véritables **sites lexicaux**. Nous faisons référence ici aux ressources orientées vers le traitement automatique de la langue, comme WordNet (Fellbaum, 1998) ou FrameNet (Baker *et al.*, 2003) — pour ne citer que les plus connues —, mais aussi aux travaux visant la construction de ressources à visées pédagogiques, comme le *Dictionnaire d'apprentissage du français langue étrangère ou seconde* DAFLES (Selva *et al.*, 2002) ou le *Diccionario de Colocaciones del Español* (Alonso Ramos, 2003). Toutes ces ressources lexicales, pourtant très différentes les unes des autres dans leur conception, reposent sur une vision non linéaire du lexique, où la modélisation de ce dernier est plus inspirée des modèles de stockage informatiques du type bases de données que des modèles « textuels » des dictionnaires. On peut très grossièrement caractériser la nouvelle vision des modèles lexicaux qui se dégage peu à peu de ces travaux comme relevant d'une organisation plus **relationnelle** des lexiques, ces derniers étant conçus comme de gigantesques graphes d'entités lexicales interconnectées (Polguère, 2006).

Bien que les auteurs du DAFLES continuent de parler de *dictionnaire* pour désigner la ressource qu'ils développent, la plupart des lexicographes et chercheurs impliqués dans la construction de la nouvelle génération de ressources lexicales préfèrent généralement utiliser d'autres termes. Tel qu'indiqué au début de la présente section, nous adoptons quant à nous le terme *site lexical* car il présente un double avantage. Tout d'abord, contrairement à *dictionnaire* (très connoté), il n'implique pas l'idée d'une organisation textuelle linéaire du modèle lexical. Ensuite, le terme *site* qu'il contient fait explicitement référence à une ressource « en ligne », qui est en évolution constante.

Le site lexical se caractérise avant tout par son architecture qui, contrairement à la structure linéaire du dictionnaire, vise à modéliser la connaissance lexicale comme un réseau d'entités interconnectées. Notons cependant que nombreux sont les projets de mise en ligne de dictionnaires « classiques » qui visent à s'affranchir de la linéarité, notamment par le recours à un balisage des données (XML ou assimilé) permettant une navigation plus souple dans les données ; tel est le cas, par exemple, du *Trésor de la Langue Française informatisé* TLFi (Pierel *et al.*, 2004). Un véritable site lexical, cependant, doit reposer sur un modèle relationnel explicite du lexique, comme celui de la Lexicologie Explicative et Combinatoire, que nous présentons brièvement dans la section suivante.

3.2. La modélisation relationnelle de la Lexicologie Explicative et Combinatoire

La Lexicologie Explicative et Combinatoire (dorénavant, LEC) est une théorie lexicologique orientée vers la description lexicographique formelle (Mel'čuk *et al.*, 1995). Outre la modélisation du sens au moyen de définitions analytiques (par genre prochain et différences spécifiques), la LEC met l'accent sur la description de deux types de liens connectant les unités lexicales dans le réseau lexical de la langue : les **liens paradigmatiques** et les **liens syntagmatiques** contrôlés par chaque unité lexicale.

Les liens paradigmatiques correspondent au phénomène de **dérivation sémantique** : une relation sémantique forte (non nécessairement accompagnée d'une relation morphologique) qui unit deux unités lexicales : *orgueil ~ amour propre, tomber ~ chute, bateau ~ capitaine, passion ~ passionné, peur ~ effrayer*, etc. Les liens syntagmatiques correspondent aux **collocations** contrôlées par les unités lexicales ; par exemple, pour *peine* : *grande, vive... peine, peine inconsolable...*, *guérir, soulager qqn. de sa peine*, etc.

