

**EUROPA, le serveur interinstitutionnel de l'Union européenne.**  
**Présentation des éléments utiles pour le traducteur et recherche d'information.**  
*Jean-Bernard Quicheron, Commission européenne*  
*Coordinateur d'EUROPPlus, intranet de la Commission*  
*Aix-en-Provence, 30 octobre 1999*

## **Introduction**

Dans ce temps qui m'est généreusement imparti, j'aimerais vous faire partager quelques-unes des ressources électroniques auxquelles vous avez accès par le biais d'Internet quasi gratuitement. Je dis bien quasi gratuitement, car il est clair que tant que les tarifs téléphoniques sont ce qu'ils sont en Europe, se connecter au réseau Internet peut être une occupation onéreuse même si l'on consulte des serveurs ou des bases de données gratuits.

Etant donné que je fais partie depuis au moins 3 ans de l'équipe qui gère EUROPA, le serveur interinstitutionnel de l'Union européenne, j'ai pensé qu'il serait judicieux de vous en présenter quelques aspects qui sont ou peuvent être d'intérêt pour les traducteurs et les interprètes.

Nous passerons en revue divers aspects tels que :

1. EUR-LEX, le journal officiel en ligne disponible dans les 11 langues communautaires
2. CELEX, base du domaine juridique disponible en ligne dans les 11 langues communautaires
3. SCADplus, base de données assez ignorée du grand public, mais qui peut être une source d'informations explicatives sur les Communautés européennes
4. EURODICAUTOM, base de données terminologiques de la Commission, disponible en ligne dans les 11 langues communautaires.
5. la recherche sur EUROPA. Nous regarderons comment il est possible d'améliorer ses performances de recherche de l'information sur un serveur comme EUROPA grâce à une meilleure maîtrise des subtilités du moteur de recherche Verity Search 97.

Nous aurions également pu regarder Systran, système de traduction automatique, mais ceci nous aurait menés probablement trop loin.

## **Internet et intranet**

Il n'est pas toujours facile de faire comprendre à l'homme de la rue et aux spécialistes de la traduction et de l'interprétation que vous êtes, les subtiles différences entre les serveurs Internet et intranet de l'Union européenne. Pourtant, je voudrais m'y attarder quelques instants, car il se peut que parfois je me réfère à l'intranet de la Commission, EUROPPlus, dont j'ai la charge.

Pour éviter tout malentendu, précisons que la Commission gère actuellement trois types de serveurs :

- **EUROPA**, qui est un serveur Internet interinstitutionnel. Ceci signifie que cette gigantesque collection de documents (1.200.000) est alimentée par toutes les institutions européennes et qu'elle est accessible gratuitement dans le monde entier par le réseau Internet. Son adresse est bien connue : <http://europa.eu.int>. En fait, chaque institution alimente son propre serveur mais les institutions partagent toutes une adresse commune.

- **EUROPPlus** est l'intranet de la Commission. Il n'est pas accessible de l'extérieur mais est consulté par tous les fonctionnaires et agents de la Commission, à Bruxelles, à Luxembourg, dans 23 villes de l'Union où la Commission dispose d'un bureau de représentation et dans plus de 100 pays du monde où la Commission dispose d'une délégation. Cela fait finalement beaucoup de monde pour un intranet.
- **EUROPATeam** est l'intranet interinstitutionnel, non accessible au grand public. Etant donné que la Commission ne souhaite pas que toutes les informations disponibles sur son intranet puissent être consultées par les fonctionnaires des autres institutions européennes, elle a créé un intranet interinstitutionnel, qui reprend bien des éléments d'EUROPPlus qu'elle est disposée à partager avec les autres institutions.

## 1. EUR-Lex

Fort heureusement, comme EUR-Lex est un produit d'information très connu des amateurs d'EUROPA, il est assez facile à trouver, il est placé sur la page d'information que l'on atteint à partir de la seconde page d'EUROPA (<http://europa.eu.int/geninfo/info-fr.htm>). La rubrique s'intitule : EUR-Lex, le droit de l'Union européenne. Vous y trouverez :

### 1. le journal officiel (publié quotidiennement dans les 11 langues)

- série C (informations, communications et actes législatifs préparatoires),
- série L (législation), un supplément et une annexe. Ces séries sont accessibles gratuitement pendant *une période glissante de 45 jours* suivant la publication. Après quoi vous les trouverez sous EUDOR en format payant. Certains particuliers ou firmes font de la veille technologique, en d'autres termes chaque jour ils vont voir les nouveautés sur le serveur et archivent gratuitement les documents qui les intéressent.

