



HAL
open science

Appropriation des savoirs et variété des langues d'usages des communautés de métier

Caroline Djambian, Sylvie Lainé-Cruzel

► To cite this version:

Caroline Djambian, Sylvie Lainé-Cruzel. Appropriation des savoirs et variété des langues d'usages des communautés de métier. GLAT 2012. Terminologies: textes, discours et accès aux savoirs spécialisés., May 2012, Genova, Italie. pp.343-356. hal-00811448

HAL Id: hal-00811448

<https://hal.science/hal-00811448>

Submitted on 10 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Djambian, C. & Laine-Cruzel, S. (2012). Appropriation des savoirs et variété des langues d'usages des communautés de métier. In *Terminologies: textes, discours et accès aux savoirs spécialisés: actes du V^e colloque international du Groupe de linguistique appliquée des télécommunications (GLAT)* (pp. 343-356). Gênes: GLAT.

Appropriation des savoirs et variété des langues d'usages des communautés de métier

Caroline Djambian*, Sylvie Lainé-Cruzel**

* LSIS, Faculté des Sciences et Techniques, Aix-Marseille Université
cdjambian@yahoo.fr

**ELICO, Université Jean-Moulin Lyon III
sylvie.laine@univ-lyon3.fr

Abstract

The specialized professional knowledges connected to the notion of business express themselves in specific linguistic forms which are not still the ones used in corpuses. These differences of user languages within the same domain are due to the coexistence of diverse communities of practice. Through an empirical study led in the nuclear engineering on the domain of severe accidents, we analyze how cultures, languages and specialized business knowledges are formed. On this example, we show how the wealth of these terminological varieties within the same domain can be taken into account in a system valuing a common conceptualization.

1 Contexte

1.1 Vers une hétérogénéité linguistique

Dans les institutions, les paysages linguistiques s'élargissent et se complexifient, créant de nouveaux besoins en termes d'accès aux savoirs. Les contextes induisant ces nouveaux besoins sont liés à des changements de modes de travail. Premièrement, la globalisation ouvrant les portes à la multiplication des partenariats internationaux et transferts de technologies, confronte subitement les institutions à la nécessité de passer de systèmes monolingues et autonomes à une communication portée par une grande variété linguistique et véhiculée par une diversité de médias. Dans notre contexte d'ingénierie nucléaire, la structure a pu jusqu'à présent s'appuyer sur une communication quasi autarcique, mais la pression actuelle des marchés l'oblige à revoir ses schémas traditionnels. L'une des unités de la structure est aujourd'hui scindée entre la région parisienne et Londres. La sûreté nucléaire est régie au plan international avec l'AIEA (Agence Internationale de l'Energie Atomique). Les

nouveaux projets de construction ne se font plus à l'échelle nationale mais européenne comme dans le cas de l'EPR (European Pressurized Reactor) où des acteurs anglais, espagnols, italiens etc., participent à la conception des centrales. Nous citerons enfin les partenariats avec les autorités locales sur les grands travaux internationaux de construction (ex : Chine) ou déconstruction (ex : Russie). Mais surtout, la structure est confrontée à une problématique de diversité langagière entre communautés de métier. Le travail par projets s'oppose à l'organisation historique de l'entreprise par métiers. D'un travail géré de façon verticale, l'entreprise est passée à une gestion transversale engageant divers métiers de façon synchrone sur un projet commun. Cela implique un nouveau mode de construction des compétences incluant une vision toujours plus globale d'un projet. Les ingénieurs sont donc fortement mobiles ou « *mouvmentables* » pour reprendre la terminologie de l'entreprise.

A ces raisons s'ajoute un important renouvellement générationnel avec de nouveaux arrivants apportant une culture axée sur la médiation et le travail en réseau. Ces réseaux dépassent les frontières de l'institution et du pays, pour impliquer une communication entre des communautés de pratiques et linguistiques très hétérogènes. Ainsi, dans le domaine des accidents graves qui nous concerne, la recherche et développement s'organisent au-delà de l'entreprise. Nous citerons le réseau Sarnet (Severe accident research network) de la Commission Européenne, sur les problématiques relatives à la fusion du cœur, qui regroupe 43 organismes impliqués dans la recherche sur la sûreté des réacteurs nucléaires (instituts de sûreté, universités, industriels, centres de recherche ...), dans 18 pays européens ainsi que les États-Unis, le Canada, la Corée du Sud et l'Inde. Mais de façon globale, la recherche et développement se construisent à travers des partenariats, comme par exemple avec l'IRSN (Institut de Radioprotection et de Sûreté Nucléaire) et le CEA (Commissariat à l'Energie Atomique), ou en inter-entreprise avec d'autres corps de métier. Les diversités des langues d'usages entre les communautés du domaine y sont très marquées, freinant l'exploitation de l'information et la mobilisation des connaissances.

