

# APPROCHE HYPERTEXTUELLE VIA HTML POUR UN OUTIL DE CONSIGNATION BILINGUE DES COMBINAISONS LEXICALES SPÉCIALISÉES

*Isabelle MEYNARD*

Département de linguistique et de traduction  
Université de Montréal

## 1. Introduction

Depuis une quinzaine d'années, l'étude des combinaisons lexicales a amené les auteurs à concevoir des outils de consignation de ces groupes de mots. Ainsi, on trouve des ouvrages sur support papier [Cohen 1986, Lainé 1993], d'autres sur bases de données [Heid & Freibott 1991]. Cependant, force est de constater que ces deux modes de consignation présentent des inconvénients non négligeables : contraintes d'espace et de redondance pour le support papier et pour les bases de données, processus de construction de la base nécessitant de la part du concepteur des connaissances en programmation.

Depuis quelques années, on a vu apparaître une nouvelle forme d'édition de documents, à savoir la publication sur Internet. Lors de nos recherches [Meynard 1997a], nous avons consulté un grand nombre de sites lexicographiques, ce qui nous a permis d'identifier trois approches principales de conception :

- Pages écrites en langage HTML où des liens ont été insérés pour renvoyer à d'autres parties de la page ou à d'autres pages ;
- Pages écrites en langage HTML qui ont été reliées à une base de données pour la recherche de résultats lors de requêtes ;
- Présentation en format PDF, c'est-à-dire dans un format spécial que l'utilisateur ne peut afficher à l'écran de son ordinateur qu'à condition que ce dernier soit muni du logiciel *Acrobat Reader*. Il s'agit, ici, d'une technique de plus en plus répandue sur Internet, pour la publication de gros volumes d'information.

Lors de notre consultation des sites lexicographiques sur Internet, nous n'avons cependant rencontré aucun site qui soit exclusivement consacré à la consignation des combinaisons lexicales relatives au domaine de spécialité qui nous intéressait, à savoir, l'autoroute de l'information. En effet, même s'il arrive que certains dictionnaires ou glossaires intègrent des combinaisons lexicales, en sous-entrées ou en contexte, aucun ne les place en entrée de manière systématique.

Nous basant sur cette lacune, nous nous sommes employée à concevoir une méthode de consignation des combinaisons lexicales spécialisées (CLS) bilingues, à la fois flexible et accessible. Nous avons donc imaginé un outil en format HTML susceptible d'offrir aux langagiers non-spécialistes du domaine de l'autoroute de l'information, un accès convivial aux usages de ce domaine.

Dans cet article, nous présentons donc l'outil que nous avons imaginé pour la consignation des combinaisons lexicales spécialisées bilingues. Cet outil a été imaginé de manière à pouvoir être adaptable à différents domaines de spécialité (économie, médecine, etc.). Pour nous assurer de la faisabilité d'un tel outil, nous avons consulté un spécialiste en génie logiciel et ergonomie des interfaces personnes-machines, qui nous a expliqué comment il est possible de construire des modèles d'écrans et de pages.

## 2. Sélection de l'information

Nous avons, dans un premier temps, sélectionné toutes les informations dont un langagier (étudiant, traducteur, terminologue, rédacteur technique, etc.) non-spécialiste du domaine pourrait avoir besoin lors de l'accomplissement de son travail :

1. Termes noyaux anglais et français ;
2. Définitions (dans les deux langues) ;
3. Sources des définitions ;
4. Catégories de combinaisons lexicales ;
5. Contextes des combinaisons lexicales : chaque combinaison lexicale est accompagnée d'une fiche où est indiqué le contexte dans lequel la combinaison a été repérée, ce qui permet, en outre, à l'utilisateur de vérifier l'usage grammatical dans lequel elle s'inscrit ;
6. Sources des contextes des combinaisons lexicales ;
7. Fiches de synthèse bilingues.

Les catégories de combinaisons lexicales que nous avons sélectionnées sont les suivantes :

- (a) nom cooccurrent + terme noyau (exemple : suppression de fichier) ;
- (b) terme noyau + verbe cooccurrent (exemple : logiciel supporte) ;
- (c) verbe cooccurrent + terme noyau (exemple : insérer un lien) ;
- (d) terme noyau + adjectif cooccurrent (exemple : requête entrante) ;
- (e) collocate noun + base noun (exemple : configuration of a session) ;
- (f) collocate verb + base noun (exemple : to edit a window) ;
- (g) base noun + collocate verb (exemple : address links to) ;
- (h) collocate adjective + base noun (exemple : parent directory).

À partir des catégories de CLS, nous avons également établi trois listes de cooccurrents dans chacune des deux langues :

- Noms cooccurrents (exemples : compression) – Collocate Nouns (exemples : design, integrity) ;
- Verbes cooccurrents (exemples : implémenter) – Collocate verbs (exemples : to initiate) ;
- Adjectifs cooccurrents (exemples : distant) – Collocate adjectives (exemples : remote).

