

# Exemple de pratique terminographique en entreprise

La constitution d'une terminologie de référence destinée à l'entreprise passe par plusieurs étapes: l'extraction des termes à partir du corpus, la validation au cours de laquelle les experts sont guidés par le terminologue et l'organisation des termes validés en domaines et sous-domaines. La dernière partie de cet article présente la difficulté de suivre certains principes de la théorie générale de la terminologie afin de produire une terminologie destinée à l'entreprise.

Termes-clés:  
corpus; domaine; extraction terminologique; pratique terminographique; théorie de la terminologie; validation.

## 1 Introduction

**S**uite à un projet de la DE (Direction de l'équipement d'EDF), une terminologie de référence a été amorcée. Cette dernière doit être composée de termes extraits de la documentation EDF et organisés en domaines et sous-domaines selon un cahier des charges établi par la DER (Direction études et recherches d'EDF).

La constitution d'une terminologie de référence entre dans le cadre du projet RMI (Référentiel méthodologique d'ingénierie) mené par la DE. Cette dernière se charge de la conception, la réalisation et l'exploitation des centrales nucléaires et mène le projet d'harmoniser et de mettre en cohérence l'ensemble des documents liés aux métiers de l'ingénierie: des doctrines, des dictionnaires, des thésaurus, des modèles de documents, etc. La DE souhaite faciliter, par le biais de la terminologie de référence, la rédaction de documents (contrôle terminologique, contrôle de la normalisation et de la cohérence et aide à la traduction) dans un premier temps et à terme, fournir des moyens avancés pour la consultation des documents techniques.

## 2 Description du corpus de référence

Le corpus que la DE a fourni et à partir duquel les termes sont extraits

est constitué d'un ensemble de documents techniques, les DSE (Dossiers de systèmes élémentaires) destinés aux constructeurs de centrales nucléaires. Ces dossiers de spécifications sont multi-auteurs. Des spécialistes de chaque métier définissent les systèmes élémentaires du palier N4 (dernière génération de centrale nucléaire) et ce sous divers angles tels que la conception, la conduite, le fonctionnement et la sûreté. Ces thématiques sont reflétées de manière très explicite dans la structure même des DSE. En effet, ils sont tout d'abord organisés en ensembles désignés par des lettres de l'alphabet. Ces lettres peuvent faire référence à une zone de la centrale nucléaire comme la lettre [1] R pour le réacteur ou bien des traitements tels que [2] J pour l'incendie. Chaque groupe est divisé en plusieurs systèmes élémentaires. Ainsi, l'ensemble R compte 17 sous-ensembles tels que le [3] RCP qui concerne le circuit élémentaire ou le [4] RHY qui indique l'hydrogénation et les besoins nucléaires. Les sous-ensembles faisant référence aux différents systèmes élémentaires présentent une certaine homogénéité. Ils se composent de rubriques qui elles-mêmes se divisent en sous-rubriques. Cette structure est commune à une majorité de DSE. À telle enseigne que [5] le document D0 contient le sommaire général. Les sous-rubriques abordent notamment l'historique de ces documents, le fonctionnement du système élémentaire, son rôle, sa base de conception et la description des matériels.