### 2. les traités :

Versions consolidées dans lesquelles ont été intégrées les modifications apportées par le traité d'Amsterdam, signé le 2 octobre 1997

- Traité sur l'Union européenne (format PDF), signé à Maastricht le 7 février 1992
  - Journal officiel C 340 10.11.1997 p. 145-172
- Traité instituant la Communauté européenne (format PDF), signé à Rome le 25 mars 1957
  - Journal officiel C 340 10.11.1997 p. 173-308

### 3. Législation:

- Législation communautaire en vigueur. Le répertoire de la législation communautaire en vigueur comprend:
  - le droit issu des relations extérieures entretenues par les Communautés européennes
  - le droit communautaire dérivé
  - le droit dérivé complémentaire (décisions des représentants des Etats membres, réunis au sein du Conseil, conventions internationales conclues entre les Etats membres en application des dispositions des traités, etc.)
- Textes consolidés
- Taux de Conversion de l'Euro (L 359, 31.12.1998)

### 5. Actes préparatoires communautaires

*Aix-en-Provence, samedi 30 octobre 1999. EUROPA, le serveur interinstitutionnel de l'Union européenne.*

Vaste catégorie de documents. EUR-Lex couvre la plupart de ces documents mais se limite à ceux pour lesquels la procédure législative n'a pas abouti et qui n'ont pas été retirés.

## 6. *Jurisprudence*

Jurisprudence récente de la Cour de justice et du Tribunal de première instance, 1997-06-17/1999-10-21

\*à l'exception des arrêts du Tribunal de première instance dans les affaires de Fonction Publique

## 2. CELEX

Celex est un système interinstitutionnel de documentation automatisée pour le droit de l'Union Européenne. Le stock documentaire de CELEX ne se limite pas aux seuls actes juridiques. Il s'étend également aux documents offrant une information complémentaire utile à une appréhension intelligible et complète des règles de droit communautaire, tels les actes préparatoires, la jurisprudence ou encore les questions parlementaires.

*Cette base offre donc une valeur ajoutée d'analyse.*

CELEX vous ouvre un accès multilingue à une large gamme d'actes juridiques :

- les traités fondateurs,
- le droit dérivé contraignant ou non,
- les avis et les résolutions des institutions et organes de l'Union européenne,
- la jurisprudence de la Cour de justice
- la jurisprudence du Tribunal de première instance des Communautés européennes.

La base de données CELEX est produite et gérée par l'Office des publications officielles des Communautés européennes.

L'on peut s'abonner à Celex par le biais d'un gateway officiel. Outre le droit éventuel d'abonnement, les coûts d'utilisation de CELEX comprennent les droits perçus sur la visualisation des documents et une charge fixe pour la recherche, comme suit :

- Visualisation d'un document dans le format notice (toutes données analytiques sans le texte) - 0,3 écu par document
- Visualisation du texte intégral, avec (format tout) ou sans données analytiques (format texte) - 1,2 écu par document
- Un droit de recherche de 0,2 écu est perçu en supplément. Cette charge forfaitaire est facturée pour toute recherche conclue par l'affichage d'une liste de résultats, sans qu'importe le choix ultérieur de visualiser, imprimer ou télécharger les documents dans l'un ou l'autre des trois formats.

## 3. SCADplus ([http://europa.eu.int/scadplus/scad\\_fr.htm](http://europa.eu.int/scadplus/scad_fr.htm))

SCADplus signifie "Service central automatisé de documentation". Il comporte de l'information générale sur les politiques de l'Union Européenne. Cette base était utilisée jusqu'il y a peu de temps exclusivement en interne. Dans le cadre de la transparence et de la généralisation des outils internet, il a été jugé utile de la rendre accessible au grand public. Bientôt, cette base comprendra

*Aix-en-Provence, samedi 30 octobre 1999. EUROPA, le serveur interinstitutionnel de l'Union européenne.* 3

des hyperliens avec Celex et EUR-Lex, ce qui permettra tout d'abord de se placer dans un contexte, puis de rechercher les textes à la base de tout le contexte juridique qui vous intéresse.

