

LE MODÈLE DE DONNÉES DU LOGICIEL *DHYDRO* APERÇU SYNTHÉTIQUE

À l'occasion de la *IX^e Université d'automne en terminologie*, nous avons présenté le tutoriel du logiciel de gestion terminologique *Dhydro*, en nous attardant tout particulièrement sur le modèle de données adopté. Le logiciel *Dhydro* a été conçu dans le cadre du programme européen *Multilingual Information Society* (MLIS) pour permettre aux rédacteurs du dictionnaire de l'Organisation hydrographique internationale d'alimenter une base de données commune au travers du réseau Internet¹. Au-delà d'une structure éditoriale propre au comité de rédaction du *Dictionnaire hydrographique*, le modèle suivi constitue un exemple typique d'approche conceptuelle des vocabulaires spécialisés multilingues et l'interface présente des caractéristiques ergonomiques intéressantes pour qui s'initie à la terminographie.

1 Le modèle de données conceptuel : vision pragmatique

Historiquement, la théorie de la terminologie a préféré le concept au signe linguistique. Nous ne reviendrons pas ici sur les nombreuses et pertinentes remises en cause d'une approche trop doctrinaire du modèle terminologique. Nous nous bornerons à observer qu'un fossé sépare souvent la prétention à décrire des concepts universels (pour autant que cela soit possible...) et la pratique observée dans de nombreuses terminographies, fussent-elles informatisées. Par souci de simplicité et de pragmatisme, nous nous contenterons d'intuitivement appréhender ici le concept, « comme l'espace de sens qui sert de terrain d'entente entre plusieurs langues » (Van Campenhoudt 2000 : 138).

Lors de la toute première université d'automne en terminologie, nous avons présenté le modèle de fiche terminologique adopté alors pour le logiciel *Termisti* (Merten, Mertens & Van Campenhoudt 1993). Dans cette communication, nous nous attachions déjà à montrer que dans le cadre d'une gestion informatisée, ce que l'on nomme *concept* (ou *notion*) est ce qui sert de point de transit garantissant une équivalence entre les termes des différentes langues.

Dans une telle perspective, les données situées au niveau dit *conceptuel* sont celles qui sont communes à l'ensemble des langues décrites (p.ex. un numéro d'ordre, une classification en domaines et sous-domaines, des informations administratives, des relations sémantiques avec d'autres concepts...). Tous les champs qui concernent la description du sens (la définition) et des termes synonymes doivent, eux, être explicitement rapportés à une langue particulière. Au sein de chaque langue, il importe, en outre, de distinguer les informations de nature sémantique de celles qui relèvent de la description de termes particuliers.

1. On trouvera une documentation détaillée sur le site officiel du projet : <http://www.loria.fr/projets/MLIS/DHYDRO/>. Les ambitions du projet ont été décrites à l'occasion de la 1^{re} *Conférence sur la coopération dans le domaine de la terminologie en Europe* (Blampain et al. 2000).