Toutes les informations que nous avons sélectionnées sont également regroupées par terme, dans une fiche de synthèse bilingue. Ces fiches présentent les mêmes informations dans les deux langues mettant ainsi en évidence les équivalences d'une langue à l'autre, seule la disposition des informations change.

Ainsi, pour le français la fiche de synthèse se présente comme suit :

Terme français	Base noun
Définition	Définition
Nom cooccurrent + terme noyau	
Verbe cooccurrent + terme noyau	
Terme noyau + verbe cooccurrent	
Terme noyau + adjectif cooccurrent	

Pour l'anglais, la fiche de synthèse se présente comme suit :

Base noun	Terme français
Définition	Définition
Base noun + collocate noun	
Collocate verb + base noun	
Base noun + collocate verb	
Collocate adjective + base noun	

### 3. Organisation de l'information à l'intérieur de l'outil

Par la suite, nous avons réparti toutes ces données en deux écrans, en fonction de la relation qui les unissait aux termes noyaux :

- Un écran est utilisé pour la consignation des données offrant à l'utilisateur une information directement reliée au terme noyau. On trouve ainsi, dans cet écran, la définition du terme noyau, les sources de cette définition, la liste des combinaisons lexicales dans lesquelles s'insère le terme noyau, la fiche de synthèse bilingue qui est consacrée au terme noyau.
- Un écran est utilisé pour la consignation des données concernant plus particulièrement les combinaisons lexicales (contexte de la combinaison lexicale et sources de ce contexte).

### 4. Structure de l'outil

Nous avons défini deux modules de navigation par lesquels l'utilisateur peut avoir accès aux informations concernant les termes noyaux ou celles concernant les combinaisons lexicales : le module de recherche et le module d'exploration.

#### 4.1. Module de recherche

Ce module offre à l'utilisateur une page dans laquelle apparaît une zone de texte où l'utilisateur saisit sa requête, une bande déroulante qui permet à l'utilisateur de déterminer quelle fonction il veut assigner au contenu de sa requête et un bouton de commande pour lancer la requête.

À l'heure actuelle, le module de recherche de l'outil n'a pas encore été programmé, cependant, son développement est en cours de réalisation. Ainsi, nous citons ici ses fonctionnalités potentielles.

- Si l'utilisateur entre un mot qui ne fait pas partie de la liste de l'outil, le moteur de recherche affiche alors une page, pour le signaler à l'utilisateur.
- Si l'utilisateur entre un mot et lui assigne la fonction de terme noyau, le moteur de recherche affiche alors la liste des cooccurrents associés à ce terme. Ces termes noyaux posséderont chacun un lien vers la fiche d'information qui les concerne. Ils seront accompagnés d'une icône permettant à l'utilisateur d'accéder à l'équivalent, dans l'autre langue, de la page «Cooccurrents associés au terme noyau». Les cooccurrents, quant à eux, posséderont un lien vers la page d'information concernant la combinaison lexicale qu'ils forment avec le terme noyau.
- Si l'utilisateur entre un mot et lui assigne la fonction de cooccurrent, le moteur de recherche affichera alors la page «Liste des combinaisons lexicales pour le cooccurrent...».

#### 4.2. Module d'exploration

Nous présentons, dans les lignes suivantes, quelques-unes des fonctionnalités du module d'exploration. Ce mode de navigation offre à l'utilisateur le choix entre une navigation en français et une navigation en anglais. Lorsque l'utilisateur a choisi la langue dans laquelle il veut naviguer, le module d'exploration affiche une liste d'index :

- pour le français : termes noyaux, noms cooccurrents, verbes cooccurrents (sujets), verbes cooccurrents (objets), adjectifs cooccurrents,
- pour l'anglais : base nouns, collocate nouns, collocate verbs (sujets), collocate verbs (objets), collocate adjectives.

##### 4.2.1 Les index de cooccurrents ou collocates

Ces index sont composés des cooccurrents en entrée, classés par ordre alphabétique [Figure1].