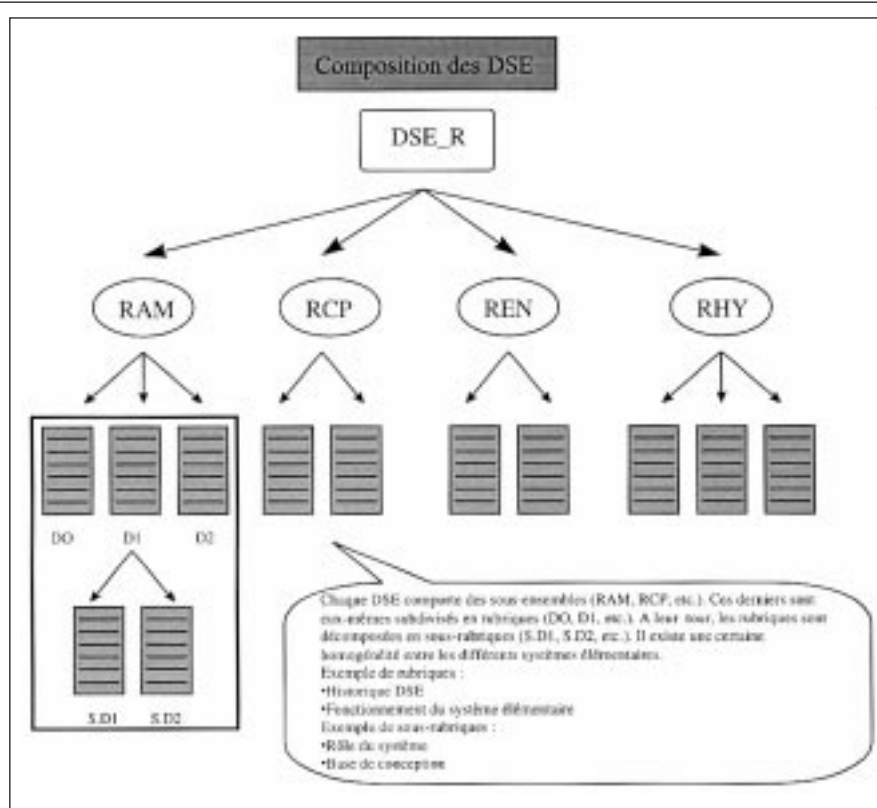


Figure 1 :  
composition des DSE

Suite à une réunion avec les experts de la DE chargés de la validation au cours de laquelle une première terminologie composée de termes extraits des DSE a été présentée, le projet a quelque peu changé d'orientation. Compte tenu de la nature très technique des documents constituant le corpus de départ, les futurs valideurs ont jugé nécessaire de construire parallèlement une seconde terminologie illustrant le savoir-faire de l'entreprise à partir d'un nouveau corpus, le *Mir* (*Manuel d'ingénierie de référence*). Ce changement de cap dénote l'évolution constante du projet et le manque de stabilisation du corpus. Bien que le traitement du *Mir* ainsi que sa structuration ne soient pas abordés dans cet article en raison des priorités données pour l'instant dans le projet

RMI à la terminologie technique basée sur les DSE, il est intéressant de noter la différence de contenu de ces deux corpus. Cette nouvelle source textuelle ne sera pas traitée sur le plan de la structure mais du contenu uniquement du fait que le corpus ne contient que des extraits du *Mir* et non la totalité.

Le *Mir* a été créé afin de répondre aux besoins de la DE essentiellement en matière de normalisation de documents d'ingénierie, de formalisation des méthodes d'ingénierie et de mise en cohérence entre projets, unités et types de documents propres à la DE. Il permet d'établir la structure documentaire de la DE et d'acquérir des connaissances au sein de cette direction.

Le *Mir* traite à la fois du savoir-faire tel que la méthode de fabrication de vannes et des objets techniques, à savoir les vannes en elles-mêmes. Il comporte tout d'abord les codes de classement des plans. Ils se présentent sous la forme d'une liste. Celle-ci est composée d'un repère (une valeur numérique) et de sa référence (du texte). Les codes désignent par exemple les terrains, les installations chantiers et les différents éléments entrant dans la construction des bâtiments. D'autre part, le *Mir* réunit des spécifications dont celle de la typologie des documents d'ingénierie. L'objectif de cette spécification est de répertorier et d'identifier les différents types de documents de contenu technique réalisés à la DE ou par des tiers.

### 3 Mise en place d'une méthodologie de travail

La constitution d'une terminologie à partir du corpus des DSE implique les trois étapes suivantes :

- Extraction des candidats termes par l'outil terminologique *Lexter* (Bourigault 1994) ;
- Validation par des experts de la DE ;
- Organisation des termes validés en domaines et sous-domaines.

#### 3.1 Extraction des candidats termes : une méthodologie incrémentale

Cette méthodologie consiste à soumettre les DSE à un outil d'extraction terminologique *Lexter* afin d'effectuer l'extraction de candidats termes. Les listes obtenues sont soumises aux experts de la DE afin d'être validées et organisées en domaines et sous-domaines. Enfin, les termes sont intégrés dans la terminologie de référence. Compte

tenu de la taille du corpus des DSE (près de 18 000 pages de documents), il est préférable que le traitement s'amorce à partir d'un sous-corpus puis que la méthodologie adoptée soit systématisée afin d'être appliquée au reste des documents.

