

QUELQUES SOLUTIONS AUX PROBLÈMES DE RECONNAISSANCE ET D'ÉQUIVALENCE DES COMBINAISONS LEXICALES À DES FINS DE TRADUCTION

Isabelle MEYNARD

Département de Linguistique et de Traduction
Université de Montréal

INTRODUCTION

Malgré les nombreuses études qui lui ont été consacrées, le phénomène des combinaisons de mots ne fait toujours pas l'objet d'un consensus. En effet, il existe toujours un flou terminologique quant à la désignation de ces groupes : *collocations*, *phraséologismes*, *combinaisons lexicales*.

En outre, les combinaisons de mots posent généralement des problèmes de traduction et de consignation dans les ouvrages lexicographiques. Reprenant les constatations que nous avons faites lors de nos recherches [Meynard : 1997a], nous nous proposons dans cet article de présenter notre approche pour le recensement des combinaisons lexicales spécialisées.

Dans un premier temps, nous nous employons à définir ce que sont ce que nous appelons les combinaisons lexicales. Par la suite, nous identifions les difficultés que peut rencontrer un langagier lors de leur extraction et de l'établissement de leurs équivalences entre langues.

1. Définition des combinaisons lexicales

Nous définissons les combinaisons lexicales comme des groupes spécifiques à un domaine de spécialité. Ces groupes ont la particularité d'être des associations de mots consacrés par l'usage, d'où l'intérêt de les recenser afin de donner aux langagiers (traducteurs, rédacteurs techniques, terminologues, etc.) la possibilité de se familiariser avec les usages spécifiques à un domaine de spécialité.

Les unités composant les combinaisons lexicales peuvent acquérir un nouveau sens dans un domaine de spécialité donné, et ce, même si elles sont associées à un autre domaine de spécialité, voire à la langue générale. Par exemple, *lancer_client* : *lancer* appartient à la langue générale et *client* appartient à plusieurs domaines de spécialité (informatique, commerce). Ainsi, dans le domaine de l'Internet, *lancer* signifie *mettre en marche* et un *client* est une "composante d'un modèle client-serveur à partir de laquelle l'utilisateur lance une application qui sera traitée en partie dans une autre composante du réseau informatique, le serveur" [Gavrilit *et al* 1996 : 169].

Par ailleurs, le sens des combinaisons lexicales peut être directement déductible du sens de chacune des unités qui les composent, par exemple : *envoyer_fichier*.

Ainsi, les combinaisons lexicales ne véhiculent pas une signification unique, ce qui les distingue des termes complexes qui représentent une seule notion terminologique. Cependant, il ne s'agit pas pour autant de formes libres puisque les éléments qui les composent sont liés par une attirance consacrée par l'usage. Par exemple, *développer un lien* ou *générer un lien* sont des formes libres que l'on ne retrouve pas dans les habitudes langagières des spécialistes du domaine de l'autoroute de l'information, qui emploient plutôt *créer un lien*.

Les combinaisons lexicales sont composées d'un terme autour duquel s'articulent d'autres mots. Puisqu'il est au centre du groupe, nous retiendrons la désignation *terme noyau* pour identifier le terme et *cooccurrent* pour identifier les mots qui lui sont associés.

En terminologie, l'idée est que tout gravite autour du *terme-nom* : cette approche nous a mené à retenir quatre catégories de combinaisons.

- Nom + terme noyau, par exemple : *format de fichier*;
- Terme noyau + verbe, par exemple : *a software runs on*;
- Verbe + terme noyau, par exemple : *lancer un client*;
- Terme noyau + adjectif, par exemple : *current document*.

Ces catégories sont génériques, c'est-à-dire qu'elles regroupent des sous-catégories de combinaisons lexicales.

1.1. Ainsi, la catégorie terme noyau + nom comprend

Pour le français, nom + préposition + terme noyau, par exemple : *compression de données*, et, pour l'anglais, terme noyau + nom, par exemple : *page creation*, et nom + préposition + terme noyau, par exemple : *creation of a page*.