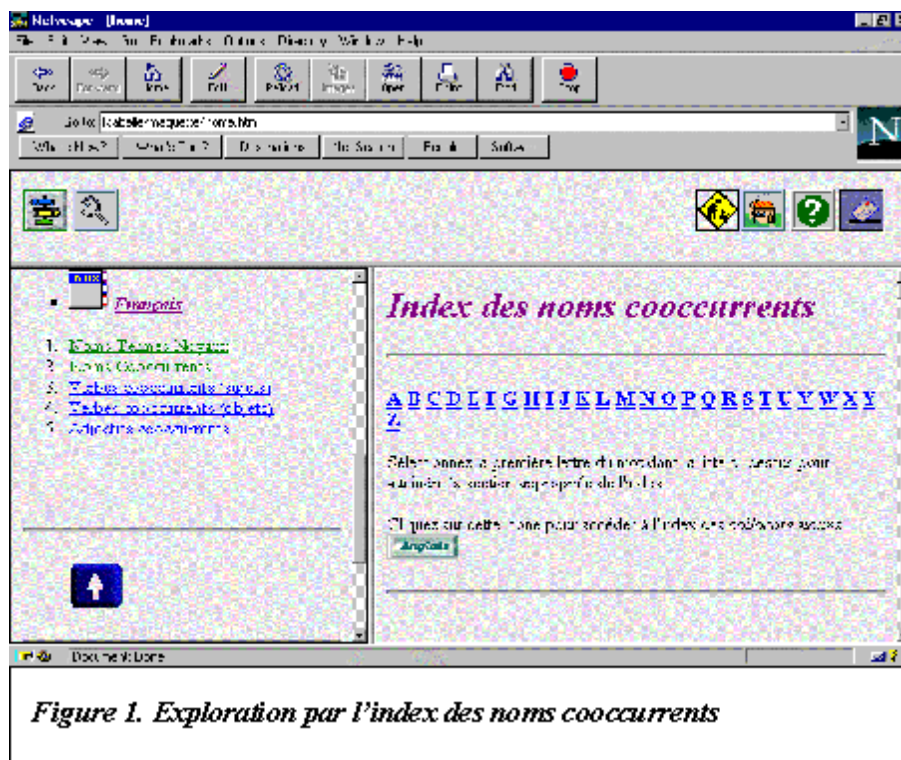


Figure 1. Exploration par l'index des noms cooccurrents

Chaque cooccurrent possède un lien vers une page affichant la liste des combinaisons lexicales qu'il forme avec les termes noyaux auxquels il s'associe. Dans chacune des combinaisons lexicales de cette liste, le terme noyau possède un lien vers la page d'information le concernant, et le cooccurrent possède un lien vers la page d'information concernant la combinaison lexicale.

À partir de la liste des combinaisons lexicales formées du cooccurrents et des termes noyaux, l'utilisateur a le choix d'avoir accès à :

- La page d'information sur la CLS, il lui suffit pour cela de cliquer sur le cooccurrent dans l'une des CLS ;
- La page d'information sur le terme noyau, il lui suffit pour cela de cliquer sur le terme noyau dans une des CLS ;
- La liste des CLS pour le cooccurrent équivalent dans l'autre langue, il lui suffit pour cela de cliquer sur l'icône indiquant la langue équivalente, dans la barre de menu ;
- La CLS équivalente dans l'autre langue, il lui suffit pour cela de cliquer sur l'icône indiquant l'autre langue et située sous la CLS.

#### 4.2.2 Les index des termes noyaux ou base nouns

Ces index sont composés des termes noyaux classés par ordre alphabétique [Figure 2]. Chacun est suivi d'une icône qui permet d'accéder directement à l'équivalent du terme dans l'autre langue. De plus, chaque terme possède un lien vers la page «Cooccurrents associés au terme noyau».

Le terme noyau possède un lien vers la page d'information qui lui est consacrée [Figure 3]. Par ailleurs, l'utilisateur choisit la fonction du cooccurrent selon la catégorie de combinaisons lexicales à laquelle il veut accéder. Chacune des fonctions possède un lien vers un signet à l'intérieur de la même page. Sous ce signet, sont alors affichés tous les cooccurrents correspondant à la fonction. Chaque cooccurrent possède également un lien vers la page d'information sur la combinaison lexicale qu'il forme lorsqu'il est associé au terme noyau.

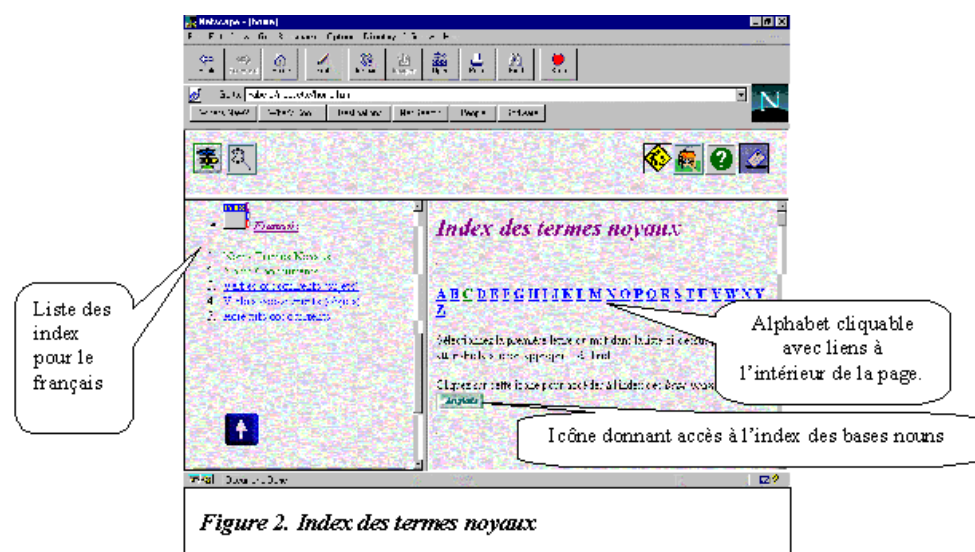
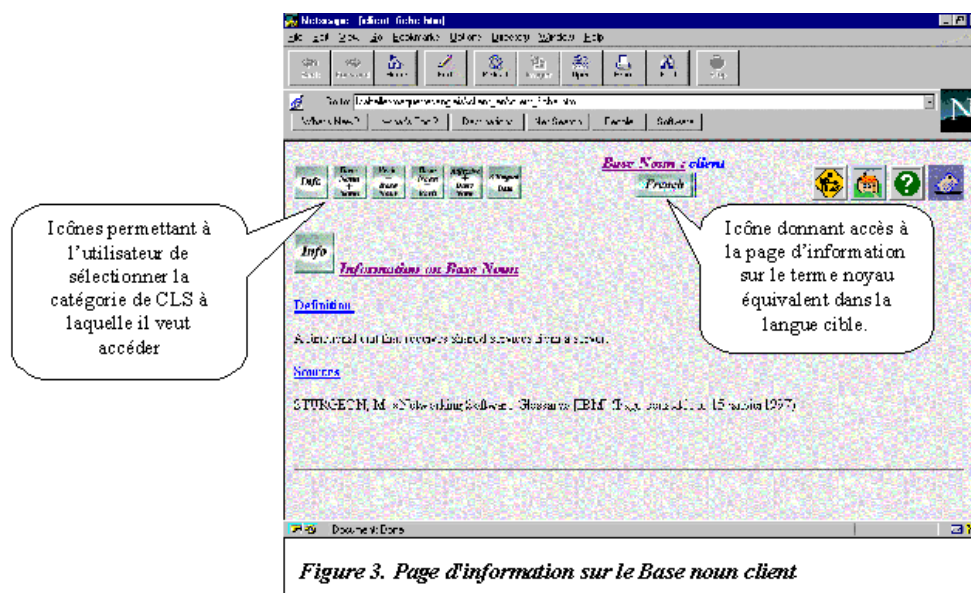