1.2 De la recherche de document à la recherche d'information

Une longue période d'informatisation à outrance des systèmes d'information a laissé croire qu'elle accompagnait ces évolutions. Mais elle n'a souvent abouti qu'à un éparpillement des ressources et des moyens (Lainé-Cruzel, 2006), ignorant les usages et le partage des savoirs de spécialité. Aujourd'hui s'impose le constat de multiples systèmes d'information figés coexistant dans l'environnement professionnel. Nos travaux dans

l'ingénierie nucléaire montrent l'exemple de l'un de ces systèmes (Djambian, 2011, 2010). La ligne verticale d'organisation traditionnelle de l'entreprise par métiers se retrouve dans une gestion de l'information diachronique suivant une autorité descendante extrêmement lourde et cloisonnée, encadrée par des obligations réglementaires et leur cohorte de procédures et mesures de traçabilité encadrant une documentation de conception excessivement importante (on compte à ce jour 10 millions de documents dans la GED (Gestion Electronique des Documents) de l'ingénierie nucléaire, certains d'entre eux comme les « *Rapports de sûreté* » faisant plus de mille pages). L'organisation par projets, les habitudes communicationnelles des nouvelles générations, aidées par un contexte d'expertise généralisé des métiers ont poussé à la création d'un système d'information parallèle dit simultané ou synchronique. Orienté métier, il répond aux exigences de concertation, de travail en temps réel et aux nouvelles pratiques. Les communautés de pratique ont ainsi créé une multitude d'outils : 400 bases de données sous des formats variés (Notes, Oracle, Access, Excel, Filemaker,...) et pas moins de 162 intranets pour une structure de 3500 personnes.

La demande actuelle vise la recherche d'information plus que de documents. Cela implique de passer d'un mode de recherche documentaire prépondérant dans le système diachronique (avec une recherche d'informations secondaires comme une référence de document, pour aboutir à l'information primaire), vers un mode de recherche contextuel plus proche des pratiques des métiers (partant d'un mot ou groupe de mots pour aboutir à un texte), et par la même occasion d'optimiser l'appariement entre requête et information. Il est irréaliste de proposer aux utilisateurs non professionnels un langage documentaire, ceux-ci se contentant généralement de requêtes d'un ou quelques mots. Or, l'ambiguïté et la redondance du langage complexifiées par la variété des langues d'usages font que les résultats d'une recherche d'information sont rarement satisfaisants et nécessitent un « *processus itératif ou incrémental* » avec des affinements et filtrages réalisés par l'utilisateur. Il convient d'« *offrir à l'utilisateur une alternative au paradigme question/réponse* » (Simoni & Fluhr, 1999). Le but est donc d'aller vers un modèle plus adapté aux tâches des métiers, qui permettra de passer de la fourniture d'une information brute à une information de niveau plus élevé, potentiellement implicite et utile (Polanco, 1999).

Il faut alors se diriger vers un système réduisant les difficultés dues à l'ambiguïté et à la redondance de la langue, décuplées par les situations de diversités langagières entre communautés de pratique, mais au-delà, axer nos réflexions sur « *comment faire ressortir et utiliser le sens donné à l'information par le métier qui l'emploie ?* ».

2 Le savoir dans l'information

2.1 La diffusion du savoir dans le collectif

Le savoir collectif spécialisé, proche de ce qu'Aristote nomme l'« *art* » (Aristote, 1986) est cette culture collective construite et validée au sein de l'institution et sur laquelle le sujet s'appuie pour authentifier son discours dans le champ social. Le savoir est approprié par et attribué à un sujet. Il s'agit dans notre travail d'avoir conscience de ces trois étapes où le savoir part d'un sujet pour rentrer dans le domaine institutionnel sous forme de savoir collectif et être ensuite réapproprié par un tiers. Dans la longue tradition de compagnonnage dont faisait l'objet la transmission de savoir au sein de l'institution qui nous intéresse, le savoir n'était pas médiaté et passait directement d'une personne à l'autre. Cette tradition étant remise en cause, la transmission du savoir repose sur cette étape transitoire, jusque là inexploitée, de passage dans l'espace public sous forme de savoir collectif décorrélé d'un sujet, avant d'être réappropriable par un autre.

X. Polanco (Polanco, 1999) déclare, « *nous ne nous intéressons pas à la connaissance en tant que savoir-faire ou expertise des sujets connaissant... Dans notre cas, l'acquisition des connaissances ne se réalise pas à partir d'un expert ou de plusieurs, mais de la connaissance existante dans les bases de données, sorte de mémoire exosomatique... La connaissance que l'ingénieur a ici à traiter n'est pas de la connaissance en action à travers l'expertise des sujets connaissant, mais de la connaissance produite par eux et stockée dans les bases de données.* » C'est ce qui est appelé de la connaissance représentée sans sujet connaissant (Popper, 1979), car enfouie dans la documentation technique et traduite en langage écrit. Ainsi, la question de l'identification des partenaires de l'échange est remise en cause par ces mutations internes : les échanges très informels qui avaient cours jusqu'alors grâce à une identification claire des partenaires de la communication, sont troublés et amenés progressivement vers des modes de communication de plus en plus médiatés. C'est toute une culture, des structures d'échange et de transmission des savoirs établies au sein d'une communauté d'usage, qui sont modifiées.

Nos travaux s'appuient donc sur le savoir médiaté par les textes techniques produits par les ingénieurs. Sa situation collective sous-entend un consensus fort sur les références qu'il véhicule. C'est donc lui qui porte les références ayant la plus haute valeur, puisque fondées sur des usages sociaux et ayant acquis une légitimité institutionnelle au cours de ces usages (construction et stabilisation au cours du temps). En cela même qu'il est du domaine collectif, ce savoir est basé sur le principe de diffusion au plus grand nombre. Outre les références, il

est aussi porteur des règles établies de transmission de ces références, également fixées par l'usage social. Le fait que ces références et leurs normes soient portées dans le discours du savoir collectif, démontre qu'elles ont été éprouvées dans leur potentiel de réappropriabilité. C'est ce qui fait leur intérêt.