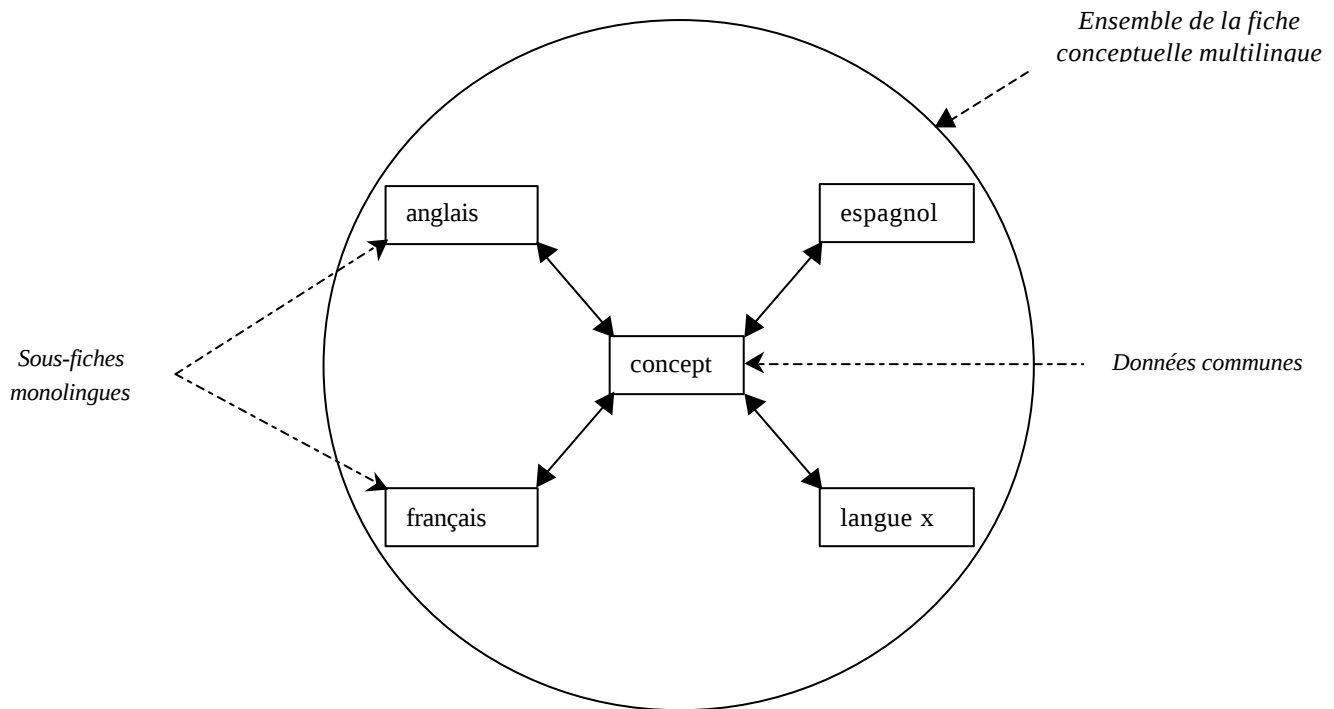


Schéma 1 : « concept » et fiche multilingue

La nécessité d'une telle structure - adoptée de longue date par le centre de recherche TERMISTI - est désormais largement reconnue : elle est plus ou moins clairement explicitée au point 8.1 de la norme ISO 12618 (1994) et est très heureusement utilisée dans le format d'échange GENETER. Au-delà de l'adoption de ce schéma, il importe également que les relations de dépendance entre les différents champs de la fiche soient explicites ; p.ex. que le champ « abréviation » soit explicitement raccroché à la forme complète du terme. À cet égard, il convient de noter que la norme ISO 12620 (1999) n'aborde guère cette question des dépendances, sinon par un classement général des informations parfois aisément critiquable.

Confrontés à la nécessité de concevoir une base de données susceptible d'accueillir un grand nombre de langues, les partenaires du projet DHYDRO ont adopté un modèle de fiche qui tente de répondre à ces exigences structurelles. Nous nous proposons de l'examiner plus avant.

2 La fiche terminologique de *Dhydro*

Le logiciel *Dhydro* n'autorise qu'un seul type de fiche : il n'est donc pas possible d'y concevoir sa propre fiche « maison ». Cette rigidité offre toutefois l'avantage de garantir que toutes les fiches suivent une même structure, ce qui est très important lorsqu'on veut échanger des données. Par ailleurs, on observera que le nombre de champs envisagés est de nature à satisfaire le plus grand nombre d'exigences et que la clarté du modèle autorise une grande souplesse d'utilisation (infinité potentielle de langues, de termes et de contextes). Dans la mesure où la structure présentée dans le schéma n° 1 est respectée, le modèle de données offre une garantie exceptionnelle de cohérence du contenu. À la différence de ce qui s'observe dans d'autres logiciels, il est, par exemple, impossible de donner des définitions différentes à deux termes présentés comme synonymes.

La place nous manque pour fournir ici le détail de chaque champ. Nous nous bornerons donc à un exposé des principes généraux, le lecteur ayant tout le loisir de consulter une documentation exhaustive sur le site du centre de recherche TERMISTI².