### Le traitement itératif des données

Il ressort du traitement d'un sous-corpus test une méthodologie permettant une extraction et une validation plus efficaces.

Les ensembles DSE sont traités un à un. Le traitement *Lexter* d'un DSE N produit une liste de termes candidats que l'on compare à la liste de termes validés du DSE précédent (N-1). Cette phase de mise en parallèle représente une pré-validation destinée à faciliter le travail des experts de la DE. Elle consiste à attribuer la validité des termes du DSE N-1 existant dans le DSE N. Ce premier élagage permet d'obtenir une

liste plus réduite qui fera l'objet d'une validation définitive toujours par les experts. Dès lors, les termes obtenus sont introduits dans la terminologie de référence.

### Initialisation du traitement : comparaison *Paluel* – *N4*

Le traitement itératif est amorcé avec les DSE\_R du palier *N4*. Le sous-corpus de départ est constitué de sept sous-ensembles du DSE\_R (RAM, RAZ, REA, REN, RPR, RRI, et RRM). Ce choix a été motivé par la phase de comparaison. Une comparaison est faite avec les 7 sous-ensembles équivalents d'un palier antérieur, le *Paluel*. Ils permettent la pré-validation des candidats termes obtenus après l'extraction terminologique. Les DSE *Paluel* sont organisés de la même manière que les DSE\_R du *N4* en ensembles de lettres et en sous-ensembles de systèmes élémentaires. Ils ont fait

l'objet d'une validation dans le cadre d'un projet EDF. Ce travail de sélection explique le choix des documents *N4* faisant référence aux mêmes systèmes élémentaires. Les termes candidats ont été extraits à l'aide de *Lexter* puis la liste a été validée puis structurée en domaines par Henry Boccon-Gibod, ingénieur expert.

Les termes potentiels issus de la comparaison *Paluel* – *N4* sont soumis aux experts. Une fois validés, ils représentent la première version de la terminologie à laquelle seront ajoutées, au fur et à mesure des traitements, de nouvelles listes de termes provenant d'autres DSE. Le traitement *Lexter* des sept sous-ensembles du DSE *Paluel* avait permis d'obtenir 11 500 termes candidats. Après la validation de l'expert Henry Boccon-Gibod, 4 000 termes avaient été validés « oui » et 7 500 validés « non ». Les sept sous-ensembles *N4* ont produit, après extraction terminologique, 24 500 termes potentiels.

La comparaison *Paluel* – *N4* permet d'obtenir les résultats suivants :

- Les termes communs aux deux paliers sont au nombre 3 500 dont 1 500 validés « oui » ;
- Les termes candidats *Paluel* absents du *N4* s'élèvent à 8 000 dont 2 500 validés « oui » ;
- Les termes candidats *N4* absents de *Paluel* atteignent le nombre de 20 000.

### 3.2 Travail de validation des experts : importance des consignes

La présentation aux experts de la DE d'un ensemble de termes extraits des DSE\_R met en évidence la nécessité de produire des consignes de validation. Une interface de validation simple permet d'afficher les termes potentiels et de leur attribuer