Dans la LEC, les relations paradigmatiques et syntagmatiques sont modélisées au moyen du système formel des **fonctions lexicales**. Chaque fonction lexicale encode une relation récurrente en langue et possède un sens général. Un peu comme une fonction mathématique, elle s'applique à un argument pour retourner une valeur (liste d'unités lexicales ou expressions exprimant la relation en question). Une application de fonction lexicale se décrit par une formule du type **f(x)=y**, où **x** est l'argument de la fonction lexicale **f** et **y** est la valeur retournée. Par exemple, la fonction lexicale **Magn**, dénotant les modificateurs d'intensification, permet de modéliser la relation syntagmatique unissant *passion* à l'ensemble de ses intensificateurs typiques : **Magn**(*passion*) = *ardente, brûlante, intense...* La relation entre *peine* et *apaiser*, quant à elle, est modélisée par la fonction **LiquFunc₀**, qui désigne un ensemble de verbes signifiants *grosso modo* 'causer que qqch. cesse' :

LiquFunc₀(*peine*) = *apaiser, calmer, dissiper, guérir, soulager la peine*. Un des plus grands avantages théoriques et pratiques du recours aux fonctions lexicales est qu'elles permettent de mettre formellement en relation des liens dérivationnels et collocationnels, souvent décrits de façon totalement séparée (Mel'čuk et Polguère, sous presse).

Outre la description fine des relations lexicales qui vient d'être présentée, la LEC répond à une approche discrète de la polysémie, c'est-à-dire une approche qui distingue explicitement toutes les acceptions d'un vocable. Cela est essentiel car, à l'intérieur d'un vocable polysémique, chacune des acceptions possède son propre réseau de relations lexicales. Par exemple, les unités lexicales **fardeau I** [*Elle déposa son lourd fardeau devant la porte.*] et **fardeau II** [*Elle ne veut pas être un fardeau pour ses parents.*] ne contrôlent pas les mêmes liens dérivationnels et collocationnels. Elles se distinguent bien entendu par leurs quasi-synonymes : **charge I**, **chargement...** dans le premier cas, et **charge II**, **poids**, **responsabilité...** dans le second. Mais elles se distinguent aussi par certaines de leurs collocations ; par exemple, *être, représenter, devenir... un fardeau* sont autant de collocations contrôlées uniquement par **fardeau II** (métaphorique).

La LEC a donné lieu à plusieurs types de modélisations, dont le *Dictionnaire Explicatif et Combinatoire* DEC, qui est un produit lexicographique de type dictionnaire papier (Mel'čuk *et al.*, 1984, 1988, 1992, 1999). Cependant, l'approche de la LEC permet aussi le développement de véritables bases de données lexicales, telles le DICE pour l'espagnol (déjà mentionné plus haut) ou le DiCo pour le français. Nous nous attarderons maintenant sur cette dernière base de données, puisqu'elle nous sert de point de référence comme source de données type pour un site lexical.

3.3. Le DiCo et sa version en ligne DiCouèbe

Le DiCo est une base de données lexicale composée d'un ensemble de « fiches », dont chacune décrit, dans des champs distincts, les propriétés sémantiques et combinatoires d'une unité lexicale donnée. Nous ne présenterons pas de façon détaillée la structure du DiCo, nous contentant d'en énumérer les caractéristiques essentielles. Le lecteur intéressé pourra consulter, par exemple, Polguère (2000) ou Jousse et Polguère (2005).

Une fiche de DiCo explicite, pour chaque unité lexicale (nous prenons ici *peine* comme exemple) :

1. ses caractéristiques grammaticales (par exemple, pour *peine*, *nom*, *fém*) ;
2. son étiquette sémantique (*sentiment négatif*) ;
3. sa structure actantielle (*peine de l'individu X causée par le fait Y*) ;
4. son tableau de régime, qui spécifie la façon dont les actants de l'unité lexicale (X, Y, Z...) s'expriment dans des structures syntaxiques qu'elle gouverne ;
5. ses dérivés sémantiques, décrits au moyen des fonctions lexicales dites « paradigmatiques » (*peiné*, *pénible...*) ;
6. les collocations qu'elle contrôle, décrits au moyen des fonctions lexicales dites « syntagmatiques » (*grande peine*, *peine passagère*, *remplir de peine...*) ;
7. des exemples d'emplois ;
8. la liste des locutions — unités lexicales multilexémiques) dans lesquelles l'unité lexicale apparaît (âme en peine, faire peine à voir...).