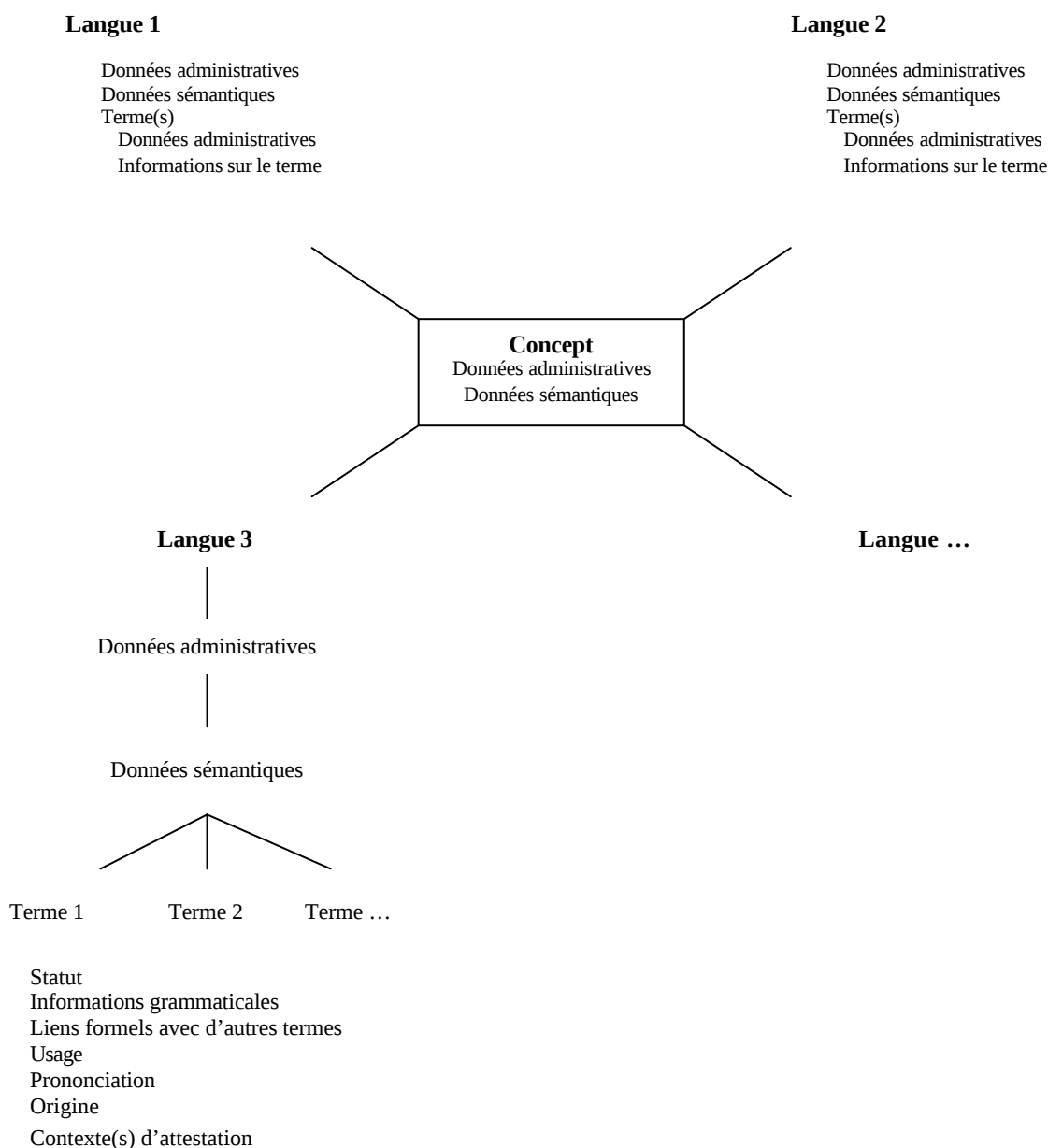


Schéma 2 : le modèle de données de la fiche « Dhydro »

2. http://www.termisti.refer.org/dhydro/aide_sommaire.html.

2.1 Description du niveau « conceptuel »

On trouve à ce niveau des informations administratives et sémantiques. Les premières sont constituées d'un identifiant d'entrée (n° de fiche), du statut de la fiche (à vérifier, validé ou retiré), de l'origine de l'information (fiche importée ou créée), du responsable de la fiche multilingue et d'une note éventuelle rédigée par celui-ci.

Les informations d'ordre sémantique qui se situent à ce niveau doivent être vraies quelle que soit la langue. Il s'agit de la classification en domaines et sous-domaines, d'une illustration éventuelle et des liens sémantiques entre concepts (espèce-genre, partie-tout, cause effet, opposition, lien indéterminé). Le principe général du schéma adopté veut que l'on situe toujours les informations au plus haut niveau possible dès lors qu'elles peuvent être héritées à un niveau inférieur : si le concept B est un type du concept A, cela doit être vrai pour l'ensemble des langues possédant des termes équivalents pour désigner lesdits concepts.