Figure 2. Index des termes noyaux



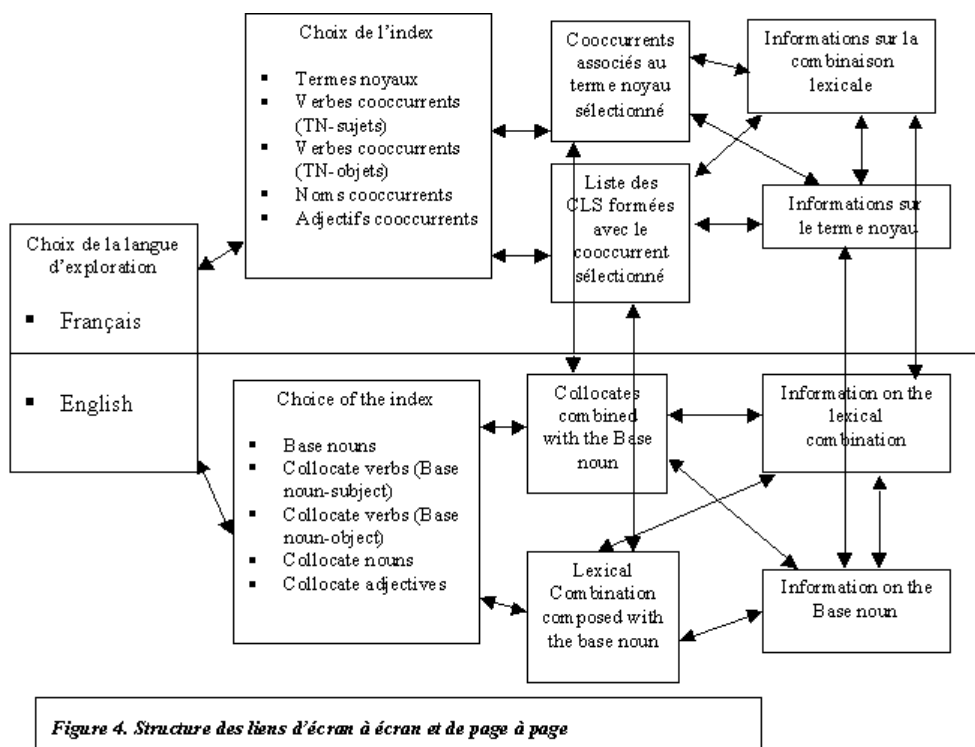
Dans cet outil, les barres de menu et donc les fonctions sont identiques dans les deux langues. Ainsi, l'utilisateur peut, à tout moment, avoir accès aux informations dans l'autre langue. Il peut également revenir à la page d'accueil ou directement commencer une nouvelle session d'exploration simplement en cliquant sur l'icône appropriée dans la barre de menu.

#### 4.3. Liens entre les pages

Nous avons établi trois catégories de liens:

- Les signets à l'intérieur des pages. Il s'agit des liens que nous avons créés pour naviguer d'une amarre dans une page à un texte dans la même page, par exemple dans la page «Cooccurents associés au terme noyau».
- Les liens de la barre de menu à une autre page. Il s'agit des liens que nous avons créés entre les icônes situées dans la barre de menu vers un texte dans une autre page du même écran, par exemple les liens reliant la barre de menu de la page d'information sur le terme noyau aux différentes catégories de combinaisons lexicales ou à la fiche de synthèse.
- Liens d'une page à une autre. Il s'agit des liens que nous avons créés pour naviguer d'une page d'un écran à une autre page d'un autre écran, par exemple les liens que possèdent les termes noyaux dans le premier écran vers la page d'information sur le terme noyau.