SCADplus comprend :

- un calendrier des réunions du Conseil des Ministres, du Parlement européen et du Comité économique et social et du Comité des Régions
- des fiches de synthèse et de suivi législatif sur les politiques de l'union, sur l'agenda 2000 (pour une Union plus forte et plus large ) et sur l'élargissement
- un guide pratique de la libre circulation des personnes
- des indications sur les programme, réseaux et observatoires en matière sociale et d'éducation
- la base SCAD qui comprend les références documentaires les plus actuelles:
- législation et publications communautaires;
- articles de périodiques provenant du monde entier;
- avis des partenaires sociaux
- le bulletin SCAD
- un glossaire sur la réforme institutionnelle
- le Traité d'Amsterdam, mode d'emploi

Ceci est une mine d'informations et de terminologie sur les activités et les politiques de l'Union. L'avantage majeur est que ce soit multilingue.

#### **4. Eurodicautom**

##### **Présentation sommaire d'Eurodicautom sur EUROPA**

Où se trouve Eurodicautom sur le serveur internet de la Commission ? Voilà une question qui n'a pas de réponse simple ni évidente d'ailleurs. Etant donné que le serveur EUROPA compte 1.153.000 documents environ, il est clair que l'on ne va pas trouver Eurodicautom sur la page d'accueil du serveur internet de la Commission. Il y a tant de services qui souhaiteraient apparaître sur cette première page !

Retenons un principe fondamental de toute recherche d'information sur internet et principalement sur les sites gouvernementaux et ceux des services publics. L'information y est souvent structurée en fonction de la hiérarchie de l'institution vers laquelle on se dirige. La Commission n'échappe à la règle.

Si EUROPA était doté d'un vrai index alphabétique ou d'un index thématique exhaustif (ce qui est particulièrement difficile à réaliser mais surtout à mettre à jour, chaque changement entraînant une quantité impressionnante de modifications), nous pourrions facilement retrouver l'information. J'ai essayé de retrouver Eurodicautom en utilisant le moteur de recherche et en tapant tout simplement dans la boîte de dialogue le mot "Eurodicautom". J'ai obtenu 490 réponses. La première pointe vers le texte fondamental de R. Goffin dont je me suis fortement inspiré pour réaliser cette partie de l'exposé et dont je recommande vivement la lecture à tous (<http://europa.eu.int/comm/sdt/edic/goffedic.doc>). En fait, la réponse 31 (du moins tel était le numéro ce jour-là) vous permet d'accéder à Eurodicautom mais il faut bien avouer que ce n'est pas la méthode la plus facile.

Je suis intervenu récemment auprès de l'équipe chargée d'EUROPA pour qu'Eurodicautom soit plus proche de la page d'accueil multilingue (<http://europa.eu.int/>). Dorénavant, il figure sur la seconde page d'accueil, par exemple celle en français (<http://europa.eu.int/index-fr.htm>). Tout en

bas de l'écran, vous trouverez une icône 'Information'. En cliquant dessus, vous arrivez à l'application Eurodicautom.

Si nous regardons la page Eurodicautom, "a translator's best friend on the web" (<http://europa.eu.int/comm/sdt/edic/indexfr.html>), nous voyons qu'elle est riche en informations.

Nous y trouvons des explications sur l'interface dite de base et sur l'interface expert, donc comment utiliser l'interface au mieux.

Puis une rubrique "comment utiliser Eurodicautom" nous amène vers des écrans expliquant comment affiner sa recherche. (<http://europa.eu.int/comm/sdt/bulletins/lingofr/lingopg1.htm>)

Vous disposez d'un guide expliquant en anglais de façon peu conviviale la façon dont l'interface a été conçue.

En cliquant sur description (<http://europa.eu.int/comm/sdt/edic/goffedic.doc>), l'on accède à un document de base capital de 37 pages qui relate toute l'histoire d'Eurodicautom. L'article, qui est un peu l'adieu du grand terminologue Roger Goffin à la profession, mérite une lecture attentive et exhaustive, car il reprend tout l'historique mais aussi tous les éléments qui ont été à la base de la naissance mais aussi de l'évolution d'Eurodicautom. Cet article s'intitule "Eurodicautom, la banque de données terminologiques multilingues de la Commission européenne, 1973-1997."