1.2. La catégorie verbe + terme noyau comprend

Pour le français, verbe + terme noyau, par exemple *implémenter un logiciel*, et verbe + préposition + terme noyau, par exemple : *naviguer sur le réseau*. Pour l'anglais, verbe + terme noyau, par exemple : *to initiate a session*, et verbe à particule + terme noyau, par exemple : *to log on a session*.

Pour cette dernière sous-catégorie, il peut parfois être difficile de déterminer s'il s'agit d'une particule ou d'une préposition. On peut donc appliquer une méthode qui consiste à essayer de mettre la sous-catégorie au passif. Ainsi, la forme passive *a session is logged on* est grammaticalement correcte, ce qui confirme que l'on a un verbe à particule et non un verbe à préposition.

Il arrive aussi que dans les ouvrages lexicographiques les verbes à particule fassent l'objet d'une entrée distincte du verbe initial, ce qui dissipe tout doute quant à savoir s'il s'agit d'une particule ou d'une préposition. Par exemple, dans le Collin, S.M.H et al [1991 : 139] on trouve *log on* en entrée avec une note : *the verbs can be spelled log on, log-on, or logon*.

1.3. La catégorie terme noyau + verbe comprend

Pour les deux langues, terme noyau + verbe, par exemple : *a page generates* (pour l'anglais) et *un serveur exécute* (pour le français). Également, toujours pour les deux langues, terme noyau + verbe + préposition, par exemple : *un client dialogue avec* ou *a software runs on*.

1.4. La catégorie terme noyau + adjectif comprend

Pour le français, terme noyau + adjectif, par exemple: *requête entrante*; et, pour l'anglais, adjectif + terme noyau, par exemple: *permanent address*.

2. Méthode de recensement des cooccurrents

2.1. Extraction automatique

À cette étape, l'informatique apporte une aide précieuse. En effet, de nombreux logiciels existent désormais sur le marché, qui permettent :

- De repérer facilement les combinaisons lexicales, en offrant à l'utilisateur la possibilité de spécifier plusieurs mots dans une même requête. Par exemple, *display 5w data* permet de chercher la combinaison lexicale *to display data*; *5w* est un opérateur booléen qui permet d'indiquer au logiciel la distance maximum (ici, 5 mots) que l'on accepte entre les éléments de la requête.
- De re-situer les combinaisons lexicales à l'intérieur du corpus par l'accès que donne le logiciel aux références des textes sources.
- D'obtenir des données statistiques sur les occurrences des combinaisons lexicales: informations importantes pour déterminer la fréquence d'emploi d'une combinaison à l'intérieur d'un corpus.

2.2. Fréquence

Le critère de fréquence d'une combinaison lexicale est fonction de l'emploi de cette combinaison au sein d'un corpus. Cet emploi doit exprimer l'usage de la langue de spécialité.

Pour notre étude, nous n'avons cependant pas fixé de seuil, car le domaine étant nouveau, on y trouve un foisonnement de néologismes dont on ne peut encore déterminer lesquels entreront dans l'usage de ce domaine de spécialité. Par exemple, nous avons retenu *to establish a link* qui n'apparaît pourtant que trois fois dans le Carroll, J. et Broadhead, R. [1995] et jamais dans les autres documents. La raison de ce choix est que ce document particulier semble être une référence dans le domaine de l'Internet, en effet, les publications du domaine mentionnent souvent des citations qui en sont tirées.

3. Problèmes de reconnaissance des combinaisons lexicales

Bien souvent, lorsqu'il s'agit de combinaisons composées du terme noyau et d'un autre nom ou d'un adjectif, le langagier peut être confronté au problème de distinction entre CLS et termes complexes, par exemple : *serveur_distant* (CLS) ou *page d'accueil* (terme complexe).

Ainsi, nous définissons un terme complexe comme étant une association de mots qui représente une notion terminologique unique. Cette association de mots ne peut être séparée par un ou plusieurs autres éléments (articles, adjectifs, verbes, etc.).