2.2 Les langages de spécialité

Tout ce processus de transmission, diffusion du savoir, et de ses références collectives, fait évidemment appel au langage. Et la mise en œuvre des langages passe par l'énonciation, qui couvre deux caractéristiques : l'individualisation, partant d'un sujet et reçue par un autre ; l'actualisation, puisqu'émetteur et récepteur lui donnent sens par des références à leur réel. Puisque le langage est relatif à un réel donné, il est représentatif de la structure sociale dans laquelle il a été énoncé, matérialisant l'information qui y circule. Si le savoir représente la trace d'un champ social, le langage est la cristallisation de cette trace.

Selon les situations de communication durant lesquelles ils sont actés, les langages de l'information peuvent prendre trois formes. *Les langages de représentation* rendent compte du réel, sous la forme de la description (ex : dans le discours scientifique), ou pour aller vers un consensus symbolique. Cela peut être sous la forme de l'appropriation de l'information par les partenaires de l'échange, qui vont faire émerger l'opinion et le savoir, ou sous la forme de l'idéologie (qui nous intéresse moins ici), dans un cadre politique, de choix ou d'engagement. *Les langages d'opération* sont les principaux langages que nous trouverons dans l'entreprise et qui s'expriment dans les textes techniques. Ils font de la référence un outil opératoire puisqu'ils visent l'obtention de résultats et de nouvelles références. Ce sont donc, par essence, les langages de la constitution du savoir collectif. On peut alors distinguer les langages qui produisent de l'action (les langages de stratégie), de ceux qui produisent du savoir. Puisque nous nous intéressons aux métiers techniques et à leurs échanges et productions d'information, nous retrouverons essentiellement ces derniers. Enfin, *les langages de symbolisation* produisent des systèmes de symboles et de représentation. Nous nous y intéresserons peu, bien qu'ils soient extrêmement présents dans les documents techniques.

3 Capter le savoir spécialisé par le langage

Si nous nous intéressons au langage, nous devons le prendre dans toute sa complexité. Il est vrai que, comme le note Bertrand Russell, dans un langage idéal il ne devrait y avoir qu'un seul mot pour un seul objet et tout objet complexe serait exprimé par une combinaison de

mots, chacun pour chaque caractéristique de l'objet (Russell, 1918-19). Mais Démocrite notait déjà dans l'antiquité que : 1) différents objets sont souvent désignés par le même nom ; 2) le même objet est souvent désigné par des noms différents ; 3) les noms désignant un objet peuvent varier dans le temps ; 4) les raisons pour lesquelles les noms sont liés aux objets présentent une grande variété (Diels, 1903). Nous pouvons ajouter : le même objet est désigné par différents termes dans différents langages.

Or, « *il existe en effet des domaines scientifiques et techniques qui nécessitent une conceptualisation du monde et la création de dénominations univoques de ses constituants* » (Roche, 2007), soit « *un moyen d'expression qui permette à la fois de prévenir les erreurs d'interprétation et d'empêcher les fautes de raisonnement* » (Frege, 1971). L'ingénierie nucléaire en fait partie. Notre but est donc de capter le savoir spécialisé à travers les langages des différentes communautés du domaine, ce qui nous conduit vers un travail terminologique. La terminologie représente le savoir spécialisé, en fixe le langage, et assure par des références stables une communication fiable. Le consensus autour de cette terminologie liée à un contexte spécifique assure une bonne transmission et appropriation de l'information et des savoirs. Mais pourra-t-on s'arrêter aux mots ? Comme Aristote le notait, dans la relation entre l'objet désigné (la référence) et le terme, on se doit d'introduire la notion de concept. Il entretient une relation naturelle avec l'objet, alors que sa relation au terme est purement conventionnelle. L'épistémologiste Chrysippos de l'école stoïcienne remarquait également la distinction et le lien entre le signe, la signification et la réalité (Heinze, 1880 ; Mates, 1961). Cette relation triptyque est reprise par nombre d'auteurs (Frege, 1892 ; Peirce, 1902 ; Gomperz, 1905 ; Saussure 1916 ; Dewey, 1925...), dont Morris qui établit que chaque signe a un « *significatum* » (concept) et un « *denotatum* », ou référence (Morris, 1938). Descartes et les auteurs de la logique de Port Royal (Arnaud & Lancelot, 1660 ; Arnaud & Nicole, 1662) soutenaient d'autre part que le langage reflète parfaitement la pensée. Il est donc un lien indéfectible entre terme et concept. On peut citer Saussure : « *La pensée, chaotique de sa nature, est forcée de se préciser en se décomposant* » (Saussure 1916) et Sapir « *le concept n'acquiert une vie distincte et indépendante que lorsqu'il a trouvé une enveloppe linguistique bien à lui* » (Sapir 1921).

Le travail terminologique est donc la première étape du travail ontologique. Notre contexte est d'autant plus intéressant que les langages scientifiques produisent des terminologies moins équivoques que les langages ordinaires : les liens du concept au terme y sont plus rigides (Marradi, 2012). Nous présentons ici la mise en place d'une base de connaissances métier d'un domaine circonscrit de l'ingénierie nucléaire, composée d'une

terminologie et d'une ontologie centrées sur le *sens métier* (Djambian, 2010). La description de nos travaux suivra trois niveaux de langage (Roche, 2007) : la langue d'intellection où les noms des concepts reflètent le système notionnel indépendamment des contingences liées aux usages ; les langues d'usage où les mots sont pris en compte de sorte à être reliés aux désignations du vocabulaire normé et à leurs différents contextes d'utilisation ; enfin la langue de représentation qui permet l'exploitation sous forme d'ontologie.