2.2 Description de chaque langue

2.2.1 LES INFORMATIONS COMMUNES À TOUS LES TERMES SYNONYMES

On retrouve d'abord des champs administratifs similaires à ceux observés au niveau du concept : en effet, dans le cadre d'un projet de coopération internationale, le rédacteur n'est pas nécessairement le même d'une langue à l'autre.

Les informations sémantiques (définition, note explicative et références bibliographiques s'y rapportant) sont situées à un niveau supérieur à celui des termes, car elles s'appliquent à tous les synonymes. Elles ne sont toutefois pas situées au niveau du concept, puisque leur formulation dépend d'une langue précise.

2.2.2 LES INFORMATIONS LEXICALES PROPRES À CHAQUE TERME

Il est possible de décrire un grand nombre de termes au travers d'une variété récurrente de champs : outre un statut de vedette ou de synonyme³, on peut enregistrer des informations grammaticales (catégorie, genre, nombre), la prononciation, l'étymologie, de multiples contextes d'attestation et leurs références bibliographiques, des informations sur l'usage (registre, fréquence, vieillissement, restriction géographique ou chronologique, note sur l'usage) et des liens lexicaux avec d'autres termes (p.ex. A est une abréviation de B).

3 Un fenêtrage conforme au modèle de données

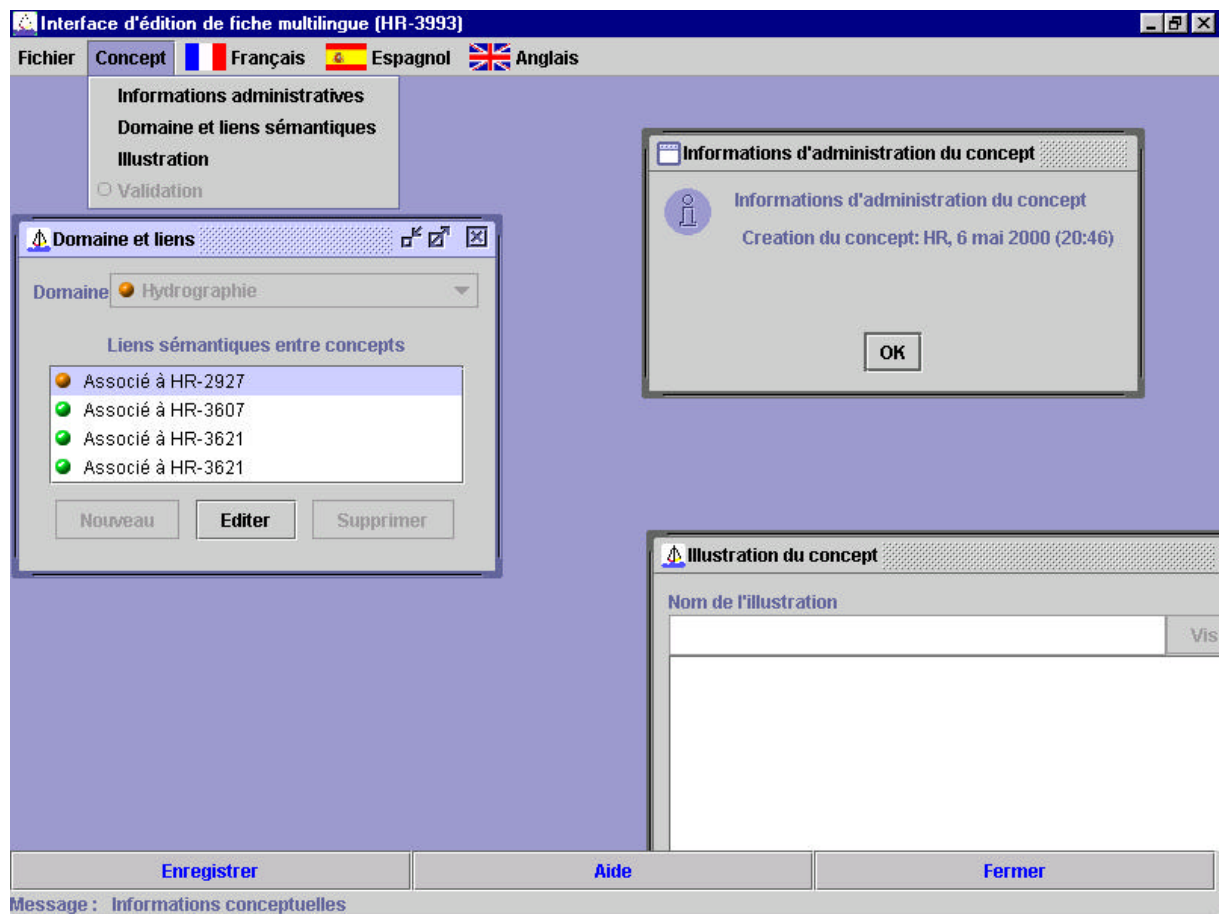
Sur la base du modèle de données proposé par le centre de recherche TERMISTI, l'interface *Dhydro* a été conçue par Jean-Luc Husson, membre de l'équipe Langue et dialogue du LORIA⁴, un vaste centre de recherche en informatique situé à Nancy. Soucieux d'ergonomie, il a veillé à conférer au logiciel un mode de fenêtrage qui rende fidèlement compte de la structure des fiche terminologiques et le dote ainsi de réelles qualités didactiques.