La plupart des liens ont été préétablis avec le modèle des pages, puisqu'ils agissaient sur les titres des sections et non sur le contenu. Cependant des liens ont dû être créés au fur et à mesure de la saisie du texte, ce sont les liens de contenu à contenu. La Figure 4 montre la structure des liens que nous avons créés d'écran à écran et de page à page.



## 5. Conclusion

Dans le présent article, nous avons mis en évidence la pertinence de notre approche étant donné l'absence, toujours d'actualité, d'ouvrages de consignation des CLS bilingues. Ainsi, nous avons étudié les quelques ouvrages qui concrétisent les approches des auteurs s'étant penchés sur le phénomène des combinaisons lexicales, et dont l'objectif principal est de fournir aux langagiers l'accès aux combinaisons lexicales correspondant à l'idée qu'ils cherchent à exprimer. De même, nous avons fait un tour d'horizon des innombrables lexiques consacrés au vocabulaire de l'autoroute de l'information, mais dont aucun ne présente les combinaisons lexicales en entrée.

Par la suite, nous avons présenté notre approche de consignation des combinaisons lexicales basée sur un outil programmé en langage HTML permettant la création de liens hypertextuels de manière à pallier les inévitables problèmes de redondance posés par la consignation sur format papier. En effet, un document édité en format HTML présente, par rapport aux autres types de supports, plusieurs avantages notables dont les principaux sont :

- La convivialité : ce critère dépend de l'aptitude de chaque concepteur de page Web, mais il est possible d'écrire des pages très faciles à consulter.
- Une interface attrayante : on peut décorer les pages que l'on crée de manière unique et personnalisée selon qu'on insère des images, qu'on change l'arrière plan de la page, que l'on utilise des polices de caractères originales, petites ou plus grandes, qu'on multiplie les couleurs dans le texte. Les combinaisons sont infinies et dépendent du goût de chacun.
- Accès à distance : Les documents HTML, sont accessibles à distance à partir d'un navigateur et ceci à travers le réseau mondial de l'Internet.
- Navigation non séquentielle et à plusieurs points d'entrée : on peut suivre différents liens pour accéder au même document et à partir d'un document on peut accéder à d'autres documents.

Par nos travaux, nous espérons pouvoir ouvrir la voie à de futurs projets d'ouvrages lexicographiques consacrés aux combinaisons lexicales spécialisées bilingues ou multilingues. En effet, notre outil doit pouvoir s'adapter à différentes langues et différents

domaines de spécialité. De même, il permet la consignation d'un volume «illimité» de données. Il possède une interface très conviviale et modulable au gré des concepteurs. De plus, il autorise des mises à jour rapides et simples.

De même, cet outil pourrait être remodelé pour répondre à différents besoins lexicographiques comme la conception d'un dictionnaire de synonymes des CLS [Meynard (b)]. Il pourrait également servir de support à l'organisation des CLS selon des critères conceptuelles en reprenant, par exemple, les fonctions lexicales telles qu'elles ont été définies par Mel'cuk *et al* [1984].

## RÉFÉRENCES DES SITES LEXICOGRAPHIQUES CONSULTÉS SUR INTERNET

- BUTTERFLY GLOSSARY (Page consultée le 23 avril 1997). [En ligne].  
Adresse URL : <http://www.rirr.cnuce.cnr.it/Glossario/ghpage.html>
- CIDIF (Page consultée le 21 avril 1997) : *Vocabulaire de l'inforoute*. [En ligne].  
Adresse URL : <http://w3.cidif.org/naviguer/vocabulaire.htm>
- CORMIER, R. (Page consultée le 21 avril 1997) : *Netglos – Français*. World Wide Language Institute. [En ligne].  
Adresse URL : <http://wwli.com/translation/netglos/glossary/french.html>
- DESCHAMP, P. (Page consultée le 23 avril 1997) : *Glossaire informatique des termes de la Commission ministérielle de terminologie informatique*. INRIA. [En ligne].  
Adresse URL : <http://www.rocq.inria.fr/qui/Philippe.Deschamp/CMTI/glossaire.html>
- ENGST, A. (Page consultée le 29 janvier 1997). Internet Starter Kit, [En ligne].  
Adresse URL : <http://www.mcp.com/hayden/iskm/iskw2/index.html>
- FREE ON-LINE DICTIONARY OF COMPUTING (Page consultée le 23 avril 1997). [En ligne].  
Adresse URL : <http://wombat.doc.ic.ac.uk/foldoc/index.html>
- INTERNET LITERACY CONSULTANT™. (Page consultée le 23 avril 1997) : *ILC. Glossary of Internet Terms*. [En ligne].  
Adresse URL : <http://www.matisse.net/files/glossary.html>
- KADOW, K. (Page consultée le 21 avril 1997) : *Kadow's Internet Dictionary*. [En ligne].  
Adresse URL : <http://msg.net/kadow/>
- UNISOFT (Page consultée le 23 avril 1997) : *Glossaire unisoft*. [En ligne].  
Adresse URL : <http://www.unisoft.fr/glossair.html>
- STURGEON M. [IBM] (Page consultée le 21 avril 1997) : *Networking Software Glossary*. [En ligne].  
Adresse URL : <http://www.networking.ibm.com/nsg/nsggls.htm>
- VECTOR COMMUNICATIONS. (Page consultée le 23 avril 1997) : *Le glossaire français d'Internet*. [En ligne].  
Adresse URL : <http://www.vector.ch/Vector/f/glossaire.html>