3.1 Le domaine et son corpus

Les « Accidents Graves » sont un domaine jeune issu de la « Sûreté Nucléaire » et où les besoins de réappropriation des connaissances sont importants. En partant de ce domaine restreint, nous souhaitons démontrer ce qui pouvait être réalisable et élargi ultérieurement. Les textes techniques étant un socle stable pour construire la Base de Connaissances métier, notre méthode part comme de nombreuses autres, de l'extraction à base de textes. Notre sélection du corpus a été réalisée avec l'aide de spécialistes, dans le souci de respecter l'application visée, une couverture maximale du domaine, l'homogénéité du corpus (du point de vue des sources et de l'usage de la langue et des concepts) et la légitimité accordée par les métiers. Notre expert a choisi huit documents qu'il estimait couvrir 90% des concepts métiers, soit un corpus de 570 pages. Ce corpus contient des notes d'études de Recherche et Développement, stratégiques ou de procédés.

3.2 Les lexiques : extractions à base du corpus de textes

L'analyse linguistique du corpus a été réalisée par l'Equipe Condillac avec l'outil LCW. Dans un premier temps, un lexique de 20 026 syntagmes nominaux a été généré. Les termes ayant une fréquence d'apparition dans le corpus inférieure ou égale à 10 n'étant pas conservés, le lexique se composait au final de 770 syntagmes nominaux. Mais les acronymes n'ayant pas été pris en compte, des notions cœurs du domaine (concernant les projets, outils, matériels, phénomènes physiques...), étant exprimées par de longues propositions, étaient éliminées. Par exemple: EPS (Etude Probabiliste de Sûreté), GIAG (Guide d'Intervention en Accident Grave), circuits RIS-EAS (injection de sécurité - aspersion enceinte) ...). Le cas des expressions contenant des chiffres ou caractères spéciaux était identique : mode α , filtre U5, Bugey post-VD3, WASH 1400, dispositifs H4-U3 ... Or ces expressions forment un moule dans lequel l'argument est figé, ce qui le rend plus solide, plus facile à comprendre et donc plus ré-appropriable par divers interlocuteurs.

Nous avons donc décidé de générer un nouveau lexique contenant uniquement les termes en majuscule, en conservant les chiffres. Ce deuxième lexique de 1710 acronymes a été extrait à partir du même corpus. Tous les niveaux d'occurrence ont été conservés. Notons que, si le lexique des syntagmes nominaux fait appel à des notions globales ou stratégiques, celui des acronymes évoque des notions de « terrain » et nomme essentiellement des matériels, actions, processus, projets... Deux tris manuels des résultats des traitements automatiques ont alors été réalisés pour obtenir 746 syntagmes nominaux et 1298 acronymes. Deux experts sont intervenus pour la validation, aboutissant à un lexique de 344 syntagmes nominaux (44% du lexique original d'occurrences inférieures à 10) et un de 343 acronymes (20% du lexique original). Ceci prouve qu'un travail purement sémasiologique n'aurait en aucun cas pu être pertinent. Puisqu'« *aucun corpus textuel n'explique toutes les connaissances que sa lecture présuppose* » (Lerat, 2009), le traitement des ressources spécialisées requiert obligatoirement des ressources externes (L'Homme, 2008).

3.3 La langue d'intellection : construction des réseaux conceptuels

Nous avons tenté de réaliser une ébauche de réseau lexical à partir de ces lexiques, à base de relations linguistiques (synonymie, d'hyponymie, ...). Mais un tel réseau est devenu très vite inextricable. Pourquoi ? « *Il convient de se rappeler que tout travail terminologique devrait être fondé sur les notions et non sur les termes* » (Felber, 1984). C. Roche rappelle que l'« *on identifie aujourd'hui trop souvent la terminologie à sa manifestation langagière [...] en oubliant que la conceptualisation et la représentation des objets du monde sont des activités centrales, si ce n'est les premières de la terminologie* » (Roche, 2007). Face à la difficulté d'élaborer une structure lexicale directement à partir des mots du domaine, nous avons pris le parti de sortir de la langue et de nous diriger directement vers le conceptuel, conservant ce « *souci constant de s'extraire de tout discours pour se référer à un « socle » stable de connaissances* » (Roche, 2007), cette part extralinguistique, externe et complémentaire aux textes.

Lorsque l'on demande à un expert de représenter son domaine, il organise des concepts en séquences. Nous avons suivi cette logique en dessinant progressivement les grandes notions du domaine pour les identifier et les classer dans des modélisations simples autour du processus du déroulement d'un Accident Grave, d'un point de vue phénoménologique. Cinq vues ont ainsi été créées : déroulement d'un accident grave, événements initiateurs, perte du confinement, rejets radioactifs, dispositions accidents graves. Nous avons pris soin de bien

distinguer les termes d'usages, des termes normés et des dénominations des concepts pour nous concentrer dès cette étape sur le niveau conceptuel en sortant de la langue des lexiques. Ce travail a été réalisé dans des conditions identiques à l'expérience d'A. Busnel et N. Aussenac-Gilles, où les ingénieurs des connaissances n'avaient pas de culture a priori du domaine à traiter et s'y sont immergés par « *une analyse systématique et approfondie des textes* » (Bourigault, 2004). Nous n'avons cependant pas complété cette imprégnation du contexte par l'utilisation d'outils de TAL (Traitement Automatique des Langues), mais par des échanges répétés avec les ingénieurs.