Si je clique sur le lien 'plus d'informations', j'aboutis à une page qui réplique partiellement l'information déjà trouvée ailleurs. Cependant, elle me donne quelques indications chiffrées sur Eurodicautom. Le 15 février 1998, la banque terminologique contenait 1.189.353 documents correspondant à environ 5 millions d'entrées.

Les différentes langues y sont assez bien représentées :

Langues	Entrées	Entrées	Langues
Allemand (DE)	670.940	885.350	Français (FR)
Anglais (EN)	851.325	851.325	Anglais (EN)
Danois (DA)	431.862	670.940	Allemand (DE)
Espagnol (ES)	407.366	477.233	Néerlandais (NL)
Finnois (FI)	133.000	474.193	Italien (IT)
Français (FR)	885.350	431.862	Danois (DA)
Grec (EL)	378.562	407.366	Espagnol (ES)
Italien (IT)	474.193	378.562	Grec (EL)
Latin (LA)	13.028	359.200	Portugais (PT)
Néerlandais (NL)	477.233	133.000	Finnois (FI)
Portugais (PT)	359.200	71.140	Suédois (SV)
Suédois (SV)	71.140	13.028	Latin (LA)

## Les principes de base d'Eurodicautom<sup>1</sup>

<sup>1</sup> Toute cette section du texte a été largement empruntée au texte remarquable de R. Goffin "Eurodicautom, la banque de données terminologiques multilingues de la Commission européenne, 1973-1997, accessible sur EUROPA à l'adresse : <http://europa.eu.int/comm/sdt/edic/backgredfr.htm>

La stratégie de recherche sur Eurodicautom a été mise au point par M. Pol François, ingénieur civil et informaticien à la Commission. Pour éviter des réponses du type tout ou rien, M. François a imaginé un système qui décompose une unité complexe en ses constituants (sauf bien entendu pour l'uniterme).

"A défaut de la réponse totale, Eurodicautom affiche les réponses partielles dans un ordre décroissant de pertinence en fonction de l'adéquation à la question posée. Le système suit donc la démarche naturelle du traducteur, lequel recherche prioritairement l'expression complète et explore ensuite, à défaut de mieux, les différents composants du multiterme recherché" (R. Goffin). Ceci s'applique au mode de recherche habituel (partial match).

Le mode de recherche "as is, full match" recherche l'occurrence des termes telle quelle et ne répond plus alors au principe énoncé ci-dessus. Le mode de recherche "all words" fournira quant à lui une pléthore de réponses, c'est-à-dire le terme dans ses multiples combinaisons.

### Les mécanismes de tris multicritères.

Citons ici Roger Goffin qui a lui-même puisé pour partie cette information dans un document interne à la Commission rédigé en 1980 par M. P. François.

*Le système va opérer une hiérarchisation automatique des réponses et les présenter dans un ordre décroissant de pertinence. La hiérarchie des réponses s'oriente autour de cinq critères intrinsèques ou permanents et de deux critères optionnels. Passons-les en revue :*

Premier critère : l'ordre des réponses est fonction du nombre d'unitermes contenus dans la question. Soit une question à trois termes ( $q = a b c$ ), les réponses se présenteront dans l'ordre suivant, d'abord la séquence  $a b c$ , puis la séquence  $a c b$ , ensuite les réponses comprenant deux des trois termes demandés  $a b$ ,  $a c$ ,  $b c$ , enfin des séquences plus longues comportant au moins deux des termes demandés.

Soit la question :

= moteur linéaire asynchrone (a b c)

séquence des réponses :

= moteur linéaire asynchrone (a b c)  
= moteur asynchrone de type linéaire (a c b)  
= moteur asynchrone (a c)  
= moteur linéaire (a b)  
= moteur asynchrone synchronisé (a c x)  
= moteur linéaire homopolaire (a b x<sub>1</sub>)  
= moteur linéaire à double inducteur (a b x<sub>2</sub>)  
= moteur linéaire synchrone (a b x<sub>3</sub>)  
= moteur linéaire à flux axial (a b x y)  
= moteur électrique linéaire (a x b)  
= moteur linéaire à simple inducteur (a b x y)  
= moteur linéaire à inducteurs (a b x)

Deuxième critère : le poids d'un terme réponse dépend de sa valeur terminologique. Un terme figurant dans le champ vedette (VE) ou dans le champ mot clé (CM) a une valeur supérieure à celle d'un terme apparaissant dans le champ phrase. L'auteur d'une fiche reconnaît l'intérêt terminologique du terme en lui donnant le statut de vedette (VE).