## BIBLIOGRAPHIE

- BÉJOINT, H. et Ph. Thoiron. (1992). «Macrostructure et microstructure dans un dictionnaire de collocations en langue de spécialité», Colloque international, Phraséologie et terminologie en traduction et en interprétation, *Terminologie et traduction*, 2/3, pp. 513-522.



- BENSON, M., E. Benson et R. Ilson. (1986). *The BBI Combinatory Dictionary of English. A guide to Word Combinations*, Amsterdam/Philadelphia, John Benjamins.
- CLAS, A. (1994). «Collocations et langues de spécialité», *Meta*, 39(4), pp. 576-580.
- COHEN, B. (1986). *Lexique des cooccurrents, Bourse et conjoncture économique*, Montréal, Linguattech.
- COHEN, B. (1992). «Méthodes de repérage et de classement des cooccurrents lexicaux», Colloque international, Phraséologie et terminologie en traduction et en interprétation, *Terminologie et traduction*, 2/3, pp. 505-511.
- GOUADEC, D. (1993). «Terminologie et phraséologie : principes et schémas de traitement», Colloque international, Phraséologie et terminologie en traduction et en interprétation, *Terminologie et traduction*, 2/3, pp. 549-563.
- HAUSMANN, F. J. (1979). «Un dictionnaire des collocations est-il possible ? », *Travaux de linguistique et de littérature*, XVII, 1, pp. 187-195.
- HEID, U. (1992). «Décrire les collocations : deux approches lexicographiques et leur application dans un outil informatisé», Colloque international, Phraséologie et terminologie en traduction et en interprétation, *Terminologie et traduction*, 2/3, pp. 523-548.
- HEID, U. (1994). "On Ways Words Work Together – Topics in Lexical Combinatorics", *Euralex '94 Proceedings*, Netherlands, pp. 227-257.
- HEID, U. et G. Freibott (1991). «Collocations dans une base de données terminologiques et lexicales», *Meta*, 36(1), pp.71-91.
- KOCOUREK, R. (1991). *La langue française de la technique et de la science. Vers une linguistique de la langue savante*, Wiesbaden, Oskar Brandstetter Verlag, 2e édition.
- LAINÉ, C. (1993). «Une approche terminologique de la phraséologie», *L'Actualité terminologique*, 26(2), pp. 14-16.
- LAINÉ, C., S. PAVEL et M. BOILEAU (1992). «La phraséologie – Nouvelle dimension de la recherche terminologique. Travaux du module canadien du Rint», *L'Actualité Terminologique*, 5(3), pp. 5-9.
- L'HOMME, M.-C. (1995). "Processing Word Combinations in Existing Term Banks", *Terminology* 2(1), pp. 141-162.
- L'HOMME, M.C. (1996). «Méthode d'accès informatisé aux combinaisons lexicales en langue technique». *Meta*, 42(1), pp. 15-23.
- MEL'CUK, I. A. et al (1984). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I*, Montréal, Presses de l'Université de Montréal.
- MEL'CUK, I. A., CLAS et A. POLGUÈRE (1995). *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve, Duculot/Aupelf-UREF.
- MEYNARD, I. (1997a). *Méthode de consignation dans un outil HTML des combinaisons lexicales spécialisées : étude basée sur 15 termes français et 15 termes anglais tirés du domaine de l'Internet*. Mémoire de maîtrise. Université de Montréal. Montréal, Canada.
- MEYNARD, I. (1997b). " Identifying a Web-based Approach concerned with a Lexical Combination Thesaurus ". À paraître dans *Proceedings of the 5th International Symposium of National Language Research Institute*, Tokyo, Japon, 27 et 28 août.
- MEYNARD, I. (1997c). «Combinaisons lexicales en langue de spécialité», Colloque Traduction et langues de spécialité, 65e Congrès de l'ACFAS, Trois-Rivières : Université de Trois-Rivières, 14 mai.
- MEYNARD, I. (1997d). «Quelques solutions aux problèmes de reconnaissance et d'équivalence des combinaisons lexicales à des fins de traduction». À paraître dans *Proceedings of the International Congress of Terminology*. Donostia-San Sebastian, Pays basque, 12 et 13 novembre.

- PAVEL, S. (1993a). «La phraséologie en langue de spécialité. Méthodologie de consignation dans les vocabulaires terminologiques», *Terminologies nouvelles*, 10, pp. 67-82.
- PICHT, H. (1987). "Terms and their LSP Environment – LSP Phraseology", *Meta*, 32(2), pp. 36-42.