Comme on peut le constater en lisant la brève présentation ci-dessus, le modèle original du DiCo, structuré par fiches lexicographiques décrivant chacune une unité lexicale spécifique, est finalement très proche du modèle classique des dictionnaires. Cependant, la nature essentiellement relationnelle du modèle sous-jacent — celui de la LEC — permet de compiler la base DiCo en des structures beaucoup moins linéaires, comme dans le cas du DiCo-SQL, accessible par l'interface DiCouèbe (Steinlin *et al.*, 2005). Le DiCouèbe, que nous ne pouvons présenter ici faute de place (Jousse et Polguère, 2005), peut être considéré comme un site lexical à part entière. Cependant, son design de type « formulaire », très proche des données elles-mêmes, fait qu'il est peu adapté à une utilisation dans un contexte pédagogique. Dans la suite de notre article, nous faisons référence par *site lexical* à un site s'appuyant sur le même type de données que celles sous-jacentes au DiCouèbe, mais permettant une navigation adaptée aux applications pédagogiques. La prochaine section est dédiée à ce problème du mode d'accès aux données lexicales.

4. Modes de navigation dans un site lexical

Il est difficile d'envisager l'utilisation d'un outil comme le DiCouèbe en contexte d'enseignement/apprentissage du lexique à l'école, notamment au primaire, en raison de la richesse du formalisme utilisé et de la rigidité du mode d'accès aux données. On doit donc chercher une alternative qui permette d'accéder de façon plus conviviale au réseau lexical encodé selon une approche relationnelle de type LEC. Ceci revient à orienter le parcours d'un élève, mais aussi d'un enseignant, dans le réseau lexical. Pour cela, il faut identifier un certain nombre de requêtes correspondant aux questions que l'on peut se poser lors de la rédaction d'un texte.

Nous avons identifié deux approches de la navigation dans l'information lexicale, qui donnent lieu à deux types d'interfaces : l'approche thématique et l'approche structurante.

4.1. L'approche thématique

L'**approche thématique** propose de grandes catégories conceptuelles dans lesquelles l'utilisateur peut naviguer pour trouver le mot adéquat, à partir d'un besoin d'expression de certains contenus très généraux. Par exemple, si un utilisateur cherche à verbaliser, autour du sens de l'unité lexicale *crainte*, l'idée de faire cesser la crainte, il peut cibler la grande catégorie thématique **phases et durée** du sentiment, dont la phase **fin** est un cas spécifique. Cette navigation lui donnera accès à des expres-

sions du type *la crainte s'estompe, dissiper la crainte*, etc. On peut ainsi dire que l'approche thématique consiste à organiser les dérivés sémantiques et les collocations par thèmes correspondant à des contenus à exprimer. Nous avons identifié quatre grands thèmes : 1) les synonymes et antonymes, 2) les phases et la durée, 3) la caractérisation du sentiment, 4) l'expression et la caractérisation des différents participants (qui correspondent aux actants de l'unité lexicale). Ces thèmes sont hiérarchisés, du plus général au plus spécifique, ce qui permet de préciser pas à pas la requête, jusqu'à trouver les expressions répondant aux besoins de communication.

Cette première approche sera maintenant illustrée à partir de l'unité lexicale *crainte*, dont la structure prédicative est 'crainte de l'individu X causée par Y'. (Les titres qui apparaissent ci-dessous sont ceux que l'on pourrait retrouver dans les pages d'un site lexical.)

1) Synonymes et antonymes de crainte

- Générique : *sentiment [de crainte]*
- Synonymes : *angoisse, appréhension, effroi, inquiétude, peur, phobie*
- Équivalent verbal : *craindre*
- Antonyme : *confiance*