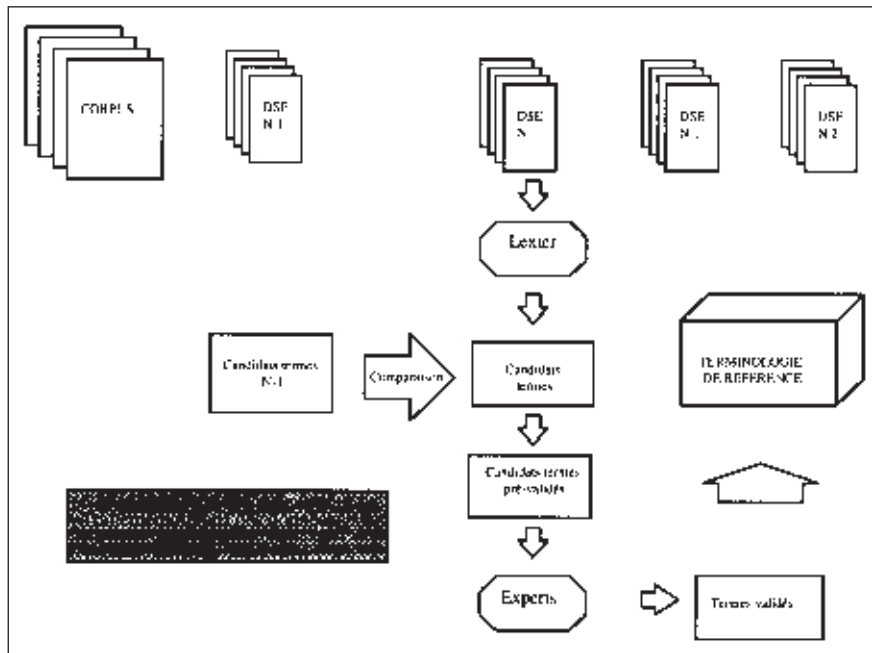


Figure 2 :  
Traitement itératif des données

quatre validités possibles: «non-validé», «oui», «non» et «à voir». Les syntagmes nominaux sont au nombre de 3 300 et sont passés en revue par trois experts. Ces derniers valident près de 500 termes en deux heures. 534 termes sont vus par au moins un expert. Les résultats du travail de validation sont les suivants:

- 111 sont validés oui/oui/oui;
- 70 sont validés non/non/non;
- 343 sont litigieux.

Au cours de la phase de validation, les experts mettent l'accent sur l'absence de consignes de validation précises. En effet, les seules recommandations transmises aux valideurs sont de sélectionner les termes candidats dans la perspective d'une intégration dans une application d'aide à la rédaction et à la compréhension de documents techniques et de ne pas en éliminer trop, certains, bien que très généraux, pouvant constituer des noeuds dans la terminologie de référence.

Les experts sont très vite confrontés aux «joies» de la validation. La première question est de déterminer s'ils doivent retenir ou rejeter les candidats termes par rapport à leurs spécialités (aspect «électricité», aspect «électronique», etc.) ou bien par rapport à l'ingénierie DE de manière générale. Les termes de gestion tels que [6] *repère signalétique* peuvent être communs à plusieurs métiers. Par contre [7] *disjoncteur de départ de tranche* figurera dans la terminologie d'un mécanicien mais pas dans celle d'un électricien. Ce dernier ne sélectionnera pas le terme [8] *pression relative de calcul* propre au métier de mécanicien.

Un autre cas de figure se pose au cours de la sélection des termes: la présence de deux termes faisant référence au même objet mais relatifs à deux paliers distincts. Ainsi, [9] *salle de conduite de tranche* est propre à *Paluel* et existe dans le *N4* sous le nom de [10] *salle de contrôle-*

*commande*. Faut-il garder une trace de ce changement de dénomination ou supprimer les anciens usages?

Le besoin de retourner aux documents afin de vérifier le contexte dans lequel le terme est utilisé se fait ressentir de manière considérable. Ainsi, [11] *activité élevée*, un terme imprécis pris hors contexte, ne peut être retenu qu'à la condition qu'il désigne un seuil bien déterminé dans un domaine particulier. Il en est de même pour [12] *haute température*. Le seul moyen de faciliter la validation et de sélectionner la bonne occurrence du terme est de sauvegarder le lien existant entre texte et terme.

Les incertitudes des experts s'estompent après discussion et sélection des applications de la terminologie de référence: l'aide à la rédaction de documents techniques. Ils appréhendent de manière plus concrète la manière dont la terminologie permet un gain en productivité et en précision sur le plan de la rédaction sans compter la cohérence et ce après avoir assisté à la démonstration de deux outils d'aide à la production de documents techniques.

En effet, la terminologie doit être représentée graphiquement à l'aide de l'outil *WorldTrek* Picard *et al.* (1999), conçu à la Direction études et recherches d'EDF. Des propriétés telles que celles souhaitées par la DE à savoir la définition, la catégorie grammaticale, l'équivalent en langue étrangère et le contexte peuvent accompagner le graphique. La DE souhaite du reste que cette organisation s'inspire des métiers de la Direction de l'équipement. Ainsi, le travail de réflexion porte dorénavant sur les domaines et sous-domaines. Bien que les valideurs s'inspirent directement de leurs spécialités au cours du travail de sélection, il n'est pas garanti que ce découpage soit en adéquation avec l'organisation proposée dans les paragraphes suivants.