3. Peu utile aux yeux de la « doctrine » conceptuelle, l'existence de ce champ est liée au cahier des charges de l'Organisation hydrographique internationale.

4. Laboratoire lorrain de recherche en informatique et ses applications, <http://www.loria.fr/equipes/led/led.html>.

3.1 Informations situées au niveau du « concept »

Des fenêtres différentes correspondent aux différentes informations : administration, illustration, domaines et liens.

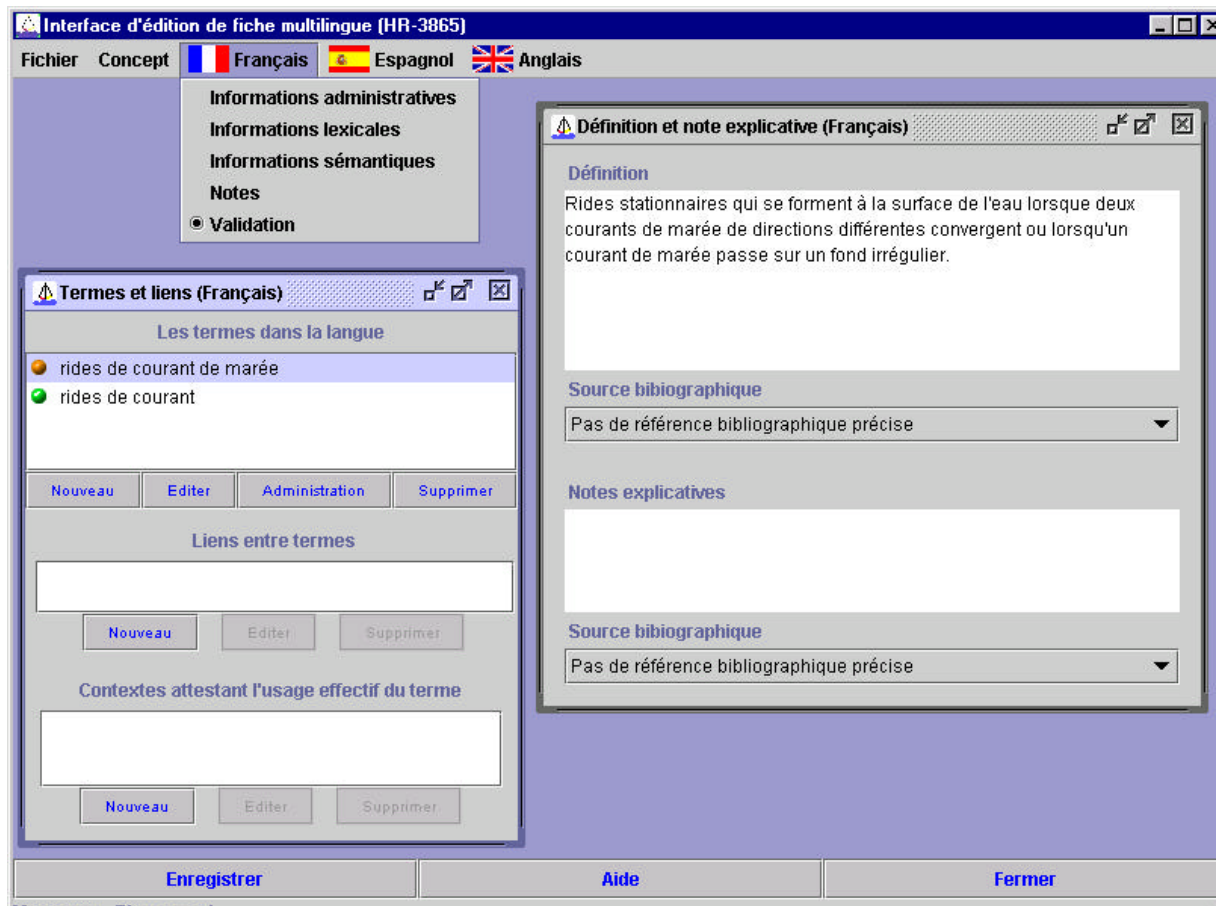


Écran 1 : fenêtrage des informations liées au niveau conceptuel

3.2 Informations dépendant des langues

3.2.1 INFORMATIONS ADMINISTRATIVES ET SÉMANTIQUES

Ces informations communes à tous les termes synonymes figurent dans des fenêtres distinctes. Le fait que les données définitoires figurent dans une fenêtre distincte de celle des termes rappelle en permanence au rédacteur que la définition doit s'appliquer à l'ensemble des synonymes.



Écran 2 : les fenêtres des informations sémantiques et lexicales sont distinctes

3.2.2 INFORMATIONS LEXICALES

Pour chaque terme, toutes les informations de nature lexicale apparaissent également dans des fenêtres particulières : description de chaque terme (écran 3), contexte(s) d'attestation (écran 4) et liens formels avec d'autres termes (écran 5).

Interface d'édition de fiche multilingue (HR-3865)

Fichier Concept Français Espagnol Anglais

Termes et liens (Français)

Les termes dans la langue

- rides de courant de marée
- rides de courant

Nouveau Editer Administration Supprimer

Liens entre termes

- Forme étendue de rides de courant (HR-3865-fr-2)

Nouveau Editer Supprimer

Contextes attestant l'usage effectif du terme

Nouveau Editer Supprimer

Description d'un terme (Français)