Troisième critère : parmi les réponses partielles, la priorité est donnée aux réponses ne contenant pas d'autres unitermes. A défaut de moteur linéaire asynchrone, on trouvera d'abord moteur asynchrone sans plus, avant moteur asynchrone synchronisé .

Quatrième critère : si les entités réponses ne possèdent pas les n termes de la question, elles se présentent selon une alternance d'apparition des différents termes composants. Soit la question a b : éolienne bipale, à défaut de réponse a b, l'ordre des réponses sera a, b, a, a, c-à-d éolienne, (hélice) bipale, éolienne, roule éolienne (plus de réponse avec bipale).

Cinquième critère : parmi les entités classées selon les critères 1 à 4, certaines sont éliminées car elles n'atteignent pas un seuil d'acceptabilité (n-1 termes). Ainsi à la question stockage inertiel par ressorts, une entité contenant uniquement stockage ne présente aucun intérêt. Par la limitation des réponses partielles, on réduit le temps d'ordinateur et on évite une réaction de rejet de la part du traducteur.

Les deux critères optionnels concernent l'indication du code matière et le choix des langues. Si l'on utilise l'option CM, le système privilégie, après classement selon les critères permanents, les fiches réponses nanties du code-matière. L'option "langue" se fait avant l'interrogation proprement dite : une langue source et une ou plusieurs langues cibles; l'utilisateur peut encore appeler la visualisation de toute la fiche par la commande "a".

Que comporte une fiche type ?

- une vedette VE
- une phrase PH
- une note NT
- un sujet ou code matière CM
- un code de fiabilité CF
- une date
- un numéro d'identification NI
- le type de document (collections) TY
- le bureau de terminologie (auteur) BE

L'on dispose dans Eurodicautom essentiellement de trois types de fiches :

1. *la fiche terme à terme* : la vedette identifiant ici la notion est un mot simple (uniterme) ou un syntagme (multiterme) avec son équivalent dans les autres langues. Ce nombre peut varier de une à douze (les onze langues officielles plus le latin). Les synonymes figurent dans le même champ. Ex : *libre circulation des travailleurs*.
2. *la fiche avec définition (DF) ou contexte illustratif* : la vedette-notion est assortie d'une définition ou d'un contexte éclairant qui illustre le sens du terme. Ex : *palplanche*.

3. *la fiche phraséologique* (PH) donne des phrases ou membres de phrases en deux ou plusieurs langues ou des contextes définitoires originaux. Ce champ permet d'illustrer le fonctionnement du terme.

Dans le cadre de cet exposé, il n'est pas possible d'aller plus loin dans la description d'Eurodicautom, car nous devrions y consacrer énormément de temps.

Je saisis l'occasion de cet exposé pour attirer votre attention sur d'autres bases terminologiques qu'EUROPA :

- Euterpe du Parlement européen ([http://muwi.trados.com:4712/MTW\\_LOGON](http://muwi.trados.com:4712/MTW_LOGON))
- TIS du Conseil des Ministres (<http://tis.consilium.eu.int/isowebtis/frames/introfsEN.htm>)
- ILOTERM, base terminologique du Bureau international du travail (<http://ilis.ilo.org/ilis/ilisterm/ilsearte.html>)
- Termiplus, du Bureau canadien de la traduction (<http://www.bureaudelatraduction.gc.ca/>)

## 5. Recherche sur EUROPA et principes de base utiles pour d'autres moteurs de recherche

La plupart des moteurs de recherche actuels se fondent sur des principes relativement communs, notamment l'application des opérateurs booléens. Par ailleurs, les règles de syntaxe, donc la façon d'écrire une commande est capitale. Ainsi, dans le moteur de recherche d'EUROPA, le fait d'utiliser les guillemets pour entourer une expression recherchée oblige le moteur à rechercher le mot tel qu'il est orthographié. Une certaine connaissance des fonctionnalités d'un moteur est donc un élément fondamental de l'aboutissement des demandes que vous poserez au moteur de recherche.