En français, ces groupes sont habituellement composés d'un terme noyau, d'une préposition et d'un nom. Pour l'anglais, en revanche, il s'agit souvent d'une seule entité graphique, par exemple : *homepage* ou *database*.

Ainsi, on conservera *serveur_distant* même si, *a priori*, selon la fréquence dans les corpus, ces deux mots semblent toujours liés de manière immuable. En effet, il est aussi possible de dire *serveur_http_distant*, sans que cela soit une erreur de langage.

Pour connaître les cas d'application de cette règle, les textes du corpus ne sont pas toujours suffisants et le langagier doit donc faire appel aux spécialistes du domaine.

4. Équivalences des combinaisons d'une langue à l'autre

La dernière étape de notre recherche a consisté à établir les équivalences des combinaisons lexicales d'une langue à l'autre.

Dans un premier temps, nous avons comparé les combinaisons lexicales anglaises et françaises pour chaque terme noyau. Par exemple, nous avons comparé les éléments de la catégorie *terme noyau + adjectif* pour le terme noyau français *logiciel* aux éléments de la catégorie *adjectif + terme noyau* pour le terme noyau anglais *software*.

Lorsque cette première démarche ne nous a pas permis de trouver toutes les équivalences, nous nous sommes fiée à notre connaissance du domaine et donc à notre intuition. Par exemple, nous avons la combinaison lexicale anglaise *to run a client*. *A priori*, cette combinaison pouvait se traduire en français par la combinaison lexicale *lancer un client*. Or nous avons déjà trouvé *to launch a client* comme équivalent de cette combinaison. Nous avons donc essayé de déterminer si *to launch a client* et *to run a client* étaient synonymes. En fait, nous savions que *to launch* implique un point de départ à une action, alors que *to run* sous-entend que l'action est en train de se dérouler; nous avons donc choisi la combinaison lexicale *exécuter un client* que nous avons déjà rencontrée dans des lectures antérieures (ne faisant pas partie du corpus) et qui exprime bien ce sens.

Dans un troisième temps, lorsque nous n'avions toujours pas toutes les équivalences, nous avons consulté les dictionnaires. Ainsi, pour la combinaison lexicale, *to flag a document*, nous avons trouvé dans le dictionnaire bilingue [Collin, S.M.H et al 1991], la traduction du verbe *to flag* soit : *marquer, baliser*. Ceci nous a permis de faire le lien avec le langage HTML qui permet d'agir sur le formatage d'un texte en le **balisant**. Ainsi, nous avons retenu la combinaison lexicale *baliser un document* plutôt que *marquer* comme équivalent de *to flag a document*, parce que nous l'avions déjà rencontrée lors de lectures antérieures.

Finalement, si après ces trois premières démarches, il manquait encore des équivalences, nous avons fait appel aux spécialistes. Ainsi, aucune des trois solutions mentionnées précédemment ne nous a permis de trouver l'équivalence de la combinaison *to resolve an address*. Nous avons donc interrogé des spécialistes du domaine de l'informatique, qui nous ont, sans hésitation, suggéré *convertir une adresse*.

CONCLUSION

Dans cet article, nous nous sommes employée à expliciter notre approche pour la reconnaissance et l'établissement d'équivalences de combinaisons lexicales spécialisées bilingues. Ainsi, nous avons tenté de faire ressortir l'importance de la constitution d'un corpus représentatif du domaine de spécialité objet de l'étude.

De même nous avons essayé d'identifier les concepts clés autour desquels s'articule notre travail d'élaboration de lexiques de combinaisons lexicales bilingues : *combinaisons lexicales, termes complexes, formes libres*.

Enfin, nous basant sur notre étude du vocabulaire de l'Internet, nous avons fait un tour d'horizon des difficultés qu'impliquent l'extraction, la traduction et la validation des combinaisons lexicales spécialisées.