La terminologie doit également être intégrée dans l'application *TermChecker* (Boccon-Gibod 1999). Cet outil facilite la vérification de terminologie lors de la rédaction d'un document à l'aide d'un logiciel de traitement de texte classique. Il fonctionne à la manière d'un correcteur orthographique et propose des termes auxquels des statuts sont affectés. Ces derniers sont choisis en fonction des besoins des utilisateurs. À ce sujet, les remarques des experts sont très précieuses. Il ne s'agit pas uniquement de présenter les termes ayant changé de dénomination avec l'évolution des techniques mais de faire une mémoire de rédaction permettant de faire la transition entre les documents propres aux différents paliers des centrales nucléaires. Les experts expriment également la volonté d'attribuer des statuts conseillant l'usage d'un terme ou au contraire l'interdisant. Cette préoccupation part d'une volonté de normalisation et de stabilisation de la langue technique de la DE à la base du projet RMI.

Il ressort du suivi de la validation des termes potentiels par les experts la mise en évidence d'un autre aspect du métier de terminologue: celui de médiateur. Il détermine, avec les futurs utilisateurs qui s'avèrent être également les valideurs, les applications et le rôle de la terminologie. De plus, il ne se contente pas d'établir des règles afin de guider les experts dans leur lourde tâche. Il a pour rôle d'observer puis d'analyser le retour d'expérience de la validation et de formaliser les consignes afin de faciliter le travail de sélection et de le mettre en cohérence notamment lorsque les experts sont plusieurs. Il prend part de manière significative au travail de normalisation réalisé par le ou les valideurs.



### 3.3 Organisation des termes DSE en domaines et sous-domaines

La dernière étape de traitement, l'organisation des termes validés extraits des DSE en domaines et sous-domaines suscite beaucoup d'intérêt auprès des experts. La première tentative s'inspire de la structure existante du thesaurus EDF.

#### *Le thesaurus EDF*

Le thesaurus EDF, réalisé par le centre de documentation de la DER tout d'abord manuellement puis de manière automatique, a été créé exclusivement à des fins d'indexation. Il réunit des mots provenant de documents techniques de natures variées, caractérisés par un profil, auxquels est attribué le statut de descripteur. Ces mots-clés sont essentiels pour la recherche d'information, le classement et la diffusion de documents. Ces termes ont fait l'objet d'une collecte par des documentalistes puis d'une validation par des experts. Le thesaurus est mis à jour régulièrement par souci de coller au corpus dont les descripteurs sont extraits.

Le thesaurus est structuré en 45 points de vue dont 15 font l'objet d'un regroupement académique. Ils varient de la biologie à la thermique en passant par le droit, l'économie, les sciences de la terre, etc. Ils se divisent en domaines tels que l'électricité, le magnétisme et l'optique pour le point de vue sciences physiques ou assurance, famille et protection sociale pour le point de vue environnement social.

#### *Inadaptation de la structure du thesaurus EDF et proposition d'une nouvelle organisation*

Appliquer le découpage du thesaurus organisé en points de vue et en domaines s'avère très rapidement

une erreur et ce sur plusieurs points. Une seconde répartition des termes est proposée aux experts : calquer les domaines et sous-domaines de la future terminologie de référence sur la structure physique des DSE notamment la division en rubriques (les documents) et sous-rubriques (les sous-documents) présentée dans la partie consacrée à la description du corpus de référence.

La complexité des DSE porte tant sur la taille que sur le contenu. La très grande masse de documents (près de 18 000 pages) motivent indirectement l'organisation. Un ensemble de 7 dossiers de systèmes élémentaires a fait l'objet d'une extraction terminologique. L'outil informatique utilisé, *Lexter*, a permis d'obtenir 24 500 termes potentiels. Respecter la structure des DSE peut permettre un traitement plus rapide à condition que cela soit en adéquation avec ce qui est attendu de la terminologie de référence.