Le moteur de recherche d'EUROPA affiche les résultats tout d'abord par niveau de pertinence (le degré de pertinence le plus élevé est de 100%). Puis, il fournit le titre ou référence du document, mention pas toujours très explicite d'ailleurs. Ensuite, il indique dans quel format est confectionné le document retrouvé (Word, Html, pdf). Enfin, selon que l'on clique sur le lien souligné en bleu ou sur la petite flèche rouge, le résultat affiché n'est pas identique. Cliquer sur le lien vous affiche le document, tandis que cliquer sur la flèche rouge vous achemine vers la première occurrence du ou des mots recherchés. Tant qu'il y a des flèches pointant vers la droite, il existe d'autres occurrences du même mot dans le document consulté.

Essayez de poser la question de la façon la plus pertinente possible afin de ne pas avoir trop de bruit de fond, donc de réponses pléthoriques et inutiles. Au-delà de 100 documents pertinents, les autres n'ont plus aucun sens, vous vous lasserez avant d'arriver au bout. Ne vous découragez pas ! Il n'est pas inintéressant d'aller voir les dernières occurrences des trouvailles, ceci permet parfois de mieux comprendre comment le moteur a réagi en fonction de votre question. Recommencez en posant la question autrement, éventuellement avec d'autres termes. N'oubliez pas qu'EUROPA comporte actuellement 1.152.946 documents !

### *Principes de base*

- écrivez en minuscules
- Verity Search 97 fait la différence entre les signes diacritiques, en d'autres termes la lettre surmontée d'un accent circonflexe est pour lui quelque chose de différent de la lettre surmontée d'un accent aigu ou grave
- mettez vos termes entre guillemets si vous voulez que le moteur de recherche trouve exactement le terme tel qu'écrit, exemple : "recyclage de bouteilles"
- les mots d'une phrase doivent bien entendu être séparés par un espace.



- si vous voulez poser des questions plus complexes, lisez les fichiers d'aide. En effet, la syntaxe à utiliser pour des opérateurs booléens tels que OU et ET (souvent d'ailleurs à écrire en anglais) est particulièrement contraignante. Les opérateurs booléens doivent être séparés des mots faisant l'objet de la requête par un espace.
- attention, sur EUROPA, vous interrogez directement en multilingue, ce qui signifie que les homographes des autres langues seront récupérés par le moteur de recherche. EUROPA ne permet pas de préciser la langue dans laquelle on recherche.

#### *Quelques fonctionnalités de base*

- si vous tapez rêve ou "rêve", vous n'obtiendrez pas les mêmes résultats. En effet, taper rêve permet de retrouver toutes les formes du verbe rêver ainsi que le mot rêverie. Tandis que taper "rêver" ne donne comme résultat vraiment que le verbe rêver à l'infinitif. J'ai fait le test : rêve donne 15.361 réponses, "rêver" n'en donne que 21.
- utiliser le joker (wild card facility) matérialisé par l'astérisque (\*) dans le moteur de recherche. Par exemple rêve\* donne 231 réponses, mais ce mode d'interrogation en vous donne pas le mot rêve en tant que tel. Vous pouvez faire vous-mêmes des tests. Si vous voulez savoir combien de fois le terme rêves (au pluriel) apparaît, tapez "rêves" et vous verrez qu'il apparaît 40 fois, dans les 231 réponses.
- si vous n'êtes pas sûrs de l'orthographe d'un mot utilisez les ?? (jokers également). Par exemple, taper "encyclo?dia" ne donnera pas les mêmes résultats que "encyclo??dia", le second terme donnant la réponse 'encyclopaedia' tandis que le premier donne la réponse 'encyclopedia'. Ceci peut être utile pour les noms propres dont on n'est pas certain de l'orthographe.
- si vous recherchez un sigle dont vous pensez qu'il est écrit en majuscules, vous pouvez forcer le moteur de recherche à ne rechercher qu'en majuscules. J'ai pris personnellement l'exemple du mot anglais leader. Si je tape "leader" j'ai 5.350 réponses, si je tape <CASE> LEADER, je n'en ai que 1.593 et je suis à peu près certain que ces réponses concernent exclusivement le programme LEADER.

### *Recherches plus poussées*

Si vous n'avez pas obtenu de réponses correspondant à votre question, vous pouvez vous essayer aux techniques plus pointues, mais vous devrez alors consulter l'aide en ligne pour y parvenir (<http://europa.eu.int/geninfo/search/index.htm#start>).