\*\*\*\*\*

## LABURPENA / RESUMEN / RÉSUMÉ / ABSTRACT

### **HTML bidezko metodo hipertestuala lexiko-elkarketa berezituen bilduma elebiduna osatzeko**

Azken hamabost urteotan, arlo espezializatueta gertatzen diren lexiko-elkarketak gero eta sarriago izan dira aztergai ikerketa askotan. Baina, hala ere, lexikografiari buruzko lan gutxi dira —paperezko euskarrian (Cohen 1986, Lainé 1993), nahiz euskarri elektronikoa (Heid 1991, l'Homme 1995, 1997)— egileen ikuspegia azaltzen dutenak. Beste alde batetik, teknologia berriek, Internet-ek, esate baterako, bide eman dute hiztegi-gintza-alorreko saio ugari egiteko. Halaz guztiz ere, guk dakigula, behintzat, soil-soilik arlo espezializatueta lexiko-elkarketak biltzen dituen lanik ez da oraindik eratu, ezta bat bakarra ere.

Lan honetan, bada, gure proposamena azalduko dugu: lexikografia-tresna elebiduna (ingelese-frantsesa) sortu dugu, bere helburua itzulpengintza izanik. Bertan, arlo espezializatu bateko lexiko-elkarketak biltzen dira, informazioaren autobideari dagozkionak hain zuzen.

Tresna berri hau HTML eta Java lengoaietan programaturik dago Internet sarean zabalkundea izan dezan. Hipertestuaren erabilpena kontu berria dugu eta honi esker orain arte bestelako hiztegi-gintza-lanetan azaltzen ziren arazoak izango dute konponbiderik. Horieta azalduz batzuk aipatzearen, esan dezagun, batetik, paperezko euskarria duten lanek lekua aprobeztatu beharra izaten dutela, eta bestetik, kontuan hartuta zenbat eta zenbat sarrera eta baliokide duzun eskuartearen, euskarri hori ez dela erraza erabiltzen. Euskarri informatikoa dutenetan, berriz, datu-basea finkatzea eta eratzeak izaten dituen zailtasunak.

Hasteko, era laburrean emanda, arlo espezializatueta lexiko-elkarketak biltzen dituen hiztegi-gintza-lanen nondik-norakoak aztertuko ditugu. Ondoren, modu zehatzagoan, gure tresnaren berri emango dugu. Bere egitura hipertestuaren bidez lotutako HTML orrialdeek osatzen dute. Orrialde batek termino-muinari buruzko informazioa ematen digu eta, besteak, berriz, lexiko-elkarketei buruzko informazioa. Horrez gain, tresna erabiltzeko modu bi aurkeztuko ditugu. 1) Termino-muinen edo inguruko osagaien zerrenda (izenondoak, aditzak, izenak) abiapuntu harturik, bila ibiltzea. 2) Bilatzeko motore baten laguntzaz, lexiko-elkarketak gako-hitzen bitartez bilatzea.

Bukatzeko, gure proposamenak aurrerantzean eskainiko dituen abantailen gainean mintzatu gara, batez ere, arlo espezializatueta lexiko-elkarketen bilduma elebidunen bilaketari dagokionez.

### **Aproximación hipertextual vía HTML a un instrumento de recopilación bilingüe de combinaciones léxicas especializadas**

Durante los últimos quince años, el fenómeno de las combinaciones léxicas en lengua de especialidad ha sido objeto de un número creciente de estudios. No obstante, son pocos los trabajos lexicográficos —bien se trate de trabajos en soporte de papel (Cohen 1986, Lainé 1993) o en soporte electrónico (Heid 1991, l'Homme 1995, 1997)— que plasman los enfoques de los autores. Por otra parte, la aparición de nuevas tecnologías, tal como Internet, ha propiciado numerosos ensayos de instrumentos lexicográficos, pero, que nosotros sepamos, no existe en la actualidad uno sólo dedicado exclusivamente a la recopilación de combinaciones léxicas en lengua de especialidad.

En este artículo presentamos, pues, nuestro enfoque que se plasma en la realización —con fines de traducción—, de un instrumento lexicográfico bilingüe (inglés-francés) de recopilación de las combinaciones léxicas de un campo de especialidad, en este caso el campo de la autopista de la información.

Este instrumento ha sido programado en lenguaje HTML y Java para la publicación por Internet. La estructura hipertextual de este instrumento le da originalidad ya que permite resolver numerosas dificultades que plantea la recopilación de combinaciones léxicas en otros formatos de obras lexicográficas. Citemos para ilustrar dichas dificultades, la atención que necesariamente se presta a los problemas de economía de espacio en las obras de soporte de papel o la incomodidad de uso que genera la redundancia de entradas y de equivalencias entre una lengua y otra, así como la complejidad de la concepción de las bases de datos para las obras informatizadas.