Sur le plan du contenu, la hiérarchisation des DSE, avant tout en sous-ensembles, à savoir les systèmes élémentaires, puis en rubriques et sous-rubriques, a pour avantage de constituer la base de travail des ingénieurs. Elle reflète des pratiques, des habitudes, une certaine vision des métiers de l'ingénierie. Cette structuration est donc directement rattachée à la réalité, à l'expérience des futurs utilisateurs de la terminologie de référence. S'inspirer du thesaurus EDF ne permet pas d'adapter la structuration à cette nécessité. Cet ouvrage concerne tout d'abord EDF en tant qu'entreprise puis la production d'énergie sous tous ses aspects. Le point de vue « sciences humaines » semble répondre difficilement à des besoins au sein de la Direction de l'équipement, plus particulièrement en ce qui concerne les domaines tels que « les arts et lettres » et « la philosophie ». Il reste très général du fait que le regroupement est

majoritairement académique et ne peut convenir à un public spécialisé. Il est important de rappeler que l'objectif « aide à la rédaction » de la terminologie de référence est très différent de celui du thesaurus conçu pour la recherche documentaire (Monteil 1995). En effet, celui-ci a pour applications la recherche documentaire et l'indexation. Ce découpage du thesaurus n'est pas adapté à la nature très technique et précise des métiers de la DE. Ces caractéristiques sont d'autant plus importantes lors de la création de documents techniques.

Le regroupement par rubriques et sous-rubriques est d'autant plus facile en raison de la sauvegarde au cours de l'extraction et de la validation du lien entre le terme et la séquence de texte dans laquelle il apparaît. Ainsi, les balises indiquant l'origine textuelle du terme reflètent les domaines et sous-domaines de la terminologie. La principale préoccupation est de connaître le taux de recouvrement des termes par rapport à la structure des DSE afin de confirmer la concordance entre les métiers de la DE et la structure des DSE. Le cas échéant, devrait-on donner la priorité à la disposition des documents en rubriques et sous-rubriques ou bien se fier davantage à la classification métiers des experts de la DE?

Il existe un autre point de divergence à propos du contenu : la nature des liens unissant les termes d'un même domaine et les domaines d'un même point de vue au sein du thesaurus. Se reporter au thesaurus EDF favorise la constitution de groupes de termes de type « générique-spécifique » majoritaires dans ce document. Ainsi, [13] les *échangeurs de chaleur* sont rassemblés par type à savoir [14] les *échangeurs à contact*, [15] les *échangeurs par mélange* et [16] les *échangeurs par surface* et non en fonction de leur rôle au sein d'un système donné

notamment le *transfert de chaleur* du [17] *cœur du réacteur* vers [18] le *circuit secondaire*. Par contre, le fait que les DSE reflètent davantage un métier implique une diversité de liens au sein d'un même domaine. Dans un DSE, la rubrique fonctionnement du système élémentaire réunit des termes décrivant le déroulement d'un processus. Le terme [19] *groupe turbo-alternateur* appartient à la même sous-rubrique « cycle de la vapeur » que les termes suivants: [20] *turbine*, [21] *sécheur-surchauffeur*, [22] *condenseur*, [23] *pompe d'extraction* et [24] *pompe alimentaire*. Ils sont unis par des liens fonctionnels. Par contre, la sous-rubrique description du matériel regroupe des termes illustrant la composition d'un système élémentaire. Dans cet exemple précis, les liens sont plutôt de type « tout-partie » et le même terme groupe *turbo-alternateur* figure avec [25] *chaudière nucléaire* et [26] *condenseur* dans le domaine « centrale nucléaire ».

Il ressort de cette analyse une nette différence entre le contenu du thesaurus EDF et celui du corpus DSE. L'élément déterminant est la personnalisation du travail linguistique qu'il s'agisse d'un thesaurus ou d'une terminologie de référence. La prise en compte de la réalité à savoir le contexte d'utilisation, le contenu des données à traiter et le public utilisateur est essentielle.

#### 4 La terminologie de référence EDF face à la doctrine

Dans le cas particulier de la constitution d'une terminologie de référence destinée à la Direction de l'équipement, certains principes de la théorie générale de la terminologie ne sont pas adaptés à la pratique terminographique.

La présentation aux experts de la DE d'une première terminologie réunissant exclusivement des termes DSE a changé le projet d'orientation. En effet, un nouveau corpus a été soumis afin d'illustrer davantage l'aspect méthodologique des métiers d'ingénierie. Le corpus est étroitement lié à cette volonté de construire une terminologie multifacettes. Déterminer les caractéristiques des documents composant le corpus permet de mieux cerner la future composition de la terminologie sur le plan des termes et de l'organisation en domaines et sous-domaines.