2) Phases et durée

2.1) Phases

- Début
 - La crainte commence à être éprouvée par X : *La crainte assaille, gagne, saisit X, La crainte s'empare de X*
 - La crainte devient plus intense : *La crainte de X augmente*
 - Qqch. rend la crainte plus intense : *Qqch. accroît, attise, conforte, excite la crainte de X*
 - Qqn. ou Qch. cause de la crainte chez X : *Qqn./Qqch. éveille, sème la crainte en X*
- Déroulement
 - X éprouve de la crainte : *X a, éprouve, nourrit, ressent de la crainte ; X conçoit, entretient de la crainte*
 - Une crainte est éprouvée par X : *Une crainte dévore, habite, hante, ronge X*
- Déroulement, à travers les réactions ou actions de X
 - X ne se laisse pas influencer par sa crainte : *X surmonte sa crainte*
 - X manifeste sa crainte par une réaction physique : *X frémit, pâlit, est/reste paralysé, est/reste tétanisé, tremble, tressaille de crainte*
 - X ne laisse pas sa crainte se manifester : *X cache, dissimule sa crainte*
 - X fait qqch. en éprouvant de la crainte : *X fait qqch. avec crainte, dans la crainte*
 - X fait qqch. sans éprouver de crainte : *X fait qqch. sans (aucune) crainte*
 - X fait qqch. du fait de sa crainte : *X fait qqch. de, par crainte de Y ; X fait qqch. dans la crainte de (voir) Y*
- Fin
 - La crainte de X devient moins intense : *La crainte de X se calme, diminue, se dissipe*
 - Qqch. rend la crainte de X moins intense : *Qqch. apaise, atténue, calme la crainte de X*
 - Qqn./Qqch. fait cesser la crainte de X : *Qqn./Qqch. apaise, balaie, calme, chasse, dissipe, lève, vainc la crainte de X*
- Résurgence
 - Une crainte est éprouvée de nouveau : *Une crainte resurgit*

- Qqn./Qqch. cause de nouveau une crainte : *Qqn./Qqch. réveille une crainte*

2.2) Durée

- Crainte qui dure : *crainte continue, perpétuelle*
- Crainte qui ne dure pas : *crainte passagère*

3) Caractérisation du sentiment

- Caractérisation positive
 - Crainte justifiée : *crainte fondée, juste, justifiée, légitime*
 - Crainte utile : *crainte salutaire*
- Caractérisation négative
 - Crainte non justifiée : *crainte chimérique, illusoire, infondée, injustifiée, irrationnelle ; fausse, vaine crainte ; crainte exagérée, excessive ; crainte irréflectie ; crainte puérile*
- Caractérisation fondée sur l'intensité du sentiment
 - Crainte intense : *folle, forte, grande, vive crainte ; crainte profonde, sérieuse, terrible*
 - Crainte peu intense : *légère, petite, sourde, vague crainte*

4) Participants

- X
 - [X] qui éprouve de la crainte : *[X] fou, plein, saisi de crainte ; [X] hanté, obsédé par la crainte*
 - [X] qui n'éprouve pas de crainte : *[X] sans crainte*
 - [X] qui tend à éprouver de la crainte : *[X] craintif, peureux*
- Y
 - Nom pour Y : *[X est] l'objet de la crainte de X*

4.2. L'approche structurante

L'**approche structurante** vise à aider le rédacteur dans le choix de la structure d'une phrase à produire. Par exemple, si c'est le sentiment de crainte que l'on veut mettre en valeur par rapport à X qui éprouve ce sentiment, on pourra dire *Une grande crainte habite X*. Si au contraire on désire mettre X de l'avant, on pourra dire *X est habité/hanté par une grande crainte*. L'approche structurante donne également au rédacteur la possibilité de choisir s'il veut utiliser une expression au style direct ou au style indirect. L'approche structurante consiste donc à donner accès à des outils d'expression véhiculant un contenu donné à l'intérieur de structures de phrase bien spécifiques. Ces outils sont avant tout des collocatifs à valeur syntaxique : des **verbes supports**. On trouvera dans la théorie de la LEC un recours systématique à de tels outils au sein de règles de paraphrasage faisant appel au système des fonctions lexicales (Mel'čuk, 1988 ; Milićević).

Parmi les outils d'expression auxquels on accède selon l'approche structurante, nous avons jusqu'à présent distingué deux cas : 1) le choix de la place et de la fonction syntaxique d'une unité lexicale et de ses actants dans la phrase ; 2) le choix d'une expression explicite ou implicite du contenu.

1) Position syntaxique du prédicat sémantique et de ses actants (X, Y, Z...) dans la phrase

La place relative de l'expression d'un prédicat et de ses actants dans la phrase contribue à ce que l'on appelle la **structure communicative** de l'énoncé — notamment pour ce qui est de la dichotomie thème/rhème (Mel'čuk, 2001).