En la primera fase plantearé resumidamente el estado de la cuestión de las obras lexicográficas de recopilación de las combinaciones léxicas en lengua de especialidad. A continuación expondremos de manera detallada la estructura que, basada en la noción de páginas HTML interrelacionadas por los lazos hipertexto, incluye una página de información sobre el término núcleo y otra sobre combinaciones léxicas. Expondremos también el funcionamiento del instrumento que se articula principalmente alrededor de dos maneras de navegación: 1) exploración a partir del índice de términos núcleos o de concurrentes (adjetivos, verbos o nombres), 2) búsqueda por palabras claves gracias a un motor de búsqueda.

Para concluir, discutiremos las perspectivas que nuestro enfoque ofrece para las futuras búsquedas en materia de recopilación bilingüe de las combinaciones léxicas en lengua de especialidad.

### **Approche hypertextuelle via HTML pour un outil de consignation bilingue des combinaisons lexicales spécialisées**

Au cours des quinze dernières années, plusieurs études ont été consacrées aux combinaisons lexicales en langue de spécialité. Pourtant, peu d'ouvrages lexicographiques —qu'il s'agisse d'ouvrages sur support papier [Cohen 1986, Lainé 1993] ou d'ouvrages sur support électronique [Heid & Freibott 1991, L'Homme 1995, 1997]— concrétisent les approches des auteurs. Par ailleurs, l'émergence de nouvelles technologies en particulier l'Internet a ouvert la voie à de nombreux essais d'outils lexicographiques, mais, à notre connaissance, aucun de ces outils n'est à l'heure actuelle exclusivement consacré à la consignation des combinaisons lexicales en langue de spécialité.

Dans cet article, nous présentons une nouvelle approche, qui s'est concrétisée par la réalisation, aux fins de traduction, d'un outil lexicographique bilingue (anglais-français) de consignation des combinaisons lexicales d'un domaine de spécialité —en l'occurrence, le domaine de l'autoroute de l'information. Cet outil a été développé en utilisant les langages

d'édition HTML et de programmation Java. La structure hypertextuelle de cet outil en fait son originalité puisqu'elle permet de résoudre de nombreuses difficultés posées par la consignation des combinaisons lexicales dans les autres formats d'ouvrages lexicographiques. Citons pour illustrer ces difficultés, l'attention nécessairement apportée, dans les ouvrages sur support papier, aux problèmes d'économie de place, au manque de convivialité qu'entraîne la redondance des entrées et des équivalences d'une langue à l'autre, ainsi que la complexité de conception des bases de données pour les ouvrages informatisés.

Dans un premier temps, nous ferons un bref état de la question des ouvrages lexicographiques de consignation des combinaisons lexicales en langue de spécialité. Par la suite, nous exposerons, de manière détaillée, la structure qui, basée sur la notion de pages HTML interreliées par des liens hypertexte, comporte une page de renseignements sur le terme noyau et une page de renseignements sur les combinaisons lexicales. Ainsi que le fonctionnement de l'outil qui offre principalement deux modes de navigation : 1) exploration à partir d'index des termes noyaux ou de cooccurrents (adjectifs, verbes ou noms), 2) recherche par mots clés grâce à un moteur de recherche.

En conclusion, nous discuterons des perspectives de recherche que notre approche ouvre en matière de consignation bilingue des combinaisons lexicales en langue de spécialité.

### **Hypertext approach via HTML for a bilingual compilation tool for specialists lexical combinations**

Over the past fifteen years a growing number of studies have been completed on lexical combinations in specialists language. Nevertheless, only a few of these lexicographical works —either in paper (Cohen 1986, Lainé 1993) or in electronic medium (Heid 1991, L'Homme 1995, 1997)— set forth the author's approach. Furthermore, the emergence of new technologies such as Internet has opened up the way to many testing on new lexicographical tools, but, as far as we know, none of this tools is exclusively devoted to compile lexical combinations in specialists language.

In this paper, we will explain our approach which has been materialized in a bilingual (English-French) lexicographical tool for translators' use conceived to compile lexical combinations within a given specialist domain —in this case the subject field of the information highway.

The mentioned tool has been programmed in HTML and Java languages so as to be published on Internet. The novelty brought about by the hypertext structure of this tool makes its originality as it enables us to solve many of the problems coming up from the compilation of lexical combinations in lexicographical works published in other formats. Among the difficulties arisen when using lexicographical works having other formats we can mention, for instance, that in works published in paper the need to economize space and the necessity of redundancy of entries and equivalents in two languages makes these works user-unfriendly. As for computerized works, their main challenge lies in the complexity to lay out their database.

To begin with, we will give an account of the existing lexicographical compilations of lexical combinations in specialists language. Then, we will thoroughly explain the structure of our tool which is based on the same concept as HTML pages interrelated by hypertext links and includes a page on the kernel term and another page which displays information about lexical combinations. We will also show how our tool works and the two different navigation modes that can be used to operate it: 1) the user can either start exploring from the index of kernel terms and co-occurrents (adjectives, verbs or nouns) 2) or she/he can set about searching key words through the searching device.

To sum up, we will discuss the prospects opened up by our approach for the forthcoming research on bilingual compilation of lexical combinations in specialists language.