L'impossibilité de revenir au texte source pose problème. Au cours de leur travail de validation, les experts souhaitent à maintes reprises se replonger dans le contexte d'utilisation d'un futur terme. Cela souligne l'importance de sauvegarder un lien entre la terminologie et le corpus, du moins le terme et la séquence de texte dont il a été extrait. Bien que la théorie générale de la terminologie préconise le détachement des termes des textes, ce besoin se fait ressentir très concrètement au cours de la validation.

L'attribution de statuts aux termes extraits des DSE tels que « accepté », « interdit » ou « à éviter » afin de guider le rédacteur pose le problème de la polysémie et de la synonymie en terminologie (Felber 1984). Selon la doctrine, les changements de sens donnent naissance à des termes très ambigus. Il est donc préférable de créer des termes nouveaux pour représenter des notions nouvelles. Or, en ce qui concerne la synonymie, elle sème la confusion et donne l'impression qu'il existe plusieurs notions. Cependant, il est difficile de faire abstraction de ces phénomènes. Ces manifestations linguistiques reflètent les conditions de création et d'utilisation d'un terme. Yves Gambier (1991: 11)

évoque les phénomènes de polysémie et de synonymie en ces termes: « Les transferts (métaphoriques) de termes répondent aux besoins pressants de dénomination, au croisement des disciplines, au rapport motivé et non pure convention entre le signe et la notion ». Il est important de donner au terminographe la possibilité d'avoir recours à toutes les manifestations du langage, à partir desquelles il fait des choix selon leur pertinence par rapport au travail terminologique à réaliser. Il reste que réduire la polysémie lors de la description d'un domaine est nécessaire afin de garantir la stabilité des significations. En même temps, le terme prend sens dans son contexte d'utilisation. Tenir compte des nuances de synonymie est donc important. Cela permet de faire état de l'évolution d'un terme au sein d'une même discipline. Ainsi, dans le cas particulier des centrales, cela permettra de garder une trace de l'évolution de technologie lors du passage d'un palier à un autre.

La prise en considération de l'interdisciplinarité au cours du travail terminologique se pose concrètement dans le cas EDF. En terminologie classique, la création de classes universelles ne permet pas de refléter ce contact entre les différentes disciplines. Or, le terme résulte de l'innovation. Selon François Gaudin (1993: 83), « l'innovation » qui constitue l'essentiel de la production terminologique « naît de réseaux transversaux et la circulation langagière, l'échange et la contamination de concepts entre les disciplines sont des moteurs puissants d'innovation ». Cette innovation résulte des contacts entre différentes communautés scientifiques. « Ce qui motive dans ce cas les interactions, et les recatégorisations, qu'elles induisent, c'est l'existence d'un but commun, d'un objectif poursuivi par un travail ». Ce phénomène de circulation des connaissances se

manifeste dans les rubriques et sous-rubriques composant les DSE. De ce fait, le terme découle directement de la pratique d'une ou plusieurs communautés. Il ne s'agit pas de partage d'universaux comme le laisse présumer la théorie générale de la terminologie à travers la constitution de grandes classes canoniques dans le but de décrire une discipline mais d'une pratique entre des acteurs prenant part à la même expérience professionnelle. Ce cheminement de l'innovation au terme permet de constater que ce dernier provient directement de l'interdisciplinarité entre métiers.

Enfin, chaque discipline peut être abordée sous des angles différents en fonction des objectifs d'une application. Comme le souligne François Rastier (1994: 62), « Les modes de structuration varient selon les domaines ». Les domaines et sous-domaines d'une terminologie sont délimités en fonction du métier, de l'application et de l'utilisateur. Au lieu d'en réduire le nombre afin d'obtenir des classes de termes très générales, il faut au contraire garder à l'esprit les spécificités de chaque pratique afin de les décrire au mieux sur le plan notionnel. Toutefois, il est important d'opérer une sélection pertinente. En effet, Rastier (1994: 77) précise que « les différents axes sémantiques seront choisis selon les conditions de la description: pour opposer métro et autobus, on peut choisir la catégorie ferré *vs* routier dans un texte technique, mais aussi lent *vs* rapide si l'on décrit les raisons du choix des usagers... Bien entendu, ces divers axes ne s'excluent pas, mais une description pertinente doit rejeter les catégories inutiles ». Il est nécessaire d'avoir une vue d'ensemble de la discipline dont on dresse la terminologie afin de ne pas éliminer les éléments pertinents.