- X est sujet

- X est sujet et *crainte* (et éventuellement Y) est complément : *X éprouve, ressent, nourrit de la crainte vis-à-vis de Y; X conçoit, entretient de la crainte vis-à-vis de X*
- Y est sujet
 - Y est sujet et *crainte* (et éventuellement X) est complément : *Y cause, engendre, provoque, suscite de la crainte chez X; Y inspire de la crainte à X*
- Le nom de sentiment est sujet
 - *Crainte* est sujet du verbe et X est complément : *La crainte dévore, habite, hante, ronge X*

2) Expression explicite ou implicite du contenu

Nous n'avons pas d'exemples impliquant crainte à citer ici. Cependant, il n'est pas rare que l'on dispose de moyens de dénoter un sentiment de façon implicite, par le biais de l'expression de la manifestation (physique) du sentiment. Par exemple, nous avons pour peur des expressions comme [*Les cheveux de X*] *se dressent sur sa tête*. On peut aussi disposer d'expressions implicites fondées sur les manifestations verbales de X, de type interjections. Par exemple, pour communiquer le sens 'X exprime son dégoût', on peut utiliser une formulation explicite, dans des expressions comme *X manifeste, montre, marque, témoigne son dégoût* ou, même, *X grimace de dégoût*. Mais on peut aussi exprimer implicitement le sens en question en mettant dans la bouche de X des interjections du type *Berk !, Beurk !* ou *Pouah !* : *Francis lut l'article qu'on lui avait envoyé. « Pouah ! », dit-il*

Pour conclure cette section, mentionnons que d'autres points d'entrée dans le lexique sont bien entendu envisageables, mais que nous ne pouvons les examiner ici faute de place. Signalons aussi que, comme l'aura peut-être remarqué le lecteur, les deux approches proposées pour la navigation dans un site lexical peuvent, dans certains cas, donner accès aux mêmes données linguistiques. Cela est voulu, et représente même à notre avis un des avantages principaux de la stratégie que nous proposons, puisqu'une même expression peut tout à fait est utilisée pour répondre à plusieurs besoins communicatifs distincts en situation de production de texte.

5. Utilisation d'un site lexical en contexte d'apprentissage scolaire

Les approches qui viennent d'être présentées servent à donner accès aux données lexicales de façon plus intuitive que par la structure tabulaire du DiCouébe. Cependant, nous croyons qu'avant d'utiliser une **ressource lexicale** quelle qu'elle soit — et pas uniquement celle dont nous présentons l'architecture —, l'usager doit avoir une bonne connaissance de la ressource elle-même et de son fonctionnement. Nous nous attacherons donc maintenant à la présentation des utilisations potentielles du site lexical proposé et surtout, à la description des connaissances lexicales mises en jeu dans l'utilisation efficace d'un tel site lexical.

5.1. Les ressources lexicales et leur exploitation

Nous distinguons trois types d'utilisateurs ayant recours à des ressources lexicales : le grand public, les élèves et les enseignants. Le grand public fait une utilisation des ressources lexicales qu'on pourrait qualifier de conventionnelle : recherche de l'orthographe ou du sens d'un mot mal maîtrisé ou inconnu, recherche de synonymes ou d'antonymes, recherche d'exemples, etc. L'élève, lui, dans les activités qu'il réalise en classe — qu'il s'agisse d'activités d'apprentissage du vocabulaire, de compréhension de texte ou de production écrite — va rechercher le même genre d'information. Cependant, des activités spécifiques le conduisent à utiliser le dictionnaire de façon plus systématique qu'un usager grand public. L'élève est ainsi amené à se familiariser progressivement avec la structure du dictionnaire et à acquérir de nouvelles connaissances qui lui permettent de mieux utiliser ce type

d'ouvrage : apprentissage de l'ordre alphabétique, de la notion de polysémie, des notions de synonymie et d'antonymie, etc. Les enseignants, quant à eux, font normalement un usage conventionnel du dictionnaire, mais ils peuvent aussi l'utiliser comme outil d'aide à la préparation d'activités sur le lexique. Simard (1990) note à cet effet que les enseignants ne sont pas suffisamment formés pour faire une telle utilisation du dictionnaire, alors que ce dernier peut bel et bien être utilisé pour développer des activités lexicales (Jousse et Tremblay, 2006).