Description d'un terme

Terme vedette

Orthographe principale

rides de courant de marée

Informations grammaticales

nom féminin

Origine

Prononciation

Usage : fréquence

Usage : localisation géographique

Usage : vieillissement

Usage : registre de langue

Usage : note

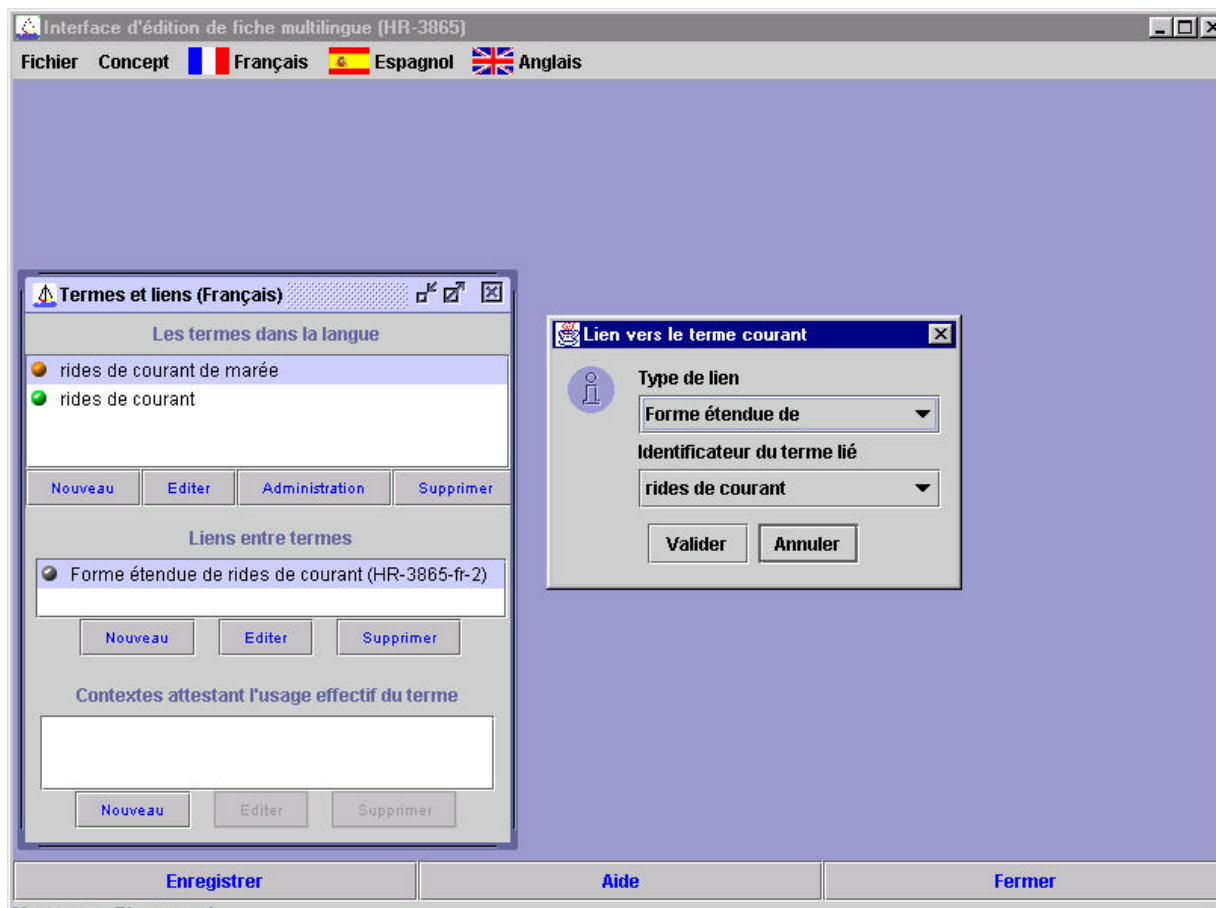
Valider Annuler

Enregistrer Aide Fermer

Écran 3 : fenêtre de description de chaque terme



Écran 4 : fenêtre de description des contextes



Écran 5 : fenêtre des liens formels avec d'autres termes

On le voit au travers des captures d'écran : les concepteurs de *Dhydro* ont tenté de créer une interface conviviale qui rende compte d'une structure pertinente. Elle devrait nécessairement amener le rédacteur, qui n'est pas toujours un linguiste, à réfléchir à la manière dont il engrange ses informations et à découvrir que l'établissement d'équivalences entre des termes spécialisés requiert une démarche analytique rigoureuse. Cette qualité sera sans doute appréciée à sa juste valeur par les enseignants soucieux d'initier les futurs traducteurs à une démarche critique plutôt que d'en faire de simples techniciens maîtrisant le seul mode d'emploi de tel ou tel logiciel à la mode.

Marc Van Campenhoudt,
Centre de recherche Termisti,
Institut supérieur de traducteurs et interprètes,
Bruxelles.

Bibliographie