Nous croyons que cette méthode sera valable pour d'autres domaines de spécialité et pourra servir de base à l'élaboration d'ouvrages de références mettant en évidence le comportement des termes en contexte.

RÉFÉRENCES DU CORPUS

- BERGERON, M. et C. KEMPA, (1995). *Vocabulaire d'Internet: vocabulaire anglais – français*. Montréal, Office de la Langue française, 62 p.
- CARROLL, J., BROADHEAD, R. (1995). *Canadian Internet Handbook*, 1995 Prentice-Hall, Edition Scarborough, 798 p.
- COLLIN, S.M.H *et al.* (1991). *Larousse informatique Dictionnaire/Dictionary*. Éditions Peter Collin Publishing & Larousse, Tours, France, 323 p.
- CONDAT, J-B., PIOCH, N. (1994). *Internet*. Montréal, Éditions J.C.I., 224 p.
- GAVRILUT, G., M. LETRANCHANT, N. ST-JACQUES, S. TELLIER. (1996). *Internet: les aides à la recherche*. Montréal. Éditions du Trécarré, 183 p.
- LALONDE, J., GOYER, N., LAURENDEAU, A. (1995). *Internet au bout des doigts*. Montréal, Éditions du Trécarré, 325 p.
- PÉLIKIS, G. (1995). *Le World Wide Web: Création de serveurs sur Internet*. Paris, Éditions Addison-Wesley France, 226 p.
- PIGOT, T. (1995). *Créez vos pages Web sur Internet*. Paris, Éditions Sybex, 207 p.
- SOHIER, D. J. (1994). *Internet: guide de survie de l'internaute: [fonctions de base]*. Montréal, Éditions Logiques, 181 p.
- SOHIER, D. J. (1995). *Internet: guide d'exploration de l'internaute: [fonctions avancées]*. Montréal, Éditions Logiques, 232 p.

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, Linguatex.
- 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. (1992). "Les unités phraséologiques verbales et leur représentation en terminographie", Colloque international, Phraséologie et terminologie en traduction et en interprétation, *Terminologie et traduction*, 2/3, pp. 493-503.
- 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). "Approche hypertextuelle via HTML pour un outil de consignation bilingue des combinaisons lexicales spécialisées". 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.
- PAVEL, S. (1993b). "Vers une méthode de recherche phraséologique en langue de spécialité", *L'Actualité terminologique*, 26(2), pp. 9-13.
- PICHT, H. (1987). "Terms and their LSP Environment – LSP Phraseology", *Meta*, 32(2), pp. 36-42.
- ROBERTS, R. P. (1993). "La phraséologie: état des connaissances", *Terminologies nouvelles*, 10, pp. 36-42.

SCHAETZEN, C. de (1993). "Un accès rapide aux collocations", *Phraséologie, Actes du séminaire international, Terminologies Nouvelles*, 10, pp. 103-114.

LABURPENA / RESUMEN / RÉSUMÉ / ABSTRACT

Itzulpen-lanetan lexiko-elkarketak identifikatu eta baliokideez ordezkatzeraoan sortzen diren arazoei aurre egiteko zenbait konponbide

Arlo espezializatueta sortzen diren lexiko-elkarketak asko ikertu dira azken aldi honetan, batez ere hamabost urtetik hona, eta arazo larri bat dagoela geratu da agerian. Kontua da itzultzaileek, terminoen baliokideak ezagutzeaz gain, arlo espezializatueta bakoitzean erabiltzen den hizkera ere ezagutu beharko dutela aurrerantzean. Horrenbestez, itzulpengintza helburu dugula, adostasunezko metodologia finkatu beharko da hizkera espezializatueta lexiko-elkarketak definitzeko, ezagutzeko (espezialitate-arloa edozein dela ere) eta zenbat diren jakiteko.

Informazioaren autobideko hiztegia oinarritzat hartuta, lexiko-elkarketa espezializatuak sailkatzerakoan eta xede hizkuntzan eurentzat baliokide egokia aurkitzerakoan sortzen diren arazoak konpondu nahi ditugu artikulu honetan.