On peut donc supposer que la façon dont l'utilisateur a recours aux différents types de ressources lexicales dépend des compétences linguistiques et métalinguistiques qu'il possède. Nous identifierons maintenant les compétences et connaissances nécessaires pour utiliser efficacement un site lexical comme celui dont nous présentons l'architecture. Celles-ci sont potentiellement les mêmes pour les différents types d'utilisateurs, que ce soit l'élève qui cherche des informations spécifiques concernant la combinatoire lexicale des unités lexicales dans son processus de rédaction de texte ou l'enseignant qui souhaite élaborer des activités d'apprentissage ciblant particulièrement le lexique.

5.2. Le rôle de la compétence lexicale dans l'utilisation d'un site lexical

La principale compétence linguistique mobilisée dans l'utilisation d'une ressource comme celle que nous présentons est sans aucun doute la **compétence lexicale**. On peut également évoquer l'existence d'une certaine **compétence métalexical**e impliquée elle aussi dans le processus de recherche et d'accès à l'information. Quelques auteurs ont formulé des définitions de la notion de compétence lexicale qui mettent en évidence les multiples facettes de celle-ci (Tréville, 2000 ; Nisubire, 2003 ; Simard, 2004). Nous proposons quant à nous de considérer la compétence lexicale comme un ensemble de connaissances, d'habiletés et d'attitudes liées au lexique. Les connaissances lexicales, élément central de la compétence lexicale, relèvent de deux phénomènes centraux dans la langue (Polguère et Tremblay, 2003), soit la **polysémie** — qui doit nous forcer à isoler les différentes acceptions d'un mot pour pouvoir les caractériser et les différencier — et la **connexion lexicale** — qui fait qu'une unité lexicale avant tout par 1) son positionnement dans le réseau lexical de la langue (liens paradigmatiques) et 2) sa combinatoire lexicale propre, qui contrôle la façon dont elle se combine aux autres unités lexicales dans la phrase (liens syntagmatiques). Si l'on n'est pas au fait de ces phénomènes, on aura du mal à naviguer efficacement dans le réseau d'informations disponibles à l'intérieur d'un site lexical. Dans la perspective de l'utilisation d'un site lexical pour l'aide à la rédaction, notamment dans le contexte de l'écriture d'un texte mobilisant des noms de sentiments (Grossmann et Boch, 2003), l'élève doit connaître le phénomène de collocation et comprendre qu'il s'agit d'une combinaison contrainte exprimant un sens très général (correspondant en fait à un patron de collocation).

Comme on peut le voir, l'expression des sentiments est soumise à différentes contraintes lexicales et il nous semble important de sensibiliser les élèves à l'existence d'un tel phénomène linguistique. Cet enseignement devrait par ailleurs être conduit dans la perspective d'une sensibilisation à l'existence de la paraphrase comme moyen linguistique pour encoder la pensée. En effet, une même idée peut être exprimée de différentes façons, notamment en jouant avec le rôle syntaxique des éléments constitutifs d'une collocation. Par exemple, le sens 'apparition/naissance d'un sentiment' peut être exprimé 1) en faisant jouer le rôle de sujet syntaxique au nom de sentiment, 2) en utilisant le premier actant du nom de sentiment comme sujet : *X s'angoisse*, *X est soulevé de dégoût*, *X conçoit du désir*, *X prend peur*, etc.

D'autres relations paraphrastiques existent, mais nous ne les examinerons pas ici, faute de place (voir les publications sur la paraphrase citées plus haut, section 4.2). Ce qu'il importe de retenir, c'est que le lexique interagit de façon privilégiée avec la syntaxe dans les manipulations présentées et qu'il existe

des schémas d'encodage qui interviennent de façon récurrente dans le travail de production linguistique.