Blampain (D.), Descotte (S.), Husson (J.-L.), Rohde (H.), Romary (L.), Van Campenhoudt (M.) et Viscogliosi (N.), 2000 : « Le projet européen DHYDRO : la normalisation à l'épreuve d'un forum terminologique », dans *Actes de la Conférence sur la coopération dans le domaine de la terminologie en Europe (Paris, 17-19 mai 1999)*, Association européenne de terminologie.

Centre de recherche TERMISTI, 2000 : *Dhydro : guide de l'utilisateur*, http://www.termisti.refer.org/dhydro/aide_sommaire.html.

Centre de recherche TERMISTI & LORIA, 1999 : *Le modèle de données et la D.T.D.*, livrable D 1.2 du projet MLIS-2009 *Dhydro*, mars 1999.

ISO 12618 (1994) : *Aides informatiques en terminologie - Création et utilisation de bases de données terminologiques et de corpus de textes*, Genève, International Organization for Standardization.

ISO 12620 (1999) : *Aides informatiques en terminologie - Catégories de données*, Genève, International Organization for Standardization.

Merten (P.), Mertens (J.) et Van Campenhoudt (M.), 1993 : « Microglossaire, réseau notionnel et gestion informatique. Une expérience de recherche en Communauté française de Belgique », dans Gouadec (D.), dir., *Terminologie & terminotique : outils, modèles & méthodes. Actes de la première université d'automne en terminologie. Rennes 2, 21-26 sept. 1992*, Paris, La Maison du dictionnaire, 1993, p. 277-293.

Van Campenhoudt (M.), 2000 : « De la lexicographie spécialisée à la terminographie : vers un "métadictionnaire" ? », dans Thoiron (Ph.) et Béjoint (H.), dir., *Le sens en terminologie*, Lyon, Presses universitaires de Lyon (Travaux du C.R.T.T.), p. 127-152.