Par exemple, dans le réseau conceptuel portant sur les *rejets radioactifs*, de nombreux noms de concepts, essentiellement de niveau inférieur, apparaissaient dans les textes. Ces concepts ont des référents concrets et perceptibles et ont donc des dénominations plus claires. La difficulté s'est posée sur leur choix, puisque mêlés à d'autres dénominations relevant de la langue d'usage, et sur leur assemblage logique (hiérarchisation et relations). Ainsi, les rejets environnementaux *terme source S3* sont présents dans les textes sous les formes : *rejet S3*, *activité S3*, *critère S3*, *terme source S3*, *rejet S'3*, *terme source*.

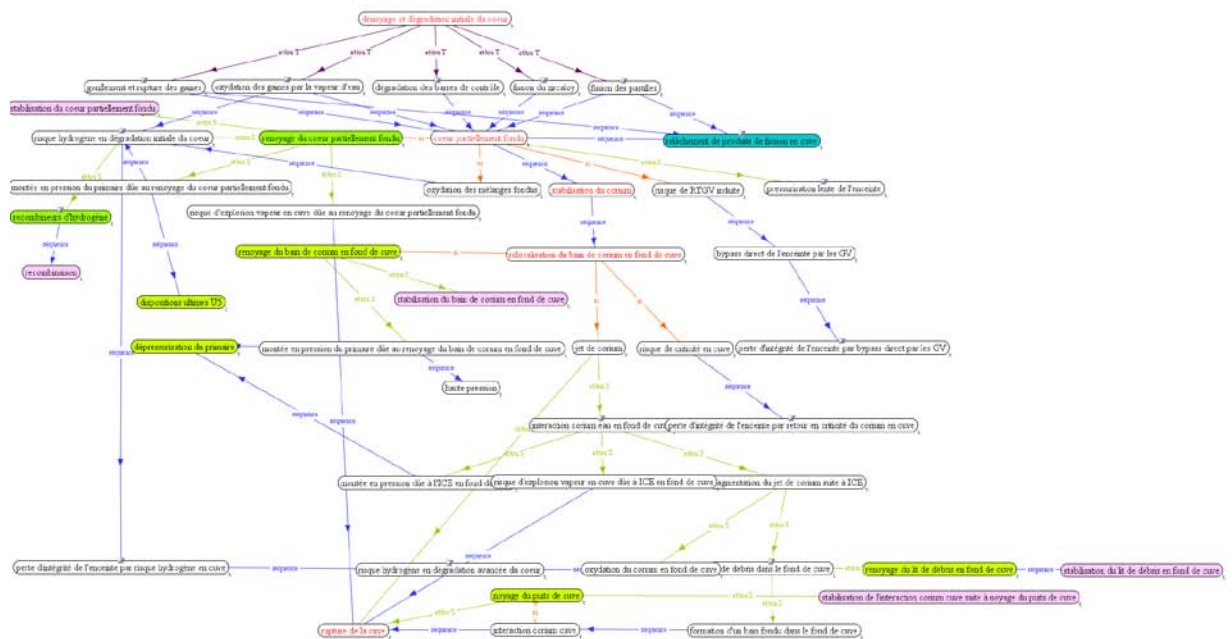


Fig. 1 – Ensemble « dénoyage et dégradation initiale en cuve », extrait du réseau conceptuel « déroulement d'un accident grave ».

L'ensemble *dénoyage et dégradation initiale en cuve* présenté ci-dessus n'est qu'une partie du réseau conceptuel complet *déroulement d'un accident grave*. Les relations y sont complexes. L'apparition des noms des concepts était parcellaire dans les textes et des créations de dénominations artificielles mais parlantes pour les métiers ont du être réalisées. Des expressions telles que *jet de corium*, *risque de RTGV induite*, *oxydation des mélanges*

fondus, pouvaient être reprises telles quelles des textes car non ambigus. Mais de nombreux termes sont utilisés de façon indifférenciée dans les parler métier, seul le contexte pouvant permettre le lien à la référence. Par exemple l'expression *montée en pression* a nécessité sur ce seul ensemble *en cuve*, la création de diverses dénominations pour lever les ambiguïtés d'usage : *montée en pression du primaire due au renoyage du cœur partiellement fondu*, *montée en pression du primaire due au renoyage du corium en fond de cuve*, *montée en pression due à l'ICE (interaction corium-eau) en fond de cuve...*

A ce stade, nous avons fait valider le travail par l'expert. La vision générale du déroulement d'un Accident Grave lui a permis de passer rapidement l'ensemble des concepts en revue. L'expert a conservé les dénominations que nous avons choisies, les jugeant plus significatives, ce qui signifie que ces notions lui parlaient indépendamment de ses usages. Nous en concluons que si le système conceptuel est établi, le système linguistique est lui mouvant et multiple et « *le constat de la variabilité des terminologies s'impose : étant donné un domaine d'activité, il n'y a pas une terminologie, qui représenterait le savoir sur le domaine* » (Bourigault, 2004).