La connaissance de l'existence des patrons de collocations permet enfin de distinguer le sens général d'une relation et sa réalisation contrainte par l'idiomaticité. Par exemple, l'expression de l'intensité d'un sentiment est un patron de collocation productif et les élèves peuvent faire des erreurs lexicales à cause d'un parasitage de collocations — ce que Polguère (à paraître) appelle *greffes collocationnelles*. Dans une expérimentation portant sur la réceptivité des élèves à des activités sur le lexique (Tremblay, 2003), des erreurs de collocations de ce type ont ainsi pu être observées : **panique bleue*, **colère rouge*, etc. Écrire de façon idiomatique est un signe de maîtrise de la langue et l'école devrait amener les élèves à ce niveau de maîtrise linguistique. Le type de ressources que nous avons présentée ici peut selon nous contribuer à améliorer la maîtrise du vocabulaire, notamment si les enseignants intègrent davantage le recours à de telles ressources lexicales dans leurs activités quotidiennes sur la langue.

5.3. En guise de conclusion : pour une utilisation accrue des ressources lexicales par les enseignants

Les enseignants jouent un rôle important dans le processus d'apprentissage de la langue par leurs élèves. Il semble cependant que, au chapitre du lexique, les enseignants dirigent peu d'activités lexicales dans une perspective de production de texte, comme le note Dreyfus (2004). En effet, les activités de vocabulaire sont le plus souvent centrées sur l'orthographe lexicale ou l'acquisition de nouveaux mots (mais pas dans une perspective rédactionnelle).

Le site lexical peut servir aux enseignants pour le développement de toutes sortes d'activités faisant intervenir des manipulations lexicales, mais aussi syntaxiques, activités qui nous semblent potentiellement productives en termes d'aide à la rédaction, notamment parce qu'elles participent au développement de la conscience réflexive face aux opérations langagières à l'œuvre dans une tâche d'écriture narrative. Le site lexical, tel que nous le concevons, prend toute son importance dans ce contexte puisqu'il fournit les différentes façons d'exprimer un même sens, ce qui peut contribuer à enrichir un texte (Boch, 2005). Ainsi, une approche lexicale de la production d'écrits pourrait se révéler particulièrement productive, notamment si les ressources lexicales telles que les dictionnaires d'apprentissage et les sites lexicaux devenaient partie intégrante des activités de planification des cours et d'enseignement proprement dit.

Remerciements

Les propositions faites dans le présent article ont trouvé leur source dans les travaux de recherche en lexicologie de l'Observatoire de linguistique Sens-Texte (OLST), financés par une subvention du *Fonds québécois de la recherche sur la Société et la culture* (FQRSC), ainsi que dans le cadre d'un projet de recherche du Lidilem (Grenoble 3) — action concertée *École et Sciences Cognitives* du Gouvernement français.

Bibliographie

Alonso Ramos M. (2003) La nature des collocatifs : leur statut en tant qu'unités lexicales. In F. Grossmann et A. Tutin (dir.), coll. « Travaux et Recherches en Linguistique Appliquée », De Werelt, Amsterdam, 45-60.

Baker C., Fillmore C., Cronin B. (2003) The Structure of the Framenet Database. *International Journal of Lexicography*, 16:3, 281-296.