Lehendabizi, lexiko-elkarketa espezializatuak zer diren zehaztuko dugu (multzo horiek zertan diren eta nola eratzen diren: termino-muina + aditza + izenondoa, etab.) Ondoren, informazioaren autobidetik zenbait adibide hartu eta itzulpen elebidunetan (ingelesa- frantsesa) lexiko-elkarketek dakartzaten arazoak ikusiko ditugu.

Algunas soluciones para los problemas de reconocimiento y equivalencia de las combinaciones léxicas en la traducción

El estudio del fenómeno de las combinaciones léxicas en lengua de especialidad, que se ha intensificado durante los últimos quince años, ha puesto en evidencia un problema importante. En efecto, los traductores, no solamente necesitan conocer las equivalencias de los términos, sino que en adelante deberán también poder acceder a los hábitos de las lenguas de especialidad. Ello trae consigo el establecimiento de una metodología consensual para la definición de las combinaciones léxicas en lengua de especialidad, para su reconocimiento, sean cuales fueren los campos de especialidad, y para el recuento de las equivalencias con la finalidad de la traducción.

Basándonos en el vocabulario de la autopista de la información, nos proponemos, en este artículo, aportar elementos de respuesta a los problemas de reconocimiento y equivalencia en lengua de destino de las combinaciones léxicas especializadas.

Dedicaremos la primera parte a definir las combinaciones léxicas especializadas (qué son estos grupos y qué formas pueden tomar: término núcleo + verbo + adjetivo, etc). En la segunda parte ilustraremos con ejemplos cogidos del campo de la autopista de la información, las dificultades de la traducción bilingüe (inglés-francés) que implican la combinaciones léxicas especializadas.

Quelques solutions aux problèmes de reconnaissance et d'équivalence des combinaisons lexicales à des fins de traduction

L'étude du phénomène des combinaisons lexicales en langue de spécialité, qui s'est intensifiée ces quinze dernières années, a mis en évidence un problème important. En effet, les traducteurs n'ont désormais plus seulement besoin de connaître les équivalences des termes, ils doivent aussi pouvoir avoir accès aux coutumes langagières des spécialistes. Ceci implique l'établissement d'une méthodologie consensuelle pour la définition des combinaisons lexicales en langue de spécialité, leur reconnaissance quels que soient les domaines de spécialité, et pour le recensement des équivalences aux fins de traduction.

Nous basant sur le vocabulaire de l'autoroute de l'information, nous nous proposons donc, dans cet article, d'apporter des éléments de réponses aux problèmes de reconnaissance et d'équivalence en langue cible des combinaisons lexicales spécialisées.

Dans une première partie, nous nous attacherons à définir les combinaisons lexicales spécialisées (ce que sont ces groupes et quelles formes ils peuvent revêtir : terme noyau + verbe, + adjectif, etc.). Dans une deuxième partie, nous illustrerons, par des exemples empruntés au domaine de l'autoroute de l'information, les difficultés de traduction bilingue (anglais-français) qu'impliquent les combinaisons lexicales spécialisées.

Some answers to the issues of recognition and equivalence of lexical combinations for translation purposes

The increase in research studies on the phenomenon of lexical combinations in specialists language over the past fifteen years has put in a prominent position an important issue. The fact is that nowadays translators not only need to know equivalent terms but they also need to have access to the specialists' language usage. This implies that a consensual methodology has to be set up in order to define what lexical combinations in specialists language are, to recognize them within whatever specialist domain and to make the inventory of equivalents for translation purposes.

On the basis of the lexicon of terms on the information highway, we will put up in this paper to the answers we propose to the issue of recognition and equivalence in target language of lexical combinations in specialists language.

In the first part of the paper we will undertake to define lexical combinations in specialists language (what those term groups are and which forms they can take on: core term + verb + adjective, etc). In the second part of the paper, we will illustrate with examples taken from the domain of information highway the difficulties that come across in bilingual translations (English-French) implying specialists lexical combinations.