3.4 La langue de représentation : construction de l'ontologie

La structure conceptuelle construite à partir d'un corpus est une bonne base de travail pour la construction d'une ontologie, mais elle ne peut constituer une ontologie en elle-même. Le passage de l'une à l'autre nécessite de lourds remaniements, les connaissances extralinguistiques permettant de lier les deux plans. Ainsi chaque aspect de l'intension d'un concept (attributs lui étant inhérents) peut être hiérarchisé, chaque aspect du concept donnant une branche de l'arbre (Marradi, 2012). Puisque nous étions à ce stade déjà dans le conceptuel, nous avons directement clarifié les relations multiples du réseau conceptuel et restructuré tous les concepts en ensembles sémantiquement liés par des relations de subsomption. Pour catégoriser essentiellement les phénomènes physiques intervenant lors d'un Accident Grave, nous avons dû créer de nouveaux concepts génériques qui n'apparaissaient alors pas dans les réseaux conceptuels. Les treize sous-ontologies obtenues ont été construites selon la méthode de différenciation spécifique inspirée des arbres de Porphyre¹ et converties en OWL (Ontology Web Language) avec l'outil OCW de l'Equipe

¹ Du philosophe néoplatonicien du III^es, Porphyre, auteur d'une introduction aux Catégories d'Aristote (l'Isagoge). Table de la coordination et de la subordination des genres et espèces, construite en partant de la substance et descendant jusqu'aux espèces particulières suivant le principe de dichotomie. La construction logique se développe comme les branches d'un arbre.

Condillac. La validation de l'expert a permis de vérifier l'organisation des relations et la présence de l'ensemble des concepts précédemment identifiés dans les réseaux conceptuels.

Illustrons avec le détail de la création de l'ontologie « événements initiateurs ». A partir du réseau conceptuel (figure 2), la hiérarchisation et les relations ont du être complétées pour que chaque dénomination de concept soit suffisamment explicite afin d'être situable dans l'ensemble de l'ontologie. Ainsi, les concepts intermédiaires *accident de perte*, *accident de rupture* et *accident de transitoire* ont été ajoutés dans la hiérarchie. En figure 3, la portion *accidents de rupture* démontre le mode de construction par différenciation spécifique.

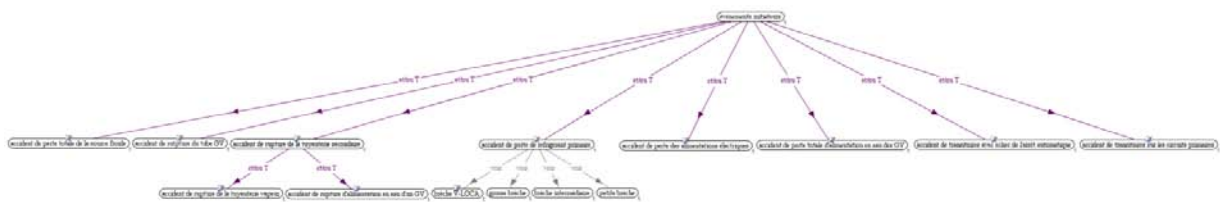


Fig. 2 – Réseau conceptuel «événement initiateur».

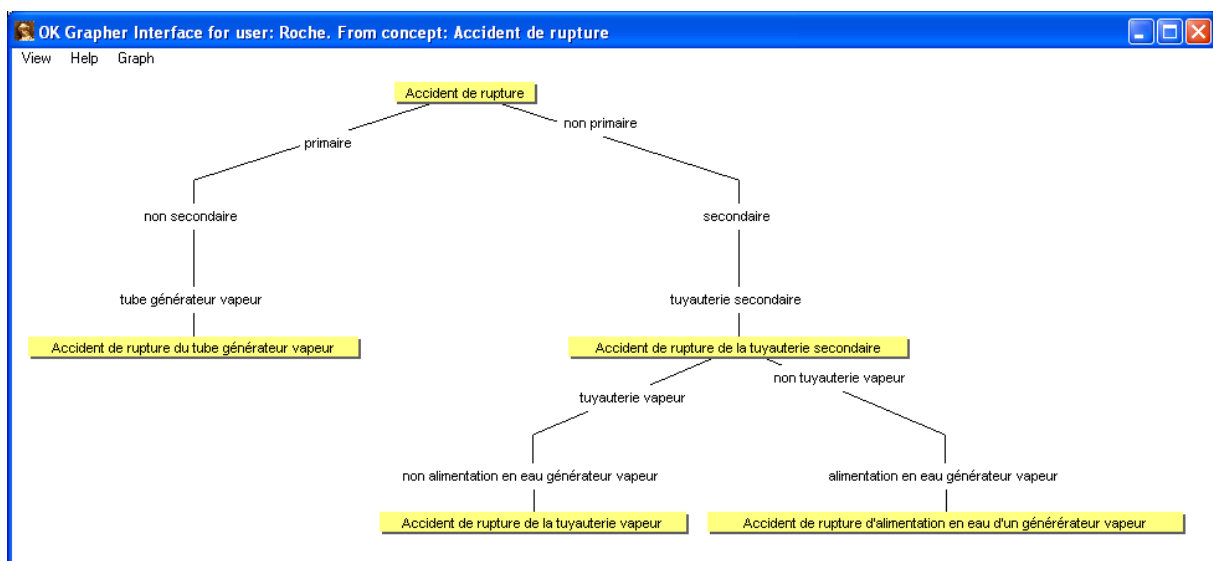


Fig. 3 – détail de la construction de l'ontologie «événement initiateur» : les « accidents de rupture »