## 5 Conclusion

Le travail sur corpus, les différents contacts avec les experts de la DE ainsi que les tentatives d'organiser les termes potentiels extraits puis validés en domaines et sous-domaines ont eu pour effet de mettre l'accent sur certains éléments importants.

L'impact du corpus sur la nature de la terminologie à produire ne s'est jamais autant confirmé. L'analyse de sa structure et de son contenu par rapport aux résultats à produire permet de reproduire cet aspect multi-domaines présent en entreprise. De plus, dans le cas particulier d'EDF, l'étude du corpus a une lourde influence sur les modèles d'organisation en domaines et sous-domaines proposés pour la terminologie de référence.

La contextualisation des termes constituant la terminologie de référence est nécessaire tout d'abord aux valideurs qui ont besoin de sélectionner un terme en fonction de son sens dans la phrase mais également aux utilisateurs. La création de liens termes-textes dans une perspective d'aide à la rédaction a pour avantage de souligner les nuances entre termes proches sur le plan sémantique.

L'inadaptation de la structure du thesaurus EDF afin d'organiser les termes en domaines et sous-domaines met l'accent sur l'importance de la pertinence en terminologie. Il faut avoir à l'esprit qu'une terminologie efficace est « éphémère » du fait qu'elle s'attache à un besoin particulier, à un moment particulier, à un public particulier.

Enfin, le travail effectué sur le projet d'une terminologie de référence renforce la vision du terminologue en tant que médiateur. L'assistance, l'aide à la décision ainsi que l'orientation des experts au cours de la validation met en relief la nécessité de coopérer

avec le terminologue afin que le travail terminologique réalisé soit le plus pertinent possible.

## Remerciements

Nos remerciements à Henry Boccon-Gibod, Didier Bourigault et Monique Slodzian pour leurs conseils éclairés et leurs précieuses relectures.

*Yasmina Abbas,*  
*Centre de recherches en ingénierie multilingue,*  
*Institut national des langues et civilisations orientales,*  
*Paris,*  
*France.*

*Marie-Luce Picard,*  
*Groupe SOAD (Statistiques,*  
*optimisation et aide à la décision)*  
*du Service IMA (Informatique et*  
*mathématiques appliquées) de la DER*  
*(Direction des études et recherches*  
*d'EDF),*  
*Clamart,*  
*France.*

## Bibliographie

Boccon-Gibod (H.), 1999: « Enjeux de la maîtrise de la terminologie pour la production et la consultation de documents – Application à la documentation technique des installations de production d'électricité », dans *DocuWorld 1999*, Paris.

Bourigault (D.), 1994: *LEXTER, Un Logiciel d'Extraction de TERminologies-Application à l'acquisition des connaissances à partir de textes*, Thèse de l'École des hautes études en sciences sociales, Paris.

EDF-DER, SID (Département systèmes d'information et de documentation), Mise à jour 1997, *Thesaurus EDF*, français-anglais.

Felber (H.), 1984: *Manuel de terminologie*, Infoterm (Centre international d'information pour la terminologie), Unesco, Paris.

Gambier (Y.), 1991 : « Terminologie et sociolinguistique », dans *Cahiers de linguistique sociale*, n° 18, p. 43-44.

Gaudin (F.), 1993 : « Socioterminologie, Des problèmes sémantiques aux pratiques institutionnelles », dans *Publications de l'Université de Rouen*, n°182, p. 83.

Monteil (M.-G.), 1995 : « Indexation automatique et manuelle, comparaison et perspectives », *IDT'95*, Paris.

Picard (M.) et Boudailler (E.), 1999 : « Worldtrek for authoring and comprehension », *IUT'99*, Redondo Beach, Californie, USA.

Rastier (F.), Cavazza (M.) et Abeille (A.), 1994 : *Sémantique pour l'analyse*, Masson, Paris.

Slodzian (M.), 1995 : « Comment revisiter la doctrine terminologique aujourd'hui? », dans *La banque des mots*, n° spécial 7, p. 11-18.