En effet, la syntaxe d'écriture de votre question est capitale !

#### Utilisation des opérateurs booléens : AND, OR, NOT, NEAR, SENTENCE, PARAGRAPH

Ces opérateurs permettent des recherches plus fines mais attention à la façon de formuler. Elle influence considérablement la nature de vos réponses.

- AND

si vous tapez : banques AND prêt AND entreprises (avec espace entre le terme et l'opérateur), le moteur de recherche vous renverra les documents dans lesquels ces trois mots se trouvent. Ces termes peuvent se trouver très loin les uns des autres, la requête a donné 1.602 réponses sur EUROPA.

- NEAR

si vous tapez : banques<NEAR>prêt<NEAR>entreprises (sans espaces), le moteur de recherche vous renverra les documents où ces trois mots se trouvent dans un certain voisinage, sans être absolument voisins. La requête a donné 1.112 réponses sur EUROPA.

- PARAGRAPH

si vous tapez : banques<PARAGRAPH>prêt<PARAGRAPH>entreprises (toujours sans espaces), le moteur de recherche vous renverra les documents où ces trois mots se trouvent dans le même paragraphe, soit dans 444 cas.

- SENTENCE

si vous tapez banques<SENTENCE>prêt<SENTENCE>entreprises (toujours sans espaces), le moteur de recherche vous renverra les documents où ces trois mots se trouvent dans la même phrase, soit dans 82 cas.

- OR

OR est un opérateur qui génère beaucoup de bruit de fond (beaucoup de réponses peu pertinentes). Je conseille de l'éviter. Il n'est utile que si vous recherchez des termes appartenant à une même famille conceptuelle.

Voici un exemple : je m'intéresse aux chevaux de façon générale. Je puis poser la question de façon spécifique. Mais comment ?

Je tape:

cheva*	2.148 réponses
"cheval"	153 réponses
"chevaux"	509 réponses
"chevalin"	7 réponses
"équidés"	406 réponses
"ongulés"	50 réponses

Je puis également taper : cheval OR chevaux OR chevalin OR équidés. Cette requête me donne 1.129 réponses. J'ai ainsi pu réduire le nombre de documents traitant de mon sujet.

- NOT

NOT est un opérateur booléen très utile lorsque l'on veut exclure un sujet. Dans le cas précédent, si je tape : cheval OR chevaux OR chevalin OR équidés NOT ongulés. Ceci me donne 1.114 réponses, soit 15 de moins que précédemment. Sur les 50 réponses relatives aux ongulés, il y en a donc 15 qui ne parlent ni de chevaux ni d'équidés.

### *Recherches très pointues*

Il est possible d'aller plus loin encore mais ceci devient vraiment un travail d'expert que nous n'aborderons pas ici. Afin néanmoins d'être plus précis, ajoutons qu'il existe des possibilités d'interroger les métadonnées des documents, mais il est clair que ceci ne concerne que les documents situés sur EUROPA qui sont formatés en HTML (hypertext markup language). Par ailleurs, ceci présuppose que les documents ont été tous dotés de métadonnées et que ce travail d'indexation a été fait correctement, ce qui est loin d'être le cas sur EUROPA.

### **Conclusions**

Le présent exposé a essayé de présenter des éléments utiles pour le traducteur et d'expliquer comment sur un serveur internet comme EUROPA l'on peut procéder à la recherche d'information. J'espère sincèrement que ces informations permettront tant au traducteur qu'à l'interprète de trouver ce dont il a besoin pour accomplir son œuvre traduisante.

Internet est une gigantesque mine d'informations, encore faut-il savoir l'utiliser intelligemment et efficacement. J'ai tenté de démontrer qu'en combinant informations purement linguistiques et renseignements de fond, le serveur EUROPA doit pouvoir faciliter le travail du linguiste.

Certes, ceci ne s'acquiert pas du jour au lendemain. Mais celui qui sait utiliser ces merveilleux outils modernes pourra en tirer profit et améliorer la qualité de ses prestations. Utilisez votre imagination, la souplesse de votre esprit et faites agir vos facultés combinatoires de la façon la plus vaste possible, alors vous trouverez plus facilement ce que vous recherchez. N'oubliez finalement pas non plus que tout ne se trouve pas encore sur EUROPA !

*Jean-Bernard Quicheron, Aix-en-Provence, 30 octobre 1999*