3.5 Les langues d'usage : définition de la terminologie

« les ontologies sont des réseaux de concepts et n'ont donc pas de signifiants. Comment alors les articuler aux langues ? Ici intervient l'articulation de la synonymie et de l'hyponymie On passe ainsi des signes aux choses (par la référence) puis aux concepts

(par la *subsumption*) » (Rastier, 2004). Il s'agit ici de retourner aux mots des lexiques pour exploiter les relations d'hyperonymie et synonymie permettant d'articuler extralinguistique et linguistique. Cette étape est également nécessaire pour faire le lien entre l'ontologie et les documents et permettre la recherche et gestion de la documentation technique. La définition de la terminologie consiste à associer chaque terme des lexiques aux concepts de l'ontologie, identifier les synonymes, les acronymes et leurs développés, etc.

progression de l'AG hors cuve	puits de cuve	TOLBIAC																		
progression du corium dans le puits de cuve	explosion																			
interaction corium-eau dans le puits de cuve	vapeur	ICE	interaction	corium-eau	interaction															
lit de débris dans le puits de cuve	assèchement	débris		solidification	lit de débris															
retour en criticité du corium hors cuve	criticité																			
interaction corium-béton	ICB	interaction	corium-béton	interaction	mode epsilon	mode ϵ	ablation	ablation	béton	du érosion	du radier	percée								
interaction corium-béton sous eau	ICB	interaction	corium-béton	configuration	de bain	PERCOLA														
interaction corium-béton à sec	ICB	interaction	corium-béton	configuration	de bain															
récupérateur de corium EPR	pic	récupérateur	de corium	core catcher	récupération	récupérateur	EPR	écart EPR	céramique	interaction	corium-	céramique	CORQUENCH							
échauffement direct de l'enceinte du à la haute pression	pic	combustion	DCH	Direct	Containment	Heating	EDE	Echauffement	Direct	de	échauffement	direct	HPME	High Pressure	Melt	Expulsion				
perte d'intégrité du puits de cuve	puits de cuve	TOLBIAC																		
risque hydrogène	H2	hydrogène	mode gamma	mode γ	production	d'hydrogène	pic	Shapiro	explosion	d'hydrogène	accélération	de	flamme							
risque hydrogène en cuve																				
risque hydrogène hors cuve	accélération	de flamme	combustion	combustion	d'hydrogène	combustion	de	combustion	de	combustion	de	monoxyde	de	concentration	hydrogène	distribution	de	l'hydrogène	répartition	de la
explosion vapeur	mode alpha	explosion	vapeur	explosion	de	explosion-	vapeur	explosion	de	explosion-	vapeur	PM	pré-mélange	PREMELANGE						
explosion vapeur en cuve	explosion	vapeur	explosion	vapeur	explosion	vapeur														

Fig. 4 – Extrait de la terminologie

Comme le rappelle Phillips, il n'y a pas de garantie a priori que deux membres de la même communauté linguistique utilisent le même mot avec le même sens dans toutes les circonstances (Phillips, 1977). Cette variété des langues d'usages, qui était au départ une entrave à l'accès aux savoirs techniques, devient en réalité une manne à exploiter pour la recherche et la gestion de l'information. Par exemple, pour le seul concept d'*interaction corium-béton*, sont identifiés dans les lexiques les termes : *ICB*, *interaction corium-béton*, *interaction*, *mode epsilon*, *mode ϵ* , *ablation*, *ablation du béton*, *érosion du radier*, *percée*, *percement*, *percée du radier*, *percement du radier*, *traversée du radier*, *radier*, *dissolution*, *configuration de bain*, *étalement*, *refroidissabilité*, *refroidissable*, *siliceux*, *silico-calcaire*, *béton*, *béton sacrificiel*, *calcaire*, *béton IET*, *béton Integral Effect Test*, *combustion du monoxyde de carbone*, *ABI*, *ARTEMIS*, *CORCON*, *GEMINI*, *MEDICIS*, *TOLBIAC-ICB*, *BETA*, *CLARA*, *COMET-L*, *COMET*, *CRATHER*, *VULCANO*. Cette terminologie a fait l'objet d'une validation par l'expert. Le résultat a été converti au format XML, compatible avec la majorité des outils de recherche et gestion de l'information actuels.

A l'issue des travaux, une plateforme de test a été mise à disposition des métiers. La Base de connaissances Accidents Graves (terminologie et ontologie) a ainsi été appliquée à la

gestion de la documentation technique de référence du domaine et à la représentation de ses connaissances. Les documents sont indexés sur l'ontologie en prenant en compte les variations d'usage des termes entre communautés de pratiques, permettant une recherche par concepts ou langage naturel avec expansion de requêtes. Par exemple, pour exprimer le concept d'*échauffement direct de l'enceinte du à la haute pression*, les métiers utilisent indistinctement : *DCH, Direct Containment Heating, EDE, Echauffement Direct de l'Enceinte, échauffement direct*. Or, l'acronyme EDE désigne également l'*espace entre enceinte*, qui ne fait pas partie du domaine des accidents graves. L'outil signalera donc l'ambiguïté entre *espace entre enceinte* et *Echauffement Direct de l'Enceinte* et les résultats concerneront l'ensemble des termes associés à ce dernier concept, à savoir : *pic, combustion, DCH, Direct Containment Heating, EDE, Echauffement Direct de l'Enciente, échauffement direct, HPME, Hight Pressure Melt Expulsion, dispersion, inflammation, CPA, RUPUICUV, CONTAIN, programme ANL, essai DISCO, DISCO-C, DISCO-H*.