- Boch F. (2005) Apprentissage du lexique et production d'écrits : une articulation féconde ? *Actes du 9^e colloque international de l'AIRDF : le français, discipline singulière, plurielle ou transversale*, Québec. [Disponible en ligne ou sur cederom : http://www.colloqueairdf.fse.ulaval.ca/actes/index.php?action=par_titre]
- Dreyfus M. (2004) Pratiques et représentations de l'enseignement du lexique à l'école primaire. Communication présentée au 9^e colloque de l'Association Internationale pour la Recherche en Didactique du Français, Québec, août 2004.
- Fellbaum C. (1998) *WordNet: An Electronic Lexical Database*, MIT Press, Cambridge, MA.
- Grossmann F., Boch F. (2003) Productions de textes et apprentissage lexical : l'exemple du lexique de l'émotion et des sentiments. *Repères*, 28, 117-135.
- Jousse A.-L., Polguère A. (2005) *Le DiCo et sa version DiCouèbe. Document descriptif et manuel d'utilisation*, Département de linguistique et de traduction, Université de Montréal.
- Jousse A.-L., Tremblay O. (2006) Le dictionnaire électronique : un outil précieux pour enseigner le lexique. *Québec Français*, 141.
- Lewis M. (1993) *The lexical approach: The state of ELT and the way forward*, Language Teaching Publications, Hove.
- Mel'čuk I. (1988) Paraphrase et lexique dans la théorie Sens-Texte. *Lexique*, 6, 13-54.
- Mel'čuk I. (2001) *Communicative Organization in Natural Language. The semantic-communicative structure of sentences*, coll. « Studies in Language Companion Series », Amsterdam/Philadelphia.
- Mel'čuk I. et al. (1984, 1988, 1992, 1999) Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques, vol. I-IV, Les Presses de l'Université de Montréal, Montréal.
- Mel'čuk I., Clas A., Polguère A. (1995) *Introduction à la lexicologie explicative et combinatoire*, Duculot, Louvain-la-Neuve.
- Mel'čuk I., Polguère A. (sous presse) Dérivations sémantiques et collocations dans le DiCo/LAF. *Langue française*, numéro spécial sur la collocation « Collocations, corpus, dictionnaires », sous la direction de P. Blumenthal et F. J. Hausmann, Larousse, Paris.
- Milićević J. (2003) Étude des aspects sémantiques et syntaxiques de la paraphrase : application à la génération automatique des phrases françaises. Thèse de doctorat, Département de linguistique et de traduction, Université de Montréal.
- Nisubire P. (2003) Développer la compétence lexicale en classe de français langue seconde. *La lettre de l'AIRDF*, 33, 2003-2, 20-26.
- Paquette G. (2002) *La modélisation des connaissances et des compétences, pour concevoir et apprendre*, Presses de l'Université du Québec, Québec.
- Picoche J., Rolland J.-C. (2002) *Dictionnaire du français usuel*, Duculot, Paris.
- Pierrel J.-M., Dendien J., Bernard P. (2004) Le TLFi ou Trésor de la Langue Française informatisé. *Actes de EURALEX 2004*, Lorient, 165-170.
- Polguère A. (2000) Towards a theoretically-motivated general public dictionary of semantic derivations and collocations for French. *Proceedings of EURALEX 2000*, Stuttgart, 517-527.
- Polguère A. (2006) Structural properties of Lexical Systems: Monolingual and Multilingual Perspectives. *Proceedings of the Post COLING/ACL-2006 Workshop: Multilingual Language Resources and Interoperability*, Sydney.
- Polguère A. (à paraître) *Soleil insoutenable et chaleur de plomb : le statut linguistique des greffes collocationnelles*. In M.-C. L'Homme et S. Vandaele (éd.) : *Les approches lexicographiques et terminologiques sont-elles compatibles ?* [titre provisoire], coll. « Pédagogie de la traduction », Presses de l'Université d'Ottawa, Ottawa.
- Polguère A., Tremblay O. (2003) Qu'y a-t-il à l'intérieur de noix ? Ou comment décortiquer les unités

lexicales. *La Lettre de l'AIRDF*, 33, 2003-3, 27-30.

Selva T., Verlinde S., Binon J. (2002) Le DAFLES, un nouveau dictionnaire pour apprenants du français. *Actes de EURALEX 2002*, Copenhague, 199-208.

Simard C. (1994) Pour un enseignement plus systématique du lexique. *Québec Français*, 92, 28-40.

Simard C. (1990) Les besoins lexicographiques du milieu de l'enseignement du Québec. In : *Dix études portant sur l'aménagement de la langue au Québec*, Conseil de la langue française, Éditeur officiel du Québec, Québec, 25-51.

Steinlin J., Kahane S., Polguère A. (2005) Compiling a "classical" explanatory combinatorial lexicographic description into a relational database. *Proceedings of the Second International Conference on the Meaning Text Theory*, Moscou, 477-485.

Tremblay O. (2004) Pour une approche structurée de l'enseignement/apprentissage du lexique. In É. Calaque et J. David (dir.) : *Didactique du lexique : contextes, démarches, supports*, coll. « Savoirs et pratiques », De Boeck, Bruxelles, 127-139.

Tréville M.-C. (2000) *Vocabulaire et apprentissage d'une langue seconde*, Les Éditions Logiques, Montréal.