Conclusion

Dans le cadre de transfert de savoir, la participation du langage à l'expérience n'est plus à prouver. Le langage classe l'expérience, l'organisant dans des catégories où elle prend sens, non seulement pour qui a vécu l'expérience, mais aussi pour les autres (Berger & Luckmann, 1966). Une fois formulées dans un code commun, les pensées peuvent être communiquées : termes et expressions sont l'aspect public du concept. La présence d'un code linguistique commun rend possible un extraordinaire enrichissement des pensées qui peuvent être communiquées. C'est à nous à en faciliter la communication et à s'assurer que le code commun soit partagé et approprié par le plus grand nombre au sein d'une communauté. Les termes vont alors agir comme des ancrages, points de départ de la conceptualisation du flot d'expérience. "*Language is not a mere inventory ... it is a creative and self-sufficient symbolic organization... which defines experience for us... Meanings are not discovered in experience but rather imposed upon it...*" (Sapir, 1931).

Bibliographie

- Aristote, *Métaphysique*, Paris: Vrin, 1986.
- Bourigault D., N. Aussenac-Gilles and J. Charlet, « Construction de ressources terminologiques ou ontologiques à partir de textes : un cadre unificateur pour trois études de cas », *Revue d'Intelligence Artificielle*, 18, 2004.
- Dewey J., *Experience and nature*, Chicago, Open Court, 1925.
- Diels H., *Die fragmente der Vorsokratiker griechisch und deutsch*. Berlin, Weidmann, 1903.
- Djambian C., Le métier : son savoir, son parler, *Actes TOTH 2011*, Annecy: Institut Porphyre, 2011.
- Djambian C., Valorisation d'un patrimoine documentaire industriel et évolution vers un système de gestion des connaissances orienté métiers, Thèse de doctorat, Université Jean Moulin Lyon3, 2010.
- Felber H., *Manuel de terminologie*, Unesco, Paris, 1984.
- Frege G., "Über Sinn und Bedeutung", In *Zeitschrift für Philosophie und philosophische Kritik*, C: 25-50, 1892.
- Frege G., *Ecrits logiques et philosophiques*, Seuil, Paris, 1971.
- Heinze M., *Zur Erkenntnislehre der Stoiker*, Leipzig, Edelmann, 1880.
- Lainé-Cruzet S., Appropriation, mutualisation, expérimentations des technologies de l'information scientifique et technique, *Revue AMETIST*, n° 0, 2006.
- Lamizet B., *Les lieux de la communication*, Liège : Mardaga, coll. « Philosophie et langage », 1992.
- Lerat P., « Propositions pour un réseau conceptuel des instruments de mesure ontologiques », *Proceedings of the Second International TOTH Conference, 5-6 June 2008, Annecy, France*, Annecy, Institut Porphyre, 73-90, 2008.
- L'Homme M. C., « Ressources lexicales, terminologiques et ontologiques : une analyse comparative dans le domaine de l'informatique », *Revue française de linguistique appliquée*, 13, 1, p. 87-118, 2008.
- Marradi A., "The concept of concept: Concepts and terms", *Knowledge organization*, 39, 1, p. 29-54, 2012.
- Mates B., *Stoic logic*, Berkeley, University of California Press, 1961.
- Morris C. W., *Foundations of the theory of signs*, Chicago, University of Chicago Press, 1938.
- Peirce C. S., "Sign", in Baldwin, James M., ed., *Dictionary of philosophy and psychology*, vol. II. New York, Macmillan, p. 527-28, 1902.
- Phillips D. L., *Wittgenstein and the scientific knowledge*, London, Macmillan, 1977.
- Polanco, X., Extraction et modélisation des connaissances : une approche et ses technologies (EMCAT), *Organisation des connaissances en vue de leur intégration dans les systèmes de représentation et de recherche d'information*, Lille : Travaux et recherches, 1999.
- Popper, K., *Objective knowledge: an evolutionary approach*, Oxford University Press, 1979.
- Rastier, F., « Ontologie(s) », *Revue d'Intelligence Artificielle*, 18, 1, p. 15-40, 2004.
- Roche, C., Le terme et le concept : fondements d'une ontoterminologie. *Actes TOTH 2007*. Annecy: Institut Porphyre, 2007.
- Russell B., "The philosophy of logical atomism", *The monist* 28: 495-527 and 29: 32-63, (version réimprimée in Russell B., *Logic and Knowledge*, London, Allen and Unwin, 1956.), 1918-9.
- Sapir E. *Language: an introduction to the study of speech*, New York, Harcourt, Brace & World, 1921.
- Sapir E. *The southern Paiute language*, Boston, The American academy of arts and sciences, 1931.
- Saussure F. de., *Cours de linguistique générale*, Paris, Payot, 1916.
- Simoni J. L. and Fluhr C. « Accès à l'information à travers les graphes de termes », In Maniez J. and Mustafa El Hadi W. (dir.) *Organisation des connaissances en vue de leur intégration dans les systèmes de représentation et de recherche d'information*, Lille, Travaux et recherches, 1999.